



Data Article

Data on preboundary lengthening in Tokyo Japanese as a function of prosodic prominence, boundary, lexical pitch accent and moraic structure



Jungyun Seo^a, Sahyang Kim^b, Taehong Cho^{a,*}

^a *Hanyang Institute for Phonetics and Cognitive Sciences of Language, Department of English Language and Literature, Hanyang University, Seoul, Korea*

^b *Department of English Education, Hongik University, Seoul, Korea*

ARTICLE INFO

Article history:

Received 26 September 2020

Revised 23 February 2021

Accepted 25 February 2021

Available online 2 March 2021

Keywords:

Preboundary lengthening

Phrase final lengthening

Tokyo Japanese

Prosodic boundary

Focus

Prominence

Lexical pitch accent

Mora

ABSTRACT

This article provides individual speakers' acoustic durational data on preboundary (phrase-final) lengthening in Japanese. The data are based on speech recorded from fourteen native speakers of Tokyo Japanese in a laboratory setting. Each speaker produced Japanese disyllabic words with four different moraic structures (CVCV, CVCVN, CVNCV, and CVNCVN, where C stands for a non-nasal onset consonant, V for a vowel, and N for a moraic nasal coda) and two pitch accent patterns (initially-accented and unaccented). The target words were produced in carrier sentences in which they were placed in two different prosodic boundary conditions (Intonational Phrase-final ('IPF') and Intonational Phrase-medial ('IPm')) and two focus contexts (focused and unfocused). The measured raw values of acoustic duration of each segment in different conditions are included in a CSV-formatted file. Another CSV-formatted file is provided with numeric calculations in both absolute and relative terms that exhibit the magnitude of preboundary lengthening across different prominence contexts (focused/unfocused and initially-accented/unaccented). The absolute durational difference was obtained as a numeric increase of preboundary lengthening of each segment produced in phrase-final

* Corresponding author.

E-mail address: tcho@hanyang.ac.kr (T. Cho).

position versus phrase-medial position (i.e., $\Delta(\text{IPf-IPm})$ where 'f' = 'final' and 'm' = 'medial'). The relative durational difference was obtained as a percentage increase of preboundary lengthening in IP-final position versus IP-medial position, which was calculated by the absolute durational difference divided by the duration of the segment in phrase-medial position and then multiplied by 100 (i.e., $(\text{Absolute difference}/\text{IPm}) \times 100$). This article also provides figures that exemplify speaker variation in terms of absolute and relative differences of preboundary lengthening as a function of pitch accent. Some theoretical aspects of the data are discussed in the full-length article entitled "Preboundary lengthening in Japanese: To what extent do lexical pitch accent and moraic structure matter?" [1].

© 2021 The Authors. Published by Elsevier Inc.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Specifications Table

Subject	Linguistics
Specific subject area	Phonetics
Type of data	Table Figure CSV file (Spreadsheet)
How data were acquired	Acoustic measurements based on speech recorded in a laboratory setting
Data format	Raw
Parameters for data collection	The acoustic duration of each segment of disyllabic target words. Experimental factors: boundary (presence vs. absence of an Intentional Phrase boundary after the target word); focus (focused vs. unfocused); moraic structure (CVCV, CVCVN, CVNCV, CVNCVN, where 'C' = a non-nasal onset consonant; 'V' = a vowel; 'N' = a nasal coda); and lexical pitch accent (initially-accented vs. unaccented).
Description of data collection	Preparation of the data involved acquisition of acoustic data via speech recording and measurement of each segment's duration
Data source location	Hanyang University, Seoul, Korea
Data accessibility	With the article Repository name: Mendeley Data Seo, Jungyun; Kim, Sahyang; Cho, Taehong (2021). "Japanese preboundary lengthening for Data in Brief", Mendeley Data, V1, https://doi.org/10.17632/ht52gbr4gk.1
Related research article	Seo, J., Kim, S., Kubozono, H., & Cho, T. (2019). Preboundary lengthening in Japanese: To what extent do lexical pitch accent and moraic structure matter?, <i>The Journal of the Acoustical Society of America</i> , 146(3), 1817–1823.

Value of the Data

- The data provided in two CSV-formatted files contain fourteen Tokyo Japanese speakers' individual acoustic measurement data on preboundary lengthening, which can be used by other researchers to explore various aspects of preboundary lengthening in Tokyo Japanese.
- In particular, the data can be used for various statistical analyses to examine speaker variation in the phonetic realization of Japanese preboundary lengthening in relation to the moraic structure, pitch accent patterns, and prosodic prominence. The data were obtained from 7 female and 7 male speakers, which would allow researchers to examine gender-related differences.

