Deguchi, M. (2015). The role of pitch contours in teaching vowel length distinctions in Japanese. In J. Levis, R. Mohammed, M. Qian & Z. Zhou (Eds). *Proceedings of the 6th Pronunciation in Second Language Learning and Teaching Conference* (ISSN 2380-9566), Santa Barbara, CA (pp. 174-183). Ames, IA: Iowa State University.

THE ROLE OF PITCH CONTOURS IN TEACHING VOWEL LENGTH DISTINCTIONS IN JAPANESE

Masanori Deguchi, Western Washington University

Despite the significant difference in duration, vowel length distinctions have perpetually troubled learners of Japanese. Some studies have found that, while traditionally described as a "quantitative" distinction, the duration itself is not a reliable cue for discerning vowel lengths (e.g., Hirata, 2004). The present study therefore explores the possibility of capitalizing on a "qualitative" distinction, particularly different pitch contours, as a more reliable cue.

INTRODUCTION

Vowel Length

There are five vowels in Tokyo Japanese: /a/, /i/, /u/, /e/, and /o/. All of these vowels contrast in terms of their length, resulting in a total of ten distinct vowel phonemes (Shibatani, 1996). Therefore, words can contrast their meanings on the basis of vowel length alone as illustrated in the minimal pairs in Table 1. For example, *obasan* 'aunt' and *obaasan* 'grandmother' are different from each other only in term of the duration of a single vowel.

Table 1

Minimal Pairs Contrasting Vowel Length

	Short Vowels	**Long Vowels**
/a/ vs. /a:/	obasan 'aunt'	obaasan 'grandmother'
/i/ vs. /i:/	ojisan 'uncle'	ojiisan 'grandfather'
/u/ vs. /u:/	suji 'tendon'	suuji 'numerals'
/e/ vs. /e:/	sekai 'world'	seekai 'correct answer'
/o/ vs. /o:/	hosoku 'supplement'	hoosoku 'natural law'

Since it is a phonemic property, the length distinction is extremely clear to native speakers of Japanese. In fact, phonetic experiments have found that Japanese long vowels are 2.4~3.2 times as long as their short vowel counterparts in their duration (Han, 1962; Ueyama, 2000).

Objectives of Present Study

Despite the substantial difference in duration, learners of Japanese, who do not have vowel length distinctions in their native language in particular, tend to have significant difficulties with such distinctions (Tajima, Kato, Rothwell, Akahane-Yamada & Munhall, 2008). Such learners generally have trouble in both production and perception (Oguma, 2000; Toda, 2003). In perception for example, learners tend to mistake long vowels for short vowels (Oguma, 2000).

While I argue that the results of the present study are applicable to production as well, I focus on discussing the perception of vowel length distinctions in the ensuing discussions. The main objectives of this paper are threefold. First, I identify and discuss the reasons for the above-mentioned difficulties; second, I propose a way to teach how to discern vowel length distinctions more effectively; third, I discuss the logic behind the proposed method of teaching.

DURATION VS. PITCH

Problems with Duration

It has often been argued that the primary cue for vowel length distinctions is duration (e.g., Fujisaki, Nakamura & Imoto, 1975). However, a fundamental problem arises from the fact that duration is a relative concept. In other words, there is no absolute long or short. A long vowel is long because it is longer than the short vowel counterpart. What this means is that it always requires comparison, specifically comparison between two tokens (i.e., minimal pairs). This is particularly problematic for Japanese since minimal pairs contrasting vowel lengths are quite limited in number (Vance, 2008). While it is true that pitch is also a relative concept (Ladefoged & Disner, 2012), comparison of pitch is done within a token, the pitch of a mora and the pitch of another mora within the same word (i.e., pitch contour). Therefore, minimal pairs are not called for when we capitalize on pitch.

