The Journal of the Acoustical Society of America

MAY 29 2024

Phrase boundaries lacking word prosody: An articulatory investigation of Seoul Korean

Jiyoung Jang; Argyro Katsika

J. Acoust. Soc. Am. 155, 3521–3536 (2024)
https://doi.org/10.1121/10.0026081

JASA ARTICLE

Phrase boundaries lacking word prosody: An articulatory investigation of Seoul Korean

Jiyoung Jang[1,a)] and Argyro Katsika[2]
[1]*Hanyang Institute for Phonetics and Cognitive Sciences of Language, Hanyang University, Seoul 04763, South Korea*
[2]*Department of Linguistics, University of California, Santa Barbara, Santa Barbara, California 93106, USA*

ABSTRACT:
This electromagnetic articulography study explores the kinematic profile of Intonational Phrase boundaries in Seoul Korean. Recent findings suggest that the scope of phrase-final lengthening is conditioned by word- and/or phrase-level prominence. However, evidence comes mainly from head-prominence languages, which conflate positions of word prosody with positions of phrasal prominence. Here, we examine phrase-final lengthening in Seoul Korean, an edge-prominence language with no word prosody, with respect to focus location as an index of phrase-level prominence and Accentual Phrase (AP) length as an index of word demarcation. Results show that phrase-final lengthening extends over the phrase-final syllable. The effect is greater the further away that focus occurs. It also interacts with the domains of AP and prosodic word: lengthening is greater in smaller APs, whereas shortening is observed in the initial gesture of the phrase-final word. Additional analyses of kinematic displacement and peak velocity revealed that Korean phrase-final gestures bear the kinematic profile of IP boundaries concurrently to what is typically considered prominence marking. Based on these results, a gestural coordination account is proposed, in which boundary-related events interact systematically with phrase-level prominence as well as lower prosodic levels, and how this proposal relates to the findings in head-prominence languages is discussed. *© 2024 Acoustical Society of America*.
https://doi.org/10.1121/10.0026081

(Received 29 July 2023; revised 19 April 2024; accepted 2 May 2024; published online 29 May 2024)
[Editor: Susanne Fuchs] Pages: 3521–3536

I. INTRODUCTION

A primary role of *prosodic structure* is governing prosodic constituency (e.g., grouping syllables into words and words into phrases). This is achieved through several phonetic dimensions, in which one is the temporal profile of the utterance. Indeed, *phrase-final* or *pre-boundary lengthening* is a phonetic dimension of the most widely attested effects of prosodic structure. These terms capture the phenomenon that speech units immediately preceding a *prosodic boundary*, i.e., at the end of a prosodic phrase, present longer durations than their phrase-medial counterparts (e.g., Byrd *et al.*, 2006; Turk and Shattuck-Hufnagel, 2007; see Fletcher, 2010). The phenomenon is so prevalent in oral and signed speech (Wilbur, 2009) that it is considered a language universal (e.g., Vaissière, 1983; Tyler and Cutler, 2009). Despite the effect's prevalence and expansive attestation, the stretch of speech affected, i.e., the scope of the effect, remains understudied and unclear. Given how fundamental prosodic boundaries are for language processing, acquisition, and communication (see Gussenhoven and Chen, 2020) and how ubiquitous the presence of pre-boundary lengthening is at those boundaries, examining the effect's scope can shed important light onto the factors that define prosodic boundaries and what this means for the grammatical structure of prosody and speech planning.

Based on the scarce work on the matter, two opposing theories have been proposed. On the one hand, the effect might be determined by a possibly language-specific, prosodic domain such as the phrase-final syllabic rhyme (Oller, 1973; Wightman *et al.*, 1992) or foot (White, 2002) or even word (Kohler, 1983). Alternatively, the effect might not target a well-defined grammatical domain but, instead, scope over a fixed interval at the boundary (Byrd and Saltzman, 2003). Regardless, stress is a factor that further fine-tunes the effect: It is initiated earlier in phrase-final words with earlier stress (e.g., Byrd and Riggs, 2008; Katsika, 2016; Kim *et al.*, 2017; see also Turk and Shattuck-Hufnagel, 2007; White, 2002). However, positions that bear stress, i.e., the marker of word prosody, are conflated with positions that can carry phrase-level prominence as the latter is marked by phrase pitch accents on stressed syllables of accented words. It is, thus, unclear which function of stress contributes to the scope of phrase-final lengthening: the marker of word prosody or the anchor for phrase-level prominence. Recent work on Japanese finds that lexical pitch accent shows patterns similar to those of lexical stress (Tsai and Katsika, 2020; Tsai, 2023), enhancing the hypothesis that word prosody determines the timing of boundary effects (see also Seo *et al.*, 2019). Yet, in that work (Tsai and Katsika, 2020; Tsai, 2023), phrase-level prominence was not considered.

Here, we turn to Seoul Korean, a language that can help us isolate the contribution of phrase-level prominence to

a)Email: jiyoungljang@hanyang.ac.kr

boundary marking. This is because Seoul Korean lacks word prosody and marks phrase-level prominence (e.g., focus) by the means of boundary tones at the edges of relatively small phrases, called Accentual Phrases (APs; Beckman and Pierrehumbert, 1986; Jun 1993). This property warrants Korean its characterization as an edge-prominence language (cf. Jun, 2014). In parallel, the initial boundary tone of APs in edge- (and head/edge-) prominence languages shares an important function with lexical stress—but not phrasal pitch accent—in stress languages, that of facilitating word segmentation (e.g., stress, Cutler and Norris, 1988; AP in Korean, Kim, 2004; Kim and Cho, 2009; AP in French, Welby, 2007; AP in Japanese, Warner *et al.*, 2010). For this reason, the study reported here is designed to examine the effect of marking phrase-level prominence/focus on the scope of phrase-final lengthening in Korean separately from that of word demarcation.

By the means of this investigation, we hope to extend the field's understanding of what factors determine the scope of phrase boundaries and how the latter interact with other grammatical domains/prosodic levels such as the syllable, prosodic word (PWd), and AP. Knowledge will also be gained as to how speech planning works to interweave different sources of prosodic and lexical/segmental information (cf. Keating and Shattuck-Hufnagel, 2002 vs Levelt, 1989). Moreover, implications can be drawn regarding how prosodic structure interfaces with information structure, specifically, focus. To better understand the kinematics of phrase boundaries, how they differ from prominence (cf. Cho, 2006), and the dynamic patterns that might give rise to them, a derived goal of this investigation is to examine duration along with position and velocity, as well as the relationship among these kinematic dimensions, following the assumptions of Articulatory Phonology (Browman and Goldstein, 1992) and the pi-gesture model of prosodic boundaries (Byrd and Saltzman, 2003), within which the current work is couched. Our ultimate goal is to contribute to a more comprehensive definition of the prosodic component of grammar and its role in linguistic cognition.

A. Prosodic structure

Theories of prosody assume that the prosodic structure is organized hierarchically with higher-level prosodic units consisting of one or more lower-level units (e.g., Beckman and Pierrehumbert, 1986; Hayes, 1989; Nespor and Vogel, 1986; Selkirk, 1984). For example, in the prosodic structure model proposed by Beckman and Pierrehumbert (1986), syllables constitute PWds; PWds constitute intermediate phrases (ips); and ips constitute intonational phrases (IPs). Prominence is marked at different levels of the structure: At the lexical level, prominent syllables are marked with stress, which indicates relative salience of these syllables compared to the unstressed syllables. At the phrasal level, accented syllables are marked with a pitch accent, which indicates that these syllables are rhythmically or conceptually more prominent than others in the phrase. Finally, prosodic structure includes information on tonal events of constituent boundaries: phrase accents marking the edges of ips and boundary tones marking the edges of IPs. These tonal markings, along with pitch accents, describe the overall intonation of the utterance.

Although there is general agreement in the structural view of prosody, models disagree in the number of hierarchical levels that they propose (see Shattuck-Hufnagel and Turk, 1996, for a review). Also, different languages are reported to exhibit distinct prosodic structures. For instance, languages, such as Japanese and Korean, are assumed to have an additional phrasal level, referred to as AP, the size of which is equal to or larger than the PWd (Beckman and Pierrehumbert, 1986; Jun, 1993). Moreover, languages may have different word prosody and phrase-level prominence systems (e.g., Jun, 2014). At the word prosody level, some languages may employ lexical stress, others may use lexical tone or lexical pitch accent, while others may not use any type of word prosody at all. At the phrase prosody level, languages are broadly categorized into head- and edge-prominence languages, where prominence is marked by a pitch accent on the head of the phrase at some phrasal level (e.g., ip) in the former and by a boundary tone at the edge of the phrase at some phrasal level (e.g., AP) in the latter. Hybrid systems have been proposed at word prosody and the phrase prosody level.

B. Korean prosody

This section summarizes key properties of Korean prosodic structure related to the current study. Korean, specifically Seoul Korean (also known as "Standard Korean" or "Pyojuneo," based on the Korean spoken in Seoul and Gyeonggi province), is an edge-prominence language known for lacking lexical stress, lexical tone, or lexical pitch accent (Jun, 1993, 2005). AP serves as the basic intonational unit, which, as proposed by Jun (1993, 2005), has an underlying tonal pattern of THLH, with H referring to High tone and L referring to Low tone. The realization of the initial tone (T) tends to depend on the laryngeal configuration of the AP-initial segment[1] (see Jun, 2005; Jeon and Nolan, 2017). This tonal pattern of AP is "not specific to a lexical item but is a property of the phrase (Jun, 1993)." Here, we adopt Jun's view that specifies the AP and IP levels above PWd in Korean (Jun, 2005, 2000). Some proposals include an intermediate ip level between the AP and IP, but as there has not been much investigation on this level and detecting it is not straightforward (Jun, 2005), we do not consider this level.

Phrase-level prominence in Korean is known to be marked by prosodic phrasing with the focused word consistently initiating an AP or higher phrase (Jun, 1993, 2005). AP boundaries between the focused item and the right edge of the IP are assumed to undergo elimination or possibly attenuation, a process referred to as dephrasing (e.g., Jun, 1993). Limited work on the phonetic correlates of prominence in Korean reports that under focus, articulatory constriction movements become longer, larger, and faster

(Shin *et al.*, 2015; Jang and Katsika, 2019; Jang, 2023). These effects vary with focus type, increasing from broad to narrow and eventually to contrastive focus (Jang, 2023). They are also stronger on the initial gesture of the focused AP and decrease with distance from AP's onset, affecting several syllables and, depending on the type of focus, even crossing word boundaries (Jang, 2023). Interestingly, IP boundaries at the right edge of the prosodic domain appear to have the same effects as those of prominence in Korean, i.e., longer, larger, and faster gestures (Kim *et al.*, 2019). This pattern, which differs from the expectation of longer, larger, but slower gestures at boundaries based on observed data (cf. Cho, 2006) and the predictions of the pi-gesture model (Byrd and Saltzman, 2003), has been attributed to the edge-prominence property of Korean: Korean boundaries present the kinematic signature of prominence because Korean uses edges, i.e., boundaries, to mark prominence (Kim *et al.*, 2019). We will return to this point below as more background information on the kinematics and modeling of prosodic boundaries is provided.

C. The phonetics of boundaries

As phrase boundaries are one of the most robust prosodic landmarks, a number of studies have looked at the phonetic events observed at or across them. Phrase-initially, constriction gestures have been reported to be temporally and spatially expanded (i.e., greater constriction for consonants and larger opening for vowels) under the notion of domain-initial strengthening (e.g., Fougeron and Keating, 1997; Cho and Keating, 2001, 2009; Keating *et al.*, 2004). Evidence indicates that the effect is cumulative, i.e., increasing with boundary strength (e.g., acoustic studies, Cho and Jun, 2000; Fougeron, 2001; articulatory studies, Byrd *et al.*, 2006; Fougeron and Keating, 1997; Keating *et al.*, 2004). The scope of this effect has been regarded to be limited to the initial segment, primarily based on articulatory data (Byrd *et al.*, 2006; Katsika, 2016; Byrd and Saltzman, 1998).

