THE EFFECTS OF VOICED STOPS ON ADJACENT VOWEL DURATION IN JAPANESE

Shoji Ishibashi[1], Shonosuke Koya[1&2]

[1]Fukuoka University Graduate School, [2] JSPS Research Fellowship for Young Scientists
ld220501@cis.fukuoka-u.ac.jp, ld220502@cis.fukuoka-u.ac.jp

ABSTRACT

Vowel duration is longer before voiced than voiceless consonants (voicing effect) and it is known as universal phenomenon across languages. However, it is still controversial whether the effect is consistently seen in languages which have vowel length contrast (long vs. short). Furthermore, while the effects of consonants on preceding vowels have often been examined, there have been few studies of those on following vowels. To address these issues, the current study conducted a production experiment in which native speakers of Japanese uttered disyllabic nonce words (CV_1CV_2 or $CV_1{:}CV_2$). Statistical analysis suggests three main results as follows: (i) a voicing effect is observed regardless of preceding vowel length; (ii) duration of the following vowel is likely to be longer when the preceding consonant is voiced than when it is voiceless; (iii) V_2 duration tends to be shorter when V_1 is phonologically long than when it is phonologically short.

Keywords: voicing effect, vowel length, Japanese, speech production

1. INTRODUCTION

A voicing distinction is one of the most common phonetic contrasts in languages, and various studies have explored voiced contrasts for decades [18, 11]. With respect to the relationship between voicing contrast and vowel duration, it is widely known that vowels followed by voiced obstruents are durationally longer than when followed by voiceless obstruents in various languages (French, Russian, Korean [4] and English [8]). In English, for example, the vowel duration in *tab* is longer than that in *tap*. This phenomenon has been called the 'pre-fortis clipping' [17] or 'voicing effect' [14] (hereafter, voicing effect). Since the effect is confirmed in several languages, it can be a cross-linguistically common tendency.

While the voicing effect is confirmed in several languages, it is also reported that a language-specific phonological system may reduce the effect. In [12], the voicing effect is not observed in Czech and the study implied that the effect may blur the phonological distinction of vowel length as Czech has short and long vowels phonemically (i.e., the voicing effect is not seen or is extremely weak in a language which has vowel length contrasts). A similar tendency is observed in Saudi Arabic, a language with phonological vowel length contrast [6].

Although several studies showed example cases in which the voicing effect may be reduced by a language-specific phonological system, there are also some studies which have different results. For instance, [1] reported that the voicing effect was confirmed in Lebanese Arabic even though it has a phonemic vowel length contrast. Thus, it is still unclear whether the degree of voicing effect is influenced by a language-specific system. To clarify this question, the voicing effect should be explored using more languages with phonological vowel length contrasts.

Japanese can be a suitable language to elucidate the relationship between the voicing effect and vowel length contrast. Japanese has phonemic vowel length distinction (e.g., /i/ 'stomach' vs. /i:/ 'good'), and the voicing effect in Japanese is controversial [15, 20]. [15] analyzed a set of four minimal pairs of words (e.g., [k**a**ka] and [k**a**ga]) uttered by ten native speakers of Japanese and reported that vowels are produced with 20% longer duration before a voiced obstruent than before a voiceless obstruent. On the other hand, [20] examined the voicing effect using the Corpus of Spontaneous Japanese and the results showed that the voicing effect was not observed. Additionally, while [15] examined the voicing effect on only short vowels, [20] examined the voicing effect on short and long vowels, and [20] reported that the voicing effect was not observed regardless of the vowel length. Even though these two studies have methodological differences, it has not yet been fully verified whether voicing effect can be reliably observed in Japanese.

In addition to the voiced obstruent effect on a preceding vowel, the effect on a following vowel is also confirmed in several languages. [13] examined minimal pairs of disyllabic words (e.g., $C_1V_1C_2V_2$; [tat**a**] and [tad**a**]) and reported that the vowel duration after voiced obstruents is longer than after voiceless ones (hereafter, called post-voicing effect) in Japanese. [1] conducted a production experiment with disyllabic nonce words and a similar phenomenon is observed in Lebanese Arabic. Moreover, the data in

[5] implies that Italian also has the post-voicing effect. Considering these reports, the post-voicing effect might be a cross-linguistically common tendency as well as the voicing effect. However, the following points should be noted. Since [1] did not consider the vowel length contrast in their analysis, it is possible that the post-voicing effect is observed only when the vowel in the first syllable is short, but not when long and vice versa. Similarly, [13] did not examine the post-voicing effect when V_1 was phonologically long. Thus, whether the post-voicing effect is observed regardless of the phonological length of the preceding vowel is still a controversial issue and should be examined in more detail.