To make vowel lengths even less reliable, long and short vowels in Japanese significantly overlap with each other in their duration at different speaking rates (Hirata, 2004). In fact, some phonetic experiments (e.g., Kinoshita, Behne & Arai, 2002) have found that native Japanese speakers rely on pitch contours in distinguishing vowel lengths when the durational cue is unreliable.

Based on the reasons that I just discussed, I argue that duration, at least duration alone, cannot provide a reliable cue for vowel length distinctions in Japanese. It is important to also note that, as with any contrastive/phonemic properties, the distribution of long and short vowels cannot be predicted by rules. What this means for Japanese instructors is that they cannot possibly teach their students in which phonological environments they expect to find long vowels to the exclusion of short vowels and vice versa.

Pitch as a More Reliable Cue

In this section, I demonstrate that, while duration does not seem reliable, there are rather systematic correlations between long vowels and the pitch contours associated with them. I further demonstrate in the next section that these characteristic pitch contours follow from general accentuation rules.

Japanese is said to be a pitch-accent language, where each mora is associated with either a high-pitch (H) or a low-pitch (L). The mora is the smallest prosodic unit, or a "beat." In Japanese, a mora can be a single vowel (V), a consonant followed by a vowel (CV), the coda nasal (N), or, most importantly for our ensuing discussion, the lengthened part of a long vowel (R). In the word *kaiin* 'members' for example, there are four morae: *ka* (CV), *i* (V), *i* (R), and *n* (N); the first mora is pronounced with low-pitch, and the rest high-pitch: LHHH. It is crucial to note here that a long vowel consists of two morae. With this background in mind, let us now discuss vowel length distinctions in three different positions in words: word-initially, word-medially, and word-finally.

As shown in figures 1 and 2 respectively, a long vowel is associated with either a high-low (HL) pitch contour (a.k.a. "pitch drop") or a low-high (LH) pitch contour (a.k.a. "pitch rise") in the word-initial position. The pitch contours are indicated by blue dotted lines in the spectrographs.

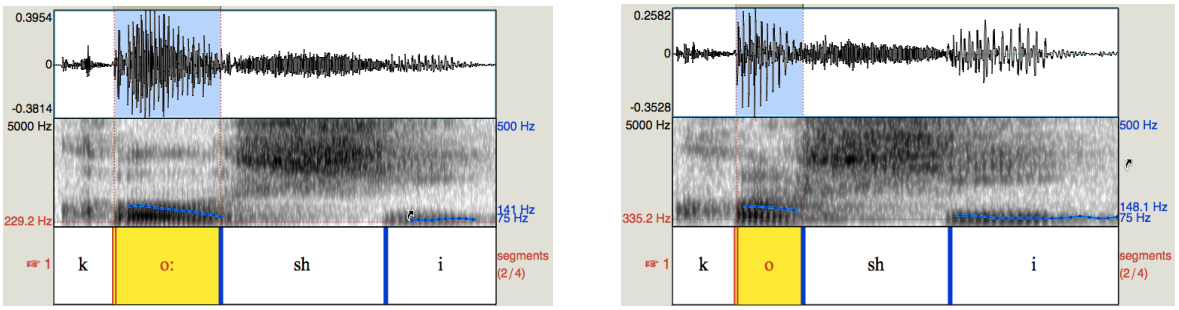

Figure 1. Length distinction word-initial position 1

Observe that, in the word *kooshi* 'lecturer' in Figure 1 (left), the pitch drops within the long vowel [o:]. Within this long vowel, the pitch dropped on an average of 44.33 Hz. Each token was recorded three times (see the appendix for the measurements). Recall that, since a long vowel is made up of two morae, a single long vowel can be associated with two distinct pitches. In *kooshi* 'lecturer,' the long vowel is associated with a high-pitch (average of 158.33 Hz) followed by a low-pitch (average of 114 Hz), resulting in a pitch drop (i.e., HL). On the other hand, no significant pitch change was observed in the short vowel in *koshi* 'old paper' in Figure 1 (right); the average pitch change within this short vowel was 11.33 Hz.

