

Vowel-to-vowel coarticulation in Japanese: The effect of consonant duration (L)

Anders Löfqvist^{a)}

Haskins Laboratories, New Haven, Connecticut 06511 and Department of Logopedics, Phoniatrics and Audiology, Clinical Sciences, Lund University, Lund, Sweden

(Received 31 January 2008; revised 24 July 2008; accepted 25 July 2008)

This paper examines vowel-to-vowel lingual coarticulation in sequences of vowel-bilabial consonant-vowel, where the duration of the oral closure for the consonant is either long or short. Native speakers of Japanese served as subjects. The linguistic material consisted of Japanese word pairs that only differed in the duration of the labial consonant, which was either long or short. Recordings were made of lip and tongue movements using a magnetometer system. It was hypothesized that there would be greater vowel-to-vowel coarticulation in the context of a short consonant, since a long consonant would allow the tongue more time to move. The overall results do not show any strong support for this hypothesis, however. Subjects modulate the speed of the tongue movement between the two vowels, making it slower during the long than during the short consonant. © 2009 Acoustical Society of America. [DOI: 10.1121/1.2973234]

PACS number(s): 43.70.Bk, 43.70.Aj [BHS]

Pages: 636–639

I. INTRODUCTION

This paper examines vowel-to-vowel coarticulation in sequences of vowel-bilabial consonant-vowel, where the duration of the oral closure for the consonant is varied for linguistic purposes, using speakers of Japanese. In Japanese, the ratio of closure duration between long and short consonants is about 2:1 (Beckman, 1982; Han, 1994; Hirata and Whiton, 2005; Löfqvist, 2005, 2006, 2007). One might thus hypothesize that there is a greater influence from an upcoming vowel on the preceding one when the intervening labial consonant is short than when it is long. In the short consonant context, there is less time for the tongue to make the transition from the first to the second vowel. A similar argument can be made for an influence of the preceding vowel on the following one, although variations in tongue movement kinematics could neutralize the effect of such a reduced temporal window for the movement. That is, the results of Löfqvist (2006) show that the duration of the tongue movement between the two vowels is longer when the consonant is long. In addition, the average speed of the tongue movement between the vowels is slower in the long consonant context, although there is no systematic difference in the magnitude of the tongue movement as a function of consonant length.

It is well known that successive sounds in speech influence each other. Such influences can occur over quite large temporal intervals (e.g., Magen, 1997) and across segment, syllable, and word boundaries. The nature and extent of such coarticulatory influences seem to depend on several factors, including speaking rate, stress, and the articulatory requirements for different segments (e.g., Modaresi *et al.*, 2004). These requirements have sometimes been indexed in terms of coarticulatory resistance, originally proposed by Bladon

and Al-Bamerni (1976), i.e., how resistant a sound is to coarticulatory influences. For example, a fricative consonant made with the the front part of the tongue in contact with the hard palate has generally been found not to be very much affected by coarticulation (e.g., Fowler and Brancazio, 2000). In the present study, the coarticulation resistance of the consonant in the VCV sequence is not a serious issue, since it is a labial nasal, not produced with the tongue, and the tongue is not rigidly coupled to the jaw. Moreover, Fowler and Brancazio (2000) found little influence of coarticulation resistance on vowel-to-vowel coarticulation.

The present study thus examines vowel-to-vowel coarticulation across a labial nasal consonant with two different durations using articulatory movement tracking with a magnetometer in native speakers of Japanese. The specific hypothesis being addressed is that any such influences will be stronger in the context of a short than of a long labial consonant; this specific issue was not examined by Löfqvist (2006).

II. METHOD

A. Subjects

Five native speakers of Japanese, three male and two female, served as subjects. They reported no speech, language, or hearing problems. They were naive as to the purpose of the study. Before participating in the recording, they read and signed a consent form. (The experimental protocol was approved by the IRB at the Yale University School of Medicine.)

