Is Acquisition of L2 Phonemes Difficult? Production of English Stress by Japanese Speakers

Mariko Kondo
Waseda University

1. Introduction

This study examined the production of English lexical stress by Japanese speakers to determine which acoustic features associated with English lexical stress are difficult for Japanese speakers to acquire. Realization of lexical accent differs between languages. English is a stress-accent language where the accent is expressed by a combination of pitch, duration, intensity and vowel quality (Fry 1955, Gay 1978, Kochanski et al. 2005 & Kochanski and Orphanidou, 2008). In contrast, Japanese is a pitch-accent language where the accent is dominantly realized by a fall in the fundamental frequency (F0) from an accented high pitched mora to the following mora, and there is very little use of intensity to mark accent (Fujisaki et al. 1986). Studies on non-native English learners have shown that problems on acquiring English stress vary depending on their first language (L1). For example, Arabic speakers show different use of F0 and formant frequencies in stressed and unstressed vowels compared with native English speakers (Zuraiq & Soreno 2007). Lee et al (2006) also reported that Japanese speakers found it difficult to achieve vowel centralization of English unstressed vowels.

This study investigated the effect of Japanese (L1) prosody on English (L2). In particular, the study focused on how Japanese speakers who speak fluent English still show the influence of Japanese phonology, and also examined whether there are any characteristics notably different from native English speakers' utterances. Different factors influence L2 speech rhythm. For example, different languages use different phonological units to keep speech rhythm. Japanese uses the mora as the fundamental unit for its speech rhythm and many studies have reported mora-based timing control for Japanese. The duration of each mora is not necessarily equal, but the duration of a word or phrase is determined by the number of morae in it (Port et al., 1987). There is mora-based segmental elasticity and durational compensation between CV which constitutes a mora rather than between V-C across a mora boundary (Campbell and Sagisaka, 1991; Sato, 1993). The effect of the mora on L2 has been reported in English (Mochizuki-Sudo & Kiritani 1991) and in French (Kondo & Shinohara, 2003 & 2006). Vowel quality is also an important cue for stress in English. Everything being equal, vowels in unstressed syllables are shorter and have centralized quality. Although the order of importance among these acoustic cues of English stress varies from study to study, all these are important cues to differentiate stressed and unstressed syllables in English (Fear et al., 1995).

Considering the prosodic differences between English and Japanese, various factors affect L2 speech rhythm. An earlier study (Kondo, 2007) found an L1 influence in English speakers' Japanese utterances. English speakers showed strong influence of lexical accent on vowel duration in their Japanese utterances. They showed two patterns: (i) large F0 increase ***but*** little durational increase, and (ii) durational increase instead of any F0 increase. In the present study most of the Japanese subjects spent some years in English speaking environments and spoke relatively good English. However, they still showed some influence of their L1, Japanese, in their English utterances.

Among these acoustic features, i.e. vowel duration, F0, intensity and quality, some features are easier to acquire than others and some take longer to acquire during the process of L2 sound acquisition.

* The study was supported by the Japanese Ministry of Education, Culture, Sports, Science & Technology, Grant-in-Aid for Scientific Research (C) No. 19520369 and Sophia University Open Research Center from MEXT.

 Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009), ed. Melissa Bowles et al., 105-112. Somerville, MA: Cascadilla Proceedings Project.

If some features are difficult to acquire, how do Japanese speakers manipulate these acoustic features while when they try to achieve English prosody? Is there any relationship between the features? The current research studies Japanese speakers' acoustic manipulation of English lexical stress and examines how Japanese speakers try to achieve English stress.

2. Experiment

2.1. Methods

In order to measure acoustic realization of English stress by Japanese speakers, eight female Japanese speakers and four female native American English speakers recorded their utterances. The Japanese subjects were students of Waseda University; they had all spent some years (up to 10 years) in English speaking environments and were fluent in English, but none of them were bilingual. The American speakers were students of Kansas University, from either Kansas or Missouri.

