VOWEL EPENTHESIS IN JAPANESE SPEAKERS' L2 ENGLISH

Kakeru Yazawa[1], Takayuki Konishi[1], Keiko Hanzawa[2&3], Greg Short[1&4], & Mariko Kondo[1,3&4]

[1]GSICCS, [2]EDU, [3]SILS, [4]LASS, Waseda University
kyazawa812@gmail.com, tkonishi@aoni.waseda.jp, keiko_hotaru@fuji.waseda.jp, iac14007@kurenai.waseda.jp, mkondo@waseda.jp

ABSTRACT

This study investigated how vowel epenthesis, a typical suprasegmental error produced by Japanese learners of English, is realized in their speech production of L2 English. We investigated two aspects: (i) whether loanword epenthesis phonology in Japanese transfers to English speech production, and (ii) how learners' English proficiency level affects vowel epenthesis. In Japanese loanword phonology, different vowels are inserted in certain phonological environments to maintain Japanese syllable structure. We analyzed speech data from the J-AESOP corpus to examine whether vowel epenthesis in L2 English of L1 Japanese learners resembles the patterns of loanword adaptation in connection with the learners' proficiency level.

The results showed that while fluent learners produced fewer epenthetic vowels, the qualities of the vowels exhibited similar patterns to loanword adaptation, regardless of learners' proficiency level. This indicates that learners' L2 proficiency level affected vowel epenthesis in quantitative rather than qualitative aspects.

Keywords: Vowel epenthesis, Japanese, English, L2 acquisition, corpus linguistics

1. INTRODUCTION

Acquisition of second language (L2) phonology involves acquisition of both segmental and suprasegmental features. While segmental features are important for accurate pronunciation of individual words, suprasegmental features such as speech rhythm strongly affect the overall intelligibility of foreign accented speech [1]. Also, instructions focusing on suprasegmental features have been revealed to be more effective for improving intelligibility than those focusing on segmental features [3]. However, what makes L2 suprasegmental acquisition difficult is, the presence of first language (L1) phonology; L2 learners tend to transfer the L1 phonological systems they have already acquired [12, 14], which makes their speech more foreign-accented and sometimes difficult to understand.

Japanese and English have different rhythmic structures. While the syllable is an important rhythmic unit in English, Japanese is based on the mora (i.e. /(C)V/ unit), and its phonotactics does not allow consonant clusters or syllable-final consonants with very few exceptions. Therefore, if English words with consonant clusters and/or word-final consonants are adapted into Japanese (e.g. *cross* /krɔs/, CCVC in English), vowel epenthesis occurs to break up consonant clusters and to avoid word-final consonants (e.g. /kurosu/, CVCVCV) because consonant deletion is not allowed [16]. Only /i, o, u/ out of the five Japanese vowels (/i, e, a, o, u/), can be inserted, but which one is inserted depends on the consonant that becomes the onset of the mora [13]. As shown in (3), the default epenthetic vowel is /u/ because it is shortest in intrinsic duration and lowest in sonority, and thus the 'least vowel-like' [8]. If the preceding segment is /tʃ, dʒ/, then /i/ is inserted because of the similarities of the places of articulation (1). The /o/ is inserted after /t, d/ because these consonants become different allophones before high vowels, /u/ or /i/ (2).

(1) ø → i / tʃ, dʒ _
(2) ø → o / t, d _
(3) ø → u / elsewhere

Vowel epenthesis in Japanese is not only restricted to loanword phonology, but is also commonly observed in Japanese learners' speech production of English due to suprasegmental L1 transfer. This often makes the resulting speech unintelligible to native English speakers [7, 15]. As L2 learners improve their proficiency, they are expected to increase pronunciation accuracy. However, with regard to vowel epenthesis, it is unclear to what extent L1 phonology transfers to L2 speech production and how this relates to the learners' proficiency. Therefore, in this study we examined how vowel epenthesis is realized in Japanese speakers' L2 English across different phonological environments, in relation to their English proficiency level.

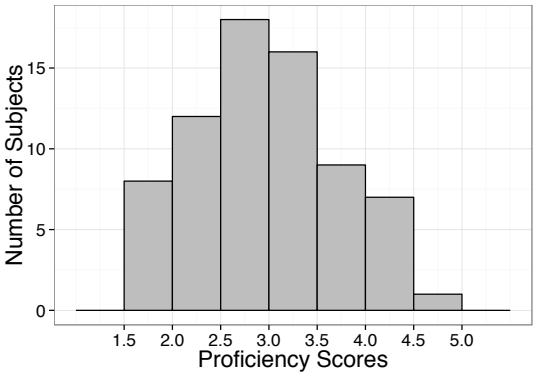
2. DATA

2.1. J-AESOP corpus

Data for analysis were obtained from the J-AESOP (Asian English Speech cOrpus Project) corpus [11].