- The data can be used to examine cross-dialectal similarities and differences, and to compare the first vs second language production, taking into account the influence of linguistic structure on the realization of preboundary lengthening.

1. Data Description

The data presented in this article contain measured acoustic durational values that can be used to examine fourteen Tokyo Japanese speakers' individual patterns of the acoustic realization of preboundary lengthening (PBL) in disyllabic words produced in various phonological and prosodic structures (such as lexical pitch accent, moraic structure and focus-induced prominence) in relation to [1].

1.1. Data files: Acoustic duration of preboundary lengthening as produced by individual speakers

Two CSV files are available in the data repository (<https://doi.org/10.17632/ht52gbr4gk.1>). One file contains measured acoustic duration (in millisecond) of each segment within target words produced by individual speakers as exemplified in Table 1. The first column of the spreadsheet indicates a subject ID number (S01-S14). The second column includes information about two disyllabic target word sets (the TAKA set and the SAKE/O set; for the list of words in each set, see Table 3). Two word sets differ in terms of segmental contexts. The 'Moraic structure' column contains information about the moraic structure of target words. Note that a vowel (V) and a nasal consonant (N) each count as one mora in Japanese. There are four types of moraic structure: CV.CV (where '.' refers to a syllable boundary) with 2 moras, CV.CVN with 3 moras (one mora in the first syllable and two moras in the second syllable), CVN.CV with 3 moras (two moras in the first syllable and one mora in the second syllable), and CVN.CVN with 4 moras. The subsequent three columns are related to prosodic contexts. The 'Pitch accent' column indicates the presence of an initial pitch accent (Initially-accented, or 'ia') or lack thereof (Unaccented, or 'ua') at the lexical level. The 'Focus' column specifies whether the target word receives focus-induced prominence or not (Focused vs. Unfocused) when produced within an utterance. The 'Boundary' column indicates whether the target word is produced in the Intonational Phrase-final position (IP-final) or in the Intonational Phrase-medial position (IP-medial). The 'Rep' column indicates the number of repetitions during the recording (ranging from 1–6 repetitions, represented as r1-r6 in the file). The remaining columns represent the duration of each segment within a target word: C stands for a non-nasal onset consonant, V for a vowel, N for a nasal coda. The numbers 1 and 2 after C, V, and N indicate the ordinal number of syllable within a target word such that C1 refers to the onset consonant in the first syllable and C2 refers to the onset consonant of the second syllable.

The second CSV file contains the derived values from the raw data in the above-mentioned CSV file. The data in the second CSV file allows one to observe the magnitude of preboundary lengthening of each segment (i.e., C1, V1, N, etc.) in both absolute and relative terms. Individual speakers' mean durational values of each segment were calculated by averaging durational values across repetitions in the IP-final and IP-medial positions within each word, mora, pitch accent, and focus conditions. These mean values are presented in the IP-final and IP-medial columns as in Table 2. The 'Absolute' column shows the numeric increase of preboundary lengthening calculated by the absolute durational difference between a segment in phrase-final position and that in phrase-medial position (i.e., $\Delta(\text{IPf-IPm})$ where 'IPf' = Intonational Phrase final, and 'IPm' = Intonational Phrase medial). The 'Percentile' column shows the relative durational difference, which is the percentage increase of preboundary lengthening from IP-medial to IP-final position. This was calculated by the absolute durational difference between IP-medial and IP-final conditions divided by the duration of the segment in phrase-medial position and then multiplied by 100 (i.e., $(\text{Absolute difference/IPm}) \times 100$).

Table 1

Part of the CSV file (raw_segmental_duration.csv) to illustrate the organization of the file with respect to experimental conditions. This sample contains the acoustic duration (in ms) of each segment from speaker S01 producing the words in the "TAKA" word set in various conditions of moraic structure, pitch accent pattern, focus and boundary (see the text for details). ' μ ' indicates the number of mora within a word, and 'r' indicates the number of repetitions during the data recording (r1-r6). C, V and N indicates a non-nasal consonant, vowel and a nasal consonant, respectively. The numbers next to the C, V, and N in the header indicates the ordinal number of syllable (i.e., the first or second syllable) to which a segment belong.