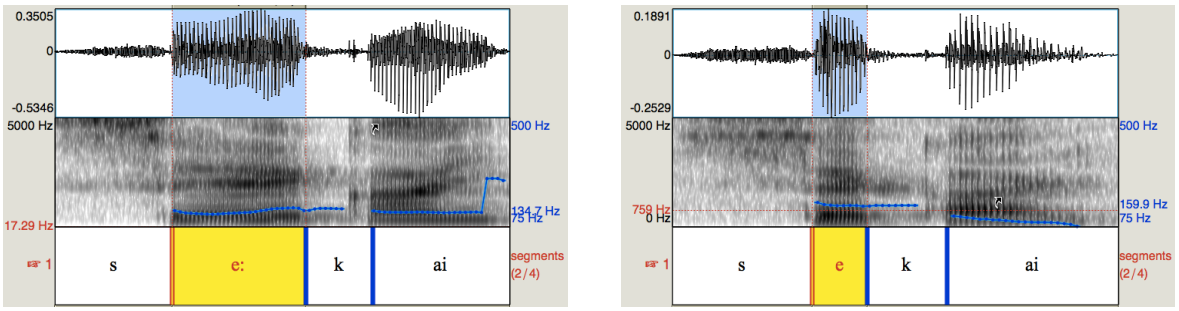

Figure 2. Length distinction word-initial position 2

Similarly, the pitch changes within the long vowel in the word *seekai* 'correct answer' in Figure 2 (left) as well; however, it changes from low to high, creating a pitch rise (i.e., LH). On average, the pitch rose by 22.67 Hz. On the other hand, there is no significant pitch change with the short vowel in *sekai* 'world' in Figure 2 (right); the average change was 8 Hz.

A similar yet distinct pattern emerges for the word-medial position. As illustrated in Figure 3 (left), a long vowel in the word-medial position is associated with a pitch drop, but not with a pitch rise.

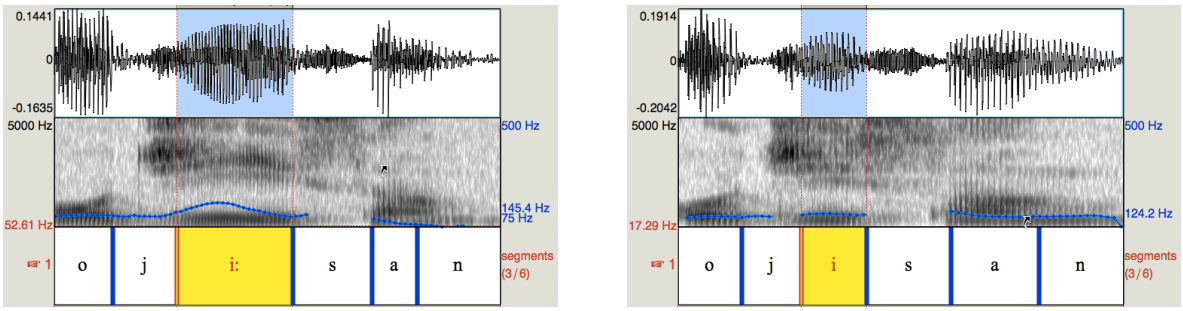

Figure 3. Length distinction in word-medial position

In the word *ojiisan* 'grandfather' in Figure 3 (left), the long vowel is realized as HL (i.e., pitch drop). The maximum pitch within this long vowel was averaged 168 Hz; the average minimum pitch was 112.64 Hz, resulting in a drop of 55.33 Hz. On the other hand, there is no significant pitch change with the short vowel in *ojisan* 'uncle' in Figure 3 (right); the average difference between the maximum and minimum pitch was 8 Hz. The crucial difference between the word-initial position and the word-medial position is that the former can be accompanied by either a pitch drop or a pitch rise, but the latter can be only associated with a pitch drop. I will discuss why such an asymmetry arises in the next section.