All the students in the study pronounced pairs of disyllabic English words with contrasting stress positions, listed in (1) below. The words were randomly presented and were pronounced three times each at a comfortable tempo in the carrier sentence '*I'll say test word now*'. The stressed syllable was marked on the reading text for both Japanese and English speakers. Stressed and unstressed vowels of each test word were examined acoustically for (i) duration, (ii) the average, peak and lowest F0, (iii) average and peak intensity and (iv) the first and second formant frequencies. For the word pairs ˈ*differ* /ˈdɪfə(r)/ - *deˈfer* /dɪˈfɜ:(r)/ *and* ˈ*permit* /ˈpɜ:(r)mɪt/ – *perˈmit* /pəˈmɪt/, measurements of the second vowel /ə(r)/ and /ɜ:(r)/ of ˈ*differ* - *deˈfer* and the first vowel /ɜ:(r)/ and /ə(r)/ of ˈ*permit* and *perˈmit* were taken including the following /r/ whenever the vowels were followed by /r/. The reason was because some speakers had acquired rhotic accent and pronounced the postvocalic /r/ while others did not.

(1) Di-syllabic words: (a) 1st syllable accented and (b) 2nd syllable accented

(a) ˈ*contract*	(noun)	/ˈkɑ:ntrækt, ˈkɒntrækt/		(b) *conˈtract*	(verb)	/kənˈtrækt/
(a) ˈ*differ*	(verb)	/ˈdɪfə(r)/	–	(b) *deˈfer*	(verb)	/dɪˈfɜ(r)/
(a) ˈ*permit*	(noun)	/ˈpɜ:(r)mɪt/	–	(b) *perˈmit*	(verb)	/pə(r)ˈmɪt/
(a) ˈ*subject*	(noun)	/ˈsʌbdʒekt/	–	(b) *subˈject*	(verb)	/səbˈdʒekt/
(a) ˈ*decrease*	(noun)	/ˈdi:kri:s/	–	(b) *deˈcrease*	(verb)	/dɪˈkri:s/

There were some problems in the experiments. Some Japanese speakers inserted a vowel in the consonant clusters in *subject* /sʌbdʒɪkt/ and *decrease* /dɪkri:s/ as /sʌbu̲dʒekt/ and /dɪku̲ri:s/ respectively. Both words were pronounced as three syllables. These utterances were eliminated from analysis. Another problem was accent placement. Many Japanese speakers as well as many native English speakers placed the lexical accent on the second syllable rather than on the first for both *decrease* (noun) and *permit* (noun) as /dɪˈkri:s/ and /pɜ:(r) ˈmɪt/. Since there were too many examples of /dɪˈkri:s/ for *decrease* (noun), the ˈ*decrease* (noun) - *deˈcrease* (verb) pair was eliminated entirely and samples of wrongly accented *permit* (noun) /pɜ:(r) ˈmɪt/ were also eliminated from analysis.

The samples were digitized at a sampling rate of 44.1 KHz, and analyzed using Praat. Statistical analysis was performed by SPSS.

3. Results and discussion

3.1. Vowel duration

Stressed vowels were longer than unstressed vowels for both native English and Japanese speakers (For the first vowel (V1): language [$F(1, 277) = 61.5$, $p < .0001$], stress position [$F(1, 277) = 117.49$, $p < .0001$]; and the second vowel (V2): language [$F(1, 277) = 30.56$, $p < .0001$], stress position [$F(1, 277) = 27.35$, $p < .0001$]) (Figures 1a, 1b). This difference occurred when either V1 or V2 was stressed, but the native English speakers reduced vowel duration significantly more than the Japanese speakers in unstressed vowels. There was a significant interaction between the language groups and the location of stress when V1 was stressed (language x stress position [$F(1, 277) = 6.43$, $p < .05$)]) showing that there was a larger difference between the durations of stressed and unstressed vowels in the native English speakers' utterances than in the Japanese speakers utterances, mainly for V1.