ID	Word set	Moraicstructure	Pitch accent	Focus	Boundary	Rep	C1	V1	N1	C2	V2	N2						
S01	TAKA	CV.CV (2 μ)	Initially- accented (ia)	Focused	IP-final	r1	47.93	71.04		92.4	94.32							
					IP-medial	r1	47.41	83.3	71.17	73.29								
				Unfocused	IP-final	r1	43.52	72.92	88.97	114.29								
			Un-accented (ua)	Focused	IP-medial	r1	37.96	74.76	65.78	59.48								
					IP-final	r1	36.51	80.84	81.77	136.11								
				Unfocused	IP-medial	r1	49.96	70.77	64.76	76.49								
		S01	TAKA	CV.CVN (3 μ)	Initially- accented (ia)	Focused	IP-final	r1	51.53	73.45	78.52	100.75	111.76					
							IP-medial	r1	44.06	69.08	78.43	99.17	73.94					
						Unfocused	IP-final	r1	49.55	73.75	72.04	90.77	152.98					
					Un-accented (ua)	Focused	IP-medial	r1	68	69.81	63.33	96.07	87.84					
							IP-final	r1	61.15	72.25	80.37	108.15	119.31					
						Unfocused	IP-medial	r1	49.78	71.39	67.71	98.21	77.86					
				S01	TAKA	CVN.CV (3 μ)	Initially- accented (ia)	Focused	IP-final	r1	51.93	94.84	70.9	95.96	110.72			
									IP-medial	r1	48.94	100.99	58.79	56.79	71.17			
								Unfocused	IP-final	r1	53.96	104.11	74.85	112.5	129.62			
							Un-accented (ua)	Focused	IP-medial	r1	45.63	90.26	65.23	63.31	58.57			
									IP-final	r1	68.47	96.64	73.66	100.7	171.5			
								Unfocused	IP-medial	r1	45.13	113.26	72.1	71.84	68.05			
						S01	TAKA	CVN.CVN (4 μ)	Initially- accented (ia)	Focused	IP-final	r2	30.06	99	74.24	88.61	171.43	
											IP-medial	r1	54.11	98.97	79.94	42.16	59.52	
										Un-accented (ua)	Focused	IP-final	r1	45.02	101.51	73.63	74.46	97.01
									IP-medial			r1	53.63	100.23	76.49	63.3	97.59	83.52
									Unfocused		IP-final	r1	54	103.23	96.27	52.51	105.94	148.8
									S01	TAKA	CVN.CVN (4 μ)	Initially- accented (ia)	Focused	IP-medial	r1	42.67	96.29	75.61
IP-final	r1							40.16						112.56	69.98	93.33	109.83	114.71
Un-accented (ua)	Focused							IP-medial					r1	36.33	113.36	83.91	54.86	111.19
								IP-final				r1	48.76	115.95	68.87	78.36	140.79	75.98
	Unfocused							IP-medial				r1	49.42	107.69	86.45	49.75	100.08	53

Table 2

Part of a CSV file (absolute_relative_duration.csv) that illustrates the organization of the data with respect to experimental conditions. This table contains the sample data values of segmental duration from speaker S01 producing the word set "TAKA". The file contains the magnitude of PBL in each segment within the target words in various phonological and prosodic contexts (moraic structure, pitch accent and focus). μ indicates a mora. Initially-accented words and unaccented words were marked with 'ia' and 'ua,' respectively. Focused and unfocused conditions were presented with 'Foc' and 'Unf,' respectively.

Speaker	Word set	Moraicstructure	Pitchaccent	Focus	Segment	IP-final	IP-medial	Absolute(IPf-IPm)	Percentile((IPf-IPm)/IPm)*100		
S01	TAKA	CVCV (2 μ)	ia	Foc	C1	46.24	53.91	-7.67	-14.23		
				Unf	C1	39.29	38.66	0.63	1.62		
			ua	Foc	C1	47.66	48.59	-0.92	-1.89		
				Unf	C1	37.6	32.07	5.53	17.25		
			ia	Foc	V1	79.66	75.35	4.32	5.73		
				Unf	V1	83.64	75.37	8.27	10.98		
			ua	Foc	V1	79.6	78.63	0.97	1.23		
				Unf	V1	81.83	78.8	3.04	3.85		
			ia	Foc	C2	79.1	70.44	8.66	12.3		
				Unf	C2	79.17	60.19	18.97	31.52		
			ua	Foc	C2	76.38	58.6	17.78	30.35		
				Unf	C2	76.89	56.67	20.22	35.69		
		ia	Foc	V2	104.08	68.61	35.47	51.7			
			Unf	V2	109.77	68.25	41.52	60.84			
		ua	Foc	V2	139.6	72.71	66.89	92			
			Unf	V2	123.53	64.27	59.26	92.22			
				CVCVN (3 μ)	ia	Foc	C1	45.92	48.33	-2.41	-4.98
						Unf	C1	46.68	48.19	-1.5	-3.12
					ua	Foc	C1	50.26	48.25	2.01	4.16
						Unf	C1	38.15	41.53	-3.38	-8.14
					ia	Foc	V1	78.36	72.07	6.29	8.72
						Unf	V1	79.65	78.59	1.06	1.35
					ua	Foc	V1	79.62	79.41	0.21	0.27
						Unf	V1	77.45	75.25	2.2	2.93
ia	Foc				C2	74.27	73.93	0.35	0.47		
	Unf				C2	73.25	57.28	15.97	27.87		
ua	Foc				C2	69.97	58.35	11.62	19.92		
	Unf				C2	67.76	64.51	3.25	5.04		
ia	Foc				V2	114.34	99.58	14.76	14.82		
	Unf				V2	100.6	102.3	-1.7	-1.66		
ua	Foc				V2	118.22	113.04	5.18	4.58		
	Unf				V2	126.85	97.54	29.31	30.05		
ia	Foc				N2	117.67	71.88	45.79	63.69		
	Unf				N2	112.8	79.7	33.1	41.53		
ua	Foc				N2	115.34	73.08	42.25	57.82		
	Unf				N2	103.25	73.35	29.89	40.75		

