University of California
Santa Barbara

Functions of prosodic boundaries in an edge-prominence language: Marking focus and phrasing in Seoul Korean

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Linguistics

by

Jiyoung Jang

Committee in charge:

Professor Argyro Katsika, Chair
Professor Louis Goldstein, University of Southern California
Professor Matthew Gordon
Professor John W. Du Bois

September 2023

The Dissertation of Jiyoung Jang is approved.

Professor Louis Goldstein

Professor Matthew Gordon

Professor John W. Du Bois

Professor Argyro Katsika, Committee Chair

September 2023

Functions of prosodic boundaries in an edge-prominence language: Marking focus and phrasing in Seoul Korean

by

Jiyoung Jang

To my family

Acknowledgements

I would like to dedicate this section to those who have guided and accompanied me throughout this intellectual journey. While no words can express my appreciation enough, I hope these acknowledgments will show at least a portion of it.

First and foremost, I would like to thank my advisor, Professor Argyro Katsika, for her guidance and support throughout my time at UCSB. I could not have asked for a better advisor, as she was not only an advisor, but a mentor, a role-model, and a friend. This work, as well as my growth as a scholar and a scientist, would not have been possible without her dedication, patience, and advice.

I would also like to thank Professor Louis Goldstein for his invaluable feedback and support. His work on Articulatory Phonology has been the foundation of my research and this dissertation. Interacting and working with him have been a great honor.

I thank Professor Matthew Gordon and Professor John W. Du Bois for serving as my committee members from the beginning to the end of my Ph.D. program as well as for their flexibility and generosity with their time. Their feedback and support have been crucial in the completion of this dissertation.

I would also like to extend my sincere gratitude to Professor Taehong Cho, my former advisor at Hanyang University. He was the reason I stepped into the world of Phonetics and the reason I was able to come this far. I was extremely fortunate to have met him as my mentor who has taught me to become a phonetician. I am also deeply thankful to Professor Sahyang Kim, as her contribution to my intellectual as well as personal growth is invaluable. Her scholarly inspiration, encouragement, and support have been instrumental in shaping my academic journey.

Special thanks go to my past and present lab members, Dr. Karen Tsai, Dr. Caroline Crouch, and Sherry Chien, for being a great companion throughout the years. I will miss the times we spent in the lab, as Caroline wrote, "working and not working". I would also like to thank all the members of the UCSB Phonetics Circle and my fellow UCSB Linguistics graduate students.

In this long journey, I have made so many dear friends who made it possible for me to enjoy the life in Santa Barbara. I would especially like to thank the "Koinonia", Dr. Jaebak Kim, Dr. Minwook Park, Dr. Yu-Jin Choi, Dr. Wonjin Yoon, Dr. Seungchan Lee, Dr. Changyun Yoo, Dr. Wonho Lee, and Hyeonsu Shin, for the countless mornings, days, and nights spent together. I thank Dr. Saejin Oh and Dr. Changmin Lee for their great company. I also thank the Korean Tennis & Surf crew as well as other members of the Korean Graduate Student Association. Many thanks go to my friends in Korea who have been supportive and encouraging despite the geographical distance.

Last but not least, I would like to express my deepest gratitude to my family. To my dad, who I will always look up to. To my mom, who is the most loving and caring person in the world. To my two brilliant sisters, who have always been by my side. To Sihoon, who took the courage to fly 9,606 kilometers to Santa Barbara with me.

The studies reported in this dissertation were supported by National Science Foundation (NSF) grant #1551428 and UCSB Academic Senate grant awarded to Argyro Katsika and UCSB Humanities & Social Sciences Research grant awarded to the author.

Curriculum Vitæ
Jiyoung Jang

Education

2023	Ph.D. in Linguistics, University of California, Santa Barbara, CA
2018	M.A. in English Linguistics, Hanyang University, Seoul
2015	B.A. in English Language Education (*Summa cum laude*, early graduation), Hanyang University, Seoul

Professional Experience

2018-2023	Graduate research assistant, Speech, Prosody, and Articulatory Kinematics Laboratory, UCSB
2022	Teaching associate (Instructor), Linguistics department, UCSB
2019-2023	Teaching assistant, Linguistics department, UCSB
2014-2018	Graduate/undergraduate research assistant, Hanyang Institute for Phonetics and Cognitive Sciences of Language, Hanyang University

Publications

Journal articles

Jang, J., Kim, S., & Cho, T. (2022, online first). Prosodic structural effects on non-contrastive coarticulatory vowel nasalization in L2 English by Korean learners. *Language and Speech.* https://doi.org/10.1177/00238309221108657

Jang, J., Kim, S., & Cho, T. (2018). Focus and boundary effects on coarticulatory vowel nasalization in Korean with implications for cross-linguistic similarities and differences. *The Journal of the Acoustical Society of America*, 144(1), EL33-39.

Kim, S., **Jang, J.**, & Cho, T. (2017). Articulatory characteristics of preboundary lengthening in interaction with prominence on tri-syllabic words in American English. *The Journal of the Acoustical Society of America*, 142(4), EL362-368.

Peer-reviewed proceedings

Jang, J., & Katsika, A. (2023). Articulatory manifestation of focus-induced prominence in Seoul Korean. *Proceedings of the 20th International Congress of Phonetic Sciences*, 1686-1690.

Katsika, A., **Jang, J.**, Krivokapić, J., Goldstein, L., & Saltzman, E. (2023). A hierarchy of prominence: The production and perception of focus in American English. *Proceedings of the 20th International Congress of Phonetic Sciences*, 1696-1700.

Gordon, M., Katsika, A., Chien, S., **Jang, J.**, & Lai, R. K.-Y. (2023). Speech rhythm metrics: a typological survey. *Proceedings of the 20th International Congress of Phonetic Sciences*, 1479-1483.

Jang, J., & Katsika, A. (2022). The coordination of boundary tones with constriction gestures in Seoul Korean, an edge-prominence language. *Proceedings of the 11th International Conference on Speech Prosody*, 145-149.

Jang, J., & Katsika, A. (2020). The amount and scope of phrase-final lengthening in Seoul Korean. *Proceedings of the 10th International Conference on Speech Prosody*, 270-274.

Katsika, A., **Jang, J.**, Krivokapić, J., Goldstein, L., & Saltzman, E. (2020). The role of focus in accentual lengthening in American English: Kinematic analyses. *Proceedings of the 10th International Conference on Speech Prosody*, 275-279.

Jang, J., Kim, S., & Cho, T. (2019). Prosodic-structural effects on voice quality associated with Korean three-way stop contrast. *Proceedings of the 19th International Congress of Phonetic Sciences*, 2416-2420.

Joo, H., **Jang, J.**, Kim, S., Cho, T., & Cutler, A. (2019). Prosodic structural effects on coarticulatory vowel nasalization in Australian English in comparison to American English. *Proceedings of the 19th International Congress of Phonetic Sciences*, 835-839.

Choi, J., Cho, T., Kim, S., Baek, Y., & **Jang, J**. (2015). Phonetic encoding of coda voicing contrast and its interaction with information structure in L1 and L2 speech. *Proceedings of the 18th International Congress of Phonetics Sciences*.

Honors and Awards

2022-2023	Humanities & Social Sciences Research Grant, UCSB
2022	International Phonetic Association (IPA) Student Award
2022	Academic Senate Doctoral Student Travel Grant, UCSB
2018-2020	Scholarship, Office of International Affairs, Hanyang University
2018-2019	Regents Fellowship in Linguistics, UCSB
2016-2017	College of Humanities Research and Education (CORE) Fellowship, Hanyang University
2016	Best Paper Award, Korean Society of Speech Sciences Conference
2015-2017	Scholarship, Hanyang University

Abstract

Functions of prosodic boundaries in an edge-prominence language: Marking focus and phrasing in Seoul Korean

by

Jiyoung Jang

This dissertation investigates the prosodic modulations, i.e., boundary- and prominence-marking, and their interactions on articulation in Seoul Korean via an electromagnetic articulography study. Prosodic structure has the function of encoding meaning by two means, one by marking phrasal boundaries and another by marking prominence. Both boundary- and prominence-marking are expressed through prosodic strengthening, which refers to "strong" articulation encompassing a spatial and/or temporal expansion of articulatory movements. Under the assumption that prosodic structure serves a dual function of marking grouping and prominence, the two kinds of prosodic strengthening are regarded to be manifested differentially in speech production. However, evidence for this largely comes from head-prominence languages, in which the two functions of prosody are more apart from each other. In this dissertation, phonetic modulations of the two prosodic functions are investigated in Seoul Korean, which is an edge-prominence language, i.e., a language that marks prominence by boundary-marking. In addition, Seoul Korean is known to not have word prosody.

The dissertation includes two research articles. The first article examines how boundaries at major phrases, i.e., Intonational Phrase (IP), are marked and how they interact with the prosodic level of Accentual Phrases (AP), in which prominence is marked in Korean. The second article examines the articulatory profile of focus-induced prominence and compares it to that of non-focused gestures. Whether different types of focus

exert different effects on articulation is also examined. In addition, a derived goal of the dissertation is to test whether the well-known interrelationships between kinematic parameters depend on prosodic function, differentiating thus the prosodic marking of phrasal boundaries from that of prominence.

Results suggest distinct articulatory modulations of boundary- and prominence-marking in Seoul Korean. Gestures at the right-edge of an IP boundary are longer, larger, and slower than phrase-medial counterparts. While phrase-final lengthening consistently affects the phrase-final syllable, its amount or scope is sensitive to focus location and the presence of factors useful to demarcating the last word of the phrase. Under focus-induced prominence, gestures are longer and larger, but not faster per se. Furthermore, it is shown that prosodic prominence is driven by focus, and is not an effect coming from domain-initial strengthening at the AP level. Different focus types systematically exert different degrees of articulatory modulation, suggesting connection between information structure and prosodic structure. Finally, throughout the dissertation, a robust effect of prosodic function on the relationship between kinematic parameters is observed: Each type of prosodic marking, i.e., IP boundary-marking and prominence-marking, presents a distinct relationship between duration and peak velocity normalized over displacement, which is often used as an estimate of stiffness.

Contents

List of Figures

List of Tables

Chapter 0

Note to readers

This is an article-based dissertation that includes two research articles as follows:

1. Article 1: Investigation of phrase boundaries in an edge-prominence language without word prosody

2. Article 2: Examination of focus-induced prominence in Seoul Korean

Article 1 and article 2 correspond to Chapter 2 and Chapter 3 in this dissertation. Each article is written as a separate article submission, and their final version might differ from the current ones as a result of the article reviewing process. Chapter 1 of this dissertation contextualizes the overarching goal and common grounds of these two articles, while Chapter 4 discusses their combined conclusions for theories of prosodic structure.

Chapter 1

Introduction

This dissertation investigates the kinematic profile of boundary- and prominence-marking in Seoul Korean. Specifically, we examine the articulatory correlates of boundaries at minor and major phrases and those of focus-induced prominence, and special attention is given in the interaction between these two prosodic factors.

Prosodic structure involves the grouping of linguistic units and the distribution of prominence in speech, contributing to the rhythmic and expressive aspects of language. As they play a crucial role in language production, perception and acquisition (see Gussenhoven & Chen, 2021), understanding how these prosodic functions work holds particular significance for research aiming to better understand language structure and communication. When linguistic units are grouped, forming a linguistic constituent, the edges of the constituent are referred to as boundaries. Theories of prosody assume that these constituents are organized hierarchically with higher-level prosodic units consisting of one or more lower-level units (e.g., Beckman & Pierrehumbert, 1986; Hayes, 1986; Nespor & Vogel, 1986; Selkirk, 1984; see Shattuck-Hufnagel & Turk, 1996 for an overview; Figure 2.1 in Articles 1 and 2). Although models propose differ in the number of levels they assume for the prosodic hierarchy, they all include two phrase levels above the level of the prosodic word. Temporal and spatial expansion of phonetic correlates at prosodic boundaries have been reported. Producing longer and larger articulatory movements, referred to as gestures, at the left-edges of phrases compared to phrase-medial ones have been discussed in the literature under the term domain-initial strengthening (e.g., Fougeron & Keating 1997; Cho & Keating 2001, 2009; Fougeron 2001; Keating, Cho, Fougeron & Hsu 2003). At the right-edges of the phrases, consistent and robust temporal expansion has been found, referred to as phrase-final lengthening (e.g., Byrd et al., 2006; Cooper & Paccia-Cooper, 1980; Katsika, 2016; Klatt, 1975; Lehiste, 1972; Oller, 1973; Turk & Shattuck-Hufnagel, 2007; Umeda, 1975; Wightman et al., 1992). For both phenomena, the strength of the effect has been shown to be correlated with the level of the prosodic phrase, which means that the effect is stronger the higher the level in the prosodic hierarchy is (e.g., domain-initial: Jun, 1993; Fougeron & Keating, 1997; Byrd & Saltzman, 1998; Cho & Jun, 2000; Fougeron, 2001; Keating et al., 2004; Byrd et al., 2006; domain-final: Byrd, 2000; Byrd & Saltzman, 1998; Cambier-Langeveld, 1997; Cho, 2006; Krivokapić, 2007).

The locus of prominence has also been discussed as another prosodic position that exhibits spatiotemporal expansion of articulatory gestures (e.g., Beckman et al., 1992; Beckman & Edwards, 1994; Byrd & Saltzman, 2003; Cho, 2006). Generally, linguistic units under prominence are assumed to be realized with stronger articulation in all dimensions, i.e., time, distance, and speed (cf. Cho, 2006, 2011). In the discussion of prominence at the phrasal level (to be contrasted with word prosody), the realization of prominence has been proposed to occur in two ways: (1) by marking the head of a prosodic unit, thus termed 'head-prominence', and (2) by marking the edge of a prosodic unit, thus termed 'edge-prominence' (Hyman 1978; Beckman 1986; Beckman & Edwards 1990; Ladd 1996; Venditti et al. 1996).

Under the assumption that prosodic structure serves a dual function of marking grouping and prominence, speakers are expected to differentiate between these two functions (cf. Cho, 2011). The two kinds of prosodic strengthening are regarded to be manifested differentially in speech production, aligning with the view that they are separately encoded in speech planning (cf. Keating, 2006; Cho & Keating, 2009; see also Cho, 2011). However, it should be noted that the seemingly clear dichotomy between these two functions of prosodic structure may arise from evidence that largely comes from head-prominence languages such as English, in which the two functions of prosody are more apart from each other (cf. Cho, 2016). Yet, even in head-prominence languages, the two functions are more interdependent than traditionally assumed. For instance, lexical stress and/or phrasal pitch accent emerges as a factor that fine-tunes the scope of phrase-final lengthening (Katsika, 2016; Kim et al., 2017; see also Turk & Shattuck-Hufnagel, 2007; White, 2002; but see Byrd & Riggs, 2008). The question that arises whether languages that mark prominence by means of grouping, i.e., edge-prominence languages, the two functions of prosody are differentiated from one another, and if yes, how and in which ways they interact with each other. In this dissertation, phonetic modulations of the two prosodic functions are investigated in Seoul Korean, which is an edge-prominence language with no word prosody (e.g., Jun, 1993). In Seoul Korean, phrasal prominence is marked by the means of Accentual Phrases (APs), with the focused linguistic unit starting an AP (or a higher phrase), thus, being AP-initially positioned (e.g., Jun, 1993, 1998, 2005, 2006). Thus, in Seoul Korean, both the function of grouping and the function of prominence are marked by the means of edges, and they may or may not be differentiated from each other in their phonetic realization. With a goal to broaden our understanding of prosodic functions and their interactions, this dissertation uses a set of articulatory studies to address the following questions in Seoul Korean:

1. How are boundaries at major phrases (i.e., at the Intonational Phrase (IP) level)

marked and how do they interact with the prosodic level of Accentual Phrases, which is also the level on which prominence is marked?

2. What is the articulatory profile of Accentual Phrases under focus and how does it compare to that of non-focused Accentual Phrases? Do different types of focus exert different effects on articulation?

An underlying goal of this dissertation is to test whether the well-known interrelationships between kinematic parameters depend on prosodic position, differentiating thus the prosodic marking of grouping and prominence in terms of motor control. It has been established that duration and stiffness show an inverse relationship, i.e., duration increases as stiffness decreases, and that displacement and peak velocity show a direct relationship, i.e., displacement increases as peak velocity increases (e.g., Byrd & Saltzman, 1998; Munhall et al., 1985). In this line of work, articulatory gestures are modeled dynamically as linear second-order critically damped models, and proxy measurements for the abstract dimension of stiffness have been are employed. One such estimate is time-to-peak velocity, with shorter time-to-peak velocity implying higher stiffness (Byrd & Saltzman, 1998). The normalized peak velocity over displacement has also been used as another empirical estimate of kinematic stiffness, capturing the observation that peak velocity increases with displacement (Munhall et al., 1985, Ostry & Munhall, 1985). Here, we also use the measure of normalized peak velocity over displacement, and we refer to it as stiffness for reasons of brevity.

Chapter 2 (Article 1) investigates the IP-final boundary kinematics. Introductions on prosodic structure, Korean prosody, as well as related discussion on the phonetic correlates of IP boundaries are presented in the chapter. Methodology of the experiment (participants, experimental procedure, experimental design and stimuli, measurements, and statistical analyses) follows. Results and discussion on the findings on IP boundary-marking in Korean concludes the chapter.

Chapter 3 (Article 2) examines the kinematics of focus-derived phrasal prominence as well as the effect of dephrasing that is observed on the linguistic units that follow the focused AP. An introduction of the state-of-the art in the phonetics of prominence and Korean prosody is given first. The chapter includes three experiments: (1) Experiment 1 looks at the kinematic effects of focus-induced prominence on the focused unit and the post-focal units; (2) Experiment 2 examines the interaction between focus-marking at the AP level and IP boundary-marking; and (3) Experiment 3 investigates the effects of different types of focus on articulation. Chapter 3 concludes with a discussion, drawing from results of the three experiments.

Chapter 4 summarizes the findings of the results and presents a general discussion on boundary- and prominence-marking in Seoul Korean. Finally, implications as well as suggestions for future directions are provided.

References

Beckman, M. E., & Edwards, J. (1994). Articulatory evidence for differentiating stress categories. *Papers in laboratory phonology III: Phonological structure and phonetic form*, 7–33.

Beckman, M. E., Edwards, J., & Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. *Papers in laboratory phonology II: Segment, Gesture, Prosody*.

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

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

Byrd, D., Krivokapić, J., & Lee, S. (2006). How far, how long: On the temporal scope of prosodic boundary effects. *The Journal of the Acoustical Society of America*, *120*(3), 1589–1599.

Byrd, D., & Riggs, D. (2008). Locality interactions with prominence in determining the scope of phrasal lengthening. *Journal of the International Phonetic Association*, *38*(2), 187–202.

Byrd, D., & Saltzman, E. (1998). Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics*, *26*(2), 173–199.

Byrd, D., & Saltzman, E. (2003). The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, *31*(2), 149–180.

Cambier-Langeveld, T. (1997). The domain of final lengthening in the production of Dutch. *Linguistics in the Netherlands*, *14*(1), 13–24.

Cho, T. (2006). Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English. *Laboratory phonology*, *8*, 519–548.

Cho, T. (2011). Laboratory phonology. *The continuum companion to phonology*, 343–368.

Cho, T. (2016). Prosodic boundary strengthening in the phonetics–prosody interface. *Language and Linguistics Compass*, *10*(3), 120–141.

Cho, T., & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, *37*(4), 466–485.

Cho, T., & Keating, P. A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of phonetics*, *29*(2), 155–190.

Cho, T., & Sun-Ah, J. (2000). Domain-initial strengthening as featural enhancement: Aerodynamic evidence from Korean. In *Chicago linguistics society* (Vol. 36, pp. 31–44).

Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and speech.* Harvard University

Press.

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

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *The journal of the acoustical society of America*, *101*(6), 3728–3740.

Gussenhoven, C., & Chen, A. (2021). *The oxford handbook of language prosody*. Oxford, UK: Oxford University Press.

Hayes, B. (1989). The prosodic hierarchy in meter. In *Rhythm and meter* (pp. 201–260). Elsevier.

Jun, S.-A. (1993). *The phonetics and phonology of Korean prosody* (PhD 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 Sciences*, *7*(1), 143–170.

Jun, S.-A. (2005). Korean intonational phonology and prosodic transcription. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 201–229). Oxford, UK: Oxford University Press.

Jun, S.-A. (2006). Intonational phonology of Seoul Korean revisited. In *Japanese-korean linguistics* (Vol. 14, pp. 15–26).

Katsika, A. (2016). The role of prominence in determining the scope of boundary-related lengthening in Greek. *Journal of phonetics*, *55*, 149–181.

Keating, P., Cho, T., Fougeron, C., & Hsu, C. (2003). Domain-initial strengthening in four languages. *Papers in laboratory phonology VI: Phonetic interpretations*, 145–163.

Klatt, D. H. (1975). Vowel lengthening is syntactically determined in a connected discourse. *Journal of phonetics*, *3*(3), 129–140.

Krivokapić, J. (2007). *The planning, production, and perception of prosodic structure* (PhD dissertation). University of Southern California.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.

Munhall, K. G., Ostry, D. J., & Parush, A. (1985). Characteristics of velocity profiles of speech movements. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(4), 457.

Nespor, M., & Vogel, I. (2007). *Prosodic phonology: with a new foreword* (Vol. 28). Mouton de Gruyter.

Oller, D. K. (1973). The effect of position in utterance on speech segment duration in English. *The journal of the Acoustical Society of America*, *54*(5), 1235–1247.

Selkirk, E. (1984). *Phonology and syntax: The relationship between sound and structure*. Cambridge, MA: MIT Press.

Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of psycholinguistic research*, *25*, 193–247.

Turk, A. E., & Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics*, *35*(4), 445–472.

Umeda, N. (1975). Vowel duration in American English. *The Journal of the Acoustical Society of America*, *58*(2), 434–445.

White, L. S. (2002). *English speech timing: a domain and locus approach* (PhD dissertation). University of Edinburgh Edinburgh.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America*, *91*(3), 1707–1717.

Chapter 2

Article 1: Boundary Kinematics

2.1 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 controlling several phonetic dimensions, one of the main ones being the temporal profile of the utterance. Indeed, one of the most widely attested phonetic effects of prosodic structure is *phrase-final* or *pre-boundary lengthening*. 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; Cooper & Paccia-Cooper, 1980; Katsika, 2016; Klatt, 1975; Lehiste, 1972; Oller, 1973; Turk & Shattuck-Hufnagel, 2007; Umeda, 1975; Wightman et al., 1992). The phenomenon is so prevalent in both oral and signed speech (Wilbur, 2009) that it is considered a language universal (e.g., Vaissière, 1983; Tyler & Culter, 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 Gussehoven & Chen, 2020, Sections V-

VII for a review), and how ubiquitous the presence of pre-boundary lengthening is in those boundaries, examining the effect's scope can shed important light onto the factors that define prosodic boundaries and what this means for both the grammatical structure of prosody and speech planning.

Based on the scarce work on the matter, two opposing sets of theories are proposed. On the one hand, there is a body of work suggesting that the effect might be determined by a, possibly language-specific, grammatical domain, such as the phrase-final syllabic rhyme or foot or even word (e.g., syllable: Nakatani et al. 1981; Oller, 1973; Wightman et al., 1992; foot: White, 2002; word: Kohler, 1983). A completely alternative proposal suggests that the effect does not target a well-defined grammatical domain, but instead, it scopes over a fixed interval at the boundary (Byrd & Saltzman, 2003). Regardless of what the domain of pre-boundary lengthening is assumed to be, stress emerges as a factor that further fine-tunes it: The effect is initiated earlier in phrase-final words with earlier stress (Katsika, 2016; Kim et al., 2017; see also Turk & Shattuck-Hufnagel, 2007; White, 2002; but see Byrd & Riggs, 2008). However, in stress languages, positions that bear stress, i.e., the word prosody marker, are conflated with positions that can carry phrase-level prominence, since the latter is marked by the means of phrase pitch accents on stressed syllables of accented words. It is thus unclear which function of stress contributes to the scope of pre-boundary lengthening; the marker of word prosody or the anchor for phrase-level prominence. Current work on boundary marking in Japanese finds that lexical pitch accent shows similar patterns to lexical stress (Tsai & Katsika, 2020; Tsai, 2023), enhancing the hypothesis that word prosody determines when boundary effects occur (see also Seo et al., 2019). However, in that work, 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 boundary marking. This is because Korean does not employ word prosody and marks phrase-level prominence (e.g., focus location and focus type) by the means of boundary tones at the edges of relatively small phrases, called Accentual Phrases (APs) (Beckman & Pierrehumbert, 1986; Jun, 1993). This property warrants Korean its characterization as an edge-prominence language (cf. Jun, 2005, 2014). It is furthermore noteworthy that the initial boundary tone of APs in edge- (and head/edge-) prominence languages share an important function with lexical stress – but not phrasal pitch accent – in stress languages, that of facilitating word segmentation (e.g., stress: Cutler & Butterfield, 1991; Cutler & Norris, 1988; AP in Korean: Kim, 2004; Kim & Cho, 2009; Kim et al., 2012; AP in French: Welby, 2007; AP in Japanese: Warner et al., 2010). The Electromagnetic Articulography study reported here is designed to examine the effect of phrase-level prominence/focus marking separately from that of word demarcation in the scope of pre-boundary lengthening in Korean.

By the means of this investigation, we hope to shed significant light onto the question 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 and AP. This study will also contribute to the discussion of how speech planning works to interweave different sources of prosodic and lexical/segmental information (cf. Keating & Shattuck-Hufnagel, 2002 vs. Levelt, 1989). Moreover, knowledge will be gained as to how prosodic structure interfaces with information structure, specifically, focus. A derived goal of this investigation is to better understand the relationship between kinematic parameters that materializes boundary-related lengthening. Our analyses examine not only the dimension of time, but position and velocity as well, since articulatory movements in the vicinity of prosodic boundaries have been found to become not only longer, but also larger and faster (e.g., Cho, 2006). Ultimately, we hope to contribute to a more comprehensive definition of the prosodic component of grammar and the role it plays in linguistic cognition.

In what follows, Section 2.1.1 introduces prosodic structure, and Section 2.1.2 the main aspects of Korean prosody. Section 2.1.3 and 2.1.4 present the state of the art with respect to the phonetic correlates of prosodic boundaries and the scope of phrase-final lengthening respectively, including previous work on Korean. Section 2.1.5 briefly describes the pi-gesture model of prosodic boundaries, and Section 2.1.6 presents our research questions and predictions.

2.1.1 Prosodic structure

Theories of prosody assume that prosodic structure is organized hierarchically with higher-level prosodic units consisting of one or more lower-level units (e.g., Beckman & Pierrehumbert, 1986; Hayes, 1986; Nespor & Vogel, 1986; Selkirk, 1984; see Shattuck-Hufnagel & Turk, 1996 for an overview). The prosodic structure of an utterance is depicted in Figure 2.1, adopting the model proposed by Beckman & Pierrehumbert (1986). As the figure illustrates, syllables constitute prosodic words (PWd); prosodic words constitute intermediate phrases (ip); and intermediate phrases constitute Intonational Phrases (IP). The figure of prosodic structure in (1) also reflects some aspects of the relative prominence of prosodic units. Prominence is marked at different levels of the structure: at the lexical level, stressed syllables, for instance, are marked by the bar between the PWd and the syllable tier (cf. Keating & Shattuck-Hufnagel, 2002), which indicates relative salience of these syllables compared to the unstressed syllables. At the phrasal level, pitch accented syllables are marked with T* (either H* for high pitch accent or L* for low 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: T- stand for phrase tones (e.g., L- or H-) marking the edges of intermediate phrases and T% for boundary tones

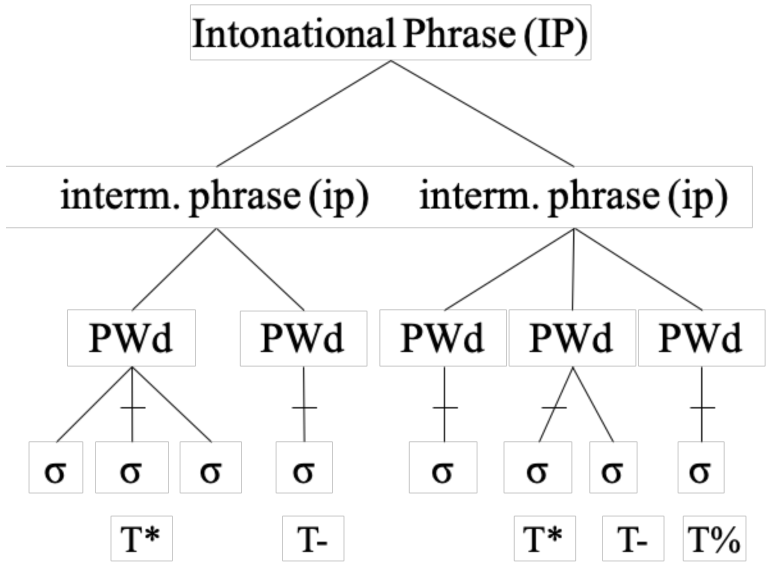

Figure 2.1: Prosodic structure, adapted from Beckman & Pierrehumbert (1986). PWd and σ refer to prosodic word and syllable respectively. Bars between the PWd and syllable tier indicate stressed syllables as in Keating and Shattuck-Hufnagel (2002). Pitch accents are represented with T*, phrase accents with T-, and boundary tones with T%.

