EFFECTS OF WORD TYPE AND SPEAKING RATE ON THE DURATIONAL STRUCTURE OF 3- & 4-MORA WORDS IN JAPANESE

Yukari Hirata[a] *& Shigeaki Amano*[b]

[a]Colgate University, USA; [b]Aichi Shukutoku University, Japan
yhirata@colgate.edu; psy@asu.aasa.ac.jp

ABSTRACT

This study examined duration of Japanese three- and four-mora words produced by native speakers at three speaking rates. Examined word pairs had single and geminate stop contrasts. Three types of words were used that differed in their moraic compositions including (1) a CV mora, (2) a moraic nasal, and (3) a long vowel. We found that the word type had differential effects on the contrasting stop closure and word duration. Speaking rate also had effects on duration of various units and ratios. In spite of these effects, however, the durational structures were well-defined according to the phonological unit of the mora.

Keywords: Japanese moras, single and geminate stops, word type, speaking rate

1. INTRODUCTION

Japanese is known as a mora-timed language, and the traditional mora theory [1] claims that the mora is isochronous and evidenced in acoustic speech signals, i.e., duration of moras tends to be roughly equal in Japanese words. A line of studies [2-5] showed that duration of Japanese words was correlated with the number of moras in the words, e.g., words of two, three, and four moras roughly corresponding to the word duration of 2:3:4. They suggest that each mora might not be in equal duration but moras within a word are adjusted to its expected duration as a whole. In studies [1, 2, 4], a geminate consonant C_2C_2 in $C_1V.\underline{C_2.C_2}V$ had duration three times longer than that of a single consonant in $C_1V.\underline{C_2}V$. The traditional mora theory [1] explains that each of Cs and Vs in $C_1V.C_2V$ receives a 0.5 mora weight, but that the geminate C_2C_2 in $C_1V.C_2.C_2V$ receives 1.5 mora weights because it included one mora plus half of a CV mora, and thus the ratio of singleton : geminate = 0.5:1.5 = 1:3. However, most studies focused on disyllables, and few studies have examined more than a limited number of longer words where the mora isochrony might be affected by having different types of moras in words (e.g., V, CV, the moraic nasal, long vowels, and geminate stops) or by having speech materials of different speaking rates. In sum, researchers are not yet in agreement about the acoustic reality of mora isochrony [6].

The present study, therefore, examined the durational structure of Japanese three- and four-mora words with different segmental compositions and speaking rates to test generalizability of earlier studies that supported the mora theory. Specifically, we examined pairs of real words that included single and geminate stop contrasts in the following three types of words classified based on their moraic composition (Table 1):

- *Word type C* included CV or V moras, e.g., *i.ka.ku* (3 moras) ‘threat’ and *i.k.ka.ku* (4 moras) ‘one stroke.’
- *Word type N* included a moraic nasal, e.g., *go.ka.n* (3 moras) ‘word connotation’ and *go.k.ka.n* (4 moras) ‘extremely cold.’
- *Word type V* included a long vowel, e.g., *ha.ke.e* (3 moras) ‘waveform’ and *ha.k.ke.e* (4 moras) ‘eight beautiful scenes.’

Table 1: Sample stimuli (of 20 words in each word type) used in the present experiment.

Word type	Singletons (3-mora words)	Geminates (4-mora words)
C CV.(<u>C.)C</u>V.CV	ka.ke.tsu so.ko.ku ma.ku.ra i.to.ku ⋮	ka.k.ke.tsu so.k.ko.ku ma.k.ku.ra i.t.to.ku ⋮
N CV.(<u>C.)C</u>V.N	sa.ki.n go.ka.n i.ka.n ka.ta.n ⋮	sa.k.ki.n go.k.ka.n i.k.ka.n ka.t.ta.n ⋮
V CV.(<u>C.)C</u>V.V	da.to.o i.ko.o ha.ke.e mi.te.e ⋮	da.t.to.o i.k.ko.o ha.k.ke.e mi.t.te.e ⋮

We examined whether these pairs of words with different numbers of moras and syllables would yield durational patterns according to the mora theory put forth in previous studies [1-5]. The following questions guided our investigation:

