

# THE PERCEPTION OF MORAE IN LONG VOWELS

## Comparison among Japanese, Korean and English speakers

Ryoji Baba, Kaori Omuro, Hiromitsu Miyazono †  
Tsuyoshi Usagawa, Masahiko Higuchi ‡

† Prefectural University of Kumamoto  
‡ Kumamoto University

### ABSTRACT

There are three kinds of *Tokushuhaku* (the specific timing morae) in the Japanese language such as the moraic nasal, the non-nasal consonant and the long vowel. Even though Japanese native speakers can perceive them perfectly, it is difficult for Japanese learners from abroad to perceive and produce them. To make it more efficient for Japanese learners to acquire them, we need to investigate how Japanese native speakers perceive *Tokushuhaku*.

This study examines the perception of a mora or morae of Japanese language specifically in the long vowel. Previous researchers (<sup>1</sup>Hiroya Fujisaki, Miyoko Sugito) have assumed that duration contrasts alone are involved in perception. However we hypothesize that the accentual change in the midst of a long vowel would mark the boundary between the morae and would affect the perception of morae by native speakers of standard Japanese.

Based on the hypothesis three kinds of perception tests with computer-edited sound stimuli were given to native Japanese, Korean and English speakers. The results of these tests suggest that native Japanese speakers count the number of morae not only by the duration of vowels, but also by the accentual change which indicates the boundary between morae. On the other hand, non-native speakers count the number of morae only by the duration of vowels.

### 1. JAPANESE SYLLABLE, MORA AND ACCENT

The phonetic units of English and Korean are syllables, and the units of Japanese are syllables and morae. 'Japanese', 'English' and 'Korean' are divided into syllables as 'ja•pa•nese', 'eng•lish' and 'ko•re•an'. The same meaning words in Korean are divided into syllables as 'il•bo•neo', 'yeo•ngeo', 'han•gu•geo' and the words in Japanese 'ni•hon•go', 'e•go' and 'kan•ko•ku•go'. The Japanese words are also divided into morae as 'ni•ho•n•go', 'e•e•go' and 'ka•n•ko•ku•go'. Japanese native speakers divide the word 'nihongo' (Japanese) into syllables as 'ni•hon•go' and also into morae as 'ni•ho•n•go'. In the former, each syllable is the sound cluster to pronounce with ease and in the latter, each mora is the unit of time. Japanese natives perceive that each mora has the same length.

The difference between 'ho•ne' (bone) vs. 'ho•n•ne' (true feeling) is the existence of the nasal *Tokushuhaku* 'n', and the difference between 'ki•te' (wear and) vs. 'ki•t•te' (stamp) is the existence of the non-nasal consonant *Tokushuhaku* 't', and 'ku•tu' (shoe) vs. 'ku•t•u' (pain) the existence of vocalic *Tokushuhaku* 'u'.

The sentences have intonations and the words have accents in the spoken languages. The functions of accent is to divide the chain of sounds into some parts that compose the sentence, and to distinguish the meanings of the words composed with the same speech sounds. English has stress accent, Japanese has pitch accent and Korean has neither stress nor pitch accent. Korean has no accent to distinguish the meanings and has liaison to clarify the relations between the parts in sentences.

The Japanese sound, '\_ga\_ k•ko•o\_ kyo\_ o•i•ku' (the pitch of the morae 'ga' and 'kyo' which follow "\_", is low, and the pitch of the morae 'k•ko•o' and 'o•i•ku' which follow "-", is high) means 'school, education' and '\_ga\_ k•ko•o\_ kyo\_ o•i•ku' means 'school education'. We have a Japanese minimal pair: '\_ha\_ si' (bridge) vs. 'ha\_ si' (chopsticks). Only the difference of the accent distinguishes these two words.

In the accent system of the standard Japanese the pitch of the first mora is always different from the pitch of the second as '\_ha\_ si', 'ha\_ si', '\_a\_ me' (candy) vs. 'a\_ me' (rain).