(continued on next page)

Table 2 (continued)

Speaker	Word set	Moraicstructure	Pitchaccent	Focus	Segment	IP-final	IP-medial	Absolute(IPf-IPm)	Percentile((IPf-IPm)/IPm)* 100
		CVNVCV (3 μ)	ia	Foc	C1	52.35	49.28	3.08	6.25
				Unf	C1	44.68	45.39	-0.71	-1.56
			ua	Foc	C1	60.42	43.36	17.06	39.35
				Unf	C1	42.18	45.24	-3.05	-6.75
			ia	Foc	V1	104.42	111.14	-6.72	-6.05
				Unf	V1	100.37	99.47	0.89	0.9
			ua	Foc	V1	117.21	107.18	10.03	9.36
				Unf	V1	103.43	100.34	3.1	3.09
			ia	Foc	N1	65.45	59.72	5.73	9.6
				Unf	N1	73.22	60.28	12.94	21.47
			ua	Foc	N1	79.61	71.06	8.54	12.02
				Unf	N1	82.03	68.33	13.69	20.04
			ia	Foc	C2	94.55	65.7	28.85	43.91
				Unf	C2	90.11	68.99	21.12	30.61
			ua	Foc	C2	81.96	71.19	10.78	15.14
				Unf	C2	78.31	57.03	21.29	37.33
			ia	Foc	V2	122.13	74.33	47.8	64.32
				Unf	V2	128.12	66.45	61.67	92.8
			ua	Foc	V2	153.98	76.69	77.29	100.78
				Unf	V2	140.99	68.55	72.44	105.67
		CVNVCV(4 μ)	ia	Foc	C1	51.01	53.9	-2.89	-5.36
				Unf	C1	40.9	45.91	-5.02	-10.92
			ua	Foc	C1	51.99	45.92	6.07	13.22
				Unf	C1	44.37	46.7	-2.33	-4.99
			ia	Foc	V1	106.33	108.05	-1.72	-1.59
				Unf	V1	107.47	103.02	4.45	4.32
			ua	Foc	V1	117.51	111.91	5.59	5
				Unf	V1	110.86	110.53	0.34	0.31
			ia	Foc	N1	77.92	70.65	7.27	10.3
				Unf	N1	84.87	74.77	10.11	13.52
			ua	Foc	N1	77.46	87.18	-9.72	-11.15
				Unf	N1	87.55	85.32	2.23	2.62
			ia	Foc	C2	72.79	71.66	1.13	1.57
				Unf	C2	63.1	69.58	-6.48	-9.31
			ua	Foc	C2	78.75	63.14	15.61	24.72
				Unf	C2	63.01	53.56	9.45	17.65
			ia	Foc	V2	116.89	105.16	11.73	11.16
				Unf	V2	116.8	89.59	27.21	30.37
			ua	Foc	V2	123.81	118.46	5.35	4.52
				Unf	V2	122.24	101.26	20.98	20.72
			ia	Foc	N2	116.71	74.13	42.58	57.44
				Unf	N2	121.21	81.68	39.52	48.38
			ua	Foc	N2	99.13	63.62	35.51	55.81
				Unf	N2	108.92	56.44	52.49	93.01

