10th International Conference on Speech Prosody 2020
25-28 May 2020, Tokyo, Japan

Pitch accent and phrase boundaries: Kinematic evidence from Japanese

Karen Tsai[1], *Argyro Katsika*[1]

[1]University of California, Santa Barbara, USA
karentsai@ucsb.edu, argyro@ucsb.edu

Abstract

Constriction gestures at phrase boundaries are longer, larger and slower. However, the scope of these effects, i.e., the stretch of speech affected, is still unclear. Previous work has mainly focused on stress languages and suggests that lexical prominence plays a key role in determining the scope of phrase-final lengthening. Less is understood about other kinematic dimensions and the interaction between lexical prominence and boundary-related effects in languages without stress, although two acoustic studies on Japanese suggest that lexical pitch accent affects the amount of phrase-final lengthening. The current Electromagnetic Articulography study examines the amount and scope of phrase-final effects in two kinematic dimensions, duration and velocity, as a function of pitch accent position in Japanese. Results show that phrase-final lengthening affects the formation of the constriction gesture in the onset of the final syllable. Words with accent on the first or second syllable show greater amount of lengthening on that formation, as well as lengthening of the same gesture's release. This is consistent with research detecting phrase-final lengthening on boundary-adjacent syllables. Analysis of normalized peak velocity revealed effects of accent, but no boundary-related slowing. The implications of these findings for prosodic structure are discussed, and typological distinctions are highlighted.

Index Terms: phrase-final lengthening, phrase-final slowing, lexical pitch accent, articulation, Japanese

1. Introduction

Prosodic boundaries are known to affect speech segments in kinematic dimensions such as duration, displacement and velocity. For instance, phrase-final constriction gestures are longer, larger, and faster than their phrase-medial counterparts (cf. [1]). Lengthening before prosodic boundaries, which is also known as phrase-final or pre-boundary lengthening, is supported by both acoustic and kinematic data derived from different languages and language varieties (e.g., acoustics: [2, 3, 4, 5, 6, 7, 8]; articulation: [9, 10, 11, 12, 13, 14, 15, 16, 17]. While this effect is well established, the scope of the effect, meaning the stretch of speech that undergoes lengthening, is unclear. Previous work locates the largest and most systematic part of the effect at the rhyme of the phrase-final syllable, a domain that seems to be important for the perception of boundaries as well [18]. However, most of this work did not look beyond that domain. The few studies that did suggest that phrase-final lengthening has small, but nevertheless systematic, effects earlier within the phrase-final word than the rhyme of its final syllable. For instance, phrase-final lengthening extends to the onset of the final syllable when the syllable's vowel is reduced or not a diphthong (English: [2], Dutch: [6]). Similarly, a small set of studies suggests that phrase-final lengthening extends away from the rhyme of the phrase-final syllable due to stress, and it might even reach the stressed syllable (e.g., Hebrew: [5], Estonian: [19], English: [20]), even when stress is more than two syllables away from the boundary (e.g., English: [8, 7, 21]). It is unclear whether the interaction of phrase-final lengthening with stress involves a single domain, for example the interval from the coda of the stressed syllable to the boundary [7], or two separate domains, i.e., the rhyme of the final syllable and the rhyme of the stressed syllable, leaving any intervening syllables unaffected [8] (but see [21]). A recent articulatory study that systematically disentangled the effect of lexical stress from that of phrasal pitch accent on boundary-related events in Greek found that phrase-final lengthening was affected by stress, and not pitch accent, and had a continuous domain [16]; the effect affected the constriction gestures that were immediately adjacent to the boundary when stress was final, but was initiated earlier the earlier the stress was within the word. These results have been accounted for using the π-gesture model [22], in which the temporal properties of constriction gestures are modulated by clock-slowing π-gestures [21, 16].

