Training Japanese Students to Recognize and Produce English Syllables

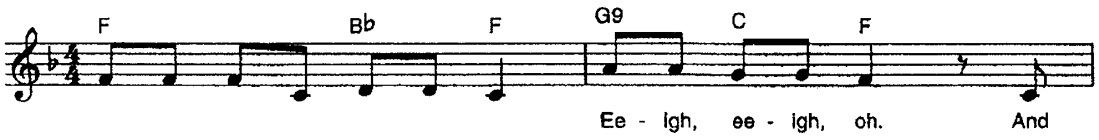
ISHIKAWA, Keiichi
Kyoto Women's University

Abstract

This study investigated how Japanese college students recognize and produce English syllables and the effects of syllable-counting training on their ability to deal with syllables. Participants improved their ability to identify syllables in English words as a result of training. Their recognition of syllables was affected by the number of syllables in words, which conceivably shows the influence of their use of moras, epenthetic vowels, bimoraic foot, and their perception of unstressed syllables. The students' ability to segment English words into syllables and to match syllables to notes in English songs was also assessed before and after training. They showed significant improvement in song performance. The results suggest that training in counting syllables is effective in improving the recognition and the production of English syllables by Japanese students.

1. Introduction

One common observation made about English spoken by Japanese students is their inappropriate timing of syllables. It is widely acknowledged that the syllable serves as the basic unit in the recognition and the production of spoken English, and that it plays an important role in forming the speech rhythm of English (e.g., Giegerich, 1992; Roach, 1991). It takes sustained effort for nonnative speakers of English to master appropriate timing and the length of English syllables (Bond & Fokes, 1985; Faber, 1986; Fokes & Bond, 1989). English syllables also seem to be very difficult to deal with for Japanese learners of English. Japanese students often have trouble singing English songs, such as, for example, when trying to sing a missing part of the following nursery rhyme with the words "Old MacDonald had a farm":



Another case showing the difficulty of identifying syllables is that experienced when Japanese people are confused about the number of syllables in a word given as a clue to the answer while playing a game like "Charades". Suppose a Japanese speaker of English has to guess the name of a sport and another player makes a batting gesture and that of a squashy ball. The Japanese player may think of the correct answer, but after he/she is told that the name has two syllables, he/she is likely to feel confused and reject their first and correct answer: "softball". According to a study by Liberman, Shanlweiler, Fischer, &

Carter (1974), 90% of the American 6-year-old children can count syllables in English words correctly.

Phonological differences in syllable structure between English and Japanese can be the main reason for Japanese people's lack of ability to deal with English syllables. English permits a wide variety of syllable shapes such as CV (consonant-vowel; e.g., "boy"), CVC (e.g., "bat"), CCVCC (e.g., "friend"), and CCCVCCCC (e.g., "strengths"), whereas Japanese syllables are highly constrained, allowing only three types: V (e.g., "i"), CV (e.g., "ta"), and CVC (e.g., "sen"). Japanese-speaking people seem to rely on moras rather than syllables when segmenting speech sound (Otake, Hatano, Cutler, & Mehler, 1993). Moras are subsyllabic units of which there are four types: a short vowel optionally preceded by a consonant (e.g., /a/, /ki/), the second part of a diphthong or a long vowel (e.g., /i/ of /tai/ or the second /i/ of /shii/), the first part of a geminate (e.g., the first /t/ of /kitte/), and the syllable final nasal (e.g., /n/ of /tonchi/). Moreover, Japanese speakers of English often insert vowels after consonants (Dopoux, Kakehi, Hirose, Pallier, & Mehler, 1999). For example, they tend to utter "su-to-rai-ku" for "strike" or "to-ra-bu-ru" for "trouble".

The syllable-counting training reported in Tajima and Erickson (2001) and Ishikawa (2002, 2003) requires participants to count the number of syllables in English words and nonwords spoken by native speakers of English. This technique has been shown to be very effective in improving Japanese listeners' ability to count English syllables and to segment words on the basis of syllables rather than moras. It can be expected that the syllable-counting task can train Japanese students to recognize and produce English words on the basis of syllables.

In English songs musical notes usually correspond to syllables. On the other hand, Japanese songs have moras matched frequently to musical notes (Kubozono, 1999; Vance, 1987). It has been shown that in English songs linguistic rhythm and syllables are closely related to musical rhythm and notes (e.g., Palmer & Kelly, 1992). This observation can lead us to consider that having Japanese students sing English songs can test their ability to produce English words based on syllabic units.