(e.g., L% or H%) marking the edges of intonational phrases. These tonal markings, along with pitch accents, describe the overall intonation of the utterance.

While there is general agreement in the structural view of prosody, models disagree in the number of hierarchical levels they propose (cf. Shattuck-Hufnagel & Turk, 1996). Also, different languages are reported to exhibit distinct prosodic structures. For instance, existence of the mora, which is a lower-level unit below the syllable, is unquestionable in Japanese, but is more controversial in other languages (cf. Shattuck-Hufnagel & Turk, 1996). Also, languages, such as Japanese and Korean, are assumed to have an additional level below the level of intermediate phrase, referred to as Accentual Phrase (AP), the size of which is equal to or larger than the prosodic word (Beckman & Pierrehumbert, 1986; Jun, 1993). Languages may have different prominence-marking systems (cf. Jun, 2006, 2014). At the word prosody level, some languages may employ lexical stress, other languages may employ lexical tone or lexical pitch accent, while others may not use any type of lexical prosody at all. At the phrase prosody level, languages are broadly categorized into head- and edge-prominence languages, with prominence being marked by a pitch accent on the head of the phrase in the former and by a boundary tone at the edge of the phrase in the latter. Hybrid systems have been proposed both at the word prosody and the phrase prosody level. Korean, which is examined here, is considered an edge-prominence language with no lexical prominence-marking system.

2.1.2 Korean prosody

This section summarizes key properties of Korean prosody related to the current study. As mentioned above, Korean, specifically Seoul Korean (also known as 'Standard Korean' or 'Pyojuneo' based on Korean spoken in Seoul and Gyeonggi area), is an edge-prominence language known for not having lexical stress, lexical tone or lexical pitch accent (e.g., Jun, 2005, 2006). The Accentual Phrase (AP), which is a post-lexical prosodic phrase, serves as the basic intonational unit, with intonation being linked to certain locations of the phrase. As proposed by Jun (1993, 2005), the AP's underlying tonal pattern is THLH, where the realization of the initial tone (T) tends to depend on the laryngeal configuration of the AP-initial segment[1] (see Jun, 1998, 2005, 2006; Jeon & 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 of prosodic structure that specifies Accentual Phrase (AP) and Intonational Phrase (IP) above the Prosodic Word (PWd) level (Jun, 1996, 1998, 2000, 2005, 2006). Some proposals include intonational phrase

[1]Tonogenetic sound changes have been observed among younger speakers, with the VOT distinction between aspirated and lenis stops merging, and F0 becoming the primary cue to distinguish the contrast (higher F0 for aspirated and lower F0 for lenis) (Bang et al., 2018; Kang, 2014). However, these changes are limited to specific segments and phrasal positions, namely stop and affricate consonants in AP-initial positions.

(ip) level between IP and AP, but since there has not been much investigation on this level, and since detecting the ip is not straightforward (Jun, 2005, 2007), we will not consider this level in further discussion.

Phrasal prominence in Korean is marked by prosodic phrasing, and the focused word consistently initiates an AP or a higher phrase (Jun, 1993, Jeon & Nolan, 2017). Often, following AP boundaries undergo elimination, or possibly attenuation, referred to as dephrasing (e.g., Jun, 1993), up to the end of the IP. Limited work on the correlates of prominence in Korean reports that under focus, articulatory constriction movements, called constriction gestures or just gestures, become longer, larger and faster (Shin et al., 2015; see also Jang & Katsika, 2019)

2.1.3 Phonetics of boundary

As boundaries are one of the most robust prosodic landmarks, a number of studies have looked at the acoustic and articulatory events that are observed at or across phrasal boundaries. Phrase-initially, constriction gestures have been reported to be temporally and spatially expanded under the notion of domain-initial strengthening (e.g., Pierrehumbert & Talkin, 1992; Fougeron & Keating, 1997; Cho & Keating, 2001, 2009; Keating et al., 2004). Acoustic and articulatory evidence indicates that domain-initial strengthening is cumulative, i.e., increasing with boundary strength (e.g., acoustic studies: Jun, 1993; Cho & Jun, 2000; Fougeron, 2001; articulatory studies: Byrd et al., 2006; Byrd & Saltzman, 1998; Fougeron & Keating, 1997; Keating et al., 2004). The scope of this effect has been regarded to be limited to the initial segment, primarily on the basis of articulatory data (e.g., Byrd et al., 2006; Byrd & Saltzman, 1998; Katsika, 2016).

At the end of phrases, i.e., before a prosodic boundary, gestures are longer, larger, and slower (cf. Cho, 2006). Temporal modification, i.e., longer durations, is the most well-known characteristic of phrase-final gestures. Studies have found consistent phrase-final lengthening in different languages (e.g., American English: Edwards, Beckman, & Fletcher, 1991; Byrd, 2000; Cho, 2006; Nakatani, O'Conor & Aston, 1981; Oller, 1973; Turk & Shattuck-Hufnagel, 2007; Kim et al., 2017; British English: Campbell & Isard, 1991; White, 2002; Korean: Kim et al., 2019; Japanese: Seo et al., 2019, Tsai & Katsika 2020; Dutch: Cambier-Langeveld, 1997; Greek: Katsika, 2009, 2016). Analogous to the property reported for domain-initial strengthening, the strength of phrase-final lengthening is also cumulative: the higher the prosodic boundary that follows, the greater the effect of phrase-final lengthening (e.g., Byrd, 2000; Byrd & Saltzman, 1998; Cambier-Langeveld, 1997; Cho, 2006; Krivokapić, 2007).

Gestures slow down at boundaries, as indicated by findings on peak velocity and time-to-peak-velocity. In Cho (2006), it is found that as boundary strength increases (i.e., the higher the level of the boundary is in the prosodic hierarchy), peak velocity decreases and time-to-peak-velocity increases for trans-boundary lip closing (vowel to consonant) movement in English. Byrd (2000) also finds that time-to-peak-velocity was greater at higher-level phrases, meaning that gestures take longer time to reach their peak velocity as boundary strength increases. Notably, peak velocity is a direct measure of velocity, while time-to-peak-velocity is considered an estimate of stiffness. Longer time-to-peak-velocity indicates a lowering of the gestures' stiffness parameter in the vicinity of the prosodic boundaries (Beckman et al., 1992; Byrd & Saltzman, 1998; Byrd, 2000), although research suggests that change in stiffness parameter alone is insufficient to capture the comprehensive kinematic variation at phrasal boundaries (Saltzman & Byrd, 2000; Byrd & Saltzman, 2003).

Finally, studies have shown that articulation in phrase-final positions demonstrates larger and less overlapped gestures (Byrd & Saltzman, 2003; Byrd et al., 2000, Cho, 2005, 2006). For instance, tongue body and jaw position have been reported to show increased magnitude of movement at stronger prosodic boundaries than weaker prosodic boundaries (Tabain, 2003). It needs to be noted, however, that studies on the spatial effects of phrase-final lengthening are not only scarce, but also tend to show much variation across speakers and segments examined (e.g., Byrd et al., 2005, 2006; Fougeron & Keating, 1997).

2.1.4 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 in earlier parts of the phrase-final word have been observed as well. For example, the effect may extend beyond the rhyme of the phrase-final syllable in Dutch depending on the vowel quality of the phrase-final vowel (Cambier-Langeveld, 1997). Another factor that emerges as relevant is the position of stress within the phrase-final word. The patterns found in Oller (1973) suggest that lengthening extends to the onset of the final syllable when lexical stress is non-final in English, while Berkovits (1994) reports that in disyllabic stress-initial words in Hebrew phrase-final lengthening affects both syllables. Both these papers report that the effect is progressive, meaning that the effect is strongest at the boundary and gradually decreasing with distance from it (see also Campbell & Isard, 1991; White, 2002, among others).

Studies that specifically explore the effect of stress/focus-marking pitch accent positions on the scope of phrase-final lengthening detect systematic interactions. For instance, in Greek, the position of lexical stress within the phase-final word affects the timing of both phrase-final lengthening and boundary tones. In particular, the earlier the stress is within the phrase-final word, the earlier these boundary-marking events are initiated (Katsika, 2016; Katsika et al., 2014). These effects hold even in de-accented phrase-final words. Similar effects of lexical stress on pre-boundary lengthening are also found in English: In an articulatory study, Kim and colleagues (2017) show that phrase-final lengthening begins earlier in the final word when lexical stress is non-final as opposed to final (but see Byrd & Riggs, 2008), whereas in an acoustic study, Turk and Shatttuck-Hufnagel (2007) detects phrase-final lengthening on the final syllable as well as on the stressed/accented syllable, leaving any intervening syllables unaffected in American English. In British English, White (2002) reports final lengthening beginning in the coda of the main-stressed (e.g., on [k] of spectre and spectacle). While this work provides great insight on the possible factors determining the activation of phrase-final lengthening, it has mainly focused on languages that employ lexical stress and mark phrasal prominence on stressed positions (head-prominence languages), leaving open questions as to the scope of phrase-final lengthening in languages with different prosodic systems.

Recently, similar investigations were carried out in 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 & Katsika, 2020), with unaccented words behaving similarly to words with final pitch accent (but see acoustic data in Seo et al., 2019 on unaccented words).

These results are taken to have typological implications, suggesting that, word prosody, regardless of type (stress/pitch accent) influences the activation and manifestation of phrase-final lengthening. Here, we turn to Korean, a language that lacks any marker of word prosody. Limited research on the matter, measuring lip movements over disyllabic words suggests that the effect is distributed across both syllables (or the whole phrase-final word), regardless of their vowels' intrinsic duration (/a/ vs. /i/) or information status (background vs. new information) (Kim et al., 2019). In addition to being longer, lip movements were also found to be larger and, contrary to expectations, faster. This is interpreted as resulting from Korean being an edge-prominence language on the premise that prominence marking is associated with longer, larger and faster movements (cf. Cho, 2006). Here, we extend the work on Korean, by examining the influence of focus position and final AP's length on the kinematic profile (i.e., on the dimensions of time, position, and velocity) of IP boundaries. The question is motivated by the effects that word prosody and/or phrase prominence have been found to have on the kinematic manifestation of prosodic boundaries in stress languages, and is centered on the typological advantage of looking at a language in which word prosody is missing and thereby disentangled from phrase-level prominence.

2.1.5 Pi-gesture framework

The patterns of longer, larger, and less overlapped gestures phrase-finally have been modelled by the means of pi-gestures (Byrd & Saltzman, 2003). The pi-gesture model is couched within Articulatory Phonology, and extends the view that phonological units are inherently dynamic to the prosodic level (e.g., Browman & Goldstein, 1992; Byrd & Riggs, 2008). Within Articulatory Phonology, the building blocks of phonology are gestures, which are "events that unfold during speech production and whose consequences can be observed in the movements of the speech articulators (Browman & Goldstein, 1992)." Constriction gestures are specified for abstract linguistic tasks executed by coordinated actions of the speech articulators (i.e., the lips, tongue tip, tongue dorsum, velum, and glottis) and triggered by internal oscillators that are coupled to each other either in-phase or anti-phase forming syllables and words (e.g., Goldstein, Byrd, & Saltzman, 2006). Contrary to constriction gestures, pi-gestures are not related to specific articulators, and

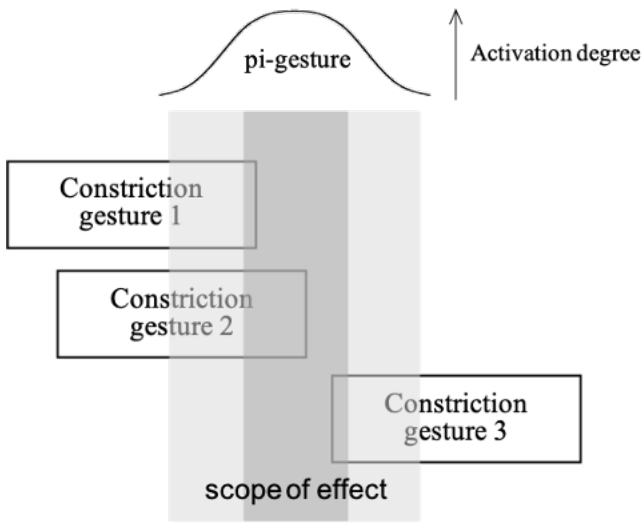

Figure 2.2: A schematic representation of a pi-gesture (adapted from Byrd & Saltzman, 2003). The gray-shaded boxes represent the effect of pi-gesture, with the darker box corresponding to the pi-gesture's maximal activation.

their task is to locally slow down the clock that controls the global pace of the utterance (e.g., Byrd & Saltzman, 2003). Clock slowing occurs locally, because it is determined by the activation interval of the pi-gesture, represented by the gray-shaded box in Figure 2.2. Pi-gestures modulate the temporal properties of (the part of) constriction gestures that overlap with their activation interval, as shown in Figure 2.2. 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 at the prosodic boundary and decreases with distance from it over a continuous domain, while the level of maximum activation depends on the strength of the prosodic boundary, capturing in this way the progressive and cumulative nature of pre-boundary lengthening, discussed in Sections 2.1.3 and 2.1.4.

Another dimension that determines the scope of pi-gesture's effect is their coordination with the constriction gestures. For instance, it has been proposed that pi-gestures in Greek are coordinated both with the phrase-final V (vocalic) gesture and the mu-gesture, meaning the modulating gesture that governs the spatio-temporal properties of stress (Saltzman et al., 2008), that corresponds to the final lexical stress of the phrase (Katsika, 2016; Katsika et al., 2014). This dual coordination accounts for the earlier initiation of boundary related events in non-final stressed words compared to final stressed words 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 non-final stress, in which the stress-related mu-gesture coordination with the pi-gesture pulls the latter towards the stressed syllable, initiating phrase-final lengthening earlier in words with non-final stress as opposed to words with final stress.

2.1.6 Hypoetheses and predictions

The goal of the current study is to investigate whether phrase-level prominence marking and/or word demarcation factors affect the kinematic manifestation (i.e., on duration, displacement, velocity) of Intonational Phrase (IP) boundaries in Seoul Korean. Our investigation is motivated by the attested findings showing position of lexical stress conditioning the scope of the phrase-final lengthening in stress languages, combined with the multi-functionality of lexical stress (word prosody marker, word demarcating cue, and anchor for phrasal pitch accent), which make it difficult to pinpoint which of these functions contribute to the exerted effect on boundary-related lengthening in stress languages. Korean lacks word prosody markers, while focus marking can be separated to an extent from word demarcation. To this end, we employ an Electromagnetic Articulography (EMA) study that manipulates the distance of the focused linguistic unit from the boundary as well as the length of the final AP (i.e., the distance of the AP-initial accent expected to cue word demarcation from the boundary).

In terms of general patterns, we expect that Korean will present phrase-final lengthening (as shown in Kim et al., 2019; also based on cross-linguistic prevalence of the effect as discussed in Vaissière, 1983; Tyler & Cutler, 2009). The effect might span over the last two syllables (or the whole phrase-final word; see Kim et al., 2019, in which the phase-final words were disyllabic), and it should be progressive, with the rhyme of the final syllable being affected the most (e.g., Byrd et al., 2006; Krivokapić, 2007; Katsika, 2016; see also acoustic studies, e.g., Berkovits,1994; Cambier-Langeveld, 1997; Campbell & Isard, 1991; Nakatani et al., 1981; Oller, 1973; Wightman et al., 1992).

Moreover, in accordance with the existing literature, we predict that the affected articulatory gestures will be not only longer but also larger (e.g., Cho, 2006; Kim et al. 2019; also, according to the pi-gesture model by Byrd & Saltzman, 2003). It has also been claimed that articulatory gestures in the vicinity of IP boundaries slow down (e.g., Cho, 2006). However, it is unclear whether this claim will hold in our data, since limited work in Korean has detected faster gestures pre-boundary (Kim et al, 2019). As faster gestures are generally associated with prominence marking, this acceleration of pre-boundary gestures was attributed to Korean being an edge-prominence language. Here, we assess the velocity profile of pre-boundary gestures by taking into account the displacement parameter, which has been shown to vary proportionally with peak velocity (Munhall et al., 1985; Ostry & Munhall, 1985), following the method in Katsika & Tsai (2021). Our goal is to assess whether phrase-final gestures remain faster than their phrase-medial counterparts despite the inherent relationship between displacement and velocity, and to also separate any conflicting influences that prominence- and boundary-marking may be exerting at the IP boundaries of an edge-prominence language. This is because the principal mechanism for gestures to become larger in prominence marking seems to be the control of velocity (the gestures are larger the faster their movement) (see discussion in Katsika & Tsai, 2021), while in boundary marking, gestures are larger because they overlap less with their neighboring gestures (see e.g., discussion on overlap in Byrd et al., 2000; Byrd & Saltzman, 1998).

Under the assumption that the anchor to phrase-level prominence affects the phonetic activation of IP boundaries, pre-boundary lengthening should be initiated earlier the earlier the focus is in the IP. If the dimension of word demarcation is or contributes to the phonetic activation of IP boundaries, pre-boundary lengthening will begin earlier in long as opposed to short final APs, since 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 (i.e., lexical stress in e.g., Greek, English, and lexical pitch accent in Japanese) that affects the span of boundary effects.

2.2 Methods

2.2.1 Participants

Seven native speakers of Seoul Korean (5 female, 2 male) between ages 21 and 29 (mean age = 24.5; median age = 23) 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 as graduate or exchange students or post-doctorate researchers at the time of the experiment, and one participant was a family member of a UCSB-affiliated researcher. The speakers were naïve as to the purpose of the current study and had no self-reported speech, hearing, or vision problems. They received financial compensation for their participation.

2.2.2 Experimental procedure

Before the experiment, the participants went through a short, 15-minute long, training session in order 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 set-up allowed participants to naturally produce the AP phrasing and place the focus contrast as intended (see 2.2.3 for details on the stimuli).

For the experimental session, kinematic data were collected using the AG501 3D electromagnetic articulograph (Carstens Medizinelektronik) 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 the means of a Sennheiser shotgun microphone set at a sampling rate of 16 kHz and positioned one foot away from the participant's face. Speech materials were presented on a computer screen, placed roughly three feet away from the participant, using custom software, developed by Mark Tiede (Haskins Laboratories). The acquired articulatory data for each trajectory was smoothed and corrected for head movement by using the reference sensors. Then, they were rotated to align the X- and Y- axis to the participants' occlusal plane. To help appropriate focus placement, a prompt sentence preceded each target sentence. Prompt sentences were presented one second before their corresponding target sentence, and were 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.

2.2.3 Experimental design and stimuli

The scope of phrase-final lengthening was examined across the test prosodic word /nɛ.maŋ.mi.nam/, which was placed in the IP-final position in a set of test stimuli. In order to create a variable stimulus set, test IPs were designed to consist of either two or three noun phrases (NPs), corresponding to 11 and 14 syllables respectively. Two- and three-NP test IPs were identical to each other except for the presence of the one-word NP /pi.mil.pu/ ('the secrecy club') at the beginning of the longer IPs (e.g., compare stimulus #3 to stimulus #4 in Table 2.1). We refer to the IP-final NP as *Final*, to the IP-penultimate NP as *Penultimate*, and, in the case of the longer IPs, to the IP-antepenultimate NP as *Antepenultimate*. The test IPs were followed by an IP boundary and then a verb phrase (VP). A set of control stimuli included the test word in IP-medial position. Test stimuli involved the same sequence of NPs as their respective control stimuli, but differed in that there was no IP boundary following the test word, which also gave rise to a difference in meaning. This is illustrated in Table 2.1 when comparing stimulus #5 to stimulus #6. The test word was purposefully selected to include nasal consonants in order to avoid segments whose laryngeal configurations would cause tonal effects. 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, 1998, 2005). Interrogative sentence constructions were used so that the IP boundary tone would consistently override the final Ha tone of the AP, resulting in LHLH% (Jun, 2000). In this way, the boundary tone could be detected and analyzed as part of a parallel project (e.g., Jang & Katsika, 2022).

In order to assess the interaction between position of focus and phrase-final lengthening, the location of the focused item within the test IP was manipulated, covering all possible combinations. Specifically, contrastive focus was on: a) the final NP (e.g., #3 in Table 2.1), or b) the penultimate NP (e.g., #2 in Table 2.1), or c) the antepenultimate NP (e.g., #1 in Table 2.1). The latter condition was possible only in three-NP IPs. Because of this, combined with the observation that phrase length affects segmental durations79, the analysis assumes five Focus Location conditions: (1) Antepenultimate in 3-NP IPs (Antepenultimate), (2) penultimate in 3-NP IPs (Penultimate (3-NPs)), (3) final in 3NP-IPs (Final (3-NPs)), (4) penultimate in 2-NP IPs (Penultimate (2-NPs)), and (5) final in 2-NP IPs (Final (2-NPs)).

To assess whether the onset of the IP-final AP plays a role in the timing of boundary-related lengthening, test IPs varied in the length of their final noun phrase (NP), which was either 4 or 7 syllables long (compare #4 to #5 in Table 2.1). The decision to construct the shorter NP to contain 4 syllables was made based on the observation that typically APs contain 3-5 syllables, with APs containing 4 syllables being more likely to yield a full AP tonal pattern (i.e., THLHa, with each syllable bearing a tone) (Jun, 2000, 2003). 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 prosodic words yielding 7 syllables, which was reported to be the upper limit of the number of syllables an AP may include, although other factors such as speech rate may come into play (Jun, 2003). In fact, ToBI analysis of the elicited data confirmed that the participants did produce these 7 syllables as 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 Figure 2.3). We refer to the two levels of the Final AP Length factors as *Short* (i.e., consisting of 4 syllables) and *Long* Final APs (i.e., consisting of 7 syllables). It is crucial to note that if dephrasing (following the focused linguistic unit) results in complete deletion of AP boundaries, this difference

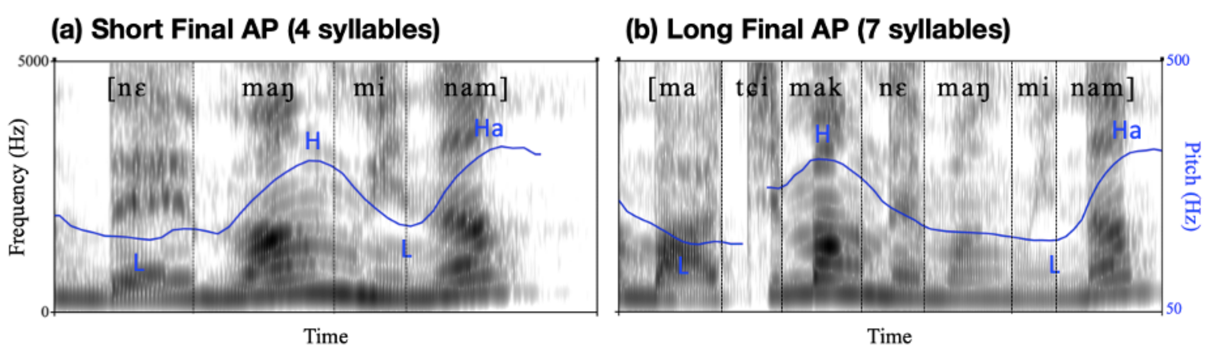

Figure 2.3: Example tonal configurations of (a) Short final AP (4 syllables) and (b) Long final AP (7 syllables) with focus on the final AP. ‘L’ and ‘H’ refer to low and high tones, respectively, and ‘a’ indicates an AP final tone (Jun, 2000).

in final AP length should hold only for IPs focused on the final NP. In IPs focused on the penultimate or antepenultimate NP, the final AP should be 11 or 14 syllables long, respectively, due to dephrasing.

A schematization of the experimental design is presented in Figure 2.4. The stimuli included two IP Position conditions (IP-final, IP-medial), five Focus Location conditions (Antepenultimate (3-NPs), Penultimate (2-NPs/3-NPs), Final (2-NPs/3-NPs)), two final AP Length conditions (Short, Long), resulting in twenty stimuli sentences. Eight stimulus blocks were constructed, each containing one repetition of the 20 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 due to technical reasons. In total, 1060 tokens were collected for the analyses reported here. The acquired data were checked for their prosodic rendition, i.e., focus placement and appropriate accentual and IP phrasing, using K-ToBI (Jun, 2000). 16 tokens were removed either due to incorrect focus placement or phrasing.

#	Focus Location	Final AP Length	IP Position	Test sentence (# = IP boundary)
(1)	Antepenultimate (3-NPs)	Short	IP-final	Prompt: (It's not the magic club.) [IP [NP **pimilpu**] [NP minami gomobuga] [NP nɛmaŋminam]]? # [IP [VP sʌntækhangʌja]]? Uncle Minam from the secrecy club is the handsome guy from Nemang? Is it decided?
(2)	Penultimate (3-NPs)	Short	IP-final	Prompt: (It's not Uncle Junseok.) [IP [NP pimilpu] [NP **minami** gomobuga] [NP nɛmaŋminam]]? # [IP [VP sʌntækhangʌja]]? Uncle Minam from the secrecy club is the handsome guy from Nemang? Is it decided?
(3)	Final (3-NPs)	Short	IP-final	Prompt: (It's not 'the handsome guy from Nowon'.) [IP [NP pimilpu] [NP minami gomobuga] [NP **nɛmaŋminam**]]? # [IP [VP sʌntækhangʌja]]? Uncle Minam from the secrecy club is the handsome guy from Nemang? Is it decided?
(4)	Final (2-NPs)	Short	IP-final	Prompt: (It's not 'the handsome guy from Nowon'.) [IP [NP minami gomobuga] [NP **nɛmaŋminam**]]? # [IP [VP sʌntækhangʌja]]? Uncle Minam is the handsome guy from Nemang (as opposed to Nowon)? Is it decided?
(5)		Long	IP-final	Prompt: (It's not the first.) [IP [NP minamiga] [NP **matɕimak** nɛmaŋminam]]? # [IP [VP sʌntækhangʌja]]? Minam is the last (as opposed to the first) handsome guy from Nemang? Is it decided?
(6)			IP-medial	Prompt: (It's not the first.) [IP [NP minamiga] [NP **matɕimak** nɛmaŋminam] [VP sʌntækhangʌja]]? Minam chose the last (as opposed to the first) handsome guy from Nemang?

Table 2.1: Example stimuli sentences organized by Focus Location (Antepenultimate (3-NPs), Penultimate (3-NPs), Final (3-NPs/2-NPs)), Final AP length (Short vs. Long), and IP Position (IP-final vs. IP-medial). Measured prosodic words are underlined, focused NPs are shown in bold. All sentences were preceded by a short IP meaning 'Really?'. See Appendix 2.2 for full set of stimuli sentences.

2.2.4 Measurements

With the exception of the coda /ŋ/, all the consonant (C) gestures of the test prosodic word /nɛ.maŋ.mi.nam/ were identified and labeled using semi-automatic custom software developed by Mark Tiede (Haskins Laboratories). The coda /ŋ/ was not included in the measurements, due to 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, with C0 being the C gesture closest to the end of the word, and consequently the IP boundary, and C4 being the most remote (Figure 2.4).

