MANIFESTATION OF LEXICAL ACCENT AND TIMING STRATEGY IN ENGLISH SPEAKERS' JAPANESE

Mariko Kondo
IALS, University of Edinburgh, Scotland

ABSTRACT

The study investigated the acquisition of Japanese speech rhythm by English speakers. Durational data of beginners and fluent speakers were compared with those of native Japanese speakers. The results showed that beginners had not yet acquired the durational characteristics of Japanese speech rhythm, which are based on the number of morae in a word, whereas fluent speakers demonstrated native-like durational control. Not only did beginners have difficulty in pronouncing words with the correct number of morae, but they also transferred English accentuation strategy, namely lengthening accented vowels and reducing unaccented vowels. Fluent speakers were able to control vowel duration separate from accent.

The results implied that speech rhythm is acquired at two separate levels; a lower level durational control at the mora level and a higher level control at the word level.

1. INTRODUCTION

Japanese speech rhythm has a durational characteristic that the duration of whole words is proportional to the number of morae [6]. Words with the same number of morae have approximately an equal duration, despite consisting of phonemes of inherently different durations and different number and structure of syllables. Presence of lexical accent or position in an utterance has little effect on duration.

This process is not easy for non-native speakers of Japanese to learn. Although there is a tendency to equalise the duration of each mora to a certain extent, the duration of each mora is not necessarily equal [1, 3, 7]. The duration of each phoneme is not consciously controllable for the speaker; it is stretched or reduced depending on its adjacent phonemes [1, 3]. It is also essential to be able to recognise every mora in a word because the duration of a whole word is determined by the number of morae in the word. Non-native speakers who are able to read Japanese can recognise a mora in Japanese, because each mora corresponds to one letter of 2 orthographic systems.

For English speaking learners of Japanese, there are 2 main problems: (1) phonological differences between Japanese and English, and (2) different durational controlling system in the 2 languages. Firstly, the mora does not have a significant status in English phonology. Therefore, although each mora is recognisable in Japanese writing, English speakers are not always aware of all morae when the numbers of morae and syllables do not match; namely the sequences including (a) long vowels and (b) moraic consonants /Q, N/, where there are 2 morae in 1 syllable [5].

Secondly, English has different durational control mechanisms in relation to speech rhythm. For instance, the duration of an accented vowels is longer than unaccented vowels. Another important rhythmic characteristic of English is that normally a strong and a weak syllables alternate, and there is a tendency of isochrony between strong syllables (i.e. the foot), though the duration of each foot is not strictly equal [2]. These problems and differences of the 2 languages affect the processes of English speakers' acquisition of Japanese speech rhythm.

In this paper, the process of acquisition of Japanese speech rhythm by native English speakers is studied. Firstly, differences in the durational features of Japanese speech rhythm will be compared between 2 groups of English speaking learners of Japanese, beginners and fluent speakers. Secondly, the processes and tactics of speech rhythm acquisition in fluent speakers will be examined.

Since speakers cannot consciously control the durations of a phoneme, mora or word, it is possible that fluent speakers use some acoustic features as cues to control durations. One of the common factors used to compare acoustic features of English and Japanese is the use of fundamental frequency (F0) which manifests the lexical accent for both production and perception. The change in F0 is one of the acoustic features which is recognised and used by speakers of both languages. F0 placement may therefore be used as a cue to time rhythm when speaking Japanese. Though accent placement in a word generally has no relation to the timing features of speech, it may function as a cue to control timing features of Japanese.

The hypotheses of the experiment are:

(i) Fluent speakers will show similar durational features as native speakers but beginners will not.

(ii) Since the durations of phoneme, mora and word are not consciously controllable, fluent speakers may use other features such as an accented vowel as a cue to adjust the timing.

2. THE EXPERIMENTAL METHODS

Fifteen people (5 native speakers of Japanese, 5 native English fluent speakers of Japanese, and 5 native English beginners of Japanese) took part in the recording. Five fluent speakers were native speakers of British English who had spent 3 to 5 years in Japan. All 5 beginners were native speakers of British English who were in the second year at Stirling University in Scotland studying Japanese and had studied Japanese for approximately 350 hours at the time of recordings.