The purpose of the present study is twofold. One is to investigate how Japanese college students recognize English syllables and what effect syllable-counting training has on their ability to count syllables in English words. The other is to examine the extent to which Japanese students improve their ability to recognize and produce English syllables by assessing whether they can divide words in English songs into syllables and whether they can assign one syllable to one musical note when they sing.

2. Method

2.1. Procedure

The participants were 20 Japanese students of English at Kyoto Women's University. The students received course credit in return for their participation. Experiments were conducted individually in a quiet room observing the following steps.

(1) Participants were told that a syllable is a unit typically comprised of one vowel, optionally preceded or followed by consonants; the examples given were *pen* as a one-

syllable word and *father* as a two-syllable word. Next, they were shown an example indicating how the words in English songs can be divided (a / fa / ther's / pen /).

(2) Students were given a sheet on which two English songs, *Humpty Dumpty* and *Old MacDonald Had a Farm*, were printed (See Appendix 1). They were then asked to divide the words into syllables with a slash, as shown in Step (1). These nursery rhymes were chosen because there is a one-to-one correspondence between syllables and notes in each song.

(3) Students listened to a piano melody of *Humpty Dumpty* which had been recorded on a mini disc. The piano melody was heard from a mini disc player.

(4) Students were requested to sing the song along with the piano melody. (Steps (3) and (4) were repeated for the other song, *Old MacDonald Had a Farm*.)

(5) Students received syllable-counting training (detailed in the subsequent section).

(6) Again, students were asked to divide the words into syllables with a slash on a new sheet on which the two songs were printed.

(7) Once more, students were requested to sing the songs along with the piano melody.

2.2. Syllable-counting training

The basic procedure of syllable-counting training was originally developed by Tajima and Erickson (2001), and later a modified version was used in Ishikawa (2002, 2003). The current version of the training is different in two respects from the one used in Ishikawa (2002, 2003). One is that nonword stimuli were deleted and more English words were added because nonwords were used in previous studies to examine the effect of syllable structure per se on the recognition of syllables and this purpose was not included in the present study. Another difference is that the number of English words used in the pretest and post-test was reduced from 200 to 100 words due to the requirements of a shorter training period. The procedure was as follows.

One hundred English words were selected as stimuli, which were varied in number of syllables from one to six, including 22 each of one-, two- and three-syllable words, 17 four-syllable words, 12 five-syllable words and 5 six-syllable words (e.g., *stamp, friendship, September, introduction, classification* and *responsibility*). Each word was read by four native speakers of English (two male British English speakers, one male and one female American English speakers), producing 400 stimuli in all. The stimuli were recorded on DAT, edited at a 16-kHz sampling rate (16-bit resolution), and stored on hard disc. The stimuli were presented auditorily over a headphone via a computer using SuperLab software (Cedrus Corp.). All the stimuli are listed in Appendix 2.

The participants were trained individually in a quiet room where they sat in front of a computer. They were instructed to listen to each word, to count the number of syllables in it, and to press the appropriate number key (from 1 to 6) on the computer keyboard. The participants initiated each trial by pressing the space bar on the keyboard. Feedback was given in such a way that if the participants pressed the correct number, a chime rang, and if they pressed the wrong number, a beep sounded and the same word was heard again by pressing the space bar. The participants could not proceed to the next trial until they

pressed the correct number.

The procedure consisted of a familiarization phase, a pretest, training session 1, training session 2, and a post-test. Feedback was given only for the familiarization phase and the two training sessions. The familiarization phase was comprised of 6 English words, which were different form the test stimuli. The pretest consisted of 100 English words spoken by one of the four English speakers. The training session used 40 English words selected from all 100 English words. Training session 1 consisted of 120 English words (3 sets of 40 English words) spoken by three English speakers. Training session 2 was the same as training session 1 except that the participants were requested to repeat aloud each word after they heard it and to then press the number key. The post-test was identical to the pretest except that 100 English words were spoken by another English speaker different from those used in the pretest and the training sessions. The total test time was about 60 minutes.

3. Results and discussion

3.1. Overall performance

The proportion of Japanese students' correct responses was 55.2% in the pretest, but it rose to 77.5% in the post-test after training, with a significant effect of training [$t(19)=7.81$, $p<.0001$]. The result indicates that Japanese tertiary students' ability to count English syllables improves significantly after a brief training with feedback, as demonstrated in previous studies.