However, this minimal work on the scope of phrase-final lengthening is mainly focused on stress languages and little is known about whether, and if yes, how languages with different prosodic systems, such as languages with lexical pitch accent, present phrase-final lengthening. It seems that while phrase-final lengthening may be universal, its specific interaction with lexical prosody may be specific to the given prosodic system, or even language. The current study investigates Japanese, a pitch accent language, in which lexical prominence is marked by a fall in F0 on the mora following the accent [23]. Japanese is an excellent language with which to examine this topic, as pitch accent in Tokyo Japanese may occur on any syllable [24], resulting in minimal sets such as /na.mi/ 'medium', /na*.mi/ 'Nami (name)' and /na.mi*/ 'wave', where the asterisk (*) indicates pitch accent. Two acoustic studies of initial-accented and unaccented disyllabic words in Japanese [25, 26] found that pitch accent affected the amount of lengthening on the final rhyme (final vowel in CV.CV words); words with pitch accent showed less boundary lengthening compared to unaccented words. However, pitch accent did not appear to affect the scope of lengthening, as lengthening was found up to the first vowel for both accent conditions.

The current study examines the interaction of pitch accent with boundary-related lengthening and slowing in two- and three-syllable words with pitch accent in all possible positions in Japanese. We aim to address the following questions: (1) when does phrase-final lengthening and slowing begin, (2) does the presence of pitch accent affect the amount and scope of these effects, and (3) does the position of lexical pitch accent affect the amount and scope of these effects? Duration is not typically considered a correlate of pitch accent in Japanese [27, 28, 29, 30, 31], so it is unlikely we would find a pattern similar to English, where the prominent syllable is lengthened as a function of boundary [1, 8, 21]. Rather, it is more plausible that the effects would be found within the final syllable. Another possibility is that similar to Greek, we might expect to

 10.21437/SpeechProsody.2020-44

see the initiation of boundary lengthening to vary as a function of pitch accent position [16]. As for velocity, we expect to find slowing of gestures in the vicinity of the boundary, based on evidence from English that lip gestures at prosodic boundaries are longer and slower [1, 17].

2. Methodology

2.1. Participants and recording apparatus

Data from four native speakers of Tokyo Japanese (1 male and 3 females, M_{age}=22 years, range 20-25 years) have been analyzed to date. Participants were not aware as to the purpose of the experiment. No speech, hearing, or vision problems were reported. Speakers were compensated for their participation.

Data were collected using the AG501 three-dimensional electromagnetic transduction device (Carstens Medizinelektronik) at the UCSB Phonetics Lab. Receiver coils were attached to five areas of interest: tongue dorsum, tongue body's center, tongue tip, upper lip and lower lip. An additional five were attached as reference points to the upper incisor, jaw, nose, and left and right ears. Audio recordings were performed simultaneously to the kinematic recordings by the means of a Sennheiser shotgun microphone set at a sampling rate of 16 kHz.

2.2. Experimental design and procedure

Two- and three-syllable words were tested. Stimuli for two-syllable words were a minimal set differing only in pitch accent: unaccented (P0), initial-accented (P1) and final-accented (P2). Stimuli for three-syllable words were a near-minimal set also differing in pitch accent (two words for P2 condition were included since we were unable to find words with alternating constriction locations). All target words are shown in Table 1. Frame sentences elicited these in two phrasal positions: phrase-medial (PhM) or phrase-final (PhF) (see Table 2 for sample sentences). All frame sentences were interrogatives in order to examine other boundary events not discussed in the present study. Context sentences were used to create short dialogues prompting target words in their frame sentences. Stimuli were presented in Kanji and Hiragana on a computer screen placed about 60 inches away from the participant. Context sentences were read silently and target sentences aloud. In total, 576 utterances were collected (2 phrasal boundaries x 3-4 pitch accent x 9 repetitions x 4 speakers), and 525 were analyzed for this paper (some tokens were excluded due to speech errors or technical reasons).

Table 1: *Target words by experimental factor*

Accent Position	Two-σ Words	Three-σ Words
Unaccented (P0)	**nami** *medium*	**namida** *Namida (name)*
Syllable 1 (P1)	**na*mi** *Nami (name)*	**na*mina** *Namina (name)*
Syllable 2 (P2)	**nami*** *wave*	**nana*me/nama*mi** *slanted/living body*
Syllable 3 (P3)		**nigami*** *bitterness*

Table 2: *Sample stimuli sentences*

Accent Position	Phrasal Position	Test Sentence
Syllable 1 (P1)	Phrase-medial (PhM)	[honto: ni **na*mi** makasita?]$_{IP}$ *really Nami defeated?* 'Really (you) defeated Nami?'
	Phrase-final (PhF)	[honto: ni **na*mi**?]$_{IP}$ [makasita?]$_{IP}$ *really Nami? defeated?* 'Really Nami? (You) defeated (her)?'