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 labeling procedure detected the following kinematic timepoints in each C gesture: onset, time of peak velocity, target, constriction maximum, release, and offset (Figure 2.5). These timepoints were identified on the basis of 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% 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 Figure 2.5).

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

1. Duration of gestural formation and duration of gestural release (in ms; a and b in Figure 2.5)
2. Gestural displacement to target (i.e., the spatial difference between onset and max-

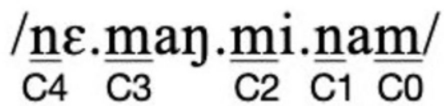

Figure 2.4: Illustration of the coding scheme for the target C gestures in the measured prosodic word /nɛ.maŋ.mi.nam/.

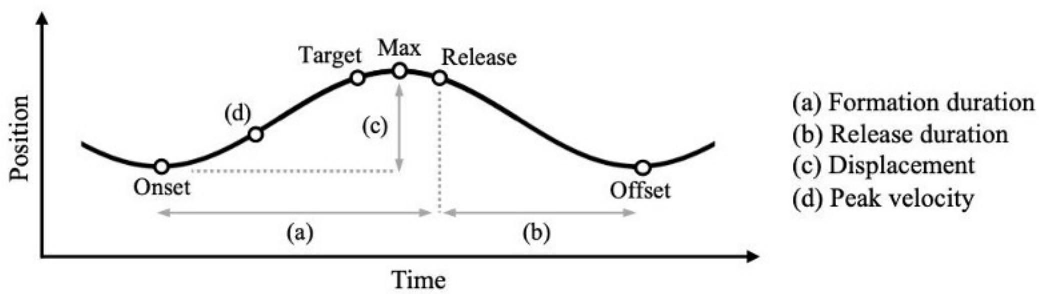

Figure 2.5: Schematized constriction gesture with kinematic measurements.

imum constrictor position) (in mm; c in Figure 2.5)

3. Gestural peak velocity to target (in cm/sec; d in Figure 2.5)

2.2.5 Statistical analysis

The retrieved data were analyzed by linear mixed-effects analysis using *lme4* package (Bates et al., 2015) in R (R Core Team, 2023). The dependent variables were formation duration, release duration, displacement, and peak velocity. For each of the dependent variables, separate linear mixed-effects models were fitted for each segment (C4, C3, C2, C1, C0). The independent variables were IP Position (IP-final, IP-medial), Final AP Length (Long, Short), and Focus Location (Antepenultimate (3-NPs), Penultimate (2-NPs/3-NPs), Final (2-NPs/3-NPs)). In addition to these variables, the following continuous predictors were fitted in each of the models in order to account for the expected covariance between them and the dependent variable (cf. Edwards et al., 1991; Byrd &

Saltzman, 2003), following the methodology in Katsika & Tsai (2021): (a) peak velocity normalized over displacement (as a reminder, we will refer to this measure henceforth as *stiffness* for reasons of brevity) 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. Note that the ratio of peak velocity to displacement has been proposed as an empirical estimate of kinematic stiffness, which captures the observation that peak velocity varies with displacement and has been shown to increase as movement duration decreases in models that assume gestures under a critically-damped second-order linear dynamical system (Munhall et al., 1985; Ostry & Munhall, 1985; used in e.g., Beckman et al., 1992; Hawkins, 1992; Roon et al., 2007). The models allowed up to three-way interactions and included random intercept of Speaker. In case of significant interactions, pair-wise comparisons were assessed by the *emmeans* package (Lenth et al., 2020) with Holm adjustment. When the interactions involved continuous predictors, pair-wise comparisons were carried out at multiple points in the continuum—10th, 25th, 50th (median), 75th, and 90th quantiles. 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, will be reported.

2.3 Results

2.3.1 Relationship between duration and stiffness

The expected relationship between duration and stiffness (estimated by the ratio of peak velocity over displacement) was observed, such that both formation and release durations decreased with increasing stiffness for all measured gestures (C0 formation: $F(1, 1028) = 344.56$, $p < 0.001$; C0 release: $F(1, 1012) = 896.56$, $p < 0.001$; C1 formation:

$F(1, 1025) = 746.25$, $p < 0.001$; C1 release: $F(1, 1026) = 5541.6$, $p < 0.001$; C2 formation: $F(1, 1025) = 345.97$, $p < 0.001$; C2 release: $F(1, 1026) = 5840.15$, $p < 0.001$; C3 formation: $F(1, 1024) = 1054.31$, $p < 0.001$; C3 release: $F(1, 1027) = 4621.34$, $p < 0.001$; C4 formation: $F(1, 1029) = 294.47$, $p < 0.001$; C4 release: $F(1, 1029) = 3625.94$, $p < 0.001$). Interestingly, an interaction between IP Position and stiffness was observed in gestures close to the prosodic boundary: C1 release ($F(1, 1025) = 399.09$, $p < 0.001$), C0 formation ($F(1, 1031) = 62.44$, $p < 0.001$), and C0 release ($F(1, 1029) = 69.37$, $p < 0.001$). Figure 2.6 demonstrates the relationship between duration and stiffness by IP Position in each gesture. In gestures that showed a significant IP Position and stiffness interaction, the absolute slope of the correlation was higher in the IP-final condition than in IP-medial condition as shown in the figure. For example, in C0 formation, for one unit ((cm/sec)/cm) of stiffness decrease, the increase in duration was 7.2 ms in the IP-final condition but just 2.7 ms in the IP-medial condition. That is, given the same decrease in stiffness, the formation of the IP-final gesture was 4.5 ms longer per unit of stiffness decrease than its IP-medial counterpart. Such difference between IP-final and IP-medial conditions were 1.2 ms/per unit of stiffness for C0 release, and 1.3 ms/unit of stiffness for C1 release, whereas the difference was at most 0.3 ms/unit of stiffness for the preceding gestures that do not show the interaction effect. This interaction implies that in Korean, stiffness, which is in an intrinsic relationship with duration (Munhall et al., 1985; Ostry & Munhall, 1985), is systematically controlled by speakers in order to give rise to the durational markers of IP position.

2.3.2 Effects of IP boundary on duration

Figure 2.7 summarizes the main effects of IP Position, demonstrating the general scope of phrase-final lengthening regardless of other focus- or AP-related factors (C0

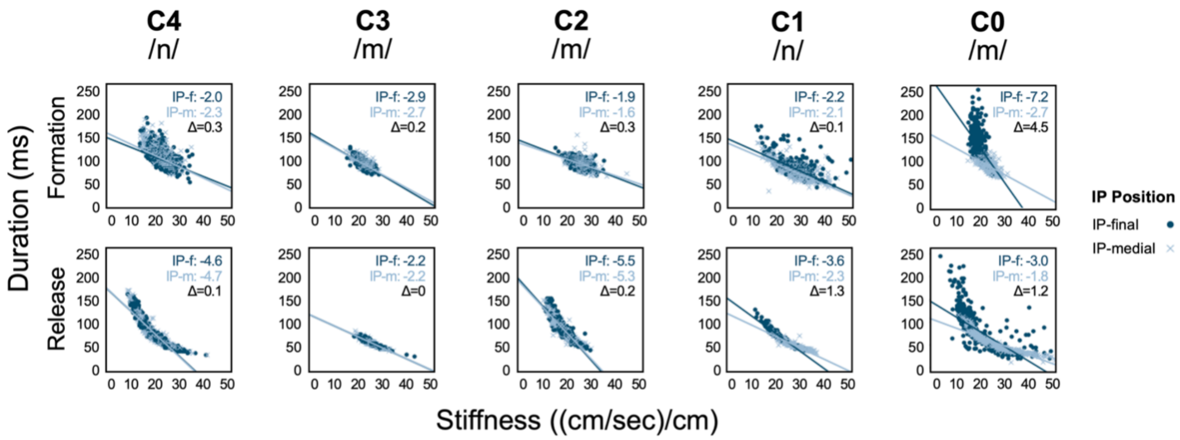

Figure 2.6: Predicted (regression line) and observed (scatter plot) relationship between duration (top panel: formation duration, bottom panel: release duration) and stiffness (measured as peak velocity over displacement) by IP Position (IP-final and IP-medial) for each gesture. The numbers in each plot represent the slope for each condition and their difference.

formation: $F(1,\ 1031) = 98.95$, $p < 0.001$, C0 release: $F(1,\ 1030) = 106.32$, $p < 0.001$, C1 formation: $F(1,\ 1020) = 5.76$, $p < 0.05$, C1 release: $F(1,\ 1025) = 478.79$, $p < 0.001$, C4 formation: $F(1,\ 1024) = 4.46$, $p < 0.05$). Phrase-final lengthening affected the C gestures of the phrase-final syllable, extending to the formation of the syllable's onset consonant. Lengthening was greater in the coda compared to the onset, with the duration difference between IP-final and IP-medial conditions for C0 being on average 22ms for formation and 8ms for release, but for C1 being on average 6ms for formation and 2ms for release. In contrast to the lengthening effect, shortening was observed at the beginning of the prosodic word, with C4 formation being shorter IP-finally as compared to IP-medially. In spite of its small size (2 ms on average), the effect held across the board, i.e., regardless of focus position or final AP length. Similar findings of such a shortening effect have been reported previously in the literature, albeit scarcely (e.g., Katsika, 2016, Byrd et al., 2006).

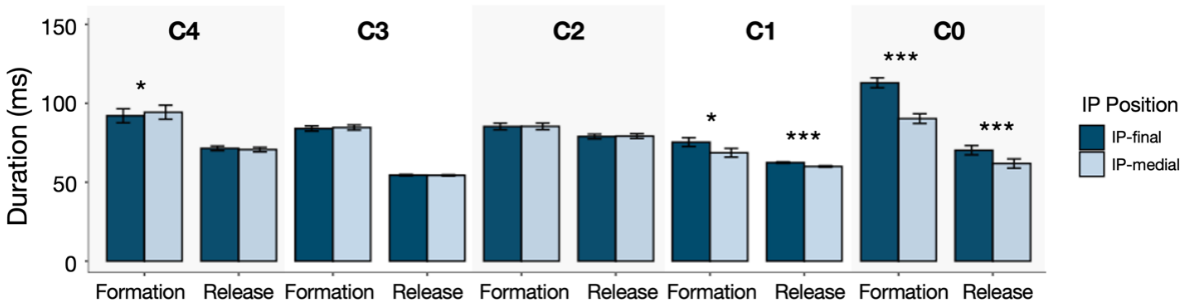

Figure 2.7: Duration (in ms) of gestural formation and release of the test C gestures (C4, C3, C2, C1, and C0) in /nɛmaŋminam/ (as a reminder, /ŋ/ is excluded from the analysis). C0 is the consonantal gesture immediately adjacent to the IP boundary and C4 is the most distant from it. Dark bars correspond to IP-final gestures and light bars to IP-medial ones. Significant pairwise comparisons are marked by the means of asterisks as follows: *** = '$p < 0.001$' and * = '$p < 0.05$'.

2.3.3 Interaction of IP boundary with Focus Location

A three-way interaction between IP Position (IP-final, IP-medial), Focus Location (Antepenultimate (3-NPs), Penultimate (2-NPs/3-NPs), Final (2-NPs/3-NPs)), and Stiffness was detected in C0 Release ($F(4, 1026) = 5.77$, $p < 0.001$) and C1 Release ($F(4, 1020) = 7.78$, $p < 0.001$). As the interaction involved a continuous factor (i.e., stiffness), post-hoc analysis included pairwise comparisons of IP Position and Focus Location effects at the following quantiles of Stiffness: 10th, 25th, and 50th (median), which represented the majority of the data at the IP-final boundary (72% for C0 release, 74% for C1 release). Figure 2.8 plots the predicted relationship between duration and stiffness by IP Position for each Focus Location, and Figure 2.9 visualizes the duration differences between IP-final and IP-medial conditions for each Focus Location at the measured quantile of stiffness. This analysis reveals that the relationship between Boundary and Stiffness found in Section 2.1 is further conditioned by the location of focus. For example, as Figure 2.8a shows, for the same decrease in stiffness, the release of IP-final C0 gestures lengthened more the further away the focus was from the boundary (i.e., in 3-NP IPs, lengthening is 3.3 ms per unit of stiffness (cm/sec) in Antepenultimate focus > 2.5 ms per unit of stiffness in Penultimate focus > 2.3 ms per unit of stiffness in Final focus; in 2-NP IPs, lengthening is 3.8 ms per unit of stiffness in Penultimate focus > 3.3 ms per unit of stiffness in Final focus). Pairwise comparisons, illustrated in Figure 2.9a, also show significantly greater duration of IP-final C0 release in Antepenultimate condition compared to both Penultimate and Final in 3-NP IPs at the 10th quantile (β = 11.2, SE = 3.2, $p < 0.01$; β = 13.83 SE = 3.2, $p < 0.001$, respectively) and the 25th quantile (β = 9.4, SE = 2.8, $p < 0.01$; β = 11.0, SE = 2.8, $p < 0.001$, respectively). Figure 2.9b plots the differences in duration between Boundary conditions, visualizing that boundary-related lengthening (i.e., durational difference between IP-final and IP-medial positions) is greater in earlier (Antepenultimate) as opposed to later focus locations (Penultimate, Final). Note that the difference between the Penultimate and Final conditions is in the same direction (i.e., Penultimate > Final), although it did not reach significance in the pairwise comparisons.

Similar results were found for the interaction in C1 Release (Figure 2.8b and Figure 2.9c, d). Post-hoc analysis at different quantiles of Stiffness detected greater difference between IP-final versus IP-medial positions in earlier as opposed to later focus locations (e.g., Antepenultimate > Penultimate > Final; Figure 2.9d). In particular, there was more lengthening in the IP-final C1 release in the Antepenultimate as opposed to the Penultimate, and also in the Penultimate as opposed to the Final. IP-final C1 release was longer in Antepenultimate as compared to Penultimate (10th quantile: β = 11.0, SE = 2.8, $p < 0.001$), although both focus conditions had similar IP-medial durations. The difference in IP lengthening between Penultimate and Final, on the other hand, stems from the fact that C1 gestures are longer IP-medially in final focus as opposed to penultimate focus (10th: β = 3.0, SE = 0.8, $p < 0.01$; 25th: β = 2.3, SE = 0.7, $p < 0.01$; 50th: β = 1.5, SE = 0.5, $p < 0.05$).

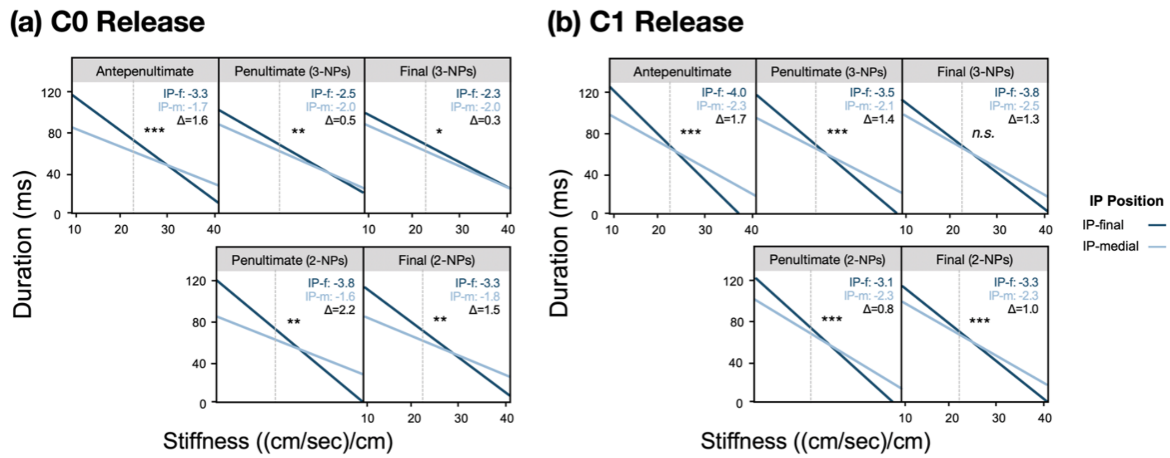

Figure 2.8: Predicted relationship between duration and stiffness by IP Position (IP-final, IP-medial) and Focus Location (Antepenultimate (3-NPs), Penultimate (2-NPs/3-NPs), Final (2-NPs/3-NPs)) for (a) C0 release and (b) C1 release. Vertical dashed lines mark the actual values of average stiffness (i.e., peak velocity/displacement) of each gesture across IP Position per Focus Location. The numbers in each plot represent the slope for each condition and their difference. Significant pairwise comparisons are marked by the means of asterisks as follows: *** = '$p < 0.001$', ** = '$p < 0.01$', * = '$p < 0.05$', and n.s. = '$p > 0.09$'.

In sum, the location of focus in the phrase systematically affected the pattern of phrase-final lengthening, such that the amount of lengthening in the gestures closest to the boundary (C1 and C0 releases) was greater when focus was further away from the boundary.

2.3.4 Interaction of IP boundary with Final AP Length

A three-way interaction between IP Position (IP-final, IP-medial), Final AP Length (Long, Short) and Stiffness was detected in C2 Release ($F(1, 1021) = 5.10$, $p < 0.05$). The post-hoc analysis detected that, in the 10th and 25th quantile of Stiffness, C2 release presented phrase-final lengthening (i.e., longer durations IP-finally than IP-medially) in long, but not in short, Final APs (Long: 10th: $\beta = 3.22$, SE = 1.13, $p < 0.01$; 25th: $\beta =$

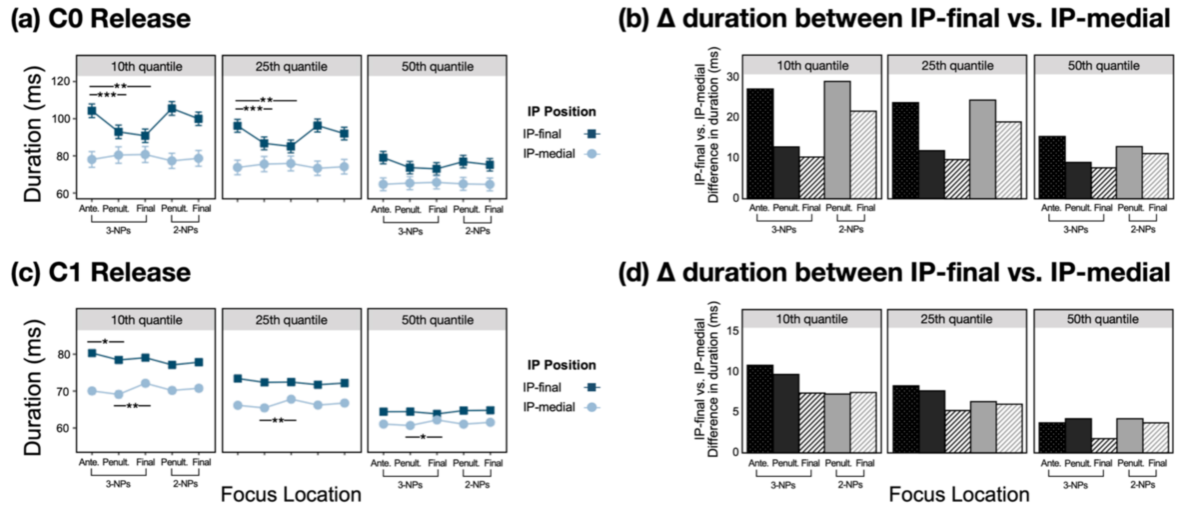

Figure 2.9: Predicted duration (in ms) of IP-final and IP-medial (a) C0 and (c) C1 releases per Focus Location at 10th, 25th, and 50th (median) quantile of Stiffness and durational difference (in ms) between IP-final and IP-medial of (b) C0 and (d) C1 releases per Focus Location. Significant pairwise comparisons are marked by the means of asterisks as follows: *** = '$p < 0.001$', ** = '$p < 0.01$', and * = '$p < 0.05$'.

1.85, SE = 0.80, $p < 0.05$; Short: 10th, 25th: $p > 0.05$). This indicates that the scope of lengthening reached the onset of the penultimate syllable (C2 release) in the Long final AP condition, extending beyond the general scope of IP-final lengthening, which was to the onset of the phrase-final syllable (C1 formation), as illustrated in Section 2.3.2. This effect holds in all focus conditions as indicated by the lack of interaction between final AP length and focus location, which would have appeared if significant enough. The fact that the length of the final AP systematically contributes to the boundary-related modulation regardless of focus location suggests that some information—albeit not tonal—of the AP's left edge remains even in cases of dephrasing. In sum, the length of the final AP systematically conditioned the scope of phrase-final lengthening, such that the effect was extended further away from the boundary the longer the final AP.

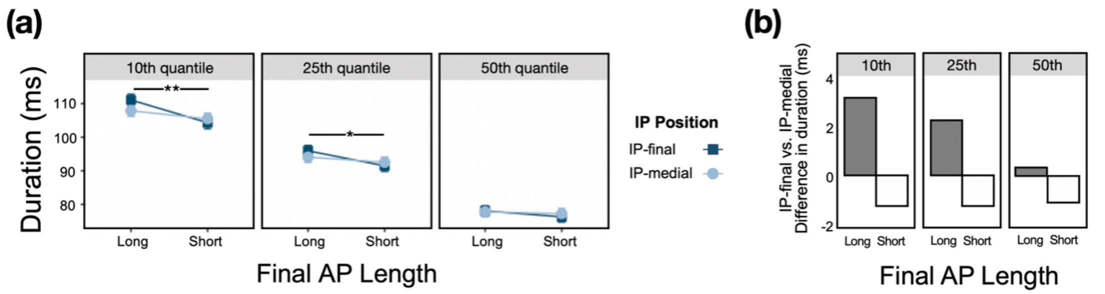

Figure 2.10: Predicted duration (in ms) of IP-final and IP-medial C2 release per Final AP Length (Long, Short) at 10th, 25th, and 50th (median) quantile of Stiffness and (b) durational difference (in ms) between IP-final and IP-medial per Final NP Length. Significant pairwise comparisons are marked by the means of asterisks as follows: ** = '$p < 0.01$' and * = '$p < 0.05$'.

2.3.5 Effects of IP boundary on displacement

Displacement was found to increase with peak velocity in all C gestures (Figure 2.11). A systematic effect of IP Position was observed on C0 ($F(1, 1028) = 96.34$, $p < 0.001$) and C2 ($F(1, 1021) = 4.18$, $p < 0.05$) with greater displacement in IP-final than in IP-medial conditions (Figure 2.12). A marginally significant effect of boundary with an opposite pattern was found in C4, with greater displacement in IP-medial positions as opposed to IP-final positions ($F(1, 1024) = 3.23$, $0.05 < p < 0.08$). Although marginal, this result is interesting in the sense that a similar, but significant, effect was observed in the formation duration of C4, which is the initial consonant of the IP-final prosodic word. Note that these results are present beyond the inherent relationship between displacement and peak velocity. To summarize, displacement increases with increased peak velocity as expected. Phrase-final boundary affects the final C gesture, such that the displacement is greater in IP-final than in IP-medial conditions. A marginal effect with an opposite pattern (i.e., greater displacement in IP-medial than in IP-final) was detected in the initial consonant of the phrase-final prosodic word.

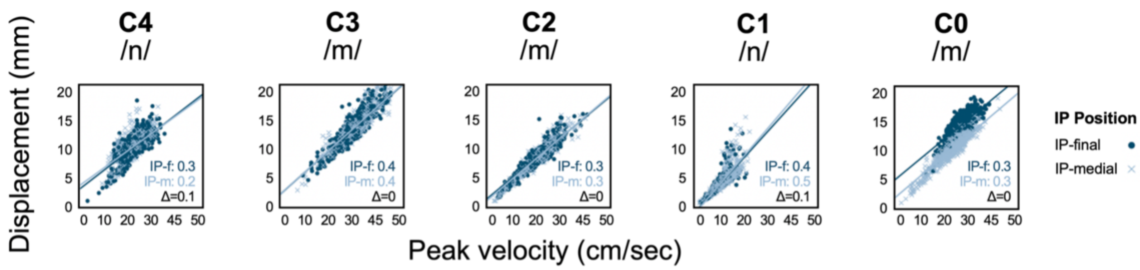

Figure 2.11: Predicted (regression line) and observed (scatter plot) relationship between peak velocity and displacement by IP Position (IP-final and IP-medial) for each gesture. The numbers in each plot represent the slope for each condition and their difference.

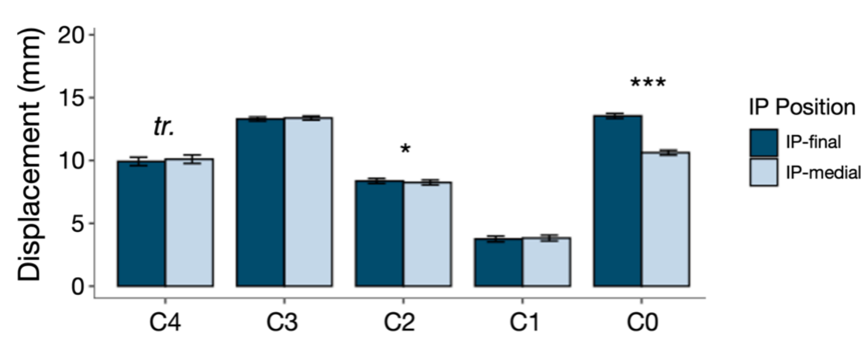

Figure 2.12: Displacement (in mm) of gestural formation of the test C gestures (C4, C3, C2, C1, and C0) in /nɛmaŋminam/ (as a reminder, /ŋ/ is excluded). Dark bars correspond to IP-final gestures and light bars to IP-medial ones. Significant pairwise comparisons are marked by the means of asterisks as follows: *** = '$p < 0.001$', * = '$p < 0.05$', and tr. = '$0.05 < p < 0.08$'.

2.3.6 Effects of IP boundary on peak velocity

Peak velocity was found to increase with increase in displacement in all C gestures (Figure 2.13). A main effect of IP Position was observed on the C0 gesture ($F(1, 1031) = 7.63$, $p < 0.01$), with lower peak velocity, indicating slower movement, in IP-final than in IP-medial positions (Figure 2.14). There was also a significant IP Position and Displacement interaction on C0 ($F(1, 1031) = 78.16$, $p < 0.001$): The slope of the regression is greater in IP-medial condition (2.1) than in IP-final (1.6) condition, indicating that given a same displacement, peak velocity of a C gesture is lower before an IP boundary. In addition, an interaction between IP Position, Focus location, and Final AP length was observed in C0 gesture ($F(4, 1026) = 2.98$, $p < 0.05$). In terms of the effect of IP Position, pair-wise comparisons revealed that peak velocity is lower in IP-final than in IP-medial conditions for all possible combinations of Focus Location and Final AP length conditions (Antepenultimate, Long: β = -4.22, SE = 0.53, $p < 0.001$; Penultimate (3-NPs), Long: β = -3.18, SE = 0.65, $p < 0.001$; Penultimate (2-NPs), Long: β = -3.18, SE = 0.58, $p < 0.001$; Final (3-NPs), Long: β = -3.48, SE = 0.53, $p < 0.001$; Final (2-NPs), Long: β = -5.51, SE = 0.56, $p < 0.001$; Antepenultimate, Short: β = -3.32, SE = 0.59, $p < 0.001$; Penultimate (3-NPs), Short: β = -5.69, SE = 0.60, $p < 0.001$; Penultimate (2-NPs), Short: β = -5.23, SE = 0.58, $p < 0.001$; Final (3-NPs), Short: β = -4.0, SE = 0.64, $p < 0.001$; Final (2-NPs), Short: β = -5.77, SE = 0.57, $p < 0.001$). The interaction seems to have stemmed from the fact that peak velocity is lower for Long than for Short final APs when Focus Location was on penultimate AP in the IP-medial conditions. In sum, peak velocity increases with increased displacement as expected. Presence of a phrase-final boundary affects the final C gesture, such that the gesture is slower in IP-final position than in IP-medial position.