Lengthening is the most well-known characteristic on phrase-final gestures, found in many different languages (e.g., American English, Oller, 1973; Turk and Shattuck-Hufnagel, 2007; British English, White, 2002; Campbell and Isard, 1991; Korean, Kim *et al.*, 2019; Japanese, Seo *et al.*, 2019; Tsai and Katsika, 2020; Dutch, Cambier-Langeveld, 1997; Greek, Katsika, 2009, 2016). Similar to domain-initial strengthening, phrase-final lengthening is also cumulative: the higher that the prosodic boundary is, the greater that the effect is (e.g., Byrd, 2000; Byrd and Saltzman, 1998; Cambier-Langeveld, 1997; Cho, 2006; Krivokapić, 2007). Phrase-final positions demonstrate larger and less overlapped gestures (Byrd and Saltzman, 2003; Byrd *et al.*, 2000; Cho, 2005, 2006). However, studies on the spatial effects of phrase-final lengthening are not only scarce but also tend to show much variation across speakers and phonemes examined (Byrd *et al.*, 2005, 2006; Fougeron and Keating, 1997). Finally, gestures slow down at boundaries, as indicated by findings on peak velocity (Cho, 2006) and time-to-peak velocity (Byrd, 2000). Nonetheless, there are findings of faster movements phrase-finally, not only from Korean as discussed in Sec. I B, but also from English, albeit on a speaker- and token-specific basis (Edwards *et al*, 1991; Byrd and Saltzman, 1998). Notably, in these studies, these faster movements were also larger. Byrd and Saltzman characteristically describe the patterns in Byrd and Saltzman (1998) as follows: "*our own experimental data have patterned variously in this regard, tending at boundaries to show slower peak velocities for comparable displacements but faster peak velocities when displacements increase, as they often do*" (p. 170).

D. The scope of phrase-final lengthening

As introduced earlier, phrase-final lengthening is a well-studied phenomenon and, yet, the scope of the effect and what determines it are still unclear issues. Cumulative work on this phenomenon suggests that the greatest and most reliable lengthening is found on the rhyme of the phrase-final syllable (e.g., Edwards *et al.*, 1991; Wightman *et al.*, 1992; Byrd *et al.*, 2006), but lengthening might affect earlier parts of the phrase-final word depending on factors such as the quality of the phrase-final vowel (Cambier-Langeveld, 1997) and the position of lexical stress (Oller, 1973; Berkovits, 1994; Katsika, 2016; Katsika *et al.*, 2014). The effect is progressive, i.e., strongest at the boundary and gradually decreasing with distance from it (see Campbell and Isard, 1991; White, 2002, among others).

Studies that specifically explore the effect of stress and/or focus-marking pitch accent on the scope of phrase-final lengthening detect systematic interactions. For instance, in Greek, the earlier the stress is within the phrase-final word, the earlier boundary tones and phrase-final lengthening are initiated, even in de-accented phrase-final words (Katsika, 2016; Katsika *et al.*, 2014). Similar effects of lexical stress on pre-boundary lengthening are also found in English: In an articulatory study, Kim *et al.* (2017) show that phrase-final lengthening begins earlier in the final word when lexical stress is nonfinal as opposed to final (but see Byrd and Riggs, 2008), whereas in an acoustic study, Turk and Shattuck-Hufnagel (2007) detect phrase-final lengthening on the final syllable as well as on the stressed/accented syllable, leaving intervening syllables unaffected (see also patterns in British English, White, 2002). This work provides great insight on the possible factors determining the activation of phrase-final lengthening, such as the marker of word prosody within the phrase-final PWd and/or the location of focus in the phrase as denoted by the nuclear pitch accent. As research on this topic has mainly focused on stress languages, in which the positions of word prosody markers and pitch accents in principle coincide, a separation of the contribution of these two factors is difficult to draw. This task becomes more problematic when considering that stress and accent, when phonetically marked, might carry the same kinematic signature of longer, larger, and faster gestures (see Katsika and Tsai, 2021, and references therein). Moreover, stress, in addition to marking a structural position and attracting pitch accents, has also been shown to function as a cue to word segmentation (Cutler and Norris, 1988), suggesting that it might not be word prosody or phrasal prominence that affects the onset of boundary marking but whichever dimension denotes in the given language "*this is the last word in the phrase*." In edge- (and head/edge-) prominence languages, this dimension is the initial boundary tone of APs (e.g., Korean, Kim, 2004; Kim and Cho, 2009; French, Welby, 2007; Japanese, Warner *et al.*, 2010). In a recent articulatory study of Japanese, a language with lexical pitch accent, a direct examination of the scope of phrase-final lengthening as a function of lexical pitch accent position in the word detected an effect similar to that of lexical stress: lengthening is initiated earlier the earlier the pitch accent (Tsai, 2023; Tsai and Katsika, 2020; cf. acoustic data comparing initial-accented to unaccented words in Seo *et al.*, 2019). We note, however, that in this study, demarcation and focus were not directly controlled.

Altogether, we take these results to have typological implications, suggesting that any of the following factors might be the main driver of or a contributor to the phonetic activation of prosodic boundaries: word prosody (e.g., lexical stress/lexical pitch accent), word demarcation (e.g., lexical stress/AP tones), and focus location (nuclear pitch accent/AP tones of focused AP). Here, we turn to Seoul Korean, capitalizing on the typological advantage of looking at a language in which word prosody is missing, disentangling itself from phrase-level prominence. We examine the role of the latter by manipulating the location of focus in the phrase while we control for the length of the final AP as an index of where the final word demarcation cue is. As mentioned above, Korean prosodic boundaries have been previously researched kinematically (Kim *et al.*, 2019). In that study, lip movements over disyllabic words were measured, and it was found that phrase-final lengthening was distributed across both syllables of the phrase-final word regardless of their information status (background vs new information) or their vowels' intrinsic duration (/a/ vs /i/). As a reminder, lip movements were also found to be larger and, contrary to expectations, faster.

E. Articulatory Phonology and the pi-gesture framework

The patterns of longer, larger, slower, and less overlapped gestures at prosodic boundaries have been modelled by the means of pi-gestures (Byrd and Saltzman, 2003). The pi-gesture model is couched within Articulatory Phonology (Browman and Goldstein, 1992; for a comprehensive review, see Iskarous and Pouplier, 2022) and extends the view that phonological units are inherently dynamic to the prosodic level (Byrd and Riggs, 2008). Within Articulatory Phonology, the atomic units of phonology are gestures. Constriction gestures are specified for abstract linguistic tasks (e.g., lip closure), executed by coordinated actions of the speech articulators (e.g., the lips, tongue tip, tongue dorsum, velum, and vocal folds), and triggered by internal oscillators that are coupled to each other either in-phase or anti-phase, forming syllables and words (Goldstein *et al.*, 2006). Mathematically, constrictions gestures are defined as critically damped second-order dynamical linear systems, which are, thus, characterized by two constants, namely, the *target* of the linguistic task (e.g., constriction degree –2 for lip closure) and *stiffness*, which controls how fast the goal is achieved. Two indirect measures of the abstract dimension of stiffness have been proposed in the literature: time-to-peak velocity (the shorter that the time-to-peak velocity is, the higher the stiffness and the faster the movement are; Byrd and Saltzman, 1998; used in, e.g., Cho, 2002, 2006; Mücke and Grice, 2014) and normalized peak velocity over displacement (capturing the observation that peak velocity increases as displacement increases; Munhall *et al.*, 1985; Ostry and Munhall, 1985; used in, e.g., Beckman *et al.*, 1992; Hawkins, 1992; Roon *et al.*, 2007). Recent proposals of nonlinear dynamical models for gestures have been put forward, but these will not be discussed here further because they have not, yet, been integrated to the rest of the theory (e.g., the prosodic model). The oscillators to which gestures are associated are mathematically expressed as limit cycle oscillators.

Unlike constriction gestures, pi-gestures are not related to specific articulators, and their task is to locally slow down the clock that controls the global pace of the utterance (Byrd and Saltzman, 2003). The mathematical expression of pi-gestures relates the amount of this slowing down to the strength of the pi-gesture, thus, being able to capture the phonological concept of prosodic levels. On the basis of longer time-to-peak velocity indicating a lowering of the gestures' stiffness parameter in the vicinity of the prosodic boundaries (Byrd and Saltzman, 1998; Byrd, 2000; Beckman *et al.*, 1992), first discussions of pi-gestures involved control of the stiffness parameter. However, change in stiffness parameter alone proved to be insufficient to capture the comprehensive kinematic variation at phrasal boundaries (Byrd and Saltzman, 2003; Saltzman and Byrd, 2000). Clock-slowing, on the other hand, provides an overarching account.

Clock-slowing occurs locally because it is determined by the activation interval of the pi-gesture, represented by the gray-shaded box in Fig. 1(a). Pi-gestures modulate the temporal properties of (the part of) constriction gestures that overlap with their activation interval, as shown in Fig. 1(a). Thus, in this framework, the scope of pre-boundary lengthening refers to the activation interval of a pi-gesture. The activation of pi-gestures reaches its maximum, the value of which depends on its strength, at the prosodic boundary and decreases with distance from it over a continuous domain, capturing the progressive and cumulative nature of boundary lengthening, which is discussed in Secs. I C and I D. As a consequence of the clock-slowing, gestures also become slower and less overlapped with each other. The larger gestures at boundaries, on the other hand, are the result of this reduction in overlap.

JASA
https://doi.org/10.1121/10.0026081

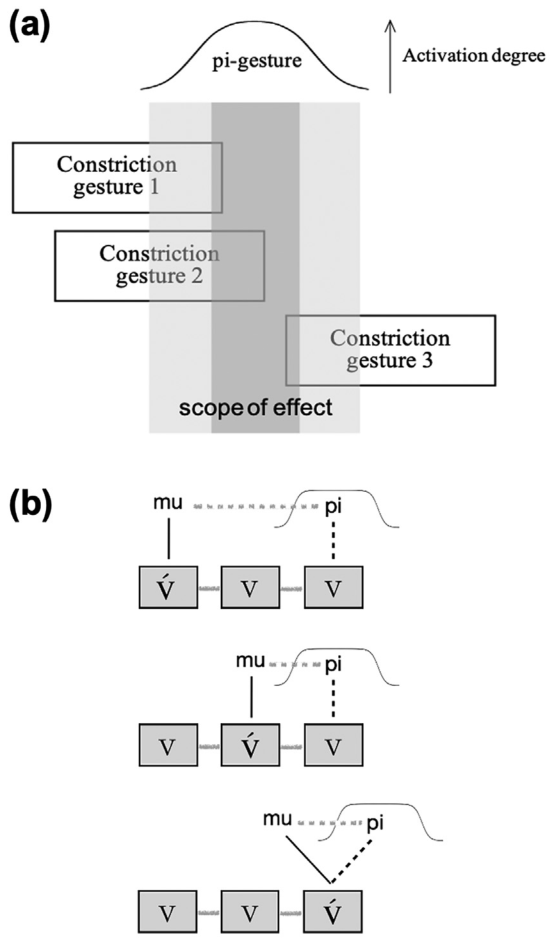

FIG. 1. (a) Schematic representation of a pi-gesture is shown (adapted from Byrd and Saltzman, 2003). The gray-shaded boxes represent the effect of pi-gesture with the darker box corresponding to the pi-gesture's maximal activation. (b) Schematic representation of the dual coordination of pi-gestures with phrase-final V and mu-gestures in trisyllabic stress-initial (top), stress-medial (middle), and stress-final (bottom) words are displayed (adapted from Katsika, 2016).

Another dimension that might determine the scope of pi-gesture's effect is its coordination with other gestures. In the original model (Byrd and Saltzman, 2003), pi-gestures were not coordinated with constriction or other gestures. However, recent work (Katsika, 2016; Katsika and Tsai, 2021) suggests that incorporation of pi-gestures and all modulation gestures, in general, in coupling relationships with the other gestures could capture the systematic effects of factors, such as the position of the final stress and/or accent on the scope of phrase-final lengthening (see Sec. I D), as well as typological distinctions in terms of word- and phrase-level prosody. Under the original proposal, the effect of phrase-final lengthening that emerges from the pi-gesture does not coincide with a grammatical domain but lasts for a specific interval at the boundary. In the coupling-based proposal, on the other hand, pi-gestures are coordinated with other phonological units and, thus, grammatical domains. However, the mode of coordination (e.g., anti-phase) along with competition from multiple coupling connections of the pi-gesture might result in the scope of the lengthening effect not corresponding to a specific grammatical domain.