Fig. 1a Duration of V1 in differing stress locations by Japanese & English speakers

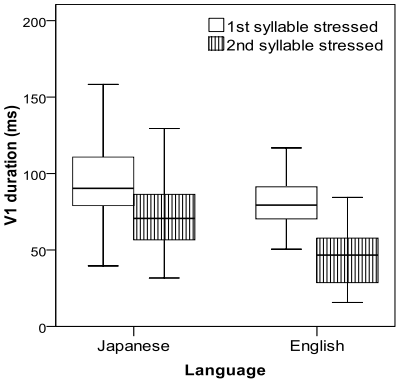
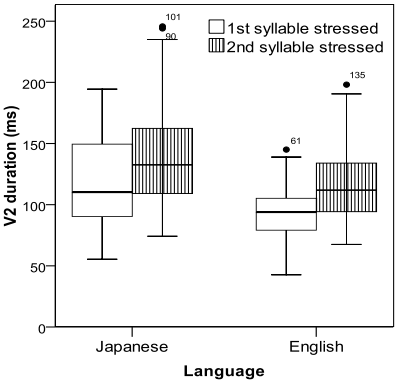


Fig. 1b Duration of V2 in differing stress locations by Japanese & English speakers



3.2. Intensity

The peak and average intensity of all vowels were measured. The ratio between the first and second vowels within the same word was obtained and analyzed by the location of stress and speakers' first language. The V1/V2 ratios for both average and peak intensities were over 100% when V1 was stressed and below 100% when V2 was stressed. This means that both language groups increased vowel intensity when vowels were stressed compared with unstressed vowels in the same word. There was no significant difference between the language groups (for average V1/V2 intensity ratio: language [$F(1, 277) = 2.55$, n.s.], stress position [$F(1,277) = 281.93$, $p < .0001$], language x stress position [$F(1, 277) = .018$, n.s.]; for peak V1/V2 intensity ratio: language [$F(1, 277) = 4.07$, $p < .05$], stress position [$F(1, 277) = 291,64$, $p < .0001$], language x accent [$F(1, 277) = .028$, n.s.]) (Figure 2a, 2b). A previous study by Fujisaki et al. (1986) found that Japanese speakers increase intensity of accented vowels in Japanese, but the degree of increase was far less than that by native English speakers speaking in English. In general increased intensity of accented vowels is a natural physiological phenomenon: intensity increase is approximately proportional to the increase of pitch. Japanese lexical accent is indicated by higher F0 and its rapid fall towards the following mora. Therefore, in Japanese vowel intensity is naturally increased when a vowel is accented because of faster vocal fold action. However, in English the increase in intensity of stressed vowels is generally greater than the increase in F0 (e.g. Fujisaki et al. 1986). Other studies have also shown that in English intensity is a more reliable cue for stress than increases in F0 (Kochanski et al., 2005 & Kochanski and Orphanidou, 2008). This indicates that in the current study, the Japanese subjects, who spoke fluent English, were able to use intensity cues to mark lexical stress when speaking in English.

Fig. 2a Average intensity ratio of V1/V2 in differing stress locations by Japanese & English speakers

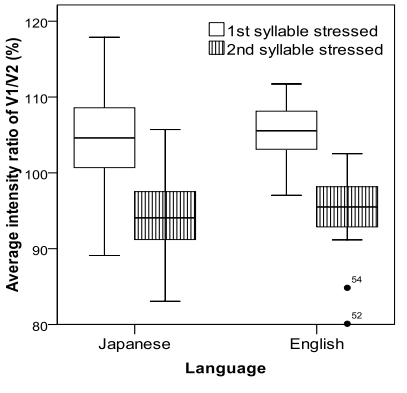
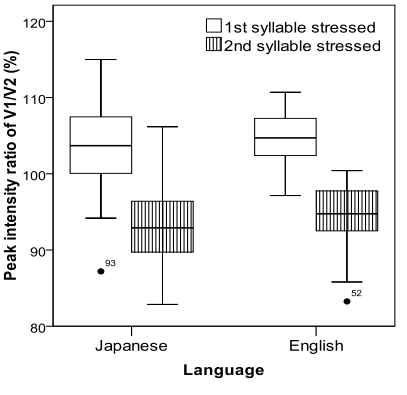


Fig. 2b Peak intensity ratio of V1/V2 in differing stress locations by Japanese & English speakers