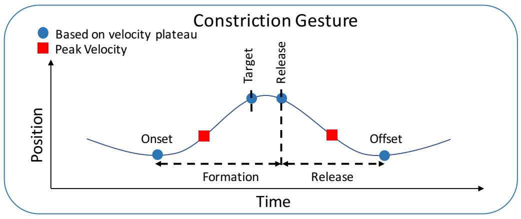

Figure 1: *Semi-automatic labeling of kinematic landmarks in consonant (C) gestures using custom software (Mark Tiede, Haskins Laboratories).*

2.3. Data analysis

Kinematic labels for all consonants (C) were made using a semi-automatic procedure that identifies constriction gestures (Mark Tiede, Haskins Laboratories). Labial C gestures (/m/) were labeled on the lip aperture tract, coronal C gestures (/t, d, n/) on the tongue tip vertical displacement tract and dorsal C gestures (/g/) on the tongue dorsum vertical displacement tract. The labeling procedure detected the following time points based on velocity criteria: onset, peak velocity, target, and release of the gesture's formation, and peak velocity and offset of the gesture's release (shown in figure 1). These labels were used to calculate the following measures: formation duration, release duration, and normalized peak velocity. Formation (F) duration was defined as the interval between the onset and release of the gesture, and release (R) duration as the interval between the release and the offset of the gesture. Peak velocity was normalized by speaker and constriction location using z-scores.

Separate linear mixed effects models with formation duration, release duration, or normalized peak velocity as response variables were fitted for each test C constriction (C1, C2 and C3 for the first, second and third consonant of the test words respectively) using the lmerTest package [32] in R. Random effects of speaker and fixed effects of phrasal position and pitch accent position were included. The relevel function in R was used to derive pairwise comparisons from the lmer output. The second evaluation of the factor pitch accent was compensated for using a Bonferroni correction.

3. Results

3.1. Formation and release duration

As Table 3 shows, significant main effects of phrase position and pitch accent, as well as interaction effects between these two factors were detected in the word-final syllable. These held for both two- and three-syllable words. Pairwise comparisons, visualized in Figure 2, revealed that the formation phase of the onset C gesture of the word-final syllable was

Table 3: *Summary of main and interaction effects for phrasal position and pitch accent from linear mixed effects models for each consonant constriction gesture in two- and three-syllable word stimuli.*

Two-Syllable Words	**C1_F**	**C1_R**	**C2_F**	**C2_R**		
Boundary	*n.s.*	*n.s.*	$\chi^2(1) = 116.2$ $p<0.001$	$\chi^2(1) = 112.99$ $p<0.001$		
Pitch Accent	*n.s.*	*n.s.*	$\chi^2(2) = 10.83$ $p<0.01$	$\chi^2(2) = 71.21$ $p<0.001$		
Boundary*Pitch Accent	*n.s.*	*n.s.*	$\chi^2(2) = 30.64$ $p<0.001$	$\chi^2(2) = 52.82$ $p<0.001$		
Three-Syllable Words	**C1_F**	**C1_R**	**C2_F**	**C2_R**	**C3_F**	**C3_R**
Boundary	*n.s.*	*n.s.*	*n.s.*	*n.s.*	$\chi^2(1) = 62.61$ $p<0.001$	$\chi^2(1) = 67.99$ $p<0.001$
Pitch Accent	$\chi^2(3) = 61.44$, p=0.05	$\chi^2(3) = 66.12$, $p<0.001$	$\chi^2(3) = 123.22$ $p<0.001$	$\chi^2(3) = 207.14$ $p<0.001$	$\chi^2(3) = 101.28$ $p<0.001$	$\chi^2(3) = 76.97$ $p<0.001$
Boundary*Pitch Accent	*n.s.*	*n.s.*	*n.s.*	*n.s.*	$\chi^2(3) = 7.72$ p=0.05	$\chi^2(3) = 21.53$ $p<0.001$

longer in phrase-final words than their phrase-medial counterparts (C2F: β=15.16, SE=1.25, $p<0.001$; C3F: β=13.60, SE=1.86, $p<0.001$).