This study used recordings from a text reading task where subjects read the Aesop fable "The North Wind and the Sun" [6], which is commonly used in English phonetic research because it contains all phonemes of English. The subjects were 71 native Japanese undergraduate and graduate students (24 male, 47 female) who were learning English as L2 at the time of recording. Each subject's English proficiency level was assessed by 8 English teachers (4 English speakers and 4 Japanese speakers) based on overall auditory impression of fluency and accuracy on a 9-point scale (1 = very poor, 2 = poor, 3 = average, 4 = good, and 5 = very good or native-like, with a 0.5 increment). The assessed values were then averaged to create a score for each subject, representing their English proficiency level. Fig. 1 shows the distribution of all subjects' scores.

Fig. 1: Assessed English proficiency scores of 71 subjects (Mean = 2.93)



2.2. Annotation

The audio data were automatically annotated using the Hidden Markov Model Toolkit (HTK) [4]. First, Hidden Markov Models (HMMs) were trained using the TIMIT speech corpus [2]. However, in order to accurately annotate Japanese speakers' English, the pronunciation dictionary of the TIMIT corpus cannot be directly used since it assumes only American English pronunciations. Thus, the dictionary was modified so that it also contained phonemes and phoneme sequences of Japanese accented English. For example, Table 1 lists pronunciations for the word "blew" in the dictionary. In addition to native English pronunciation [blu:], it contains different types of pronunciations reflecting both segmental (mispronouncing [l] as [r]) and suprasegmental (/u/ epenthesis) errors in Japanese accented English.

Using the TIMIT acoustic model and the modified dictionary, forced alignment was performed on all of the data. The automatic alignment output was further modified manually in order to exclude cases of inaccurate recognition by HTK and to make the alignment more reliable.

Table 1: Pronunciations for "blew" used for automatic annotation.

TIMITBET	IPA
b l uw	blu:
b r uw	bɹu:
b uh l uw	bʊlu:
b uh r uw	bʊɹu:

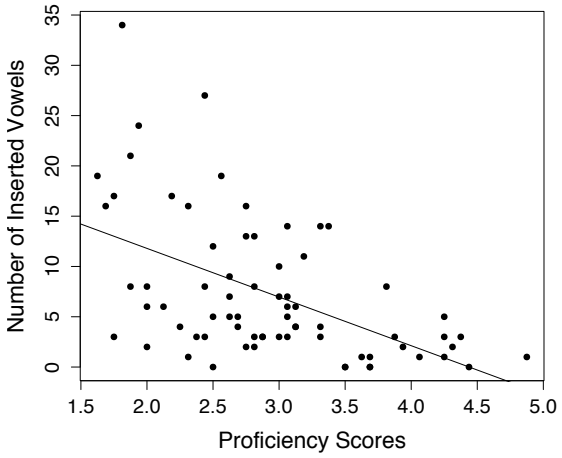
3. ANALYSIS AND RESULTS

Most of the subjects (65 out of 71) produced at least one epenthetic vowel. A total of 518 cases of observed vowel epenthesis were used for the analysis. In order to test the effect of proficiency level, subjects were divided into 2 groups based on the median proficiency score of 2.875: high (n = 35) and low (n = 36) proficiency groups.

3.1. Number of epenthetic vowels and L2 proficiency

Pearson's correlation test found an intermediate negative correlation (r = -0.5124, $p < .001$) between the number of epenthetic vowels and the learners' proficiency scores. However, it is notable that even advanced learners still produced some epenthetic vowels (See Fig. 2). This implies that the influence of Japanese CV syllable structure persists, making acquisition of English syllable structure difficult. Previous research on perception has also shown the rigidity of L1 syllable structure, demonstrating that Japanese CV structure is acquired by as early as 14 months of age [10].

Fig. 2: Number of epenthetic vowels in relation to the speaker's proficiency scores.



In our study, this was also shown by the subjects who produced epenthetic vowels both within consonant clusters and after word-final consonants. Other studies have shown that the number of epenthetic vowels in consonant clusters decreases as the learners' proficiency level increases [9]. Therefore, we conducted Poisson regression analysis to determine whether this is also the case for word-final consonants, and which phonological environment (within consonant clusters and after word-final consonants) is more likely to elicit vowel epenthesis. The analysis examined the number of epenthetic vowels given the phonological environment and proficiency level (high and low). The results showed that proficiency level was the only significant contributor ($p < .001$); the effects of phonological environments and the interaction between environments and proficiency levels were not significant. This suggests that while the number of epenthesis itself decreased as the learners' proficiency level increased, they still produced some epenthetic vowels under both phonological environments regardless of their proficiency level. This suggests that vowel epenthesis in L2 English of Japanese occurs to maintain the /CV/ structure, similarly to the patterns of loanword phonology.