1) *Stop closure duration*. How do the word type and the speaking rate affect the duration of single and geminate stop closure? Does the ratio of singletons and geminates remain at 1:3, regardless of word types and speaking rates?
2) *Word duration*. How do the word type and the speaking rate affect the duration of three- and four-mora words containing single and geminate stops? Does the ratio remain at 3:4, regardless of the two factors?
3) *Subword duration*. We used the term "subword" to refer to the portion of the three- and four-mora words that excludes the final mora for word types C and N, e.g., the portion, *i.ka* or *i.k.ka*, in the pair *i.ka.ku* 'threat' and *i.k.ka.ku* 'one stroke.' Do subwords show durational patterns similar to those of two- and three-mora words as in [4]? This would be the case if the traditional mora theory holds where mora isochrony is assumed. If duration is adjusted only for the entire word while individual moras are not isochronous, then the truncated portion should show a durational pattern different from that of two- and three-mora words as in [4].

2. METHOD

2.1. Stimuli

Four adult native speakers of Japanese, born and residing in Tokyo metropolitan areas, participated in recording. For each of the three types of words, C, N, and V (Table 1), there were 10 three-mora words and 10 four-mora words including single and geminate stops. All words were meaningful, real words, with various consonants and vowels [/k(ː) t(ː)], [k t g d ts ɟʑ s ɕ h m n ɾ], and [i e ɑ o ɯ eː oː/]. Pitch accents were matched between 6, 7 and 8 word pairs in C, N, and V, respectively, while those of other pairs were not.

These words were spoken in a carrier sentence [soɾe wɑ ____ dɑ to omoimɑsɯ̥] 'I think that it is ___," spoken at slow, normal, and fast rates. Speaking rates were defined as follows: "as slowly as possible but still connected" for *slow*, "relaxed and comfortable" as *normal*, and "as fast as possible without making excessive errors" for *fast* rates. The actual rate of speech was self-determined by each speaker. Speakers read a randomized list of words first at the normal rate, followed by the slow and then fast rates. The order of words was different for each repetition. A total of 2160 tokens (= 60 words x 4 speakers x 3 rates x 3 repetitions) was recorded. Out of 2160 tokens, 10 tokens were discarded due to slurring.

2.2. Analyses

Stimuli were digitally recorded and their duration was measured using PRAAT. Duration of contrasting stop closure (the underlined (C)C portion in Table 1), the preceding and following vowels, the target word, as well as the word-initial consonant if the word begins with one. For word types C and N, "subword," i.e., the duration of the target word excluding the final mora, was also measured. "Subword" was not measured for word type V since we cannot reliably divide the last two moras that include a long vowel.

First, an analysis of variance (ANOVA) was conducted for each of stop closure, word, and subword durations, and reported separately in three result sections (3.1-3.3). Quantity (singleton, geminate), speaking rate (slow, normal, fast), and word type (C, N, V) (V excluded from subword analysis) were within-subjects factors. Second, three durational ratios were calculated by rate and by word type: (1) the ratio of geminate to singleton stop closure, (2) the ratio of four-mora to three-mora words with the three-mora word duration standardized by three (3:X), and (3) the ratio of geminate to singleton subword with the singleton subword duration standardized by two (2:X).

3. RESULTS

3.1. Closure duration

Closure duration was affected by all three factors of quantity, speaking rate, and word type. The mean duration was significantly longer for geminates (218 ms) than singletons (74 ms) [$F(1,3)=525.2$, $p<0.001$]. There was a significant interaction of quantity and rate [$F(2,6)=96.3$, $p<0.001$]. As shown in Table 2, the duration differences between singletons and geminates were greater for slower rates. The closure duration was also significantly affected by the word type [$F(2,6)=38.1$, $p<0.001$]. The mean closure duration in word type V (158 ms) was significantly longer than that in word type C (134 ms) [$p=0.005$], and marginally longer than that in word type N (147 ms) [$p=0.053$], while the difference between types C and N was not significant [$p>0.1$].

Table 2: Mean closure duration (ms) of single (S) and geminate stops (G) by rate and word type.

Speaking rate	Quantity	Word Type		
		C	N	V
Fast	S	37	45	53
	G	90	105	115
Normal	S	56	73	85
	G	165	185	200
Slow	S	93	108	119
	G	362	364	378

Table 3 shows the mean ratios of geminate to singleton stop closure for each rate and word type. The normal rate ratio for word type C showed a value closest to 3, which was what the traditional mora theory had predicted. However, the ratios were observed to be smaller for faster rates, and this pattern was observed for all three word types.