We also need to refer to the tendency that vowel duration after geminates or a long vowel tends to be shorter than after singletons or short vowel in Japanese [3, 7, 10]. [10] conducted a production experiment using disyllabic words ($C_1V(:)_1C_2V_2$) and reported that vowel duration in the syllable following the long vowel tends to be shorter than following the short one. As [10] used only voiceless stops for word-internal consonants in their stimuli, however, it should also be explored with voiced stops.

Based on the results of the previous studies, the present study addressed the following three research questions by conducting a production experiment in which speakers uttered disyllabic nonce words ($C_1V(:)_1C_2V_2$):

(1) Is the voicing effect observed in not only short vowels but also long vowels?
(2) Is the post-voicing effect observed regardless of the phonological length of the preceding vowel?
(3) Is the V_2 duration shortened when V_1 is a long vowel regardless of the voicing of the word-internal consonant?

2. METHOD

2.1. Participants

Our ten participants, five male and five female, were all native speakers of Japanese with a mean age of 22.2 (range: 18-28). The experiment was conducted at the Phonetic Laboratory of Fukuoka University. All participants were first introduced to the purpose of the experiment and the procedure before signing a consent form. Demographic data, such as age, sex, and birthplace, were described before the experiment started. Participants received monetary compensation.

2.2. Procedure

Materials used in the present study were two-syllable words of the form ‘$pV_1(:)CV_2$’ (V_1 = a, e, o; C = p, t, k, b, d, g; V_2 = a, e, o); the total number of words was 108 (3 vowel qualities of V_1 (/a/, /e/, /o/) × 2 vowel lengths of V_1 (short, long) × 3 types of C_2 (labials, alveolar, velar) × 3 vowel qualities of V_2 (/a/, /e/, /o/). These words were written in *katakana* and were embedded in a carrier phrase /karewa ___ to itta/ (“he said ___”). The participants read the 108 words within a carrier phrase five times, and a total of 5400 tokens (108 words × 5 repetitions × 10 participants) were recorded using a digital audio recorder (Sony PCM-D100) in a soundproof room. 52 tokens from the entire data sets were discarded due to errors of utterances; thus, a total of 5348 tokes were used in the present analysis.

2.3. Analysis

Five acoustic measures were made for each token: the VOT of C_1 (**p**V(V)CV), V_1 duration (p**V(V)**CV), the C_2 closure duration (pV(V)**C**V), the VOT of C_2 (pV(V)**C**V), and the V_2 duration (pV(V)C**V**). The durations were segmented by the visual inspection of the waveform and spectrogram with reference to periodicity and formant energy bands using Praat (ver. 6.1.51) speech analysis software [2], following the segmentation criteria used in [9] and [10]. The criteria are as follows: the VOT of C_1 and C_2 were measured from the onset of the release burst of C_1 and C_2 to the onset of the voicing of V_1 and V_2, respectively. The durations of V_1 and V_2 were determined as the points where the formants abruptly appeared, and the offset of V_1 and V_2 was determined to be the points where the formants abruptly ceased. The C_2 closure duration was measured between the offset of V_1 and the onset of the release burst of C_2.

3. RESULT

Word example	C_1 VOT	V_1	C_2 closure	C_2 VOT	V_2
papo	30.4	58.4	70.6	19.1	69.0
pabo	36.9	76.3	41.2	18.4	83.5
paapo	28.8	173.7	79.7	22.8	58.4
paabo	35.3	201.9	47.1	16.1	72.0

Table 1: Mean duration (ms) for C_1 VOT, V_1, C_2 closure, C_2 VOT, and V_2

Table 1 summarizes the mean duration of C_1 VOT, V_1, C_2 closure, C_2 VOT, and V_2. The mean duration of V_1 and V_2 is illustrated in Figure 1 and 2 above. These show that the voicing effect and post-voicing effect is likely to be observed.

