Quantitative Analysis of English Rhythm Spoken by Japanese Learners ——Focusing on Vowel Duration——[1)]

JUNKO SUGIMOTO

1. Introduction

Rhythm plays one of the key roles in a spoken language. When people study the pronunciation of a foreign language, they usually start with the practice of words, mainly focusing on vowels and consonants, and then move on to bigger units such as sentences. Unfortunately, they often do not have enough chances and time to practice the rhythm of the language. However, the acquisition of rhythm should be one of the most crucial goals because the correct rhythm of a language may facilitate the listener's understanding of the messages.

The term rhythm can be defined as the regular recurrence of beats (cf. Couper-Kuhlen 1993). Isochrony is the key notion in the description of speech rhythm. In the study of rhythm, the binary classification between stress-timed rhythm and syllable-timed rhythm is widely known and acknowledged. Abercrombie (1967: 97) stated that every language in the world is spoken with either of the two kinds of rhythm. In stress-timed rhythm, stresses are said to recur isochronously, whereas in syllable-timed rhythm, syllables are said to recur isochronously.

English is recognized as a representative of stress-timed rhythm. In stress-timed rhythm, stresses recur at approximately equal intervals. When a stressed syllable is followed by several unstressed syllables, unstressed syllables will be compressed. On the other hand, when a stressed syllable is directly followed by another stressed syllable, the syllable will be pronounced relatively slowly. The duration of syllables is controlled in this way so that the isochrony of stresses will be maintained.

In contrast, Japanese belongs to the group of syllable-timed rhythm. Strictly speaking, Japanese is often described as mora-timed rhythm. A mora is a unit that is smaller than a syllable. In mora-timed rhythm, morae recur at approximately equal intervals. In other words, each mora is pronounced with approximately the same duration. Since a mora coincides with a syllable in many cases, mora-timed rhythm can be considered as a subgroup of syllable-timed rhythm.

English and Japanese are described as having different types of rhythm. Many researchers agree and get the impression that rhythm of English and Japanese are different. In the comparison of two languages, however, it is necessary to use an objective criterion rather than relying on impressionistic judgement. Since acoustic instruments such as the sound spectrograph have been available, many researchers have attempted to find physical evidence for the dichotomy of rhythm. They measured inter-stress intervals for stress-timed rhythm, and measured syllable duration for syllable-timed rhythm.

For instance, Roach (1982) measured inter-stress intervals and syllable duration of six languages: three so-called stress-timed rhythm languages, English, Russian, and Arabic, and three so-called syllable-timed rhythm languages, French, Telugu, and Yoruba. Roach's results showed no statistical difference in the inter-stress intervals and syllable duration between stress-timed and syllable-timed rhythm languages. Roach concluded that the distinction between stress-timed and syllable-timed rhythm is based only on auditory and subjective impression.

Dauer (1983) carried out a similar experiment. Dauer measured inter-stress intervals for five languages: English, Thai, Spanish, Italian, and Greek. Among the five languages, English is described as having stress-timed rhythm. However, Dauer's results showed that inter-stress intervals of English were not significantly more isochronous than other languages. Many acoustic studies failed to quantitatively prove the difference between stress-timed and syllable-timed rhythm. Then, where does this difference in rhythm among languages come from?

People are said to have a perceptual tendency towards isochrony. Allen

(1975: 76) discovered that people have a tendency to underestimate the duration of long intervals and overestimate that of short ones in order to balance the intervals and perceive them as isochronous. Similarly, Lehiste (1979) found out that when people listen to speech, they tend to hear stresses as more isochronous than they really are. Because of these reasons, the notion of isochrony is still considered as good evidence for the dichotomy of rhythm. Nevertheless, is it truly impossible to find the physical criterion to describe speech rhythm? There is the need for the method to quantify rhythm of various languages.

2. Pairwise Variability Index

The Pairwise Variability Index (PVI) is one method which makes it possible to compare rhythmic differences across languages objectively. It was devised by researchers at the University of Cambridge. PVI focuses on the difference of duration between adjacent vowels. If adjacent vowels are similar in duration, PVI value will come out small. On the other hand, if the duration of adjacent vowels differs greatly, PVI value will come out large. PVI can be interpreted as "Rhythm Index" because vowel duration is strongly related to rhythm. When a speaker wants to control the duration of a syllable, s/he accomplishes it by controlling the vowel duration. If successive vowels are similar in duration, it suggests that successive syllables are similar in duration. Thus, a smaller PVI value corresponds to syllable-timed rhythm. On the other hand, a larger PVI value corresponds to stress-timed rhythm because the duration of successive vowels tend to vary in order to keep the isochrony of stresses.