Table 3: Mean ratios of geminates to singletons by rate and word type.

Speaking rate	Word Type		
	C	N	V
Fast	2.43	2.36	2.17
Normal	2.96	2.51	2.36
Slow	3.88	3.38	3.19

3.2. Word duration

Word duration was also affected by the three factors of quantity, speaking rate, and word type. The mean word duration was significantly longer for those including geminates (i.e., four-mora words, 642 ms) than singletons (three-mora words, 497 ms) [$F(1,3)=252.1$, $p<0.001$]. There was a significant interaction of quantity and rate [$F(2,6)=103.1$, $p<0.001$]. As shown in Fig. 1, the duration differences between singletons and geminates were greater for slower rates. The word duration was also significantly affected by the word type [$F(2,6)=23.4$, $p=0.001$]. The mean word duration for type N (593 ms) was significantly longer than those for type V (562 ms) [$p=0.01$] and for type C (553 ms) [$p=0.03$], while the difference between types V and C was not significant [$p>0.1$].

Table 4 shows the mean ratios of four- to three-mora words with the three-mora word duration standardized by three for each rate and word type. Word type C at the normal rate showed exactly the predicted ratio of 3:4. However, as in the results of stop closure duration, the ratios were observed to be smaller for faster rates, and this pattern was observed for all three word types.

Figure 1: Mean duration of three- and four-mora words including single (S) and geminate stops (G). C, N, V represent three word types.

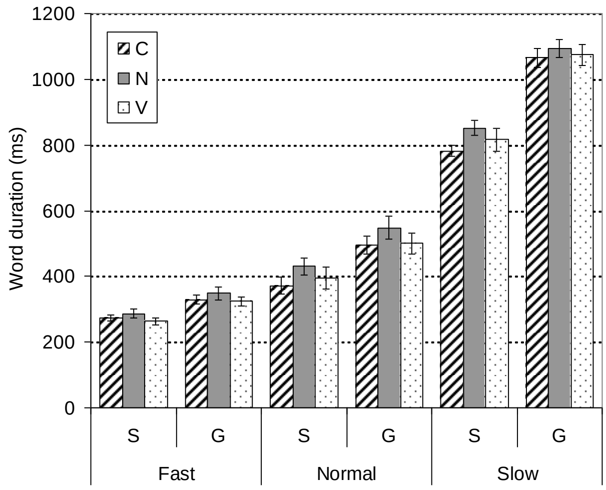


Table 4: Mean ratios of four- to three-mora words, with three-mora word duration standardized by three (three-mora: four-mora = 3 : X).

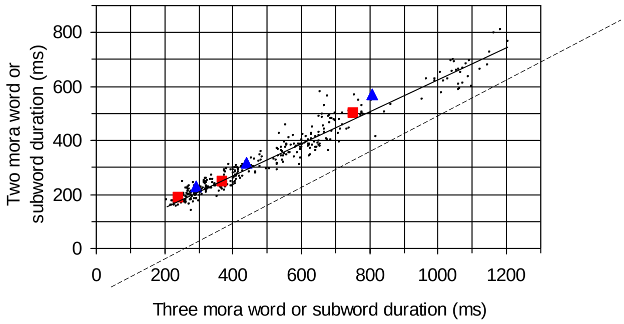
Speaking rate	Word Type		
	C	N	V
Fast	3.60	3.65	3.70
Normal	4.00	3.82	3.80
Slow	4.09	3.85	3.95

3.3. Subword duration

The ANOVA for subword duration showed main effects of quantity and rate, and an interaction of these two factors [$F(2,6)=109.6$, $p<0.001$], similar to those in the word duration above. There was also a significant effect of word type [$F(1,3)=30.0$, $p=0.012$]: the mean subword duration was longer for word type N (443 ms) than C (383 ms).