In order to determine (i) whether or not V_1 duration systematically varies depending on the voicing of C_2, (ii) whether or not V_2 duration varies depending on the voicing of C_2 and (iii) whether or not V_2 duration varies depending on the V_1 length, a mixed ANOVA was conducted separately for V_1 and V_2 duration. The fixed factors were V_1 length (phonologically short vowel, long vowel), voicing of

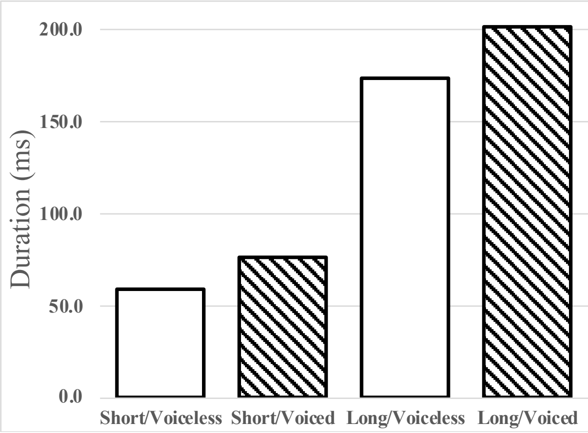

Figure 1: Mean duration of V_1 for each length and voicing of C_2

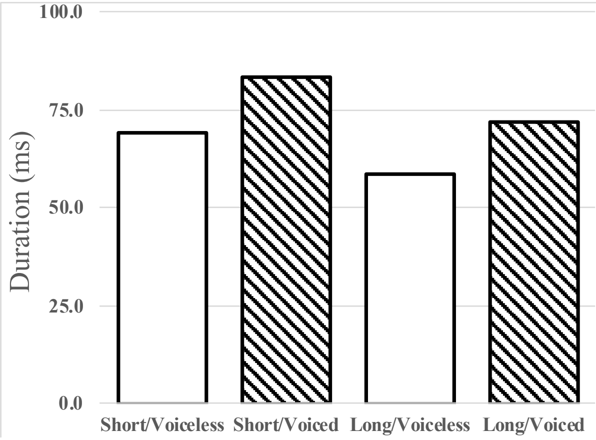

Figure 2: Mean duration of V_2 for each length and voicing of C_2

C_2 (voiceless, voiced), V_1 quality (/a/, /e/, /o/), C_2 types (labial (/p/, /b/), alveolar (/t/, /d/), velar (/k/, /g/)), and V_2 quality (/a/, /e/, /o/); random factors were participants and repetitions. Results for interactions other than V_1 length × C_2 voicing are omitted due to space limitations.

3.1. V_1 duration

A mixed ANOVA was conducted with V_1 duration as the dependent variable and V_1 length, V_1 quality, voicing of C_2, C_2 types, and V_2 quality as fixed factors; random factors were participants and repetitions. All main effects except for V_2 quality were significant, and interaction between C_2 voicing and V_1 length were significant (F (1, 5301) = 23.581, $p < 0.01$). Since the interaction between V_1 length and C_2 voicing were significant, statistical analyses were conducted for each vowel length (short and long vowel).

The V_1 (short vowel) duration is statistically longer when C_2 is voiced than when C_2 is voiceless (C_2 voicing: F (1, 2656) = 2444.457, $p < 0.01$).

The V_1 (long vowel) duration is statistically longer when C_2 is voiced than when C_2 is voiceless (C_2 voicing: F (1, 2650) = 1466.972, $p < 0.01$).

3.2. V_2 duration

A mixed ANOVA was conducted with V_2 duration as the dependent variable and V_1 length, V_1 quality, voicing of C_2, C_2 types, and V_2 quality as fixed factors; random factors were participants and repetitions. All main effects were significant, and interaction between C_2 voicing and V_1 length were significant (F (1, 5301) = 11.605, $p < 0.01$). Since the interaction between V_1 length and C_2 voicing was significant, statistical analyses were conducted for each vowel length (short and long vowel).

