Epenthesis, intrusion, or deletion? Vowel alternation in consonant clusters by Japanese ESL learners

Akitsugu Nogita
University of Victoria
akitsugu@uvic.ca

This study investigates whether vowel insertion in English consonant clusters produced by Japanese English-as-a-second-language (ESL) learners is due to misarticulation or misinterpretation. Intermediate level Japanese ESL learners read aloud written English words and mimicked auditory English words. The results showed that the participants inserted a vowel in consonant clusters notably less frequently in the mimicking task than in the reading task, suggesting that the participants can perceive and produce consonant clusters. The participants were also asked to divide each stimulus word into syllables, and they often clearly pronounced extra vowels: e.g. 'ba-d*o*-min-ton' for 'badminton.' I conclude that vowel insertion is not because of their inability to articulate consonant clusters, but their misinterpretation that there is a vowel where there is actually not. When they mimicked auditory stimuli, they phonetically deleted such vowels, but vowels still existed in their phonological representations.

1 Introduction

This study investigates whether the vowel insertion in English consonant clusters produced by Japanese English-as-a-second-language (ESL) learners is due to misarticulation or misinterpretation that there is a vowel between the consonants. It is well attested that second language (L2) learners may insert a vowel into a consonant cluster that is illegal in their first language (L1). There are three possible motivations, 'epenthesis', 'intrusion', and 'incorrectly perceived L2 input.' Briefly, epenthesis is a lexical vowel insertion which occurs to satisfy lexical syllabification, whereas intrusion does not involve a syllabic or moraic lexical vowel but a vowel-like sound occurs between a consonant cluster when the first consonant is released before the second consonant starts (Hall, 2006; Davidson & Stone, 2004). Another way to say this is that an epenthetic vowel is intentional, while an intrusive vowel is likely to be unintentional but is due to a gestural timing issue. I assume that Japanese ESL learners' vowel insertion is neither of these, but 'incorrectly perceived L2 input'; they misinterpret that there

is supposed to be a vowel. For example, they may assume that 'badminton' is supposed to be 'bad*o*minton.'

Dehaene-Lambertz, Dupoux, and Gout (2000) stated that Japanese speakers paid little attention to whether the vowel [u][1] is present or absent in phonotactically illegal consonant clusters, such as 'igmo' vs. 'igumo'. Funatsu et al. (2008) argued that Japanese speakers can detect consonant clusters that are illegal in Japanese. In their experiment, novice Japanese English-as-foreign-language (EFL) learners heard and mimicked English words with mostly [t/d] + [ɹ] clusters. The participants generally pronounced the clusters correctly without vowel insertion. Funatsu et al. also mentioned that occasional short vowel insertion was vowel intrusion, or gestural mistiming. However, according to their data, the inserted vowels were mostly [o] and [u], both of which are common Japanese epenthetic vowels. If these were really intrusion, these would have been [ə]-like sounds. Therefore, I assume that these vowels were not intrusion, but occurred at a more phonological level. Besides, Funatsu et al.'s consonant clusters were mostly in word-initial positions, and [t/d] and [ɹ] are distant in sonority; stops like [t/d] are low while approximants like [ɹ] are quite high in sonority hierarchy. Since word-initial segments are salient and consonant sequences with great sonority distance are less marked, I will use word-medial clusters with smaller sonority distance, such as obstruent-obstruent sequences, which are more marked.

2 Experiment

2.1 Participants

I recruited eight lower-intermediate to upper-intermediate Japanese ESL learners in Canada with relatively short length of residence (3 to 13 months) in an English speaking country. All of them were females in their 20's. They were all from Kantô or Chûbu regions. None of them reported a hearing problem.