If moras are isochronous, we would predict the subword duration to resemble two- and three-mora words found in [3], having the ratio of 2:3. In the present data, the mean ratios of the two-mora vs. three-mora subwords for word type C were 2:3.03 (slow), 2:3.03 (normal), and 2:2.57 (fast), and those for word type N were 2:2.84 (slow), 2:2.76 (normal), and 2:2.56 (fast), showing a decreasing tendency with increased rates. Fig. 2 shows the mean two- versus three-mora subword durations plotted and compared with data obtained for two- and three-mora words in [4]. As shown in this figure, the present subword data are in very close overlap with the two- and three-mora words in [4].

Figure 2: Mean values of two- vs. three-mora subword duration of the present data plotted over two- vs. three-mora word duration in [4].



■ Wtype C subwords (fast, normal, and slow rates)
▲ Wtype N subwords (fast, normal, and slow rates)
· 2- & 3-mora words from [4]
— Regression line in [4]
- - - Hypothetical function predicted from [5]

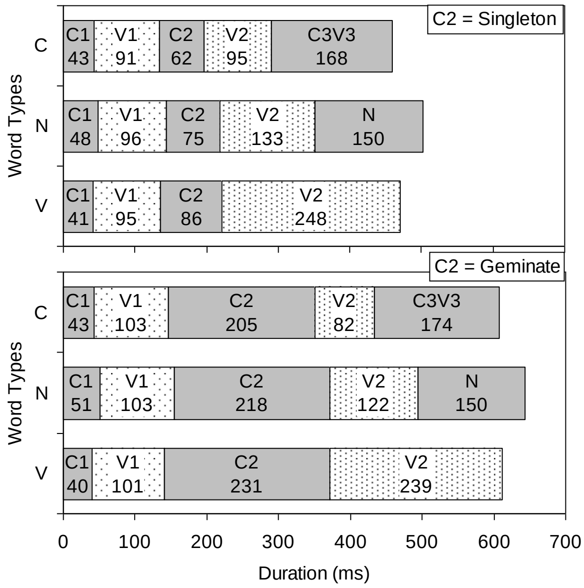
4. DISCUSSION

Fig. 3 shows the summary of word internal structures for three- (upper panel) and four-mora words (lower panel) by word type, averaged across rates. We can see that the contrasting stop closure duration (C2) was systematically affected by the word type (in 3.1), but that the distinction of singletons and geminates was still maintained clearly. This quantity distinction was well reflected on the overall word duration (in 3.2). As shown in 3.1 and Fig. 3, the stop closure was systematically longer in word type V than C. This could well be to compensate for the fact that the final mora in the long vowel in word type V was not as long as a full C3V3 mora. In this case, we could say that mora compensation was successful since word duration of the two word types ended up being similar (i.e., not significantly different).

In Fig. 3 (also in 3.3) we observe that subword duration is longer for word type N than C. However, it is interesting to note that this difference is well within the durational variation of two- and three-mora words that is caused by speaking rate differences in [4] (Fig. 2). We can say that the duration of subwords with single versus geminate stops is well defined in terms of the corresponding phonological mora counts.

Taken together, the present findings are not inconsistent with the idea that word-internal segments and moras tend to be organized so that their sum matches with the duration predicted by the mora theory, supporting [1-5]. Duration of moras is not perfectly isochronous due to various constraints but is predictable and well organized towards isochrony.

Figure 3: Word-internal duration structures of word types C, N, and V. The number in each cell shows the mean duration (ms) of each segment.



5. ACKNOWLEDGMENTS

This research was supported by Colgate University and NTT Communication Science Laboratories.

6. REFERENCES

[1] Han, M. 1962. The feature of duration in Japanese. *The Study of Sounds* 10, 65-80.

[2] Han, M. 1994. Acoustic manifestations of mora timing in Japanese. *J. Acoust. Soc. Am.* 96(1), 73-82.

[3] Hirata, Y. 2004. Effects of speaking rate on the vowel length distinction in Japanese. *J. Phonetics* 32, 565-589.

[4] Hirata, Y., Whiton, J. 2005. Effects of speaking rate on the single/geminate stop distinction in Japanese. *J. Acoust. Soc. Am.* 118(3), 1647-1660.

[5] Port, R.F., Dalby, J., O'Dell, M. 1987. Evidence for mora timing in Japanese. *J. Acoust. Soc. Am.* 81(5), 1574-1585.

[6] Warner, N., Arai, T. 2001. Japanese mora-timing: A review. *Phonetica* 58, 1-25.