3.2. How do Japanese students recognize syllables and what effect does the training have on their syllable identification?

The following sections will examine how Japanese students recognize English syllables and how their recognition ability is improved through the training by analyzing error responses in view of the number of syllables in the words. The specific questions asked are (1) whether Japanese students recognize more number of syllables in the words than are present, (2) which words, from 1-syllable to 6-syllable words, are difficult to identify, and (3) how many syllables Japanese students tend to count in each of the 1-syllable to 6-syllable words. For each question, the results will be compared between the pretest and the post-test to assess the effects of training.

3.2.1. Do Japanese students recognize more syllables or less syllables than are present?

The error responses were classified into two categories: "More" and "Less". "More" cases included responses like "two" or more for 1-syllable words like *brand*, and "five" or "six" for 4-syllable words like *concentration*. "Less" cases contained answers such as "one" for 2-syllable words like *program*, and "four" or less for 5-syllable words like *classification*. Figure 1 shows the average number of errors for each subject in which "More" and "Less" syllables were recognized in the pretest and the post-test.

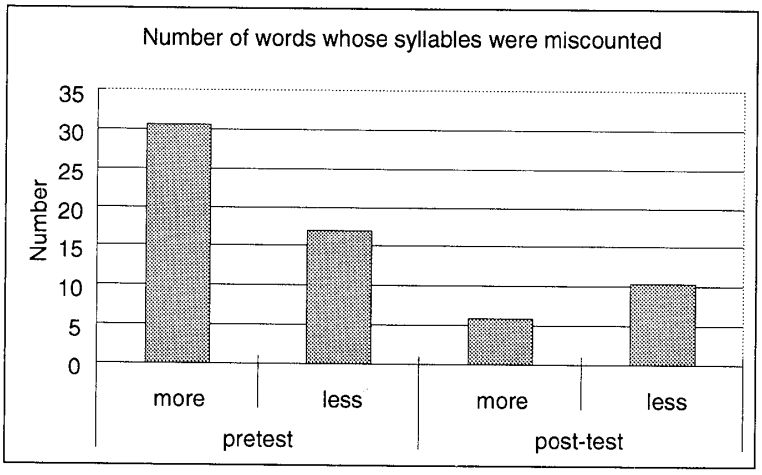


Figure 1. Average number of errors in which "More" and "Less" syllables were recognized.

The cases in which "More" syllables were recognized were significantly more frequent than those where "Less" syllables were recognized [$t(38)=2.46$, $p<.02$] in the pretest. However, in the post-test, "Less" cases were more numerous than "More" cases [$t(38)=2.83$, $p<.01$]. In fact, for example, many participants answered "two" for 3-syllable words like *department* and *scholarship*. The fact that Japanese speakers recognized less syllables than were present in the post-test suggests that after syllable-counting training they had not used moras to count syllables in words.

3.2.2. Which words are difficult to identify?

Figure 2 shows the proportion of errors for each of 1 to 6-syllable words.

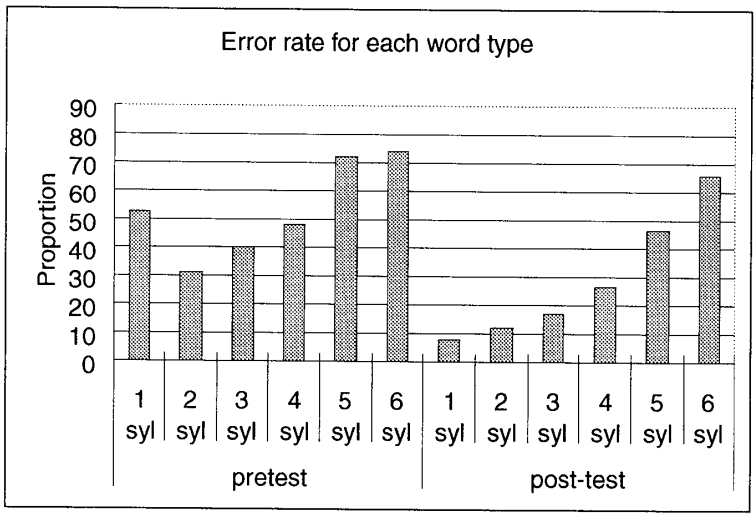


Figure 2. Proportion of errors for each type of English word.