Modulation gestures, called mu-gestures, essentially extend the concept of pi-gestures to incorporate prosodic modeling of the spatiotemporal effects of stress, accent, and, in general, prominence (Saltzman *et al.*, 2008). Two types of modulation gestures have been proposed, i.e., temporal and spatial. The goal of the temporal gestures is to modulate the evolution of time as determined by the global clock of the utterance, whereas the goal of the spatial gestures is to affect the target parameter of concurrent constriction gestures. However, only the temporal gestures have been defined mathematically; they are determined by the same differential equation as the pi-gesture. Although mu- and pi-gestures are fundamentally the same object, in the remainder of this text, modulation gestures related to IP boundaries will be referred to as pi-gestures, and those related to focus will be referred to as mu-gestures.

Returning to the issue of pi-gesture coordination, it has been proposed that pi-gestures in Greek are coordinated with the phrase-final V (vocalic) gesture and the last stress-instantiating mu-gesture (Saltzman *et al.*, 2008). This competitive dual coordination accounts for the differences in scope that phrase-final lengthening has as a function of stress in Greek. In stress-final words, the two coordination nodes—phrase-final V gesture and mu-gesture—coincide in the final syllable. However, this is not the case in words with nonfinal stress, in which the stress-related mu-gesture pulls the pi-gesture toward the stressed syllable, initiating phrase-final lengthening earlier in words with nonfinal stress as opposed to words with final stress [Fig. 1(b)].

F. Research questions and predictions

The first and main question ($Q1$) addressed here is: *What is the scope of phrase-final lengthening in Seoul Korean, and is it fine-tuned by the position of phrase-level prominence and the word demarcation function of the phrase-final AP, or both?* In terms of general patterns, we expect that Korean will present phrase-final lengthening, which might span over the last two syllables or the whole phrase-final word, as shown in Kim *et al.* (2019). The effect should be progressive, where the rhyme of the final syllable is affected the most (Byrd *et al.*, 2006; Krivokapić, 2007; Katsika, 2016; see also acoustic studies, Berkovits, 1994; Cambier-Langeveld, 1997; Oller, 1973; Wightman *et al.*, 1992). As previous work has mainly dealt with stress languages, the effect of phrase-level prominence and/or word demarcation function of final AP on the scope of phrase-final lengthening in Korean, which is a language free from the multi-functionality of lexical stress (i.e., word-level prosody marker, word demarcation cue, and anchor for phrasal pitch accent), constitutes an exploratory question. Under the assumption that the location of phrase-level prominence affects the onset of phrase-final lengthening, pre-boundary lengthening should be initiated earlier the earlier the focus is in the IP. If the dimension of word demarcation plays a role, lengthening will begin earlier in long as opposed to short final APs because in the former, the word demarcation cue connected to the AP's left edge is farther away from the IP boundary. Finally, a pattern in which the scope of the effect is not adjusted with respect to either AP length or focus position would suggest that it is the marker of word prosody within the phrase-final word *per se* that has been found to affect the span of phrase-final lengthening in the previously reported languages (e.g., lexical stress in Greek and English and lexical pitch accent in Japanese).

The second question (*Q*2) of this study is: *What is the displacement and velocity profile of IP boundaries in Seoul Korean?* We predict that phrase-final gestures will be larger (Cho, 2006; Kim *et al.*, 2019). However, it is unclear whether they will be faster as previously reported for Korean and assumed to be a characteristic of edge-prominence (Kim *et al.*, 2019; but see also English data in Byrd and Saltzman, 1998) or slower (Byrd, 2000; Cho, 2006), as predicted by the pi-gesture model, especially when their displacement is taken into consideration as suggested in Byrd and Saltzman (2003).

Extending *Q*2 and to specifically address the latter point, i.e., the relationship between displacement and velocity, we include a third question (*Q*3): *How is the relationship between kinematic parameters modulated by IP boundaries vs prominence?* We include this question for two reasons: first, to understand the relationship between kinematic dimensions and the dynamical systems potentially responsible for the modulation at IP boundaries, as well as the modulation under prominence (although secondarily as the kinematic profile of prominence is not the main topic of this paper), and second, to evaluate the conflicting predictions regarding the velocity profile at IP boundaries, specifically, whether IP-final gestures are faster or slower in Korean as in *Q*2. It has been established that (a) duration increases as stiffness (measured either as normalized peak velocity over displacement or time-to-peak velocity) decreases, and (b) displacement increases as peak velocity increases (e.g., Byrd and Saltzman, 1998; Munhall *et al.*, 1985). Nonetheless, what remains to be seen is how these relationships are affected by the prosodic-structural modulations. Of particular interest is examining what happens when one source is predicted to modulate the gestures to move slower (e.g., IP boundary) while the other requires them to move faster (e.g., prominence). Thus, we aim to bring together the two effects and investigate how each modulation interacts kinematically with the other.

II. METHODS

A. Participants

Seven native speakers of Seoul Korean (five females and two males) in their 20s ($age_{mean} = 24.5$ years old; $age_{median} = 23$ years old) participated in the present study. Data collection could not be extended due to COVID-19. Six participants were affiliated with the University of California Santa Barbara (UCSB) as graduate or exchange students or postdoctorate researchers at the time of the experiment, and one participant was a family member of a UCSB-affiliated researcher. The speakers were naive as to the purpose of the study and had no self-reported speech, hearing, or vision problems. They gave informed consent, approved by the UCSB Institutional Review Board, and received financial compensation for their participation.

B. Experimental procedure

Before the experiment, the participants went through a short, 15-min long, training session to be familiarized with the speech materials and the experimental procedure. To contextualize the speech materials, participants were instructed to imagine a context in which they were preparing a school play with friends, and the target words were introduced as names of roles in the play. This setup allowed participants to naturally produce the AP phrasing and place the focus contrast as intended (see Sec. II C for details on the experimental design and stimuli).

For the experimental session, kinematic data were collected using the AG501 3D electromagnetic articulograph (Carstens Medizinelektronik, Bovenden, Germany) at UCSB SPArK (Speech, Prosody, and Articulatory Kinematics Laboratory). Ten receiver coils were attached to the participants' head and vocal tract as follows: tongue dorsum, tongue tip, midway between tongue dorsum and tongue tip, upper lip and lower lip, upper and lower incisor, left and right ear, and nose. The last five sensors served as references. Audio recordings were performed simultaneously with the kinematic recordings by means of a Sennheiser shotgun microphone (Wedemark, Germany) set at a sampling rate of 16 kHz and positioned 1 ft away from the participant's face. Speech materials were presented on a computer screen, placed roughly 3 ft away from the participant using custom software, which was developed by Mark Tiede (Haskins Laboratories, New Haven, CT). The acquired articulatory data for each trajectory were smoothed and corrected for head movement by using the reference sensors. Then, they were rotated to align the *X* and *Y* axes to the participants' occlusal plane. To help appropriate focus placement, a prompt sentence preceded each target sentence. Prompt sentences were presented 1 s before their corresponding target sentence and read silently by the participant. Target sentences were read aloud. All target sentences had the form of a question, which participants were instructed to read as if they were talking to a friend.

C. Experimental design and stimuli

The scope of phrase-final lengthening was examined across the test PWd /nɛ.maŋ.mi.nam/ [a compound noun that means "*a handsome guy from Nemang* (name of a village)"], which was placed in the final position of the target IP. Two sets of test stimuli were constructed. One set (set *A*) consisted of three IPs as follows: adverbial phrase (AdvP) $\#_{IP}$ two noun phrases (NPs) $\#_{IP}$ verb phrase (VP; # indicates prosodic boundaries; see Table I). The target IP is the second IP with the two NPs. The AdvP in the first IP was always /tɕintɕ*aʔ/ ("Really?"). We refer to the IP-final NP of the target IP as the *ultimate* and to the IP-penultimate NP

https://doi.org/10.1121/10.0026081

TABLE I. Example stimuli with Penultimate focus are organized by Final NP length (short, long) and IP Position (IP-final, IP-medial). Measured PWds are underlined, focused NPs are displayed in bold. All sentences were preceded by an AdvP /tɕintɕ*a?/ (“Really?”).

Test sentence (# = IP boundary)
Penultimate focus, Short final NP, IP-final (Prompt: It’s not uncle Junseok.) [$_{\text{NP}}$ **minami** gomobuga] [$_{\text{NP}}$ nɛmaŋminam] # [$_{\text{VP}}$ sʌntækhangʌja] Minam uncle-NOM Nemang handsome guy decide-INT Uncle **Minam** from the secrecy club is the handsome guy from Nemang? Is it decided?
Penultimate focus, Short final NP, IP-medial (Prompt: It’s not uncle Junseok.) [$_{\text{NP}}$ **minami** gomobuga] [$_{\text{NP}}$ nɛmaŋminam] [$_{\text{VP}}$ sʌntækhangʌja] Minam uncle-NOM Nemang handsome guy decide-INT Uncle **Minam** from chose the secrecy club is the handsome guy from Nemang?
Penultimate focus, Long final NP, IP-final (Prompt: It’s not Junseok.) [$_{\text{NP}}$ **minamiga**] [$_{\text{NP}}$ matɕimak nɛmaŋminam] # [$_{\text{VP}}$ sʌntækhangʌja] Minam-NOM last Nemang handsome guy decide-INT **Minam** is the last handsome guy from Nemang? Is it decided?
Penultimate focus, Short final NP, IP-medial (Prompt: It’s not Junseok.) [$_{\text{NP}}$ **minamiga**] [$_{\text{NP}}$ matɕimak nɛmaŋminam] [$_{\text{VP}}$ sʌntækhangʌja] Minam-NOM last Nemang handsome guy decide-INT **Minam** chose the last handsome guy from Nemang?

as the *penultimate*. See Table I for a list of the example stimuli for set *A*. Another set of stimuli (set *B*) were constructed as to be identical with set *A*, except that they had an additional one-word NP */pi.mil.pu/* (“*the secrecy club*”) at the beginning of the target IP, i.e., the target IP contained three, instead of two, NPs: AdvP #$_{\text{IP}}$ three NPs #$_{\text{IP}}$ VP. For instance, the equivalent in set *B* of the set *A* stimulus listed first in Table I would be /pimilpu minamigomobuga nɛmaŋminam? sʌntaekhangʌja?/. Again, the target IP is the second IP with the three NPs. Stimulus set *B* was constructed to increase variability in the stimuli as well as to be analyzed as part of a parallel study on AP boundaries reported in Jang and Katsika (2023).

In the test stimuli, the target IPs were followed by an IP boundary and then a VP, which constituted its own IP. For instance, the first example listed in Table I illustrates that the target IP /minamigomobuga nɛmaŋminam?/ is followed by a VP /sʌntækhangʌja?/, which is an IP. A set of control stimuli included the test word in IP-medial position. Control stimuli involved the same sequence of words as their respective test stimuli but differed in that there was no IP boundary following the test word, which also gave rise to a difference in meaning (compare the first stimulus to the second stimulus in Table I).

The test word was purposefully selected to include nasal consonants to avoid segments whose laryngeal configurations would cause tonal effects (e.g., Jun, 1998). The goal was to elicit a typical AP tonal pattern of Seoul Korean, i.e., LHLHa where LHa marks the end of an AP (Jun, 2005). IP-finally, boundary tones in Seoul Korean are analyzed as overriding the final tone of the AP (Jun, 2000). To construct IP-final stimuli that is comparable to the typical LHLH pattern of the AP, interrogative sentences were used as these require a high boundary tone (H%), which, thus, matches in terms of pitch movement the high tone on the right side of the AP (Ha), resulting in LHLH% (Jun, 2000). In addition, the H% of interrogatives allowed for boundary tones to be detected and analyzed as part of a parallel investigation on boundary tone coordination (Jang and Katsika, 2022).