### 2. THE EXPERIMENT USING A SEQUENCE OF VOWELS INTERRUPTED BY A CONSONANT

#### 2.1. Procedure

The nonsense words [\_a\_ tade\_s], [\_a\_ taade\_s] pronounced by an adult male standard Japanese speaker were used for the experiment as the sound stimuli. The original sounds were recorded into a digital audio tape and the sound stimuli were created by <sup>2)</sup>PICOLA (Pointer Interval Control OverLap and Add).

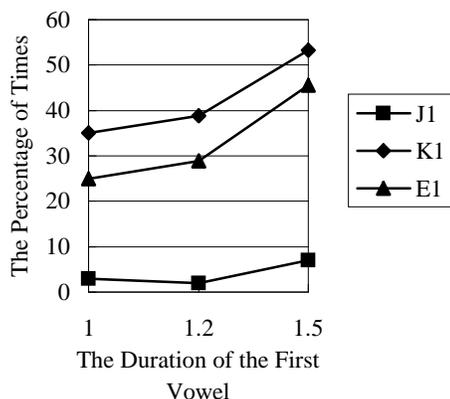
The duration of the first vowel [a] of [atades] was shortened in 3 stages (1.0, 0.6, 0.3), and at the same time the duration of the second vowel [a] was lengthened in 5 stages (1.0-2.0). The duration of the first vowel of [ataades] was lengthened in 3 stages (1.0, 1.2, 1.5), and at the same time the duration of [aa] after the consonant [t] was shortened in 6 stages (1.0-0.375). The total number of the sound stimuli was 33, and 10 sets of these stimuli were recorded into a digital audio tape at random.

The subjects were divided into the three groups; 10 native Japanese speakers (group J1), 9 native Korean speakers (group K1), 9 native English speakers (group E1). The subjects were asked to listen to the tape using headphones and to mark the one of 4 choices closest to what they perceived; "atadesu", "aatadesu", "ataadesu", and

“aataadesu” written in Japanese Hiragana for J1, Korean Hankul alphabet for K1 and romanization for E1 on the answer sheets.

## 2.2. Results

Figure 1 shows the percentage of times that subjects marked “aataadesu” when they listened to the stimuli created from the sound [ataades] by lengthening the first vowel in 3 stages (1.0, 1.2, 1.5). The duration of the second and third vowels was shortened to 0.375 in all cases.



**Figure 1:** The percentage of times that subjects marked “aataadesu” when they listened to the stimuli created from the sound [ataades] by lengthening the first vowel in 3 stages with duration of the second and third vowels shortened to 0.375.

Even though the first vowel was lengthened to 1.5, J1 did not perceive the first vowel of the stimuli as [aa] (two morae). On the other hand, K1 perceived it as [aa] more than 5 times of 10 of the stimuli, and those in E1 did as many times as K1 did. The more the first vowel lengthened, the more they perceived the stimuli as [aa] (two morae).

As Fig. 1 shows, J1 rarely perceived the stimuli as [aataades] even though the duration of the first vowel was lengthened. On the other hand, both K1 and E1 perceived them as [aataades] when the first vowel was lengthened. In standard Japanese, however, the accent of the first mora must be different from the accent of the second, there was no accentual change in the lengthened first vowel [a] of the stimuli. Thus the subjects in J1, who perceive morae by the accentual change, did not perceive the stimuli as [aataades]. On the other hand, non-native Japanese speakers who perceive morae by the duration perceived it as [aataades].

## 3. EXPERIMENT USING A SEQUENCE OF VOWELS WITHOUT A CONSONANT

Two experiments examined how the pitch change affects the perception of morae in a sequence of vowels without a consonant. In our previous experiment we assumed that there were two factors

influencing Japanese native speakers perception of the number of morae of the first vowel; a consonant [t] and a pitch change. In the following experiments since sound stimuli without consonants were used, the basis for perceiving the first vowel is only the pitch change.