All subjects pronounced Japanese words composed of 1 to 6 morae. 8 test words were chosen for each number of morae, yielding 48 test words. The test words were presented in a random order in a carrier sentence “Kore wa ___ to iimasu” (This is pronounced ___) written in Japanese orthography. Pronunciation was provided for all test words. No kanji characters that the beginners subjects had not studied in the course were used in the text. Most test words contained sequences such as long vowels and geminate consonants, as well as a /CjV/ sequence. A total of 720 samples (48 words x 5 subjects x 3 groups) were collected. The position of lexical
accent varied in each set of 8 test words, from having the accent on the first mora to having no accent.

All words were taken from the two Japanese textbooks that the beginners were using for their studies [4, 10]. The subjects were able to familiarise themselves with the texts and practise if they wished before the recording. The subjects were instructed not to insert a pause between the test word and the carrier sentence.

The recordings took place in an anechoic room and samples were digitised at a sampling rate of 16 KHz, and analysed using Waves+ software on a SUN workstation. Measurements were taken of: (i) duration of a whole word, (ii) duration from the beginning of a word to the beginning of an accented vowel, and (iii) the durational ratio between (i) and (ii) above. The measurement (i) was obtained to examine the acquisition of speech rhythm, because a word is the minimum unit of utterance in which the Japanese speech rhythm is clearly manifested. The measurement (ii) was taken from the point where periodic waveform of the last vowel /a/ of the carrier sentence /korewa/ to the beginning of periodic waveform of the accented vowel of the test word. When the accented vowel followed a stop consonant, the voice onset of the vowel was taken to be the beginning of the vowel. The vowel on which a subject actually put an accent in the recording was taken as 'accented' rather than the vowel which should theoretically be accented.

The durational ratio (iii) was obtained in order to examine the function of lexical accent in terms of duration and timing adjustment and also to normalise individual difference in speaking rate. Since phonetic environments for accented vowels were not controlled in this experiment, it is not appropriate to compare the durations of lexically accented vowels in the test words. Therefore, to measure the effect of lexical accent on the duration, the ratio (iii) above was used.

3. RESULTS

3.1. Entire Word Duration

The complete durations of 720 recorded samples of the test words by 15 subjects were measured and the results of each group were compared with those of the other groups. The word durations were categorised not by the number of morae in a correct pronunciation, but by the number of morae in actual utterances. For example, when a subject mispronounced a 3-mora word as 2-mora, e.g. */kino/ (2 morae) for /kinoo/ (3 morae) (yesterday), it was categorised as a 2-mora word instead of 3-morae. The results of all three groups are presented in Figure 1.

Two factor analysis of variance was performed taking word duration as a dependent variable and the number of morae in a word and fluency as independent variables. The duration of whole words increased with increasing numbers of morae in a word for all 3 groups [F(5,701)=681, p<.001]. The duration of words with the same number of morae were generally longest in beginners' utterances. As shown in Figure 2, in the native speaker group, word duration increased nearly linearly with the number of morae in a word. The rate of increase for fluent speakers was very similar. In the beginner group, however, the overall rate of change was greater than both native and fluent speakers [F(2,720)=99.951, p<.001].

3.2. Manifestation of Lexical Accent

The ratio between (i) duration of a whole word and (ii) duration from the beginning of a word to the beginning of an accented vowel, obtained by (ii) divided by (i), was calculated. Figures 3(a) ~ 2(f) show the mean durational ratios of the three fluency groups for words with different numbers of morae. The results were analysed by the position of the accented mora rather than by accented syllable. This was because, although the accent bearing unit in Japanese is the syllable not the mora [8, 9], accents were often placed on the wrong mora in a syllable, e.g. the accent on the /N/ instead of /hNkii/ 'weather'. No subject placed an accent on the final mora of 4-mora or 5-
mora words in any groups. Therefore, there is no data for the 4th mora or 5th mora in Figures 3(d) and 3(e), respectively.

Non-accented words, such as *koocha* ‘tea’ /kootja/, *fuutoo* ‘envelope’ /fuutoo/, and *kuruma* ‘car’ /kuruma/, were eliminated from the data. When an accented vowel was devoiced in words such as *ki* ‘tree’ /ki/ (to iimasu)/, *yuki* ‘snow’ /juk/ (to iimasu)/, and *su* ‘nest’ /so/ (to iimasu)/, the words were also eliminated from the data. The ratios were categorised by the three fluency groups and the mean ratio of each group was calculated.