A number of studies have been conducted using PVI by the researchers at the University of Cambridge. The first attempt was to compare rhythm between accents of English. Low *et al.* (2000) used PVI to observe the rhythmic difference between British English and Singapore English. Singapore English is said to have syllable-timed rhythm. Low *et al.* prepared two sets of sentences: F set and R set. F set is the sentences whose vowels are all full vowels. On the other hand, R set is the sentences in which a full vowel and a reduced vowel appear alternatively. The important point is that reduced vowels are shorter than full vowels in

English. Low *et al.* (2000)'s results showed that PVI of F set was low for both British and Singapore English, whereas PVI of R set differed between British and Singapore English. PVI of R set was high for British but not for Singapore English, meaning that Singapore English speakers pronounced reduced vowels longer than they should be. This is the reason why Singapore English is perceived as having syllable-timed rhythm. Nolan and Hoshino (2000) compared British English and Caribbean English by using exactly the same procedure. They found out that the R set of Caribbean English showed lower PVI, which confirms that Caribbean English is often perceived as having syllable-timed rhythm.

PVI can be applied not only to the comparison between accents of English but also to the comparison between languages. Grabe *et al.* (1999) compared the acquisition of rhythm between English and French by using PVI. They measured vowel duration of French utterances and calculated PVI. Their results revealed that French has lower PVI than English, which supports the idea that French has syllable-timed rhythm. Results of the three works are summarized in Figure 1 below.

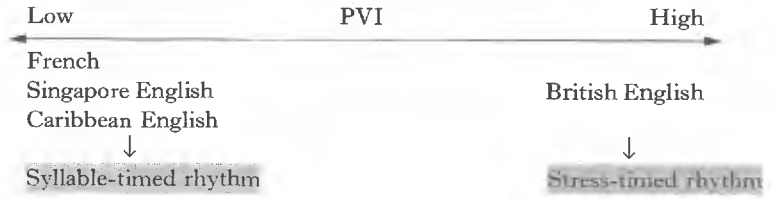

Figure 1 PVI comparison between stress-timed and syllable-timed rhythm

PVI has a number of methodological advantages over the measurement of inter-stress intervals and syllable duration. One of the difficulties in measuring inter-stress intervals is the identification of stresses. Stresses are realized differently among languages. For instance, in English, stresses are realized by the combination of pitch, intensity, and duration. Hearers can only recognize that a specific syllable is stressed by comparing it with other syllables. In addition, although there are some basic rules, the locations of stresses are not fixed in English sentences. Because of these

reasons, the identification of stresses is very difficult.

Syllable boundaries are also difficult to locate. The syllabification rule of a language differs depending on researchers. Even if you agree on the number of syllables and their boundaries on the phonemic level, it is often difficult to locate the boundaries on the phonetic level because there is combination of sounds which is difficult to segment. In the calculation of PVI, the only requirement is the measurement of vowel duration. Unlike stresses or syllable boundaries, vowels are easy to identify. The demarcation of vowels is generally easier and simpler than the demarcation of syllables. It is also relatively easy to control the phonetic contexts of vowels. Most of all, PVI enables us to measure and calculate the rhythm index by using exactly the same criterion for all languages. Therefore, PVI is a very useful method to quantify rhythm of various languages.

In this paper, the same procedure as Low *et al.* (2000) will be used to analyze rhythm of Japanese and English spoken by Japanese learners. The purposes of the experiment are (i) to compare rhythmic differences of Japanese and English by using PVI, and (ii) to investigate the characteristics seen in the rhythm of English spoken by Japanese by comparing it with English spoken by British speakers.

3. Materials and methods

3.1. Sentences used in the experiment

English vowels are divided into two groups: strong vowels and weak vowels. Strong vowels of English include short vowels, long vowels, and diphthongs. Long vowels and diphthongs have longer duration than short vowels when compared in the same phonetic contexts. Weak vowels include /ɪ/ and /ʊ/ as well as /ə/, /i/ and /u/. In contrast, Japanese has only five vowels all of which are short vowels. Thus, English has a wider variety of vowels than Japanese.

As for English sentences used for this experiment, eight sentences were used in total. According to the study by Low *et al.* (2000) done for Singapore English, English sentences were divided into two groups: S set and S-W set[2]. All vowels of the four sentences in S set are strong vowels. On the other hand, the vowels of the four sentences in S-W set were made

of alternations of strong and weak vowels. Each English sentence consisted of eight vowels.

As for Japanese sentences, eight sentences were used. Since Japanese does not have the distinction between strong and weak vowels as English, eight Japanese sentences cannot be divided into subgroups. Each Japanese sentence consisted of twelve vowels. In Japanese, high vowels, /i/ and /u/, are often phonologically devoiced between voiceless consonants. Devoiced vowels are phonetically realized in different ways: they may be devoiced, be shorter in duration, or be completely deleted. Devoiced vowels are usually very difficult to identify and segment. In addition, the degree of vowel devoicing varies according to various factors such as speakers, accents, rate of speech, and style. To avoid such difficulties, the phonetic contexts which induce vowel devoicing were excluded.

For both English and Japanese, grammatically and semantically simple sentences were selected. Also, all sentences were relatively short so that speakers could read each sentence in one breath group. In order to measure the duration of vowels, segmentation is crucial. Sequences of vowels were avoided because it is often impossible to locate the boundary between two vowels. For the same reason, a sequence of a vowel and a semivowel, or a sequence of a semivowel and a vowel were avoided as far as possible. All sentences used in the experiment are listed in Table 1.

