Mechanisms of vowel epenthesis in consonant clusters: an EMA study

S. Funatsu[a] and M. Fujimoto[b]

[a]Prefectural University of Hiroshima, 1-1-71 Ujinahigashi Minami-ku, 734-8558 Hiroshima, Japan
[b]National Institute for Japanese Language and Linguistics, 10-2 Midori-machi, 190-8561 Tachikawa, Japan
funatsu@pu-hiroshima.ac.jp

The mechanisms of vowel epenthesis in consonant clusters were investigated using an electromagnetic articulograph (EMA). The target languages were Japanese and German. Japanese does not allow consonant clusters, while German does. Two Japanese speakers and two German speakers participated in this experiment. For Japanese speakers, normalized tongue tip displacements from the first consonant to the second consonant in clusters (/bn/, /pn/) were significantly larger than those of German speakers ($p<0.001$). Also, the normalized moving times of tongue tip for Japanese speakers were significantly longer than those for German speakers ($p<0.001$). These results suggested that the coarticulation between first and second consonant for Japanese speakers would be weaker than that for German speakers. Moreover, from the measurement of tongue back movement, the timing between articulatory movement and vocal fold vibration would affect the vowel epenthesis.

1 Introduction

It is well known that second language learners have difficulties in pronunciation of some phoneme sequences which do not exist in their native languages. In these cases, epenthesis, segment change or deletion may occur. For instance, English speakers pronounce “zd” and “zb” in Czech as /zəd/ and /zəb/, respectively [1]. Davidson pointed out that vowel epenthesis for native English speakers would arise from mis-timing in C1C2 sequence or the insertion of a lexical vowel [2]. It is well known that Japanese speakers often insert vowels in consonant clusters (e. g. English word “dry” may be pronounced as /dorai/), possibly because Japanese does not allow the consonant sequences due to its phonotactics. Consonant clusters and closed syllables are not allowed in Japanese. Therefore, in loan words vowel epenthesis frequently occurs, and vowel /o/ is inserted after /t/ and /d/, vowel /i/ is inserted after /tʃ/ and /dʒ/, and vowel /u/ is inserted after other consonants [3,4]. Vowel epenthesis in consonant clusters might arise out of articulatory timing (see Davidson [1.2]) or perception. Dupoux et al. reported that native Japanese speakers perceived the “illusory vowel /u/” in consonant clusters [5,6]. From this result, they concluded that vowel epenthesis in consonant clusters in Japanese originated from the perception of an “illusory vowel”. Consequently, vowel epenthesis in Japanese speakers is caused by perception, whereas for English speakers it is caused by production. Is an origin of vowel epenthesis language specific? The present study aimed to elucidate the mechanisms of vowel epenthesis in consonant clusters using electromagnetic articulograph (EMA).

2 Methods

2.1 Speakers and speech samples

Two Japanese male speakers (Jp1, Jp2), one German female speaker (Ge1), and one German male speaker (Ge2) participated in this experiment. Speech samples were 28 nonsense words in both Japanese and German. These words (X) were embedded in a German sentence “Sage X.” (“Say X.”). Japanese speakers pronounced these sentences 8 times respectively and German speakers pronounced these sentences 10 times respectively.

Speech samples:
bjaht, blaht, bnaht, bunaht, bvaht, gijaht, gjaht, glaht, gnaht, gujaht, gulaht, gunaht, guvaht, gvaht, gzaht, kjaht, klaht, knaht, ksaht, kvaht, pjaht, plaht, pnaht, pvaht, skaht, tskaht, zgaht, zugaht

2.2 Recordings

Articulatory movement and acoustic data were recorded simultaneously using an EMA (AG-500, Carstens). In this articulatory measurement sensor coils were placed on the tongue tip (TT), tongue middle (TM), tongue back (TB), incisors of the lower jaw (LJ), upper and lower lip (UL, LL), as well as on the nasion and left and right retroauricle for head movement correction. The movements of each sensor were recorded separately by X and Y axis. The X axis is the anterior-posterior axis, and the Y axis is the superior-inferior axis. Both axes are on the mid-sagittal plane. In the present measurement, articulatory movement was analyzed in following 6 words, bnaht, pnaht, gnaht, knaht, bunaht, and gunaht.

