Non-native productions of Japanese single stops that are too long for one mora unit

Yasuyo Minagawa-Kawai and Shigeru Kiritani

Department of Cognitive Sciences, Graduate School of Medicine, University of Tokyo

myasuyo@m.u-tokyo.ac.jp

Abstract

The difficulty for non-native speakers in producing Japanese geminate stops with a long enough closure has been pointed out and extensively studied. However, the reverse problem exists for particular language speakers such as Chinese. For these speakers, production of Japanese intervocalic single stops sound like geminate stops. This study aimed to show acoustic evidence of this problem in the production of Japanese voiceless stops by learners of Japanese, and to compare speakers from different language groups, English, Chinese and French speakers. The results of the experiment indicated that since some of the stop productions by Americans and Chinese were too long for one mora unit, the Japanese heard a geminate (2 mora) instead of a single. Because French subjects had a small difference in the stop durations from the Japanese value, only a few of their productions of single stops were perceived as geminates. The language differences in timing pattern for Japanese stops were interpreted to reflect differences in timing implementation of native languages that is stress-timed and syllable-timed, at least concerning the difference between English and French speakers.

1. Introduction

Rhythmic patterning, such as stress and syllable timing, varies according to language. In general, it is difficult to acquire a second language which is rhythmically distinct from the native language. Accordingly, for many learners of Japanese, mastering mora timing in Japanese often turns out to be a problem and many studies have examined this topic both acoustically and perceptually. What these production studies mainly captured was the difficulty in producing Japanese special mora (geminate stop, long vowel, and moraic nasal) with a long enough duration to distinguish them from their short counterparts. However, particularly for geminate stop the reverse problem exists for particular language speakers such as English and Chinese speakers. For these speakers, production of Japanese single stops is not well controlled to adjust the mora timing and they sound like geminate stops, as in *saQkana* (CVCCVCV) for *sakana* (fish) (CVCVCV). On the other hand, speakers of some other languages such as French, Spanish and Finnish appear to be able to realize smooth mora timing in CV sequences containing stops. This problem seems to be related to timing control in the native language. This study aims to show acoustic evidence of this problem in the production of Japanese voiceless stops by learners of Japanese, and to compare different language groups, English, Chinese and French speakers.

2. Methods

Production test Three mora words containing a voiceless stop consonant in the second mora were used for the materials. They were natural words and non-words of which the first vowel (V_1) is /a/ and the pitch accent pattern is LHH. The syllable structures were of 2 types, (C)VCVCV and CVCVV. The words used for the materials were apapa, sapota (nonsense words), atama (head), Nakano (a place name), satoo (the left party), katee (process), yakee (night view). 20 subjects, 5 Americans, 5 Chinese (Beijing), 5 French, and 5 Japanese participated in the experiment. All the subjects, apart from the Japanese, were beginning learners of Japanese and their experience of learning Japanese was almost the same. The subjects were asked to read each word on the list three times at normal speed. The speech material was recorded on a cassette recorder using a microphone in a quiet room. The recorded material was digitized onto a computer with a sampling rate of 22 kHz. This was then analyzed via wideband spectrograms (8 ms/125 Hz) and waveforms to measure the segmental duration.

Gemination perception test To judge whether the utterances by non-native speakers were heard as geminated stop or a single stop by native Japanese speakers, a gemination perception test was carried out by 5 native Japanese speakers. 420 randomized stimuli were presented to the subjects in a soundproof room. The subjects were asked to judge whether the intervocalic single stop in each stimulus was geminated or not at 3 levels.

3. Results

The results of the acoustic measurement are shown in Figure 1. In the figure, the average durations of each segment in 4 words (atama, sapota, satoo, katee) are indicated according to the language group. First of all, closure duration (CD) and duration of closure plus voice onset time (CVD) produced by the American and Chinese speakers demonstrated longer values in every word than for Japanese speakers. On the other hand, French productions of CD and CVD were only a little longer than the Japanese productions. For segments other than CD and CVD, the V_1 duration of Americans was found to be always shorter than Japanese V_1 ($p<.0001$.) in spite of the longer duration of the total word length for Americans. These observed tendencies for the 4 words in the Figure 1 are common for all the words.

