

Durational Cues and Pitch Cues in Japanese Mora

Tomoko Kozasa

Department of Linguistics
University of Hawai‘i at Manoa
kozasa@hawaii.edu

Abstract

The goal of this paper is to examine the functions of durational cues and pitch cues in Japanese long vowels and their role in the organization of prosody.

There are two dimensions in the prosodic organization of speech. One is the quantitative or temporal organization, and the other is the qualitative organization, which involves the accentuation/stress system of the language.

The categories of quantitative dimension of prosody at the prosodic-word level used to classify languages are mora-syllable-, and stress-/foot- timed. Isochronous distribution of these units in a language has been challenged for several decades; however, the temporal dimension of speech prosody can be measured in phonetic terms as duration. The categories of qualitative dimension of prosody at the prosodic-word level used to classify languages are pitch-accent, tone, and stress. They are captured as the movement of pitch, or the movement of fundamental frequency (*F0*) in phonetics.

Temporal and accentual dimensions of prosody are bound together in a complex way; however, they can be analyzed as two distinct phonetic properties: duration and *F0*. In order to fully understand the prosodic organization of a language, we must investigate how these two phonetic signals interact with each other in natural speech.

Two production experiments were conducted. One involved both duration and pitch. The other involved only duration.

The results from these production experiments show that these two phonetic signals influence each other when native speakers of the Tokyo dialect of Japanese produce phonologically long vowels.

1. Introduction

Japanese is a language in which the mora plays an important role in prosody. The mora is generally used in two different ways: as a temporal unit or as a tone-bearing unit. It is used as a means of determining the phonological length of a segment. For instance, vowel length is distinctive in Japanese and long vowels carry two moras. Therefore, the Japanese word *biru* ‘building’ has two moras and with a long vowel *biiru*, which is a totally different word meaning ‘beer’, carries three moras.

The mora, as a temporal unit, has four different manifestations in Japanese. It may be realized as 1) a vowel in a light syllable (CV), 2) the second element of a long vowel or a diphthong (CVV or CV̤), 3) a coda nasal (CVN), and 4) the first half element of a geminate consonant in the coda position (CVC̤=C̤).

In addition, Japanese mora may bear either high (H) or low (L) tone. Japanese is a pitch-accent language; that is, the position and/or presence of pitch-accent changes the lexical

meaning of a word. When a word is accented, there is a HL pitch sequence within the word in the Tokyo dialect. For example, the word *kaki-ga* with a HLL pitch pattern means ‘oyster-nominative (NOM);’ with a LHL pitch pattern, it means ‘fence-NOM,’ and with a LHH pitch pattern, which is unaccented, it means ‘persimmon-NOM.’ An accented word containing a long vowel, such as *paaru* (HLL) ‘pearl’ has a HL sequence within the syllable in the Tokyo dialect [11]. The mora realized only as a vowel or a coda nasal can function as tone/pitch-bearing unit. In this study, I will focus on the mora realized as a vowel, since it occurs frequently in speech and it is therefore easy to find material for experiments.

The relationship between temporal and accentual dimensions of Japanese prosody has not been thoroughly investigated. The majority of phonetic investigation on Japanese prosody involves the durational aspect of Japanese mora [1, 3, 5, 6, 8, 15, 16]. Several studies have been done on the pitch-accent or accent patterns of Japanese [7, 17], on the effects of pitch-accent on prosodic segmentation [13], and on the effects of sound recognition [4]. A few studies have examined the effects of pitch-accent on the duration of vowels [2, 9, 10, 12]. Some report that pitch-accent does not have a significant effect on the duration of Japanese vowels [2, 10]. Others report that pitch-accent does in fact influence the perception of vowel length [9, 12].

This paper reports the results from two production experiments. The results indicate that speakers intensify the distinction between short and long vowels when they produce unaccented long vowels by lengthening the duration, which suggests that pitch-accent does have some effect on the production of long vowels.

2. Experiments

The purpose of the experiments was to investigate how native speakers of the Tokyo dialect use durational and pitch cues in producing Japanese vowels. In order to examine the interaction between the two acoustic cues, two separate production experiments were conducted. For the first experiment, only accented vowels were used and, in the second one, only unaccented vowels were used.

2.1. Participants

Since pitch-accent patterns have so many dialectal variations in Japanese, the participants in the experiments for this study were restricted to the Tokyo dialect speakers. There were four participants (two females and two males) in each experiment. Their age range was from the late 20s to the late 30s. There was a male participant in each of the experiments. All participants claimed that either one or both of their parents were also native speakers of the Tokyo dialect or lived most of their lives in a community where the Tokyo

dialect was spoken. Consequently, the pitch patterns of the participants were not influenced by other dialects.