Table 1 Sentences used in the experiment

(a) English sentences

Set	Code	Sentences
S set	S1	This big chair takes up too much space.
	S2	I know Ted came back late last night.
	S3	I'd like fourteen first-class stamps, please.
	S4	Miss Smith baked cakes on my birthday.
S-W set	S-W1	Everybody skates in winter.
	S-W2	Jack believes in ghosts and fairies.
	S-W3	John decided not to buy them.
	S-W4	Nancy says she wants to see him.

(b) Japanese sentences (with English translation)

Code	Sentences (Japanese/English translation)
J1	*Asano sanpoga sobono nikkada.* (My grandmother's daily routine is to take a walk in the morning.)
J2	*Kadono misega kujini shimaru.* (The shop at the corner closes at nine.)
J3	*Kagono nakamiha kudamonoda.* (There are fruits inside the basket.)
J4	*Sorosoro shingakkiga hajimaru.* (The new term starts soon.)
J5	*Shimaneni sumukotoni kimeta.* (I decided to live in Shimane.)
J6	*Aneha sakunen hatachini natta.* (My elder sister became twenty last year.)
J7	*Denwaga nattanode mega sameta.* (I woke up because the telephone rang.)
J8	*Kazokude sapporoni dekakeru.* (Our family will go to *Sapporo*.)

3.2. Subjects

Two groups of English and Japanese speakers were asked to participate in the recording. Each group consisted of five speakers: two female speakers and three male speakers. The first group was native speakers of British English (BF1, BF2, BM1, BM2 and BM3). All speakers come from Southeast area of England. The average age of this group was 31.8.

The second group was native speakers of Japanese (JF1, JF2, JM1, JM2 and JM3). All speakers come from Tokyo area of Japan, and had studied English for more than ten years in Japan. At the time of recording, they had lived in London for more than ten months but less than twenty-two months. They had no previous experience of living abroad. Although they did not have difficulty in communicating in English, all speakers were judged to have a strong Japanese accent by the author. The average age of this group was 29.4. Japanese subjects were asked to read both English and Japanese sentences.

3.3. PVI formula

PVI was calculated for (i) Japanese sentences spoken by Japanese, (ii) English sentences spoken by Japanese, and (iii) English sentences spoken by British. For English sentences, PVI was calculated for S set and S-W set separately for both Japanese and British speakers.

The formula of PVI is given below (Low *et al.* 2000: 383):

$$\mathrm{PVI} = 100 \times \left[\sum_{k=1}^{m-1} \left| \frac{d_k - d_{k+1}}{(d_k + d_{k+1})/2} \right| / (m-1) \right]$$

(where m = number of vowels in utterance;
d = duration of the k^{th} vowel)

The formula shows that PVI is calculated by taking the durational difference between successive vowels, and the mean duration of the successive vowels. The durational difference between successive vowels is divided by their mean duration. The absolute values are summed and divided by the number of differences. The formula normalizes the speaking rate of speakers. Therefore, the difference of speaking rate among speakers will not interfere with the results. In order to make the final PVI value easier to handle, the output is multiplied by 100 (cf. Grabe *et al.* 1999; Low *et al.* 2000).

3.4. Recording

Recording session was carried out in an anechoic room of the Department of Phonetics and Linguistics at University College London. The sentences read by speakers were recorded onto DAT tape.

Enough time to read through the sentences was given to all speakers before the recording. The speakers were instructed to read each sentence in one breath group without a pause or hesitation in between, and to read in comfortable speaking rate. They were also instructed to read each sentence without imagining any specific context. If a subject made a mistake or violated the instruction, s/he was asked to repeat the sentence from the beginning.

Speakers pronounced the words and read the sentences in different ways. One case is the realization of vowels. The examples are "chair" ([tʃeə] or [tʃeː]) and "believe" ([bə'liːv] or [bɪ'liːv]). The qualities of vowels used by Japanese speakers when they read English sentences were different from those used by British speakers. Another case is the difference in use of an intonation pattern in few sentences. Such inter-speaker and inter-group realizational differences of vowels and intonation were ignored because the main focus of this experiment is to measure vowel duration.

3.5. Measurements

The data was transferred onto a PC and analyzed by the SFS program (Speech Filing System © University College London). The vowel duration was established by using the measurement box. The wide-band spectrogram was mainly used, but also the narrow-band spectrogram and the amplitude waveform were used when they were necessary.

Whether the segmentation is going to be easy or difficult depends on the types of consonants which surround the vowel. The most important point in segmentation is to use the consistent criteria for all sentences and for all speakers. The most difficult groups of consonants to segment were liquids and semivowels. In both cases, the most noticeable cue of segmentation was the energy drop. The finishing point of formant transition was also used as a cue where it was observed. The criteria of segmentation were mainly based on Peterson and Lehiste (1960).

Two Japanese speakers inserted an extra vowel incorrectly. Since the main point is to observe how Japanese speakers control English vowels, these inserted vowels were not taken into consideration. Even though the major vowel devoicing contexts were avoided, vowel devoicing and vowel

elision were still observed. They were excluded from the calculation of PVI.