A completely different picture emerges for the word-final position. As illustrated in Figure 4, the pitch goes down for both the long vowel (left) and the short vowel (right), and it does so very steadily.

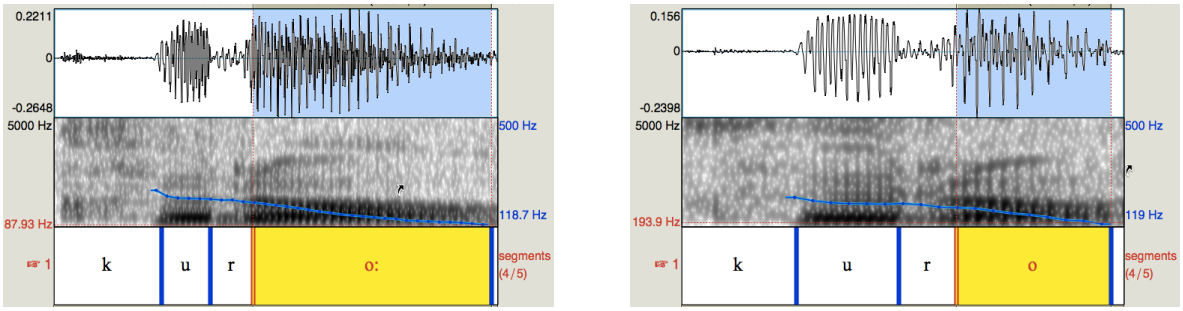

Figure 4. Length distinction in word-final position

I argue that this "decrease of pitch" must be distinguished from "pitch-drop" that we discussed above. In a pitch-drop (e.g., *ojiisan* 'grandfather' in Figure 3, left), two morae in a long vowel are linked with two distinct pitches: the first mora with H and the second mora with L. As a result, we observe a "bump," where the high pitch peaks. In contrast, we observe a smooth steady decrease of pitch in the word-final position as if there were only one mora. I surmise that the two morae of a word-final long vowel are associated with a single pitch (i.e., LL or HH), and that the (gradual) decrease of pitch is due to the decrease of air in the lungs, approaching the end of an utterance. This second point is also consistent with the fact that the pitch decrease occurs with a short vowel as well in the word-final position despite the fact there is only one mora (thus only one pitch). For these reasons, I argue that long vowels in the word-final position are characterized by a "flat" pitch. What this implies is that vowels in the word-final position have a very similar, if not identical, pitch pattern whether they are short vowels or long vowels, which, in turn, implies that vowel lengths cannot be distinguished in terms of pitch in the word-final position. While this may seem to undermine my proposal, it, in fact, supports it. It has been observed that the phonemic distinctions between long and short vowels are often neutralized in word-final positions in Japanese (Kubozono, 2002). For example, in words such as *ohayoo* 'good morning,' *hontoo* 'really,' and *sensee* 'teacher,' the long vowels can be pronounced short, "blurring" the differences in duration. In fact, words, such as *ohayoo* and *hontoo*, are sometimes even spelled with a short vowel, which never happens with long vowels in the word-initial or word-medial positions. While native speakers of Japanese are very keen to vowel length distinctions in word-initial and word-medial positions, they are much less sensitive in the word-final position. This asymmetry can be attributed to the fact that a pitch cue is available in the former but not in the latter. The characteristic pitch contours for long vowels in different positions are summarized in Table 2.

Table 2

Summary of Characteristic Pitch Contours of Long Vowels

Position	Pitch Contour	Example
Word-initial	"Pitch drop" (HL)	ko.o.shi (HLL) 'lecturer'
	"Pitch rise" (LH)	se.e.ka.i (LHHH) 'answers'
Word-medial	"Pitch drop" (HL)	o.ji.i.sa.n (LHLLL) 'grandfather'
Word-final	"Flat" (LL)	ku.ro.o (HLL) 'hardship'
	"Flat" (HH)	ru.bi.i (LHH) 'ruby'

As shown in Table 2, long vowels are associated with specific pitch contours in the word-initial and word-medial positions. I discuss in the next section that these pitch contours follow from general accentuation rules in Japanese.

