



# FACTORS AFFECTING NATIVE JAPANESE SPEAKERS' PRODUCTION OF INTRUSIVE (EPENTHETIC) VOWELS IN ENGLISH WORDS

*Keiichi Tajima\**, *Donna Erickson†*, and *Kyoko Nagao‡*

\*Information Sciences Division, ATR International, Kyoto, Japan

†Gifu City Women's College, Gifu City, Japan

‡Konan University, Kobe, Japan

## ABSTRACT

A salient characteristic of English spoken by native Japanese is the tendency to produce intrusive vowels between consonants or following word-final consonants (e.g., English "stress" as /su.to.re.su/). This study conducted speech analysis of a large set of Japanese-accented English utterances to evaluate the effects of various factors on vowel epenthesis. Results reveal effects of speaking rate and phonetic environment, but little effect of word familiarity and frequency.

## 1 INTRODUCTION

Adult native speakers of Japanese typically undergo six years of mandatory English education during secondary school (age 12–18), in which the primary focus is on grammar rather than on conversation. Thus, native Japanese speakers typically have better skills with written English than with the production and perception of spoken English. A salient characteristic of English produced by Japanese is the production of intrusive, or epenthetic, vowels within consonant clusters and following word-final consonants. For example, for English words such as "rugby" and "strange", it is not uncommon for Japanese speakers to realize them as /ragubi/ and /storenɕi/, respectively.

The common occurrence of epenthetic vowels in Japanese-accented English is usually attributed to phonological differences between Japanese and English, particularly with respect to constraints on syllable structure. In fact, many English words have been borrowed into Japanese as loanwords. But in this process the English words are assimilated to the syllable structure constraints of Japanese, which is based primarily on open (consonant-vowel) syllables (Katayama, 1998). It is possible that English words that exist as Japanese loanwords may be produced in a "loanword-like" fashion by Japanese speakers.

While it is true that spoken English is not taught extensively at school, it is not an exaggeration to claim that Japanese are constantly exposed to spoken English in other contexts, e.g., through the mass media (imported music,

radio shows, etc). It is therefore possible that production skills in English may be related to the degree of familiarity Japanese speakers have with spoken English words.

In addition to the above effects, vowel epenthesis may also be influenced by certain conditions under which productions are elicited from Japanese speakers. It is possible, for example, that a slower speaking rate leads to more instances of vowel epenthesis than a normal rate, since a slow rate of speech might induce a clear and careful articulation of individual consonants, accompanied by an audible vowel-like release following each consonant. It is also possible that Japanese speakers are better able to produce nativelike utterances if the utterance consists of a single isolated word than if it consists of a sentence.

Finally, vowel epenthesis may be influenced by phonetic properties of the neighboring consonants. For example, it is possible that audible and visible epenthetic vowels are more likely when surrounded by voiced consonants than by voiceless consonants, since vowels surrounded by voiceless consonants may undergo vowel devoicing, a common phonological process in Japanese. Also, stop consonants may be more likely to be released and followed by a noticeable vowel-like segment than affricates or fricatives, since the articulation of stops involves build-up of air pressure behind the constriction but does not involve generating sustained frication noise after constriction release. Likewise, it is possible that bilabial stops are more likely followed by a vowel-like segment than alveolar or velar stops, since an anterior place of articulation allows greater pressure build-up inside the vocal tract than a posterior one.

Even though vowel epenthesis in Japanese-accented English is well-known, few studies have investigated how various experimental and stimulus-related factors influence this process. There is evidence that vowel epenthesis and other deviations in syllable structure in foreign-accented English lead to reduced intelligibility by native English listeners (Tajima et al., 1997). However, before we can address the processes by which English syllable structure is acquired by Japanese, it is necessary to understand the extent to which various factors influence the production of epenthetic vowels. The aim of this study is to investi-

gate vowel epenthesis in Japanese-accented English on the basis of speech analysis, and to evaluate the effect of the following factors: speaking condition, word familiarity and frequency, and phonetic environment.