2.2 Stimuli

The stimuli were 12 real words and 12 nonsense words that had consonant clusters [b, d, g] + [obstruent, n, m, l] in a word-medial position, which are phonotactically illegal in Japanese. The real words were considered familiar to Japanese speakers as loanwords. All the real words had primary stress on the first syllable. Nonsense English words were made based on English phonotactics. I avoided the vowel /u/, which is phonetically similar to the default

[1] The high non-front vowel in standard Japanese is typically realized as [ɯ] which does not have lip rounding or lip protruding (Tsuzuki, 1996). Since in standard Japanese [ɯ] and [u] are not phonologically contrastive, Funatsu et al. (2008) and Dehaene-Lambertz (2000) used [u]. When I cite them, I use [u]. Otherwise, I use [ɯ].

Japanese epenthetic vowel /ɯ/ (Strange, et al. 2008), as well as /oʊ/ after /d/ which is phonetically similar to the common Japanese epenthetic vowel /o/ after /d/ as in 'bad*o*minton.' There were also eight real and eight nonsense word fillers. Since most of the crucial items were disyllabic, I made most fillers not disyllabic. Including fillers, there were 20 real words and 20 nonsense words.

These 20 real words were randomized in order, and so were the nonsense words, and the written stimuli and audio stimuli were in different order. The written stimuli were printed on a sheet. The audio stimuli were produced by a phonetically trained female native speaker of Canadian English. The stimuli were recorded in a booth in the UVic Linguistics Speech Research Lab. Following are the stimuli. Parenthesized words are fillers.

1. Real words: subject, webmail, webnet, tablet, foodbank, badminton, Sydney, badly, rugby, eggman, magnet, ugly, (avocado, banana, coconut, fruit, grape, ice-cream, strawberry, vegetable)

2. Nonsense words: ebdet, gabmee, gabno, cabla, idgay, cadma, pednay, edlee, agday, egmad, hegneb, agla, (ba, cantukpeg, gamboozee, jeejee, ma, muzz, smecks, sna)

2.3 Procedure

There were four tasks: two production tasks and two syllabification tasks. In the first task, the participants were recorded reading aloud the written real words and nonsense words. The recording was done with the software Audacity set at 44100 Hz and 32-bit float in the UVic Phonetics Lab. In the second task, the participants were asked whether they know what 'syllable' (or '*onsetsu*' in Japanese) was, and they were asked to separate each written word into syllables. I demonstrated how to divide the Japanese word '*wasabi*' into 'wa-sa-bi' by making a pause between syllables, and explained that the monosyllabic Japanese word '*ka*' (mosquito) could not be divided into a smaller unit. The participants who did not know syllables were instructed to syllabify according to their impression. In the third task, they listened to each stimulus without looking at written cues and immediately mimicked the stimulus. To avoid practice effects, the participants heard each stimulus only once, except for a few cases when the participants could not say anything. In the fourth task, the participants listened to each word and divided it into syllables without looking at written cues.

3 Results and Discussion

3.1 Results of the Production Tasks

The participants generally tended to insert a vowel in consonant clusters but there were different tendencies between the reading and mimicking tasks. All the participants inserted a vowel more frequently and mean duration of the vowels tended to be longer in the reading task. These results agree with Funatsu et al.'s (2008) study. The participants sometimes clearly released or aspirated the first consonant although there were acoustically no periodic pitch pulses. They occasionally made the first consonant and an inserted vowel coalescent, as in [g] and the inserted vowel in 'rug(ɯ)by' coalescing to form [ɣ]. Release, aspiration, and coalescence tended to occur more frequently in the mimicking task. Such productions were counted as incorrect productions but not as vowel insertions. Overall, all the participants correctly pronounced consonant clusters more frequently in the mimicking task. Table 1 and 2 show the number of vowel insertion, mean duration of inserted vowels, the number of release, aspiration, or coalescence, and the number of correct productions of consonant clusters in both reading and mimicking tasks. Irrelevant productions, such as pronouncing a wrong word, were discarded.

Table 1.
Productions of consonant clusters in the real words. Nb: 'Re' = reading task; 'Mi' = the mimicking task; '# of V Ins.' = the number of vowel insertion; 'M. Duration' = mean duration of the inserted vowels; '# of Rel, Asp, Coalescence' = the number of release, aspiration, or coalescence; 'Correct' = the number of correct productions.