3.3. The fundamental frequency of stressed and unstressed vowels

The average F0 of all vowels, peak F0 of stressed vowels and lowest F0 of unstressed vowels were measured. The ratio between stressed and unstressed vowels within the same word was obtained for average F0 and peak and lowest F0s, and analyzed by the location of stress and speakers' first language. The V1-V2 ratios were over 100% for both V1 and V2 for both language groups (Figures 3a, 3b). This means that when vowels were stressed, their F0s increased. The V1-V2 ratio was significantly higher in V1 than V2 for both average F0 ratio and the ratio between peak and lowest F0s for both language groups (for average V1-V2 F0 ratio: language [$F(1, 277) = .14$, n.s.], stress position [$F(1, 277) = 89.46$, $p < .0001$; for peak F0/lowest F0 ratio: language [$F(1, 277) = .000$, n.s.], stress position [$F(1, 277) = 130.36$, $p < .0001$]). However, there was a significant interaction between language group and stress location: [$F(1, 277) = 10.81$, $p < .001$] for the average F0 ratio and [$F(1, 277) = 4.7$, $p < .05$] for the peak-lowest ratio. This indicates that although both Japanese and English speakers increased F0 of stressed vowels the English speakers changed F0 more depending on whether syllables were stressed or unstressed.

The lower V1-V2 ratio when the second vowel in a word was stressed indicated that the increase in F0 was less when the second vowel of each word was stressed compared with when the first vowel was stressed. This was probably due to the effect of natural declination of F0, which means that F0 is higher at the beginning of utterances and is gradually lowered towards the end of utterances. The difference between language groups suggested that Japanese speakers may tend to mainly use F0 change to indicate stress whereas English speakers rely on intensity increase more than F0 increase to mark lexical stress.

Fig. 3a Average F0 ratio of V1-V2 in differing stress locations by Japanese & English speakers

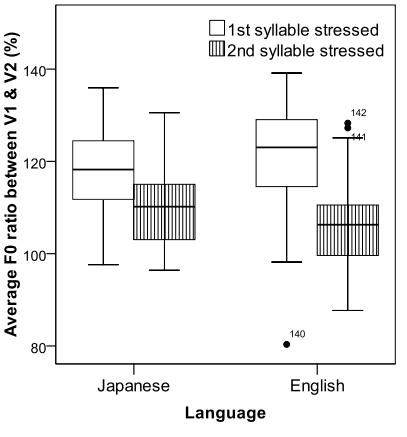
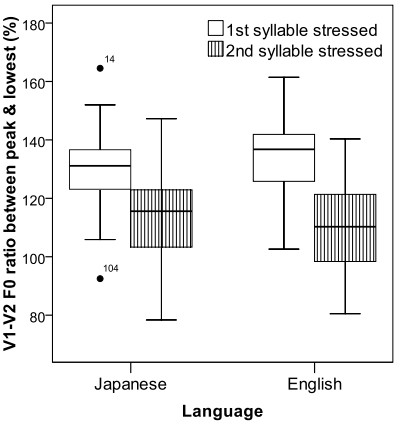


Fig. 3b Peak & lowest F0 ratio of V1-V2 in differing stress locations by Japanese & English speakers