## 2 METHODS

### 2.1 Database Collection

A list of 90 English words were compiled. The words were selected such that the word set as a whole contained representative consonant clusters of English. Consonant clusters were classified according to word position, number of consonants in the cluster, as well as voicing, manner, and place of each consonant.

Some of the words existed in Japanese as loanwords, written using katakana script (e.g., basketball, platform). Others were expected to be English words that Japanese students encounter during secondary-school English education (e.g., children, explain). Other words were expected to be unfamiliar English words (e.g., glimpsed, thrive).

Participants were 26 monolingual native Japanese students from Doshisha University, Kyoto, Japan (15 females and 11 males, between 18 and 21 years of age). Most subjects spoke the Kansai dialect of Japanese. The subjects have received six years of English education in secondary school, in which the primary focus is on grammar rather than on conversation. None of the subjects had experience living in an English-speaking community. In addition, a control group of 6 native American/Canadian English speakers also participated (3 females, 3 males, between 22 and 42 years of age). They were visiting researchers at ATR Laboratories, Kyoto, Japan. Their length of residency in Japan ranged from one to five years.

The recordings took place in an anechoic chamber at ATR Laboratories. During each subjects' recording, (s)he was given a randomized list of the words to be produced. For each item, subjects heard the item produced by a male native American English speaker; this ensured that the subjects knew the pronunciation of each word. Subjects read through the 90 words in isolation at two speaking rates, normal and slow. In addition, 30 of the 90 words were embedded in the carrier sentence "I say \_\_\_ in the list", and were read at a normal rate. The order of the three speaking conditions was counterbalanced across subjects. The utterances were recorded onto a DAT, and later saved as audio files at 22050-Hz sampling rate.

Speaking rate difference between the slow and normal conditions was estimated by measuring word duration in tokens of ten test words selected at random. On average, word duration was 25.8% longer ( $SD = 24.3$ ) in the slow-rate condition than in the normal-rate condition.

### 2.2 Speech Analysis

Judgment concerning the presence vs. absence of epenthetic vowels in the utterances was obtained separately from three trained phoneticians (the authors of this paper). One of them is a native English speaker, one is a native Japanese speaker, and one is a bilingual in both

languages. All received a graduate degree in Linguistics in the United States.

For each utterance, a waveform, spectrogram, fundamental frequency contour, amplitude envelope, and probability of voicing profile were displayed using ESPS and waves+. The general criterion for marking an epenthetic vowel was that it had to be both audible and visible on the waveform or spectrogram. Care was taken to ensure that epenthetic vowels were clearly visible, by enforcing such specific criteria as the following: (1) Clear vowel-like formant structure must be visible on the spectrogram, particularly the first formant. (2) The waveform must show periodicity from phonation. (3) Very short epenthetic vowels must contain at least two pitch periods. (4) Epenthetic vowels that are adjacent to another consonant with formant structure (e.g., a nasal or liquid) must show evidence that identifies them as vowels (e.g., presence of a noticeable change in amplitude or a rapid formant transition).

The judges were instructed to put a label on any "extra" vowel that was not a proper part of the word, which met the criteria above, including vowels appearing within consonant clusters and vowels following word-final consonants.

The Japanese speakers' utterances presented many "borderline" cases in which the presence vs. absence of an epenthetic vowel was not clear-cut even after careful inspection. Out of 17978 possible positions in the speakers' utterances in which an epenthetic vowel could potentially occur (i.e., between every pair of adjacent consonants and following word-final consonants), the three judges agreed on 90.2% of the judgments. That is, for 90.2% of the time, all judges agreed that an epenthetic vowel was present in a given position or that it was absent. For the remaining 9.8%, the judges gave mixed responses. Therefore, the three judges were generally in good agreement with one another.