### 3.1. The Experiment with the Sound Controlled by the Speech Velocity

The two nonsense sounds [aade\_s], [aade\_s] were used for the experiment. A male standard Japanese speaker pronounced these two sounds with 6 degrees of speech velocity and with 3 degrees of pitch change; flat(no pitch change), normal (small pitch change), and emphatic(large pitch change). The pitch change occurs between the first mora and the second mora. The stimuli were recorded into a digital audio tape. The total number of the stimuli was 36, and 10 sets of these stimuli were recorded into a digital audio tape at random.

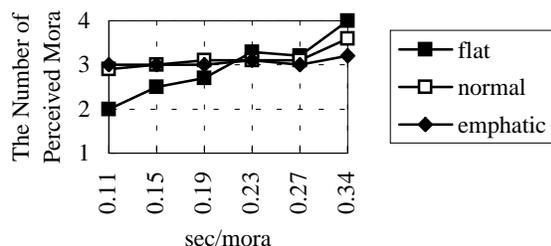
The subjects were divided into the three groups; 8 native Japanese speakers who have learned Japanese phonology (group J2a), 10 native Japanese speakers who had not learned Japanese phonology (group J2b), 21 native Korean speakers (group K2). The choices on the answer sheets were “adesu”, “aadesu”, “aadesu”, and “aaadesu” written in Japanese Hiragana or Korean Hankul.

### 3.2. Results

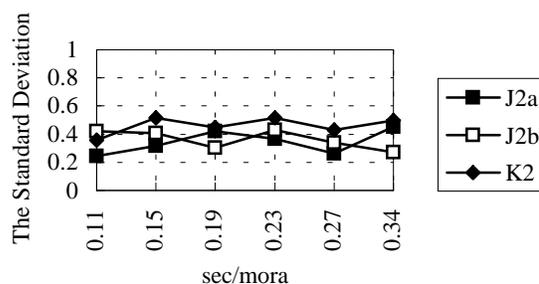
Fig. 2, 3 show the typical results of an subject in each group J2a and group K2. The stimuli is [aaades], and pronounced with 6 degrees of speech velocity (0.11-0.34 sec/syllable) and 3 degrees of pitch change. Based on our hypothesis that the pitch change affects the perception of morae, it is assumed that even if the speech velocity was slowed and thus the total duration of the word became longer, the number of morae in the long vowel would be perceived as having many morae as the original sound (3 morae; [aaades]).

As Fig. 2 shows, in the flat pitch change, J2a perceived the original sound as [aaades] when it was pronounced fast; and the slower the stimuli was pronounced, the great were the number of perceived morae. On the other hand, the stimuli with normal and emphatic pitch change were perceived as having 3 morae [aaa] and even when the stimuli was pronounced slower, the number did not increase as much as when the stimuli were pronounced with flat pitch. Furthermore, the number of perceived morae of the stimuli with emphatic pitch change did not increase much compared to the stimuli in the flat and normal pitch change.

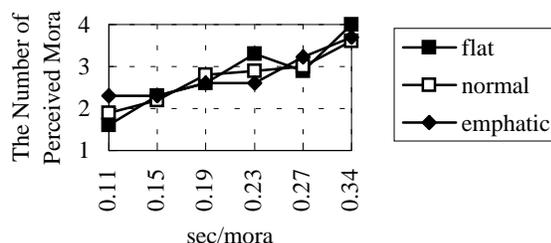
In Fig. 3, the Korean subject did not perceive the stimuli in 0.11 sec/mora velocity as 3 morae, and the number of perceived morae increased constantly when the stimuli was pronounced slower. Regardless of the difference of the pitch level, the slower the speech velocity became, the more the number of perceived morae increased.



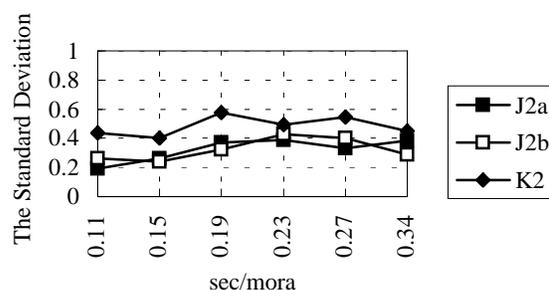
**Figure 2:** The typical results of an subject in group J2a. The original sound of the stimuli is [aaades], and pronounced with 6 degrees of speech velocity and 3 degrees of pitch change.



