THE EFFECTS OF VOICED STOPS ON ADJACENT VOWEL DURATION IN JAPANESE

Shoji Ishibashi¹, Shonosuke Koya¹&²

¹Fukuoka University Graduate School, ²JSPS 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 (CV₂CV₂ or CV₁CV₂). 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) V₂ duration tends to be shorter when V₁ 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. /iː/ ‘good’), and the voicing effect in Japanese is controversial [15, 20]. [15] analyzed a set of four minimal pairs of words (e.g., [kaka] and [kasa]) 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., C₁V₁C₂V₂; [tata] and [tada]) 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
2. Speech Acoustics

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(V)CV’ (V = a, e, o; C = p, t, k, b, d, g; V = a, e, o); the total number of words was 108 (3 vowel qualities of V1 (/a/, /e/, /o/) × 2 vowel lengths of V1 (short, long) × 3 types of C1 (labials, alveolar, velar) × 3 vowel qualities of V2 (/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 tokens were used in the present analysis.

2.3. Analysis

Five acoustic measures were made for each token: the VOT of C1 (pV(V)CV), V1 duration (pV(V)CV), the C2 closure duration (pV(V)CV), the VOT of C2 (pV(V)CV), and the V2 duration (pV(V)CV). 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 C1 and C2 were measured from the onset of the release burst of C1 and C2 to the onset of the voicing of V1 and V2, respectively. The durations of V1 and V2 were determined as the points where the formants abruptly appeared, and the offset of V1 and V2 was determined to be the points where the formants abruptly ceased. The C2 closure duration was measured between the offset of V1 and the onset of the release burst of C2.

3. Result

<table>
<thead>
<tr>
<th>Word example</th>
<th>C1 VOT</th>
<th>V1</th>
<th>C2 closure</th>
<th>C2 VOT</th>
<th>V2</th>
</tr>
</thead>
<tbody>
<tr>
<td>pabo</td>
<td>36.9</td>
<td>76.3</td>
<td>41.2</td>
<td>18.4</td>
<td>83.5</td>
</tr>
<tr>
<td>pabo</td>
<td>28.8</td>
<td>173.7</td>
<td>79.7</td>
<td>22.8</td>
<td>58.4</td>
</tr>
<tr>
<td>pabo</td>
<td>35.3</td>
<td>201.9</td>
<td>47.1</td>
<td>16.1</td>
<td>72.0</td>
</tr>
</tbody>
</table>

Table 1: Mean duration (ms) for C1 VOT, V1, C2 closure, C2 VOT, and V2.
2. Speech Acoustics

C2 (voiceless, voiced), V1 quality (/a/, /e/, /o/), C2 types (labial (/p/, /b/), alveolar (/t/, /d/), velar (/k/, /g/)), and V2 quality (/a/, /e/, /o/); random factors were participants and repetitions. Results for interactions other than V1 length × C2 voicing are omitted due to space limitations.

3.1. V1 duration

A mixed ANOVA was conducted with V1 duration as the dependent variable and V1 length, V1 quality, voicing of C2, C2 types, and V2 quality as fixed factors; random factors were participants and repetitions. All main effects except for V2 quality were significant, and interaction between C2 voicing and V1 length were significant (F(1, 5301) = 23.581, p < 0.01). Since the interaction between V1 length and C2 voicing was significant, statistical analyses were conducted for each vowel length (short and long vowel).

The V1 (long vowel) duration is statistically longer when C2 is voiced than when C2 is voiceless (C2 voicing: F(1, 2650) = 1466.972, p < 0.01).

3.2. V2 duration

A mixed ANOVA was conducted with V2 duration as the dependent variable and V1 length, V1 quality, voicing of C2, C2 types, and V2 quality as fixed factors; random factors were participants and repetitions. All main effects were significant, and interaction between C2 voicing and V1 length were significant (F(1, 5301) = 11.605, p < 0.01). Since the interaction between V1 length and C2 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 V2 duration (with short V1) as the dependent variable and voicing of C2, V1 quality, C2 types, and V2 quality as the fixed factors; random factors were participants and repetitions. The V2 duration (with short V1) is statistically longer when C2 is voiced than when C2 is voiceless (C2 voicing: F(1, 2650) = 1688.961, p < 0.01).

A mixed ANOVA was conducted with V2 duration (with long V1) as the dependent variable and voicing of C2, V1 quality, C2 types, and V2 quality as the fixed factors; random factors were participants and repetitions. The V2 duration (with long V1) is statistically longer when C2 is voiced than when C2 is voiceless (C2 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 V2 duration (with voiceless C2) as the dependent variable and V1 length, V1 quality, C2 types, and V2 quality as the fixed factors; random factors were participants and repetitions. The V2 duration is statistically shorter when V1 length is long than when V1 length is short (V1 length: F(1, 2656) = 789.771, p < 0.01).

A mixed ANOVA was conducted with V2 duration (with voiced C2) as the dependent variable and V1 length, V1 quality, C2 types, and V2 quality as the fixed factors; random factors were participants and repetitions. The V2 duration is statistically shorter when V1 length is long than when V1 length is short (V1 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