## 3 RESULTS

### 3.1 Effect of Speaking Condition

We first examine how vowel epenthesis in Japanese-accented English is influenced by the speaking condition. The rate of occurrence of epenthetic vowels was calculated for 30 test words that were produced in all three speaking conditions. Percentage of tokens that contained an epenthetic vowel within consonant clusters was calculated for each judge. Epenthetic vowels produced word-finally were not included, since they are difficult to detect adjacent to the carrier phrase in the embedded condition. Mean rate of occurrence of epenthetic vowels (and standard deviations across the three judges) are shown in Table 1 for the three conditions and for the English and Japanese speakers.

The table indicates that compared to the normal-rate condition, vowel epenthesis is more frequent in the slow-rate condition, but not much more frequent in the embedded condition. Thus, while speaking slowly induced speakers to produce intrusive vowels within consonant clusters,

**Table 1:** Rate of occurrence of epenthetic vowels in words produced under the three speaking conditions. Means (and standard deviations) across the three judges are shown.

Condition	English	Japanese
Isolated/Normal	3.0 (0.3)	21.5 (3.5)
Isolated/Slow	5.4 (0.8)	28.2 (4.0)
Embedded/Normal	3.9 (1.7)	22.5 (4.5)

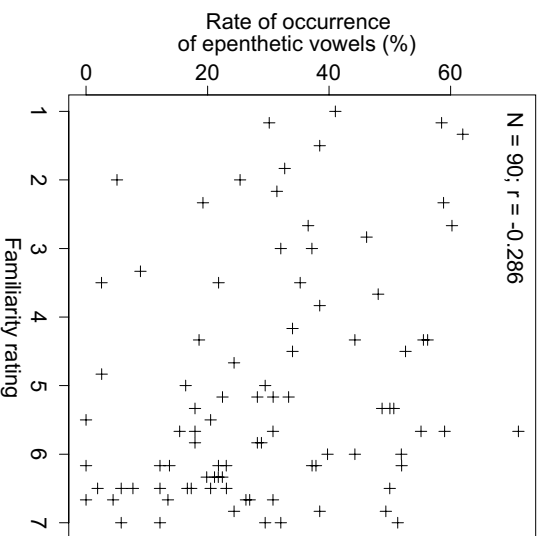
producing words in isolation did not lead to fewer cases of vowel epenthesis.

### 3.2 Effect of Word Familiarity and Frequency

It is possible that Japanese speakers’ familiarity with native-speaker pronunciation of English words enables them to produce the words in a relatively native-like fashion, with fewer instances of vowel epenthesis. To test this, six native Japanese speakers (4 females, 2 males, between 24 and 31 years of age, all of whom were different from the 26 speakers in the database) were asked to rate the degree to which they were familiar with each English word on a scale from 1 (very unfamiliar) to 7 (very familiar). Each word was shown visually on the computer screen; also, a male native speaker’s production of the word was presented through headphones. Subjects were not given specific instructions to rate their familiarity with the written forms or the spoken forms of the words.

Figure 1 plots the mean familiarity rating for each word on the  $x$ -axis, and the rate of occurrence of epenthetic vowels for the word on the  $y$ -axis. The latter was calculated based on tokens in the isolated/normal and isolated/slow conditions. The percentage of tokens that contained an epenthetic vowel was computed for each word, averaged across the judges. The figure shows that points are more densely distributed near the lower right-hand corner of the plot than other regions. There was a slight but significant tendency for words that were judged to be familiar to have fewer instances of vowel epenthesis than those judged unfamiliar [Pearson’s  $r = -0.29$ ;  $t(88) = -2.65$ ;  $p = .006$ ].