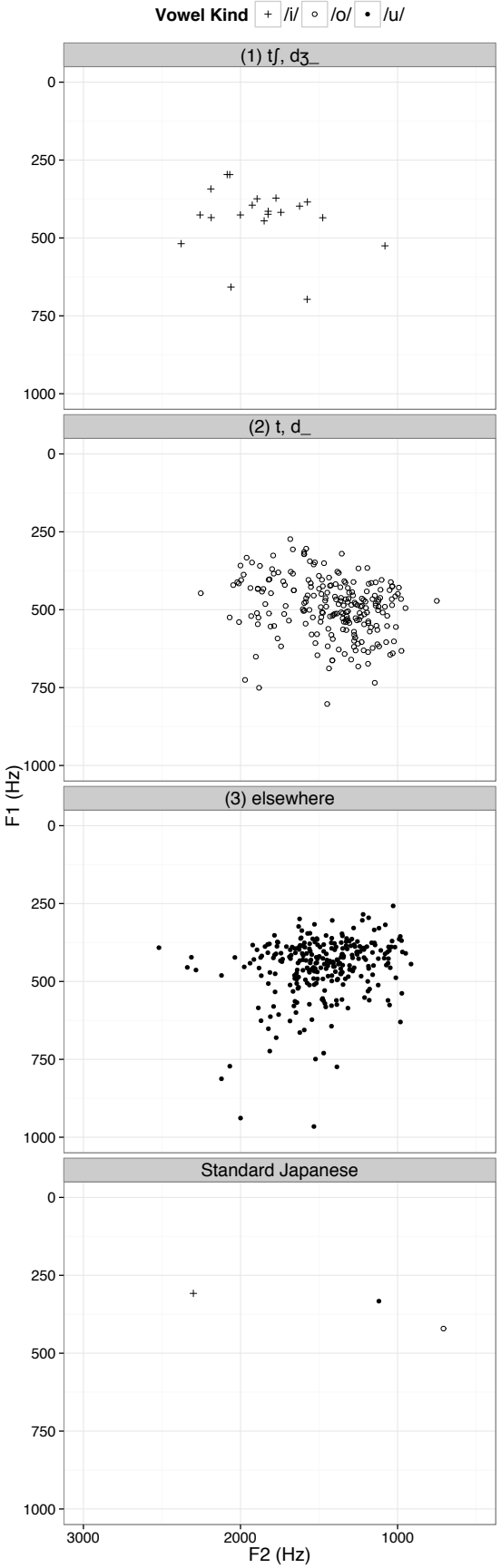
3.2. Vowel space of epenthetic vowels

Two-way repeated measures ANOVAs were conducted to examine the effect of preceding segments and proficiency level on the quality of epenthetic vowels (F1 and F2 values, measured by LPC). Preceding segments were classified into 3 context groups based on Japanese loanword phonology (See (1) - (3)): /i/ group (n = 20), /o/ group (n = 221), and /u/ group (n = 277). There were fewer /i/ insertions because the reading text contained only 2 words ("which" and "obliged") that have /tʃ/ and /dʒ/. The proficiency level groups (high and low) were the same as described previously.

The analyses showed that preceding context had a significant effect on both F1 ($F(2, 449) = 11.9$, $p < .001$) and F2 ($F(2, 449) = 20.02$, $p < .001$). However, there was no significant effect of proficiency level, nor any significant interaction between context and proficiency level. These results suggest that the preceding segment affects epenthetic vowel quality, but the quality does not change according to learners' proficiency.

Post-hoc pairwise comparisons with Bonferroni correction found significant differences in vowel quality between the three contexts. F1 values were significantly different between /o/ and /u/ groups ($p < .001$) and between /o/ and /i/ groups ($p < .05$).

Fig. 3: F1 and F2 values of vowels inserted after preceding consonants (1) - (3), and typical values of standard Japanese vowels [5].



There were also significant differences in F2 values between /i/ and /u/ groups ($p < .001$) and between /i/ and /o/ groups ($p < .001$). These results suggest that the quality of epenthetic vowels follows the pattern of Japanese loanword phonology, i.e. that an epenthetic vowel has a quality close to [o] when it follows /t, d/, or close to [i] after /tʃ, dʒ/, and close to [ɯ] if it follows any other segment.