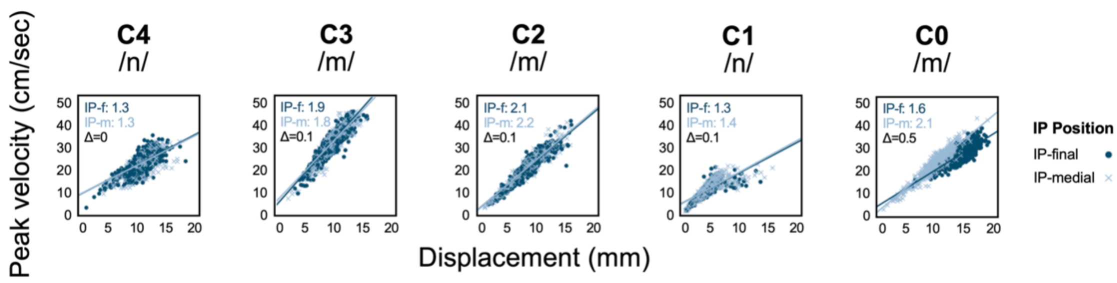

Figure 2.13: Predicted (regression line) and observed (scatter plot) relationship between peak velocity and displacement by IP Position (IP-final and IP-medial) for each gesture. The numbers in each plot represent the slope for each condition and their difference.

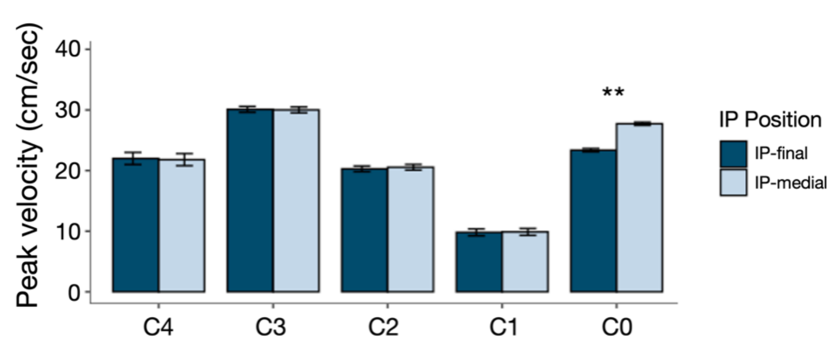

Figure 2.14: Peak velocity (in cm/sec) of the gestural formation of the test C gestures (C4, C3, C2, C1, and C0) in /nɛmaŋminam/ (as a reminder, /ŋ/ is excluded). Dark bars correspond to IP-final gestures and light bars to IP-medial ones. Significant pairwise comparisons are marked by the means of asterisks as follows: ** = ‘$p < 0.01$’.

2.4 Discussion

2.4.1 The kinematics of boundaries in Korean

One of the main goals of the present study was to examine the scope of IP boundary lengthening in Korean. Results showed that gestures preceding IP boundaries in Korean were indeed longer in duration than their counterparts in phrase-medial positions. This confirms the predominant prediction for phrase-final lengthening on the basis of numerous findings in previous literature, according to which linguistic units (gestures, segment, syllables, words) are lengthened phrase-finally (e.g., Byrd et al., 2006; Edwards et al., 1991; Fletcher, 2010). The general pattern is for pre-boundary lengthening in Korean to extend over the final syllable of the phrase-final word. Similar to previous findings reported in the literature, the final rhyme of the phrase was affected the most (e.g., Byrd et al., 2006; Byrd & Riggs, 2008; Krivokapić, 2007; Katsika, 2016; see also acoustic studies, e.g., Berkovits,1994; Cambier-Langeveld, 1997; Campbell & Isard, 1991; Nakatani et al., 1981; Oller, 1973; Wightman et al., 1992). In our data, lengthening systematically extended beyond the final rhyme, reaching the onset of the final syllable. This pattern slightly diverges from previous work on Korean, in which lengthening was found to affect a longer stretch, encompassing not only the final syllable, but the antepenultimate syllable as well. The test words there were disyllabic, and the difference in the patterns found in that study and ours seems to stem from the different prosodic and rhythmic structures of the stimuli, since as found here and discussed in detail in the next section, the prosodic factors of focus location and AP boundary distribution cause pre-boundary lengthening to stretch further away from the IP boundary and towards the beginning of the phrase-final word. Regardless, both studies indicate that phrase-final lengthening in Korean can influence a relatively long stretch of speech, extending over one or even two syllables.

In terms of kinematic manifestation at the boundary, gestures, in addition to being longer, were also larger and slower. Coda consonants of phrase-final words presented the most systematic patterns, showing effects on all three dimensions: longer duration, larger displacement, and slower peak velocity as opposed to their phrase-medial counterparts. Our data adds to previous findings (e.g., Byrd & Saltzman, 2003; Byrd et al., 2000, Cho, 2005, 2006), further indicating that, similarly to head-prominence languages, edge-prominence languages mark IP boundaries by the means of longer, larger and slower gestures. This might seem to contradict previous work on Korean by Kim et al. (2019) that has reported longer, larger but faster gestures IP-finally. However, we attribute the difference between these results to the fact that the analyses used in the current study take into account known relationships of interdependence between kinematic parameters (i.e., duration-to-stiffness and peak velocity-to-displacement). With respect to peak velocity specifically, our analyses detected that boundary types had distinct peak velocity-to-displacement profiles (see e.g., separate peak velocity/displacement slopes by boundary type for C0 in Figure 13), and it is when considering these distinct profiles that gestures are slower IP-finally as opposed to IP-medially.

Kinematic signatures of longer, larger (a.k.a. less overlapped), and slower gestures found phrase-finally have been accounted for by the pi-gesture model (Byrd & Saltzman, 2003). As a reminder, a pi-gesture is not connected to a specific track variable, but is instead assumed be a modulation gesture slowing down the time course of all constriction gestures co-active with it. This causes affected gestures to be longer (increased duration) and slower (lower peak velocity) resulting in change in both duration and relative timing, shown as decreased gestural overlap (cf. simulations in Byrd & Saltzman, 2003). In turn, decrease in overlap grants constriction gestures more time to faithfully realize their targets without being truncated, resulting in larger displacement. Although these properties of the pi-gesture can account for the patterns of longer, larger, and slower gestures found in our data, it is unclear how they can address the distinct duration-to-stiffness and peak velocity-to-displacement profiles that clearly characterize Korean boundaries. Similar questions for the pi-gesture model have also been raised by Iskarous and Pouplier (2022), where it is pointed out that the positive correlation between displacement and peak velocity in empirical research contradicts the pi-gesture's prediction of inverse relation between the two parameters that results from clock-slowing. Moreover, framed under Articulatory Phonology (Browman & Goldstein, 1992) and Task Dynamics (Saltzman & Munhall, 1989), pi-gesture affect gestures that are modeled as linear dynamical systems (cf. Sorensen & Gafos, 2016). However, the results of the present study also add to the discussion of whether gestures should be represented as a non-linear autonomous dynamical models rather than linear non-autonomous ones (Sorensen & Gafos, 2016; see Iskarous & Pouplier, 2022 for a review). As also discussed in Iskarous and Pouplier (2022), these are not reasons to reject the pi-gesture model, the benefits of which are significant for research in prosody, but should be used as pointers for the mathematical development of the model.

2.4.2 Interactions with Focus Location and Final AP Length

Another important question for the current study was to assess if and how the functions that are served by lexical stress in stress languages (i.e., phrase-level prominence anchor, word demarcation, word prosody marker) interact with the kinematic marking of IP boundaries in Seoul Korean. To that end, we examined the effect of focus location as an index of phrase-level prominence and the number of syllables in the final AP as an index of word demarcation. Word prosody was not considered per se since Korean does not have any. Interestingly, our results reveal a degree of interaction of IP-final length-

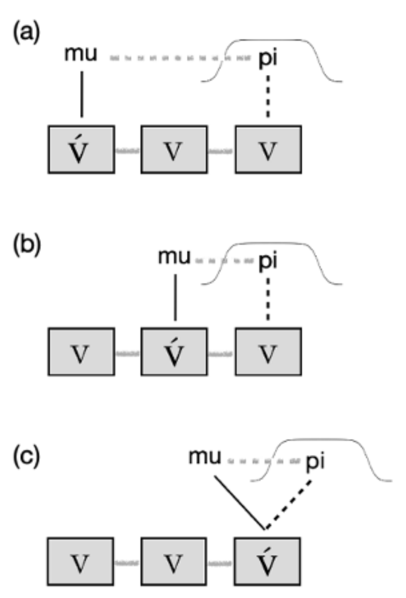

Figure 2.15: Schematic representation of the dual coordination of pi-gestures with phrase-final V and mu-gestures in trisyllabic stress-initial (a), stress-medial (b), and stress-final (c) words (adopted from Katsika, 2016 and Katsika & Tsai, 2021). Solid lines represent in-phase relationships and dashed lines represent anti-phase relationships.

ening with both these factors. The direction of both effects is similar to that of lexical stress in stress languages: IP-final lengthening begins earlier in phrases with earlier focus location/final AP boundary/stress. In previous work, the effect of stress was accounted for through a dual coordination of the pi-gesture with a) the final syllable's V gesture and b) the mu-gesture triggering lexical stress (cf. Figure 2.15, as adopted by Katsika, 2016). Competition between these two coordination relationships was proposed to cause the earlier initiation of boundary related events in non-final stressed words compared to final stressed words.

To extend this account to Korean, we first need to assume that both focus and AP boundary marking are instantiated by their own mu-gestures—actually, the AP boundary would itself be more accurately captured by a pi-gesture. There is evidence indicating that phrase-level prominence/focus is marked by kinematic modulations, and specifically by the means of longer, larger and faster gestures (e.g., Jang & Katsika, 2023; Im et al., 2023), and these effects can rise from mu-gestures. We propose that these mu-gestures be in-phase with the focused AP's initial constriction gesture, since the kinematic effects of prominence are first observed in the initial segment of the focused AP. As for AP boundaries, although they are undoubtedly related to a tonal form, little is known about their kinematic profile, and connecting them to a pi-gesture requires further investigation. Yet, it is worth noting that limited current data suggest that pre-focal and dephrased parts of a phrase don't differ kinematically from each other (e.g., Jang & Katsika, 2023; Im et al., 2023), suggesting that dephrasing does not result in complete removal of AP boundaries, but that it mainly affects the tonal dimension, allowing for kinematic markers of AP boundaries. The latter could in principle be captured by a pi-gesture coordinated in-phase with the AP- initial constriction gesture. That dephrasing is not complete is further corroborated by the findings here, according to which the length of the last AP affected the scope of pre-boundary lengthening regardless of its dephrasing status (i.e., the effect held in all focus conditions, even when the last AP was not focused). This suggests that AP boundaries may be available to the speakers, albeit in an attenuated form, in the context of dephrasing.

Assuming the two prosodic gestures (i.e., a prominence-related mu-gesture and an AP-related pi-gesture), the pi-gesture for the IP boundary is then coordinated with both of them, but also with the phrase-final syllable. Figure 2.16 illustrates these coordination relationships. There are some points that the reader should keep in mind about the figure. First, the coordination with the focus mu-gesture and that with the final AP pi-gesture are provided in different panels in order to highlight their individual effects, but the reader should assume that these coordination links are in effect simultaneously. Second, only C gestures are presented, because the analyses here addressed only those. The contribution of the V gestures to these coordination relationships is still to be assessed. Finally, based on currently available data, it is unclear what type of coordination could characterize the timing relationship between the IP pi-gesture and either the focus mu-gestures or the AP pi-gesture. To denote this uncertainty, Figure 2.16 presents those coordination links as thick, grayed out lines.

A coordination relationship with the phrase-final syllable is proposed because pre-boundary lengthening systematically affects it as a whole. Thus, contrary to the pi-gesture model for Greek boundaries, which involved anti-phase coordination of the pi-gesture with the phrase-final V gesture, the Korean model considers in-phase coordination with the constriction gestures in the onset (and possibly in the nucleus as well) of the phrase-final syllable. As in the model of Greek boundaries, the coordination of the pi-gesture with the phrase-final syllable is stronger as compared to its coordination to either of the prosodic gestures (i.e., the prominence-instantiating mu-gesture and the AP pi-gesture). This weaker coordination with the prominence-instantiating mu-gesture can account for the observation that the formation of the final syllable's coda becomes longer when focus is further away from the IP boundary, i.e., when focus is not final: As a result of this weaker coordination, the pi-gesture is "attracted" towards the focus-marking mu-gesture, but does not fully overlap with it. This attraction causes the pi-gesture to be initiated earlier as compared to when focus is final, and, consequently, to reach their 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 would thus lengthen more (see Figure 2.16a).

As for the interaction with the final AP, the weaker coordination between the IP and AP pi-gestures (as opposed to that between the IP pi-gesture and the final syllable) can explain the shifts in the scope of pre-boundary lengthening: This weaker coordination "attracts", but does not pull fully, the IP boundary-instantiating pi-gesture towards the beginning of the final AP, and as a result pre-boundary lengthening extends further away from the IP boundary when the final AP starts further away from it (i.e., when the final AP consists of more syllables). Figure 2.16b illustrates these shifts. The coordination between the two pi-gestures seems to be stronger than that between the IP pi-gesture and the focus-inducing mu-gesture, since the former results in a more extended scope of the effect.

Finally, the fact that no interaction was observed between the effect of IP boundary and focus location or the size of the final AP on the dimensions of displacement and (peak) velocity indicates that any coordination is at the level of temporal modulation (i.e., it concerns pi-gesture and temporal mu-gestures; regarding the distinction between temporal and spatial mu-gestures, the reader is referred to Saltzman et al., 2008).

In sum, in Korean, IP boundary marking has an intricate relationship with both prominence marking and boundary marking at other prosodic levels. Assuming that the marking of the AP level is connected to word demarcation, then the relationship between pre-boundary lengthening and the final AP would suggest that the prosodic structure has early access to the lexical items of the utterance under planning. Further support for this conclusion comes from the patterns of boundary-related shortening, discussed in the following section.

2.4.3 IP boundary-related shortening

In addition to the IP-final lengthening extending over the final syllable of the phrase-final word (and to the onset of the penultimate syllable when the final AP was long), a small but significant IP-related shortening effect was detected on the initial C gesture of the phrase-final word. Boundary-related shortening effects have previously been reported in the literature on either side of the boundary (e.g., Kim et al., 2017; Katsika, 2016; Byrd et al., 2006). Post-boundary, the shortening effect has been shown to be system-

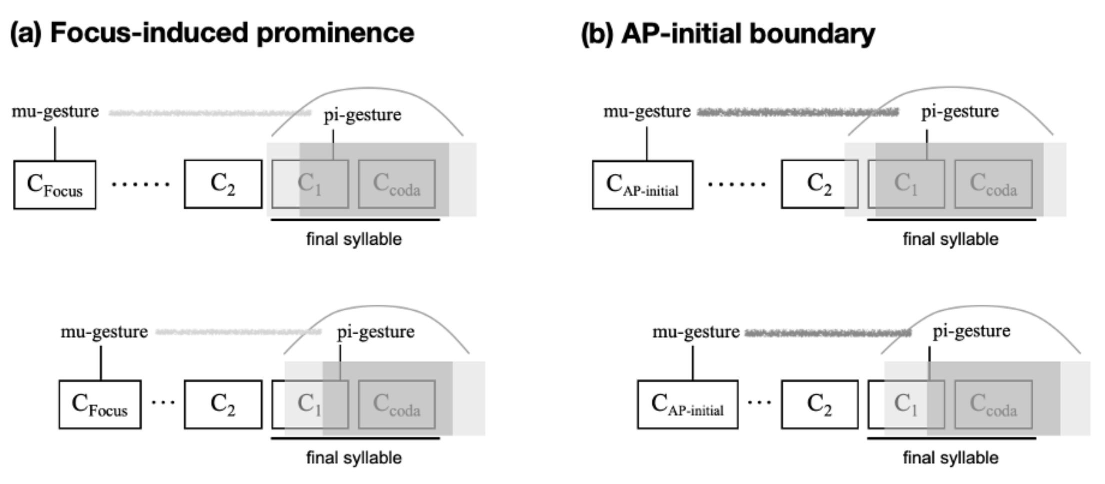

Figure 2.16: Schematic representation of the dual coordination of pi-gestures with phrase-final coda C and mu-gestures: (a) focus-induced prominence; (b) AP-initial boundary. The gray-shaded boxes indicate the effect of pi-gesture, with the darker box representing stronger activation.

atic and has been 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 byproduct of the coordination between the stress-instantiating mu-gesture and the pi-gesture, which attracted the two gestures towards each other. In our data, shortening is not conditioned by the position of the equivalent factors, i.e., the position of focus and/or the position of the final AP's initial boundary. Instead, the effect is consistently found on the initial consonant of the phrase-final prosodic word (i.e., initial /n/ in /nɛmaŋminam/). Recall that the onset of the prosodic word showed, in addition to this significant shortening effect, a marginal effect on the dimension of displacement, and once again that held regardless of other factors examined. This result may suggest that the lexical items that feed into the domain of the prosodic word are accessible early to prosodic speech planning (cf. Keating & Shattuck-Hufnagel, 2002 vs. Levelt, 1989).

2.5 Conclusion

The present study examined the articulatory correlates of IP-final boundary in Seoul Korean, with the goal to investigate the scope of the boundary-related effect and how it interacts with other grammatical domains and/or prosodic levels. Gestures in Korean are in general longer, larger, and slower at boundaries beyond inherent interdependencies between kinematic parameters, i.e., the relationship between duration and stiffness and that between displacement and peak velocity. Pre-boundary lengthening systematically affected the phrase-final syllable, with both focus location and final AP fine-tuning its manifestation. The amount of IP-final lengthening was greater the further away the focus was from the boundary, and the scope of lengthening extended to the onset of the penultimate syllable of the phrase-final prosodic word when the final AP was long. Combined the results have implications for our understandings of the prosodic control of gestures that materializes boundary-related lengthening. In particular, they suggest that control of boundary-related lengthening reported in previous studies is not restricted to word-level prosody per se (Katsika, 2016; Katsika & Tsai, 2021), but that it extends to both phrase-level prominence as well as factors that undertake the function of word demarcation in languages with no word prosody.

2.A Appendix

Focus Location	Final AP Length	IP Position	Test sentence (# = IP boundary)
Ante-penultimate (3-NPs)	*Long*	IP-final	Prompt: (It's not the magic club.) [IP [NP **pimilpu]** [NP minamiga] [NP matɕimak nɛmaŋminam]]? # [IP [VP sʌntækhangʌja]]? Minam from **the secrecy club** is the last handsome guy from Nemang? Is it decided?
		IP-medial	Prompt: (It's not the magic club.) [IP [NP **pimilpu]** [NP minamiga] [NP matɕimak nɛmaŋminam] [VP sʌntækhangʌja]]? Minam from **the secrecy club** chose the handsome guy from Nemang?
	Short	IP-final	Prompt: (It's not the magic club.) [IP [NP **pimilpu]** [NP minami gomobuga] [NP nɛmaŋminam]]? # [IP [VP sʌntækhangʌja]]? Uncle Minam from **the secrecy club** is the handsome guy from Nemang? Is it decided?
		IP-medial	Prompt: (It's not the magic club.) [IP [NP **pimilpu]** [NP minami gomobuga] [NP nɛmaŋminam] [VP sʌntækhangʌja]]? Uncle Minam from **the secrecy club** chose the handsome guy from Nemang?
Penultimate (3-NPs)	Long	IP-final	Prompt: (It's not Junseok.) [IP [NP pimilpu**]** [NP **minamiga**] [NP matɕimak nɛmaŋminam]]? # [IP [VP sʌntækhangʌja]]? **Minam** from the secrecy club is the last handsome guy from Nemang? Is it decided?
		IP-medial	Prompt: (It's not Junseok.) [IP [NP pimilpu**]** [NP **minamiga**] [NP matɕimak nɛmaŋminam] [VP sʌntækhangʌja]]? **Minam** from the secrecy club chose the handsome guy from Nemang?

	Short	IP-final	Prompt: (It's not Uncle Junseok.) [$_{IP}$ [$_{NP}$ pimilpu] [$_{NP}$ **minami** gomobuga] [$_{NP}$ nɛmaŋminam]]? # [$_{IP}$ [$_{VP}$ sʌntækhangʌja]]? Uncle **Minam** from the secrecy club is the handsome guy from Nemang? Is it decided?
		IP-medial	Prompt: (It's not Uncle Junseok.) [$_{IP}$ [$_{NP}$ pimilpu] [$_{NP}$ **minami** gomobuga] [$_{NP}$ nɛmaŋminam] [$_{VP}$ sʌntækhangʌja]]? Uncle **Minam** from the secrecy club chose the handsome guy from Nemang?
Final (3-NPs)	Long	IP-final	Prompt: (It's not the first.) [$_{IP}$ [$_{NP}$ pimilpu] [$_{NP}$ minamiga] [$_{NP}$ **matɕimak** nɛmaŋminam]]? # [$_{IP}$ [$_{VP}$ sʌntækhangʌja]]? Minam from the secrecy club is **the last** handsome guy from Nemang? Is it decided?
		IP-medial	Prompt: (It's not the first.) [$_{IP}$ [$_{NP}$ pimilpu] [$_{NP}$ minamiga] [$_{NP}$ **matɕimak** nɛmaŋminam] [$_{VP}$ sʌntækhangʌja]]? Minam from the secrecy club chose **the last** handsome guy from Nemang?
	Short	IP-final	Prompt: (It's not 'the handsome guy from Nowon'.) [$_{IP}$ [$_{NP}$ pimilpu] [$_{NP}$ minami gomobuga] [$_{NP}$ **nɛmaŋminam**]]? # [$_{IP}$ [$_{VP}$ sʌntækhangʌja]]? Uncle Minam from the secrecy club is the handsome guy from **Nemang**? Is it decided?
		IP-medial	Prompt: (It's not 'the handsome guy from Nowon'.) [$_{IP}$ [$_{NP}$ pimilpu] [$_{NP}$ minami gomobuga] [$_{NP}$ **nɛmaŋminam**] [$_{VP}$ sʌntækhangʌja]]? Uncle Minam from the secrecy club chose the handsome guy from **Nemang**?
Penultimate (2-NPs)	*Long*	IP-final	Prompt: (It's not Junseok.) [$_{IP}$ [$_{NP}$ **minamiga**] [$_{NP}$ matɕimak nɛmaŋminam]]? # [$_{IP}$ [$_{VP}$ sʌntækhangʌja]]? Minam from (as opposed to Junseok) is the last handsome guy from Nemang? Is it decided?

		IP-medial	Prompt: (It's not Junseok.) [$_{IP}$ [$_{NP}$ **minamiga**] [$_{NP}$ matɕimak <u>nɛmaŋminam</u>] [$_{VP}$ sʌntækhangʌja]]? Minam from (as opposed to Junseok) chose the last handsome guy from Nemang?
	Short	IP-final	Prompt: (It's not Uncle Junseok.) [$_{IP}$ [$_{NP}$ **minami** gomobuga] [$_{NP}$ <u>nɛmaŋminam</u>]]? # [$_{IP}$ [$_{VP}$ sʌntækhangʌja]]? Uncle Minam is the handsome guy from **Nemang** (as opposed to Nowon)? Is it decided?
		IP-medial	Prompt: (It's not Uncle Junseok.) [$_{IP}$ [$_{NP}$ **minami** gomobuga] [$_{NP}$ <u>nɛmaŋminam</u>] [$_{VP}$ sʌntækhangʌja]]? Uncle Minam chose the handsome guy from **Nemang** (as opposed to Nowon)?
Final (2-NPs)	*Long*	IP-final	Prompt: (It's not the first.) [$_{IP}$ [$_{NP}$ minamiga] [$_{NP}$ **matɕimak** <u>nɛmaŋminam</u>]]? # [$_{IP}$ [$_{VP}$ sʌntækhangʌja]]? Minam is **the last** (as opposed to the first) handsome guy from Nemang? Is it decided?
		IP-medial	Prompt: (It's not the first.) [$_{IP}$ [$_{NP}$ minamiga] [$_{NP}$ **matɕimak** <u>nɛmaŋminam</u>] [$_{VP}$ sʌntækhangʌja]]? Minam chose **the last** (as opposed to the first) handsome guy from Nemang?
	Short	IP-final	Prompt: (It's not 'the handsome guy from Nowon'.) [$_{IP}$ [$_{NP}$ minami gomobuga] [$_{NP}$ <u>**nɛmaŋminam**</u>]]? # [$_{IP}$ [$_{VP}$ sʌntækhangʌja]]? Uncle Minam is the handsome guy from **Nemang** (as opposed to Nowon)? Is it decided?
		IP-medial	Prompt: (It's not 'the handsome guy from Nowon'.) [$_{IP}$ [$_{NP}$ minami gomobuga] [$_{NP}$ <u>**nɛmaŋminam**</u>] [$_{VP}$ sʌntækhangʌja]]? Uncle Minam chose the handsome guy from **Nemang** (as opposed to Nowon)?

Table 2.2: Stimuli sentences organized by Focus Location (Antepenultimate (3-NPs), Penultimate (3-NPs), Final (3-NPs/2-NPs)), Final AP length (Short vs. Long), and IP Position (IP-final vs. IP-medial). Measured prosodic words are underlined, focused NPs are shown in bold. All sentences were preceded by a short IP meaning 'Really?'.

References

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

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: 10.18637/jss.v067.i01

Beckman, M. E., Edwards, J., & Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology ii: Gesture, segment, prosody* (pp. 68–89). Camdrige, UK: Cambridge University Press.

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

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

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, *49*(3-4), 155–180.

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

Byrd, D., Kaun, A., Narayanan, S., & Saltzman, E. (2000). Phrasal signatures in articulation. In M. B. Broe & J. B. Pierrehumbert (Eds.), *Papers in laboratory phonology v: Acquisition and the lexicon* (pp. 70–87). Cambridge, UK: Cambridge University Press.

Byrd, D., Krivokapić, J., & Lee, S. (2006). How far, how long: On the temporal scope of prosodic boundary effects. *The Journal of the Acoustical Society of America*, *120*(3), 1589–1599.

Byrd, D., Lee, S., Riggs, D., & Adams, J. (2005). Interacting effects of syllable and phrase position on consonant articulation. *The Journal of the Acoustical Society of America*, *118*(6), 3860–3873.

Byrd, D., & Riggs, D. (2008). Locality interactions with prominence in determining the scope of phrasal lengthening. *Journal of the International Phonetic Association*, *38*(2), 187–202.

Byrd, D., & Saltzman, E. (1998). Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics*, *26*(2), 173–199.

Byrd, D., & Saltzman, E. (2003). The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, *31*(2), 149–180.

Cambier-Langeveld, T. (1997). The domain of final lengthening in the production of Dutch. *Linguistics in the Netherlands*, *14*(1), 13–24.

Campbell, W. N., & Isard, S. D. (1991). Segment durations in a syllable frame. *Journal of phonetics*, *19*(1), 37–47.

Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /a, i/ in English. *The Journal of the Acoustical Society of America*, *117*(6), 3867–3878.

Cho, T. (2006). Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English. *Laboratory phonology*, *8*, 519–548.

Cho, T., & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, *37*(4), 466–485.

Cho, T., & Keating, P. A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of phonetics*, *29*(2), 155–190.

Cho, T., & Sun-Ah, J. (2000). Domain-initial strengthening as featural enhancement: Aerodynamic evidence from Korean. In *Chicago linguistics society* (Vol. 36, pp.

31–44).

Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and speech.* Harvard University Press.

Cutler, A., & Butterfield, S. (1991). Word boundary cues in clear speech: A supplementary report. *Speech Communication*, *10*(4), 335-353.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(1), 113–121.

Edwards, J., Beckman, M. E., & Fletcher, J. (1991). The articulatory kinematics of final lengthening. *the Journal of the Acoustical Society of America*, *89*(1), 369–382.

Fletcher, J. (2010). The prosody of speech: Timing and rhythm. In *The handbook of phonetic sciences* (p. 521-602). John Wiley & Sons, Ltd.

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

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *The journal of the acoustical society of America*, *101*(6), 3728–3740.

Goldstein, L., Byrd, D., & Saltzman, E. (2006). The role of vocal tract gestural action units in understanding the evolution of phonology. *Action to language via the mirror neuron system*, 215–249.