2.2. Procedure

The procedure followed in both experiments was the same. The participants read the material sentences in a sound attenuated recording studio. Each material sentence was read at three different speech rates: fast, normal, and slow, which was dependent on the judgment of the participant. The participants were instructed not to insert any pause into a sentence, because a pause would affect the duration of utterances and pitch movement. A pause resets prosodic constituents; thus, it may change the pitch movement. The target vowels and the speech rates were randomly distributed. The utterances were tape-recorded on TANBERG TCR522 cassette-recorder through a microphone and digitized by using Pitchworks at a sampling rate of 11,025 Hz.

Measurements were taken from the digitized sound file. The duration of the sentences and words were measured as well as the duration of vowels to ensure that sentences were read at the desired speech rate. For this study, only the data from fast and slow speech were used. The beginning and the end of the vowels were determined by presence of the second (F_2) and higher formants. Pitch fall was the value of the highest F_0 in the target vowel minus the lowest F_0 in the following syllable.

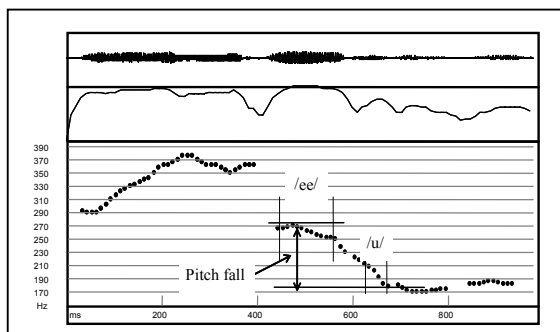


Figure 1: Measurement points of pitch fall:
Kawaii be_{er}u dane 'What a pretty veil.'

2.3. Materials

The pitch-accent pattern of target words was checked against a dictionary of Japanese accent pattern [13]. The material sentences were written in regular Japanese orthography; that is, a mixture of *Kanji*, or Chinese characters and *Hiragana/Katakana*, and Roman letters. The desired speech rate was displayed in a smaller font at the upper left corner of each sentence.

2.3.1. Experiment 1

The target vowels in the first experiment were all accented. A total of 25 minimal pair sentences were used for this experiment (five minimal pair for each of the five vowels in Japanese), such as *Ima* 'aru' (HL) *to iimashita* 'I said "to exist" now' and *Ima* 'aaru' (HHL) *to iimashita* 'I said "R" now,' and *Kawaii beru* (HL) *dane* 'What a cute bell!' and *Kawaii beeru* (HLL) *dane* 'What a pretty veil!'

2.3.2. Experiment 2

The target vowels in the second experiment were all unaccented. A total of 18 minimal or near minimal pair words, such as *sokai* (HHH) 'evacuation' and *sookai* (HHHH) 'general meeting,' and 9 words containing an unaccented long vowel, such as *yuukai* (HHHH) 'kidnap' and *meekyappu* (HHHHL) 'make-up' were used. Since it is rare for Japanese words to have an unaccented long vowel, some long vowels crossed over morpheme boundaries, such as *akaari* (LHHH), 'a red ant,' *aka* 'red' + *ari* 'an ant,' and *keshiin*, *keshi* 'to erase' + *in* 'mark', 'a postmark.' In this experiment, all target words were read in a carrier sentence *Ima* _____ *to iimashita* 'I said _____ now.'

3. Hypotheses

If speech rates have an effect on the duration of vowels, then the duration of vowels should be shorter in fast speech regardless of whether they are short or long, or accented or unaccented. However, if pitch-accent has any effect on the duration of vowels, the results from each experiment will be different from each other.

3.1. Experiment 1

- If each mora carries a tone, a long vowel must be approximately twice as long as a short vowel in both fast and slow speech, as it has a HL pitch sequence within the vowel. In other words, ratio between short and long vowels should not be affected by speech rates.
- If speech rates have an effect on pitch fall, its range would be smaller in fast speech than that in slow speech. However, if there is not enough pitch fall, the vowel will lose pitch-accent. Therefore, speakers will maintain a certain amount of pitch fall to avoid losing the pitch-accent from the target vowel.

3.2. Experiment 2

- Since speakers cannot rely on a pitch cue in unaccented vowels, the relative duration of vowels should not be affected by speech rates. In other words, the duration of long vowels should not be shortened in fast speech. Furthermore, the duration of short vowels should not be lengthened in slow speech, because if speakers lengthen the duration of a short vowel, it simply becomes a long vowel.

4. Results

There are some differences between the results from the two experiments which indicate that pitch-accent had some effect on the duration of vowels.