Grabe *et al.* (1999) excluded the final vowel in the calculation of PVI because they wanted to remove the influence of final syllable lengthening. In this study, two versions of PVI were calculated for all sentences. The difference between PVI with all vowels and PVI without the final vowels was very small. Statistically, there was no significant difference between PVI with all vowels and PVI without the final vowel ($p \geq 0.5$). Thus, the influence of final syllable lengthening was not observed in this experiment. In this study, PVI with all vowels were used.

4. Results and discussion

4.1. Hypotheses

From the descriptions and previous works on rhythm, some predictions can be made. The following four hypotheses are presented with brief explanations.

<Hypothesis 1> Japanese sentences will show lower PVI than English sentences. Since mora-timed rhythm of Japanese is a subgroup of syllable-timed rhythm, Japanese is predicted to show lower PVI value than English which is stress-timed.

<Hypothesis 2> PVI of English spoken by Japanese will be higher than that of Japanese but lower than that of English spoken by British. PVI of English spoken by Japanese is predicted to show influence of Japanese rhythm to a certain extent.

<Hypothesis 3> PVI of English will differ between S set and S-W set for British speakers but not for Japanese speakers. For British speakers, PVI of S-W set should be higher than that of S set. On the other hand, for Japanese speakers, PVI of S-W set and S set may be the same because Japanese does not have weak vowels.

<Hypothesis 4> For English, Japanese speakers' PVI of S-W set will be lower than British speakers'. British speakers' PVI of S-W set is expected to be high because weak vowels are shorter than strong vowels. In contrast, Japanese speakers' PVI of S-W set may not be high as British speakers' because Japanese does not have weak vowels.

4.2. A comparison using all sentences: Japanese vs. English spoken by Japanese vs. English spoken by British

Table 2 summarizes PVI of Japanese and English sentences used in this study. Each speaker's PVI is shown by the average of all sentences. Figure 2 shows the comparison among PVI of Japanese (40.3 ± 3.5), English spoken by Japanese (41.6 ± 6.5) and English spoken by British (56.7 ± 4.3).

Table 2 PVI of Japanese, English by Japanese, and English by British

Japanese Speakers			British Speakers	
Code	PVI of Japanese (by Japanese)	PVI of English (by Japanese)	Code	PVI of English (by British)
JF1	39.5	36.6	BF1	60.1
JF2	44.2	53.2	BF2	58.5
JM1	34.7	35.2	BM1	51.7
JM2	39.3	39.0	BM2	61.9
JM3	43.8	43.8	BM3	51.5
Average	40.3	41.6	Average	56.7
SD	±3.5	±6.5	SD	±4.3

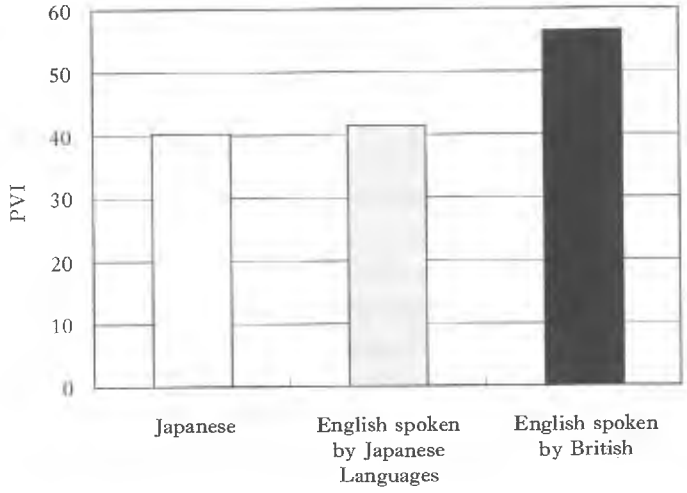

Figure 2 PVI comparison among Japanese, English spoken by Japanese, and English spoken by British

Firstly, speech rhythms between Japanese spoken by Japanese and English spoken by British were compared. PVI of Japanese (40.3 ± 3.5) was significantly lower than that of English spoken by British (56.7 ± 4.3) (t test, $p \leq 0.001$). In other words, the vowel duration varies more in English than in Japanese. This result supports Hypothesis 1: Japanese will show lower PVI than English, meaning that Japanese is more syllable-timed than English. The difference in PVI value reflects the difference of rhythm between English and Japanese. The perceptual and subjective impression that English and Japanese have different kinds of rhythm is supported quantitatively.

Secondly, the characteristics of rhythm of English spoken by Japanese can be investigated by comparing it to that of English spoken by British. PVI values were from small to large in the order of Japanese, English spoken by Japanese, and English spoken by British. Although the PVI of English spoken by Japanese was slightly higher than that of Japanese, no significant difference between the two existed (t test, $p \geq 0.5$). In contrast, the PVI of English spoken by Japanese was significantly lower than that of English spoken by British (t test, $p \leq 0.001$). Interestingly, English spoken by Japanese is more similar to Japanese than English spoken by British as far as PVI is concerned.