Gussenhoven, C., & Chen, A. (2021). *The oxford handbook of language prosody.* Oxford, UK: Oxford University Press.

Hawkins, S. (1992). An introduction to task dynamics. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology ii: Gesture, segment, prosody* (pp. 9–25). Camdrige, UK: Cambridge University Press.

Hayes, B. (1989). The prosodic hierarchy in meter. In *Rhythm and meter* (pp. 201–260). Elsevier.

Im, S., Kim, S., & Cho, T. (2023). Some asymmetrical pre- versus post-focal effects on articulatory realization of prominence distribution in Korean: A preliminary report. In *Proceedings of hanyang international symposium on phonetics and cognitive sciences of language (hisphoncog) 2023* (pp. 206–207).

Iskarous, K., & Pouplier, M. (2022). Advancements of phonetics in the 21st century: A critical appraisal of time and space in articulatory phonology. *Journal of Phonetics*, *95*, 101195.

Jang, J., & Katsika, A. (2019). Articulatory characteristics of focus in Korean. *The Journal of the Acoustical Society of America*, *146*(4), 3084–3084.

Jang, J., & Katsika, A. (2022). The coordination of boundary tones with constriction gestures in Seoul Korean, an edge-prominence language. *Proceedings of the International Conference on Speech Prosody 2022*, 145–149.

Jang, J., & Katsika, A. (2023). Edgy articulation: the kinematic profile of accentual phrase boundaries in Seoul Korean. In *Proceedings of hanyang international symposium on phonetics and cognitive sciences of language (hisphoncog) 2023* (pp. 204–205).

Jeon, H.-S., & Nolan, F. (2017, January). Prosodic marking of narrow focus in Seoul Korean. *Laboratory Phonology*, *8*(1), 1–30.

Jun, S.-A. (1993). *The phonetics and phonology of Korean prosody* (PhD 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 Sciences*, *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* (pp. 483–486).

Jun, S.-A. (2005). Korean intonational phonology and prosodic transcription. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 201–229). Oxford, UK: Oxford University Press.

Jun, S.-A. (2006). Intonational phonology of Seoul Korean revisited. In *Japanese-korean linguistics* (Vol. 14, pp. 15–26).

Jun, S.-A. (2014). Prosodic typology: By prominence type, word prosody, and macro-rhythm. In *Prosodic typology ii: The phonology of intonation and phrasing* (pp. 520–539). Oxford, UK: Oxford University Press.

Kang, Y. (2014). Voice onset time merger and development of tonal contrast in Seoul Korean stops: A corpus study. *Journal of Phonetics*, *45*, 76–90.

Katsika, A. (2009). Boundary- and prominence-related lengthening and their interaction. *The Journal of the Acoustical Society of America*, *125*, 2572-2572.

Katsika, A. (2016). The role of prominence in determining the scope of boundary-related lengthening in Greek. *Journal of phonetics*, *55*, 149–181.

Katsika, A., Krivokapić, J., Mooshammer, C., Tiede, M., & Goldstein, L. (2014). The coordination of boundary tones and its interaction with prominence. *Journal of Phonetics*, *44*, 62–82.

Keating, P., Cho, T., Fougeron, C., & Hsu, C. (2003). Domain-initial strengthening in four languages. In J. Local, R. Ogden, & R. Temple (Eds.), *Papers in laboratory phonology vi: Phonetic interpretations* (pp. 145–163). Camdrige, UK: Cambridge University Press.

Keating, P., & Shattuck-Hufnagel, S. (2002). A prosodic view of word form encoding for speech production. *UCLA working papers in phonetics*, 112–156.

Kim, J. J., Baek, Y., Cho, T., & Kim, S. (2019). Preboundary lengthening and boundary-related spatial expansion in Korean. In *Proceedings of the 19th international congress of phonetic sciences* (pp. 1–5).

Kim, S. (2004). *The role of prosodic phrasing in Korean word segmentation* (PhD dissertation). University of California, Los Angeles.

Kim, S., Broersma, M., & Cho, T. (2012). The use of prosodic cues in learning new words in an unfamiliar language. *Studies in Second Language Acquisition*, *34*(3), 415–444.

Kim, S., & Cho, T. (2009). The use of phrase-level prosodic information in lexical segmentation: Evidence from word-spotting experiments in Korean. *The Journal of the Acoustical Society of America*, *125*(5), 3373–3386.

Kim, S., Jang, J., & Cho, T. (2017). Articulatory characteristics of preboundary lengthening in interaction with prominence on tri-syllabic words in American English. *The Journal of the Acoustical Society of America*, *142*(4), EL362–EL368.

Klatt, D. H. (1975). Vowel lengthening is syntactically determined in a connected discourse. *Journal of phonetics*, *3*(3), 129–140.

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* (PhD dissertation). University of Southern California.

Lehiste, I. (1972). The timing of utterances and linguistic boundaries. *The Journal of the Acoustical Society of America*, *51*(6B), 2018–2024.

Lenth, R. V. (2020). *emmeans: Estimated marginal means, aka least-squares means.* Retrieved from `https://CRAN.R-project.org/package=emmeans` (R package)

Levelt, W. J. (1993). *Speaking: From intention to articulation.* Cambridge, MA: MIT press.

Munhall, K. G., Ostry, D. J., & Parush, A. (1985). Characteristics of velocity profiles of speech movements. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(4), 457.

Nakatani, L. H., O'Connor, K. D., & Aston, C. H. (1981). Prosodic aspects of American

English speech rhythm. *Phonetica*, *38*(1-3), 84–105.

Nespor, M., & Vogel, I. (2007). *Prosodic phonology: with a new foreword* (Vol. 28). Mouton de Gruyter.

Oller, D. K. (1973). The effect of position in utterance on speech segment duration in English. *The journal of the Acoustical Society of America*, *54*(5), 1235–1247.

Ostry, D. J., & Munhall, K. G. (1985). Control of rate and duration of speech movements. *The Journal of the Acoustical Society of America*, *77*(2), 640–648.

Pierrehumbert, J., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology ii: Gesture, segment, prosody* (p. 90–127). Camdrige, UK: Cambridge University Press.

R Core Team. (2021). *R: A language and environment for statistical computing.* Vienna, Austria. Retrieved from https://www.R-project.org/

Roon, K. D., Gafos, A. I., Hoole, P., & Zeroual, C. (2007). Influence of articulator and manner on stiffness. In *Proceedings of the 16th international congress of phonetic sciences* (pp. 409–412).

Saltzman, E., & Byrd, D. (2000). Task-dynamics of gestural timing: Phase windows and multifrequency rhythms. *Human Movement Science*, *19*(4), 499–526.

Saltzman, E., Nam, H., Krivokapic, J., & 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* (pp. 175–184).

Selkirk, E. (1984). *Phonology and syntax: The relationship between sound and structure.* Cambridge, MA: MIT Press.

Seo, J., Kim, S., Kubozono, H., & Cho, T. (2019). Preboundary lengthening in japanese: To what extent do lexical pitch accent and moraic structure matter? *The Journal of the Acoustical Society of America*, *146*(3), 1817–1823.

Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of psycholinguistic research*, *25*, 193–247.

Shin, S., Kim, S., & 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.*

Sorensen, T., & Gafos, A. (2016). The gesture as an autonomous nonlinear dynamical system. *Ecological Psychology*, *28*(4), 188–215.

Tabain, M. (2003). Effects of prosodic boundary on/ac/sequences: articulatory results. *The Journal of the Acoustical Society of America*, *113*(5), 2834–2849.

Tsai, K. (2023). *Edgy prosody: An articulatory investigation of the role of lexical pitch accent in Tokyo Japanese boundary marking* (PhD dissertation). University of California, Santa Barbara.

Tsai, K., & Katsika, A. (2020). Pitch accent and phrase boundaries: Kinematic evidence from Japanese. In *Proceedings of the 10th international conference on speech prosody 2020* (pp. 215–219).

Turk, A. E., & Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics*, *35*(4), 445–472.

Tyler, M. D., & Cutler, A. (2009). Cross-language differences in cue use for speech segmentation. *The Journal of the Acoustical Society of America*, *126*(1), 367–376.

Umeda, N. (1975). Vowel duration in American English. *The Journal of the Acoustical Society of America*, *58*(2), 434–445.

Vaissière, J. (1983). Language-independent prosodic features. In *Prosody: Models and measurements* (pp. 53–66). Springer.

Warner, N., Otake, T., & Arai, T. (2010). Intonational structure as a word-boundary cue in Tokyo Japanese. *Language and Speech*, *53*(1), 107–131.

Welby, P. (2007). The role of early fundamental frequency rises and elbows in French word segmentation. *Speech Communication*, *49*(1), 28–48.

White, L. S. (2002). *English speech timing: A domain and locus approach* (PhD dissertation). University of Edinburgh Edinburgh.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America*, *91*(3), 1707–1717.

Wilbur, R. B. (2009). Effects of varying rate of signing on ASL manual signs and nonmanual markers. *Language and speech*, *52*(2-3), 245–285.

Xu, Y., & Wang, M. (2009). Organizing syllables into groups—evidence from f0 and duration patterns in Mandarin. *Journal of phonetics*, *37*(4), 502–520.

Chapter 3

Article 2: Focus Kinematics

3.1 Introduction

Prosodic structure serves two important linguistic roles. One is grouping linguistic units into larger phrases by marking the boundaries of these constituents. The other is marking the relative saliency of constituents that are important for either rhythmic or conceptual reasons (e.g., to mark stressed and accented syllables respectively in stress languages). We refer to this function of prosody by the term *prominence*. In the discussion of prosodic typology, prominence pertains to how relative saliency is expressed both through word prosody (i.e., the realization of prominence relations within words) and phrasal prosody (i.e., the realization of prominence relations among words) (cf. Jun, 2006). While specific criteria for categorizing word prosody remain unsettled (see Beckman, 1986; van Der Hulst, 1999; Fox, 2002 for various proposals on types of word prosody and their definitions), widely acknowledged categories are—tone, stress, and pitch-accent (e.g., Beckman, 1986; Ladd, 1996; Cruttenden, 1997). On the other hand, the realization of phrase-level prominence has been proposed to occur in two ways: a) by marking the head of a prosodic unit, and b) by marking the edge of a prosodic unit (Hyman, 1978;

Beckman, 1986; Beckman & Edwards, 1990; Ladd, 1996; Venditti et al., 1996). Prominence realization via the heads of a prosodic unit, thus termed 'head-prominence' (Jun, 2006), involves manipulating prosodic features, such as pitch, duration, and amplitude to render a syllable or word prominent, represented as a phrasal pitch accent. The specific realization of the phrasal pitch accent depends on the prosodic features employed for word prosody in the language (Jun, 2006). Alternatively, prominence can be realized via marking the edges when the prominent word is positioned at specific locations within a prosodic unit (e.g., the beginning or the end), and a phrasal tone marks the unit's edges. Languages that employ this type of prominence-marking are termed 'edge-prominence languages' (Jun, 2006).

Work on the phonetic realization of these prosodic positions has shown that prominence marking by accent (either sentence stress or nuclear pitch accent) is acoustically manifested with expansion of duration, pitch, and amplitude (cf. Fletcher, 2010). Articulatorily, accented gestures are known to be produced with longer, larger, and faster movements (e.g., Beckman et al., 1992; Beckman & Edwards, 1994; Byrd & Saltzman, 2003; Cho, 2006).

According to limited reports on English and German, it is also suggested that focus derived from information structure may show a categorical effect depending on its type (acoustic studies: Gussenhoven, 1983; Baumann et al., 2006; see Breen et al., 2010; articulatory studies: Hermes et al., 2008; Roessig & Mücke, 2019; Mücke & Grice, 2014; Katsika et al., 2020). These studies examined different types of focus (e.g., broad focus, narrow focus, contrastive focus, unfocused) and found that phonetic dimensions, such as acoustic F0 and kinematic duration and displacement, increased across focus types. In combination, this work indicates that the phrase-level prominence may be encoded in a way that signals information beyond simple presence or absence of prominence.

However, evidence for results come mainly from head-prominence languages, such as

English, in which phrase-level prominence is usually marked by placing a pitch accent on the stressed syllable of the prominent word, while there might be typological differences and the effects of the above-mentioned issues might differ on a language-dependent basis. Moreover, while limited evidence suggests that focus-induced prominence in Korean is kinematically manifested similarly to head-prominence languages, i.e., via longer, larger and faster gestures (e.g., Shin et al., 2015), a full understanding of the effect and its source is lacking. For instance, the scope of the effect, its interaction with boundary-marking, and the effect of focus type on it have yet been discussed. Thus, the goal of this article is devoted to examining the effect of focus-induced prominence in Seoul Korean, answering the following questions. First, we test the kinematic effects of focus-induced prominence on the focused unit as well as the post-focal linguistic units, which undergo a phenomenon referred to as *dephrasing*. Second, the interaction between focus-marking and IP boundary marking is investigated, as Korean marks prominence via boundaries exhibiting thus an interesting relationship between the two prosodic functions. Finally, it is examined whether different types of focus exert different effects on articulation in Seoul Korean.

In the following sections, 3.1.1 discusses the reported phonetic correlates of prominence, and 3.1.2 the proposed prosodic typology focusing on head- vs. edge-prominence languages. Section 3.1.3 introduces Korean prosody as well as presents the questions and goals of the present article.

3.1.1 Phonetics of prominence

As mentioned above, linguistic units under prominence and syllables adjacent to the prosodic junctures are important loci in speech. Thus, extensive work has focused on examining the phonetic realization of these prosodic landmarks. For instance, segments under prominence (including accent or stress) are produced with longer duration, greater intensity, greater pitch movement, and unreduced vowel quality under prominence in the acoustic domain (Cho, 2011, Katsika & Tsai, 2021; cf. Lehiste, 1970; Beckman, 1986; see Fletcher, 2010, for an overview). In the articulatory domain, prominence is accompanied by increased respiratory effort (Cho, 2011; see Lehiste, 1970, for discussion). Also, gestures under prominence have found to be produced with longer movement duration, larger spatial displacement, and faster articulatory movement (Beckman & Edwards, 1994; Beckman et al., 1992; Cho, 2006; de Jong, 1991, 1995; Fowler, 1995).

However, it has been reported in the literature that all of these three kinematic dimensions—i.e., longer duration, larger displacement, and faster velocity—may not necessarily be used concurrently to differentiate between accented and unaccented gestures. For example, Beckman et al. (1992) found that accented syllables are produced with larger and longer jaw opening gestures compared to their unaccented counterparts, but there was no significant difference in velocity. Also, Cho (2006) finds the three kinematic patterns consistently for the English consonantal lip opening movement, but the lip closing movement show context-dependent results: when following an accented vowel, the movement is longer, larger, and faster, but when preceding an accented vowel, the movement is larger, but not faster or longer. While these variations may arise due to different granularity of the measurements taken, it may also come from language-specific modulation associated with the prominent gestures. It is important to note, however, that these results on the phonetics of prominence come mainly from head-prominence languages, such as English, in which phrase-level prominence is usually marked by placing a pitch accent on the stressed syllable of the prominent word (cf. Beckman & Pierrehumbert, 1986).

Recent work presents evidence that it is not simply the accentuation per se, but focus as derived from information structure that drives prominence-related modifications (acoustic properties: Gussenhoven, 1983; Baumann et al., 2006; see Breen et al., 2010; articulatory properties: Hermes et al., 2008; Roessig & Mücke, 2019; Mücke & Grice, 2014 for German; see also Katsika et al., 2020, 2023 for American English). Acoustically, it was found that speakers mark distinctively the focus range, such that narrow focus shows longer duration, greater intensity, and higher F0 than broad focus (Breen et al., 2010). More recent articulatory work looked at different types of focus (i.e., contrastive focus, narrow focus, broad focus, unfocused) and found that phonetic dimensions, such as acoustic F0 and kinematic duration and displacement, increased across focus types. While there might be typological and/or language-specific differences, findings demonstrate that phrase-level prominence is realized both by modifying multiple phonetic dimensions, but also by showing categories beyond the simple accented vs. unaccented distinction. Thus, it is hypothesized that phrasal prominence may be organized hierarchically, with levels of the hierarchy possibly representing focus types. If this is the case, it is unclear which focus types are encoded and in what order and whether this tendency emerges as a language-universal pattern.

3.1.2 Head-prominence vs. edge-prominence languages

Here, we adopt the typological distinctions proposed by Jun (2006, 2014), as introduced in this section. In the Autosegmental-Metrical (AM) framework, prosody is characterized in two perspectives: the prosodic structure of an utterance and the relationships of prominence within the structure (cf. Beckman 1996; Ladd, 1996; Shattuck-Hufnagel & Turk, 1996). Prosodic structure is organized hierarchically with higher- level prosodic units (e.g., Intonation Phrase, intermediate intonation phrase) consisting of one or more lower-level units (syllable, prosodic word). As Figure 3.1 illustrates, syllables constitute prosodic words (PWd); prosodic words constitute intermediate phrases (ip);

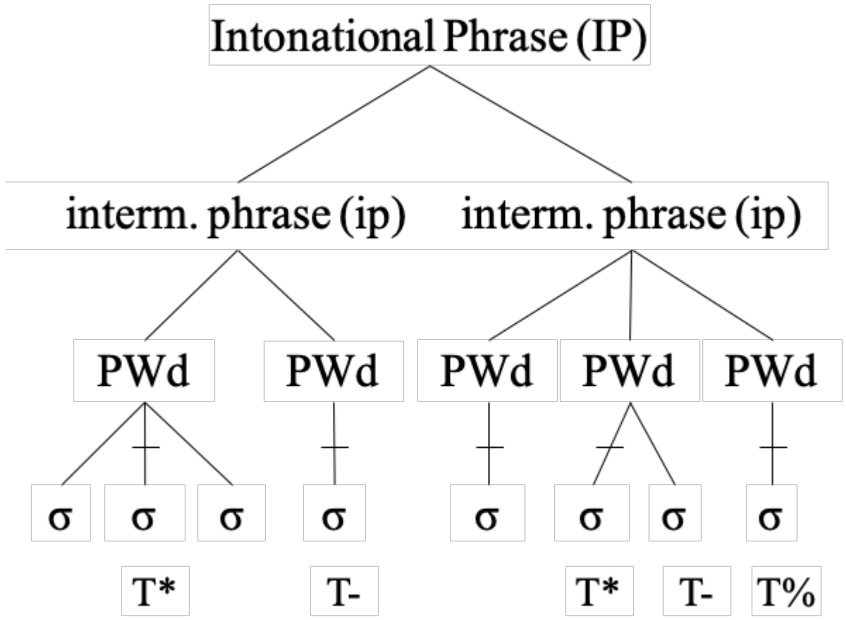

Figure 3.1: Prosodic structure, adapted from Beckman & Pierrehumbert (1986). PWd and σ refer to prosodic word and syllable respectively. Bars between the PWd and syllable tier indicate stressed syllables as in Keating and Shattuck-Hufnagel (2002). Pitch accents are represented with T*, phrase accents with T-, and boundary tones with T%.

and intermediate phrases constitute Intonational Phrases (IP). Prominence is marked at different levels of the structure. At the level of words, some syllables are more salient, or prominent, than others, and at the levels of phrases, some words are more prominent than others. In this structure, the realization of prosodic structure and prominence is achieved by manifesting prosodic modifications (e.g., pitch, duration, and/or amplitude) as well as segmental properties of consonants and vowels.

At the word level, languages are categorized as stress (e.g., English, Greek), pitch-accent (e.g., Japanese, Basque), or tone languages (e.g., Mandarin, Hausa). Note that languages may combine more word prosody systems (e.g., both stress and lexical pitch accent: Chickasaw, Swedish) or use none of them (e.g., Seoul Korean). At the phrasal level, two prominence systems are proposed based on data (Jun, 2006). Phrasal prominence can be realized by marking peaks or heads, via a local manipulation of prosodic features such as pitch, duration, and/or amplitude, making a syllable or word more prominent. Languages employing this type of prominence are called head-prominence languages. In the AM model, a phrasal head is represented as a pitch accent, notated with an asterisk (T*, where T stands for tone, e.g., L*, H*). The manipulation of these phrasal pitch accents is shaped by the prosodic features utilized by the word prosody of a given language (Jun, 2006). For example, in English where lexical stress is realized by both duration and amplitude of the syllable, the phrasal pitch accent also employs changes in these dimensions. Other language might use a different subset of the phonetic correlates for prominence, as the exact acoustic or articulatory correlates of phrasal pitch accents seem to be language-specific (Jun, 2006; but also, Cho, 2011, 2015). On the other hand, there are languages in which phrasal prominence is realized by marking the edges, i.e., by placing the prominent word either at the beginning or the end of a prosodic phrase, the level of which is language-dependent. These are called edge-prominence languages. In turn, a boundary tone marks the edges of the prominent prosodic unit. Some examples of tones marking prominence at different prosodic levels include the Prosodic Word boundary tone in Serbo-Croatian (%L), the Phonological Phrase boundary tone in Bininj Gun-wok (Lp), and the Accentual Phrase boundary tone in Korean (Ha) (Jun, 2006). For example, when a word receives contrastive focus in Korean, a prosodic boundary is inserted before or after the focused word, dephrasing the postfocal items (see Jun, 1996, 2003; Jun & Lee, 1998). Languages may use both heads and edges to mark phrasal prominence, being thus called head/edge-prominence languages (Jun, 2014).

3.1.3 Korean prosody and the current study

In this work, we investigate Seoul Korean with the goal to examine the articulatory correlates of focus-induced prominence in an edge-prominence language. Examining an edge-prominence language would be particularly informative in terms of understanding how the two functions of prosody, i.e., prominence and grouping, interact, since in such a language the two functions merge in focus-marking, where grouping, i.e., prosodic boundaries, is used to mark prominence. Moreover, Seoul Korean does not employ word prosody at all, i.e., it does not have lexical stress, lexical tone, or lexical pitch accent (Jun, 2005, 2006). Figure 3.2 illustrates the intonation model of Seoul Korean. Phrase-level prominence is marked by the means of Accentual Phrases (APs), which serves as the basic intonational unit and is marked by a particular pitch contour (Jun, 1998, 2005). AP's proposed underlying tonal pattern is THLH, where the realization of the initial tone (T) tends to depend on the laryngeal configuration of the AP-initial segment (Jun, 1993, 1998, 2005). Focus-induced prominence in Korean employs this AP level, as the focused word consistently starts an AP (or a higher phrase), and any following AP boundaries up to the end of the Intonational Phrase (IP) typically undergo elimination, or possibly attenuation, referred to as dephrasing (Jun, 1993) (Figure 3.3).

As such, APs serve multiple functions, being the basic intonational unit as well as marking phrase-level prominence in Seoul Korean. However, findings are scarce on the phonetic dimensions of these prosodic landmarks. Limited work on the correlates of prominence in Korean reports longer, larger, and slightly faster vocalic movements under focus in Korean (Shin et al., 2015; see Cho, 2022 for a review). However, the scope of the effect as well as the phenomenon of dephrasing that arises from it is unclear. We use Electromagnetic Articulography (EMA) to examine these issues in Korean (Experiment 1). Our second research question involves the interaction of prominence-marking with

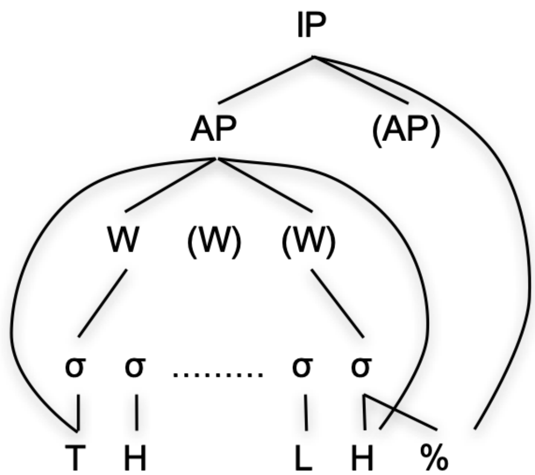

Figure 3.2: Intonation model of Seoul Korean, adapted from Jun (2000).

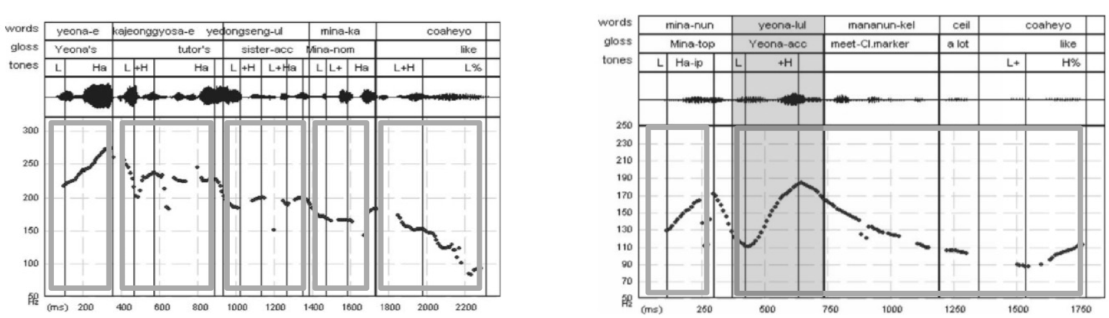

Figure 3.3: Examples of pitch tracking from Jun (2011). Example on the left shows an utterance with multiple APs marked by line boxes demonstrating no dephrasing. Example on the right shows dephasing after the focused word marked by the shaded box.

marking higher-level IP boundaries, as they are both demarcative in nature. In particular, we examine the interaction between focus marking (left-edge AP boundary) and IP marking (right-edge IP boundary), with a goal to broaden our understanding on the interplay between information structure and prosodic structure (Experiment 2). Finally, the effect of different types of focus on articulation is tested to shed light on the hierarchy and degrees of prominence in Korean (Experiment 3).

We also investigate whether the established interrelationships between kinematic parameters under the assumptions of a critically damped second-order linear gestural dynamical system (i.e., duration increases as stiffness decreases, and displacement increases proportionally to velocity (e.g., Byrd & Saltzman, 1998; Munhall et al., 1985)), show different profiles as a function of the gesture's position in the phrase (e.g., initial in focused AP vs. final in IP). Thus, understanding the relationship between kinematic parameters as a function of prosodic modulation is an additional derived goal of this article.

3.2 Experiment 1: Kinematics of focus-induced prominence and dephrasing

As a reminder, the main goal of Experiment 1 is to examine the kinematic manifestation of constriction gestures encoding different types of focus information. Specifically, test words vary with respect to focus status, being: 1) focused, i.e., initial in a focused AP; 2) unfocused, but not dephrased, i.e., initial in a pre-focal AP; or 3) unfocused and expected to be dephrased by virtue of following a focused AP. We predict gestures under focus to be longer, larger, and faster, based on previous articulatory research on phrasal prominence mainly in head-prominence languages (cf. Cho, 2006) and limited work on Korean (Shin et al., 2015, see Cho, 2022 for a review). As far as the stretch of speech affected by the focus effect as well as the kinematic manifestation of dephrasing, these are still unclear issues. Based on findings of the extended scope of prominence effect beyond the stressed syllable (head) in head-prominence languages, we expect the scope of the focus-induced prominence to span more than the initial syllable of the AP. Nonetheless, to which syllable the effect extends needs to be tested.

Regarding dephrasing, one hypothesis is that the tonal attenuation observed in dephrasing is accompanied by articulatory attenuation, i.e., shorter and smaller gestures as compared to their unfocused (but not dephrased) counterparts. Initial gestures of prosodic domains have been assumed to undergo 'domain-initial strengthening', i.e., stronger articulation (e.g., Fougeron & Keating 1997; Cho & Keating 2001, 2009; Fougeron 2001; Keating, Cho, Fougeron & Hsu 2003). If the unfocused AP-initial gestures are under the effect of domain-initial strengthening, they would show 'stronger' articulation compared to the dephrased (AP-medial) ones. However, if no significant difference is detected between the dephrased and unfocused gestures, it may suggest that that AP-initial gestures do not show domain-initial strengthening and that dephrasing is not accompanied by articulatory attenuation.

The basic hypothesis with respect to the interrelationships between the kinematic parameters is that duration will increase as stiffness decreases, and displacement will increase proportionally to velocity as shown in previous studies within Task Dynamics (Byrd & Saltzman, 1998; Munhall et al., 1985). However, it is unclear whether these patterns will differ systematical as a function of focus status.