To assess the interaction between position of focus and phrase-final lengthening, each NP of the target IP was focused and manipulated: contrastive focus was either on the ultimate or the penultimate NP. To assess whether the onset of IP-final AP affects the timing of phrase-final lengthening, the IP-final NP was either four or seven syllables long. The decision to construct the shorter NP to contain four syllables was made based on the observation that typically APs contain 3–5 syllables, where APs containing four syllables are more likely to yield a full AP tonal pattern (i.e., THLHa, with each syllable bearing a tone; Jun, 2000). For the longer AP, it was crucial that its length differed from the short AP but only to an extent that allowed it to be produced in a single AP. Therefore, the longer AP was constructed to contain two PWds yielding seven syllables (Jun, 2003). Analysis of the elicited data confirmed that the participants produced the long NPs (seven syllables) within one AP with initial LH tone (of LHLHa) falling on the initial two syllables and LHa tone (of LHLHa) falling on the final two syllables (see Fig. 2, although the Ha of final AP tone is overridden by H%). We refer to the two levels of the final NP length factor as *short* (i.e., consisting of four syllables) and *long* final NPs (i.e., consisting of seven syllables). It is crucial to note that if dephrasing (following the focused linguistic unit) results in complete deletion of AP boundaries, this difference in final AP length should hold only for IPs focused on the ultimate NP. In IPs focused on the penultimate NP, the final AP should be 11 syllables long as a result of dephrasing.

The stimuli included two IP positions (IP-final, IP-medial), two focus locations (penultimate, ultimate), and two final NP lengths (short, long) in two- and three-NP long IPs, resulting in 16 stimulus sentences. Eight stimulus blocks were constructed, each containing 1 repetition of the 16 stimulus sentences in a randomized order. These blocks were intermixed with blocks containing stimuli constructed for other experiments. Note that for one speaker, five blocks were collected because of technical reasons. In total, 848 tokens were collected for the analyses reported here.

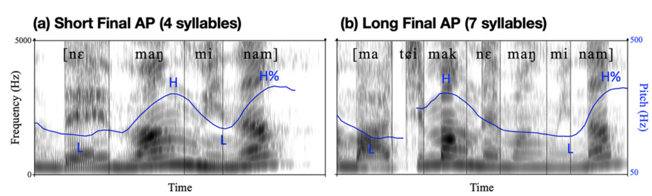

FIG. 2. (Color online) Example tonal configurations of (a) short final AP (four syllables) and (b) long final AP (seven syllables) with focus on the ultimate NP are displayed. *L* and *H* refer to low and high tones, respectively, and *H%* indicates boundary tone (Jun, 2000).

The acquired data were checked for their prosodic rendition, i.e., focus placement and appropriate AP and IP phrasing, using K-ToBI (Jun, 2000) principles by J.J. Either due to focus placed on unintended locations or boundaries inserted before the target word, 16 tokens were removed.

D. Measurements

With the exception of the coda /ŋ/ in the second syllable, all the consonant (C) gestures of the test PWd /nɛ.maŋ.mi.nam/ ("*a handsome guy from Nemang*") were measured using semi-automatic custom software called mview, developed by Mark Tiede (Haskins Laboratories). The coda /ŋ/ was not included in the measurements because of its degree of blending with the neighboring vowel. The remaining five consonants, i.e., the onset consonant of each syllable and the coda consonant of the final syllable, are referred to here as C4, C3, C2, C1, and C0, where C0 is the C gesture at the end of the word and, consequently, adjacent to the IP boundary, and C4 is at the beginning of the PWd and the most remote from the boundary.

For C1 and C4, both of which are /n/, C (consonant) constrictions were detected on the tongue tip vertical displacement trajectory. For C0, C2, and C3, which were all /m/, lip aperture was used. The labelling procedure detected the following kinematic timepoints in each C gesture: onset, time of peak velocity, target,[2] constriction maximum, release, and offset (Fig. 3). These timepoints were identified based on velocity criteria, i.e., peak velocity for the homonymous timepoints, velocity minima for constriction maxima, and velocity plateaus for the other timepoints. Velocity plateaus were detected based on a set threshold of 20%—the default in mview—of the velocity range between two consecutive alternating velocity extrema (i.e., one minimum and one maximum). Each gesture consisted of two phases: the formation (*F*), which corresponded to the interval between onset and release, and the release (*R*), meaning the interval between release and offset (see Fig. 3).

On the basis of these timepoints, the following measures were calculated:

(1) duration of gestural formation (i.e., time difference from onset to release) and duration of gestural release [i.e., time difference from release to offset; in ms; see Figs. 3(a) and 3(b), respectively];
(2) gestural displacement to target [i.e., the spatial difference between onset and maximum constrictor position; in mm; see Fig. 3(c)];
(3) gestural peak velocity to target [in cm/sec; see Fig. 3(d)]; and

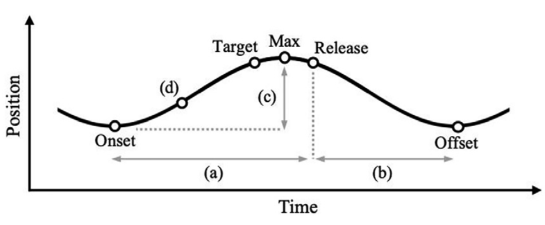

FIG. 3. Schematized constriction gesture with kinematic measurements.

(4) normalized peak velocity over displacement (i.e., peak velocity divided by displacement) as an approximate of *stiffness* (Munhall *et al.*, 1985; Ostry and Munhall, 1985; used in, e.g., Beckman *et al.*, 1992; Hawkins, 1992; Roon *et al.*, 2007), based on Articulatory Phonology's model of articulatory gestures as a critically damped second-order linear dynamical systems. This approximation to stiffness was used here instead of the alternative measure of time-to-peak velocity because it is derived from the other two kinematic measures calculated here, allowing for a more direct comparison between the assessed kinematic dimensions. However, preliminary analysis also examined the other measure of stiffness, i.e., time-to-peak velocity, and the two analyses reached similar conclusions. Note that stiffness was analyzed as a continuous predictor included in the model, rather than as a dependent variable (see Sec. II E for details.)

E. Statistical analysis

The retrieved data were Box-Cox transformed (Gurka *et al.*, 2006) to conform to the normality assumption for linear mixed-effects analysis and analyzed by linear mixed-effects analysis using *lme4* package (Bates *et al.*, 2015) in *R* (R Core Team, 2023). For assessing *Q*1 (i.e., the scope of pre-boundary lengthening and the effects of focus position and final AP length on it), the dependent variables of the models were formation duration and release duration, respectively. For each of the dependent variables, separate linear mixed-effects models were fitted for each segment (C4, C3, C2, C1, and C0). The independent variables were IP position (IP-final, IP-medial), final NP length (long, short), and focus location (penultimate, ultimate). The maximal models included three-way interactions, and *drop1* function in *lmerTest* package (Kuznetsova *et al.*, 2017) was used to determine which interactions and/or factors were to be eliminated to yield the optimal model. The maximal model included random intercept of speaker and IP length (longer, shorter IPs) and random slopes of IP position, focus location, and final NP length as long as the model converged. In case of significant interactions, pairwise comparisons were assessed by the *emmeans* package (Lenth, 2020) with Bonferroni adjustment. To answer *Q*2 (i.e., the displacement and velocity profile of IP boundaries), we used the same model structure as described above for *Q*1, but the dependent variables were displacement and peak velocity. For the purpose of the present study, results that are directly related to the research questions, i.e., main effects of IP position and/or interactions involving it, are reported. To understand the relationship between kinematic dimensions and the dynamical systems that might account for them and also assess the contradictory predictions in terms of the velocity profile (i.e., faster vs slower IP-final gestures in Korean) as targeted by *Q*3, a second set of linear mixed-effects models were fitted. The data corresponded to the stimuli with short ultimate NPs such that C4 would coincide with the beginning of the AP when the ultimate NP is focused. As the focus

https://doi.org/10.1121/10.0026081

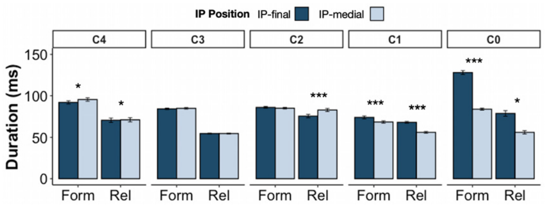

FIG. 4. (Color online) Formation and release duration (in ms) of the test C gestures (C4, C3, C2, C1, and C0) in /nɛmaŋminam/ "*a handsome guy from Nemang*" (as a reminder, /ŋ/ is excluded from the analysis) by IP position. C0 is the consonantal gesture immediately adjacent to the IP boundary and C4 is the most distant from it. Asterisks mark significance from the model with Box-Cox transformed data. ($^{***} = p < 0.001$, $^{**} = p < 0.01$, $^{*} = p < 0.05$.)

effect in Seoul Korean is found to be the strongest at the initial syllable of the focused linguistic unit and decreases with distance (Jang, 2023), the long ultimate NPs would not be suitable for examining the focus effect as the measured interval starts from the fourth syllable of the NP. The models included Box-Cox transformed (Gurka *et al.*, 2006) formation duration, displacement, and peak velocity as dependent variables and the following continuous predictors, respectively, following the methodology applied in Katsika and Tsai (2021): (a) stiffness (i.e., normalized peak velocity over displacement) for the models of formation duration and release duration, (b) peak velocity for the models of displacement, and (c) displacement for the models of peak velocity. To decode the locus of focus- vs boundary-related effects over the test word, IP position (IP-final, IP-medial), focus location (penultimate, ultimate), and consonant position (C4, C3, C2, C1, and C0) were also included as fixed factors. Random intercepts of speaker, trajectory, and IP length and random slopes of IP position, focus location, and consonant position were included in the model as long as the model converged.

III. RESULTS

A. Scope of phrase-final lengthening

Figure 4(a) summarizes the main effects of IP position on formation and release duration in each gesture, demonstrating the general scope of phrase-final lengthening regardless of other focus- or NP-length-related factors. Phrase-final lengthening affected the C gestures of the phrase-final syllable, extending to C1 formation [C0 formation, $F_{(1,6)} = 53.2$, $p < 0.001$; C0 release, $F_{(1,5)} = 12.3$, $p < 0.05$; C1 formation, $F_{(1,797)} = 32.0$, $p < 0.001$; C1 release, $F_{(1,829)} = 525.0$, $p < 0.001$]. That is, the effect of IP-final lengthening goes beyond the final rhyme to the onset of the final syllable. Lengthening was greater in the coda (C0) as compared to the onset (C1) with 42 and 18 ms of lengthening, respectively, in the formation and release of C0, but 5 ms and 11 ms, respectively, for the formation and release of C1. In parallel to lengthening, effects of shortening were observed at earlier gestures of the PWd: C2 release [$F_{(1,832)} = 45.6$, $p < 0.001$], which is in the onset of the syllable before lengthening begins, and C4 formation [$F_{(1,5)} = 8.8$, $p < 0.05$] and release [$F_{(1,15)} = 5.2$, $p < 0.05$], which compose the onset of the word-initial syllable, are shorter IP-finally as compared to IP-medially. Similar findings of pre-boundary shortening have been reported previously in the literature, albeit scarcely (English, Byrd *et al.*, 2006; Greek, Katsika, 2016).

Figure 5 summarizes the results of pairwise comparisons for significant IP position and focus location interactions. Significant two-way interactions between IP position and focus location (penultimate, final) were detected on the measures of formation duration [$F_{(1,823)} = 6.7$, $p < 0.01$] for C0, i.e., the boundary-adjacent consonant [Fig. 5(A)]. C0 gesture was longer in words medially in IPs with final as opposed to penultimate focus ($p < 0.05$). This, in turn, indicates that C0 gestures undergo stronger boundary effects (i.e., longer gestures) when focus was on the penultimate as opposed to the ultimate NP, as illustrated in Fig. 5(A).