The effects of the number of syllables in words were significant in both the pretest and the post-test [$F(5, 95)=12.85$ & 38.66, $p<.0001$]. In the pretest, post hoc Tukey-Kramer tests with a .05 significance level revealed that 5 and 6-syllable words produced more errors

than 1-, 2-, 3-, and 4-syllable words, and that 1-syllable words made more errors than 2-syllable words. These results suggest that before Japanese students receive syllable-counting training, the longer English words are, the more difficult it becomes to recognize syllables, except for 1-syllable words. One-syllable words seem to be rather difficult to recognize as a whole unit. In the post-test, post hoc tests showed that 6-syllable words produced more errors than the other types of words, suggesting the strong effect of the number of syllables. The notable difference from the pretest was that 1-syllable words were identified more accurately. This indicates that syllable-counting training enables Japanese students to recognize a monosyllable word as a single unit.

3.2.3. How many syllables do Japanese students recognize in each type of word?

Figure 3 shows the average number of error responses separately for 1-, 2-, 3-, 4-, 5-, and 6-syllable English words.

One-syllable English words, in the pretest, were incorrectly recognized as having 2 syllables most often ($p<.05$), followed by 3 syllables. In the post-test, they were also wrongly identified as consisting of 2 syllables most often ($p<.05$), while there were no wrong responses to 3 to 6 syllables, indicating a narrower range of possible incorrect responses. This result provides further evidence that syllable-counting training can weaken the tendency for Japanese students to identify more syllables in words. For 2-syllable words, in the pretest they were wrongly recognized as having 3 syllables most often ($p<.05$), followed by 4 and 1 syllable, then a small number of 5 and 6 syllables. In the post-test, they were incorrectly identified as having 3 and 1 syllable most often ($p<.05$) to the same extent. For 3-syllable words, in the pretest, they were wrongly recognized as having 2 syllables most often ($p<.05$), followed by 4 syllables, and then 5 and 6 syllables. In the post-test, they were perceived as 2 syllables most often ($p<.05$), followed by 4 syllables. For 4-syllable words, they were incorrectly recognized as having 3 syllables most often ($p<.05$), followed by 2 and 5 syllables, and similarly in the post-test 3 syllables had the most errors, followed by 5 syllables. For 5-syllable words, they were wrongly identified as 4 syllables most often ($p<.05$), followed by 3 syllables in both the pretest and the post-test. For 6-syllable words, they were incorrectly recognized as having 5 syllables most often ($p<.05$), followed by 4 syllables in both tests.

These results can be summarized as follows. In the pretest, 1 and 2 syllable English words tended to be wrongly recognized as having 2 and 3 syllables respectively in most cases, suggesting that they were likely to be identified as having one more syllable. In contrast, 3, 4, 5, and 6 syllable words tended to be recognized as having 2, 3, 4, and 5 syllables respectively most often, indicating that they were likely to be identified as having one less syllable. It seems likely that Japanese students tend to recognize less syllables in English words when the words to be counted have three or more syllables. In these cases students do not seem to rely on moras as the unit for counting English syllables because, if they had done so, they should have identified more syllables in the words. Although the same trend can be observed in the post-test, the noticeable difference from the pretest