RELIABILITY OF PITCH CONTOURS

I demonstrate in this section that the distribution of the characteristic contours I discussed above are predictable by general rules of accentuation although the distribution of long and short vowels itself is not predicable.

Recall that a long vowel in the word-initial position is always characterized by either a "pitch drop" or a "pitch rise." This fact follows from an accentual rule, the "initial lowering" rule (Haraguchi, 1977). According to this rule, the first mora of a word is always low-pitched in Japanese unless it is accented. It is important to note that the accented mora is marked H, and the remaining morae are marked L. With this background in mind, first consider a situation where the first mora is accented. If the first mora is accented, then the first mora is H, and the second mora is L, creating a pitch drop. If, on the other hand, the first mora is not accented, the first mora is L, and the second mora has to be H, resulting in a pitch rise. What this means is that the pitch of the first mora and that of the second mora must be always different from each other. In other words, the first two morae of a word must be either HL (i.e., pitch drop) or LH (i.e., pitch rise).

Let us now examine the word-medial position. Recall that, unlike in the word-initial position, long vowels in the word-medial position are always characterized by a pitch drop. In other words, long vowels cannot be associated with a pitch rise in word medial positions. This observation is consistent with the accentuation rule that "special morae" can never be accented. Special morae include the moraic nasal and moraic obstruents as well as the lengthened part of a long vowel. With this background in mind, let us examine two scenarios. First, in order for a pitch drop to obtain, the first mora of a long vowel has to be H and the second mora has to be L. This pitch contour is consistent with the pitch rule since the second mora (the lengthened part of the long vowel), which is a special mora, is not accented. Second, a pitch rise occurs only when the first mora of a long vowel is L and the second mora is H. Since the second mora of a long

vowel cannot be accented because it is a special mora, the LH pitch contour (i.e., pitch rise) is never available in the word-medial position.

SUMMARY

Since vowel length is a phonemic distinction in Japanese, the distribution of long and short vowels is unpredictable. This presents serious difficulty for both learners and instructor of Japanese. In this paper, I demonstrated that, while the length distinctions themselves are not predictable, the pitch contours associated with them are predictable, by showing that they follow from general rules of accentuation.

Given the predictability of pitch contours and the unpredictability of length distinctions, I argue that it is more effective to capitalize on the pitch rather than the duration in teaching vowel length distinctions in Japanese. More specifically, I suggest that instructors draw learners' attention to the lack or the presence of a pitch contour for short or long vowels, respectively. Since learners tend to mistake long vowels for short vowels (Oguma, 2000), training learners to identify the characteristic pitch contours of long vowels is crucial. In addition, since these distinctions often disappear in word-final positions, it is best to practice vowel length distinctions in the word-initial and word-medial positions.

ACKNOWLEDGMENTS

I would like to thank the participants of the 6th Annual PSLLT conference and the anonymous reviewer for their many constructive comments. Thanks are also due to the College of Humanities and Social Sciences and the Department of Modern and Classical Languages at Western Washington University for the funds made available to participate in the conference.

ABOUT THE AUTHOR

Masanori Deguchi (masanori.deguchi@wwu.edu) is an Associate Professor of Japanese and Linguistics at Western Washington University in Bellingham, Washington. He received his Ph.D. in Linguistics (major) and Philosophy (minor) from Indiana University, Bloomington in 2003. Ever since, he has been teaching at Western Washington University. He primarily teaches Japanese courses, including Japanese Phonetics, and he also teaches linguistics (e.g., phonology and semantics) and language pedagogy. His main area of expertise is in the syntax and semantics of Japanese. His representative publications include "Two Indirect Passive Constructions in Japanese" in Alexiadou and Schäfer (eds.) *Non-canonical passives*, John Benjamins.