Final NP length also interacted with IP position, as plotted and summarized in Fig. 5(B). Such interactions were detected on C0's formation duration [$F_{(1,821)} = 12.7$, $p < 0.001$]. According to the pairwise comparisons, this interaction was the result of C0 formation being longer when belonging in long as opposed to short NPs IP-medially ($p < 0.001$). IP-finally, there was no difference in terms of formation duration between the long and short final NP conditions of C0. Taken together, this means that the amount of IP-final lengthening on C0 was greater in short than in long final NPs. In earlier consonants, IP position and final NP length interactions were significant in C2 release [$F_{(1,832)} = 16.1$, $p < 0.001$] and C4 formation [$F_{(1,802)} = 8.5$, $p < 0.01$]. The interactions were related to the IP-related shortening or compression effect, which was systematically observed in the short but not in the long final NPs. Specifically, C2 release and C4 formation was significantly longer IP-medially as compared to IP-finally only in short final NPs ($p < 0.001$, in C2 release and C4 formation). In long final NPs, IP-related shortening was observed only in C2 release (*marginally significant*,

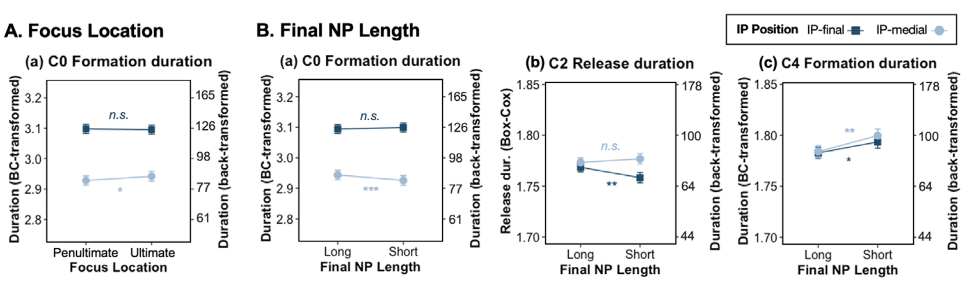

FIG. 5. (Color online) (A) Predicted duration for significant IP position × focus location interaction in (a) C0 formation duration and (B) predicted duration for significant IP position × final NP length interactions in (a) C0 formation duration, (b) C2 release duration, and (c) C4 formation duration are shown (footnote 3). ($^{***} = p < 0.001$, $^{**} = p < 0.01$, tr. $= 0.05 < p < 0.07$, n.s. $= p > 0.09$.)

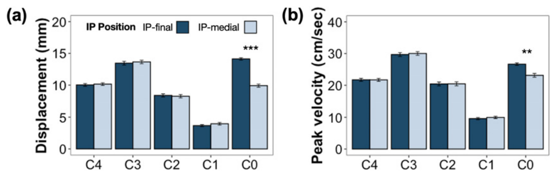

FIG. 6. (Color online) (a) Displacement (in mm) and (b) peak velocity (in cm/s) of the test C gestures (C4, C3, C2, C1, and C0) in /nɛmaŋminam/ "*a handsome guy from Nemang*" (as a reminder, /ŋ/ is excluded from the analysis) by IP position are shown. C0 is the consonantal gesture immediately adjacent to the IP boundary and C4 is the most distant from it. Asterisks mark significance from the model with Box-Cox transformed data. ($^{***} = p < 0.001$, $^{**} = p < 0.01$.)

$0.05 < p < 0.07$), and not in C4 formation ($p > 0.09$). Note that C2 and C4 are the initial consonants of each noun in the compound noun that consist of the final PWd (/nɛmaŋ/ "Nemang (name of a village)" + /minam/ "a handsome guy" in /nɛmaŋminam/ "*a handsome guy from Nemang*").

In sum, the scope of IP-final lengthening extends to the onset of the phrase-final syllable with greater lengthening observed in the coda as compared to the onset consonant. Focus location systematically affects the IP-final lengthening: The lengthening effect of the IP-final gesture was greater when focus was further away from the IP boundary. The length of the phrase-final syntactic phrase also matters but not as predicted. As a reminder, we expected that the length of the final NP would interact with pre-boundary effects only when that NP would be focused and, thus, would be equal to the final AP of the phrase as well, bearing an initial AP tone that functions as a word demarcation cue (e.g., Kim, 2004; Kim and Cho, 2009). Instead, our findings show that regardless of whether the final syntactic phrase bears AP tones or not (or else regardless of whether it is dephrased or not), phrase-final lengthening and boundary-related shortening are greater in short as opposed to long NPs. Interestingly, the shortening effects were observed on the consonants that coincided with the beginning of each noun component of the compound noun. We take these findings to suggest that dephrasing, even if it removes the intonational contour of the dephrased APs, might not eliminate other aspects of the AP boundaries, and a discussion of short vs long NPs is, in fact, a comparison of APs.

B. Displacement and velocity of IP boundaries

Displacement and peak velocity, as illustrated in Figs. 6(a) and 6(b), respectively, showed a main effect of IP only in the IP boundary-adjacent consonant [C0 displacement, $F_{(1,6)} = 119.7$, $p < 0.001$; C0 peak velocity, $F_{(1,5)} = 30.0$, $p < 0.01$]. C0 gestures under the effect of IP boundary were larger and faster than their IP-medial counterparts. The result of faster IP-final gesture is in agreement with previous findings in Korean (Kim *et al.*, 2019; see also Edwards *et al.*, 1991; Byrd and Saltzman, 1998 for English).

Figure 7(a) summarizes the results of pairwise comparisons for significant IP position and focus location interactions on displacement and peak velocity. Significant two-way interactions between IP position and focus location (penultimate, final) were detected on boundary-adjacent C0 displacement [$F_{(1,825)} = 6.6$, $p < 0.05$] and peak velocity [$F_{(1,794)} = 6.9$, $p < 0.01$]. IP-medial C0 gesture was larger and faster when focus location was final as opposed to the penultimate NP (displacement, $p < 0.001$; peak velocity, $p < 0.001$). This, in turn, indicates that C0 gestures show stronger boundary effects (i.e., larger and faster gestures) when focus was on the penultimate as opposed to the ultimate NP, which is similar to the effect found for C0 formation duration in Sec. II A.

Final NP length also interacted with IP position, as plotted and summarized in Fig. 7(B). Interactions were detected on C0 peak velocity [$F_{(1,796)} = 9.4$, $p < 0.01$] and C1 displacement [$F_{(1,792)} = 5.9$, $p < 0.05$]. According to the pairwise comparisons, the interaction on peak velocity was the result of C0 being slower in IP-medial long as opposed to short NPs ($p < 0.01$), which goes hand in hand with it being longer (Sec. II A). In addition, C1 was larger IP-medially as opposed to IP-finally (*marginally significant*, $0.05 < p < 0.08$) only in the short NPs.

In sum, phrase-final C gestures in the coda were larger and faster. Focus location systematically affects the displacement and peak velocity of IP-final gesture: The strengthening effect—larger and faster—of the IP-final gesture was greater when focus was further away from the IP boundary, which is similar to the effect found for IP-final lengthening. Final NP length also affected the kinematic profile such that boundary-related compression is greater in short than in long NPs.

C. Relationships between kinematic parameters modulated by IP boundaries and prominence

This section reports the results of the linear mixed-effects analysis, considering the relationship between

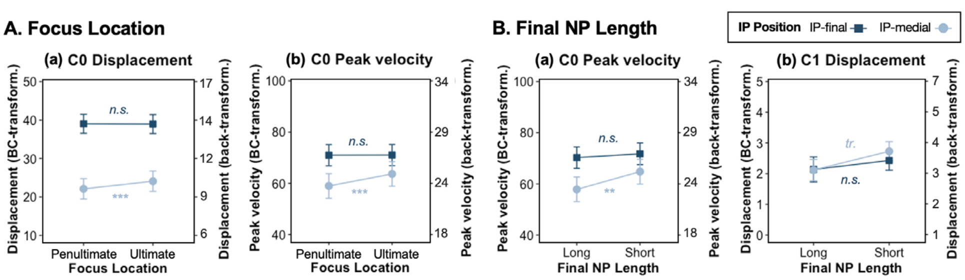

FIG. 7. (Color online) (A) Predicted results are shown for significant IP position × focus location interaction in (a) C0 displacement and (b) C0 peak velocity. (B) Predicted results are displayed for significant IP position × final NP length interactions in (a) C0 peak velocity and (b) C1 displacement. ($^{***} = p < 0.001$, $^{**} = p < 0.01$, tr. $= 0.05 < p < 0.07$, n.s. $= p > 0.09$.)

JASA https://doi.org/10.1121/10.0026081

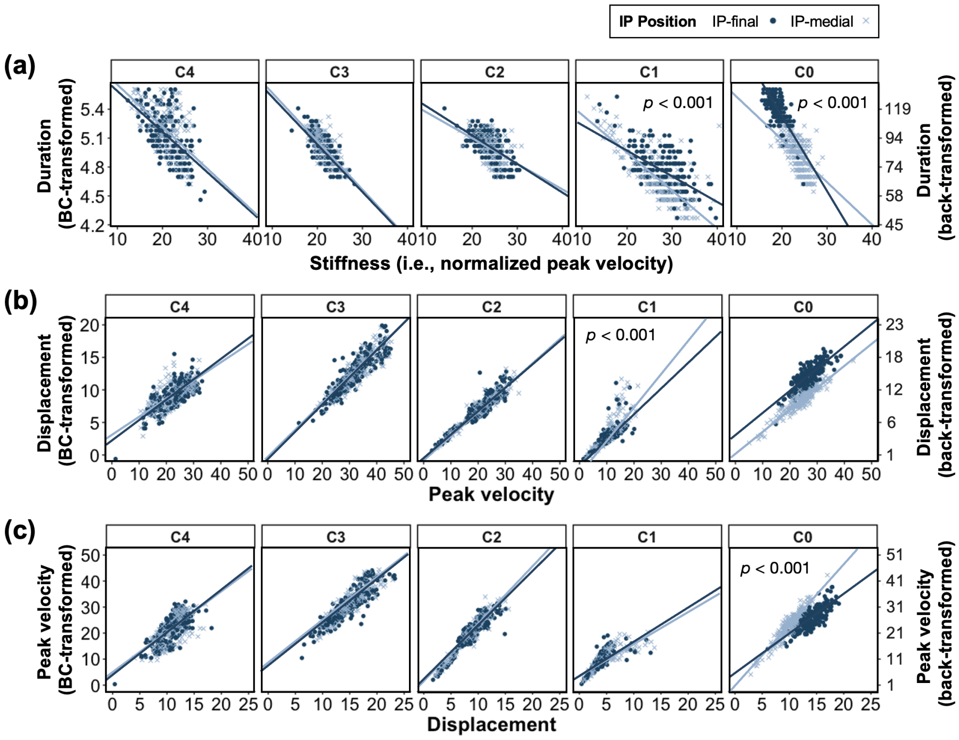

FIG. 8. (Color online) Predicted regression lines are displayed for (a) Box-Cox transformed formation duration and stiffness (i.e., normalized peak velocity), (b) Box-Cox transformed displacement and peak velocity, and (c) Box-Cox transformed peak velocity and displacement as a function of IP position by consonant position. Back-transformed duration (in ms), displacement (in mm), and peak velocity (in cm/s) are indicated on the right y axis. p-values mark significant pairwise comparisons on slopes of each condition.

kinematic dimensions. The predicted relationship between each measurement and continuous factors for each consonant are plotted by IP position in Fig. 8 and focus location in Fig. 9. As predicted, formation duration and stiffness (i.e., normalized peak velocity over displacement) showed an inverse relationship such that formation duration increased with decrease in stiffness [$F_{(1,2052)} = 931.2$, $p < 0.001$]. Three-way interactions of IP position, consonant position, and stiffness were significant for formation duration [$F_{(4,2049)} = 6.7$, $p < 0.001$]. As illustrated in Fig. 8(a), pairwise comparisons on predicted regressions revealed that the slope values of the regressions were significantly different in IP-final vs IP-medial positions for C1 and C0 gestures ($p < 0.001$ for both C's). In particular, the slope of the regressions in IP-final C0 is slightly steeper than that in IP-medial position. That is, given the same decrease in