In addition to degree of familiarity, the existence of Japanese loanwords for some of the English words may have led Japanese speakers to produce the words in a ‘loanword-like’ fashion, resulting in more instances of vowel epenthesis. To test this, we took 58 of the 90 test words which contained no inflectional suffixes (e.g., past tense or plural morphemes), and determined whether a Japanese loanword existed for each word. This was done by searching in the NTT Lexical Properties of Japanese database (Amano and Kondo, 2000), which contains a text corpus of 14 years of Japan’s Asahi newspaper articles. Among the words that appeared in the database as loanwords ( $N = 34$ ), mean rate of occurrence of epenthetic vowels was 20.4% (SD = 13.1). Among words that were not found as loanwords ( $N = 23$ ), the mean was 30.7% (SD = 14.8). The difference was marginally significant [ $t(55) = 2.11$ ;  $p = .039$ ]. Thus, the results suggest a trend



**Figure 1:** Rate of occurrence of epenthetic vowels in each word as a function of native Japanese speakers’ subjective familiarity with the word.

in the opposite direction; vowel epenthesis was on average more common in words that did not exist as loanwords in Japanese than in words that did.

From the words that did exist as Japanese loanwords, we further obtained frequency counts from the same database. Frequency ranged from 1 (e.g., ‘chapter’, ‘photograph’) to greater than 2000 (e.g., ‘tank’, ‘rugby’). No significant correlation was found between  $\log(\text{frequency})$  and rate of occurrence of epenthetic vowels for each word [Pearson’s  $r = -0.20$ ;  $t(35) = -1.22$ ;  $p = 0.23$ ].

Altogether, the results suggest a moderate tendency for high-familiarity words to exhibit fewer instances of epenthetic vowels than low-familiarity words. However, the presence of Japanese loanwords or their frequency of use are not good predictors of the rate of occurrence of epenthetic vowels.

### 3.3 Effect of Neighboring Consonants

We now examine the degree to which vowel epenthesis in Japanese-accented English is influenced by properties of the neighboring consonants. To evaluate these factors, we compare the rate of occurrence of epenthetic vowels among 10 words in the database which contained word-final 2-consonant clusters varying in consonant voicing, manner, and place (‘grape<sup>s</sup>’ /ps/, ‘accept’ /pt/, ‘magnets’ /ts/, ‘six’ /ks/, ‘looked’ /kt/, ‘cube<sup>s</sup>’ /bz/, ‘described’ /bd/, ‘add<sup>s</sup>’ /dz/, ‘legs’ /gz/, ‘bragged’ /gd/). In Table 2(a)–(c), the clusters are organized according to different consonantal properties. The value in each cell was calculated by first computing for each judge the mean rate of occurrence of epenthetic vowels across all the clusters in the given row, then taking the mean across the three judges. The values in parentheses show the standard deviation across the judges. For the present purpose, we ignore the presence of morpheme boundaries in some of the clusters in

**Table 2:** Rate of occurrence of epenthetic vowels within word-final 2-consonant clusters, compared with respect to different consonantal properties. Means (and standard deviations) across the three judges are shown.

Cluster	English		Japanese	
(a) voiceless vs. voiced				
ps, pt, ts, ks, kt	0.0	(0.0)	4.6	(1.7)
bz, bd, dz, gz, gd	4.4	(2.5)	31.3	(5.6)
(b) labial vs. alveolar vs. velar				
ps, bz	4.2	(0.0)	27.2	(4.7)
ts, dz	0.0	(0.0)	0.6	(1.1)
ks, gz	5.6	(4.8)	15.4	(3.9)
(c) stop-fricative vs. stop-stop				
ps, bz, ks, gz	0.7	(1.2)	23.3	(3.6)
pt, bd, kt, gd	4.9	(2.4)	21.3	(2.9)

**Table 3:** Rate of occurrence of epenthetic vowels following singleton word-final consonants.

Cluster	English		Japanese	
(a) voiceless vs. voiced				
p, t, k, f, s, ʃ	0.5	(0.8)	2.8	(3.7)
b, d, g, v, z, ʒ	19.0	(6.4)	14.7	(2.3)
(b) labial vs. alveolar vs. velar				
p, b	20.8	(4.2)	16.7	(5.3)
t, d	12.5	(7.2)	11.9	(2.0)
k, g	16.7	(8.3)	9.6	(0.6)
(c) stop vs. affricate vs. fricative				
t, d	12.5	(7.2)	11.9	(2.0)
ʃ, ʒ	1.4	(2.4)	2.3	(2.4)
s, z	0.0	(0.0)	6.4	(1.1)

this table.