Participants	P1		P2		P3		P4		P5		P6		P7		P8	
Stimuli	Re	Mi	Re	Mi	Re	Mi	Re	Mi	Re	Mi	Re	Mi	Re	Mi	Re	Mi
# of V Ins. M. Duration	9 32ms	7 25ms	8 24ms	5 22ms	6 43ms	1 7ms	11 26ms	7 27ms	9 59ms	5 30ms	10 32ms	7 25ms	12 55ms	5 36ms	2 17ms	2 15ms
# of Rel, Asp, Coalescence	1	2	3	3	1	3	0	4	2	2	2	1	0	4	2	0
Correct	**1**	**3**	**1**	**4**	**5**	**8**	**1**	**1**	**1**	**4**	**0**	**4**	**0**	**3**	**8**	**10**

Table 2.
Productions of consonant clusters in the nonsense words.

Participants	P1		P2		P3		P4		P5		P6		P7		P8	
Stimuli	Re	Mi	Re	Mi	Re	Mi	Re	Mi	Re	Mi	Re	Mi	Re	Mi	Re	Mi
# of Ins. M. Duration	9 21ms	1 20ms	5 29ms	1 27ms	9 71ms	0	12 30ms	2 22ms	8 68ms	5 17ms	8 49ms	4 35ms	12 47ms	6 29ms	9 26ms	1 35ms
# of Rel, Asp, Coalescence	0	4	1	4	1	3	0	6	1	1	4	3	0	3	0	2
Correct	**3**	**6**	**3**	**7**	**2**	**9**	**0**	**4**	**2**	**6**	**0**	**5**	**0**	**3**	**3**	**9**

3.2 Results of Syllabification Tasks

Reportedly, only P1 and P8 had received formal instruction on English syllabification in Canada, but not in Japan, about a month and half a year prior to the experiment respectively. The summary of the number of errors is shown in Table 3. The results show that only P8 correctly syllabified all the words, and P1 made fewer errors than the others. P8 performed better probably because she had known syllable for longer than P1. According to Ueyama (2003) Japanese ESL learners with more than five years of residence in the U.S. may not naturally acquire English syllabification. However, these results suggest that formal instruction can greatly help Japanese ESL learners aware of it.

Table 3.
Errors in the syllabification tasks. Nb: P1 and P8 had explicit knowledge of English syllabification.

	*P1	P2	P3	P4	P5	P6	P7	*P8
Real words reading – mimicking	4 – 4	17 – 8	4 – 7	7 – 8	8 – 7	16 – 16	6 – 9	0 – 0
Nonsense words reading – mimicking	1 – 1	17 – 11	4 – 3	2 – 4	6 – 2	17 – 17	3 – 5	0 – 0
Overall reading – mimicking	5 – 5	34 – 19	8 – 10	9 – 12	14 – 9	33 – 33	9 – 14	0 – 0

The participants' error patterns were quite inconsistent, such as mora-based, foot-based, morpheme-based, or others. For example, P2 divided 'webnet' into 'we-b(ɯ)-ne-t(^{h})' referring to a mora while dividing 'icecream' into 'i-ce-cream'. This implies that the participants did not know what to do in the syllabification task. I interpret that their errors were random, or pre-systematic errors. What is interesting is the way of their pronunciation. For example, P6 correctly pronounced the consonant cluster in 'eggman' without releasing /g/ in the mimicking task, while when syllabifying the audio stimulus, she pronounced [ɛ-gɯ-mæn] with clear [ɯ] after [g]. The other participants also tended to add clear [ɯ] or [o] after voiced consonants and add aspiration or devoiced vowels [ɯ̥] or [o̥] after voiceless consonants: e.g. [sʌ-bɯ-dʒɛ-kɯ̥-to̥] 'subject.' This implies that the participants were able to pronounce target-like consonant clusters at the phonetic level, but there were vowels at their phonological underlying level, or in their mind.