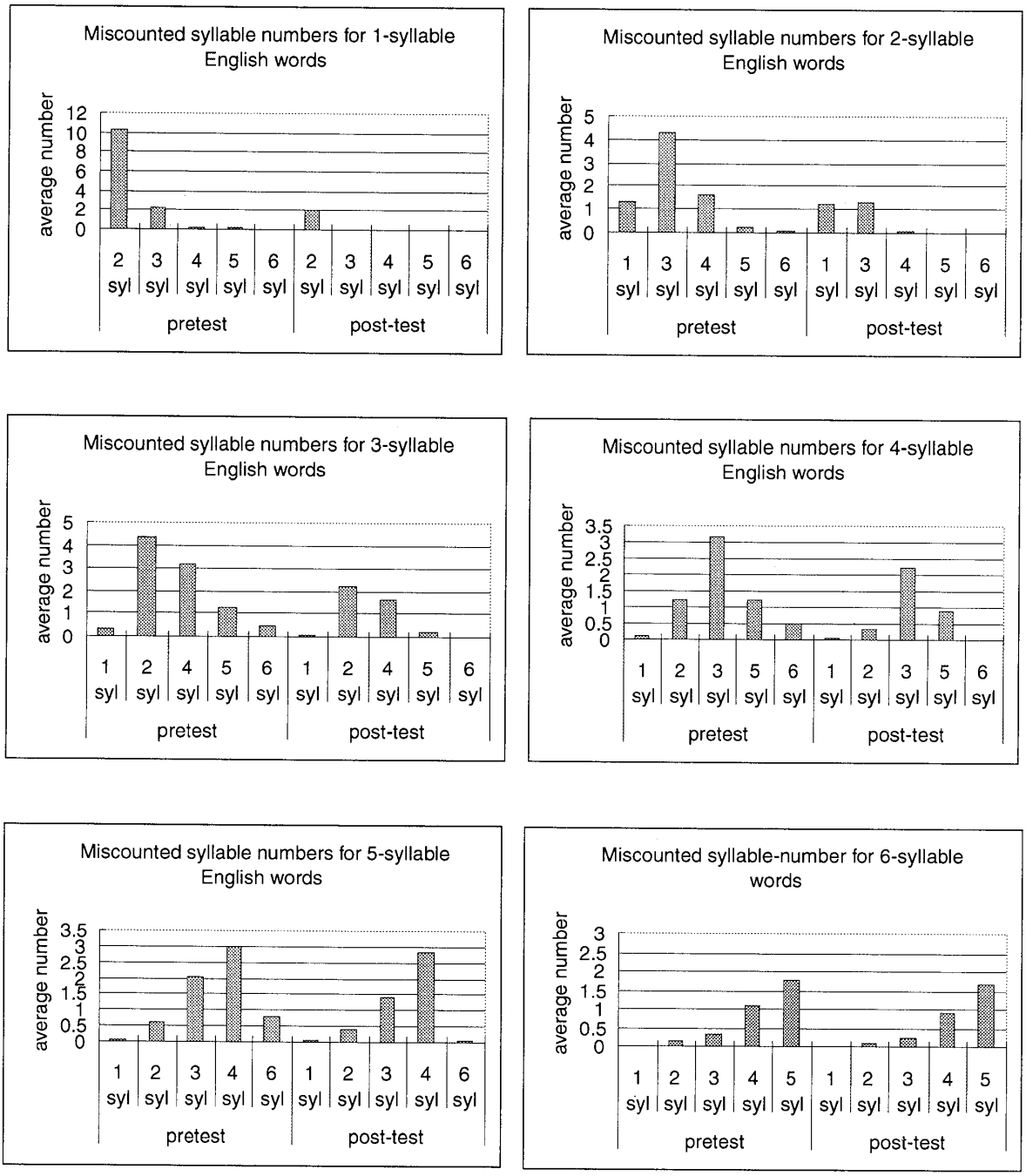


Figure 3. Average number of error responses for 1-, 2-, 3-, 4-, 5-, and 6-syllable words.

was that the range of possible wrong responses was significantly limited due to the effect of syllable-counting training.

3.3. Segmentation of English syllables on the text

In this section, Japanese students' ability to divide English words into syllables on the sheet (See Appendix 1) will be compared prior to and after syllable-counting training. Figure 4 shows error rates of wrong segmentations for two songs.

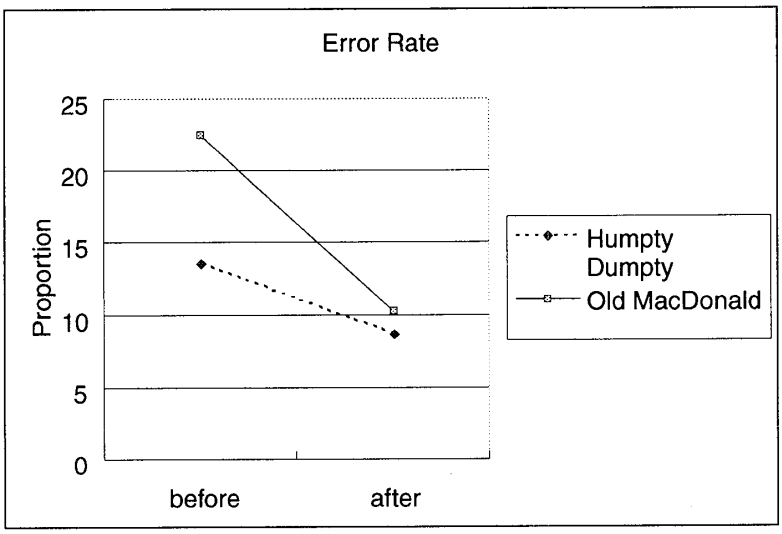


Figure 4. Proportion of wrong segmentations for two songs.