These results do not support Hypothesis 2: PVI of English spoken by Japanese will be higher than that of Japanese but lower than that of English spoken by British. In fact, there was no statistical difference between PVI value of Japanese and that of English spoken by Japanese. In foreign language acquisition, the native language of the learners often influences the target language. Similar levels of PVI of Japanese and English spoken by Japanese strongly reflect the transfer of Japanese rhythm to English by Japanese learners.

4.3. A comparison focusing on S set and S-W set of English

The rhythm of Japanese and English were compared in more detail by calculating PVI separately for S set and S-W set of English sentences. Table 3 summarizes British speakers' and Japanese speakers' PVI of S set and S-W set.

Table 3 PVI of S set and S-W set of English spoken by British and by Japanese

English spoken by British			English spoken by Japanese		
Speaker	S set	S-W set	Speaker	S set	S-W set
BF1	36.5	83.8	JF1	29.0	44.3
BF2	31.2	85.7	JF2	43.3	63.1
BM1	29.7	73.6	JM1	25.6	44.7
BM2	38.5	85.3	JM2	32.4	45.6
BM3	32.2	70.7	JM3	35.2	52.3
Average	33.6	79.8	Average	33.1	50.0
SD	±3.3	±6.4	SD	±6.0	±7.2

4.3.1. Japanese vs. S set and S-W set of English spoken by British

Figure 3 shows the comparison among PVI of Japanese (40.3 ± 3.5), S set by British speakers (33.6± 3.3), and S-W set by British speakers (79.8± 6.4).

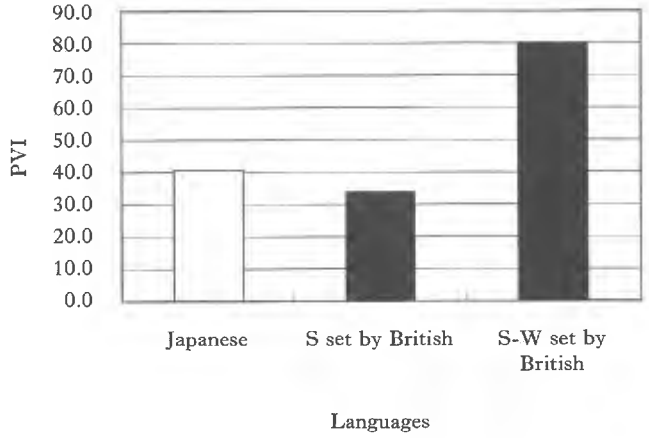

Figure 3 PVI comparison among Japanese, S set, and S-W set of English spoken by British

PVI was from large to small in the order of S-W set, Japanese, and S set. An interesting point is revealed when PVI of S set and S-W set were compared respectively with that of Japanese. PVI of S set by British speakers (33.6 ± 3.3) was significantly lower than that of Japanese (40.3 ± 3.5)

(t test, $p \leq 0.025$), whereas that of S-W set (79.8 ± 6.4) was significantly higher than that of Japanese (t test, $p \leq 0.001$).

The results point out that Japanese sentences are similar to S set of English, but are quite different from S-W set of English with regard to PVI. In other words, the rhythm of Japanese and English are similar if English sentences consist of strong vowels alone. The fact that PVI of S set was similar to that of Japanese is an interesting point. Japanese only has five short vowels. On the other hand, English has a wider variety of strong vowels which include not only short vowels but also long vowels and diphthongs. Therefore, the duration of vowels among English may be expected to differ more greatly than Japanese. However, in fact, PVI of S set of English was lower than that of Japanese. On the other hand, PVI of S-W set of English was much higher than that of Japanese. The rhythmic difference between Japanese and English becomes prominent especially when both strong and weak vowels are used. Weak vowels are considered to be the key to differentiate between Japanese rhythm and English rhythm. How Japanese speakers pronounce weak vowels in S-W set is an important point.

PVI of S set and S-W set of English spoken by British were also compared in Figure 3. PVI of British speakers' S-W set (79.8 ± 6.4) was significantly higher than that of S set (33.6 ± 3.3) (t test, $p \leq 0.001$). This result is consistent with that of Low *et al.* (2000) and of Hoshino and Nolan (2000). British speakers' PVI of S-W set was higher because the duration of weak vowels are notably shorter than that of strong vowels in English.

4.3.2. Japanese vs. S set and S-W set of English spoken by Japanese

Figure 4 shows the comparison among PVI of Japanese (40.3 ± 3.5), S set by Japanese speakers (33.1 ± 6.0), and S-W set by Japanese speakers (50.0 ± 7.2).