3.3 Discussion

In Nogita (2010) I found that many Japanese ESL learners are never taught the basic English symbol-sound correspondence rules. Therefore, it is likely that Japanese ESL learners built their own English symbol-sound correspondence rules. Considering the fact that they often added an extra vowel, Japanese

learners' perception of the English orthography is like abugida, or alphasyllabary, rather than alphabet; each consonant letter, probably except for <N>, has a default following vowel which is pronounced every time the consonant requires a following vowel according to Japanese phonotactics; for example, the italicized consonant letters in 'su*b*je*c*t' and 'we*b*mai*l*' are not followed by a vowel letter, but are pronounced as [bɯ], [kɯ], [to], [bɯ], and [lɯ] (or [ɾɯ]) respectively.

As for the mimicking task, the participants more frequently produced consonant clusters correctly. This indicates that at least Japanese ESL learners with several months of exposure to native English can produce consonant clusters correctly. As mentioned in §3.1, the participants sometimes released the first consonants in clusters, or devoiced /b/, /d/, and /g/ and aspirated them, especially in the mimicking task. Goad et al. (2003) reported the same tendency. My participants' aspiration in a consonant cluster is considered as a voiceless vowel, often [ɯ̥]. This interpretation is consistent with the interpretation that there is a vowel at their' underlying representation (UR). Urbanczyk (1996) reported a similar phenomenon in Salish; there is syllabic aspiration, which is actually a voiceless schwa. When the participants released the first consonant, such as [b] in 'subject', the release was actually /ɯ/ in their mind, which was phonetically minimized. Figure 1 shows the comparison among 'epenthesis', 'intrusion', and Japanese ESL learners' interpretation of English words where a vowel exists in the first place, which is allophonically *weakened* or *deleted*. For example, if a learner knows that 'gb' in 'rugby' is a consonant cluster but adds a lexical vowel [ɯ] in order to make it fit in his/her L1 phonotactics, it is epenthesis. If a learner tries to produce [gb] but fails to coordinate two consonants in production, it is intrusion. If a learner misinterprets 'gb' as /gɯb/ but this /ɯ/ is altered to [ɯ̥], ø and so forth at the surface level, it is weakening or deletion.

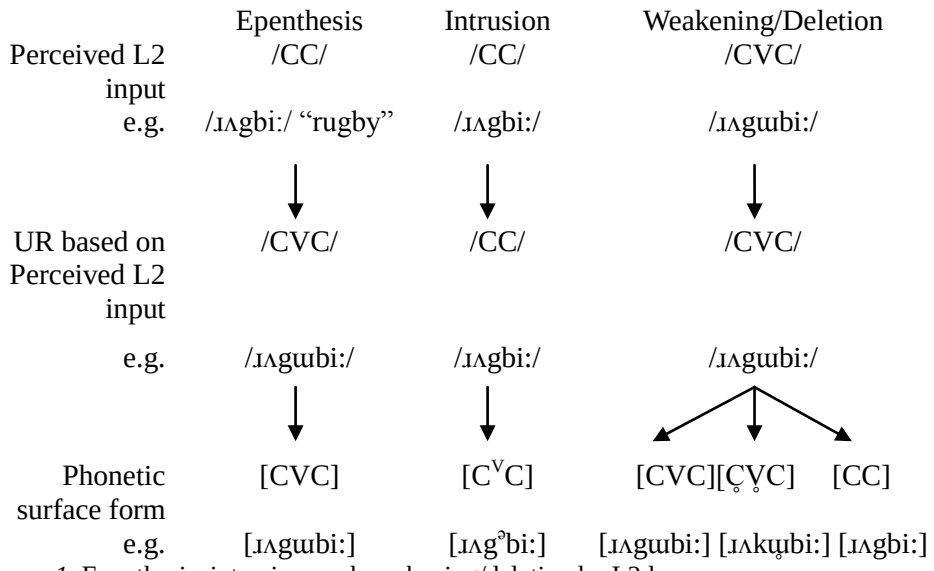

Figure 1. Epenthesis, intrusion, and weakening/deletion by L2 learners.