**Figure 5 :** The standard deviation in normal pitch change.

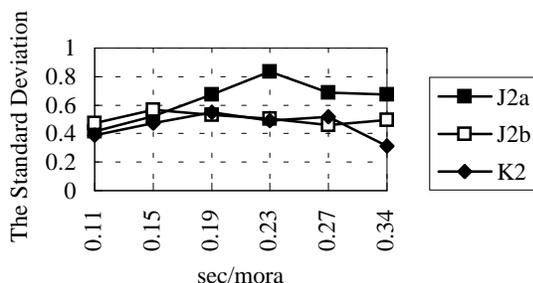


**Figure 3 :** The typical results of an subject in group K2.



**Figure 6 :** The standard deviation in emphatic pitch change.

Fig. 4-6 show the standard deviation in 3 degrees of pitch change. The perception in the normal and emphatic pitch change of J2a is stable compared to that in the flat pitch change. It means that the accentual change affects the perception of morae of native Japanese speakers. On the other hand, there was no difference among the pitch levels in the results of K2. It means that the pitch change does not affect the perception of K2. If the pitch change is clearly indicated, the subjects in J2a can perceive the number of morae more easily, and if the pitch change is not clearly indicated, it is difficult for them to perceive it.



**Figure 4 :** The standard deviation in flat pitch change.

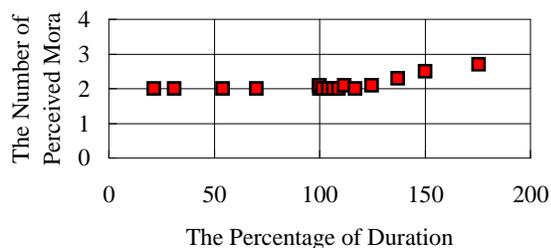
### 3.3. The Experiment with the Sound Controlled by the Duration

The two nonsense words [aades], [aaades] pronounced by a male standard Japanese speaker were used for the experiment. An early part of the ‘aa’ vowel in [aades] was shortened by computer synthesizer in 4 stages (24% min.) creating 4 stimuli. 10 stimuli were created in the same way by lengthening a late part of the vowel in 10 stages (175% max.). From [aaades] 6 stimuli were created by lengthening an early part of the ‘aaa’ vowel in 6 stages (293% max.) and 6 stimuli were created by shortening a late part of the vowel (24% min.). The original sounds were recorded on a digital audio tape, and were analyzed and synthesized by DSP Sona-Graph 5500 and LPC Parameter Manipulation and Synthesis Program by Kay Elemetrics Corp. The total number of sound stimuli was 31 including the two original sounds of [aades] and the three original sounds of [aaades]. 10 sets of these stimuli were recorded into a digital audio tape at random.

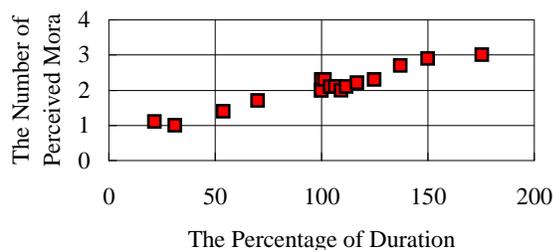
The subjects were divided into the three groups; 8 native Japanese speakers who had learned Japanese phonology (group J2a), 10 native Japanese speakers who had not learned Japanese phonology (group J2b), 21 native Korean speakers who had learned Japanese (group K2a), 9 native Korean speakers who had not learned Japanese (group K2b). The subjects in group J2a, J2b and K2a were the same as those in the previous examination. The choices are “adesu”, “aadesu”, “aaadesu”, and “aaaadesu”.