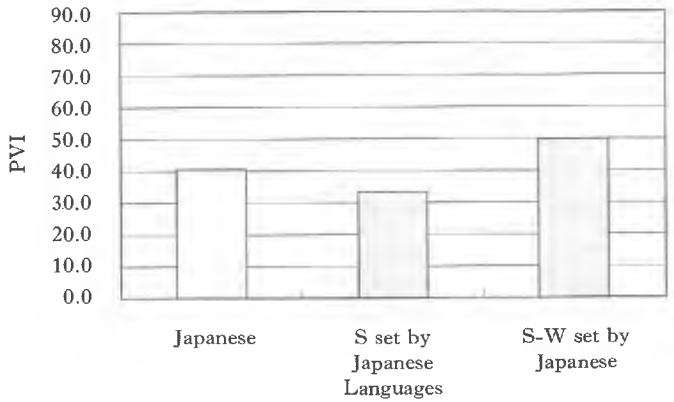

Figure 4 PVI comparison among Japanese, S set, and S-W set of English spoken by Japanese

PVI was from large to small in the order of S-W set, Japanese, and S set. PVI of Japanese speakers' S-W set (50.0 ± 7.2) was significantly higher than that of Japanese (40.3 ± 3.5) (t test, $p \leq 0.025$), whereas there was no significant difference between Japanese speakers' PVI of S set and that of Japanese (t test, $p \geq 0.100$). Furthermore, Japanese speakers' S-W set (50.0 ± 7.2) was significantly higher than that of S set (33.1 ± 6.0) (t test, $p \leq 0.01$).

These results only partly support Hypothesis 3: PVI will differ between S set and S-W set for British speakers but not for Japanese speakers. PVI of S set and S-W set differ not only for British speakers but also for Japanese speakers. In the same way as British speakers, Japanese speakers' PVI of S-W set was higher than that of S set. Since Japanese does not have weak vowels, Japanese speakers may have difficulties in pronouncing weak vowels. Nevertheless, the high PVI of S-W set by Japanese speakers proves that they are making an effort to reduce the duration of weak vowels as native speakers do. As in British speakers, PVI of S-W set was higher than that of Japanese, and PVI of S set was actually lower than that of Japanese although there was no significant difference.

4.3.3. S set and S-W set by British vs. S set and S-W set by Japanese

Finally, PVI of S set and S-W set between Japanese speakers and British speakers were compared. Figure 5 shows PVI of S set and S-W set of Japanese and British speakers.

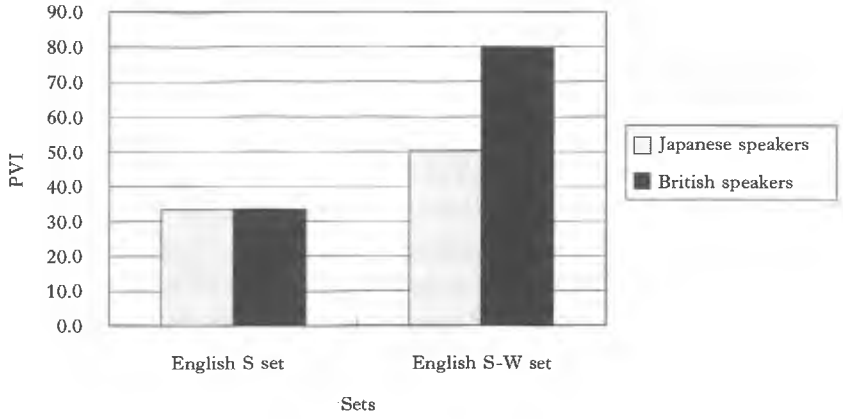

Figure 5 PVI comparison of S set and S-W set between Britisn and Japanese speakers

Interestingly, as for S set, there was not much difference between the two groups of speakers. Though Japanese speakers' PVI (33.1 ± 6.0) was slightly lower than British speakers' (33.6 ± 3.3), there was no significant difference (t test, $p \geq 0.5$). As for the comparison of S-W set, British speakers' PVI of S-W set (79.8 ± 6.4) was significantly higher than Japanese speakers' (50.0 ± 7.2) (t test, $p \leq 0.001$).

These results support Hypothesis 4: Japanese speakers' PVI of S-W set will be lower than British speakers'. Although Japanese speakers' PVI of S-W set was higher than their PVI of S set, it was not high enough to reach the level of British speakers'. In other words, although Japanese speakers are making an effort to reduce the duration of weak vowels, their effort is not enough. In spite of the difference between Japanese and English vowels, as far as the durational variation is concerned, Japanese speakers seem to have no problem in the pronunciation of strong vowels because PVI value of S set was not different between Japanese and British.

The results altogether point out that the weak point of Japanese speakers is the pronunciation of weak vowels.

4.4. Characteristics of English rhythm spoken by Japanese

So far, PVI comparison was made only between the averages of speaker-groups or between S set and S-W set. However, when the average is compared, the inter-speaker differences are ignored. Each speaker has his/her individual tendency and weak points. This section focuses on such inter-speaker variations. Inter-speaker variation of English spoken by Japanese was observed in detail. The control of vowel duration of S-W set is focused in this section because it appears to be the weak point of Japanese speakers when speaking English.

Among Japanese speakers, JF2 showed the highest PVI (average 60.7) and JF1 showed the lowest PVI (average 42.6) for S-W set among all Japanese speakers. This result indicates that JF2's PVI of S-W set is the closest to British speakers' (average 79.8). Does this mean that JF2 is good at controlling the vowel duration like British speakers? The sentences S-W3 is picked up and the duration of vowels are observed in detail below.

Table 4 shows the measured vowel duration of S-W3 by British speakers (average), JF1, and JF2. Table 4 is transformed to Figure 6. The vowel duration is shown as a percentage by taking the total vowel duration of the sentence as 100% to avoid that the difference of speaking rate among speakers affecting the comparison.