The value to normalize the difference of total word length for CD and CVD (N value) was calculated by dividing CD and CVD by the total word length and multiplying by 100. Figure 2 indicates

the N value of CD and CVD. As was observed in Figure 1, the American and Chinese data show higher values for both CD and CVD than the Japanese values. These differences are even more prominent in Figure 2 than Figure 1. As for French speakers, in most of the words the N values of CD as well as CVD are longer than those for Japanese speakers, but the differences are small. The number of words that have statistical differences ($p<.05$) from native productions in N value of CD and CVD are 6 (CVD) and 5 (CD) for Americans, 7 and 5 for Chinese and 4 and 2 for French. From Figure 2 again, one notices a tendency for both CD and CVD to be longer before a long vowel such as in "yakee". Although this characteristic is common to every language, a language difference in the extent of CD and CVD lengthening is observed. Figure 3 shows the results of the geminate perception test. The vertical axis indicates the percentage rate at which native Japanese speakers perceived a gemination (geminate perception rate). The geminate perception rates vary according to the language types. Corresponding to the measurement results, American and Chinese rates are outstanding for their higher rate. French has a relatively small rate. Apart from the language differences, differences in word types are also observed. The geminate perception rates tend to be higher in the words that include a long vowel (CVCVV) than in the words consisted of a single vowel (CVCVCV).

Since duration of voiceless stops plays a crucial role in the perception of Japanese geminate stops (Fujisaki et al., 1973), the relation between N values of stop duration (CVD) and the geminate perception rate of the utterances was investigated. In Figure 4, stop duration in every utterance by Americans and Chinese, and the geminate perception rate are plotted. In the figure where all the tokens of 7 words are included, both the American and Chinese data exhibit a strong correlation between N value of stop duration and the geminate perception rate. The longer the stop duration of each token, the higher geminate perception rate. To take a word variation in the perception of geminate stops, Pearson correlation between N values of stop duration and geminate perception rate for every word in the American and Chinese productions were analyzed. It was found that N values of stop duration were highly correlated ($p<.01$) with the geminate perception rate for every word.

4. Discussion

The phenomenon which motivated the present study is the non-native productions of Japanese single stops that are perceived as geminated stops by native speakers of Japanese. It is known that this problem is seen in particular language speakers such as Chinese. Comparing the speakers of 3 languages, the results of the present study confirmed such general impressions acoustically and revealed that the English and Chinese speakers showed longer stop durations than those of Japanese speakers. About 35% of their single stops in 3 mora words were perceived as geminate stops. However the French speakers exhibited only slightly longer durations for stops than the Japanese, resulting in an 8% geminate perception rate. The correlation between duration of the stops produced by Americans and Chinese and the geminate perception rate was examined and found to be strong for

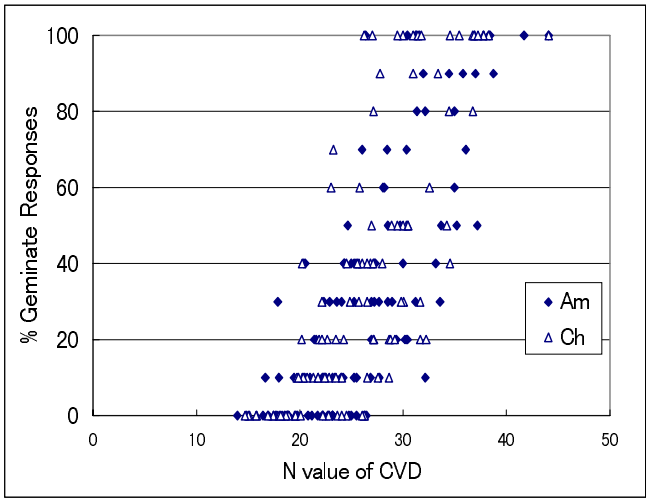

Figure 4. Scatter plot of the geminate perception rate against N value of CVD for each utterance by American and Chinese subjects**.**

every word. This indicated that since some of the stop productions by English and Chinese speakers were too long for one mora unit, Japanese heard a geminate (2 mora) instead of a single (1mora). Timing implementation of Japanese stops by the French speakers was rather similar to that of Japanese speakers when compared to the Americans and Chinese, most of the productions of one mora by French speakers being heard as one mora by the native Japanese speakers.