3.2.1 Methods

Participants

Seven native speakers of Seoul Korean (5 female, 2 male) in their 20s (mean age = 24.5; median age = 23; age range = 21 and 29) participated in the present study. Data collection was interrupted due to COVID-19. Out of seven speakers, six participants were affiliated with the University of California Santa Barbara as graduate or exchange students or post-doctorate researchers at the time of the experiment, and one participant was a family member of a UCSB-affiliated researcher. The speakers had no self-reported speech, hearing, or vision problems and were naïve as to the purpose of the current study. They were rewarded with financial compensation for their participation.

Experimental procedure

Before the experimental session began, a short, 15-minute long, training session was held in order to familiarize the participants with the speech materials and the experimental procedure. Participants were instructed to imagine a context in which they were preparing a school play with friends to contextualize the speech materials (see section Experimental design and stimuli for more details). This set-up was designed to facilitate participants naturally produce the phrasing and place the focus contrast as intended (see the next section for details on the stimuli).

During the experimental session, kinematic data were collected using the AG501 3D electromagnetic articulograph (Carstens Medizinelektronik) 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 (Figure 3). The last five sensors were used as references.

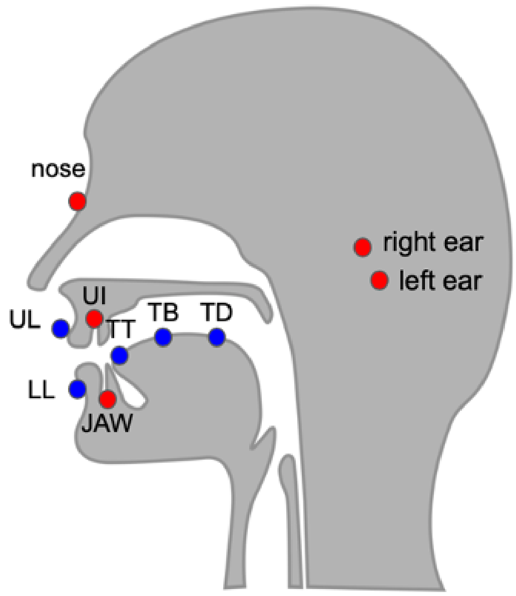

Figure 3.4: EMA sensor placement.

Simultaneous audio recordings were performed with the kinematic recordings using a Sennheiser shotgun microphone set at a sampling rate of 16 kHz, which was positioned one foot away from the participant's face. Speech materials were presented on a computer screen, placed roughly three feet away from the participant, using custom software, developed by Mark Tiede (Haskins Laboratories). The articulatory data acquired from the experiment was smoothed and corrected for head movement by using the reference sensors for each trajectory. Then, they were rotated to align the X- and Y- axis to the participants' occlusal plane. To help appropriate focus placement, a prompt sentence preceded each target sentence. Prompt sentences were presented one second before their corresponding target sentence, and were read silently by the participant. Target sentences were read aloud, which participants were instructed to read as if they were talking to a friend.

Experimental design and stimuli

Example stimuli are presented in Table 3.1. Participants were asked to imagine a situation where there are two uncles, both named Minam, one in a magic club and another in a secrecy club, and two uncles, both named Junseok, one in a magic club and another in a secrecy club. They were also asked to imagine that they are in the middle of deciding who would play the roles of 'the handsome guy from Nemang', the IPA transcription of its Korean equivalent being /nɛmaŋminam/, and 'the handsome guy from Nowon', transcribed as /nowonminam/. Test prosodic word /minami/ ('Minam (proper name)' + '-i (NOM)') was embedded in stimuli sentences that consisted of three Noun Phrases (NPs). Position of contrastive focus was varied, as prompted by mini dialogues, so as to yield the following focus types on the test word:

1) focused, when focus is on the test word (stimulus #1 in Table 3.1)

2) unfocused, when focus is on the next NP following the test word (stimulus #2 in Table 3.1)

3) dephrased, when focus is on the initial NP (stimulus #3 in Table 3.1).

To examine the effects of IP boundary (see Article 1), a set of test IPs were designed to match the focus placement of the stimuli but without the IP boundary in the middle. The presence or absence of the IP boundary in this case was not significant in the initial analyses, and thus, a random structure for frame sentence was added in the model.

Eight repetitions of each condition, randomized along with other stimuli examining other aspects of Korean prosody, were collected, except for one speaker, by whom five repetitions were recorded due to technical reasons. The acquired data were checked for their prosodic rendition, which, among other dimensions, confirmed tonal attenuation due to dephrasing. Figure 4 shows example pitch trackings of the stimuli by Focus status.

Focus	Example sentences (# = IP boundary)
(1) Focused	Prompt sentence: ‘It’s not uncle Junseok.’ Test sentence: [AP pimilpu] [AP **minami** gomopuga] [AP nɛmaŋminam]? # [sʌntækhangʌja]? ‘**Uncle Minam** of the secrecy club is the handsome guy from Nemang? Is it decided?’
(2) Unfocused	Prompt sentence: ‘It’s not the handsome guy from Nowon.’ Test sentence: [AP pimilpu] [AP minami gomopuga] [AP **nɛmaŋminam**]? # [sʌntækhangʌja]? ‘Uncle Minam of the secrecy club is **the handsome guy from Nemang**? Is it decided?’
(3) Dephrased	Prompt sentence: ‘It’s not the one in the magic club.’ Test sentence: [AP **pimilpu**] [AP minami gomopuga] [AP nɛmaŋminam]? # [sʌntækhangʌja]? ‘Uncle Minam of **the secrecy club** is the handsome guy from Nemang? Is it decided?’

Table 3.1: Example stimuli sentences by Focus status (Focused, Unfocused, Dephrased). Focused words are in bold and measured intervals are underlined.

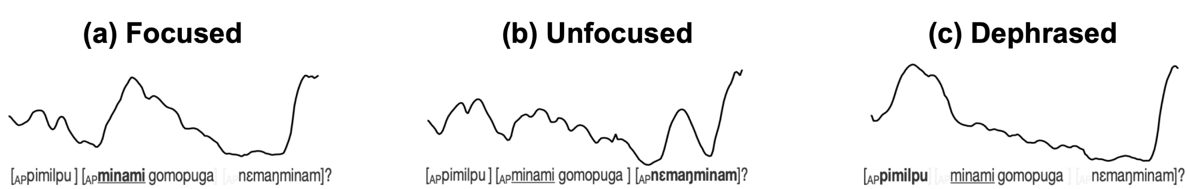

Figure 3.5: Examples of pitch tracking by Focus status.

Measurements

Consonant (C) gestures of the test prosodic word /minami/ were identified and labeled using semi-automatic custom software developed by Mark Tiede (Haskins Laboratories). The consonants are referred to here as C1, C2, and C3, from left to right. For C1 and C3, both of which are /m/, consonant (C) constrictors were detected on the lip aperture trajectory. For C2, which is /n/, tongue tip vertical displacement trajectory was used. The labeling procedure detected the following kinematic timepoints in each C gesture: onset, time of peak velocity, target, constriction maximum, release, and offset (Figure 3.5). These timepoints were identified on the basis of 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% 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 Figure 3.5).

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

1. Duration of gestural formation (in ms; a in Figure 3.5)
2. Gestural displacement to target (i.e., the spatial difference between onset and maximum constrictor position) (in mm; c in Figure 3.5)
3. Gestural peak velocity to target (in cm/sec; d in Figure 3.5)

Statistical analysis

The retrieved data were analyzed by linear mixed-effects analysis using *lme4* package (Bates et al., 2015) in R (R Core Team, 2022). The dependent variables were formation duration, displacement, and peak velocity. For each of the dependent variables, separate

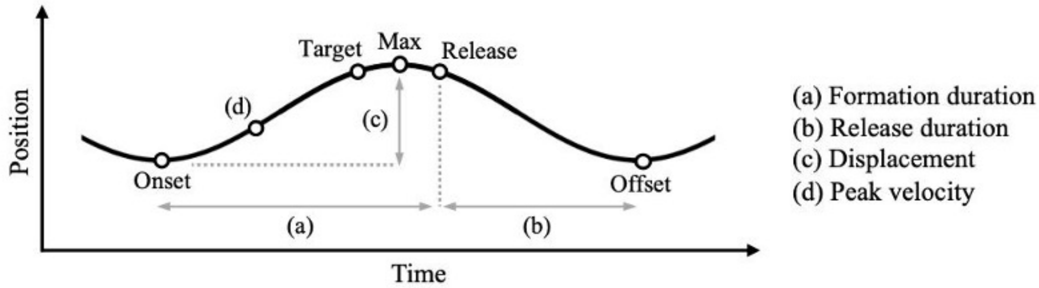

Figure 3.6: Schematized constriction gesture with kinematic measurements.

linear mixed-effects models were fitted for each segment (C1, C2, C3). The independent variable was Focus status (Focused, Unfocused, Dephrased). In addition, the following continuous predictors were fitted in each of the models in order to account for the expected covariance between them and the dependent variable as well as to see whether the gestures affected by prominence also show a distinct profile of these interrelationships (cf. Edwards et al., 1991; Byrd & Saltzman, 2003), applying the methodology in Katsika & Tsai (2021): (a) peak velocity normalized over displacement, referred to as *stiffness* henceforth for reasons of brevity, (b) peak velocity for the models of displacement, and (c) displacement for the models of peak velocity. Note that the ratio of peak velocity to displacement has been proposed as an empirical estimate of kinematic stiffness, which captures the observation that peak velocity varies with displacement and has been shown to increase as movement duration decreases in research that views gestures as critically damped second-order linear system (Munhall et al., 1985; Ostry & Munhall, 1985; used in e.g., Beckman et al., 1992; Hawkins, 1992; Roon, Gafos, Hoole, & Zeroual, 2007). The models included random intercept of Speaker and frame sentence. In case of significant interactions, pair-wise comparisons were assessed by the *emmeans* package (Lenth et al., 2020) with Holm adjustment.

3.2.2 Results

The expected relationship between duration and stiffness (estimated by the ratio of peak velocity over displacement) was observed in all C measures, such that formation duration decreased with increasing stiffness (C1: $F(1, 626) = 171.2$, $p < 0.001$; C2: $F(1, 623) = 364.7$, $p < 0.001$; C3: $F(1, 634) = 144.3$, $p < 0.001$) (3.7b). Displacement and peak velocity also demonstrated the expected relationship, such that displacement was found to increase with peak velocity in all C gestures (C1: $F(1, 635) = 2318.6$, $p < 0.001$; C2: $F(1, 610) = 2198.9$, $p < 0.001$; C3: $F(1, 615) = 1152.8$, $p < 0.001$) (3.8b, 3.9b).

Effects on duration

Figure 3.7a shows the results on duration in each consonant. Significant main effects of Focus status were detected in C1 ($F(2, 629) = 43.2$, $p < 0.001$) and C2 ($F(2, 620) = 20.5$, $p < 0.001$). Pair-wise comparison revealed that the formation duration of C1 was greater in Focused condition than in both Unfocused and Dephrased conditions ($\beta = 5.7$, SE = 0.8, $p < 0.001$; $\beta = 4.3$, SE = 0.8, $p < 0.001$, respectively). No pair-wise difference was found between Unfocused and Dephrased C1 gestures. Similar results were found for C2: Formation duration was greater when C2 was focused than either unfocused or dephrased ($\beta = 8.1$, SE = 0.9, $p < 0.001$; $\beta = 9.0$, SE = 0.9, $p < 0.001$, respectively), and no difference was found between the Unfocused and Dephrased conditions. C3 did not show a main effect of Focus status.

There were also significant interactions between Focus status and Stiffness in C1 ($F(2, 624) = 24.6$, $p < 0.001$) and C2 ($F(2, 620) = 10.2$, $p < 0.001$). As shown in Figure 3.7b, the absolute slope of the linear relationship between duration and stiffness was higher in Focused than the other two conditions (i.e., Unfocused, Dephrased). In C1, for each unit ((cm/sec)/cm) of stiffness decrease, the increase in duration was 1.5

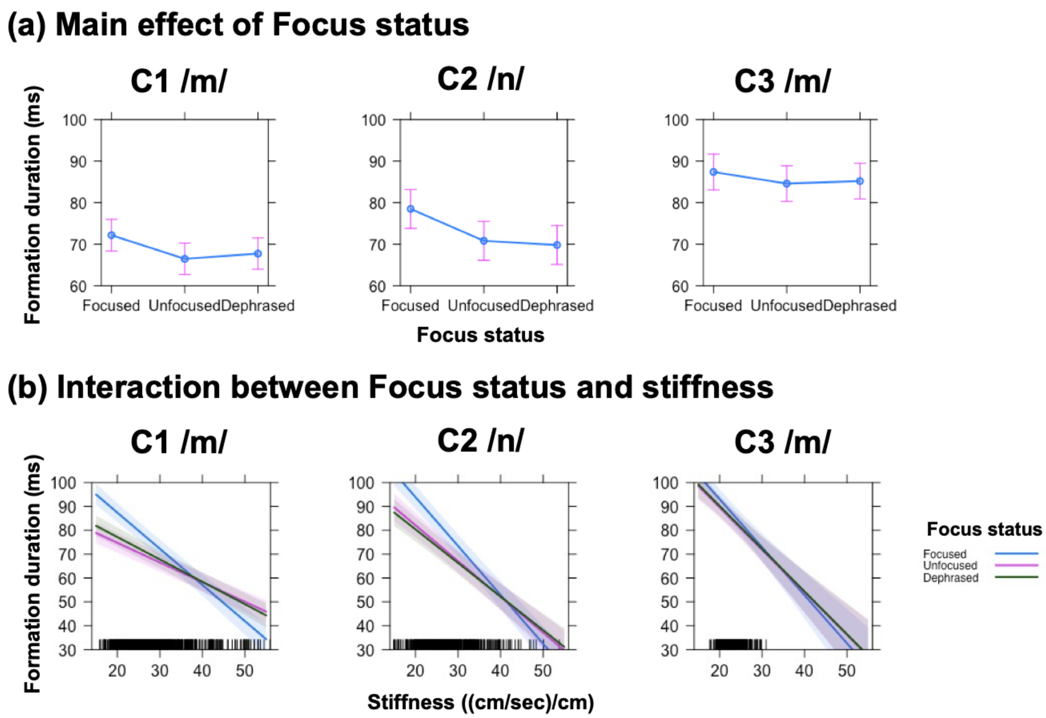

Figure 3.7: Results for duration (in ms) in (a) and its relationship with stiffness ((cm/sec)/cm) in (b) as a function of focus status.

ms in the Focused condition but just 0.8 and 0.9 ms in the Unfocused and Dephrased conditions, respectively. In C2, for the same decrease in stiffness, the duration increase was 2.1 ms under focus but 1.5 and 1.4 ms in the Unfocused and Dephrased conditions, respectively. No further effect of Focus status was found in C3.

In sum, gestures under focus are longer than their unfocused counterparts. The focus effect extends to the second syllable of the word. No difference between the two non-focused conditions, whether it is unfocused (but not dephrased) or dephrased (due to it being post-focal). An interaction between Focus status and stiffness was found, such that gestures affected by focus (i.e., C1, C2) showed distinct stiffness profiles.

Effects on displacement

Displacement showed main effects of Focus status in C1 ($F(2, 629) = 3.8$, $p < 0.05$) and C2 ($F(2, 620) = 3.8$, $p < 0.05$) (Figure 3.8). Pair-wise comparison indicated that displacement of C1 was greater in Focused condition than in both Unfocused and Dephrased conditions ($\beta = 0.3$, SE = 0.03, $p < 0.001$; $\beta = 0.2$, SE = 0.03, $p < 0.001$, respectively). Between Unfocused and Dephrased conditions, the displacement was smaller when unfocused than dephrased (β = -0.06, SE = 0.03, $p < 0.05$). Also in C2, displacement was greater in Focused condition than in both Unfocused and Dephrased conditions ($\beta = 0.2$, SE = 0.04, $p < 0.001$; $\beta = 0.2$, SE = 0.04, $p < 0.001$, respectively), while Unfocused and Dephrased were not different from each other. In C3, while the main effect of Focus status was not significant, displacement showed a significant interaction between Focus status and peak velocity ($F2, 631) = 8.7$, $p < 0.001$). Post-hoc analysis of the interaction revealed that the absolute slope of the displacement-peak velocity linear relationship was small but significantly higher in Focused than the other two conditions (i.e., Unfocused, Dephrased). Specifically, for each unit (cm/sec) of decrease in peak velocity, the increase in displacement was 0.4 mm in the Focused condition but 0.3 mm in the Unfocused and Dephrased conditions. The difference between Unfocused and Dephrased was not significant.

In sum, prominent gestures are larger in displacement than their unfocused or dephrased counterparts. The scope of the effect spans over three consonants, indicating that the effect extends to the onset of the (resyllabified) third syllable. The onset of the first syllable showed a significant difference between the two unfocused conditions, such that the displacement was smaller in Unfocused than in Dephrased conditions. This result may suggest a spillover effect from the previous focused linguistic unit in the dephrased condition, but further investigation is needed.

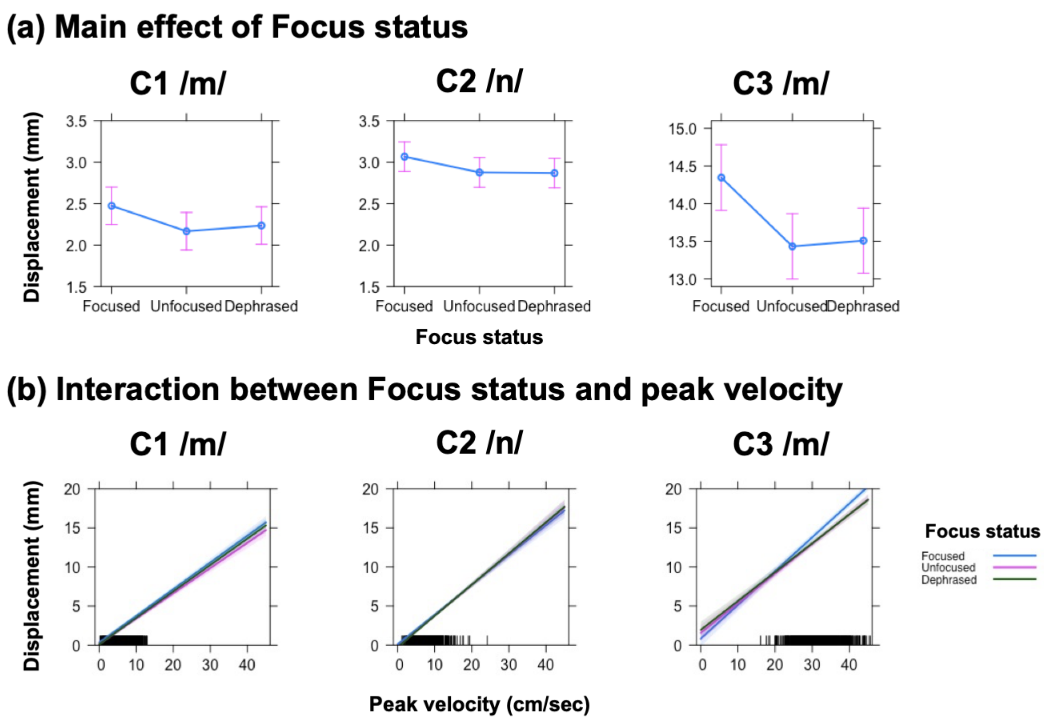

Figure 3.8: Results for displacement (in mm) in (a) and its relationship with peak velocity (cm/sec) in (b) as a function of focus status.

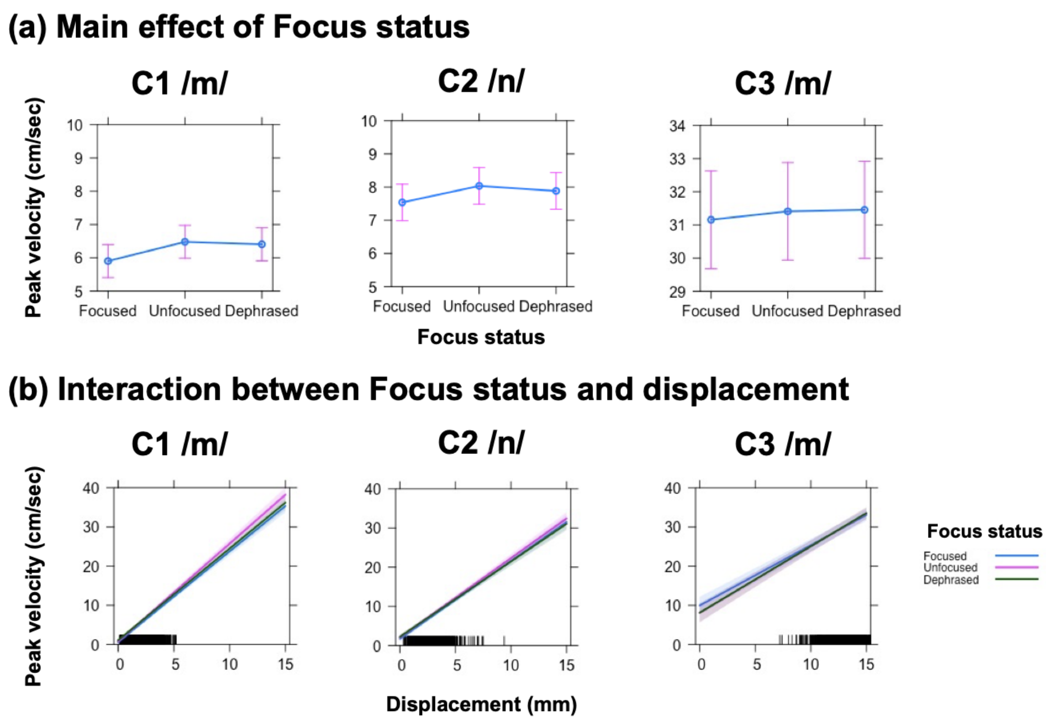

Figure 3.9: Results for peak velocity (in cm/sec) in (a) and its relationship with displacement (mm) in (b) as a function of focus status.

Effects on peak velocity

Interestingly, no main effect nor any interaction involving Focus status was found for peak velocity (Figure 3.9).

3.2.3 Interim discussion

The goal of Experiment 1 was to examine (1) the effect of focus-induced prominence on the articulation, (2) the effect of dephrasing (thus AP-medial) on articulation as compared to an unfocused (but AP-initial) counterpart, and (3) the interrelationship between kinematic parameters across the prosodic word of Seoul Korean.

Results indicate that the kinematic profile of phrasal prominence in Korean involves larger and longer, but not necessarily faster, gestures. The scope of the effect expands beyond the boundary-adjacent gesture. Specifically, the effect of focus reached the (re-syllabified) onset of the third syllable. While longer and larger movements are similar to the effects found in other stress languages (e.g., English), no effect of focus in peak velocity diverges from most findings of previous studies. We attribute this difference to the fact that the analyses used in the current study takes into account interdependences between kinematic parameters (i.e., duration-to-stiffness and peak velocity-to-displacement). Considering this, the lack of an effect on the peak velocity measure in our data may indicate that its interrelationship with displacement may be sufficient to account for any variations in velocity. Nonetheless, readers are reminded that displacement showed both significant effects of focus and correlated with peak velocity, which means that the larger gestures in focused units were also faster.

Furthermore, results suggest that dephrasing consists more of a tonal attenuation than articulatory attenuation. By design, the Unfocused condition was pre-focal and was constructed to be AP-initial, whereas the Dephrased condition was constructed to be post-focal and thus dephrased (AP-medial). If dephrasing, as deletion or attenuation of an AP boundary, accompanied articulatory attenuation, we might have expected to see shorter and smaller gestures in the Dephrased condition as compared to their Unfocused counterparts. However, the two non-focused conditions were largely similar to each other. This would, then, indicate that there is no articulatory evidence of domain-initial strengthening in AP-initial gestures (cf. Keating et al., 2004). If anything, the displacement of the word-initial C gesture showed a larger movement when it was dephrased as opposed to when it was unfocused (but not dephrased). This pattern may be due to a spill-over effect from the immediately preceding focused unit.

Lastly, results showed expected interrelationship between parameters—duration increases with decreased stiffness and displacement increases with peak velocity, in line with findings of head-prominence languages (Edwards et al., 1991; Byrd & Saltzman, 2003). However, it was found that this interrelationship is further modulated by prominence, such that the slope of the correlation between duration and stiffness was steeper when the gesture was under focus as opposed to when the gesture was unfocused or dephrased. The relationship between displacement and peak velocity, however, was not further affected by focus-induced prominence.

3.3 Experiment 2: Focus-induced prominence and its interaction with IP boundary

Extending the investigation from Experiment 1, Experiment 2 aims to examine the interaction between focus marking (left-edge AP boundary) and IP marking (right-edge IP boundary), with a goal to broaden our understanding on the interplay between information structure and prosodic structure. The prediction on the effect of focus would be that the gestures under focus are longer and larger, but not faster, based on previous research on phrasal prominence (Shin et al., 2015) as well as the results in Experiment 1. However, whether this prominence-related effects are affected by an upcoming IP boundary, and if yes, how, is unknown. We also extend the investigation of the interrelationships between kinematic parameters by seeing whether these differ word-initially (where focus effects are expected) from word-finally (where IP effects are expected, as also shown in Article 1). The current experiment also examines whether the pattern or the degree of dephrasing differ with distance of focus from the right edge of the IP. We expect those to decrease as the distance between focus and IP boundary increases.

3.3.1 Methods

Participants and experimental procedure

The participants and experimental procedure were the same as in Experiment 1 (see Section 3.2.1 for details).

Experimental design and stimuli

The same stimuli as in Experiment 1 were used for Experiment 2. However, the effect of focus was examined across the test interval /nɛ.maŋ.mi.nam/ for Experiment 2. The prosodic word /nɛ.maŋ.mi.nam/ is a compound noun consisting of two words /nɛ.maŋ/ ('Nemang (name of a place)') and /mi.nam/ ('handsome guy') forming one prosodic word. This target word /nɛ.maŋ.mi.nam/ was framed to receive contrastive focus in a set of test stimuli comprising an Intonational Phase (IP). Focused test words in those stimuli were compared to two sets of control stimuli that included the test word in unfocused condition; in one set, focus was placed on the preceding content word, proximate to the test word, and in the other, focus was on the distant content word, initial of the IP. We refer to these conditions as Proximate and Distant, respectively. Having these two different unfocused conditions allows examining the degree of dephrasing. To examine the interaction between focus marking (left-edge) and IP-boundary (right-edge), test stimuli were further manipulated so that test words were either IP-final or IP-medial. The combination of Focus distance (Focused, Proximate, Distant) and Within-IP Position (IP-final, IP-medial) yielded six conditions in total, which occurred in an experimental block along with stimuli for other experiments.

Same as Experiment 1, each block was randomized and repeated eight times. For one speaker, five repetitions were collected due to interruption of the experimental session for technical reasons. In total, 318 tokens were included in the analyses reported here.

The acquired data were checked for their prosodic rendition, i.e., focus placement and prosodic boundaries, using K-ToBI.

Measurements

Same measurement protocol was used as Experiment 1. Consonant (C) gestures in the interval /nɛ.maŋ.mi.nam/ were analyzed, except the coda of the second syllable. Coda /ŋ/ was excluded from the analysis because of its degree of blending with the neighboring vowels. We will refer to the measured consonants as C1, C2, C3, C4, and C5. All of these were onsets, except for C5 which is the coda of the fourth syllable.

Statistical analysis

The retrieved data were analyzed in the same way as in Experiment 1. To test the effect of Focus distance and its interaction with Within-IP Position, each C gesture was examined for the dependent variables of formation duration, displacement, and peak velocity. Fixed effects of Focus distance (Focused, Proximate, Distant) and Within-IP Position (IP-final, IP-medial) were included. Same with the models in Experiment 1, the following continuous predictors were fitted in each of the models in order to account for the expected covariance between them and the dependent variable (cf. Edwards et al., 1991; Byrd & Saltzman, 2003): (a) stiffness, as calculated by peak velocity over displacement, for the models of formation duration, (b) peak velocity for the models of displacement, and (c) displacement for the models of peak velocity. Random effect of speaker was added.