The proportions of wrong segmentations for *Humpty Dumpty* were 13.5% before syllable-counting training and 8.6% after training, with a significant effect of training [$t(19)=2.03$, $p<.06$], and the error rate for *Old MacDonald Had a Farm* was also reliably reduced from 22.5% (before training) to 10.2% (after training)[$t(19)=4.74$, $p<.001$]. These results suggest that syllable-counting training is very effective in improving the students' ability to divide English words into syllables. The following are examples of how their segmentations were improved after training:
Hum/p/ty → Hump/ty, Mac/Do/na/ld → Mac/Do/nald , far/m → farm, du/cks → ducks, qua/ck → quack, he/re → here, toge/ther → to/ge/ther, horses → hors/es

3.4. Song performance

In this section we will assess Japanese students' ability to sing English songs on the basis of syllables, that is, whether they can assign one syllable to one musical note. As mentioned in the Introduction and the Method sections of this paper, in English songs musical notes mostly coincide with syllables, and the two English songs used in the present research show a perfect one-to-one correspondence between syllables and notes. Therefore, having students sing these songs can test their ability to produce English words in terms of syllables.

Regarding scoring, two independent raters (the researcher and a native American English speaker who teaches English at a Japanese university) evaluated the students' performance. The raters listened to the songs sung by the students heard from a mini disc player. The order of the songs heard from the player was randomized across the students and the tests. The raters were required to check how the students' timing of syllable beats corresponds to the timing of musical notes, and to note that their performance regarding melodies should not be considered. They rated the students' performance on a scale of 1 to 5 in the following criteria: Five points (Very good) are given if the number of errors is between 0 and 1. Four points (Good) are given if the

number of errors is 2 to 3. Three points (Satisfactory) are given if the number of errors is 4 to 6. Two points (Poor) are given if the number of errors is 7 to 9. One point (Very poor) is given if the number of errors is more than 10. The inter-rater correlation coefficient was .76, suggesting a reasonably high reliability. Average scores of the two raters were used in the statistics.

Figure 5 shows the average scores for syllabic timing in songs sung by the twenty students who had received syllable-counting training.

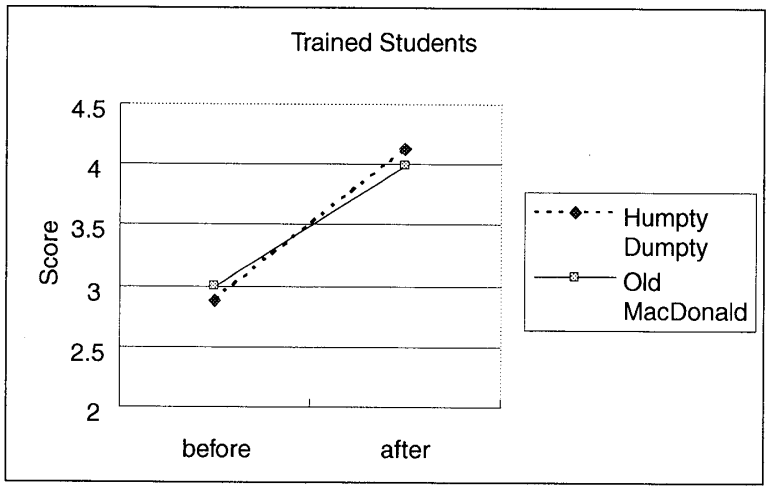


Figure 5. Average scores for trained students.

For both songs, the scores were significantly increased after training [$t(19)=6.34$, $p<.0001$ for *Humpty Dumpty* & t $(19)=4.87$, $p<.0001$ for *Old MacDonald Had a Farm*]. The results suggest that Japanese students can improve their ability to produce English words in terms of syllables through syllable-counting training. However, there is a possibility that the students' ability was improved just because they sung the songs twice, and not because they received syllable-counting training. In order to test this possibility, another 12 students of English at Kyoto Women's University were recruited, and were asked to sing these songs twice consecutively, without the training. All the other steps in the procedure were the same with the trained students, including a brief explanation of the syllable and the segmentation of English syllables on the text.

Figure 6 shows the average scores for syllabic timing in songs sung by the twelve students who had not been trained to count syllables in English words.