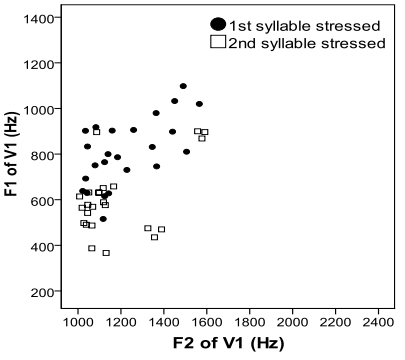


3.4. Vowel quality

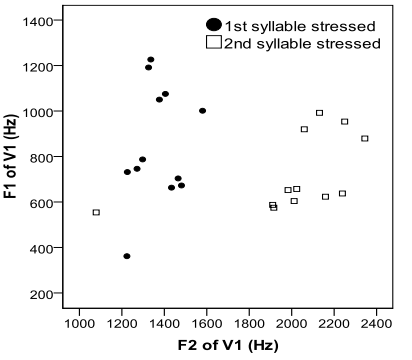
The first and second formants of all vowels were measured at the middle point of their steady state and their formant values were plotted for each vowel by the presence or absence of stress and the results are shown in Figures 4a-4h. English speakers generally showed a clear difference between stressed and unstressed vowel quality, especially when the vowels in the same position were clearly different phonemically, e.g. the underlined vowels in *contract* /ˈkɑ̲ːntrækt, ˈkɒ̲ntrækt/ vs. /kə̲nˈtrækt/, *permit* /ˈpɜ̲ː(r)mɪt/ vs. /pə̲(r)ˈmɪt/, and *subject* /ˈsʌ̲bdʒekt/ vs. /sə̲bˈdʒekt/. F1 and F2 values of phonemically the same vowels in English speakers' production were similar. In other words, whenever there was phonemic difference, English speakers produced vowels of different quality. However, Japanese speakers did not show any clear difference in vowel quality based on stress assignment. The F1 and F2 values of stressed and unstressed vowels produced by Japanese speakers tended to be merged together (Figures 4a, 4c, 4e, and 4g).

Figures 4a-4h. F1 & F2 of stressed and unstressed vowels by Japanese and native English speakers

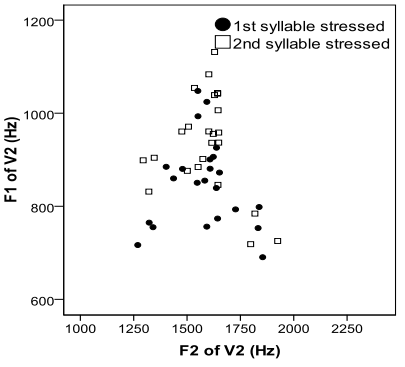
4a(i) Japanese: F1 & F2 of V1 of *contract*



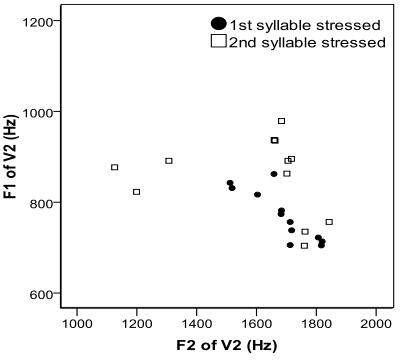
4b(i) English: F1 & F2 of V1 of *contract*



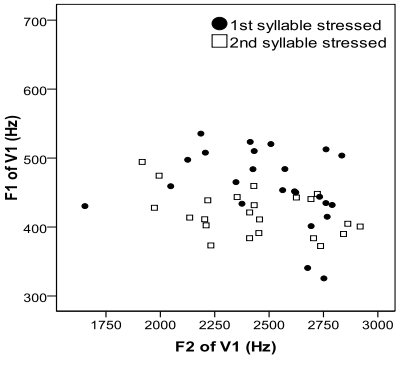
4a(ii) Japanese: F1 & F2 of V2 of *contract*



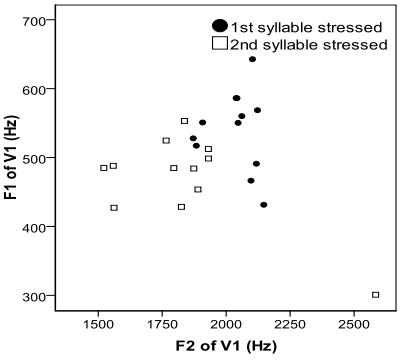
4b(ii) English: F1 & F2 of V2 of *contract*



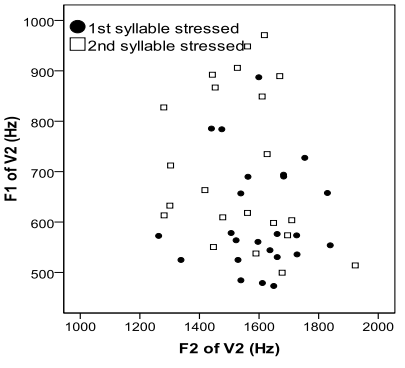
4c(i) Japanese: F1 & F2 of V1 of *differ/defer*



