

JANUARY 09 2025

Articulatory correlates of consonantal length contrasts: The case of Japanese mimetic geminates

Francesco Burroni  ; Shigeto Kawahara  ; Jason A. Shaw 



JASA Express Lett. 5, 015201 (2025)

<https://doi.org/10.1121/10.0034762>



Articles You May Be Interested In

The articulatory kinematics of final lengthening

J Acoust Soc Am (January 1991)

Phrase boundary effects on the temporal kinematics of sequential tongue tip consonants

J. Acoust. Soc. Am. (June 2008)

Intersegmental coordination adjacent to multiple prosodic boundaries

J Acoust Soc Am (May 1998)



ASA

Advance your science and career as a member of the **Acoustical Society of America**

[LEARN MORE](#)

Articulatory correlates of consonantal length contrasts: The case of Japanese mimetic geminates

Francesco Burroni,^{1,a)}  Shigeto Kawahara,²  and Jason A. Shaw³ 

¹Institute for Phonetics and Speech Processing, Ludwig Maximilian University, Munich, Bavaria 80799, Germany

²The Institute of Cultural and Linguistic Studies, Keio University, Minato-ku, Tokyo 108-8345, Japan

³Department of Linguistics, Yale University, New Haven, Connecticut 06520, USA

francesco.burroni@phonetik.uni-muenchen.de, kawahara.research@gmail.com, jason.shaw@yale.edu

Abstract: This study investigates the articulatory correlates of consonantal length contrasts in Japanese mimetic words using electromagnetic articulography data. Regression and dynamic time warping analyses applied to intragestural timing, kinematic properties, and intergestural timing reveal that Japanese geminates are characterized by longer closure phases, longer gestural plateaus, higher tongue tip positions, larger movements, and lower stiffness. Geminates also exhibit distinct timing relationships with adjacent vowels, specifically, longer times to target that allow for longer preceding vowels. These findings shed light on the articulatory mechanisms underlying Japanese geminate production, their relationship to acoustics, and their characterization in a broader cross-linguistic perspective. © 2025 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

[Editor: Peggy Mok]

<https://doi.org/10.1121/10.0034762>

Received: 21 August 2024 Accepted: 11 December 2024 Published Online: 9 January 2025

1. Introduction

In many of the world's languages, consonant duration can be employed contrastively.^{1,2} A well-known example is Japanese. The production of Japanese long or “geminate” consonants has been primarily studied through their acoustic manifestations.³ Acoustically, Japanese geminates are distinguished from singletons mainly by their longer constriction durations.⁴ Geminates are also accompanied by slightly longer preceding vowels and by other non-durational cues: for example, larger intensity differences surrounding geminates, larger pitch-accent f_0 drops across geminates, vowels with lower F_1 after geminates, etc. (cf. Ref. 3).

Unlike their acoustic correlates, the articulatory mechanisms involved in Japanese geminate production have not been thoroughly explored.^{5–12} Three key areas require further investigation: (i) intragestural timing properties, which encompass the duration and timing characteristics of geminate articulation; (ii) kinematic properties, which are responsible for the shape of the trajectories of articulatory movements during geminate production; and (iii) intergestural timing properties, i.e., the timing relationships between geminate articulation and the articulation of surrounding sounds.

With respect to intragestural timing, one understudied aspect is how the longer closure duration of geminates is implemented kinematically. Two main strategies have been reported. The first strategy is the realization of geminates by holding constriction targets for longer periods of time: in effect, lengthening the plateau of kinematic trajectories.⁵ A second strategy is the realization of geminates by slowing down articulator movements, specifically, tongue movements for lingual consonants.⁹

With respect to kinematic features, one study of Japanese bilabial geminate plosives⁷ has uncovered that geminates display (i) higher articulator positions, (ii) larger movement amplitude, (iii) similar closure peak velocities, and (iv) lower stiffness. These findings point to distinct kinematic specifications for articulator control during geminate production, beyond simply longer durations (cf. also Refs. 3 and 13). More extended articulator contacts, compatible with more constricted targets, have also been reported for lingual consonants using electropalatography.^{13,14} However, more investigations of kinematic features in various manners and places of articulation are necessary to fully characterize geminate articulation in Japanese.

Finally, with respect to intergestural timing, pioneering comparative work on Japanese and Italian geminates by Smith^{5,6} has uncovered a potential kinematic basis for the longer duration of vowels before geminates in Japanese. This longer duration may be in part due to intrinsically longer kinematic movements, as also suggested in subsequent work.⁹ Crucially, however, Japanese geminates also display longer times to target compared to singletons, which may allow preceding vowels to be acoustically manifested for a longer period of time. This timing feature sets Japanese apart from other

^{a)} Author to whom correspondence should be addressed.

languages, for example, from Italian. In Italian, geminates are realized by anticipating closure onset during the preceding vowel, thus, in effect shortening the acoustic manifestation of the vowel.^{5,6,15}