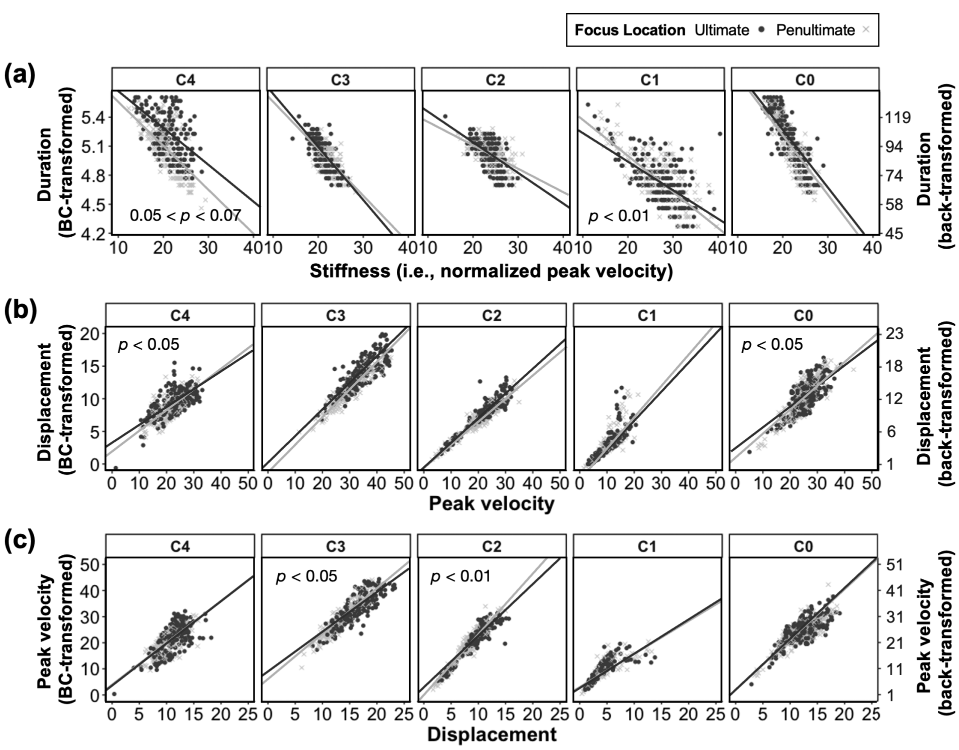

FIG. 9. Predicted regression lines are shown for (a) Box-Cox transformed formation duration and stiffness (i.e., normalized peak velocity), (b) Box-Cox transformed displacement and peak velocity, and (c) Box-Cox transformed peak velocity and displacement as a function of focus location by consonant position. Back-transformed duration (in ms), displacement (in mm), and peak velocity (in cm/s) are indicated on the right y axis. p-values mark significant pairwise comparisons on slopes of each condition.

stiffness, the increase in duration is greater when the gesture is followed by an IP boundary than when it is not. Moreover, for C0 and C1, at average stiffness, formations were longer in IP-final than in IP-medial position ($p < 0.001$ in both C's). C4 was shorter IP-finally than IP-medially ($p < 0.01$), albeit difference in slope was not detected. There was also a marginally significant three-way interaction of focus location, consonant position, and stiffness for formation duration [$F_{(4,2049)} = 2.2$, $0.05 < p < 0.07$]. Pairwise comparison revealed that the regressions for ultimate vs penultimate are marginally different in C4 and significantly different in C1 ($0.05 < p < 0.06$; $p < 0.01$, respectively), as shown in Fig. 9(a). At average stiffness, gestures were longer when focused (i.e., in ultimate) as opposed to unfocused (i.e., in penultimate) in C4 and C0 ($p < 0.001$; $p < 0.01$, respectively), which are the gestures at the edges of the target PWd.

For displacement, there was a positive relationship with peak velocity such that displacement increased with increase in peak velocity, as expected [$F_{(1,2024)} = 3584.9$, $p < 0.001$]. There was significant three-way interaction between IP position, consonant position, and peak velocity for displacement [$F_{(4,2139)} = 4.1$, $p < 0.01$]. As shown in Fig. 8(b), regression lines for IP-final and IP-medial were significantly different in C1 ($p < 0.001$). Although the slopes for IP-final and IP-medial C0 gesture were not different, the regression line for IP-final C0 is above that for IP-medial, indicating that displacement in C0 is consistently larger in the vicinity of an IP boundary regardless of its peak velocity. At average peak velocity, C0 was larger IP-finally than IP-medially ($p < 0.001$), but C1 and C4 were smaller IP-finally than IP-medially ($p < 0.001$; $p < 0.05$, respectively). There was also a significant three-way interaction of focus location, consonant position, and peak velocity for displacement [$F_{(4,2139)} = 2.4$, $p < 0.05$]. As demonstrated in Fig. 9(b), the regression lines for ultimate and penultimate focus location conditions were significantly different in C4 and C0 ($p < 0.05$ for both C's), which are gestures at both edges of the target PWd. At average peak velocity, gestures were larger in ultimate than penultimate conditions in C4, C3, C2, and C0 ($p < 0.05$, $p < 0.001$, $p < 0.001$, and $p < 0.05$, respectively). The results on duration and displacement are in accordance with the findings in Secs. III A and III B, which showed that gestures are longer and larger when preceding an IP boundary. Additionally, in these gestures, we see distinct slopes of IP-final and IP-medial conditions in the relationships between the parameters, i.e., duration to stiffness and displacement to peak velocity.

Finally, an interesting result is detected for the peak velocity measure. There was a significant IP position, consonant position, and displacement interaction for peak velocity [$F_{(4,2135)} = 12.5$, $p < 0.001$]. Whereas peak velocity, in general, shows expected positive relationship with displacement (i.e., peak velocity increases with increase in displacement), as shown in Fig. 8(c), the interaction reveals that steeper slope value for IP-medial C0 gestures than IP-final gestures ($p < 0.001$), suggesting that given the same increase in displacement, the increase in peak velocity is less in IP-final than in IP-medial positions. Thus, at average displacement, peak velocity of IP-final C0 is lower (i.e., slower) than its IP-medial counterpart ($p < 0.001$), confirming the hypothesis stated in Byrd and Saltzman (2003). There was a significant three-way interaction between focus location, consonant position, and displacement [$F_{(4,2135)} = 2.8$, $p < 0.05$], where slopes for regressions are significantly different between ultimate and penultimate conditions for C3 ($p < 0.05$) and C2 ($p < 0.001$) as can be observed in Fig. 9(c). In particular, the slopes of the regressions in C3 and C2 are slightly steeper in penultimate than in ultimate condition. At average displacement, no significant pairwise comparison was detected.

This co-analysis of kinematic dimensions has offered important insights for our understanding of the kinematic profile of prominence and boundaries, as well as how these interact: First, it helped us see that there is slowing IP-finally, as predicted. This slowing interacts with a parallel increase in peak velocity due to increase in displacement (as indicated by the finding that peak velocity increases with displacement), and the strength of this interaction is such that when peak velocity is viewed independently of displacement, as in Sec. III B, IP-final gestures appear faster than their counterparts. The interaction further indicates that IP-final gestures become larger, not solely by virtue of less overlap with adjacent gestures, but also from an independent prosodic source, which is presumably related to prominence (more on this in Sec. IV). This conclusion is enhanced by the patterns driven by focus: Under focus, we observe longer, larger, and faster articulatory movements (Kim *et al.*, 2019). Importantly, although the effect spans over several syllables, it is mainly located on the gestures on both edges of the focused word. Finally, the distinct slopes that are related to either focus or IP boundaries call the attention of attempts to dynamical modeling of prosody on the possibility of different set of values for the constants of targets and stiffness by prosodic position, pertaining also to the discussion of how continuous vs categorical these positions are. It is also crucial to note that although the current analysis did not account for individual variance, future work examining how speaker-specific signatures of boundary-related modulations emerge would be informative.

IV. DISCUSSION

A. The scope of phrase-final lengthening in Korean

One of the main goals of the present study was to examine the scope of IP boundary lengthening in Korean (*Q*1). Our findings indicate that IP-final gestures are longer than their counterparts in IP-medial positions, as expected (e.g., Byrd *et al.*, 2006; Edwards *et al.*, 1991; Fletcher, 2010). Lengthening extends over the phrase-final syllable, where the rhyme is affected the most (as in, e.g., Byrd *et al.*, 2006; Byrd and Riggs, 2008; Krivokapić, 2007; Katsika, 2016; see also acoustic studies, e.g., Berkovits, 1994; Cambier-Langeveld, 1997; Oller, 1973; Wightman *et al.*, 1992). This pattern slightly diverges from previous work on Korean in which lengthening was found to affect both syllables of disyllabic words (Kim *et al.*, 2019). This difference in pattern most likely arises from the different prosodic and rhythmic structures of the stimuli. In addition to that, it is unclear how much of this discrepancy is caused by natural limitations of this type of work. For example, our results are based on data from seven speakers of Seoul Korean, focusing on alveolar and labial consonants of utterances produced in a question intonation. Extensions of this line of research should consider a wider range of segment combinations, intonation contours, and population. Regardless, we can conclude that phrase-final lengthening in Korean influences a relatively long stretch of speech, extending over at least one syllable.

B. The interaction of phrase-final lengthening with phrase-level prominence

As part of $Q1$, we examined whether the IP-final lengthening is fine-tuned by the position of phrase-level prominence. The position of focus in the phrase affects the degree of lengthening. The effect is greater when focus is further away from the boundary as opposed to when it is boundary adjacent. The direction of the effects is like that of lexical stress in stress languages in which IP-final lengthening is greater the earlier the stress in the phrase-final word (Katsika, 2016; Katsika *et al.*, 2014). In previous work, the effect of stress was accounted for through a dual coordination of the pi-gesture with the final syllable's V gesture and the mu-gesture triggering lexical stress by making the stressed gestures longer, larger, and faster [see Fig. 1(b)]. Competition between these two coordination relationships causes the earlier initiation of IP boundary-related events in nonfinal-stressed words compared to final-stressed words.

This account can be extended to Korean. Focus in Korean is marked by kinematic patterns similar to those of lexical stress and phrasal accent in head-prominence languages, i.e., by longer, larger, and faster gestures (Jang and Katsika, 2023; see also Sec. III C). The effect is stronger at the left edge of the focused AP and wanes out with distance from it. Note that the profile of stress is similar: Stressed syllables are affected the most while effects also spillover to neighboring syllables. It can, thus, be assumed that the effects of focus in Korean can rise from a mu-gesture coordinated in-phase with the focused AP's initial syllable because the kinematic effects of prominence are first and mainly observed there. In turn, the pi-gesture, i.e., the modulation gesture associated with the IP boundary, has a dual coordination, i.e., in-phase coordination with the phrase-final syllable (as the whole syllable is affected) and, possibly, anti-phase coordination with the mu-gesture (the reader is directed to Katsika and Tsai, 2021 for a discussion of the possible coordination types and their strengths of pi-/mu-gestures and explanatory power on typology). Competition between these two coordination relationships causes *attraction* of the pi-gesture toward the focus-marking mu-gesture. As the pi-gesture is attracted away from the IP boundary, it is activated earlier and, consequently, it reaches its level of maximum activation earlier with respect to the constriction gestures, which means that a larger part of the affected constriction gestures would overlap with that maximum level and, thus, lengthen more (see Fig. 10).

Furthermore, the analyses in Sec. III C show that focus has a second, very local, domain of influence: The initial and final gestures of the focused AP are also systematically affected on all three kinematic dimensions of duration, displacement, and velocity (see Dimitrova and Turk, 2012 for effects of focus on word edges in English). On the basis of these findings, which present systematic kinematic effects at the beginning and end of the focused AP, we propose that focus in Korean is related to two mu-gestures, one on each edge of the focused AP. The question would then be *why two mu-gestures?*. The hypothesis that we put forward is that the ultimate function of prosodic modulation gestures (mu-/pi-gestures) is the incorporation of tone and intonation in the utterance; the kinematic effects emerge from the process. This is essentially what pi-gestures do in the account of stress-boundary interactions: They trigger boundary tone gestures (Katsika *et al.*, 2014; Katsika, 2016). Two mu-gestures, each on both sides of the AP, correspond well with the intonational contour of Korean APs, which posits initial and final AP tones, and because AP is the level in which focus is marked in Korean, the edges of the focused constituent coincide with AP boundaries with dephrasing occurring in between. That the right AP boundary coincides with the IP boundary results in overlap between the AP-final mu- and pi-gestures, yielding what is described as "overriding" of the AP tone by the boundary tone, resulting from coarticulation of the two tones and, also, the conflicting effects on velocity, which were detected in Sec. III C: The pi-gesture slows down motion during the same time that the mu-gesture accelerates it. Thus, depending on the point of view one takes via the methods that they apply, IP-adjacent gestures might appear either slower or faster, with both conclusions being accurate (compare, e.g., Sec. III B to Sec. III C). We argue that the analysis applied in Sec. III C is a tool that can assist in separating these concurrent opposing forces.