In three-syllable words only, effects of pitch accent were found on word-final C formations (C3F), where accented final syllables (P3) were longer than any other accent position, regardless of phrasal position (β_{P0}=-19.24, SE=2.9, β_{P1}=-26.87.58, SE=2.83, β_{P2}=-16.88, SE=2.27, $p<0.001$). However, a similar effect of pitch accent on duration was not found in two-syllable words, suggesting that other rhythmic reasons possibly involving polysyllabic shortening may also be relevant.

With respect to the interaction of boundary and pitch accent position, phrase-medially, word-final C formations in two-syllable words (C2F) were longer when unaccented (P0) than accented (β_{P1}=-11.45, SE=2.06, β_{P2}=-11.21, SE=1.96, $p<0.001$). In phrase-medial three-syllable words, the word-final C formation (C3F) was shortest in P1 (β_{P0}=13.06, SE=3.79, β_{P2}=14.21, SE=3.1, β_{P3}=32.58, SE=3.55, $p<0.01$), and longest in P3 (β_{P0}=-19.51, SE=3.65, β_{P1}=-32.58, SE=3.55, β_{P2}=-18.36, SE=2.89, $p<0.001$). Also, in the formation of the final consonant (C2F) in two-syllable words, there was more boundary-related lengthening in accented words (P1, P2) than in unaccented words (P0), possibly related to C2F being shorter in accented words in the phrase-medial condition (β_{P1}=-11.45, SE=2.06, β_{P2}=-11.21, SE=1.96, $p<0.001$).

Turning to C release durations of the word-final syllables, these were longer phrase-finally as compared to phrase-medially, except in unaccented words, in which case no lengthening was detected (two-syllable: β_{P1}=38.55, SE=3.32, $p<0.001$, β_{P2}=32.63, SE=3.13, $p<0.001$; three-syllable: β_{P1}=17.72, SE=4.69, $p<0.001$, β_{P2}=26.33, SE=2.9, $p<0.001$). Similarly, there was no lengthening on the release of the final C gesture in three-syllable words with accent on the final syllable (P3). This latter exception is presumably due to the interaction with pitch accent position in C3, where C release durations in the final-accented (P3) condition were also longer phrase-medially (β_{P0}=-29.38, SE=4.48, $p<0.001$, β_{P1}=-29.47, SE=4.35, $p<0.001$, β_{P2}=-27.22, SE=3.55, $p<0.001$). Phrase-finally, release phases for C3R saw P2/P3 and P0/P1 patterning together, with P2 and P3 being longer than P0 and P1 (P2: β_{P0}=-20.22, SE=4.05, $p<0.001$, β_{P1}=-10.87, SE=4.03, $p<0.05$, P3: β_{P0}=-26.53, SE=4.48, $p<0.001$, β_{P1}=-17.18, SE=4.47, $p<0.001$). It is not immediately clear why these conditions

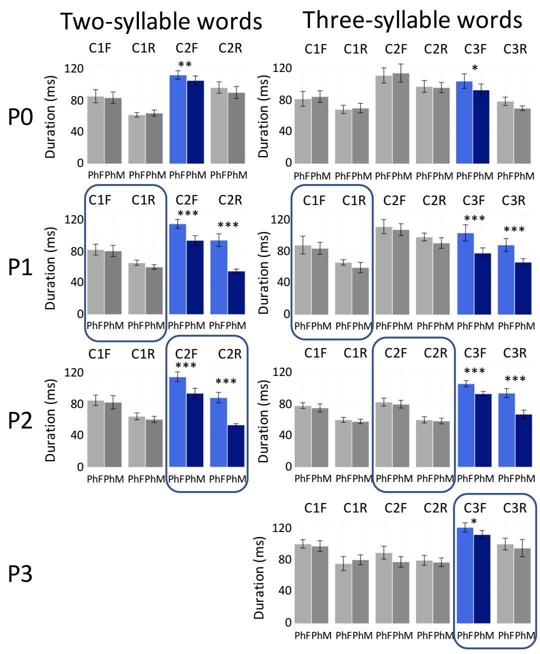

Figure 2: *Duration (in ms; with standard error) of formation (F) and release (R) of test C gestures by boundary (PhM, PhF) and accent (P0, P1, P2, P3). Syllables with pitch accent are boxed. Significance values: *** p<0.001, ** p<0.01, * p<0.05.*

pattern in this way for three-syllable words, which raises the possibility of articulator- or word-specific effects.