Different timing pattern based on the rhythmic structure of the speakers' native language seems to have emerged in Japanese production and caused the durational difference among language groups, at least between English and French speakers. There is a difference in the rhythmic structures of English and French, namely English is a stress-timed language while French is a syllable-timed. In stress-timed English, syllables are lengthened or shortened depending on the stress position, namely syllables are prolonged when stress is assigned and shortened when they precede or are followed by a stressed syllable. Regarding consonants, some studies have noted the effects of stress on the duration of stop consonants. Crystal & House (1987) reported that when CV in the context of –VCV- is stressed, the consonant, particularly a stop is extended about 1.4 times (for closure), 1.75 times (for closure plus VOT) compared to an unstressed CV. Likewise, according to Stathopolous & Weismer (1983), the stop duration (closure) in stressed CV is 1.35 times longer than that in unstressed. English stress seems to prolong consonants as well as vowels. It is considered that American productions of Japanese stops were lengthened due to the interference of this stress factor from English. In the typical foreign accented 3-mora Japanese word by English speakers, stress-like accent is assigned to the second mora and this did happen in this experiment too. This tendency may derive from English stress. It seems that the second mora was erroneously stress-accented and this stress-like accent might lead to the consonant lengthening. For English

speakers, we concluded that some of the stop productions in Japanese were lengthened as a consequence of the assignment of stress-like accent, and lengthened stops resulted in native Japanese perceiving a geminate stop. Apart from the stop duration, V_1 durations for Americans were notable for their shorter duration than those of any group including Japanese. This may be interpreted as a compensatory shortening for the stress lengthening. A V_1 that precedes a stress-accented syllable is assumed to be shortened in the same pattern as has been observed in the native language (Hoequist, 1983) even in the production of a second language. Besides V_1, another adjacent vowel V_2, also seems to be slightly shortened compensating the long stop durations for Americans.

Syllable timing rhythm has been defined as a syllable recurring at a regular timing interval. Although perfect isochrony has been questioned for syllable timing, compared to stress-timed English, in syllable-timed French, syllable duration seems to be less affected by word contexts such as stress. In fact many studies have verified that the degree of stress-induced segmental lengthening is lower in French than in English (Fant, Kruckenberg & Nord, 1991). Accordingly French consonants are also expected not to be extended much by a stress factor and this can be confirmed by some studies that show consonantal data (O'Shaughnessy, 1984). According to these studies, consonants in the word final syllable were the almost the same length as in word non-final unstressed position. Because of French syllable timing that is basically free from stress-induced duration increase, the French speakers were successful in adjusting the duration of a single stop within a one mora unit.

For Chinese, one of the reasons for the longer stop duration in this study may be the consonant ratio to vowel rates in CV (consonant ratio). As is often the case with Chinese learners of Japanese, most of the Chinese subjects in this experiment substituted a Chinese aspirated stop for a Japanese voiceless stop in the recording. In Chinese, the consonant ratio of voiceless aspirated stop is high 1.2-1.5 (Feng, 1985) while the consonant ratio for Japanese voiceless stop is only 0.7 (Sagisaka & Tohkura, 1984). Due to the longer portion of stop in Chinese timing control, the influence of Chinese may have caused the longer stops in Japanese production. For Chinese speakers, further study is required to understand the mechanism of their timing control in Japanese by investigating the timing pattern in the Chinese language.