Fig. 3 shows the F1 and F2 values of vowels inserted after (1) /tʃ, dʒ/, (2) /t, d/ and (3) other consonants. The typical F1 and F2 values of Standard Japanese vowels /o, i, u/ are also plotted (/i/: F1 = 308Hz, F2 = 2300Hz; /o/: F1 = 421Hz, F2 = 708Hz; /u/: F1 = 333Hz, F2 = 421Hz). As can be seen from the figure, the distributions of epenthetic vowels in the F1-F2 vowel space under three contexts were close to that of Standard Japanese vowels /o, i, u/.

4. CONCLUSIONS

The present study found that vowel epenthesis in Japanese learners' L2 English exhibited similar patterns to loanword adaptation in Japanese. While the number of vowel epenthesis decreased as the learners' proficiency level increased, even fluent speakers produced some epenthetic vowels both in consonant clusters and after word-final consonants, suggesting a rigidity of the L1 syllable structure. The quality of epenthetic vowels was found to be affected by L1 phonology, at least to some extent, regardless of the learners' proficiency level. The results indicate that the relationships between vowel epenthesis in Japanese speakers' L2 English and their level of proficiency are quantitative rather than qualitative.

5. ACKNOWLEDGEMENTS

This research was supported by the Waseda University Grant for Special Research Projects No. 2014B-439 and Grant-in-Aid for JSPS Fellows No.26-04006.

6. REFERENCES

[1] Anderson-Hsieh, J., Johnson, R., Koehler, K. 1992. The Relationship Between Native Speaker Judgments of Nonnative Pronunciation and Deviance in Segmentais, Prosody, and Syllable Structure. *Language Learning* 42(4), 529–555.

[2] The DARPA TIMIT Acoustic-Phonetic Continuous Speech Corpus (TIMIT). http:/www.ldc.upenn/readme_files/timit.readme.html

[3] Derwing, T. M., Munro, M. J., & Wieve, G. 2003. The effects of pronunciation instruction on accuracy, fluency, and complexity of L2 accented speech. *Applied Language Learning* 13(1), 1-17.

[4] Hidden Markov Model Toolkit (HTK) htk.eng.cam.ac.uk

[5] Imaishi, M., Miwa, J. 1989. Acoustic characteristics of vowels: differences in dialects. In: Sugito, M. (ed.), *Japanese Language and Japanese Language Education Vol. 2. Japanese Phonetics and Phonology (1)* (in Japanese). Meiji Shoin: Tokyo. 85-108.

[6] International Phonetic Association. 1999. *Handbook of the International Phonetic association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

[7] Kashiwagi, A., Snyder, M. 2008. American and Japanese listener assessment of Japanese EFL speech: Pronunciation features affecting intelligibility. *The Journal of Asia TEFL*. 5(4), 22-47.

[8] Kubozono, H. 1999. *Japanese phonetics*. Tokyo, Japan.

[9] Masuda, H., Arai, T. 2010. Processing of consonant clusters by Japanese native speakers: Influence of English learning backgrounds. *Acoustical Science and Technology* 31(5), 320 – 327.

[10] Mazuka, R., Cao, Y., Dupoux, E., Christophe, A. 2011. The development of phonological illusion: A cross-linguistic study with Japanese and French infants, *Developmental Science* 14(4), 693-699.

[11] Meng, H., Tseng, C. Kondo, M., Harrison, A., Visceglia, T. 2009. Studying L2 Suprasegmental Features in Asian Englishes; A Position Paper, *Proc. Interspeech 2009* Brighton, 1715-1718.

[12] Osterhout, L., Poliakov, A., Inoue, K., McLaughlin, J., Valentine, G., Pitkanen, I., Frenck-Mestre, C., Hirschensohn, J. 2008. Second-language learning and changes in the brain. *Journal of Neurolinguistics* 21(6), 509–521.

[13] Otaki, Y. 2012. A phonological account of vowel epenthesis in Japanese loanwords: Synchronic and diachronic perspectives. *Phonological studies* 15, 35-42.

[14] Patkowski, M. 1989. Age and accent in a second language: a reply to James Emil Flege. *Applied Linguistics* 11, 73-89.

[15] Raux, A., Kawahara, T. 2002. Automatic Intelligibility Assessment and Diagnosis of Critical Pronunciation Errors for Computer-Assisted Pronunciation Learning. *Proc. 7th ICSLP* Denver, 737-740.

[16] Shinohara, S. 2004. Emergence of universal grammar in foreign word adaptations. In: Kager, R., Pater, J., Zonneveld, W. (eds.), *Constraints in Phonological Acquisition*. 292-320.