3.3.2 Results

In the following sections, the effects of Focus distance and Within-IP Position are presented by the parameters, i.e., formation duration, displacement, and peak velocity. Then, the interrelationships between the kinematic parameters and the effects of prosodic structure on this relationship will be discussed at the end.

Effects on duration

A significant main effect of Focus distance as well as interaction between Focus distance and stiffness were found for C1 ($F(2, 523) = 11.5$, $p < 0.001$; $F(2, 529) = 144.3$, $p < 0.001$, respectively) and C2 ($F(2, 526) = 6.7$, $p < 0.01$; $F(2, 526) = 5.4$, $p <0.001$, respectively). For the main effect on C1, pair-wise comparison indicated that the C1 formation duration was greater in focused condition compared to both unfocused conditions, regardless of whether focus was distant or proximate ($\beta = 15.5$, SE = 1.4, $p < 0.001$; $\beta = 16.5$, SE = 1.1, $p < 0.001$, respectively). Between the two unfocused conditions, formation duration was not significantly different. In terms of the interaction between Focus and stiffness, the slope of the linear relationship between duration and stiffness was higher in Focused than Proximate and Distant conditions. For one unit ((cm/sec)/cm) of stiffness decrease, the increase in duration was 3.3 ms in the IP-final condition, 2.9 in the Distant condition and 2.5 in the Proximate condition (Figure 3.13).

For C2, pair-wise comparison revealed that C2 duration was longer in focused condition compared to unfocused conditions, both when focus was distant or proximate ($\beta = 1.8$, SE = 0.7, $p < 0.05$; $\beta = 1.8$, SE = 0.4, $p < 0.01$, respectively). The two unfocused conditions were not significantly different from each other. The interaction between Focus distance and stiffness in C2 is reflected in the higher slope in Focused than the other two conditions (i.e., Proximate, Distant). For one unit ((cm/sec)/cm) of stiffness decrease, the increase in duration was 3.2 ms in the IP-final condition but just 2.4 and 2.5 ms in the Proximate and Distant condition, respectively.

In terms of Within-IP Position, a significant main effect was detected in C1 ($F(1, 523) = 6.3$, $p < 0.05$) with the formation duration of C1 being shorter in IP-final position than in IP-medial position (β = -4.3, SE = 1.1, $p < 0.001$). This IP-related shortening has been reported in the literature as *preboundary shortening* (Krivocapić, 2014; Katsika, 2016; Kim et al., 2017). Also, a main effect of Within-IP Position ($F(1, 531) = 34.4$, $p < 0.001$) with IP-final gestures being longer than IP-medial gestures (β = 23.2, SE = 2.4, $p < 0.001$) as well as an interaction between Within-IP Position and stiffness was detected for C5, which is the IP-final gesture ($F(1, 531) = 21.5$, $p < 0.001$, respectively). This interaction shows different slopes of the relationship between duration and stiffness for IP-final and IP-medial gestures, 7.5 and 2.9, respectively, and suggests that stiffness profiles are distinct for IP-final and IP-medial gestures (see Article 1, Section 4.1 for related discussion). No effect of Focus distance or Within-IP Position was found for C3 and C4.

In sum, the duration of consonantal constriction gestures is longer under focus compared to those when unfocused. The effect reaches the onset of the second syllable. When dephrased, durations are not affected by whether focus is further away or proximate. IP boundary marking affects the duration of the final coda consonant, i.e., C5, such that the presence of the following IP boundary lengthens its duration.

Effects on displacement

Displacement showed a main effect of Focus distance in C1 ($F(2, 523) = 3.1$, $p < 0.05$), C2, although marginally ($F(2, 527) = 2.9$, $0.05 < p < 0.06$), and C5 ($F(2, 526) = 6.9$, $p < 0.01$). Pair-wise comparison found that C1 in focused condition was significantly larger than in the Proximate condition (β = 0.3, SE = 0.1, $p < 0.01$) and marginally

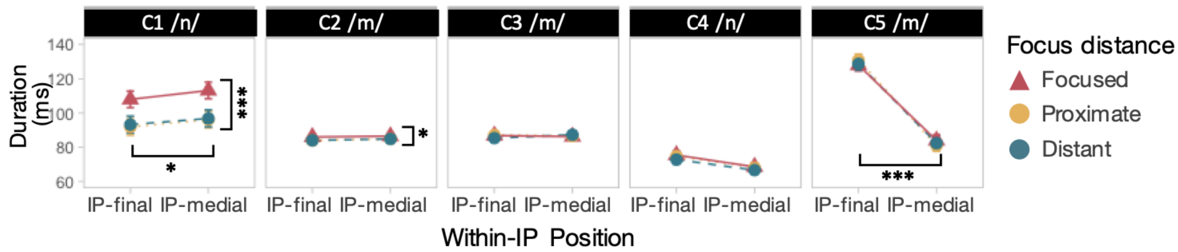

Figure 3.10: Results on duration (in ms) by Focus distance and Within-IP Position.

larger than in the Distant condition (β = 0.2, SE = 0.1, 0.05 < p < 0.06). Similarly, in C2, displacement was larger when the test word was focused than when focus was either distant or proximate (β = 0.9, SE = 0.1, p < 0.001; β = 1.1, SE = 0.1, p < 0.001, respectively). In both cases, the two unfocused conditions were not significantly different from each other. Although C5 showed a significant main effect of Focus distance, pair-wise comparison revealed that the three conditions were not significantly different from each other.

Interestingly, a marginally significant interaction between Focus distance and Within-IP Position was found in C1 displacement ($F(2, 523) = 2.6$, 0.05 < p <0.07). Pair-wise comparisons revealed that the displacement of C1 was smaller in IP-final position than in IP-medial position when focus was proximate (β = -0.3, SE = 0.1, p < 0.05) or distant (β = -0.4, SE = 0.2, 0.05 < p < 0.09 (marginal effect)). However, this effect was not found when C1 was focused (p > 0.09). Finally, a main effect of Within-IP Position was detected on C5 ($F(1, 527) = 42.9$, p < 0.001), with IP-final gestures being larger than IP-medial ones.

In short, gestures are larger when focused compared to unfocused (i.e., focus is distant or proximate). The effect extends to C2, which is in the second syllable of the prosodic word. No difference is found between the two unfocused conditions. An interaction between focus and IP boundary was detected such that in unfocused words, displacement

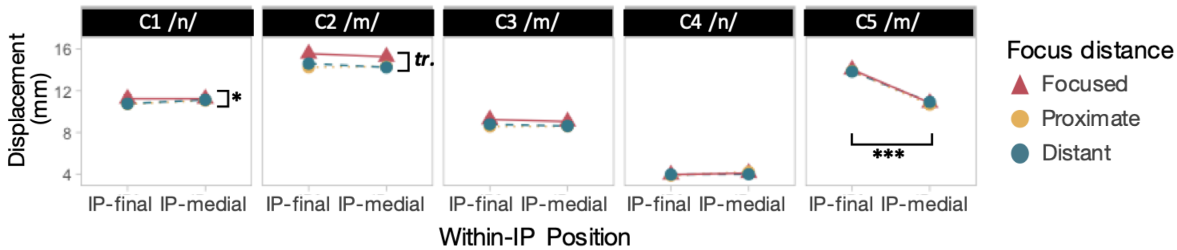

Figure 3.11: Results on displacement (in mm) by Focus distance and Within-IP Position.

was smaller at IP-final position as opposed to IP-medial position.

Effects on peak velocity

With respect to peak velocity, C1 and C2 did not show any effect of Focus distance. In C5, both the main effect of Within-IP Position and its interaction with displacement were significant ($F(1, 531) = 4.0$, $p < 0.05$; $F(1, 529) = 38.0$, $p < 0.001$, respectively). Pair-wise comparison of the main effect found that the peak velocity of C5 was greater IP-medially than IP-finally (β = -4.5, SE = 0.3, $p < 0.001$). As for the interaction, the slope of the peak velocity to displacement regression is greater in the IP-medial condition (2.0) than in IP-final (1.5) condition, meaning that given one unit (mm) of increase in displacement, increase in C5's peak velocity is 1.5 cm/sec IP-finally as opposed to 2.0 cm/sec IP-medially (Figure 3.13).

In sum, peak velocity did not show any effect of Focus distance. Instead, the final coda consonant of the word, i.e., C5, was slower when it was preceding an IP boundary than not.

Relationships between kinematic parameters and interaction with focus

Figure 3.13 demonstrates the relationship between formation duration and stiffness by Focus distance in each consonant position. As was shown in Experiment 1, expected

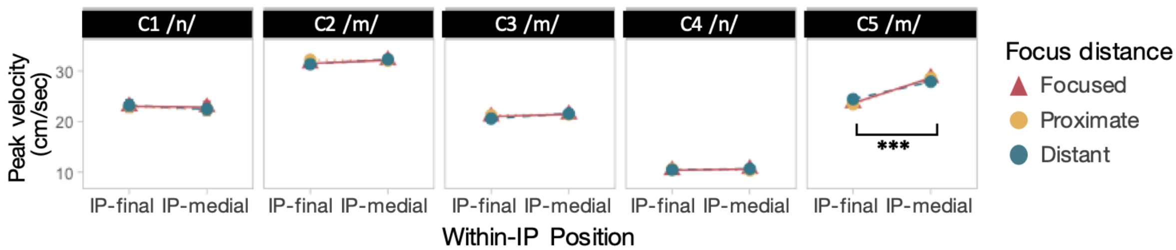

Figure 3.12: Results on peak velocity in cm/sec by Focus distance and Within-IP Position.

covariance between formation duration and stiffness was observed, such that duration decreased as stiffness increased (cf. Edwards et al., 1991; Byrd & Saltzman, 2003). However, this relationship further interacted with Focus distance. Specifically, the duration and stiffness relationship showed significant interactions in C1 and C2 (see Section 3.2.1 for statistics). As can be seen in Figure 3.13, the slopes were steeper when the gestures were under focus compared to unfocused. This indicate that the intrinsic relationship between duration and stiffness is further modulated by focus-induced prominence. In addition to the different stiffness profile for the focused conditions, note that C5 has the highest slope value, even compared to the focused gestures in C1 and C2. C5, which is the gesture that is adjacent to the IP boundary, seems to be displaying yet another profile of stiffness, presumably coming from the IP boundary marking (see Article 1, Discussion section for related discussion). On the other hand, the linear relationship between displacement and peak velocity was not affected by the presence or absence of focus, showing similar scope of the interrelationship across the entire word.

3.3.3 Interim discussion

The goal of the Experiment 2 was to investigate whether and how focus-marking (left-edge AP boundary) interacts with IP marking (right-edge IP boundary) in Seoul Korean. As with the previous experiment, we find longer and larger, but not independently faster,

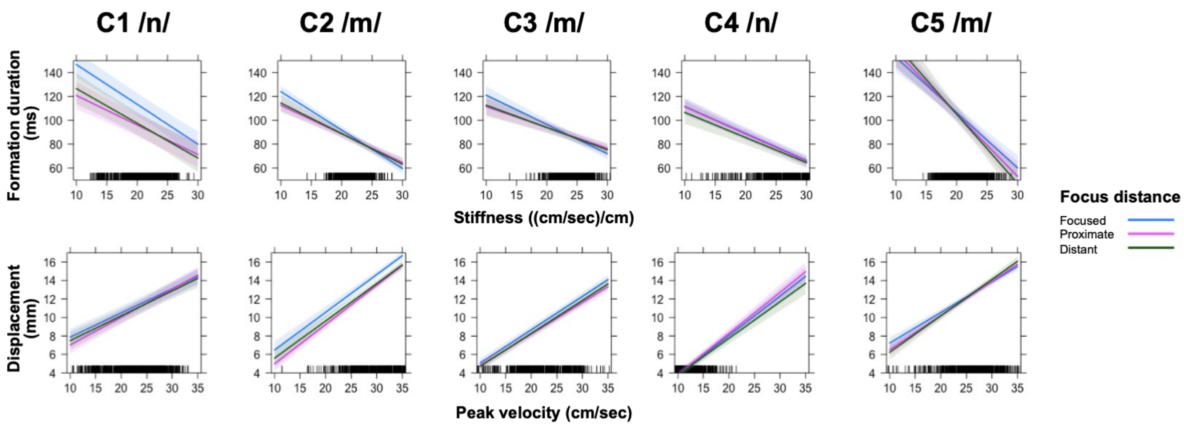

Figure 3.13: Relationships between kinematic parameters: (top) duration (in ms) by stiffness (in (cm/sec)/cm); (bottom) displacement (in mm) by peak velocity (in cm/sec).

gestures under focus-induced prominence. The effect extends to the onset of the second syllable. In general, the effect of focus-marking extends to the second syllable, while the effect of IP-boundary-marking is retained in the last syllable. However, we indeed find an interaction between focus-marking and boundary-marking, indicating an intricate relationship between phrase-level prominence-marking and IP boundary-marking. The smaller displacement found in IP-final position (as opposed to IP-medial position) is found only when the test word did not receive focus. This may suggest that this effect of IP boundary is constrained by prominence-marking, such that it emerges only when there is absence of prominence.

Another goal of the present study was to examine the phenomenon of dephrasing that arises from the effect of focus-induced prominence. In particular, we tested whether the patterns of dephrasing differ with distance of focus from the right edge of the IP. Contrary to predictions, our results suggest that the two unfocused conditions, whether the focused linguistic unit being proximate or distant, were not systematically different from each other.

Lastly, we also aimed to extend the investigation of the interrelationship between kinematic parameters in more depth, looking at this relationship as a function of prosodic position. An interesting pattern was found in terms of the relationship between duration and stiffness, as measured by the empirical estimate of normalized peak velocity over displacement. While all consonants show the same direction of the relationship, i.e., increase in duration as stiffness decreases, the presence of focus further modulates this interrelationship. Gestures under focus have a steeper slope of correlation compared to the unfocused conditions, again regardless of the distance of the focused linguistic unit. Moreover, IP boundary also exerts a distinct modulation on this relationship, such that gestures at IP-final boundary show the steepest slopes. Thus, it is suggested that the stiffness profiles are distinctively modulated by the prosodic functions of prominence- and IP boundary-marking.

3.4 Experiment 3: Focus types

The goal of Experiment 3 is to examine whether different types of focus are differentiated in their effect on articulation. The question is whether prosodic structure encodes presence or absence of focus or whether different types of focus exert more fine-grained effect of focus on articulation. A recent study on the role of focus in American English reports that, in addition to the presence of accent, focus structure further encodes the amount of accentual lengthening, presenting four degrees of phrasal prominence: contrastive focus > narrow focus > broad focus > absence of focus/accent (Katsika et al., 2019, 2023; also see Roessig & Mücke, 2019; Mücke & Grice, 2014 for similar results reported on German). Here, we expand this investigation with a language with no heads and no phrasal pitch accents.

3.4.1 Methods

Participants

Six native speakers of Seoul Korean (1 female, 4 male) in their 20s to 30s (mean age = 30.8; median age = 31; age range = 29 and 33) participated in the present study. All participants were affiliated with the University of California, Santa Barbara as graduate students or post-doctorate researchers at the time of the experiment. Out of the six participants, two speakers have also participated in Experiment 1. The speakers had no self-reported speech, hearing, or vision problems. They were naïve as to the purpose of the current study and received financial compensation for their participation.

Experimental procedure

Kinematic data were collected using the AG501 3D electromagnetic articulograph (Carstens Medizinelektronik) 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 were used as references. Simultaneous audio recordings were performed with the kinematic recordings using a Sennheiser shotgun microphone set at a sampling rate of 16 kHz, which was positioned one foot away from the participant's face. Speech materials were presented on a computer screen, placed roughly three feet away from the participant, using custom software, developed by Mark Tiede (Haskins Laboratories). The articulatory data acquired from the experiment were smoothed and corrected for head movement by using the reference sensors for each trajectory. Then, they were rotated to align the X- and Y- axis to the participants' occlusal plane. To help appropriate focus placement, a prompt sentence preceded each target sentence. Prompt sentences

List of target words		
/mapu/ ('horseman')	/napi/ ('butterfly')	/katɕi/ ('eggplant')
/madiɨɾidɨ/ ('Madrid')	/namunɨlpo/ ('sloth')	/katɕitɕiki/ ('pruning')

Table 3.2: List of target words.

were presented one second before their corresponding target sentence, and were read silently by the participant. Target sentences were read aloud.

Experimental design and stimuli

The test words used consisted of either two or four syllables with /m, n/ or /k/ as the onset consonant (Table 3.2). To test the effect of different focus types on articulation, the test words were embedded in frame sentences that yielded the following five focus types: contrastive (corrective) focus (CF); narrow focus (NF); broad focus (BF); unfocused with a contrastive focus on a different word (UC); and unfocused with a narrow focus on a different word (UN). Example sentences are presented in Table 3.3.

Eight repetitions of each condition, randomized, were collected. The acquired data were checked for their prosodic rendition, which, among other dimensions, confirmed tonal attenuation due to dephrasing.

Measurements

Same measurement protocol as Experiment 1 was used. Consonant (C) gesture of the initial syllable were analyzed. For /m/, /n/ and /k/, consonant constrictors were detected on the lip aperture, tongue tip vertical displacement, and tongue dorsum vertical displacement trajectories, respectively.

Focus	***Example sentences***
CF	Prompt sentence: 'Did Minam visit the farmer?' Test sentence: [minamiga **mapu**lul paŋmunhesʌ] Minam-NOM **horseman**-ACC visit-PAST 'Minam visited **the horseman**.'
NF	Prompt sentence: 'Who did Minam visit?' Test sentence: [minamiga **mapu**lul paŋmunhesʌ] Minam-NOM **horseman**-ACC visit-PAST 'Minam visited **the horseman**.'
BF	Prompt sentence: 'What happened?' Test sentence: [**minamiga mapulul paŋmunhesʌ**] **Minam-NOM horseman-ACC visit-PAST** '**Minam visited the horseman**.'
UC	Prompt sentence: 'Did Junseok visit the horseman?' Test sentence: [**minam**iga mapulul paŋmunhesʌ] **Minam**-NOM horseman-ACC visit-PAST '**Minam** visited the horseman.'
UN	Prompt sentence: 'Who visited the horseman?' Test sentence: [**minam**iga mapulul paŋmunhesʌ] **Minam**-NOM horseman-ACC visit-PAST '**Minam** visited the horseman.'

Table 3.3: Example stimuli sentences by Focus Type (CF, NF, BF, UC, and UN) with the test word /mapu/ ('horseman'). Focused words are in bold and measured intervals are underlined.

Statistical analysis

The retrieved data were analyzed in the same way as in Experiment 1. To test the effect of Focus Type, C gesture was examined for the dependent variables of formation duration, displacement, and peak velocity. Fixed effects of Focus Type (CF, NF, BF, UC, UN) and random effect of speaker as well as trajectory was added. Again for each model, the following continuous predictors were fitted in each of the models in order to account for the expected covariance between them and the dependent variable (cf. Edwards et al., 1991; Byrd & Saltzman, 2003): (a) peak velocity normalized over displacement, referred to as *stiffness*, for the models of formation duration, (b) peak velocity for the models of displacement, and (c) displacement for the models of peak velocity.

3.4.2 Results

A Main effect of Focus Type was significant in all three measured dimensions, i.e., formation duration ($F(4, 1461) = 91.4$, $p < 0.001$), displacement ($F(4, 1460) = 10.5$, $p < 0.001$), and peak velocity ($F(4, 1461) = 6.9$, $p < 0.001$). Results are plotted in Figure 3.14 and pair-wise comparisons are listed in Table 3. As demonstrated in the table, pair-wise comparisons gestures were longer and larger with the presence of focus, but focus type further modulating the amount of the effect. Results indicate that phonetic dimensions differentiated primarily among focus types and not simply between presence (in all focused conditions) and absence (in all unfocused conditions) of focus. However, phonetic dimensions differed in the number of focus types, and thus degrees of prominence they distinguished. Figure 3.15 schematizes these categories by parameters. In terms of formation duration, four degrees were distinguished: CF > NF > BF > UN, UC (Figure 3.15a). Displacement distinguished two degrees: CF, NF > BF, UN, UC (Figure 3.15b). Note that despite having a main effect, pair-wise comparison of peak velocity did not

		CF	NF	BF	UN
NF	Duration	β=4, SE=1.5, $p < 0.05$			
	Displacement	β=-0.1, SE=0.1, *n.s.*			
	Peak velocity	β=0.1, SE=0.2, *n.s.*			
BF	Duration	β=18, SE=1.5, $p < 0.001$	β=16, SE=1.5, $p < 0.001$		
	Displacement	β=0.4, SE=0.1, $p < 0.05$	β=0.5, SE=0.1, $p < 0.001$		
	Peak velocity	β=0.3, SE=0.2, *n.s.*	β=-0.4, SE=0.2, *n.s.*		
UN	Duration	β=24, SE=1.5, $p < 0.001$	β=20, SE=1.5, $p < 0.001$	β=5, SE=1.5, $p < 0.01$	
	Displacement	β=0.5, SE=0.1, $p < 0.001$	β=0.6, SE=0.1, $p < 0.001$	β=0.1, SE=0.1, *n.s.*	
	Peak velocity	β=-0.2, SE=0.2, *n.s.*	β=-0.3, SE=0.2, *n.s.*	β=0.1, SE=0.2, *n.s.*	
UC	Duration	β=25, SE=1.5, $p < 0.001$	β=21, SE=1.5, $p < 0.001$	β=7, SE=1.5, $p < 0.001$	β=2, SE=1.5, *n.s.*
	Displacement	β=0.5, SE=0.1, $p < 0.001$	β=0.6, SE=0.1, $p < 0.001$	β=0.1, SE=0.1, *n.s.*	β=0, SE=0.1, *n.s.*
	Peak velocity	β=-0.2, SE=0.2, *n.s.*	β=-0.3, SE=0.2, *n.s.*	β=0.1, SE=0.2, *n.s.*	β=0, SE=0.2, *n.s.*

Table 3.4: Results of pair-wise comparisons by Focus Type for formation duration, displacement, and peak velocity.

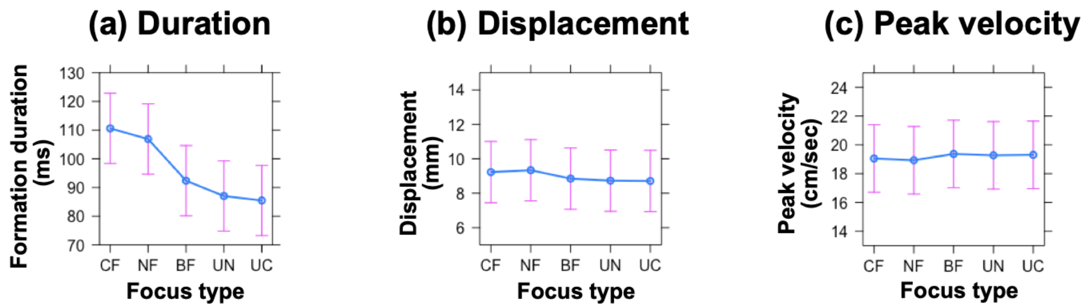

Figure 3.14: Results on (a) duration (in ms), (b) displacement (in mm), and (c) peak velocity (in cm/sec).

show any significant difference between focus type pairs (Figure 3.15c).

In addition, all measures showed significant interactions with the continuous factors, namely stiffness for formation duration, peak velocity for displacement, and displacement for peak velocity. Figure 3.16 shows the interaction between the measured parameters and the continuous factors. Interestingly, the slopes of the relationships were further modulated by focus type. Specifically, duration showed three degrees of focus: CF, NF > BF > UN, UC. CF and NF had higher slope values, 3.3 and 3.4, respectively, than BF, which had a slope value of 2.3. BF, in turn, had higher slope values than UN and UC, which had 1.8 as their slope values. For displacement, two degrees were distinguished: CF, NF > BF, UN, UC. Again, higher slope values were found for the group of CF and NF (0.4 for both) compared to the group of BF, UN, and UC (0.3 for all). However, for peak velocity, while it also distinguished two degrees: CF, NF < BF, UN, UC, the directionality of the effect was opposite to that of duration and displacement—slopes were shallower for CF and NF (1.4 for both) and steeper for BF, UN, and UC (1.6 for BF, 1.7 for both UN and UC). That is, for an increase of one displacement unit (mm), the increase in peak velocity was greater for BF, UN, and UC (either 1.6 or 1.7 cm/sec) than CF and NF (1.4 cm/sec).

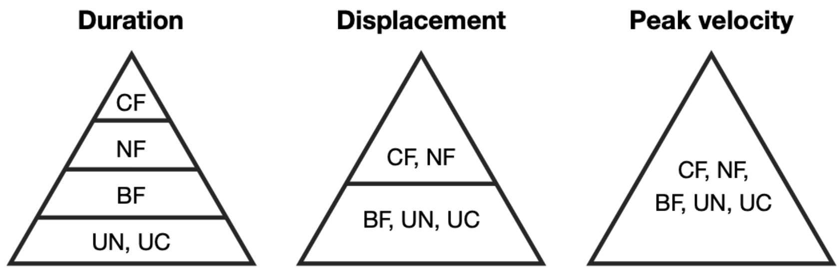

Figure 3.15: Schematic representation of the focus types categorized by kinematic parameters.

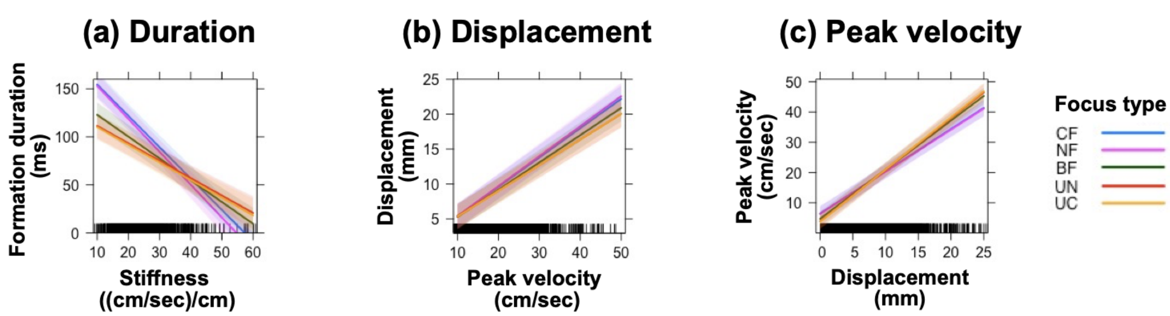

Figure 3.16: Interrelationship between kinematic parameters by focus type: (a) duration (in ms) by stiffness (in (cm/sec)/cm), (b) displacement (in mm) by peak velocity (in cm/sec), and (c) peak velocity by displacement.

3.4.3 Interim discussion

In Experiment 3, it was investigated whether different types of focus, i.e., contrastive focus, narrow (but not contrastive) focus, broad focus, and unfocused, are differentiated in their effect on articulation. The results confirm that prominence encodes the information focus structure and not simply on the basis of presence or absence of focus (cf. Katsika et al., 2023; Roessig & Mücke, 2019; Mücke & Grice, 2014). The different focus types indeed were differentiated in their modulations in articulation and different kinematic parameters exhibit different levels. Specifically, duration showed the most levels of focus, meaning that it distinguished the most levels of focus with contrastive focus having the greatest duration and the two unfocused conditions having the least duration. Displacement distinguished two levels, with the distinction drawn between narrow focus and broad focus. Peak velocity was not affected by focus nor showed any effect of focus type, aligning with the results in Experiment 1 and Experiment 2 which suggested that peak velocity may not be a parameter that is directly controlled by prominence.