In three-syllable words, effects of pitch accent on duration were also seen on the C constrictions of non-final syllables. Pairwise comparisons revealed some significant patterns across pitch accent conditions (all $p<0.05$). In the formation gesture of C1, medial-accented words (P2) were shortest, and final-accented words (P3) were longest. In the formation gesture of C2, accented syllables (P2) and words with accent on the following syllable (P3) were shorter than unaccented (P0) and words with preceding accent (P1). The release gesture of C2 also had accented syllables (P2) being shortest, and unaccented (P0) and initial-accented words (P1) being longest.

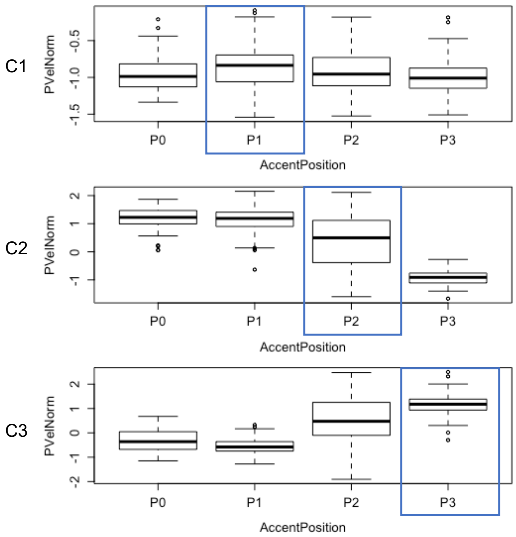

Figure 3: *Normalized peak velocity (PVelNorm) of test C gestures per pitch accent position (P0, P1, P2, P3) in three-syllable words. Boxes indicate the accented syllable.*

3.2. Normalized peak velocity

Two-syllable words showed a main effect of boundary in C2 ($\chi^2(1) = 7.12$, $p<0.01$), and slower C gestures phrase-finally in accented conditions. However, after using a Bonferroni correction, there was no significant difference between phrasal positions. Three-syllable words also unexpectedly showed no main effects of boundary. In three-syllable words only, there was a main effect of pitch accent in all accented consonants (C1: $\chi^2(3) = 7.77$, p=0.05, C2: $\chi^2(3) = 188.98$, $p<0.001$, C3: $\chi^2(3) = 199.71$, $p<0.001$). Pairwise comparisons did not reveal any significant contrasts for C1. In C2, final-accented words (P3) were the slowest, medial-accented (P2) was in the middle, and unaccented (P0) and initial-accented (P1) were fastest ($p<0.0001$). In C3, the effect was reversed, with final-accented (P3) being the fastest, medial-accented (P2) again in the middle, and unaccented (P0) and initial-accented (P1) being the slowest (p<0.0001) (see Figure 3). Again, this unsystematic patterning of accent conditions that only appears in three-syllable words suggests the possibility of rhythmic, or articulator- or word-specific effects. Since current data are limited to one or two words per accent condition, we cannot separate directly address this latter issue, although analysis of additional data is underway. Most interestingly, together with the duration results of three-syllable words, C3 of the final-accented (P3) condition is found to be both the fastest and the longest gesture, which is characteristic of phrasal pitch accent in stress languages. No interaction effects between phrasal position and pitch accent position were found in two- or three-syllable words.

4. Discussion

Across two- and three-syllable word experiments, we found that boundary lengthening begins in the formation gesture of the final consonant. Furthermore, presence, but not position, of lexical pitch accent affects the scope and amount of boundary lengthening in Japanese. In words with accent on the first or second syllable, there was a greater amount of boundary-related lengthening in the formation of the onset C gesture of the phrase-final syllable, and the effect also extended further rightward to the release phase of that consonant. While final-accented three-syllable words did not show boundary-related lengthening on the final C's release (C3R), this could have been due to the concurrent accent-related lengthening since these gestures were longest when accented (P3), regardless of their phrasal position. However, duration is not considered a correlate of lexical pitch accent in Japanese [27, 28, 29, 30, 31], and our results do not show a robust correlation, as two-syllable words did not show a similar effect. Thus, other rhythmic considerations may be needed, such as word length and polysyllabic shortening, or perhaps even the relative rarity of final-accented three-syllable nouns in Tokyo Japanese (2% final-accented compared to 52% unaccented and 42% initial-accented) [24]. There may also be other phonological pressures imposing limits on lengthening, since Japanese has length distinctions on consonants and vowels (see similar discussion in e.g., [6, 33]). Nevertheless, it is interesting to note that the kinematic profile of word-final accented syllables (P3) in three-syllable words is consistent with what we expect from results on phrasal pitch accent from stress languages.