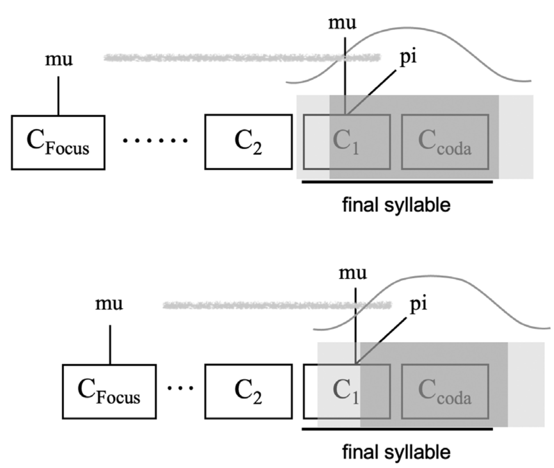

FIG. 10. Schematic representation of the dual coordination of pi-gestures with the onset C gesture of the phrase-final syllable and the focus-marking mu-gesture. The gray-shaded boxes indicate the effect of pi-gesture with the darker box representing stronger activation. Only C gestures are presented because V gestures have not been analyzed yet.

Therefore, in the account provided here, the kinematic signature of prominence found in the boundary-adjacent gestures is attributed to intonational dimensions in this case of the AP level and may not be a feature specific to edge-prominence marking. For instance, the larger and faster movements that have been reported by some studies for English (Edwards *et al*, 1991; Byrd and Saltzman, 1998) might be related to the phrase accent of ips. Under this hypothesis, languages that use AP to mark focus but do not involve an AP tone on the right edge of the AP are not expected to show these prominence-like kinematic effects at the IP boundary. That mu-/pi-gestures might be associated with phrase-level tonal events in Korean is further supported by the fact that boundary tones have been found to undergo a similar influence by focus as phrase-final lengthening. In other words, boundary tones in Korean are initiated earlier the earlier the focus is in the IP (Jang and Katsika, 2022). It is, however, important to note that this is specific to a boundary tone contour, i.e., *H%*, examined in the current study. Further studies are called for to further test and expand the proposed account.

C. The interaction of phrase-final lengthening and word demarcation in Korean

Another component of *Q*1 examined whether the word demarcation function of the phrase-final AP affects phrase-final lengthening. The length, as indexed by the number of syllables, of focused phrase-final APs did not further affect the scope of boundary, suggesting that the demarcation function of the AP-initial accent might not be at play here. Instead, robust effects of the length of focused and dephrased APs were detected with the shorter APs presenting the more extreme effects. These patterns indicate that dephrased APs, despite losing their tonal specification, might maintain other prosodic dimensions (in terms of kinematics and general timing patterns) that still characterize them as APs. Assuming that dephrased APs maintain their AP status allows us to account for the observed boundary effects through prosodic hierarchy because we expect that lengthening decreases with size of prosodic constituent (Lehiste, 1972). This would also work well within a view of prosody as a network of nested oscillators, in which it is not only the case that higher-level oscillators affect lower-level oscillators, but also vice versa, i.e., lower-level oscillators, AP in this case, affect the duration of higher oscillators (e.g., Nam *et al.*, 2008). Evidence for this type of relationship also comes from the PWd level, which is found to be the domain of boundary-related shortening, as further discussed below.

D. IP boundary-related shortening

A small but significant IP-related shortening effect was detected on the earlier C gestures of the phrase-final word. IP boundary-related shortening effects have previously been reported in the literature on either side of the boundary (Byrd *et al.*, 2006; Katsika, 2016; Kim *et al.*, 2017). Post-boundary, the shortening effect has been shown to be systematic and characterized as compensatory in nature. Pre-boundary, shortening has been found to be less systematic and speaker specific (Katsika, 2016). Furthermore, in Greek, the location of the shortening effect was also related to the position of lexical stress in the phrase-final word (Katsika, 2016) and, for that reason, the effect was considered a by-product of the coordination between the stress-instantiating mu-gesture and the pi-gesture, which attracted the two gestures toward each other. In our data, the effect is consistently found on the initial consonant of each noun in the compound noun that consist of the final PWd (/n̲ɛmaŋm̲inam/ "*a handsome guy from Nemang* (name of a village)"). Another factor associated with boundary-related shortening is the length of the AP (dephrased or not), with the effect being more enhanced in shorter as opposed to longer APs, a phenomenon that falls in the account provided in Sec. IV C. These results in combination may suggest that the lexical items that feed into the domain of the PWd are accessible early to prosodic speech planning (cf. Keating and Shattuck-Hufnagel, 2002 vs Levelt, 1989).

E. The kinematic profile of prosodic boundaries

As a response to Q2 that assesses the displacement and velocity profile of IP boundaries in Seoul Korean, we can conclusively say that gestures at the IP boundary are larger. Whether they are faster or slower, though, depends on one's methodological point of view. Sections III C and IV B present arguments that this mismatch on the velocity profile stems from concurrent, competing demands on velocity in IP-final gestures. These gestures are affected simultaneously by a mu-gesture to trigger the AP tone and a pi-gesture to trigger the boundary tone. This competition yields gestures with velocities that increase as their target increases, but the rate of this increase is slowed down in IP-final positions. This observation also has another important consequence: The enlargement of gestures at boundaries cannot be solely the outcome of less inter-gestural overlap. The mu-gesture associated with the AP-final tones may better account for these larger movements.

V. CONCLUSION

The present study examined the articulatory correlates of IP-final boundary in Seoul Korean using electromagnetic articulography (EMA) with the goal to investigate the scope of the boundary-related effect and how it interacts with other grammatical domains and/or prosodic levels. At IP boundaries, gestures in Korean are, in general, longer, larger, and, considering their displacement, slower as well. Pre-boundary lengthening systematically affected the phrase-final syllable with focus location fine-tuning its manifestation. The amount of IP-final lengthening was greater the further away focus was from the boundary. Furthermore, the effect was sensitive to the size of constituents at the AP and PWd levels. The results have implications for our understanding of the prosodic effects exerted by boundaries and phrase-level prominence. They suggest that effects of boundary-related lengthening reported in previous studies are not driven by word-level

JASA
https://doi.org/10.1121/10.0026081

prosody *per se* (Katsika, 2016; Katsika and Tsai, 2021), but they involve phrase-level prominence as well, presumably because of the requirement for incorporation of phrase-level tones that IP boundaries and focus-marking pose.

ACKNOWLEDGMENTS

We would like to thank the Korean speakers who participated in the study. We are also extremely grateful to the associate editor and three anonymous reviewers for their constructive feedback and Louis Goldstein, Matthew Gordon, and John Du Bois for their comments on earlier versions of the manuscript. This work was supported by the National Science Foundation Grant No. 1551428.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Ethics Approval

The authors hereby state that this project has obtained approval from the Institutional Review Board (IRB) to ensure compliance with ethical guidelines.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

[1]Tonogenetic sound changes have been observed among younger speakers, where the voice onset time distinction between aspirated and lenis stops merges, and *F*0 becomes the primary cue for the contrast (higher *F*0 for aspirated and lower *F*0 for lenis; Kang, 2014; Bang *et al.*, 2018). However, these are limited to specific segments and phrasal positions, i.e., stop and affricate consonants in AP-initial positions.

[2]Here, the term *target* does not have the dynamical meaning defined in Sec. I E, according to which targets can only be approached and not achieved. Rather, the term here is synonymous to constriction.

[3]We plot the predicted values of the linear mixed-effects models following the guideline in Gries (2021, p. 308), which advises focusing on the predictions of the model when summarizing a regression model to control for all variables rather than using descriptive plots of observed data points. In certain cases, such as mixed-effects models, the two types of values will not be identical, and the predicted values are likely to be more useful and realistic (Gries, 2021, p. 308).

Bang, H. Y., Sonderegger, M., Kang, Y., Clayards, M., and Yoon, T. J. (**2018**). "The emergence, progress, and impact of sound change in progress in Seoul Korean: Implications for mechanisms of tonogenesis," J. Phonet. **66**, 120–144.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (**2015**). "Fitting linear mixed-effects models using lme4," J. Stat. Soft. **67**(1), 1–48.

Beckman, M., Edwards, J., and Fletcher, J. (**1992**). "Prosodic structure and tempo in a sonority model of articulatory dynamics," in *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, edited by G. Docherty and D. Ladd (Cambridge University Press, Cambridge, UK), pp. 68–89.

Beckman, M. E., and Pierrehumbert, J. B. (**1986**). "Intonational structure in Japanese and English," Phonol. Yearb. **3**, 255–309.

Berkovits, R. (**1994**). "Durational effects in final lengthening, gapping, and contrastive stress," Lang. Speech **37**(3), 237–250.

Browman, C. P., and Goldstein, L. M. (**1992**). "Articulatory phonology: An overview," Phonetica **49**, 155–180.

Byrd, D. (**2000**). "Articulatory vowel lengthening and coordination at phrasal junctures," Phonetica **57**(1), 3–16.

Byrd, D., Kaun, A., Narayanan, S., and Saltzman, E. (**2000**). "Phrasal signatures in articulation," in *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, edited J. Kingston and M. E. Beckman (Cambridge University Press, Cambridge, UK), pp. 70–87.

Byrd, D., Krivokapić, J., and Lee, S. (**2006**). "How far, how long: On the temporal scope of phrase boundary effect," J. Acoust. Soc. Am. **120**(3), 1589–1599.

Byrd, D., Lee, S., Riggs, D., and Adams, J. (**2005**). "Interacting effects of syllable and phrase position on consonant articulation," J. Acoust. Soc. Am. **118**(6), 3860–3873.

Byrd, D., and Riggs, D. (**2008**). "Locality interactions with prominence in determining the scope of phrasal lengthening," J. Int. Phonet. Assoc. **38**, 187–202.

Byrd, D., and Saltzman, E. (**1998**). "Intragestural dynamics of multiple phrasal boundaries," J. Phonet. **26**, 173–199.

Byrd, D., and Saltzman, E. (**2003**). "The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening," J. Phonet. **31**(2), 149–180.

Cambier-Langeveld, T. (**1997**). "The domain of final lengthening in the production of Dutch," Linguist. Neth. **14**, 13–24.

Campbell, W. N., and Isard, S. D. (**1991**). "Segment durations in a syllable frame," J. Phonet. **19**(1), 37–47.

Cho, T. (**2002**). "The effects of prosody on articulation in English," Ph.D. dissertation, University of California, Los Angeles.

Cho, T. (**2005**). "Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /a,i/ in English," J. Acoust. Soc. Am. **117**(6), 3867–3878.

Cho, T. (**2006**). "Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English," in *Papers in Laboratory Phonology VIII: Varieties of Phonological Competence (Phonology and Phonetics)*, edited by L. Goldstein, D. H. Whalen, and C. T. Best (Mouton de Gruyter, Berlin, Germany), pp. 519–548.

Cho, T., and Jun, S.-A. (**2000**). "Domain-initial strengthening as featural enhancement: Aerodynamic evidence from Korean," Chicago Linguist. Soc. **36**, 31–44.

Cho, T., and Keating, P. (**2001**). "Articulatory and acoustic studies on domain-initial strengthening in Korean," J. Phonet. **29**(2), 155–190.

Cho, T., and Keating, P. (**2009**). "Effects of initial position versus prominence in English," J. Phonet. **37**(4), 466–485.