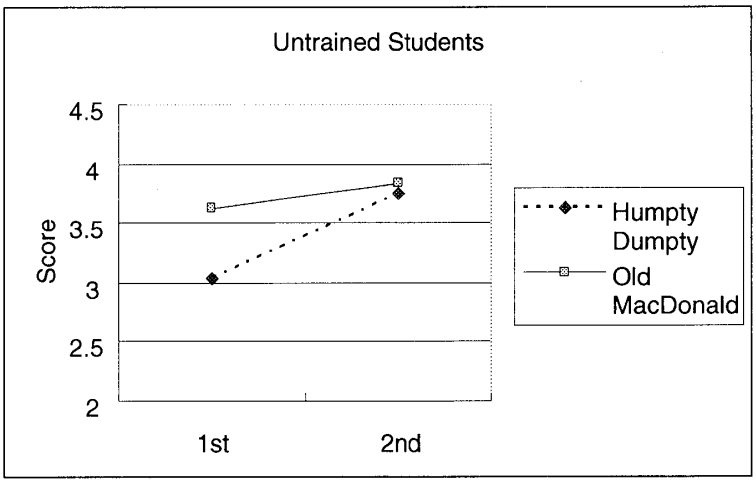


Figure 6. Average scores for untrained students.

Regarding *Old MacDonald Had a Farm*, the scores were not improved significantly from the first time to the second time [$(t(11)=1.05, p>.10)$. This result supports the claim that the improvement of the trained students' ability to produce English syllables was due to the effect of syllable-counting training. As for *Humpty Dumpty*, however, a significant difference was detected between 1st and 2nd time [$t(11)=2.38, p<.05$], indicating an improvement of the students' ability to assign syllables to notes, although the degree of improvement was smaller than was the case with the trained students. The latter result may also be attributed to the observation that *Humpty Dumpty* is relatively easy to sing, perhaps because it is a short song with a slow tempo, in comparison with *Old MacDonald Had a Farm*. At any rate, we cannot deny the possibility that having students sing English songs repeatedly might improve their ability to sing them based on syllables if the songs are relatively slow and short.

4. Conclusion

The present study investigated Japanese students' recognition and production of English syllables. Japanese students were trained to count syllables in English words, and their performance on counting tasks and singing English songs was compared prior to and after training.

The first finding is that syllable-counting training can improve Japanese students' ability to identify English syllables and that the number of syllables in English words influences their recognition of syllables. The following characteristics were made clear by analyzing error responses in view of the number of syllables in words. Japanese listeners tend to recognize more syllables than are present before training, but after training they are likely to perceive less syllables than are present; they have difficulty identifying syllables in longer words except for 1-syllable words, which can be recognized more easily after training; and they tend to recognize 1 and 2 syllable words as having more syllables, while they are likely to identify 3 and more syllable words as having less syllables, narrowing the range of wrong answers after training.

Why did Japanese students often recognize less syllables when the words had three or more syllables? If Japanese speakers had used moras to identify syllables in words (e.g., Otake et al, 1993), they should have recognized more syllables than were present. Moreover, if they had inserted vowels after consonants (e.g., Dupoux et al, 1999), they also should have perceived more syllables than are present. Therefore, in the case of English words with three and more syllables, we need to deny their use of moras and epenthetic vowels. How can such results be accounted for? Two possibilities can be suggested. The first possibility may be to assume that Japanese listeners cannot perceive unstressed syllables because these syllables are weak and often reduced. For example, when they counted syllables in three-syllable words like *scholarship*, they tended to answer "2", probably because they could not perceive at least one of the two unstressed syllables: "lar" and "ship". The second possible explanation could be to speculate that they use two syllables as one unit, an idea which is a similar notion to bimoraic foot, i.e., two moras as a single unit or a minimal phonological word (e.g., Ito, 1990; Mori, 2002; Poser, 1990). For instance, three-syllable words like *grandmother*, which was often recognized as having two syllables, could be considered as consisting of "grand" and "mother" because the latter, "mother", comprised of the second and the third syllables could be perceived as a single unit. The above findings suggest that factors other than moras and epenthetic vowels may exercise an effect on the recognition of English syllables by Japanese students.

The second finding is that syllable-counting training can improve Japanese students' ability to divide English words into syllables and to match each syllable to one musical note when they sing English songs. This kind of counting task seems to be very effective in helping Japanese learners of English properly recognize and produce English syllables. Counting syllables may possibly lead to the development of a clearer and deeper understanding of English syllable structure and syllable shapes.

Further work is needed to explore the cognitive mechanism which develops the relationship between counting and producing English syllables. It would also be of primary interest to look into what other language skills (including listening ability) the competence developed by syllable-counting training may extend to. It is hoped that the kind of training reported in this paper will be critically and positively evaluated so that it can be put into practical use in the English learning environment.