In general, boundary-related lengthening in Tokyo Japanese extends over a continuous interval, similar to [7] and unlike [8]. Contrary to previous work that locates the effect in the rhyme of the phrase-final syllable [14, 15, 21], our data indicate that in Tokyo Japanese phrase-final lengthening also affects the onset of the phrase-final syllable. The effect possibly reaches the vowel of the penultimate syllable, as in [25, 26]. However, further analysis is needed to confirm this, since our current report does not include vowels.

How pitch accent position affects the scope of boundary lengthening in Japanese seems to differ from the interactions found in English and Greek. Unlike English, the lexically prominent syllable does not systematically undergo lengthening as a function of boundary [8]. Unlike Greek, the scope of lengthening does not expand towards the prominent syllable [16]. The fact that Japanese differs from patterns in these stress languages may be due to typological differences; further research on pitch accent languages is needed to clarify whether this difference may be attributed to the different types of lexical prominence systems. What does remain consistent with English and Greek is that lexical prominence in Japanese is found to interact with the scope of phrase-final lengthening.

In terms of velocity, the analysis of normalized peak velocity detected no slowing on any of the C gestures composing the phrase-final words and no interaction with pitch accent. This finding was unexpected (cf. [1, 17], and challenges the π-gesture model, which predicts that phrasal boundaries induce slower movements [11, 22].

5. Conclusion

In conclusion, this study has examined lexical pitch accent and its role in the phonetic realization of prosodic structure in Japanese. We have examined kinematic dimensions of duration and velocity, showing intricate interactions between pitch accent and boundary events.

6. Acknowledgements

This work was supported by NSF Grant #1551428. We would like to thank everyone who participated in our experiments, as well as the UCSB Phonetics Circle for their valuable feedback.

7. References

[1] T. Cho, "Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English," *Laboratory phonology*, vol. 8, pp. 519–548, 2006.

[2] D. K. Oller, "The effect of position in utterance on speech segment duration in English," *The Journal of the Acoustical Society of America*, vol. 54, no. 5, pp. 1235–1247, 1973.

[3] R. Berkovits, "Progressive utterance-final lengthening in syllables with final fricatives," *Language and Speech*, vol. 36, no. 1, pp. 89–98, 1993.

[4] R. Berkovits, "Utterance-final lengthening and the duration of final-stop closures," *Journal of Phonetics*, vol. 21, no. 4, pp. 479–489, 1993.

[5] R. Berkovits, "Durational effects in final lengthening, gapping, and contrastive stress," *Language and Speech*, vol. 37, no. 3, pp. 237–250, 1994.

[6] T. Cambier-Langeveld, "The domain of final lengthening in the production of Dutch," *Linguistics in the Netherlands*, vol. 14, no. 1, pp. 13–24, 1997.

[7] L. White, "English speech timing: a domain and locus approach," Ph.D. dissertation, University of Edinburgh, 2002.

[8] A. E. Turk and S. Shattuck-Hufnagel, "Multiple targets of phrase-final lengthening in American English words," *Journal of Phonetics*, vol. 35, no. 4, pp. 445–472, 2007.

[9] J. Edwards, M. E. Beckman, and J. Fletcher, "The articulatory kinematics of final lengthening," *the Journal of the Acoustical Society of America*, vol. 89, no. 1, pp. 369–382, 1991.

[10] M. E. Beckman and J. Edwards, "Intonational categories and the articulatory control of duration," *Speech perception, production, and linguistic structure*, pp. 359–375, 1992.

[11] D. Byrd and E. Saltzman, "Intragestural dynamics of multiple prosodic boundaries," *Journal of Phonetics*, vol. 26, no. 2, pp. 173–199, 1998.