4.1. Experiment 1

As I have predicted, speech rates had a significant effect on the duration of both short and long vowels (both [$p < .0001$]). That is, the duration of vowels was lengthened in slow speech. Furthermore, long vowels were significantly longer than short vowels in both fast and slow speech. The ratio of mean duration of short vowels versus long vowels was 1:1.78 in fast speech and 1:1.91 in slow speech. Long vowels were almost as twice as long as short vowels in both speech rates.

Table 1: Mean duration of accented vowels

Speech Rate	Short (ms)	Long (ms)	Ratio (Short : Long)
Fast	68.3	121.6	1 : 1.78 [$p < .0001$]
Slow	88.3	168.5	1 : 1.91 [$p < .0001$]
Ratio (Fast : Slow)	1 : 1.29 [$p < .0001$]	1 : 1.39 [$p < .0001$]	

However, speech rates did not have an effect on pitch fall ([$p = .8253$] in short vowels, [$p = .0512$] in long vowels). In other words, speakers maintain a certain amount of pitch fall in both fast and slow speech when they produce accented long vowels. Moreover, the mean value of pitch fall was slightly greater in fast speech than slow speech.

Table 2: Mean pitch fall of accented vowels

Speech Rate	Short (Hz)	Long (Hz)
Fast	58.0	76.0
Slow	56.5	91.4

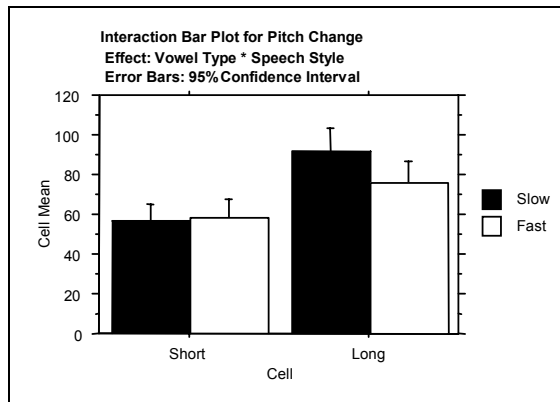


Figure 2: Interaction bar plot for Pitch fall

4.2. Experiment 2

The duration of both short and long unaccented vowels was also significantly affected by speech rates ([$p < .0001$]). The

mean duration of unaccented vowels was shorter than that of accented vowels, except for long vowels in slow speech. However, ratios between short and long vowels were greater than those in accented vowels. In particular, long vowels in slow speech were exceedingly longer.

Table 2: Mean duration of unaccented vowels

Speech Rate	Short (ms)	Long (ms)	Ratio (Short : Long)
Fast	60.8	117.4	1.00 : 1.93 [$p < .0001$]
Slow	83.0	201.9	1.00 : 2.43 [$p < .0001$]
Ratio (Fast : Slow)	1 : 1.37 [$p < .0001$]	1 : 1.72 [$p < .0001$]	

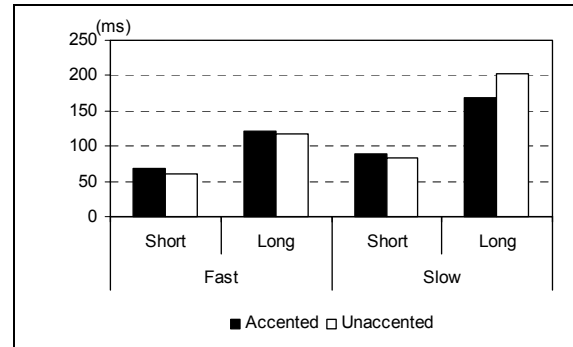


Figure 3: Comparison of vowel durations

5. Discussion

The results from these experiments indicate that duration is sensitive to speech rates. The duration of vowels varies significantly according to speech rates regardless of whether they are accented or unaccented.

We need to account for the excessively long duration of unaccented long vowels in slow speech. This result suggests that speakers lengthen the duration of unaccented long vowels to make sure that they produce long vowels. In contrast, speakers do not lengthen the duration of unaccented short vowels. Since there are only durational cues in unaccented vowels, if speakers lengthen the duration of an unaccented short vowel, it becomes an unaccented long vowel, which could neutralize the contrast between distinct words (*hosoo* 'pavement' vs. *hoosoo* 'broadcasting'). However, the ratio between the short and long of unaccented vowels was greater than that of accented vowels. These results indicate that since speakers have only a durational cue in producing unaccented vowels, they must maintain the durational distinction between short and long vowels.

As for the pitch fall in accented vowels, speech rates do not have an effect on pitch fall. Even in fast speech where the duration of vowels is shortened, speakers maintain a certain amount of pitch fall in the production of accented long vowels. One study [9] suggests that when pitch cue is not present in an accented long vowel, listeners discriminate the vowel length at a chance level. Another perception study [12] also suggests that when a vowel has its pitch fall within the vowel, the vowel tends to be perceived as a long vowel, even though the duration is not as long as a long vowel is

supposed to be. Moreover, speakers highlighted the pitch-accent of short vowels in fast speech by increasing pitch fall. Pitch-accent may not have a significant effect on the intrinsic duration of vowels, as some studies conclude [2, 10]; nevertheless, speakers/listeners seem to rely on the pitch cue to produce and/or determine vowel length.