Table 4 Vowel duration of S-W3 "John decided not to buy them." by British speakers, JF1 and JF2

	/ɒ/	/ɪ, ə/	/aɪ/	/ɪ, ə/	/ɒ/	/ə/	/aɪ/	/ə/
British speakers	0.1152*	0.0412	0.1336	0.0506	0.0820	0.0380	0.1978	0.0591
JF1	0.1448	0.0573	0.1318	0.0753	0.1188	0.1008	0.1772	0.1783
JF2	0.1591	0.0380	0.1034	0.0955	0.1006	0.0853	0.1790	0.2741

*second

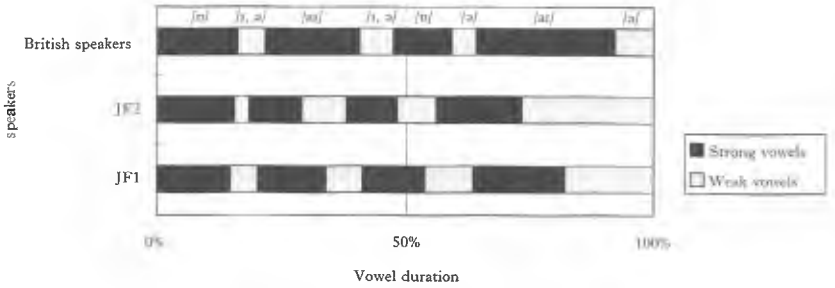

Figure 6 Comparison of vowel duration of S-W3 "John decided not to buy them."

British speakers' average clearly showed the alternation of long strong vowels and short weak vowels. Strong vowels were apparently longer than weak vowels. As a result, British speakers' PVI value became large. JF1 and JF2 showed some characteristics seen in rhythm of English spoken by Japanese. One characteristic is that there is not much difference in duration between strong and weak vowels. In JF1's case, a strong vowel /ɒ/ of "not" and a weak vowel /ə/ of "to," a strong vowel /aɪ/ of "buy" and a weak vowel /ə/ of "them" were similar in duration. In the same way, in JF2's case, a strong vowel /aɪ/ and a weak vowel /ɪ, ə/ of "decided," and a strong vowel /ɒ/ of "not" and a weak vowel /ə/ of "to" were similar in duration. Another characteristic of Japanese English rhythm is that the last weak vowel is very long. For both JF1 and JF2, the last weak vowel /ə/ of "them" was longer than the preceding strong vowel /aɪ/ of "buy." In English, a function word such as "them" is usually not stressed. Since JF1 and JF2 pronounced a weak vowel of "them" longer than it should be, it may sound unnatural to native speakers of English. These results also indicate that the similarity in PVI values between British speakers (average) and JF2 did not truly reflect the actual duration of vowels. Since PVI focuses on the difference of duration between successive vowels, PVI of S-W3 does not differ much between that of JF2 and that of British speakers, in spite of this important difference. This is a weak point of PVI method because it only focuses on the absolute value of the difference of duration between successive vowels: even if the duration of two vowels is reversed,

PVI values could be the same.

5. General discussion

Rhythm of languages has long been considered to consist two distinct categories: stress-timed and syllable-timed rhythm. By using PVI, speech rhythm can be represented as a continuum. The next important step is to plot each language along the continuum. PVI makes it possible to quantify rhythm of languages using one objective criterion: the durational variation of vowels. In case of English and Japanese, PVI values were significantly far apart along the rhythm continuum.

If an English sentence happens to consist only of strong vowels, as in S set of the experiment, that English sentence will have rhythm similar to syllable-timed rhythm. In fact, PVI of S set spoken by British speakers showed close resemblance to PVI of Japanese. Japanese speakers' PVI of S set was not much different from British speakers'. In other words, Japanese speakers were good at pronouncing syllables of approximately equal length, which they are used to doing when they speak their native language. In reality, sentences which only consist of strong vowels are rare in English since the frequency of weak vowels is very high.

Vowels are one of the major factors which characterize rhythm of languages. In other words, the inventory and distribution of vowels affect the rhythm of a language. Dauer (1983) explained that vowel reduction is one of the important phonological features which determine rhythm of a language. If a language has vowel reduction, it tends to have stress-based rhythm. On the other hand, if a language does not have vowel reduction, it tends to have syllable-based rhythm. English and Japanese match with this description. English has weak vowels and Japanese does not. Does a language which has stress-timed rhythm happen to have weak vowels, or does a language which has weak vowels in its vowel system happen to have stress-timed rhythm? It may be a circular argument. One thing which is clear is that weak vowels are the ones which make English rhythm sound very different from Japanese rhythm.

As far as PVI is concerned, Japanese learners had no problem in S set but in S-W set. Therefore, Japanese learners' weak point is to control the duration of weak vowels differently from that of strong vowels. Japanese learners need to study what weak vowels are, and where they are used in English sentences. For instance, they must practice the stress pattern of polysyllabic words, and study the division between function and content words. In this experiment, Japanese learners' effort to reduce the duration of weak vowels was recognized. However, their effort was not enough. They must bear in mind that English weak vowels are even shorter than they think they are.