Another interesting finding is that the interrelationships between kinematic parameters, i.e., the relationship between duration and normalized peak velocity over displacement (referred to as *stiffness*), and the relationship between displacement and peak velocity, are further modulated by this focus type. Especially for duration, which showed the greatest distinction between focus types, its correlation with stiffness was also distinguished by focus type, such that three groups were formed: (1) contrastive focus and narrow focus, (2) broad focus, and (3) two unfocused conditions. This result adds to the finding of the previous experiment (Experiment 2) that the correlation between duration and stiffness is not only controlled by the presence or absence of prominence, but also by the information structure. In terms of displacement and peak velocity, two groups were distinguished: (1) contrastive focus and narrow focus and (2) broad focus and two unfocused conditions. For displacement, the direction was that gestures receiving narrow focus, regardless of whether it is contrastive or not, had steeper slope than the broad focus or unfocused conditions. Peak velocity had an opposite direction such that the more prominent group had shallower slope than the non-prominent group, showing less increase in peak velocity, given the same increase in displacement.

Thus, results of Experiment 3 suggest a hierarchical structure of the focus types exerting cumulative effects on the articulatory realization in Korean, similar to the results reported for head-prominence languages (e.g., Katsika et al., 2023; Roessig & Mücke, 2019; Mücke & Grice, 2014).

3.5 Discussion

The main goal of this chapter was to examine the kinematic realization of focus-induced prominence in Seoul Korean, an edge-prominence language. To this end, the following three questions were assessed: The first question was what the articulatory modulation of prominence as well as its interaction with IP boundary-marking in Seoul Korean was. The prediction was that the gestures would be longer, larger, and faster under focus, in accordance with the patterns found in other languages (e.g., Beckman & Edwards, 1994; Beckman et al., 1992; Cho, 2006; de Jong, 1991, 1995; Fowler, 1995) as well as in Korean (Shin et al., 2015). As Korean marks prominence by the means of phrasing, it may be predicted that an interaction between prominence- and IP boundary-marking is found. However, in what way it would interact was an open question. Second, we examined the articulatory profile of dephrasing, a phenomenon of tonal effects on the post-focal parts of the utterance, which are well established (e.g., Jun, 1993, 2005, 2011; Jun & Lee, 1998). The hypothesis tested was that the attenuation on the tonal dimension would be accompanied by reduction on the articulatory dimensions as well. In other words, we expected shorter and smaller gestures with respect not just to focused gestures but also to unfocused gestures belonging in the pre-focal part of the utterance. With respect to dephrasing, we also investigated whether the degree of articulatory attenuation would vary as a function of focus position. Our expectation was that articulatory dephrasing would be stronger in the vicinity of the focused unit. Third, we examined the effect of focus type on articulation, and whether prominence-induced articulatory modulation would simply encode presence or absence of focus or whether it would, instead, reflect different focus type. As a reminder, based on limited evidence from head-prominence languages (Hermes et al., 2008; Roessig & Mücke, 2019; Mücke & Grice, 2014 for German; see also Katsika et al., 2020, 2023 for American English), we expected to find different degrees of prominence reflecting focus type in the following order from low to high: broad focus, then narrow focus, and ultimately contrastive focus. An explorative thread underlying all three questions was whether established relationships between kinematic parameters under the assumptions of a linear critically damped second-order dynamcial system (Byrd & Saltzman, 1998; Munhall et al., 1985) varied systematically depending on prosodic position (e.g., dephrased or initial in focused AP or IP-final). If such systematicities were to be found, implications for prosodic modulation on the kinematic profile of utterances could be drawn.

Combined results suggest that gestures under focus are longer and larger than their unfocused counterparts. These patterns are in line with previous literature, which reported temporal and spatial expansion of the prominent gestures (Beckman et al., 1992; Cho, 2006; de Jong, 1991, 1995). This body of work also detected faster gestures under prominence. The current experiments, however, showed that although peak velocity increased with displacement, as expected in models under Task Dynamics (e.g., Munhall et al., 1985; Ostry & Munhall, 1985), it was not further modulated by focus-induced prominence. This finding raises the question whether the faster movements found in other languages would still be maintained if the interrelationship between peak velocity and displacement was considered.

As far as the scope of focus-induced prominence is concerned, the effect extends away from the left-edge of the focused AP, spanning over at least two syllables. In Experiment 2, the effect even reached the resyllabified onset of the third syllable. As in both cases of Experiments 1 and 2 the effect spans up to the first morphological boundary of the test words, it may be suggested that prominence marking in Korean affects the first morpheme of the focused linguistic unit. The current studies were not designed to test this, and calls for further investigation on this matter. However, it needs to be noted that the patterns detected in Experiment 1 also indicated that focus-marking effects might be spilling over across a word boundary, a possibility that the next steps of this work will explore further.

Moreover, the expansion of the effect beyond the local AP boundary suggests that this is indeed an effect driven by focus-induced prominence, and not an effect demarcating the AP-initial boundary. Domain-initial strengthening, i.e., stronger articulation of gestures at the left-edge of prosodic domains, is a well-established phenomenon (Fougeron & Keating 1997; Cho & Keating 2001, 2009; Fougeron 2001; Keating, Cho, Fougeron & Hsu 2003; Cho & McQueen, 2005; see Cho, 2016 for a review). As phrase-level prominence in Korean is realized by inserting an AP boundary before the focused word (e.g., Jun, 1993), the longer and larger gestures found here may be attributable to AP-initial strengthening. However, the scope of the domain-initial strengthening has been assumed to be generally localized on the domain-initial gesture (cf. Cho, 2016). Here, however, we find an effect that affects two, possibly, three syllables, indicating focus-marking and AP-initial strengthening are distinct effects, even in an edge-prominence language. This conclusion is further supported by both the patterns of dephrasing and the effects of focus type found here, as further discussed below.

An intricate relationship between phrase-level prominence marking and IP boundary marking is suggested in Korean. Specifically, the initial gesture of the test word is the smallest when IP-final and unfocused. At this point, it cannot be confirmed whether the smaller displacement in IP-final position as opposed to IP-medial position may be coming from a similar mechanism to the durational effect of preboundary shortening (Krivocapić, 2014; Katsika, 2016; Kim et al., 2017) or from word-boundary demarcation. Nonetheless, it seems that this effect involving IP boundary-marking is constrained by focus-marking, such that the effect only appears when there is absence of the effect of prominence.

In terms of dephrasing, results suggest that the effect may not be an articulatory attenuation, but more so a tonal modification, of the post-focal linguistic units. Experiment 1 found that the two unfocused conditions, one being AP-initial and the other being AP-medial (due to dephrasing), were largely similar to each other. This finding is congruent with the conclusion that the articulatory strengthening observed in the initial gestures of the focused AP encode focus rather than the AP-initial boundary. The only significant difference was contrary to predictions, since the word-initial C gesture showed a larger displacement when it was dephrased as opposed to unfocused. This is presumably due to a spillover effect of focus, which in these data was in the adjacent noun phrase. However, further investigation is needed to confirm this possibility.

Results also show that focus type exerts different degrees of articulatory modulation, presenting cumulative effects from no focus to contrastive focus, similar to effects found in head-prominence languages (Hermes et al., 2008; Roessig & Mücke, 2019; Mücke & Grice, 2014 for German; see also Katsika et al., 2020, 2023 for American English). In the duration measure, gestures receiving contrastive focus were longer compared to those receiving narrow focus, which were in turn longer than those receiving broad focus. Unfocused gestures, because of the presence of a contrastive or narrow (but not contrastive) focus on another word, were the shortest. Displacement distinguished two degrees of prominence, one with contrastive and narrow focus, and another with broad focus and no focus. Like in Experiments 1 and 2, peak velocity was not affected by any focus type. Thus, similarly to the little evidence coming from head-prominence languages, the Korean data examined here suggest a specific order, and thus, a hierarchical structure in the phonetic realization of the focus types. However, it is yet unclear where this hierarchical structure comes from. One possibility is that this reflects a hierarchical order in information structure, which interfaces with prosodic structure. Another alternative could be that the hierarchy emerges from different degrees of trade-offs between linguistic and prosodic information demanded by the different focus types, as it would be predicted, for instance, by the Speech Signal Redundancy Hypothesis (cf. Aylett & Turk, 2004). That is, information structure might not be directly encoded in the prosodic structure of language, and the intensity of the prosodic effect might result from the degree of linguistic redundancy of the target word in a given focus type. In experimental set-ups that require the participants to use the same word order to express different focus conditions, linguistic redundancy of the focused word is expected to decrease from broad to narrow and from narrow to contrastive focus, and speakers can only rely on prosodic redundancy in order to express prominence. It is yet to be determined if similar granularity of focus marking is to be observed when speakers are left free to use other, linguistic devices of focus marking, such as different word orders or morphological markers of focus.

As current theories of prosody stand, a specific account of the connection between information structure to prosodic structure is lacking. For instance, in the framework of Articulatory Phonology, modulation gestures, or mu-gestures, have been proposed to model prominence-related manifestation of gestures (Saltzman et al., 2008). In this model, the cumulative effect of focus type could in principle be captured by mu-gestures of different strengths, similarly to the control of pi-gesture strength that has been used to model effects of boundaries at different levels of the prosodic hierarchy (cf. Byrd & Saltzman, 2003). However, an equivalent hierarchy of prominence in which focus type is encoded has not been established for prominence (see for instance hierarchies proposed in metrical phonology; e.g., Liberman & Prince, 1977; Nespor & Vogel, 1986; Pierrehumbert & Beckman, 1988; Selkirk, 1984). The non-linear dynamical model of prominence marking proposed in Roessig and Mücke (2019) captures the cumulative (kinematic and F0) effects of prominence across focus types by controlling a single dynamical parameter. This model views prominence as a gradient system, predicting an increasing degree in prominence and a set of phonetic parameters marking them that work in tandem. A direct link to either information or prosodic structure levels/categories is not assumed. Why however these degrees would correspond to the specific focus types and in that order in languages is not clear. The results of this article combined with the results of the previous article which showed that focus affects IP boundary marking as well calls for further development of theories on the connection between information structure and prosodic structure.

The cumulative effect of focus type further indicates that such modulations are prominence-marking in nature and not simply domain-initial strengthening. This is because in all focused conditions, the focus item begins an AP, but the degree of strengthening varies with focus type.

Finally, the present study finds robust effects of prosodic position on the relationship between duration and stiffness, estimated by normalized peak velocity over displacement. The slope of the regression line between duration and stiffness is greater for gestures at the beginning of the focused AP as opposed to gestures medial in the AP. This pattern is even more pronounced when focus type is considered, with slope values at the beginning of focused APs distinguishing three degrees of prominence: (1) contrastive and narrow focus, (2) broad focus, and (3) no focus. Slopes are the steepest for gestures at the IP-final position (see Section 2.3.4 of Experiment 2). Combined with the results of Article 1 on the kinematic profile of IP boundaries, it may be concluded that the prosodic functions of prominence- and IP boundary-marking present distinct stiffness profiles on articulation, suggesting that such prosodic modulations might be arising at least partly from controlling gestural stiffness (see discussion in Cho, 2006). Stiffness could not be the only parameter controlled though, since this would not be able to capture the displacement effects. For that, we would need to control the gestural target. Note, however, that the relationship between displacement and peak velocity does not present the same dependency on prosodic position as the relationship between duration and stiffness. According to Cho (2006), shrinking (or less shrinking), defined as "a change in both target and stiffness", may account for the effect found for duration and displacement, but not in peak velocity. In the mu-gestural model, we would need both a spatial and a temporal mu-gesture to be co-active in order to get the longer and larger gestures found here. However, mu-gestures would not be able to capture the distinct stiffness profiles presented by prominent and IP-final positions (see also discussion in Article 1 and in Iskarous & Pouplier, 2022). On the other hand, the similarity of the patterns found here to those in Roessig and Mücke (2019) combined with the distinct profiles that the relationship between duration and peak velocity normalized over displacement (called *stiffness*) shows by focus type in the current chapter and by prosodic position (in the previous chapter) may be considered as supporting proposals for re-consideration of the dynamics of gestures and their prosodic modulations by the means of non-linear dynamics (cf. Roessig & Mücke, 2019; Sorensen & Gafos, 2016; Iskarous & Pouplier, 2022).

In sum, this article examines the articulatory modulations of focus-induced prominence in Seoul Korean and reports that gestures under focus are longer, larger, but not necessarily faster than their unfocused counterparts. Focus is marked by kinematic patterns distinguished from the IP boundary-related effects. This indicates that in an edge-prominence language, such as Korean, the functions of prominence- and IP boundary-marking are distinct, despite the fact that both employ prosodic boundaries. Prominence-marking does not simply encode presence or absence of focus, but also different focus types, showing a cumulative effect from no focus to broad, narrow, and finally, contrastive focus. Furthermore, gestures under prominence and at IP boundaries demonstrated distinct profiles of stiffness.

References

Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and speech*, *47*(1), 31–56.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: 10.18637/jss.v067.i01

Baumann, S., Grice, M., & Steindamm, S. (2006). Prosodic marking of focus domains-categorical or gradient. In *Proceedings of the 3rd international conference on speech prosody 2006* (Vol. 2006, pp. 301–304).

Beckman, M. E. (1996). The parsing of prosody. *Language and cognitive processes*, *11*(1-2), 17–68.

Beckman, M. E., & Edwards, J. (1990). Lengthenings and shortenings and the nature of prosodic constituency. In J. C. Kingston & M. E. Beckman (Eds.), *Papers in laboratory phonology i:, between the grammar and physics of speech* (Vol. 152, p. 178). Cambridge, UK: Cambridge University Press.

Beckman, M. E., & Edwards, J. (1994). Articulatory evidence for differentiating stress categories. *Papers in laboratory phonology III: Phonological structure and phonetic form*, 7–33.

Beckman, M. E., Edwards, J., & Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology ii: Gesture, segment, prosody* (pp. 68–89). Camdrige, UK: Cambridge University Press.

Beckman, M. E., & Pierrehumbert, J. (1988). *Japanese tone structure.* MIT Press Cambridge.

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

Breen, M., Fedorenko, E., Wagner, M., & Gibson, E. (2010). Acoustic correlates of information structure. *Language and cognitive processes*, *25*(7-9), 1044–1098.

Byrd, D., & Saltzman, E. (1998). Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics*, *26*(2), 173–199.

Byrd, D., & Saltzman, E. (2003). The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, *31*(2), 149–180.

Cho, T. (2006). Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English. *Laboratory phonology*, *8*, 519–548.

Cho, T. (2011). Laboratory phonology. *The continuum companion to phonology*, 343–368.

Cho, T. (2015). Language effects on timing at the segmental and suprasegmental levels. *The handbook of speech production*, 505–529.

Cho, T. (2016). Prosodic boundary strengthening in the phonetics–prosody interface. *Language and Linguistics Compass*, *10*(3), 120–141.

Cho, T. (2022). Linguistic functions of prosody and its phonetic encoding with special reference to Korean. *Japanese/Korean Linguistics*, *29*.

Cho, T., & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, *37*(4), 466–485.

Cho, T., & Keating, P. A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of phonetics*, *29*(2), 155–190.

Cho, T., & McQueen, J. M. (2005). Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, *33*(2), 121–157.

Cruttenden, A. (1997). *Intonation.* Cambridge, UK: Cambridge University Press.

De Jong, K. J. (1991). *The oral articulation of English stress accent* (PhD Dissertation). The Ohio State University.

De Jong, K. J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *The journal of the acoustical society of America*, *97*(1), 491–504.

Edwards, J., Beckman, M. E., & Fletcher, J. (1991). The articulatory kinematics of final lengthening. *the Journal of the Acoustical Society of America*, *89*(1), 369–382.

Fletcher, J. (2010). The prosody of speech: Timing and rhythm. In F. Gibbon, W. J. Hardcastle, & J. Laver (Eds.), *The handbook of phonetic sciences* (p. 521-602). John Wiley & Sons, Ltd.

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

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *The journal of the acoustical society of America*, *101*(6), 3728–3740.

Fowler, C. A. (1995). Acoustic and kinematic correlates of contrastive stress accent in spoken English. *Producing speech: Contemporary issues: For Katherine Safford Harris*, 355–373.

Fox, A. (2002). *Prosodic features and prosodic structure: The phonology of 'suprasegmentals'*. Oxford, UK: Oxford University Press.

Gussenhoven, C. (1983). Testing the reality of focus domains. *Language and speech*, *26*(1), 61–80.

Hawkins, S. (1992). An introduction to task dynamics. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology ii: Gesture, segment, prosody* (pp. 9–25). Camdrige, UK: Cambridge University Press.

Hermes, A., Becker, J., Mücke, D., Baumann, S., & Grice, M. (2008). Articulatory gestures and focus marking in German. *Proceedings of the 4th international conference on speech prosody 2008*, 457–460.

Hyman, L. M. (1978). Word demarcation. *Universals of human language*, *2*, 443–470.

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

Jun, S.-A. (1993). *The phonetics and phonology of Korean prosody* (PhD 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 Sciences*, *7*(1), 143–170.

Jun, S.-A. (2005). Korean intonational phonology and prosodic transcription. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 201–229). Oxford, UK: Oxford University Press.

Jun, S.-A. (2006). Intonational phonology of Seoul Korean revisited. In *Japanese-korean linguistics* (Vol. 14, pp. 15–26).

Jun, S.-A. (2011). Prosodic markings of complex NP focus, syntax, and the pre-/postfocus string. In *Proceedings of the 28th west coast conference on formal linguistics* (pp. 214–230).

Jun, S.-A. (2014). Prosodic typology: By prominence type, word prosody, and macro-rhythm. In *Prosodic typology ii: The phonology of intonation and phrasing* (pp. 520–539). Oxford, UK: Oxford University Press.

Jun, S.-A., & Lee, H.-J. (1998). Phonetic and phonological markers of contrastive focus in Korean. In *Proceedings of the 5th international conference on spoken language processing.*

Katsika, A., Jang, J., Krivokapić, J., Goldstein, L., & Saltzman, E. (2020). The role of focus in accentual lengthening in American English: Kinematic analyses. In *Proceedings of the 10th international conference on speech prosody 2020* (pp. 275–279).

Katsika, A., Jang, J., Krivokapić, J., Goldstein, L., & Saltzman, E. (2023). A hierarchy of prominence: The production and perception of focus in American English. In *Proceedings of the international congress of phonetic sciences 2023* (pp. 1696–1700).

Katsika, A., & Tsai, K. (2021). The supralaryngeal articulation of stress and accent in Greek. *Journal of Phonetics*, *88*, 101085.

Keating, P., Cho, T., Fougeron, C., & Hsu, C. (2003). Domain-initial strengthening in four languages. In J. Local, R. Ogden, & R. Temple (Eds.), *Papers in laboratory phonology vi: Phonetic interpretations* (pp. 145–163). Camdrige, UK: Cambridge University Press.

Kim, S., Jang, J., & Cho, T. (2017). Articulatory characteristics of preboundary lengthening in interaction with prominence on tri-syllabic words in American English. *The Journal of the Acoustical Society of America*, *142*(4), EL362–EL368.

Krivokapić, J. (2007). *The planning, production, and perception of prosodic structure* (PhD dissertation). University of Southern California.

Ladd, D. R. (2008). *Intonational phonology.* Cambridge University Press.

Lehiste, I. (1970). *Suprasegmentals.* Cambridge, MA: MIT Press.

Lenth, R. V. (2020). *emmeans: Estimated marginal means, aka least-squares means.* Retrieved from `https://CRAN.R-project.org/package=emmeans` (R package)

Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic inquiry*, *8*(2), 249–336.

Mücke, D., & Grice, M. (2014). The effect of focus marking on supralaryngeal articulation–is it mediated by accentuation? *Journal of Phonetics*, *44*, 47–61.

Munhall, K. G., Ostry, D. J., & Parush, A. (1985). Characteristics of velocity profiles of speech movements. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(4), 457.

Nespor, M., & Vogel, I. (2007). *Prosodic phonology: with a new foreword* (Vol. 28). Mouton de Gruyter.

Ostry, D. J., & Munhall, K. G. (1985). Control of rate and duration of speech movements. *The Journal of the Acoustical Society of America*, *77*(2), 640–648.

R Core Team. (2021). *R: A language and environment for statistical computing.* Vienna, Austria. Retrieved from `https://www.R-project.org/`

Roessig, S., & Mücke, D. (2019). Modeling dimensions of prosodic prominence. *Frontiers in Communication*, *4*, 44.

Roon, K. D., Gafos, A. I., Hoole, P., & Zeroual, C. (2007). Influence of articulator and manner on stiffness. In *Proceedings of the 16th international congress of phonetic sciences* (pp. 409–412).

Saltzman, E., Nam, H., Krivokapic, J., & 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* (pp. 175–184).

Selkirk, E. (1984). *Phonology and syntax: The relationship between sound and structure.*

Cambridge, MA: MIT Press.

Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of psycholinguistic research*, *25*, 193–247.

Shin, S., Kim, S., & 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.*

Sorensen, T., & Gafos, A. (2016). The gesture as an autonomous nonlinear dynamical system. *Ecological Psychology*, *28*(4), 188–215.

Van der Hulst, H. (1999). Word accent. *Empirical Approaches to Language Typology*, 3–116.

Venditti, J. J., Jun, S.-A., & Beckman, M. E. (2014). Prosodic cues to syntactic and other linguistic structures in Japanese, Korean, and English. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 299–324). Psychology Press.

Chapter 4

General Discussion

This dissertation investigated the prosodic modulations, i.e., boundary- and prominence-marking, and their interactions on articulation in Seoul Korean. The following questions were assessed via an electromagnetic articulography study:

1. How are boundaries at major phrases (i.e., at the Intonational Phrase (IP) level) marked and how do they interact with the prosodic level of Accentual Phrases, which is also the level on which prominence is marked?

2. What is the articulatory profile of Accentual Phrases under focus and how does it compare to that of non-focused Accentual Phrases? Do different types of focus exert different effects on articulation?

3. Do interrelationships between kinematic parameters depend on prosodic position, differentiating the prosodic marking of grouping from at of prominence?

One of the larger goals of this dissertation was to examine whether the two functions of prosodic structure, i.e., grouping- and prominence-marking, would be differentiated in their manifestation in an edge-prominence language, where both functions seem to be utilizing *edges/phrasal boundaries* as the means. Results of the current dissertation suggest distinct articulatory modulations of boundary- and prominence-marking in Seoul Korean.

In terms of boundary-related effects, gestures at the right-edge of an IP boundary are longer, larger, and slower in Seoul Korean, in line with previous literature (e.g., Byrd et al., 2006; Edwards et al., 1991; Fletcher, 2010). These effects hold beyond the relationships between kinematic parameters, i.e., inverse relationship between duration and stiffness, and direct relationship between displacement and peak velocity. Phrase-final lengthening consistently affected the phrase-final syllable. Both phrasal prominence and factors that undertake the function of word demarcation affect the scope of IP-final lengthening: Lengthening was greater the further away focus was from the boundary, and the effect extended to the onset of the penultimate syllable of the phrase-final prosodic word when the further away the left edge of the final AP was from IP's end.

Under focus-induced prominence, on the other hand, gestures are longer and larger, but not faster per se. Temporal and spatial expansion of the prominent gestures is in line with previous literature (Beckman et al., 1992; Cho, 2006; de Jong, 1991, 1995). However, while previous studies also report faster gestures under prominence, results of the current dissertation indicated that peak velocity was not modulated by focus-induced prominence beyond the expected interrelationship with displacement (e.g., Munhall et al., 1985; Ostry & Munhall, 1985). This finding calls for further examination of the faster movements reported in other languages to test whether the effect is sustained even when the relationship between peak velocity and displacement is taken into account.

The patterns of the effects of focus suggest that prosodic prominence is driven by focus, and not an effect coming from domain-initial strengthening at the AP level. While the scope of domain-initial strengthening has been assumed to be localized to the domain-initial gesture (cf. Cho, 2016), focus-driven effects in the current study extends over two, or possibly, three syllables. Moreover, the fact that unfocused AP-initial gestures were not so different from unfocused AP-medial ones also adds to the argument that AP-initial gestures do not undergo domain-initial strengthening and that instead any kinematic effects on the left edge of APs are driven by focus.

Different focus types systematically exercise different degrees of articulatory modulation. The effects suggest a hierarchical structure, showing a cumulative pattern with contrastive focus demonstrating the strongest effect and no focus demonstrating the least effect. Similar patterns are reported in head-prominence languages (e.g., German: Hermes et al., 2008; Roessig & Mücke, 2019; Mücke & Grice, 2014; American English: Katsika et al., 2020, 2023), suggesting that a connection between focus structure and the prosodic function of prominence-marking may be a general property of language. However, it is yet unclear where this hierarchical structure comes from and a specific account of the connection between information structure to prosodic structure is lacking in current theories of prosody. Thus, results of this dissertation call for further development of theories on the connection between information structure and prosodic structure.

Finally, throughout the dissertation, robust effect of prosodic structure on the relationship between duration and stiffness, as approximated by nornalized peak velocity over displacement, has been found. The slope of the regression line between duration and stiffness is greater for gestures in prosodically prominent positions. Under focus, the gesture's slope values are higher than their unfocused counterparts. Different focus types also exert different effects, distinguishing multiple levels of prominence: (1) contrastive and narrow focus; (2) broad focus; and (3) no focus. This fine-tuning is further differentiated by the source of the prosodic effect—that is, different slopes are found for prominence- vs. boundary-markings. IP-final gestures show the steepest slopes. Combining the results of both Articles 1 and 2 (Chapters 2 and 3), it is concluded that the prosodic functions of prominence- and IP boundary-marking present distinct stiffness profiles on articulation.

To conclude, this dissertation examined the phonetic modulation of grouping- and prominence-marking in Seoul Korean and found that the two functions exert distinct manifestations on articulation. This finding adds to the existing literature by illustrating that even in an edge-prominence language like Korean, in which prominence is marked by the means of grouping, the two functions of prosodic structure are encoded distinctively in the phonetic realization of the language. Furthermore, evidence for an interaction between the two prosodic functions was detected, analogous to effects found in head-prominence languages (cf. Katsika, 2016). Thus, the analyses of gestures in this dissertation are fruitful as they provide us with patterns that are comparable across examined languages, as well as offer the ground for insightful discussion of the interface between prosodic structure and information structure.

References

Beckman, M. E., Edwards, J., & Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology ii: Gesture, segment, prosody* (pp. 68–89). Camdrige, UK: Cambridge University Press.

Byrd, D., Krivokapić, J., & Lee, S. (2006). How far, how long: On the temporal scope of prosodic boundary effects. *The Journal of the Acoustical Society of America*, *120*(3), 1589–1599.

Cho, T. (2006). Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English. *Laboratory phonology*, *8*, 519–548.

Cho, T. (2016). Prosodic boundary strengthening in the phonetics–prosody interface.

Language and Linguistics Compass, *10*(3), 120–141.

De Jong, K. J. (1991). *The oral articulation of English stress accent* (PhD Dissertation). The Ohio State University.

De Jong, K. J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *The journal of the acoustical society of America*, *97*(1), 491–504.

Edwards, J., Beckman, M. E., & Fletcher, J. (1991). The articulatory kinematics of final lengthening. *the Journal of the Acoustical Society of America*, *89*(1), 369–382.

Fletcher, J. (2010). The prosody of speech: Timing and rhythm. In F. Gibbon, W. J. Hardcastle, & J. Laver (Eds.), *The handbook of phonetic sciences* (p. 521-602). John Wiley & Sons, Ltd.

Hermes, A., Becker, J., Mücke, D., Baumann, S., & Grice, M. (2008). Articulatory gestures and focus marking in German. In *Proceedings of the 4th international conference on speech prosody 2008* (pp. 457–460).

Katsika, A., Jang, J., Krivokapić, J., Goldstein, L., & Saltzman, E. (2020). The role of focus in accentual lengthening in American English: Kinematic analyses. In *Proceedings of the 10th international conference on speech prosody 2020* (pp. 275–279).

Katsika, A., Jang, J., Krivokapić, J., Goldstein, L., & Saltzman, E. (2023). A hierarchy of prominence: The production and perception of focus in American English. In *Proceedings of the international congress of phonetic sciences 2023* (pp. 1696–1700).

Mücke, D., & Grice, M. (2014). The effect of focus marking on supralaryngeal articulation–is it mediated by accentuation? *Journal of Phonetics*, *44*, 47–61.

Munhall, K. G., Ostry, D. J., & Parush, A. (1985). Characteristics of velocity profiles of speech movements. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(4), 457.

Roessig, S., & Mücke, D. (2019). Modeling dimensions of prosodic prominence. *Frontiers in Communication*, *4*, 44.