Masanori Deguchi
Department of Modern and Classical Languages
516 High Street, Miller Hall 217E
Bellingham, WA 98226, USA
Phone: +1 360 650 4155

REFERENCES

Fujisaki, H., Nakamura, K., & Imoto, T. (1975). Auditory perception of duration of speech and non-speech stimuli. In G. Fant & M. A. A. Tatham (eds). *Auditory analysis and perception of speech* (pp. 197–219). London: Academic Press.

Haraguchi, S. (1977). *The tone pattern of Japanese: an autosegmental theory of tonology*. Tokyo: Kaitakusha.

Han, M. (1962). The feature of duration in Japanese. *Onsei no Kenkyu, 10*, 65-80.

Hirata, Y. (2004). Effects of speaking rate on the vowel length distinction in Japanese. *Journal of Phonetics, 32*, 565-589.

Kinoshita, K., Behne, D., & Arai, T. (2002). Duration and F0 as perceptual cues to Japanese vowel quantity. In J. Hansem & B. Pellom (eds). *7th International Conference on Spoken Language Processing* (pp. 757–760), Baixas, France: ISCA.

Kubozono. H. (2002). Temporal neutralization in Japanese. In C. Gussenhoven & N. Warner (eds). *Laboratory Phonology* 7 (pp. 171–201). Berlin & New York: Mouton de Gruyter.

Ladefoged, P., & Disner, S. F. (2012). *Vowels and consonants*. West Sussex, U.K.: Wiley-Blackwell.

Oguma, R. (2000). Perception of Japanese long vowels and short vowels by English-speaking learners. *Japanese-Language Education around the Globe, 10*, 43-54.

Shibatani, M. (1996). *The languages of Japan*. Cambridge: Cambridge University Press.

Tajima, K., Kato, H., Rothwell, A., Akahane-Yamada, R., & Munhall, K. G. (2008). Training English listeners to perceive phonemic length contrasts in Japanese. *Journal of Acoustic Society of America, 123*(1), 397-413.

Toda, T. (2002). *Second language speech perception and production: acquisition of phonological contrasts in Japanese*. Maryland: University Press of America.

Ueyama. M. (2000). *Prosodic transfer: An acoustic phonetic analysis of Japanese-accented English.* Ph.D. dissertation, University of California at Los Angeles.

Vance, T. (2008). *The sounds of Japanese*. Cambridge: Cambridge University Press.

APPENDIX: Pitch Measurements

Kooshi 'lecturer'

	Recording 1	Recording 2	Recording 3	**Average**
Max	155	159	161	**158.33**
Min	115	113	114	**114.00**
Difference	40	46	47	**44.33**

Koshi 'old paper'

	Recording 1	Recording 2	Recording 3	**Average**
Max	164	153	153	**156.67**
Min	158	139	139	**145.33**
Difference	6	14	14	**11.33**

Seekai 'correct answer'

	Recording 1	Recording 2	Recording 3	**Average**
Max	146	149	143	**146.00**
Min	120	124	126	**123.33**
Difference	26	25	17	**22.67**

Sekai 'world'

	Recording 1	Recording 2	Recording 3	**Average**
Max	160	163	154	**159.00**
Min	150	155	148	**151.00**
Difference	10	8	8	**8.00**

Ojiisan 'grandfather'

	Recording 1	Recording 2	Recording 3	**Average**
Max	169	167	168	**168.00**
Min	117	107	114	**112.67**
Difference	52	60	54	**55.33**

Ojisan'uncle'

	Recording 1	Recording 2	Recording 3	**Average**
Max	128	126	126	**126.67**
Min	118	122	121	**120.33**
Difference	10	4	5	**6.33**