In this study, in examining the correlation between the geminate perception rate and stop duration, stop duration was CVD that includes VOT. The threshold CVD values from single to geminate perception (the point where geminate perception rate is 50%) in this experiment agreed with the averaged threshold values in the previous studies (Americans 171ms, Chinese 175ms, previous studies 183ms). However, if the threshold is calculated in terms of the closure (CD) only, the threshold CD values for American and Chinese utterances show lower values than those of the previous Japanese studies (Americans 128ms, Chinese 141ms, previous studies 169ms). These discrepancies in the threshold values between foreign utterances and Japanese native utterances derive from the length of VOT. Since Japanese has a very small VOT length, acoustic role of VOT in the geminate perception has not been researched. However, from our results showing that the threshold values were only agreeable in terms of CVD, it seems that Japanese speakers use VOT in the perception of a geminate. The acoustic role of VOT in the perception of Japanese geminate stops should be further examined through perception experiment using synthetic stimuli.

5. Conclusions

The present study provided acoustic evidence that the durations of some Japanese stops at the second mora produced by Americans and Chinese were longer than those of Japanese speakers and French speakers produced slightly longer stops than the Japanese speakers. Consequently, the proportion of productions perceived as geminate was rather high for stops by American and Chinese, while it was small for French productions. This language difference in timing pattern in Japanese stops was interpreted to reflect a difference in timing implementation of native language, at least concerning the American and French differences.

6. Acknowledgments

This work was supported by JSPS grant no.7849.

7. References

1. Crystal, T.H. & House, A.S. “Signal durations in connected speech: syllabic stress”. *Journal of Acoustical Society of America*, 83, 1574-1585, 1987.

2. Fant, G., Kruckenberg, A. & Nord, L. “Durational correlates of stress in Swedish, French and English”. *Journal of Phonetics*, 19, 351-365, 1991.

3. Feng, R.“Duration of initials , finals, and tones in Beijing dialect”. *Working papers in experimental phonetics*, 131-195. Beijing University press. (in Chinese) 1985.

4. Fujisaki, H., Nakamura, K. & Imoto, T. “Auditory perception of duration of speech and non-speech stimuli”. *Annual Bulletin No.7*, 45-64. University of Tokyo: RILP, 1973.

5. Hoequist, C. “Syllable duration in stress-, syllable- and mora-timed languages”. *Phonetica*, 203-237, 1983.

6. O’Shaughnessy, D. “A multiplespeaker analysis of durations in real French paragraphs”. *Journal of Acoustical Society of America*, 76(6), 1664-1672, 1984.

7. Sagisaka, Y. & Tohkura, Y. “Phoneme duration control for speech synthesis by rule”. *Journal of EICE*, 67-A. 629-636, 1984.