B. Linguistic material

The linguistic material consisted of Japanese words with a sequence of vowel-labial nasal-vowel. These words formed minimal pairs, where the only difference between the pairs was the duration of the labial consonant. The words were designed to require a substantial amount of tongue move-

^{a)}Present address: Haskins Laboratories, 300 George St., New Haven, CT 06511. Electronic mail: lofqvist@haskins.yale.edu

ment from the first to the second vowel. The following words were used: /kami, kammi/ (“god,” “sweets”), /kamee, kammee/ (“participation,” “impression”) /kema, kemma/ (“Kema, place name,” “polish”). The linguistic material was organized into randomized lists and presented to the subjects in Japanese writing, with the words occurring in a short frame sentence. Fifty repetitions of each word were recorded.

C. Movement recording

The movements of the lips, the tongue, and jaw were recorded using a three-transmitter magnetometer system (Perkell *et al.*, 1992). Receivers were placed on the vermilion border of the upper and lower lip, on three positions of the tongue, referred to as tip, blade, and body, and on the lower incisors at the gum line. Two additional receivers placed on the nose and the upper incisors were used for the correction of head movements. Two receivers attached to a plate were used to record the occlusal plane by having the subject bite on the plate during the recording. All data were subsequently corrected for head movements and rotated to bring the occlusal plane into coincidence with the x -axis. This rotation was performed to obtain a uniform coordinate system for all subjects (cf., Westbury, 1994).

The articulatory movement signals were sampled at 500 Hz after low-pass filtering at 200 Hz. The resolution for all signals was 12 bits. After voltage-to-distance conversion, the movement signals were low-pass filtered using a 25 point triangular window with a 3-dB cutoff at 14 Hz; this was done forwards and backwards to maintain phase. To obtain the instantaneous velocity of the tongue receivers, the first derivative of the position signals was calculated using a three-point central difference algorithm. For each tongue receiver, its speed [$v = \sqrt{(x^2 + y^2)}$] was also calculated. The velocity and speed signals were smoothed using the same triangular window. The acoustic signal was pre-emphasized, low-pass filtered at 4.5 kHz and sampled at 10 kHz.

The horizontal and vertical positions of the tongue body during the first and second vowels were defined algorithmically in the tongue body speed signal as minima during the first and second vowels, see Fig. 1(a); Fig. 1(b) shows the movements of all receivers during this interval. They correspond to the onset and offset of the tongue movement between the two vowels. We should note that at these points in time, the horizontal and vertical velocities of the tongue are usually not zero. This is partly because the kinematic signals are expressed in a maxilla-based coordinate system, thus the recorded tongue body movement also includes the contribution of the jaw. Such a coordinate system is appropriate when we are interested in the tongue as the end effector

T-tests were used to assess differences between the long and short consonants for each subject. Given the large number of comparisons, an α -level of 0.001 was adopted based on dividing the standard alpha level of 0.05 by the number of comparisons.

III. RESULTS

The duration of the oral closure for the labial consonant showed a robust difference with no overlap between the val-

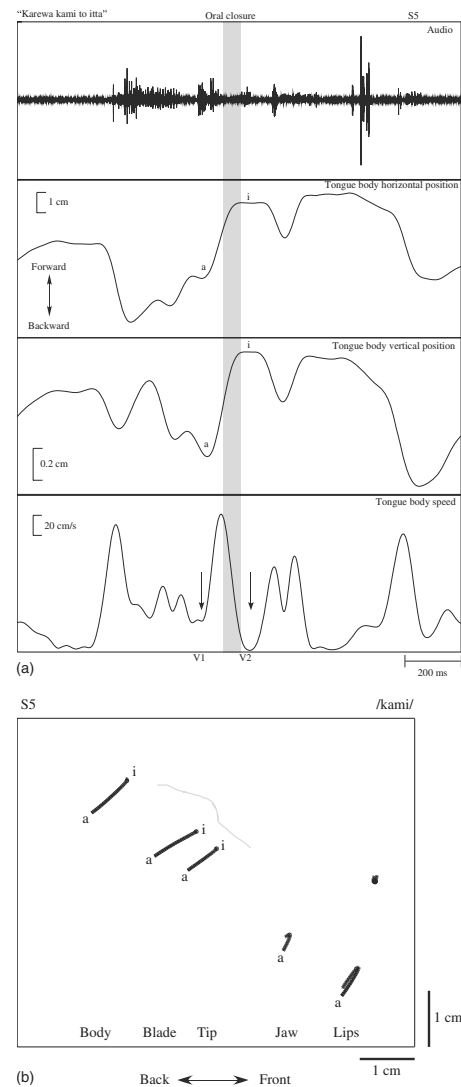


FIG. 1. (a) Audio and tongue body signals for the word /kami/ with arrows showing the points used in the speed signal for defining the onset and offset of the tongue movement between the two vowels. The baseline in the bottom panel with the speed signal represents zero speed. (b) Articulatory movements from the first to the second vowel in “kami.” The gray line represents a tracing of the hard palate.

ues for the short and long ones. The range of durations for the short consonants was 54–95 ms, while that for the long ones was 119–165 ms (Löfqvist, 2006).