6. Conclusion

We have examined the results from two production experiments. Native speakers of the Tokyo dialect employ both durational cues and pitch cues to produce Japanese long vowels. When a long vowel is not accented, speakers lengthen the duration of the vowel to ensure that the vowel is long, since speakers rely solely on durational cues. However, when they produce an accented vowel, pitch cues are more stable than durational cues, as they are not affected by speech rates.

Acoustic signals sometimes contain several redundant cues. For instance, the duration of a vowel preceding a voiced stop tends to be longer than a vowel preceding an unvoiced stop, and vowel quality may be slightly altered depending on adjacent consonants. Japanese accented vowels have two cues (durational and pitch), while unaccented vowels have only the durational cue. Speakers use phonetic cues efficiently to produce desired speech sounds. Consequently, when speakers produce accented long vowels in fast speech, they retain the more stable cue (pitch cue), and when they produce unaccented long vowels, they emphasize the only cue (durational cue). Although it is not exactly clear how durational cues and pitch cues interact with each other, it is difficult to deny to the effect of pitch-accent on the duration of vowels in Japanese.

7. Acknowledgment

I would like to thank participants in my experiments and proofreaders of this paper. I also thank Victoria Anderson, Patricia Donegan, Michiko Nakamura, Kenneth Rehg, Amy Schafer, and Naoko Takahashi for their collaboration and feedback. This study is supported by the Department of Linguistics Endowment Fund at the University of Hawai'i at Manoa.

8. References

- [1] Beckman, Mary E., 1982a. Segment duration and the 'Mora' in Japanese. *Phonetica* 39, 113-135.
- [2] Beckman, Mary E., 1982b. Effects of accent on vowel duration in Japanese. *The Journal of the Acoustical Society of America*, 71, S23.
- [3] Bloch, B., 1950. Studies in colloquial Japanese IV: Phonemics. *Language*, 26, 86-125.
- [4] Cutler, Ann and Takeshi Otake, 1999. Pitch accent in spoken-word recognition in Japanese. *The Acoustical Society of America*, 105, 1877-1888.
- [5] Han, M.S., 1962. The feature of duration in Japanese. *Onsei no Kenkyuu*, 10, 65-75.
- [6] Han, M.S., 1994. Acoustic manifestations of mora timing in Japanese. *The Journal of the Acoustical Society of America*, 96, 73-83.
- [7] Haraguchi, Shosuke, 1988. Pitch accent and intonation in Japanese. In *Autosegmental Studies on Pitch Accent*, Harry van der Hulst and Norval Smith (eds.). Dordrecht, Netherlands, 123-150.
- [8] Homma, Y., 1981. Durational relationship between Japanese stops and vowels. *Journal of Phonetics*, 9, 273-381.
- [9] Kozasa, Tomoko, 2002. Duration and F0 Cues for vowel length in Japanese. *The Journal of the Acoustical Society of America*, 111, 2365.
- [10] Larish, Michael D., 1989. Converging sources of evidence again the phonetic reality of the mora in Japanese. *University of Hawai'i Working Papers in Linguistics*, 19 (2), 103-151.
- [11] McCawley, James D., 1968. *The phonological component of a grammar of Japanese*. The Netherlands: Mouton.
- [12] Nagano-Madsen, Yasuko, 1992. Influence of fundamental frequency pattern on the perception of the vowel mora in Japanese. *Phonum*, 1, 31-35.
- [13] Nihon Hoosoo Kyookai, 1966. *Nihongo hatsuon akusento jiten* (Dictionary of Japanese accent pronunciation), Tokyo, Japan: Nihon Hoosoo Shuppan Kyookai.
- [14] Pierrehumbert, Janet and Mary Beckman, 1988. *Japanese Tone Structure*. Cambridge, MA: MIT Press.
- [15] Port, R.F., J. Dalby, and M. O'Dell, 1987. Evidence for mora timing in Japanese. *The Journal of the Acoustical Society of America*, 81, 1574-1585.
- [16] Warner, Natasha and Takayuki Arai, 2001. The role of the mora in the timing of spontaneous Japanese speech. *The Journal of the Acoustical Society of America*, 109, 1144-1156.
- [17] Weitzman, Raymond S., 1970. *Japanese accent: An analysis based on acoustic-phonetic data*. Doctoral dissertation, Los Angeles, CA: University of Southern California.