2.3 Data analyses

Tongue tip movement was measured as follows; tongue tip displacement (D mm) and moving time (T ms) were measured at the burst point of first consonants (/b/, /p/, /g/, /k/) to the complete articulation point of second consonants (maximum of tongue height in /n/). Taking into account for speech rate, T was normalized by word length (L ms) which was defined as the time from the burst of the first consonant to the burst of the last consonant /t/, i. e. normalized moving time (Fig. 1).

$$T_n = \frac{T}{L} \qquad (1)$$

Considering about the size of oral cavity or articulator of each subject, Dx and Dy were also normalized. Dx was normalized by the difference between maximum value of X (Xmax) and minimum value of X (Xmin) in each utterance.

$$D_{xn} = \frac{D_x}{(X_{max} - X_{min})} \qquad (2)$$

Dyn, same as Dxn,

$$D_{yn} = \frac{D_y}{(Y_{max} - Y_{min})} \qquad (3)$$

$$D_n = \sqrt{(D_{xn}^2 + D_{yn}^2)} \qquad (4)$$

Epenthetic vowels were detected by visual inspection on a spectrogram. Moreover, the trajectories of TT and TB in whole sentences were measured.

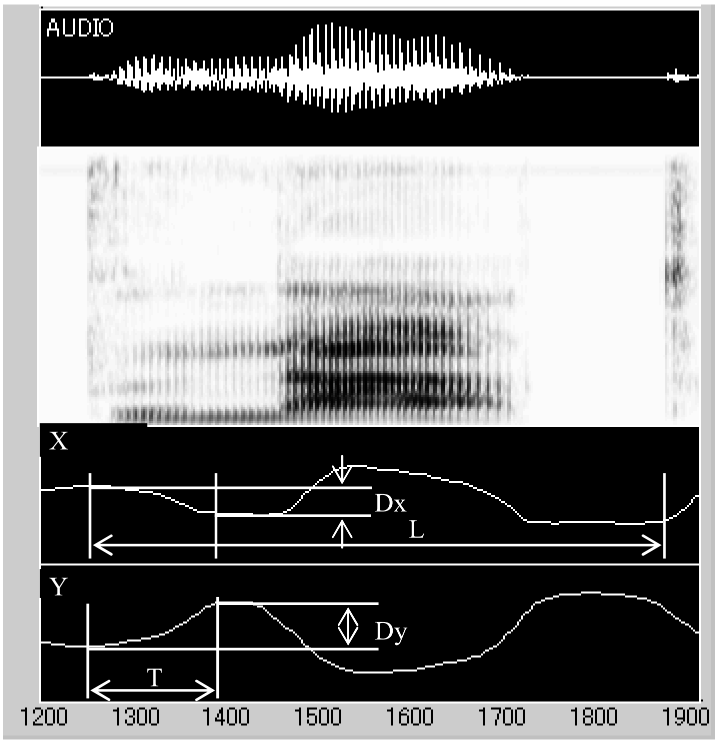


Figure 1: Acoustic wave form and articulatory analysis.

3 Results

3.1 Tongue tip displacement analyses

Table 1 shows mean values of the rate of vowel epenthesis. In the case that first consonants are voiced stops (bnaht and gnaht), vowel epenthesis rate is higher than the case where first consonants are voiceless stops (pnaht and knaht) in both Japanese and German speakers. In all words, vowel epenthesis rates in Japanese speakers were higher than those for German speakers. In the case of bnaht with German speakers, vowel epenthesis rate (0.8) is higher than other words, and this rate is slightly lower than Japanese speakers. This might be because consonant cluster /bn/ does not exist, but other clusters (/pn/, /gn/, and /kn/) exist in German.

Table 2 shows the mean values of tongue tip displacement (D), normalized tongue tip displacement (Dn), tongue tip moving time (T), and normalized tongue tip moving time (Tn) in Japanese and German speakers. For bnaht and pnaht, place of articulation changes from bilabial to dental or alveolar. While, in gnaht and knaht, it changes from velar to dental or alveolar. In bnaht and pnaht (bilabial), Dns in Japanese speakers were significantly larger than those in German speakers ($p<0.001$). In gnaht and knaht (velar), Dns in Japanese speakers were larger than those in German speakers, but not significant difference. In bnaht, pnaht, gnaht, and knaht, Tns in Japanese speakers were significantly longer than those in German speakers (pnaht: $p=0.0018$, others: $p<0.0001$).