A. Tongue body position during the first vowel

Figure 2 presents the tongue body positions during the first vowel. According to the hypothesis, there should be a difference in the tongue positions between the long and short vowel contexts. In particular, it was expected that the tongue would be in a higher and more front position in the words with a short consonant /kami, kamee/ than in the words with a long consonant /kammi, kammee/ due to the influence of the front second vowel. In the words /kema, kemma/, the hypothesis predicts a lower and more retracted tongue position for the first vowel in /kema/ than in /kemma/ due to the influence of the second, back, vowel. An inspection of Fig. 2 suggests that the prediction for the words /kami, kammi, kame, kammee/ is partly supported: The unfilled circles and

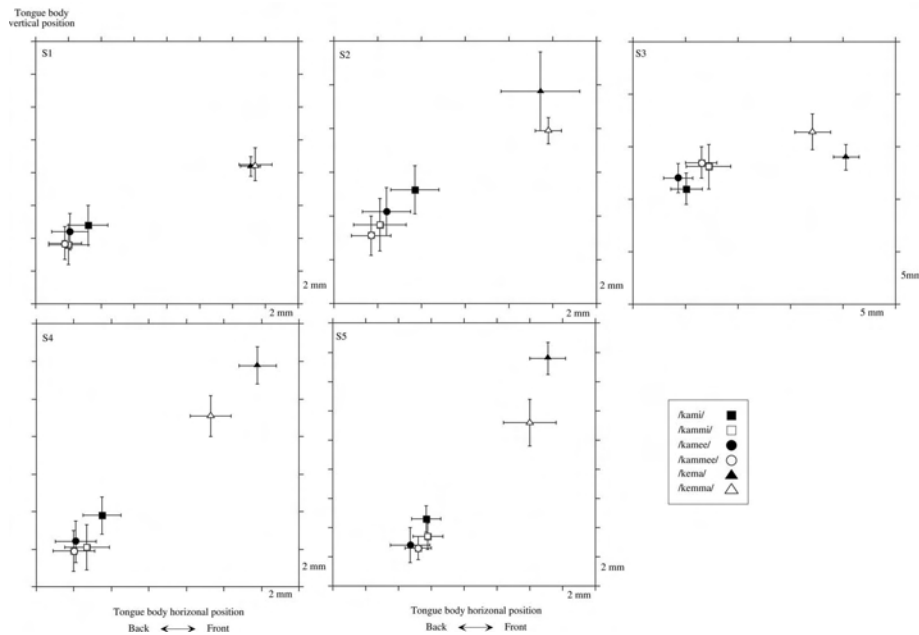


FIG. 2. Tongue body positions during the first vowel (mean and standard deviation).

squares are in a lower and more posterior position than their filled counterparts for four of the subjects 1, 2, 4, and 5. That is, the vowels in the words with a long consonant are less influenced by the second vowel than those in the words with a short consonant. However, for subject 3, the opposite is the case. For the words /kema, kemma/, the prediction is not supported for most subjects, since the filled triangles (/kema/) tend to be in a higher and more anterior position than their unfilled counterparts for subjects 2, 3, 4, and 5. The statistical analysis for /kami, kammi/ showed a significant difference in horizontal tongue position for subjects 1, 2, 3, and 4 ($t=6.4, 10.89, -7.97, 4.32$, with $p<0.001$ in all cases). Note however, that for subject 3, the results are opposite to the predicted ones. Subject 5 showed no statistical difference ($t=-0.62$ ns). For the vertical tongue position, all subjects showed a significant difference ($t=4.81, 7.3, -5.81, 7.73, 7.28$, $p<0.001$), again with the results of subject 3 opposite to the predicted ones.