Acknowledgments

This article is based on the presentation made at the 43rd JACET annual convention, Nagoya, September 2004. I am grateful to the audience for their helpful comments. This research was supported in part by a grant from the Japanese Ministry of Education, Culture, Sports, Science and Technology (No. 14510637).

References

Bond, Z.S., & Fokes, J. (1985). Non-native patterns of English syllable timing. *Journal of Phonetics*, 13, 407–420.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1568–1578.

Faber, D. (1986). Teaching the rhythms of English: A new theoretical base. *IRAL*, 24, 205–216.

Fokes, J., & Bond, Z.S. (1989). The vowels of stressed and unstressed syllables in nonnative English. *Language Learning*, 39, 341–373.

Giegerich, H.J. (1992). *English phonology: An introduction*. Cambridge, U.K.: Cambridge University Press.

Ishikawa, K. (2002). Syllabification of intervocalic consonants by English and Japanese speakers. *Language and Speech*, 45, 355–385.

Ishikawa, K. (2003). Effects of syllable-counting training on Japanese listeners' perception of English syllables. *Essays & Studies*, 48, 6–18. Kyoto Women's University.

Ito, J. (1990). Prosodic minimality in Japanese. *CLS* 26, Vol. 2: The parasession on the syllable in phonetics and phonology, 213–239.

Kubozono, H. (1999). Mora and syllable. In N. Tsujimura (Ed.), *The handbook of Japanese linguistics* (pp.31–61). MA, USA: Blackwell.

Lieberman, I.Y., Shankweiler, D., Fischer, F.W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, 18, 201–212.

Mori, Y. (2002). Lengthening of Japanese monomoraic nouns. *Journal of Phonetics*, 30, 689–708.

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language*, 32, 258–278.

Palmer, C., & Kelly, M.H. (1992). Linguistic prosody and musical meter in song. *Journal of Memory and Language*, 31, 525–542.

Poser, W. J. (1990). Evidence for foot structure in Japanese. *Language*, 66, 78–105.

Roach, P. (1991). *English phonetics and phonology: A practical course* (2nd ed.). Cambridge, U.K.: Cambridge University Press.

Tajima, K., & Erickson, D. (2001). Syllable structure and the perception of second-language speech. In Spoken Language Working Group (Ed.), *Speech and Grammar III*. Tokyo: Kuroshio.

Vance, T. (1987). *An introduction to Japanese phonology*. New York: State University of New York Press.

Appendex 1: The sheet used by participants to segment the words and sing the two songs chosen for this study

Humpty Dumpty

Humpty Dumpty sat on a wall,

Humpty Dumpty had a great fall.

All the king's horses and all the king's men

Couldn't put Humpty together again.

Old MacDonald Had a Farm

Old MacDonald had a farm, ee-igh, ee-igh, oh.

And on his farm he had some ducks, ee-igh, ee-igh, oh.

With a quack, quack here, Quack, quack there,

Here a quack, there a quack, everywhere a quack, quack,

Old MacDonald had a farm, ee-igh, ee-igh, oh.

Appendix 2: English words used for syllable-counting training

6 Practice Words

track	British	tragedy	capacity	possibility	interrelationship

100 English Words for Training

1 syllable	2 syllables	3 syllables	4 syllables	5 syllables	6 syllables
blast	against	agreement	astonishment	accommodation	accountability
brand	anxious	atmosphere	concentration	administration	availability
camp	bankrupt	brotherhood	confirmation	classification	identification
climb	blanket	confidence	conveniently	encyclopedia	responsibility
drain	bridegroom	criterion	diplomatic	multicultural	revolutionary
dress	classroom	department	distribution	nationality	
drink	craftsman	displacement	entertainment	necessarily	
France	danger	dramatic	identity	opportunity	
glimpse	entrance	employment	information	organization	
ground	extreme	frequency	introduction	preparatory	
prompt	frankly	government	investigate	probability	
quest	friendship	grandmother	mentality	semiconductor	
script	instinct	important	professional		
shrimp	neglect	instrument	reluctantly		
splash	precise	photograph	spectacular		
spring	program	pleasantly	speculation		
stamp	Scotland	prescription	transportation		
straight	Spanish	scholarship			
strength	splendid	September			
task	sweetness	skeptical			
throat	training	specialist			
trend	trumpet	statistics			