3.2.1. Vowel duration after voiceless/voiced obstruents

A mixed ANOVA was conducted with V_2 duration (with short V_1) as the dependent variable and voicing of C_2, V_1 quality, C_2 types, and V_2 quality as the fixed factors; random factors were participants and repetitions. The V_2 duration (with short V_1) is statistically longer when C_2 is voiced than when C_2 is voiceless (C_2 voicing: F (1, 2656) = 1688.961, $p <$ 0.01).

A mixed ANOVA was conducted with V_2 duration (with long V_1) as the dependent variable and voicing of C_2, V_1 quality, C_2 types, and V_2 quality as the fixed factors; random factors were participants and repetitions. The V_2 duration (with long V_1) is statistically longer when C_2 is voiced than when C_2 is voiceless (C_2 voicing: F (1, 2650) = 1381.640, $p <$ 0.01).

3.2.2. Vowel in the syllable after a phonologically short/long vowel

A mixed ANOVA was conducted with V_2 duration (with voiceless C_2) as the dependent variable and V_1 length, V_1 quality, C_2 types, and V_2 quality as the fixed factors; random factors were participants and repetitions. The V_2 duration is statistically shorter when V_1 length is long than when V_1 length is short (V_1 length: F (1, 2656) = 789.771, $p < 0.01$).

A mixed ANOVA was conducted with V_2 duration (with voiced C_2) as the dependent variable and V_1 length, V_1 quality, C_2 types, and V_2 quality as the fixed factors; random factors were participants and repetitions. The V_2 duration is statistically shorter when V_1 length is long than when V_1 length is short (V_1 length: F (1, 2650) = 1087.816, $p < 0.01$).

3.3. Summary of the results

The results of the present study can be summarized as follows:

(i) Vowel duration is longer when the following consonant is voiced than when it is voiceless regardless of the vowel length (voicing effect).

(ii) Vowel duration is longer when the preceding consonant is voiced than when it is voiceless regardless of the preceding syllable's vowel length (post-voicing effect).

(iii) In disyllabic words ($C_1V_1C_2V_2$), vowel duration is shorter when the phonological length of V_1 is long than when V_1 is short regardless of the voicing of C_2.

4. DISCUSSION AND CONCLUSION

In order to examine the effects of voiced stops on adjacent vowel duration, the present study conducted a production experiment in which ten native speakers of Japanese uttered disyllabic words.

The first question of this study was whether the voicing effect is seen in Japanese. The results showed that the voicing effect is observed in Japanese regardless of the vowel length. In the discussion of the voicing effect in Japanese, the reports by previous studies were contradicted. While [15] indicated that the voicing effect is seen in Japanese, [20] said that it is not. In light of the above, our result seems to correspond to the result of [15]. This means our result contradicts the implication of [12], i.e., [12] implied that the voicing effect is not seen or is extremely weak in a language which has vowel length contrast, but our result showed the voicing effect in Japanese which does have this contrast.

Although the results of the present study and [20] are contradicted, it should be noted that there is a methodological difference between [20], a study using a corpus, and the present study, which is an experimental study. Given the difference in the characteristics of corpus and experimental data [19, 16], methodological differences may cause such conflicting results between the [20] and the present study, and, thus, it is not necessarily true that the result of [20] is not valid. Further study is required.

Our second interest was the voicing effect on following vowel (post-voicing effect). The results showed that the post-voicing effect was observed in Japanese regardless of the preceding vowel length. In light of the post-voicing effect, [13] reported that the vowel duration after voiced obstruents is longer than after voiceless ones. However, the research was limited to only the case when the preceding vowel was phonologically short. The current study confirms that the post-voicing effect is seen regardless of the phonological length of the preceding vowel. To clarify to what extent the post-voicing effect is a phonetic tendency, other languages should be explored. Furthermore, although we only used the vowel length contrast on V_1, a relationship between V_2 length contrast and the post-voicing effect is also interesting. These questions should be addressed in future studies.