Table 3 shows Dn and Tn classified by vowel epenthesis. In this table, “E” means the case that vowel epenthesis occurred and “N” means the case that it did not occurred. Each value is the mean of Japanese and German speakers. In Dn in all words, “E” was larger than “N”, especially in the case of pnaht and gnaht, “E” was significantly larger than “N”. In Tn, “E” was significantly larger than “N”, except in bnaht (in bnaht, “E” was larger than “N” but not significant). This might be because, mentioned above, consonant cluster /bn/ does not exist, but other clusters (/pn/, /gn/, and /kn/) exist in German.

Table 1: Mean values of the rate of vowel epenthesis.

Word	Subject	Epenthesis rate
bnaht	G	0.80
	J	0.88
pnaht	G	0
	J	0.60
gnaht	G	0.65
	J	0.88
knaht	G	0
	J	0.44

G: German, J: Japanese.

Table 2: Mean values of tongue tip displacement and moving time of Japanese and German speakers.

Word	Subject	D (mm)	Dn	T (ms)	Tn
bnaht	G	5.40	0.36	52.4	0.11
	J	**7.96***	**0.67****	**118.9****	**0.28***
pnaht	G	6.37	0.39	97.7	0.18
	J	8.01	**0.66****	**129.1***	**0.23***
gnaht	G	4.95	0.35	58.8	0.11
	J	5.70	0.45	**131.6****	**0.24****
knaht	G	5.32	0.36	88.9	0.16
	J	5.91	0.45	**143.1****	**0.25****

**: $p<0.0001$, *: $p=0.0018$.

Table 3: Dn and Tn classified by vowel epenthesis.

word	Epenthesis	Dn	Tn
bnaht	E	0.54	0.17
	N	0.29	0.12
pnaht	E	0.78	0.25
	N	**0.43****	**0.19***
gnaht	E	0.46	0.19
	N	**0.20***	**0.10***
knaht	E	0.42	0.28
	N	0.40	**0.18***