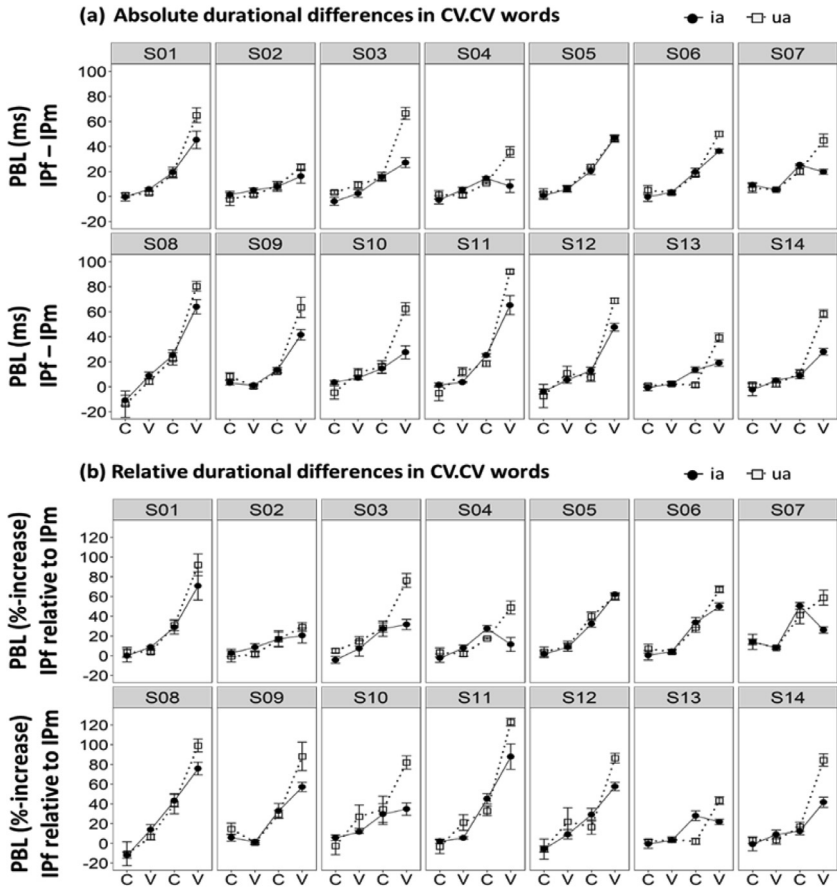


Fig. 1. Line-point plots for CV.CV words showing the magnitude of preboundary lengthening (PBL) across 14 speakers. (a) shows PBL in terms of the absolute increase from the duration (ms) of a segment in the Intonational Phrase medial (IPm) position to that in the Intonational Phrase final (IPf) position. (b) shows PBL in terms of the relative (percent) increase in duration of a segment in the IP-final position relative to its duration in the IP-medial position. Data are plotted as a function of pitch accent pattern: initially accented ('ia') vs. unaccented ('ua').

1.2. Figures: Degree of preboundary lengthening produced by individual speakers in different pitch accent and focus contexts

The following figures (Figs. 1-4) show how each individual speaker's production of preboundary lengthening varies as a function of pitch accent (i.e., initially-accented versus unaccented) in four different moraic structures (i.e., CVCV, CVCVN, CVNVCV, CVNVCVN). See the text in the figure caption for each figure for further explanation of the illustrated data.

2. Experimental Design, Materials and Methods

2.1. Participants

Fourteen speakers of Tokyo Japanese in their 20 s (7 females and 7 males, Mage=24.2 years, range 19-29 years) paid to participate in the recording. Twelve speakers were born and raised in