Table 2(a) compares the mean rate of occurrence of epenthetic vowels in voiceless-voiceless vs. voiced-voiced clusters. The table clearly shows that Japanese speakers produce many more instances of epenthetic vowels in voiced environments than in voiceless environments.

Similarly, Table 2(b) compares the mean percentages for clusters with different places of articulation of the first consonant. For Japanese speakers, the rate of occurrence of epenthetic vowels is high for /-ps/ and /-bz/ clusters, low for /-ts/ and /-dz/ clusters, and intermediate for /-ks/ and /-gz/ clusters. The low percentages for the alveolar clusters may be due to the existence of /ts/ and /dz/ as affricates in Japanese.

Table 2(c) compares the means for clusters ending in an alveolar stop (/t d/) vs. fricative (/s z/). No clear differences are found between these types of clusters.

Vowel epenthesis in Japanese-accented English occurs not only within consonant clusters but also following word-final consonants. The rate of occurrence of epenthetic vowels following word-final singleton consonants was com-

pared for 12 different words, (“raindrop” /p/, “admit” /t/, “sneak” /k/, “photograph” /f/, “actress” /s/, “stretch” /ʃ/, “tribe” /b/, “ski” /d/, “clog” /g/, “thrive” /v/, “spies” /z/, “oblige” /ʒ/). In Table 3(a)–(c), the percentages are collapsed with respect to various consonantal properties. Table 3(a) shows that voicing again has a strong effect on the rate of occurrence of epenthetic vowels. Table 3(b) reveals somewhat more instances of epenthesis after bilabial stops than after alveolar or velar stops. Finally, Table 3(c) shows a small tendency for stops to be followed by epenthetic vowels more often than affricates or fricatives. These effects are witnessed not just in the Japanese productions but also in the native productions.

## 4 DISCUSSION

In this study, production of epenthetic vowels was analyzed in a corpus of English words produced by native Japanese speakers. Results indicate that the rate of occurrence of epenthetic vowels is not dramatically influenced by how familiar Japanese speakers are with a given English word, or whether and how often an English word appears as a loanword in Japanese. On the other hand, vowel epenthesis is substantially influenced by speaking rate. It is also influenced by properties of the neighboring consonants, such as their voicing, manner, and place, with voicing having a particularly strong effect. Altogether, this study provides experimental evidence for the common occurrence of epenthetic vowels in Japanese-accented English, and provides quantitative estimates of the effect of various factors on the phenomenon.

It seems that the high rate of occurrence of epenthetic vowels in Japanese-accented English reflects the strong influence of the native phonological system on the production of speech in a second language. An analogous result has been found in the perception of English syllable structure by native Japanese listeners. We have found that Japanese speakers have difficulty counting the number of syllables in spoken English words, and that their level of difficulty increases as the number of consonants in the words increases (Tajima et al., 2000). Future work is needed to relate the perception results with the results from the present study, and to further understand the mechanisms underlying acquisition of prosodic properties of a second language.

## REFERENCES

- Amano, S. and Kondo, T. (2000). *Nihongo-no Goi-Tokusei (Lexical Properties of Japanese) Volume 7*. Sanseido, Tokyo. (In Japanese)
- Katayama, M. (1998). *Optimality Theory and Japanese Loanword Phonology*. Doctoral dissertation, University of California, Santa Cruz.
- Tajima, K., Erickson, D., and Akahane-Yamada, R. (2000). Training Japanese to perceive syllables in spoken English. In *Proceedings of the 2000 Spring Meeting of the Acoustical Society of Japan*, pages 347–348, Funabashi City, Japan.
- Tajima, K., Port, R. F., and Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign-accented English. *Journal of Phonetics*, 25:1–24.