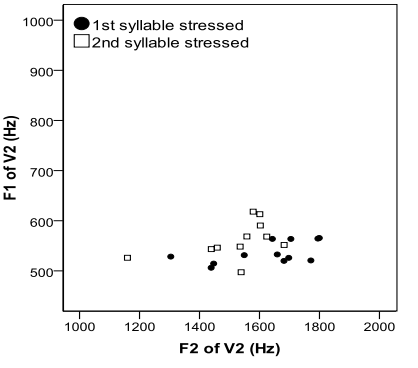
4d(i) English: F1 & F2 of V1 of *differ/defer*



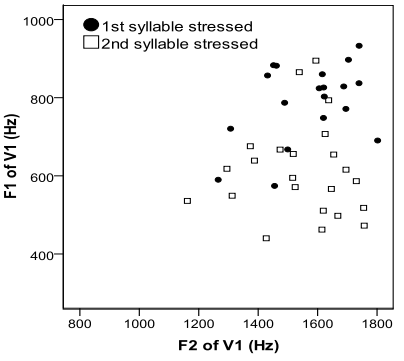
4c(ii) Japanese: F1 & F2 of V2 of *differ/defer*



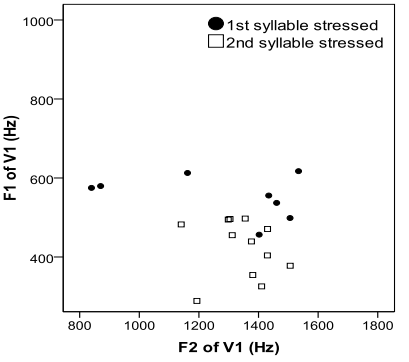
4d(ii) English: F1 & F2 of V2 of *differ/defer*



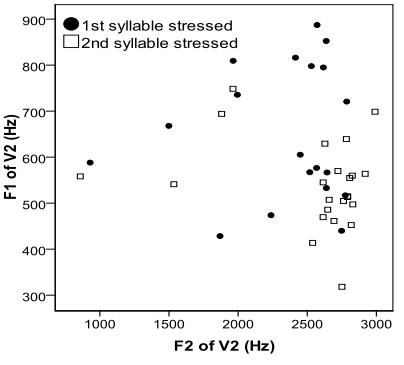
4e(i) Japanese: F1 & F2 of V1 of *permit*



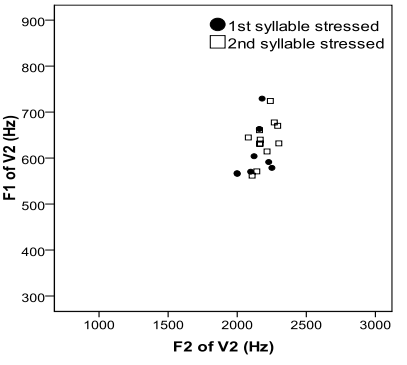
4f(i) English: F1 & F2 of V1 of *permit*



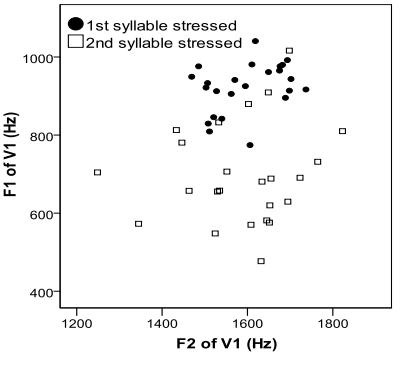
4e(ii) Japanese: F1 & F2 of V2 of *permit*



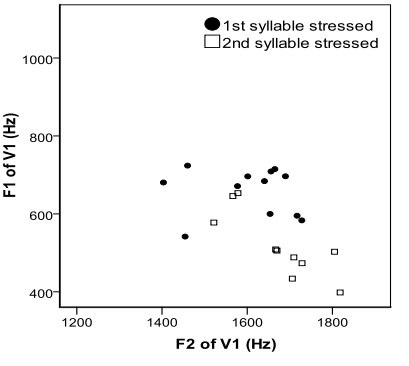
4f(ii) English: F1 & F2 of V2 of *permit*