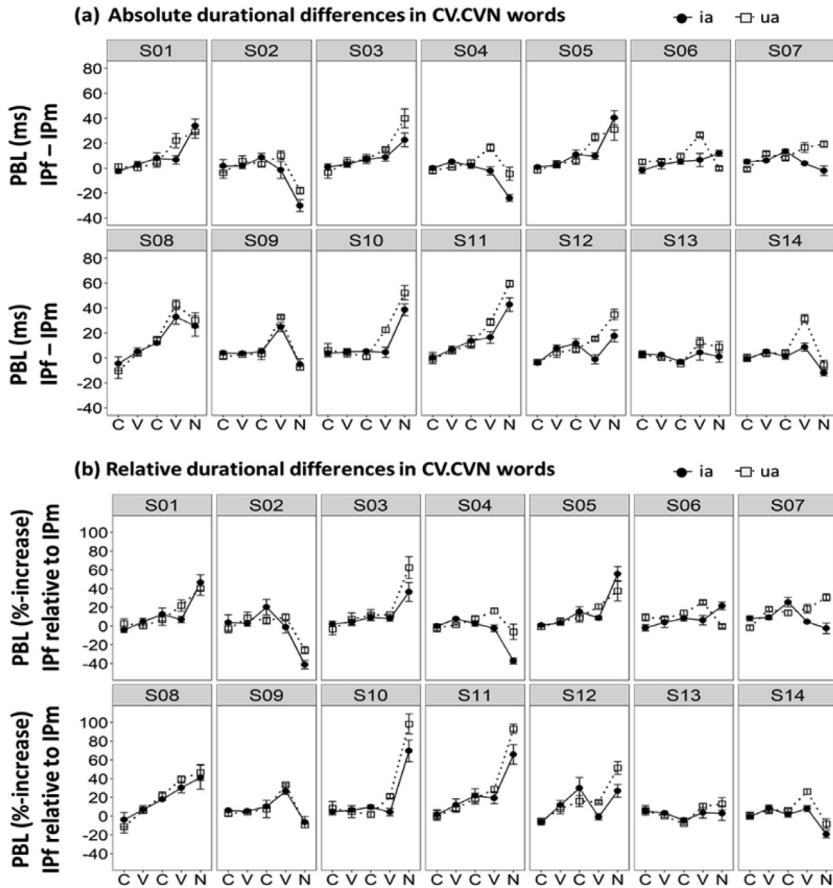


Fig. 2. Line-point plots for CV.CVN words showing the magnitude of preboundary lengthening (PBL) across 14 speakers. (a) shows PBL in terms of the absolute increase from the duration (ms) of a segment in the Intonational Phrase medial (IPm) position to that in the Intonational Phrase final (IPf) position. (b) shows PBL in terms of the relative (percent) increase in duration of a segment in the IP-final position relative to its duration in the IP-medial position. Data are plotted as a function of pitch accent pattern: initially accented ('ia') vs. unaccented ('ua').

Tokyo and two speakers in Kanagawa and Saitama Prefecture (S15 and S11, respectively) where Tokyo Japanese is spoken. All participants were temporary residents in Korea, studying at universities as exchange students, and had resided in Korea for less than three years at the time of recording. The consent was informed and the participants signed a consent form to participate in the research; and the obtained acoustic data were analysed anonymously.

2.2. Speech materials for acoustic recordings

Two sets of target words were used (TAKA set and SAKE/SAKO set) for the recording. Each word set contained eight words with four different moraic structures. There were CV.CV words with two moras (e.g., taka), CVN.CV words with three moras (e.g., tankan), CV.CVN words with three moras (e.g., takan), and CVN.CVN words with four moras (e.g., tankan). In addition, two different lexical pitch accent patterns (unaccented ('ua') vs. initially accented ('ia')) were employed. The two word sets differed in terms of their segmental makeup. The TAKA set contained words

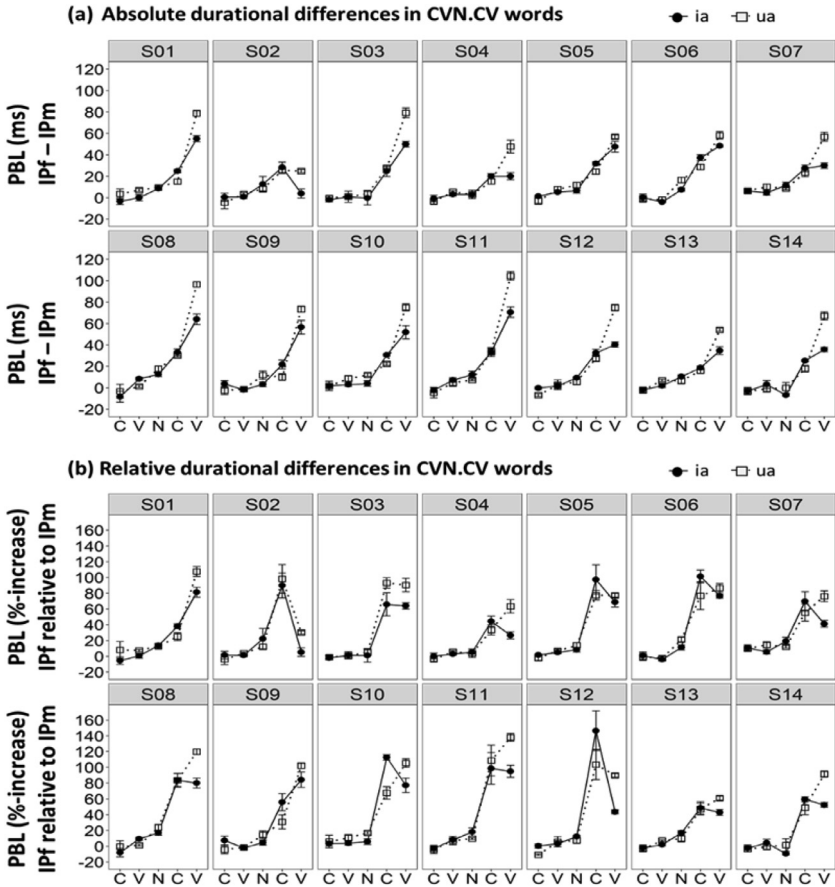


Fig. 3. Line-point plots for CVN.CV words showing the magnitude of preboundary lengthening (PBL) across 14 speakers. (a) shows PBL in terms of the absolute increase from the duration (ms) of a segment in the Intonational Phrase medial (IPm) position to that in the Intonational Phrase final (IPf) position. (b) shows PBL in terms of the relative (percent) increase in duration of a segment in the IP-final position relative to its duration in the IP-medial position. Data are plotted as a function of pitch accent pattern: initially accented ('ia') vs. unaccented ('ua').