In spoken English, the distribution of strong and weak vowels is not regular but is very complicated. Japanese learners can start practice by using the controlled sentences which have a regular alternation of strong and weak vowels. Some examples are given below (S = strong vowel; W = weak vowel).

SW	Jane and Lucy go to school together.
WS	She studies English every week.
SWW	This is a present for Jonathan.
WWS	Can you tell me the way to the library?

Duration of vowels is strongly related to stresses. If the duration of a vowel is longer than surrounding vowels, that syllable tends to be perceived as stressed. If a Japanese learner pronounces a weak vowel longer than it should be, the native speakers of English may perceive the syllable as stressed. The correct control of vowel duration may lead to the correct realization of stresses, thus, to the correct acquisition of English stress-timed rhythm.

To summarize, the rhythm of English spoken by Japanese was studied using the Pairwise Variability Index (PVI) which reflects the durational difference of successive vowels. The results of the experiment indicate that Japanese speakers have a problem in controlling the duration of weak vowels in contrast with strong vowels. The lack of weak vowels in the vowel system of Japanese is considered to be one of the main causes of this problem. The durational control of weak vowels could be the key to improve English rhythm of Japanese learners.

REFERENCES

Abercrombie, D. 1967. *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Allen, G. D. 1975. "Speech rhythm: its relation to performance universals and articulatory timing." *Journal of Phonetics* 3: 75–88.

Beckman, M. 1982. "Segment duration and the 'mora' in Japanese." *Phonetica* 39: 113–35.

Couper-Kuhlen, E. 1986. *An introduction to English prosody*. London: Edward Arnold.

Couper-Kuhlen, E. 1993. *English Speech Rhythm: Form and Function in Everyday Verbal Interaction*. Amsterdam: John Benjamins Publishing Company.

Darwin, C. J. and A. Donovan. 1979. "The perceived rhythm of speech." *Proceedings of the ninth International Congress of Phonetics Sciences*. vol. 2 268–74.

Darwin, C. J. and A. Donovan. 1980. "Perceptual studies of speech rhythm: isochrony and intonation." In J. C. Simon ed. *Spoken Language Generation and Understanding*. Dordrecht: D. Reidel Publishing Company.

Dasher, R. and D. Bolinger. 1982. "On pre-accentual lengthening." *Journal of the International Phonetic Association* 12: 58–69.

Dauer, R. M. 1983. "Stress-timing and syllable-timing reanalyzed." *Journal of Phonetics* 11: 51–62.

Gimson, A. C. 1994. *Pronunciation of English*. 5th ed. Revised by Alan Cruttenden. 1st ed. in 1962. London: Edward Arnold.

Grabe, E., B. Post, and I. Watson. 1999. "The acquisition of rhythmic patterns in English and French." *Proceeding of 14th International Congress of Phonetic Sciences*. 1201–1204.

Hoequist, Charles Jr. 1982. "The perceptual center and rhythm categories." *Language and Speech* 26: 367–76.

Huggins, A. W. 1979. "Some effects on intelligibility of inappropriate temporal relations within speech units." *Proceedings of the Ninth International Congress of Phonetic Sciences*. vol. 2. 283–9.

Laver, J. 1994. *Principles of phonetics*. Cambridge: Cambridge University Press.

Lehiste, I. 1977. "Isochrony reconsidered." *Journal of Phonetics* 5: 253–63.

Low, E. L., E. Grabe and F. Nolan. 2000. "Quantitative Characterizations of Speech Rhythm: Syllable-Timing in Singapore English." *Language and Speech* 43(4): 377–401.

Mortimer, C. 1985. *Elements of Pronunciation: Intensive Practice for Intermediate and More Advanced Students*. Cambridge: Cambridge University Press.

Nespor, M. 1990. "On the rhythm parameter in phonology." In I. M. Roca ed. *Logical Issues in Language Acquisition*. Dordrecht: Foris Publications. 157–175.

Nolan, F. and N. Hoshino. 2000. "Rhythm in Caribbean English." BAAP (British Association of Academic Phoneticians) Colloquium at Glasgow.

Peterson, G. E. and I. Lehiste. 1960. "Duration of syllable nuclei in English." *The Journal of the Acoustical Society of America* 32: 693–703.

Ramus, F., M. Nespor, and J. Mehler. 1999. "Correlates of linguistic rhythm in the speech signal." *Cognition* 73: 265–292.

Roach, P. 1982. "On the distinction between 'stress-timed' and 'syllable-timed' languages." In D. Crystal ed. *Language Controversies*. London: Edward Arnold.

Studdert-Kennedy, M. 1979. "Speech perception." *Proceedings of the ninth International Congress of Phonetic Sciences*. vol. 1. 59–81.

NOTES

1) This paper is a slightly revised version of my MA thesis submitted to University College London in 2000. I am very grateful to my supervisor Mr. John Maidment for his suggestions and comments on my original thesis.

2) S set in this experiment corresponds to F set, and S-W set corresponds to R set of Low *et al.*'s (2000) experiment.