My participants' process was 'Weakening/Deletion'; the participants started from /CVC/ which could end up with [CVC] with a full vowel, [C̥V̥C] (or [C̥^{h}C]) with a voiceless vowel, [CC] with release of the first consonant, or [C$^{\neg}$C] without release or the target-like cluster, by phonetically adjusting the vowel. These alternations are considered as free allophonic variations. Japanese ESL learners incorrectly assume that there is a vowel in the first place, which causes vowel insertion despite their ability to produce consonant clusters. Examples of the participants' vowel alternations are shown in Figure 2 to 5.

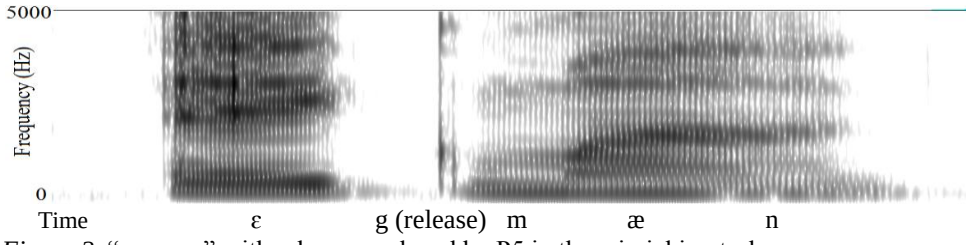

Figure 2. "eggman" with release produced by P5 in the mimicking task

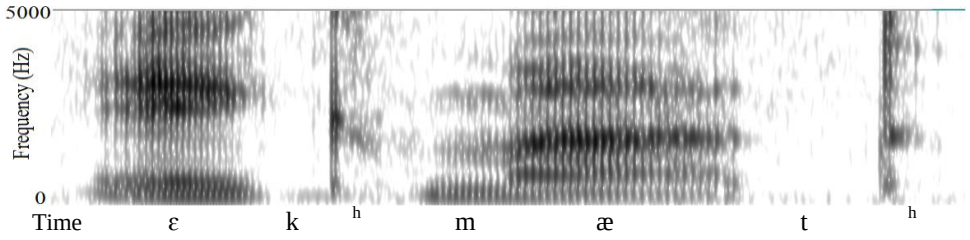

Figure 3. "egmad" with aspiration produced by P4 in the mimicking task

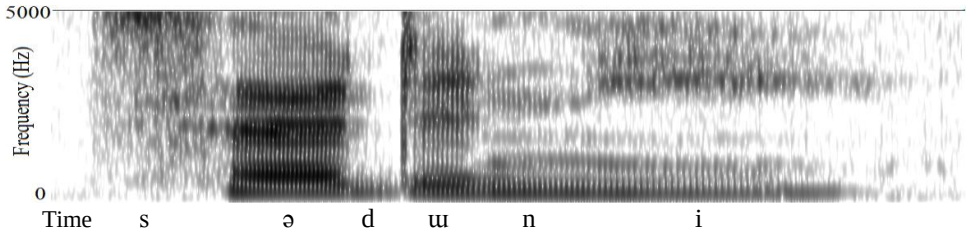

Figure 4. "Sydney" with clear [ɯ] insertion produced by P3 in the reading task

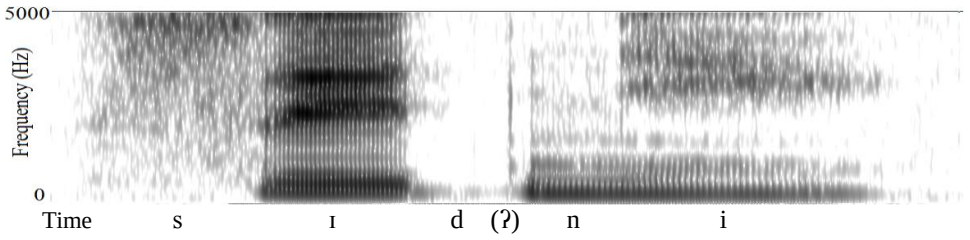

Figure 5. "Sydney" with no vowel insertion produced by P3 in the mimicking task

4 Conclusion

Despite a relatively short length of residence, all the Japanese participants perceived and produced English consonant clusters at the phonetic level. However, at least seven of my participants misinterpreted that there was a vowel where there was actually not: e.g. they assumed that 'badminton' was supposed to be 'bad*o*minton'. This misconception is considered to come from a lack of formal instruction of English syllabification and basic symbol-sound correspondence rules. Due to a lack of explicit knowledge, they may have developed their own system and considered English alphabet as abugida in which each consonant letter has a default vowel. This causes their vowel insertion in consonant clusters. At the phonetic level, in order to imitate native English speakers' production, they weakened or deleted a vowel that existed in the first place in their own English phonological interpretation. Such vowel weakening/deletion was free allophonic variations, but did not function phonemically, which was the real problem for those Japanese ESL learners. In short, Japanese ESL learners know *how* to pronounce consonant clusters, but do not know *when* to pronounce them. For future studies, it is necessary to design formal instruction for helping Japanese ESL learners *understand* (not articulate and perceive) English consonant clusters and syllabification.