Cutler, A., and Norris, D. (**1988**). "The role of strong syllables in segmentation for lexical access," J. Exp. Psychol.: Hum. Percept. Perform. **14**(1), 113–121.

Dimitrova, S., and Turk, A. (**2012**). "Patterns of accentual lengthening in English four-syllable words," J. Phonet. **40**(3), 403–418.

Edwards, J., Beckman, M. E., and Fletcher, J. (**1991**). "The articulatory kinematics of final lengthening," J. Acoust. Soc. Am. **89**, 369–382.

Fletcher, J. (**2010**). "The prosody of speech: Timing and rhythm," in *The Handbook of Phonetic Sciences*, 2nd ed., edited by W. J. Hardcastle, J. Laver, and F. E. Gibbon (Blackwell, Oxford, UK), pp. 523–602.

Fougeron, C. (**2001**). "Articulatory properties of initial segments in several prosodic constituents in French," J. Phonet. **29**(2), 109–135.

Fougeron, C., and Keating, P. (**1997**). "Articulatory strengthening at edges of prosodic domains," J. Acoust. Soc. Am. **101**, 3728–3740.

Goldstein, L., Byrd, D., and Saltzman, E. (**2006**). "The role of vocal tract gestural action units in understanding the evolution of phonology," in *Action to Language via the Mirror Neuron System*, edited by M. A. Arbib (Cambridge University Press, Cambridge, UK), pp. 215–249.

Gries, S. T. (**2021**). *Statistics for Linguistics with R: A Practical Introduction*, 3rd ed. (Mouton de Gruyter, Berlin, Germany).

Gurka, M. J., Edwards, L. J., Muller, K. E., and Kupper, L. L. (**2006**). "Extending the Box–Cox transformation to the linear mixed model," J. R. Statist. Soc. Ser. A: Stat. Soc. **169**(2), 273–288.

Gussenhoven, C., and Chen, A. (eds). (**2020**). *The Oxford Handbook of Language Prosody* (Oxford University Press, Oxford, UK).

Hawkins, S. (**1992**). "An introduction to task dynamics," in *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, edited by G. Docherty and D. Ladd (Cambridge University Press, Cambridge, UK), pp. 9–25.

Hayes, B. (**1989**). "The prosodic hierarchy in meter," in *Phonetics and Phonology: Rhythm and Meter*, edited by P. Kiparsky and G. Youmans (Academic, New York), Vol. 1, pp. 201–259.

Iskarous, K., and Pouplier, M. (**2022**). "Advancements of phonetics in the 21st century: A critical appraisal of time and space in Articulatory Phonology," J. Phonet. **95**, 101195.

Jang, J. (**2023**). "Functions of prosodic boundaries in an edge-prominence language: Marking focus and phrasing in Seoul Korean," Ph.D. dissertation, University of California, Santa Barbara.

Jang, J., and Katsika, A. (**2019**). "Articulatory characteristics of focus in Korean," J. Acoust. Soc. Am. **146**(4), 3084.
Jang, J., and Katsika, A. (**2022**). "The coordination of boundary tones with constriction gestures in Seoul Korean, an edge-prominence language," in *Proceedings of the 11th International Conference on Speech Prosody 2022*, Lisbon, Portugal (May 23–26, 2022), pp. 145–149.
Jang, J., and Katsika, A. (**2023**). "Edgy articulation: The kinematic profile of accentual phrase boundaries in Seoul Korean," in *Proceedings of the Hanyang International Symposium on Phonetics and Cognitive Sciences of Language (HISPhonCog) 2023*, Seoul, South Korea (May 26–27, 2023), pp. 204–205.
Jeon, H. S., and Nolan, F. (**2017**). "Prosodic marking of narrow focus in Seoul Korean," Lab. Phonol. **8**(1), 2.
Jun, S.-A. (**1993**). "The phonetics and phonology of Korean prosody," Ph.D. dissertation, Ohio State University.
Jun, S.-A. (**1998**). "The accentual phrase in the Korean prosodic hierarchy," Phonology **15**(2), 189–226.
Jun, S.-A. (**2000**). "K-ToBI (Korean ToBI) labelling conventions," Speech Sci. **7**(1), 143–170.
Jun, S.-A. (**2003**). "The effect of phrase length and speech rate on prosodic phrasing," in *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, Spain (August 3–9, 2003), pp. 483–486.
Jun, S.-A. (**2005**). "Korean intonational phonology and prosodic transcription," in *Prosodic Typology*, edited by S.-A. Jun (Oxford University Press, Oxford, UK), pp. 201–229.
Jun, S.-A. (**2014**). "Prosodic typology: By prominence type, word prosody, and macro-rhythm," in *Prosodic Typology II: The Phonology of Intonation and Phrasing*, edited by S.-A. Jun (Oxford University Press, Oxford, UK), pp. 520–539.
Kang, Y. (**2014**). "Voice onset time merger and development of tonal contrast in Seoul Korean stops: A corpus study," J. Phonet. **45**, 76–90.
Katsika, A. (**2009**). "Boundary- and prominence-related lengthening and their interaction," J. Acoust. Soc. Am. **125**, 2572.
Katsika, A. (**2016**). "The role of prominence in determining the scope of boundary-related lengthening in Greek," J. Phonet. **55**, 149–181.
Katsika, A., Krivokapić, J., Mooshammer, C., Tiede, M., and Goldstein, L. (**2014**). "The coordination of boundary tones and their interaction with prominence," J. Phonet. **44**, 62–82.
Katsika, A., and Tsai, K. (**2021**). "The supralaryngeal articulation of stress and accent in Greek," J. Phonet. **88**, 101085.
Keating, P., Cho, T., Fougeron, C., and Hsu, C. (**2004**). "Domain-initial articulatory strengthening in four languages," in *Papers in Laboratory Phonology VI: Phonetic Interpretation* (Cambridge University Press, Cambridge, UK), pp. 143–161.
Keating, P., and Shattuck-Hufnagel, S. (**2002**). "A prosodic view of word form encoding for speech production," UCLA Working Papers in Phonetics, pp. 112–156.
Kim, J., Baek, Y., Cho, T., and Kim, S. (**2019**). "Preboundary lengthening and boundary-related spatial expansion in Korean," in *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, Australia (August 5–9, 2019), pp. 1019–1023.
Kim, S. (**2004**). "The role of prosodic phrasing in Korean word segmentation," Ph.D. dissertation, University of California, Los Angeles.
Kim, S., and Cho, T. (**2009**). "The use of phrase-level prosodic information in lexical segmentation: Evidence from word-spotting experiments in Korean," J. Acoust. Soc. Am. **125**(5), 3373–3386.
Kim, S., Jang, J., and Cho, T. (**2017**). "Articulatory characteristics of preboundary lengthening in interaction with prominence on tri-syllabic words in American English," J. Acoust. Soc. Am. **142**(4), EL362–EL368.
Kohler, K. J. (**1983**). "Prosodic boundary signals in German," Phonetica **40**(2), 89–134.
Krivokapić, J. (**2007**). "The planning, production and perception of prosodic structure," Ph.D. dissertation, University of Southern California, Los Angeles, CA.
Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (**2017**). "lmerTest Package: Tests in linear mixed effects models," J. Stat. Soft. **82**(13), 1–26.
Lehiste, I. (**1972**). "The timing of utterances and linguistic boundaries," J. Acoust. Soc. Am. **51**(6B), 2018–2024.
Lenth, R. (**2020**). "emmeans: Estimated marginal means, aka least-squares means," R package version 1.4.5, available at https://CRAN.R-project.org/package=emmeans (Last viewed April 12, 2024).
Levelt, W. J. M. (**1989**). *Speaking: From Intention to Articulation* (MIT Press, Cambridge, MA).
Mücke, D., and Grice, M. (**2014**). "The effect of focus marking on supralaryngeal articulation—Is it mediated by accentuation?," J. Phonet. **44**, 47–61.
Munhall, K. G., Ostry, D. J., and Parush, A. (**1985**). "Characteristics of velocity profiles of speech movements," J. Exp. Psychol.: Hum. Percept. Perform. **11**, 457–474.
Nam, H., Saltzman, E., Krivokapić, J., and Goldstein, L. (**2008**). "Modeling the durational difference of stressed vs. unstressed syllables," in *Proceedings of the 8th Phonetic Conference of China*, Beijing, China (April 18–20, 2008).
Nespor, M., and Vogel, I. (**1986**). *Prosodic Phonology* (Foris, Dordrecht, Netherlands).
Oller, K. D. (**1973**). "The effect of position in utterance on speech segment duration in English," J. Acoust. Soc. Am. **54**, 1235–1247.
Ostry, D., and Munhall, K. (**1985**). "Control of rate and duration of speech movements," J. Acoust. Soc. Am. **77**, 640–648.
R Core Team. (**2023**). "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria, available at https://www.R-project.org/ (Last viewed April 12, 2024).
Roon, K. D., Gafos, A. I., Hoole, P., and Zeroual, C. (**2007**). "Influence of articulator and manner on stiffness," in *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken, Germany (August 6-10, 2007), pp. 409–412.
Saltzman, E., and Byrd, D. (**2000**). "Task-dynamics of gestural timing: Phase windows and multifrequency rhythms," Hum. Mov. Sci. **19**(4), 499–526.
Saltzman, E., Nam, H., Krivokapić, J., and Goldstein, L. (**2008**). "A task-dynamic toolkit for modeling the effects of prosodic structure on articulation," in *Proceedings of the 4th International Conference on Speech Prosody 2008*, Campinas, Brazil (May 6–9, 2008), pp. 175–184.
Selkirk, E. (**1984**). *Phonology and Syntax: The Relation between Sound and Structure* (MIT Press, Cambridge, MA).
Seo, J., Kim, S., Kubozono, H., and Cho, T. (**2019**). "Preboundary lengthening in Japanese: To what extent do lexical pitch accent and moraic structure matter?," J. Acoust. Soc. Am. **146**(3), 1817–1823.
Shattuck-Hufnagel, S., and Turk, A. (**1996**). "A prosody tutorial for investigators of auditory sentence processing," J. Psycholinguist. Res. **25**(2), 193–247.
Shin, S., Kim, S., and Cho, T. (**2015**). "What is special about prosodic strengthening in Korean: Evidence in lingual movement in V#V and V#CV," in *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, UK (August 10–14, 2015).
Tsai, K. (**2023**). "Edgy prosody: An articulatory investigation of the role of lexical pitch accent in Tokyo Japanese boundary marking," Ph.D. dissertation, University of California, Santa Barbara.
Tsai, K., and Katsika, A. (**2020**). "Pitch accent and phrase boundaries: Kinematic evidence from Japanese," in *Proceedings of 10th International Conference on Speech Prosody 2020*, Tokyo, Japan (May 25–28, 2020), pp. 215–219.
Turk, A. E., and Shattuck-Hufnagel, S. (**2007**). "Multiple targets of phrase-final lengthening in American English words," J. Phonet. **35**(4), 445–472.
Tyler, M. D., and Cutler, A. (**2009**). "Cross-language differences in cue use for speech segmentation," J. Acoust. Soc. Am. **126**(1), 367–376.
Vaissière, J. (**1983**). "Language-independent prosodic features," in *Prosody: Models and Measurements*, edited by A. Cutler and D. R. Ladd (Springer, Berlin), pp. 53–66.
Warner, N., Otake, T., and Arai, T. (**2010**). "Intonational structure as a word-boundary cue in Tokyo Japanese," Lang. Speech **53**(1), 107–131.
Welby, P. (**2007**). "The role of early fundamental frequency rises and elbows in French word segmentation," Speech Commun. **49**(1), 28–48.
White, L. (**2002**). *English Speech Timing: A Domain and Locus Approach* (The University of Edinburgh, Edinburgh, UK).
Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., and Price, P. J. (**1992**). "Segmental durations in the vicinity of prosodic phrase boundaries," J. Acoust. Soc. Am. **91**, 1707–1717.
Wilbur, R. B. (**2009**). "Effects of varying rate of signing on ASL manual signs and nonmanual markers," Lang. Speech **52**(2-3), 245–285.