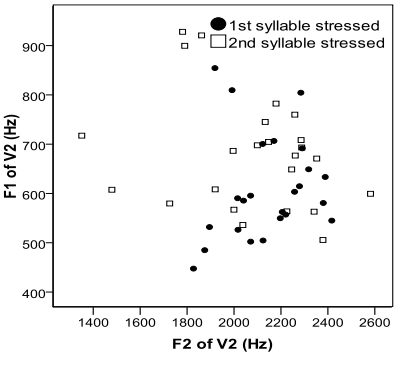
4g(i) Japanese: F1 & F2 of V1 of *subject*



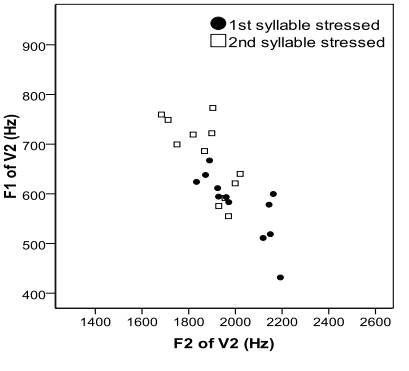
4h(i) English: F1 & F2 of V1 of *subject*



4g(ii) Japanese: F1 & F2 of V2 of *subject*



4h(ii) English: F1 & F2 of V2 of *subject*



3.5. Lexical stress manipulation by Japanese speakers

The Japanese and native English speakers showed similar patterns for (i) vowel duration (ii) average and peak intensity ratios of stressed and unstressed vowels in each word, and (iii) average F0 ratio and peak & lowest F0 ratio of vowels in differing stress locations (Figures 1-3). Vowel durations of unstressed vowels were significantly shorter than their stressed counterparts in the same words for both language groups, but the native English speakers reduced vowel duration significantly more in the unstressed vowels. However, there was a significant difference between Japanese and native English speakers in terms of F1 and F2 of vowels, i.e. vowel quality. The native English speakers varied their F1 and F2 values depending on the presence or absence of stress: unstressed vowels were centralized. The differences in vowel quality were greatest when vowels differed phonemically. However, stress had little effect on the vowel quality of the Japanese speakers' utterances.

The results suggest that Japanese speakers are more sensitive to duration and F0, which are phonologically important in Japanese; Japanese lexical accent is indicated mainly by F0 change. Therefore, it is natural that Japanese speakers also use F0 to indicate lexical stress in English. Japanese differentiate vowels and consonants by their duration both phonetically and phonemically: short vowel /V/ [V] and long vowels /VV/ [V:], and singleton /C/ and geminate voiceless obstruents /CC/ [C:]. Therefore, the Japanese speakers may have been sensitive to the change of vowel duration when the syllable was accented. In terms of intensity, the Japanese subjects showed a similar pattern of intensity increase as native English speakers when vowels were in stressed syllables. As explained in the Introduction, intensity is not associated with accent in Japanese, but the Japanese subjects controlled intensity in a similar way to the native English speakers. However, the Japanese speakers did not differentiate vowel quality indicating that they were not very sensitive to changes in vowel quality associated with the presence or absence of stress in English. This lack of sensitivity is expected to a certain extent because accent does not affect vowel quality in Japanese. Moreover, Japanese has smaller number of vowel phonemes than English. If stressed and unstressed vowel phonemes in English fall in to the same phoneme category in Japanese it is unlikely that Japanese speakers will differentiate these vowels. The Japanese subjects, despite speaking fluent English, may have not acquired English vowel phonemes completely because they perceived both stressed and unstressed vowels to be the same phoneme, and could only perceive differences based on duration, intensity and F0. Consequently, they expressed English accent only by duration, F0 and intensity.

4. Conclusion

The results suggest that the Japanese speakers were able to achieve good control of F0, intensity and duration of stressed and unstressed vowels in English, but even though they were fluent speakers they could not control native-like vowel quality. This implies that it is very difficult for Japanese to acquire L2 phonemes. The results supported the findings of Lee et al (2006). It indicates that the phonemic category of L1 is fixed at an early stage in native language development, and is not likely to be altered, and is therefore transferred to L2 phonology.