The third of our questions is the relationship between a long vowel and the duration of following vowel. In conducting a production experiment in which participants uttered disyllabic words ($C_1V_1(:)C_2V_2$), [10] reported that vowel duration in the syllable following a long vowel is likely to be shorter than that following a short vowel. However, [10]'s research only used voiceless consonants as C_2. The current study investigated if the tendency is seen when C_2 is voiced. The results show that V_2 duration tends to be shorter when V_1 is phonologically long than when V_1 is phonologically short regardless of the voicing of C_2. The results reinforce the relationship between a long vowel and the following vowel's duration, which is suggested by [10]. In order to elucidate the generality of this phenomenon, experimental research on other languages which have phonologically long vowels is required.

ACKNOWLEDGMENTS

Many thanks to Professor Howe Stephen Mark, Professor Eiji Yamada, Associate Professor Hajime Takeyasu and two anonymous referees for their helpful comments on earlier versions. This work was partly supported by JSPS KAKENHI Grant Number JP 22J22198. The survey and experiments for this paper were also supported by the Phonetics Laboratory of Fukuoka University. All remaining errors are ours.

REFERENCES

[1] Al-Tamimi, J., Khattab, G. 2018. Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate stops. *Journal of Phonetics*. 71, 306–325.

[2] Boersma, P., Weenink, D. 2021. Praat: doing phonetics by computer (Version 6.1.54). Amsterdam. Retrieved from http://www.fon.hum.uva.nl/praat/.

[3] Campbell, N. 1999. A study of Japanese speech timing from the syllable perspective. *Journal of the Phonetic Society of Japan*. 3, 29–39.

[4] Chen, M. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica*. 22, 129–159.

[5] Esposito, A., Benedetto, M-G. D. 1999. Acoustical and perceptual study of gemination in Italian stops. *J. Acoust. Soc. Am*. 106, 2051–2062.

[6] Flege, J. E. 1979. Phonetic Interference in Second Language Acquisition. PhD. dissertation, Indiana University, Bloomington.
[7] Han, M. S. 1994. Acoustic manifestations of mora timing in Japanese. *J. Acoust. Soc. Am*. 96, 73–82.
[8] House, A. S. 1961. On vowel duration in English. *J. Acoust. Soc. Am*. 33, 1174–1178.
[9] Hussain, Q., Shinohara, S. 2019. Partial devoicing of voiced geminate stops in Tokyo Japanese. *J. Acoust. Soc. Am*. 145, 149–163.
[10] Ishibashi, S. 2022. Effects of adjacent segments on the duration of the vowel and consonant within a disyllabic word in Japanese. *Phonological Studies* 25, 11–20.
[11] Keating, P., Linker. W., Huffman. M. 1983. Patterns in allophone distribution for voiced and voiceless stops. *Journal of phonetics* 11, 277–290.
[12] Keating, P. 1985. Universal phonetics and the organization of grammars. In: Victoria F. A. (ed), *Phonetic linguistics: Essays in honor of Peter Ladefoged*. Academic Press, 115–132.
[13] Kubozono, H. 2005. Nihongo ni okeru jikanseigyo no syoso (in Japanese). Inritsu ni tyakumoku sita onseigengojohosyori no kodoka (A report of MEXT Grants-in-Aid for Scientific Research). Retrieved from http://www.govo.t.u-Tokyo.ac.jp/tokutei_pub/.
[14] Mitleb, F. M. 1984. Voicing effect on vowel duration is not an absolute universal. *Journal of Phonetics* 12, 23–27.
[15] Port, R. F., Dalby, J., O'Dell, M. 1987. Evidence for mora timing in Japanese. *J. Acoust. Soc. Am*. 81, 1574–1585.
[16] Wagner, P., Trouvain, J., Zimmerer, F. 2015. In defense of stylistic diversity in speech research. *Journal of Phonetics*. 48, 1–12.
[17] Wells, J. C. 1990. Syllabification and allophony. In: Ramsaran S. (ed), *Studies in the pronunciation of English: A commemorative volume in honour of A. C. Gimson*. Routledge, 76–86.
[18] Williams, L. 1977. The voicing contrast in Spanish. Journal of phonetics. 5, 169–184.
[19] Xu, Y. 2010. In defense of lab speech. *Journal of phonetics*. 38, 329–336.
[20] Yoneyama, K., Kitahara, M. 2014. Voicing effect on vowel duration: corpus analyses of Japanese infants and adults, and production data of English learners. *Journal of the Phonetic Society of Japan*. 18, 30–39.