**: p<0.0001, *: p<0.001

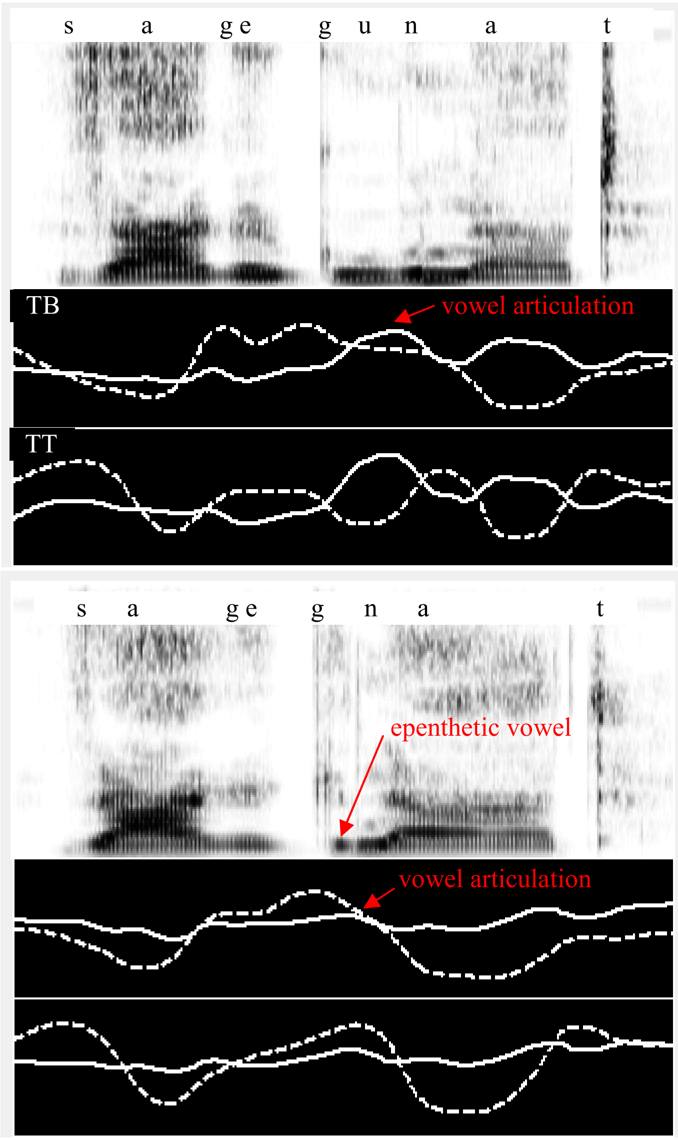
3.2 Articuratory movement analyses

Tongue trajectories of 5 utterances (whole sentences) were shown in Figure 2. Top panel shows gunaht. Second and third panels show examples of vowel /u/ epenthesis between /g/ and /n/. In the second panel, X movement in TB shows vowel /u/ articulation (cf. top panel gunaht), and spectrogram shows epenthetic vowel /u/ between /g/ and /n/. While in the third panel, the spectrogram shows epenthetic vowel between /g/ and /n/, but X movement in TB does not show vowel articulation. Fourth and bottom panels do not show vowel epenthesis. In 4th panel, X movement in TB shows vowel articulation (cf. top panel gunaht), but spectrogram shows no epenthetic vowel. Bottom panel shows no epenthetic vowel in spectrogram and TB movement. These phenomena were observed in bnaht, pnaht, and knaht in Japanese and German speakers in partial.

Hence, we classified consonant cluster pronunciation into following 4 cases.

1. Vowel epenthesis and articulatory movement for vowels both occur.
2. Vowel epenthesis occurs, but articulatory movement for vowels does not.
3. Vowel epenthesis does not occur, but articulatory movement for vowels occurs.
4. Both vowel epenthesis and articulatory movement for vowels do not occur.

From these results, vowel epenthesis in consonant cluster will occur not only by articulatory movement but by vocal fold vibration.