with the same segments for unaccented and initially accented conditions. The SAKE/SAKO set, however, had different vowels in the second syllable as it was impossible to find words with exact same segmental makeups and lexical pitch accent patterns. The unaccented words had /e/ and the initially-accented words had /o/. (See [1] for the complete list of words.) The mora structure and the lexical pitch accent patterns were included as experimental factors in order to examine whether and how the degree of preboundary lengthening would vary depending on the number of moras within a syllable and the presence or absence of pitch accent (initially-accented vs. unaccented).

Each target word was produced in a mini dialogue consisting of a question (prompt) sentence and a target-bearing sentence as an answer. The target-bearing sentences were produced in different prosodic contexts of boundary (Intonational Phrase (IP)-final and IP-medial) and focus (focused and unfocused) types. (See [1] for a set of example sentences.)

For the boundary condition, target words were placed either in the Intonational Phrase final position (i.e. IP-final), or in the Intonational Phrase medial position (i.e. IP-medial). The target word in the IP-final condition was always the final word in a sentence, such that speakers

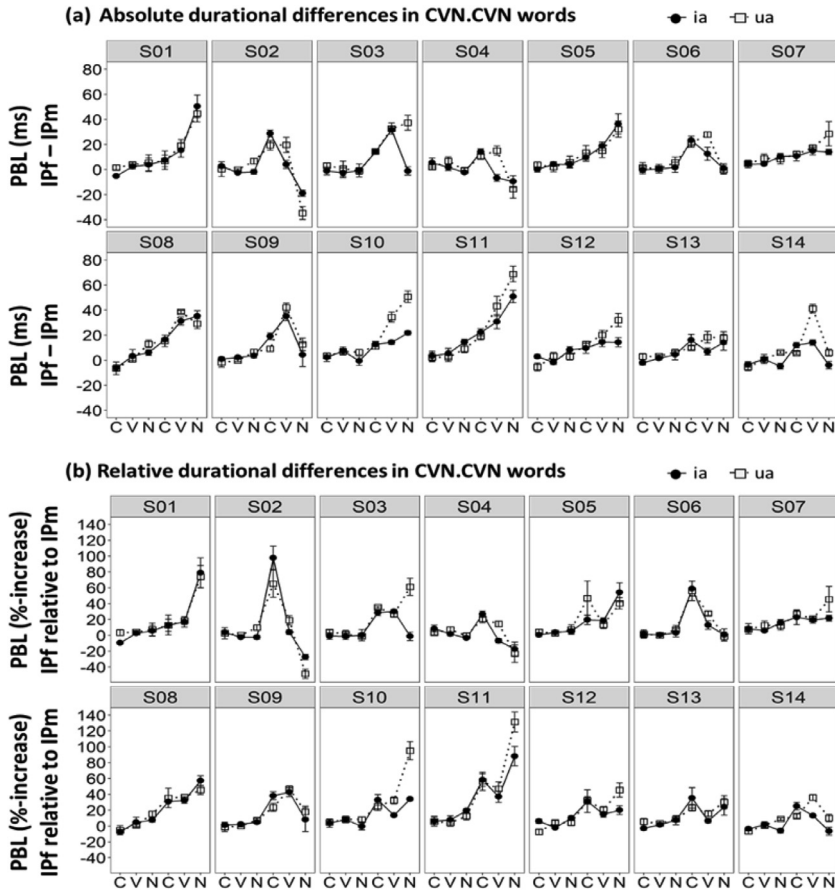


Fig. 4. Line-point plots for CVN.CV N words showing the magnitude of preboundary lengthening (PBL) across 14 speakers. (a) shows PBL in terms of the absolute increase from the duration (ms) of a segment in the Intonational Phrase medial (IPm) position to that in the Intonational Phrase final (IPf) position. (b) shows PBL in terms of the relative (percent) increase in duration of a segment in the IP-final position relative to its duration in the IP-medial position. Data are plotted as a function of pitch accent pattern: initially accented ('ia') vs. unaccented ('ua').

naturally inserted a pause, clearly marking the end of an Intonational Phrase after the target word. Note that in the IP-medial condition, target words were produced with a following particle which was encliticized with the target word, so that no phrase boundary would be inserted after the target word.