8. Stathopoulos, E. & Weismer, G. “Closure duration of stop consonants”. *Journal of Phonetics*, 11, 395-400, 1983.

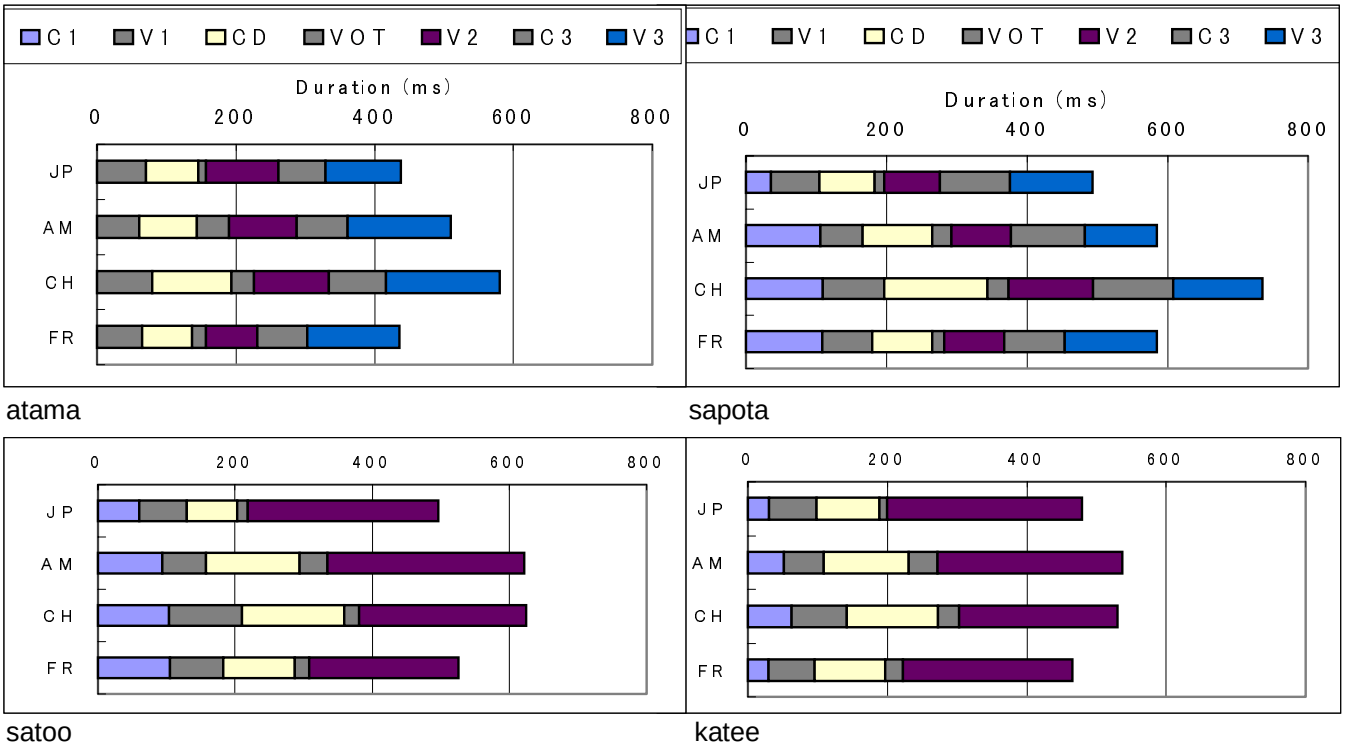

Figure 1. Mean segmental duration in 4 words for each of the 4 language groups.

Figure 2. Averaged N value of closure duration and VOT in 7 words for each of the 4 language groups.

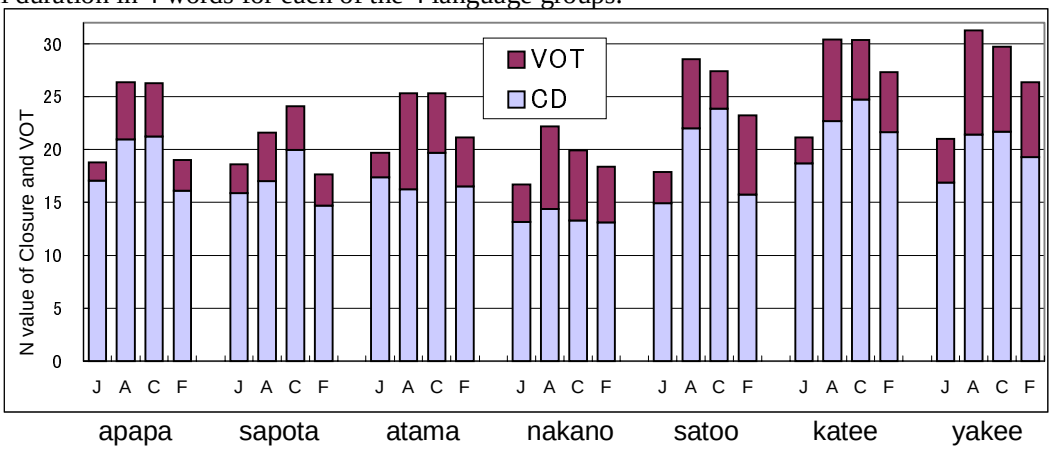

Figure 3. Geminate perception rate in 7 words for each of the 4 language groups