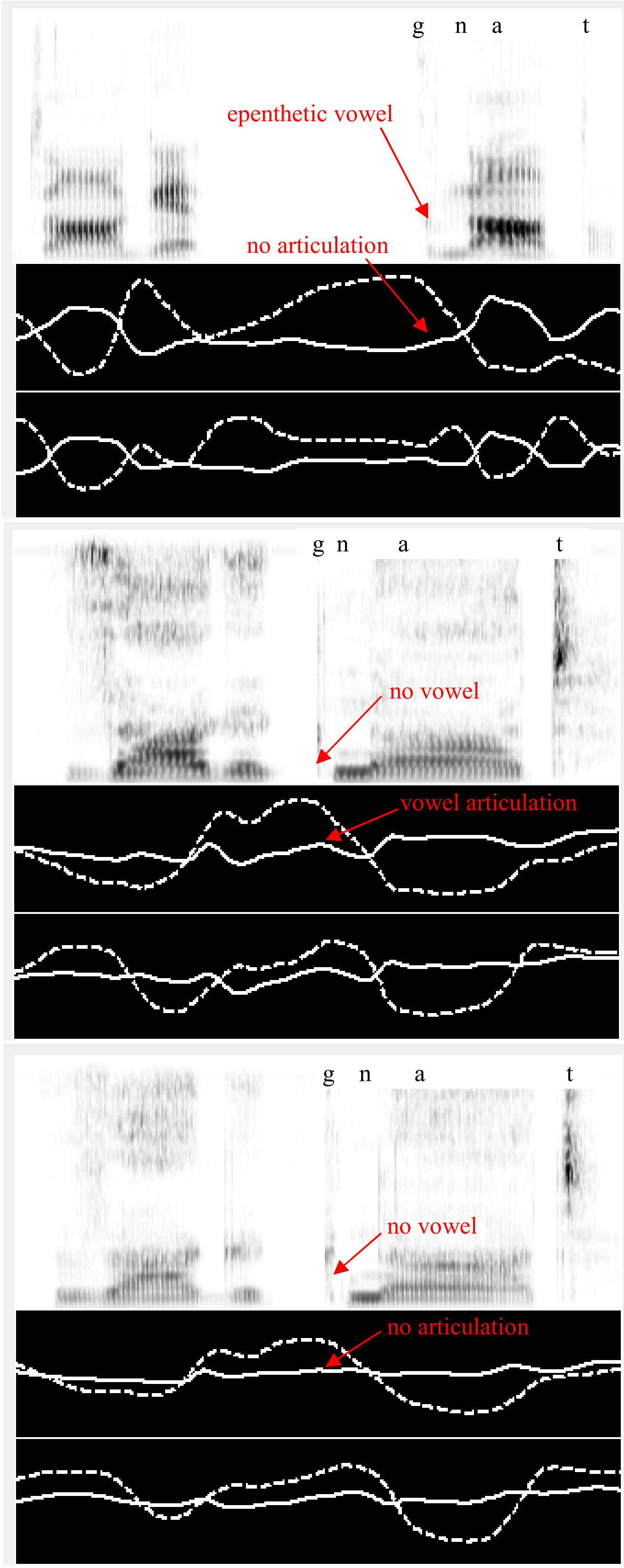


Figure 2: Tongue movement of one utterance of gunaht and 4 utterances of gnaht. Top panel is uttered by Ge1, second is Ge1, third is Jp2, 4th is Ge 1, and bottom is Ge1. In each panel, first row shows spectrogram, second row TB movement, and third row TT movement. Solid line means X movement, dashed line means Y movement.

4 Discussion

Vowel epenthesis rates in Japanese speakers were higher than those in German speakers for all words. In pnaht (bilabial), the normalized displacements of tongue tip in Japanese speakers were significantly larger than those of German speakers. Moreover, normalized tongue tip moving times in Japanese speakers were significantly longer than those in German speakers. This is because German has consonant clusters and German speakers will prepare for the next consonant articulation when they articulate the first consonant in a consonant cluster. Japanese, on the other hand, does not have consonant clusters in general. Accordingly, Japanese speakers will not prepare for the next consonant when they articulate the first consonant. In other words, the coarticulation between consonants would be weaker for Japanese speakers. However, bnaht is excluded from this case by the reason mentioned earlier.

While in gnaht and knaht (velar), the normalized displacements of tongue tip in Japanese speakers were larger, but not significant than those for German speakers. This might be because the difference of the place of articulation in first consonants between velars and bilabials. When bilabials are articulated the tongue tip is free from articulation, however when velars are articulated the tongue back moves to posterior-superior and the tongue tip also moves to posterior-superior. Namely, the tongue tip is not so free as in the case of bilabials. When bilabials are articulated, Japanese speakers can move their tongue different from German speakers, whereas when velars are articulated, Japanese and German speakers have to move their tongue similar to each other. Therefore, the normalized displacements of tongue tip would not significantly differ between Japanese and German speakers.

Vowel epenthesis in consonant cluster occurs by articulatory movement and vocal fold vibration. The timing between articulation and vocal fold vibration would be important for vowel epenthesis. In the case that first consonant is voiceless (/p/, /k/), if vocal fold vibration occurs before the second consonant (/n/), vowel is inserted. If the onset of vocal fold vibration, however, is late for next consonant articulation, vowel epenthesis could not occur. Whereas, in the case where first consonant is voiced (/b/, /g/), if the interruption of vocal fold vibration does not occur between the first and second consonant, a vowel is inserted. Conversely speaking, when vocal fold vibration does not occur or interrupts before second consonant, vowels are not inserted in a consonant cluster. These interpretations would be applicable for Japanese and German speakers.

5 Conclusions

Vowel epenthesis rates of Japanese speakers were higher than those of German speakers. One of the mechanisms of vowel epenthesis in consonant clusters for Japanese speakers is that tongue displacement from first to second consonant in Japanese was larger than that in German. In other words, this result suggests that coarticulation between first and second consonant would be weak for Japanese speakers. Moreover, the interpretation that the gap of timing between articulatory movement and vocal fold vibration would generate vowel epenthesis would be applicable to Japanese and German.

Acknowledgments

This study was supported by a Grant-in-Aid for Scientific Research (No. 20520394, 23520539) from the Japan Society for the Promotion of Science.

References

[1] Davidson, L., "Phonology, phonetics, or frequency: Influences on the production of non-native sequences", *Journal of Phonetics*, 34(1), 104-137, (2006).

[2] Davidson, L., "Addressing phonological questions with ultrasound", *Clinical Linguistics & Phonetics*, 19(6/7), 619-633, (2005).

[3] Quakenbush, H., and Ohso, M., "Formation of loan words and its education", (in Japanese), The National Institute for Japanese Language, (1990).

[4] Kubozono, H., "Prosodic structure of loanwords in Japanese: Syllable structure, accent and morphology", *Journal of Phonetic Society of Japan*, 6, 79-97, (2002).

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

[6] Dupoux, E., Pallier, C., Kakehi, K., and Mehler, J., "New evidence for prelexical phonological processing in word recognition", *Language & Cognitive Processes*, 16, 491-505, (2001).