In the words /kamee, kammee/, only subjects 2 and 3 showed a difference in the horizontal position ($t=4.81$ and -8.71 , $p<0.001$), but not subjects 1, 4, and 5 ($t=1.17, 0.26$, and -1.85 ns). Again, the results for subject 3 are opposite to the predicted ones. For the vertical position, only subjects 2 and 3 showed a significant difference ($t=5.94$, and -5.06 , $p<0.001$) but not subjects 1, 4, or 5 ($t=3.22, 2.55, 0.94$ ns). Finally, for the words /kema, kemma/, the horizontal position was different for subjects 3, 4, 5 ($t=10.13, 8.62$, and 4.49 , $p<0.001$), but not for subjects 1 and 2 ($t=-1.17$, and -2.53 ns). For the vertical position, subjects 2, 3, 4, and 5 showed a difference ($t=6.68, -8.46, 12.66, 15.87$, $p<0.001$), but not subject 1 ($t=-0.34$ ns). Note, however, that this difference was opposite to the predicted one.

The results for the first vowel thus show that for the low back vowel /a/, three of the subjects showed some qualified support for the hypothesis that there would be less vowel-to-vowel coarticulation across a long consonant. However, for

the front vowel /e/, the opposite pattern was found in four subjects. Overall, the support for the hypothesis is very weak.

B. Tongue body position during the second vowel

Figure 3 shows the tongue body position during the second vowel. Here, the prediction is that the front vowels /i, e/ would have a lower and more retracted tongue position in the words /kami/ and /kamee/ due to the influence of the low back first vowel. For the low back vowel /a/ in /kema, kemma/ the prediction is that it will have a higher and more advanced tongue position due to the influence of the first vowel /e/.

The overall results suggest that there is no reliable difference in the tongue position during the second vowel as a function of consonant length. Subject 2 showed the horizontal positions in the words /kamee, kammee/ to be significantly different ($t=-3.77$, $p<0.001$). For subject 3, there was a significant difference in the horizontal position for the words /kame, kammee/ ($t=4.46$, $p<0.001$), and also for the horizontal position for the words /kema, kemma/ ($t=4.58$, $p<0.001$). Subject 4 had differences in the horizontal position for the words /kame, kamme/ ($t=-3.48$, $p=0.001$), and the vertical position for the words kema, kemma/ ($t=5.15$, $p<0.001$). Finally, subject 5 only showed a difference in the horizontal position for the words /kema, kemma/ ($t=8.89$, $p<0.001$). Of these statistically reliable differences, only three of them were in the predicted direction, subjects 2, 4, and 5, but in no case was the difference found for both the horizontal and vertical positions. For subject 3, the two reliably different results were opposite to the predicted ones. Overall, the results for the second vowel show no consistent differences in tongue position as a function of consonant length.

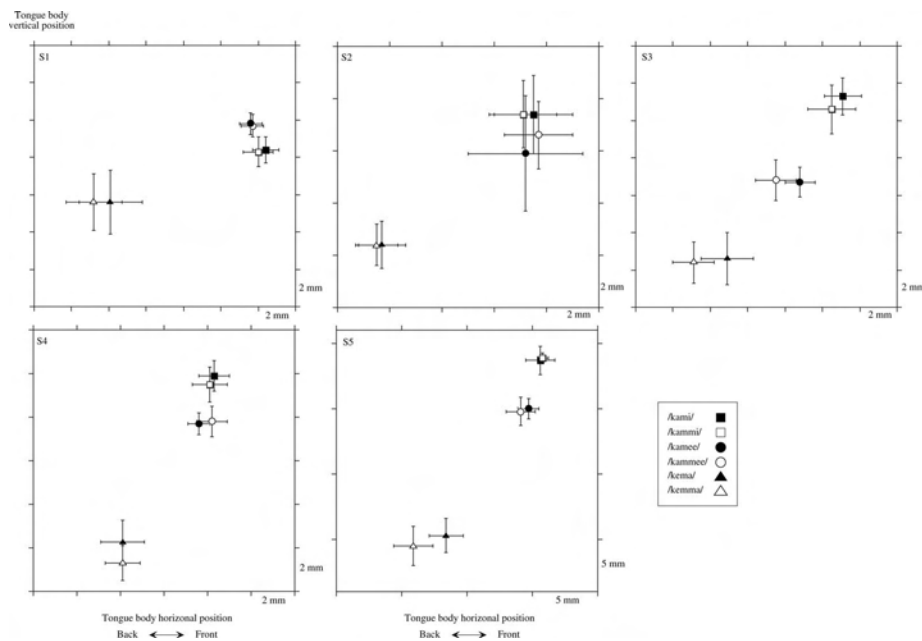


FIG. 3. Tongue body positions during the second vowel (mean and standard deviation).