### 3.4. Results

Fig. 7 shows the typical result of an subject in J2a and Fig. 8 shows one in K2a. The original sound of the stimuli is [aades]. We hypothesized<sup>3)4)</sup> that native Japanese speakers would perceive the boundary of the first mora in a long vowel by the accentual change, and use the duration of the (accentually determined) first mora as a measure to determine the number of morae in the vowel. Based on the hypothesis, it is assumed that when the duration of the early part of the vowel is shortened, the sound will be perceived as [aaades] (3 morae) and that when the duration of the late part is lengthened, the stimuli will also be perceived as [aaades].



**Figure 7 :** The typical result of an subject in J2a. The original sound of the stimuli is [aades] and an early part of the ‘aa’ vowel was shortened in 4 stages creating 4 stimuli. Ten stimuli were created by lengthening a late part of the vowel in 10 stages.



**Figure 8 :** The typical result of an subject in K2a.

As Fig. 7 shows, even when the duration of the early part of the vowel is shortened, the number of morae perceived by J2 does not decrease. On the other hand, as Fig. 8 shows the subject in K2a perceived a smaller number of morae when the first vowel was shortened. That means the Korean subjects perceived the morae based on the total duration of the stimuli, but the Japanese subject did not. We assumed that since the pitch change affects the perception of the Japanese subjects, the number of perceived morae of the stimuli did not decrease even though the duration was shortened. If Japanese subjects perceive the morae only by duration, they should perceive fewer morae when the duration of the first vowel is shortened.

## 4. DISCUSSION

In the examination using the sequence of vowels interrupted by a consonant, when the duration of the first vowel of the sound [ataades] was lengthened, Japanese subjects did not perceive it as [aatades] nor [aataades] even though non-native Japanese subjects did. Thus it appears that if there is no accentual change in the sequence of vowels, native Japanese speakers do not perceive the sequence as bimoraic. In other words, duration contrasts alone affect the perception of non-native Japanese speakers, and accentual change affects the perception of native Japanese speakers more than the duration does.

In the experiments using the sequence of vowel without a consonant, the accentual change was the only cue to perceive the first mora. As the experiment with the sound controlled by the speech velocity indicated, if there is a strong accentual change in the sequence of vowels, native Japanese speakers perceive the sequence more clearly and easily as multi-moraic.

## 5. CONCLUSION

Previous researchers (<sup>1)</sup> Hiroya Fujisaki and Miyoko Saito) insisted that only the duration of the vowel affects the perception of the number of morae. Based on the results of our research, however, it is clear that not only the duration of the vowel sequence but also the accentual change has an important influence on the perception of Japanese subjects.

Japanese learners, in whose mother tongue length is not a distinctive feature of vowels and no plural units are perceived in a long vowel, must acquire the ability to catch the accentual change in the midst of a long vowel in order to perceive the number of morae of the vowel.

## 6. ACKNOWLEDGMENT

We would like to thank to Prof. Morita and Prof. Itakura of Nagoya University for their generous permission to use PICOLA.

## 7. REFERENCES

1. Fujisaki, H. and Sugito, M., *Onsei no Buturiteki Seisitu: Iwanami Koza Nihongo 5*, Iwanami Syoten, 65-105, 1977.
2. Morita, N. and Itakura, F., “Time-Scale Modification Algorithm for Speech by Use of Pointer Interval Control OverLap and Add (PICOLA) and Its Evaluation,” Proc. of Annual Meeting of Acoust. Soc. of Jpn., Oct., 1986.
3. Kai, T., Baba, R. and Others, “A Hearing a Mora in a Long-Vowels: Between Japanese and Koreans,” Technical Paper of Hearing Committee of Acoust. Soc. of Jpn., H-94-80, Dec., 1994
4. Omuro, K., Baba, R. and Others, “The Perception of Morae in Long Vowels: Comparison among Japanese, Korean and English speakers,” Technical Paper of Hearing Committee of Acoust. Soc. of Jpn., H-95-88, Dec., 1995