For the focus conditions, the target words were either focused or unfocused. The focus factor was included to examine how the potential lexical level prominence effect due to lexical pitch accent might be further modulated by the phrase-level prominence induced by focus. In order to induce the focus-induced prominence on the target word, speakers were asked to correct the wrong information given in bold in the prompt sentence. This guided them to produce the target word with focus, contrasting the word in the answer sentence with a word in the prompt sentence. For the unfocused condition, a word preceding the target word was focused, so that the focus did not fall on the target word.

The data collection took place in a sound-treated room at Hanyang Institute for Phonetics and Cognitive Sciences of Language (HIPCS). The data were recorded with a Tascam HD-P2 digital recorder and a SHURE KSM 44 microphone at a sampling rate of 44 kHz. Before recording the

data, participants went through a pre-training session of about 10 min to familiarize themselves with the words and sentences used in the experiment. Prompt questions were pre-recorded by a female native speaker of Tokyo Japanese. During the data collection, a participant sat in front of a PC, heard a prompt question through a loudspeaker and simultaneously saw the sentence that was visually presented on the monitor. Speakers were asked to listen to the prompt questions and answer them by reading aloud the target-bearing sentences presented on the monitor with the meaning contrast in mind. At the time of recording, the experimenter who was a trained prosody transcriber monitored the production carefully, and asked the participant to read the sentence a few more times when the utterance contained any mispronunciation or hesitation.

Each recording session took about 90–120 min, including two 10-minute breaks. Target-bearing sentences were repeated 6 times per speaker in a pseudo-randomized order. A total of 5376 tokens were collected: 2 target word sets (TAKA vs. SAKE/SAKO), 4 mora structures (CV.CV vs. CV.CVN vs. CVN.CV vs. CVN.CVN), 2 pitch accents (initially accented vs. unaccented), 2 focus conditions (focused vs. unfocused), 2 boundary conditions (IP-final vs. IP-medial), 6 repetitions and 14 speakers. The prosodic information of each utterance (boundary and prominence) was initially examined by the author as a trained Japanese ToBI (Tone and Break Indices) transcriber. Two other trained phoneticians checked if the utterances were produced with intended prosodic renditions in terms of prominence and boundary. As a result, 406 tokens that deviated from the intended prosodic renditions as agreed upon by the transcribers were excluded from further analyses, leaving 4970 tokens for acoustic analyses.

2.3. Measurements

The duration of each segment in the target word was measured by comparing the waveform and spectrogram using Praat [2]. Three types of the durational measures were included as follows:

- Closure duration of the consonant (C1, C2) was measured from the end of F2 of the preceding vowel to the closure release.
- Vowel duration (V1, V2) was measured as an interval from the release of the stop closure to the end of the vowel as marked by F2 in CV (an open syllable) or by the onset of nasal murmur in CVN. Note that we defined vowel duration in such a way that it included VOT as suggested by [3], especially because VOT for consonants in our Japanese data was generally very short and often negligible (see [1] for further discussion on this point.)
- Duration of the nasal coda consonant (N1, N2) was measured from the onset to the offset of the nasal energy (murmur) as indicated by an overall weakening of (especially higher) formants which are reorganized due to nasal zeros and poles displayed on the spectrogram.

Ethics Statement

Informed consent was obtained from speakers who participated in the recording.

CRedit Author Statement

Jungyun Seo: Conceptualization, Methodology, Formal analysis, Investigation, Data Curation, Writing-Original draft preparation, Visualization; **Sahyang Kim:** Conceptualization, Methodology, Investigation, Writing- Review & Editing, Funding acquisition; **Taehong Cho:** Conceptualization, Methodology, Investigation, Writing- Review & Editing, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

Acknowledgments

We thank the Japanese speakers for their participation in the recording. A special thanks goes to Haruo Kubozono for his guidance for constructing speech materials. This work was supported by the Ministry of Education of the Republic of Korea and the [National Research Foundation of Korea](#) (NRF-2018S1A5A2A03036736).

References

- [1] J. Seo, S. Kim, H. Kubozono, T. Cho, Preboundary lengthening in Japanese: to what extent do lexical pitch accent and moraic structure matter? *J. Acoust. Soc. Am.* 146 (2019) 1817–1823, doi:[10.1121/1.5122191](https://doi.org/10.1121/1.5122191).
- [2] P. Boersma, D. Weenink, Praat: Doing Phonetics by Computer [computer program], 2018 <http://www.praat.org/>.
- [3] A. Turk, S. Nakai, M. Sugahara, Acoustic segment durations in prosodic research: a practical guide, in: S. Sudhoff, D. Lenertova, R. Meyer, S. Pappert, P. Augurzkzy, I. Mleinek (Eds.), *Methods in Empirical Prosody Research*, Mouton de Gruyter, Berlin, 2006, pp. 1–28.