These results imply that in the process of L2 acquisition acoustic features of stress are ranked, and for Japanese speakers it is difficult to acquire L2 phonemes.

References

Campbell, W. Nick and Sagisaka, Yoshinori (1991). Moraic and syllable-level effects on speech timing. *Journal of the IEICE*, SP 90-107, 35-40.

Fear, Beverley D., Cutler, Anne and Butterfield, Sally (1995). The strong/weak syllable distinction in English. Journal of the Acoustical Society of America, Volume 97, Issue 3, March 1995, 1893-1904.

Fry, Dennis B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustic Society of America,* 27, 765–768.

Fujisaki, Hiroya, Hirose, Keikichi., and Sugito, Miyoko (1986). Comparison of acoustic features of word accent in English and Japanese. *Journal of the Acoustic Society of Japan* (E) 7: 57–63.

Gay, Thomas (1978). Physiological and acoustic correlates of perceived stress. *Language and Speech*, 21, 347-353.

Kochanski, Greg, Grabe, Ester, Coleman, John and Rosner, Burton (2005). Loudness predicts prominence: fundamental frequency lends little. *Journal of the Acoustic Society of America,* 118:2, 1038-1054.

Kochanski, Greg and Orphanidou, Christina (2008). What marks the beat of speech?. *Journal of the Acoustic Society of America*, Vol. 124: 5, 2780-2791.

Kondo, Mariko and Shinohara, Shigeko (2003). Durational evidence of the psychological reality of the mora in Japanese speakers' French. In M.J. Solé, D. Recasens and J. Romero (eds.), *The Proceedings of the 15th ICPhS*, Barcelona, 463–466.

Kondo, Mariko and Shinohara, Shigeko (2006). 'Phonological parsing of foreign sound sequences and the effect on speech rhythm', *Language and Language Processing*, 265-274, Karnowski, Paweł & Szigeti, Imre (Eds.), Peter Lang: Bern, Switzerland.

Kondo, Mariko (2007). Acoustic Realization of Lexical Accent and its Effects on Phrase Intonation in English Speakers' Japanese, *The Proceedings of the 16th International Congress of Phonetic Science,* Saarbrücken, Germany. Paper ID 1306, 1649-1652**.**

Lee, Borim, Guion, Susan G. & Harada, Teruo (2006). Acoustic analysis of the production of unstressed English vowels by early and late Korean and Japanese bilinguals, *Studies in Second Language Acquisition*, 28, 487-513.

Mochizuki-Sudo, Michiko and Kiritani, Shigeru (1991). Production and perception of stress-related durational patterns in Japanese learners of English. *Journal of Phonetics* 19: 231–248.

Port, Robert F., Dalby, Jonathan, and O'Dell, Michael (1987). Evidence for Mora Timing in Japanese. *Journal of the Acoustic Society of America,* 81, 1574-1585.

Sato, Yumiko (1993). The durations of syllable-final nasals and the mora hypothesis in Japanese, *Phonetica* **50**, 44-67.

Zuraiq Wael & Sereno, Joan A (2007). English lexical stress cues in native English and non-native Arabic speakers, *The Proceedings of the 16th International Congress of Phonetic Science*, 829-832.

Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)

edited by Melissa Bowles, Tania Ionin, Silvina Montrul, and Annie Tremblay

Cascadilla Proceedings Project Somerville, MA 2009

Copyright information

ISBN 978-1-57473-434-8 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Kondo, Mariko. 2009. Is Acquisition of L2 Phonemes Difficult? Production of English Stress by Japanese Speakers. In *Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)*, ed. Melissa Bowles et al., 105-112. Somerville, MA: Cascadilla Proceedings Project.

or:

Kondo, Mariko. 2009. Is Acquisition of L2 Phonemes Difficult? Production of English Stress by Japanese Speakers. In *Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)*, ed. Melissa Bowles et al., 105-112. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2261.