The durational ratios showed considerable variation between the three speaker groups. The position of the accented mora significantly affected the durational ratio in words of all numbers of morae (p<0.0001). The effect of fluency was only significant in 1-mora words [F(1,94)=5.85, p<0.005], 2-mora words [F(2,106)=11.858, p<0.0001], and 4-mora words [F(2,106)=8.611, p<0.0005]. In most cases, the native speakers’ ratios were the highest, and the beginners’ ratios were the lowest.
The fluent speakers’ ratios were generally much closer to those of the native speakers than beginners. The interpretation of the ratios in 1-mora words is not straightforward, because in 1-mora words, the ratios are largely dependent upon the combination of phonemes. The results suggest that although fluent speakers failed to place the accent on the correct mora in nearly half of the samples (45.42%), they still controlled the timing feature of the lexical accent in the same way as native speakers.

Beginners also incorrectly placed the lexical accent on the wrong mora in over 56% of samples. However, beginners’ ratios were smaller than those of the other two groups, especially native speakers, for most accent positions in the test word groups of all mora numbers. The results also showed that entire word durations in the beginners groups were much longer than the mora numbers. The results also showed that entire word durations in the beginners groups were much longer than the durations in the other two groups. This indicates that, unlike the fluent speakers, the beginners reduced the duration before the accented vowel or lengthened after the accented vowel.

4. DISCUSSION

These results demonstrated that fluent non-native speakers showed the predicted durational feature of Japanese rhythm. Beginners also increased word duration with increasing morae number, but unlike the other 2 groups, beginners did not show a liner increase of duration.

These results suggest three problems for English speaking beginners of Japanese: (1) identifying morae, (2) pronunciation difficulty, and (3) manifestation of timing.

Problem (1) means that English-speaking beginners have a problem with either identifying the second mora in Japanese syllables /CjV/ and /CjVC/, or if it is identified, pronouncing it fully, so as to adjust word duration in relation to the number of morae in a word [3, 7].

The second problem (2) was a difficulty in pronunciation caused by unfamiliar sequences. The speakers of both non-native groups altered /CjV/ and /CjVC/ sequences, either inserting /i/ between /C/ and /j/ or deleting /i/ from /CjV/ sequence. These mistakes alter the number of syllables and morae in a word and, subsequently, the duration of the word.

The third problem (3) was demonstrated by the ratio between the duration from the beginning of a word to the beginning of an accented vowel and the duration of a whole word. Figures 3(a)–3(f) showed that the ratios of fluent speakers were either close to those of native speakers or came between the ratios of native speakers and beginners in most categories. The native speakers’ ratios were higher than the beginners’ ratios.

This means that the beginners either lengthened the duration of an accented vowel in a word, or reduced the duration before the accented vowel, or did both. In English, a unit of rhythm in terms of duration is a ‘foot’, which usually consists of one strong and one weak syllable. If a syllable does not have an accent, i.e. it is in weak position, its duration is reduced. Transferring English durational strategies, if the beginners interpreted an accented vowel as a strong syllable in a foot, the preceding and following syllables become weak syllables. Therefore, beginners lengthened the strong syllable, i.e. the accented syllable, and shortened the preceding syllables.

Fluent speakers, on the other hand, did not have problems (1) and (3), though some of them did have difficulty in (2) pronouncing /CjV/ and /CjVC/ sequences. They did not necessarily locate an accent in the right syllable, but they were able to control the duration of an accented vowel in a strong syllable, and the durations of unaccented vowels in a weak syllable, not lengthening the former and shortening the latter.

5. CONCLUSIONS

The results indicated that beginners have not yet acquired Japanese speech rhythm. Not only did they have problems in pronouncing certain sequences, but they were also unable to control the effect of the lexical accent on vowel duration. They transferred English accent and duration strategies, lengthening accented vowels and shortening unaccented vowels before the accented vowel in a word. This made it difficult to achieve Japanese speech rhythm.

In contrast, fluent speakers have learned how to control durational features in Japanese, even though they have not fully acquired the prosodic features of Japanese. While they may have placed an accent on the wrong syllable the durations of segments were controlled, so that the duration of the whole word was proportional to the number of morae in that word. This is similar to the way that a native speaker might deliberately pronounce the word with the accent on the same wrong position. This means fluent speakers did not use accented vowels as a cue to time the duration as hypothesised. Instead, they can switch timing manipulation from English to Japanese.

The results implied that the acquisition of speech rhythm is achieved at two separate levels, namely a lower level durational control at the mora level and a higher level control at the word level.

However, in order to understand fully how English speaking learners acquire Japanese speech rhythm, it is essential to investigate other acoustic features of lexical accent, such as amplitude and F0 change, as well as the duration in more controlled environment.

REFERENCES