IV. DISCUSSION

This study examined the influence of consonant duration on vowel-to-vowel coarticulation in Japanese. It was hypothesized that a short intervocalic labial consonant would allow more coarticulation than a long consonant, in particular since the duration of a long consonant is about twice as long as that of a short consonant. The overall results do not show any strong support for this hypothesis, however. Three of the subjects showed the expected influence of a following high vowel on a preceding low back vowel, but one subject showed the opposite results. There were no effects on the second vowel. Thus, there was some limited evidence for more anticipatory influences than carryover effects.

The most likely reason for the small effects is that Japanese speakers adjust the speed of the tongue movement to maintain a similar, but not identical, coordination of lip and tongue movements for long and short consonants (Löfqvist, 2006). That is, the onset of the tongue movement occurs before the oral closure for the consonant, and its offset occurs after the oral release. As a consequence, the tongue positions for the vowels in the context of the long and short consonants are very similar. A further consequence is that the duration of the tongue movement between the two vowels is longer when the intervening consonant is long than when it is short. The magnitude of the movement path between the two vowels did not vary systematically with consonant duration. Not modulating the speed of the tongue movement would result in the tongue reaching the intended target for the second vowel well before the release of the long consonant and it might then have to stop moving. Such a movement pattern would involve successive accelerations and decelerations of the tongue that would involve a higher cost of effort. Thus, speakers avoid excessive accelerations and decelerations of the tongue by keeping it moving.

ACKNOWLEDGMENTS

The author is grateful to Mariko Yanagawa for help with the Japanese material and running the experiments. This work was supported by the National Institute on Deafness and Other Communication Disorders, National Institutes of Health Grant No. DC-00865.

- Beckman, M. (1982). "Segment duration and the mora in Japanese," *Phonetica* **39**, 113–135.
- Bladon, A., and Al-Bamerni, A. (1976). "Coarticulation resistance in English /l/," *J. Phonetics* **4**, 137–150.
- Fowler, C. A., and Brancazio, L. (2000). "Coarticulation resistance of American English consonants and its effects on transconsonantal vowel-to-vowel coarticulation," *Lang Speech* **43**, 1–41.
- Han, M. (1994). "Acoustic manifestations of mora timing in Japanese," *J. Acoust. Soc. Am.* **96**, 73–82.
- Hirata, Y., and Whiton, J. (2005). "Effects of speaking rate on the single/geminate stop distinction in Japanese," *J. Acoust. Soc. Am.* **118**, 1647–1660.
- Löfqvist, A. (2005). "Lip kinematics in long and short stop and fricative consonants," *J. Acoust. Soc. Am.* **117**, 858–878.
- Löfqvist, A. (2006). "Interarticulator programming: Effects of closure duration on lip and tongue coordination in Japanese," *J. Acoust. Soc. Am.* **120**, 2872–2883.
- Löfqvist, A. (2007). "Tongue movement kinematics in long and short Japanese consonants," *J. Acoust. Soc. Am.* **122**, 512–518.
- Magen, H. (1997). "The extent of vowel-to-vowel coarticulation in English," *J. Phonetics* **25**, 187–205.
- Modaresi, G., Sussman, H., Lindblom, B., and Burlingame, E. (2004). "An acoustic analysis of the bidirectionality of coarticulation in VCV utterances," *J. Phonetics* **32**, 291–312.
- Perkell, J., Cohen, M., Svirsky, M., Matthies, M., Garabieta, I., and Jackson, M. (1992). "Electromagnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements," *J. Acoust. Soc. Am.* **92**, 3078–3096.
- Westbury, J. (1994). "On coordinate systems and the representation of articulatory movements," *J. Acoust. Soc. Am.* **95**, 2271–2273.