Acknowledgements

I am grateful to Dr. Ewa Czaykowska-Higgins for her guidance in this study and expertise in phonology, and to my colleague, Yanan Fan, for her creative ideas.

References

Coulmas, S. (2003). *Writing System*. New York: Cambridge University Press.

Dehaene-Lambertz, G., Dupoux, E. & Gout, A. (2000). Electrophysiological correlates of phonological processing: A cross-linguistic study. *Journal of Cognitive Neuroscience, 12*(4), 635.

Davidson, L., & Stone, M. (2004). Epenthetsis Versus Gestural Mistimimg in Consonant Cluster Production: An Ultrasound Study. In G. Garding & M. Tsujimura (eds.), *WCCFL 22 Proceedings*. Cambridge, MA: Cascadilla Press, 165-178.

Funatsu, S., Imaisumi, S., Fujimoto, A., Hashizume, M., & Kurisu, K. (2008). Do Japanese speakers perceive nonexistent vowels in non-native consonant clusters?. *The Journal of the Acoustical Society of America, 123*(5), 3072.

Goad, H. & Kang, H. (2003). Word-final syllabification in L2 acquisition with emphasis on Korean learners of English. In Juana M. Liceras, Helmut Zobl & Helen Goodluck (eds.) *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference*. Somerville, Mass.: Cascadilla Press. 122–129.

Hall, N. (2006). Cross-linguistic patterns of vowel intrusion. *Phonology 23*. The United Kingdom: Cambridge University Press. 387-429.

Nogita, A. (2010). *Is [si] really problematic?: Examination of the [si] and [ʃi] confusion by Japanese ESL learners*. Saarbrücken, Germany: VDM Verlag.

Strange, W., Akahane-Yamada, R., Kubo, R., Nishi, K., & Trent-Brown, S. (2008). Acoustic and perceptual similarity of Japanese and American English vowels. *The journal of the Acoustical Society of America, 124*(1), 576-588.

Tsuzuki, M. (1996). An allophonic study of Japanese vowels and consonants. *The Journal of Aichi Gakuin University. Humanities & sciences, 43*(2), 41-66.

Vance, T. J. (2008). *The sound of Japanese*. New York: Cambridge University Press.

Urbanczyk, S. (1996). *Patterns of Reduplication in Lushootseed*. Ph.D. dissertation, University of Massachusetts.

Ueyama, M. (2003). Awareness of L2 syllable structure: The case of L2 Japanese and L2 English. *Journal of the Phonetic Society of Japan, 7*(2), 84-100.