[12] D. Byrd, "Articulatory vowel lengthening and coordination at phrasal junctures," *Phonetica*, vol. 57, no. 1, pp. 3–16, 2000.

[13] D. Byrd, S. Lee, D. Riggs, and J. Adams, "Interacting effects of syllable and phrase position on consonant articulation," *The Journal of the Acoustical Society of America*, vol. 118, no. 6, pp. 3860–3873, 2005.

[14] D. Byrd, J. Krivokapić, and S. Lee, "How far, how long: On the temporal scope of prosodic boundary effects," *The Journal of the Acoustical Society of America*, vol. 120, no. 3, pp. 1589–1599, 2006.

[15] J. Krivokapic, *The planning, production, and perception of prosodic structure*. University of Southern California, 2007.

[16] A. Katsika, "The role of prominence in determining the scope of boundary-related lengthening in Greek," *Journal of Phonetics*, vol. 55, pp. 149–181, 2016.

[17] S. Kim, J. Jang, and T. Cho, "Articulatory characteristics of preboundary lengthening in interaction with prominence on trisyllabic words in American English," *The Journal of the Acoustical Society of America*, vol. 142, no. 4, pp. EL362–EL368, 2017.

[18] C. W. Wightman, S. Shattuck-Hufnagel, M. Ostendorf, and P. J. Price, "Segmental durations in the vicinity of prosodic phrase boundaries," *The Journal of the Acoustical Society of America*, vol. 91, no. 3, pp. 1707–1717, 1992.

[19] D. Krull, "Prepausal lengthening in Estonian: Evidence from conversational speech," in *Estonian prosody: Papers from a symposium*. Institute of Estonian Language and Authors Tallinn, 1997, pp. 136–148.

[20] L. H. Nakatani, K. D. O'Connor, and C. H. Aston, "Prosodic aspects of American English speech rhythm," *Phonetica*, vol. 38, no. 1-3, pp. 84–105, 1981.

[21] D. Byrd and D. Riggs, "Locality interactions with prominence in determining the scope of phrasal lengthening," *Journal of the International Phonetic Association*, vol. 38, no. 2, pp. 187–202, 2008.

[22] D. Byrd and E. Saltzman, "The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening," *Journal of Phonetics*, vol. 31, no. 2, pp. 149–180, 2003.

[23] Y. Sugiyama, "Acoustic correlates of Japanese pitch accent," *The Journal of the Acoustical Society of America*, vol. 118, no. 3, pp. 1902–1902, 2005.

[24] H. Kubozono, "Japanese pitch accent," *The Blackwell Companion to Phonology*, pp. 1–29, 2011.

[25] J. Seo, S. Kim, H. Kubozono, and T. Cho, "Interaction between rhythmic structure and preboundary lengthening in Japanese," in *Proceedings of the 19th International Congress of Phonetic Sciences*, M. T. . P. W. Sasha Calhoun, Paola Escudero, Ed., Aug. 2019, pp. 637–641.

[26] J. Seo, S. Kim, H. Kubozono, and T. Cho, "Preboundary lengthening in Japanese: To what extent do lexical pitch accent and moraic structure matter?" *The Journal of the Acoustical Society of America*, vol. 146, no. 3, pp. 1817–1823, 2019.

[27] R. S. Weitzman, *Word accent in Japanese*. Management Information Services, 1970, vol. 9.

[28] W. J. Poser, "The phonetics and phonology of tone and intonation in Japanese," Ph.D. dissertation, Massachusetts Institute of Technology, 1984.

[29] M. E. Beckman, *Stress and non-stress accent*. Walter de Gruyter, 1986, vol. 7.

[30] J. Pierrehumbert and M. Beckman, "Japanese tone structure," 1988.

[31] M. Sugito, "Nihongo onsei no kenkyu 5: "hana" to "hana" [Studies on Japanese Phonetics 5]," *Osaka: Izumi Shoin*, 1998.

[32] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest package: tests in linear mixed effects models," *Journal of Statistical Software*, vol. 82, no. 13, 2017.

[33] A. Cutler and T. Otake, "Pitch accent in spoken-word recognition in Japanese," *The Journal of the Acoustical Society of America*, vol. 105, no. 3, pp. 1877–1888, 1999.