Smith^{5,6} also reported longer trans-consonantal lags between preceding and following vowels across geminates than singletons. This finding suggests a more sequential organization for vowels across a geminate in Japanese than in Italian, in which the trans-consonantal vocalic lag is unaffected or perhaps even shortened, at least with bilabial geminates.^{5,6,15} Smith's^{5,6} works offered an intriguing basis for the intergestural timing organization of geminates. However, subsequent work on Japanese geminates did not show clear intergestural timing differences^{8,9} and has also called into question the relationship between a longer time to target and duration of the preceding vowel.¹¹

With these issues in mind, we investigated the articulatory correlates of geminate production in Japanese using electromagnetic articulography (EMA). We analyzed geminate production along the three main lines discussed above, namely, in terms of their intragestural timing, kinematic properties, and intergestural timing. The questions we addressed are as follows: (1) What are the kinematic correlates that underlie the longer duration of Japanese geminates? (2) What kinematic features differentiate the articulatory trajectories of singletons and geminates? (3) How are geminates timed with respect to surrounding segments, and is there a kinematic basis for the longer duration of preceding vowels?

Given that both acoustic and articulatory properties can be influenced by a host of lexical factors,^{16–19} we studied the properties of Japanese geminates by comparing them to singletons with the same lexical item. This is possible because Japanese mimetic (i.e., sound symbolic) words can be produced with gemination of a medial consonant to express emphasis.²⁰ Japanese mimetic words thus offer a controlled testing ground for geminate articulation in a situation that is largely independent of lexical influences.

2. Methods

2.1 Participants and materials

Seven native Japanese speakers (3 female, 4 male) participated in the experiment. All reported normal speech and hearing and daily use of the Tokyo dialect.

The speech material consisted of 21 existing Japanese mimetic words, with three items for each of seven consonantal types (Table 1). Each target consonant was produced as a singleton or a geminate (for emphasis).²⁰

Given that previous work has focused on bilabial consonants^{5–8,12} and our interest also lies in the interaction between geminates and preceding vowels, we focused on consonants produced with the same articulator as vowels, the tongue. Thus, we examined apical/laminal consonants, as they are produced with the front of the tongue, which is straightforward to track with EMA.

In each trial, participants produced singleton mimetic words in isolation, followed by their emphatic geminate variant, also produced in isolation. Items were elicited in isolation to minimize total experiment duration and confounds that could arise from different prosodic realizations of a carrier phrase. The target consonants were never adjacent to word boundaries, thus, also limiting the need for a carrier phrase. The order of lexical items was fully randomized. In total, 21 unique words × 2 realizations (singleton/geminate) × 10 repetitions × 7 speakers yielded a total of 2940 tokens for analysis.

2.2 Experimental procedure

Articulatory movements were captured using an NDI Wave System [NDI (Northern Digital, Inc.), Waterloo, Canada] sampling at 100 Hz (S1–S3, 1F, 2M) and a Carstens AG501 system (Carstens, Chicago, IL) sampling at 1250 Hz (S4–S7, 2F, 2M). Acoustic data were simultaneously recorded at 22.05 kHz. The two datasets were first analyzed separately. Since no substantial differences were found, we present combined analyses of the two datasets. Our supplementary material contains scripts demonstrating our analyses for the datasets, both separately and combined, for the interested readers.

Both systems used a nearly identical sensor setup: five tongue sensors (three on the sagittal midline, two parasagittally). Only the most anterior sensor on the tongue, attached less than 1 cm from the tongue apex, entered into the analyses reported here. Additional sensors were placed on the lips, jaw, left/right mastoid processes, nasion, and maxilla (for AG501 sessions only).

Table 1. Experimental stimuli.

Target	Item 1	Item 2	Item 3
[t]	kat(α)a–kata	gat(α)a–gata	pet(α)a–peta
[d]	kud(α)o–kudo	gud(α)a–guda	od(α)o–odo
[r]	par(α)a–para	per(α)a–pera	dor(α)o–doro
[z]–[j]	giz(α)a–giza	oz(α)u–ozu	uj(α)i–uji
[s]	kos(α)o–koso	kas(α)a–kasa	ϕus(α)a–ϕusa
[ts]	ϕut(α)su–ϕutsu	kat(α)su–katsu	gut(α)su–gutsu
[t̚ e̞]	net(̄a)ei–net̄ei	kat(̄a)ea–kat̄ea	gut(̄a)ea–gut̄ea

Participants sat in a sound-attenuated room; stimuli were displayed in Japanese orthography on a monitor ~25 cm outside the EMA field. Stimulus presentation was controlled using EPRIME (NDI Wave) or MATLAB²¹ (AG501), enabling real-time monitoring for errors. Rare hesitations or mispronunciations were marked and reinserted randomly among the remaining items. Following the main recording session, we also recorded the bite plane of each participant by having them hold a rigid object, with three (NDI Wave) and two (AG501) sensors attached to it, between their teeth. Head movements were corrected computationally after data collection with reference to three sensors on the head and the two/three sensors on the bite plane and maxilla. The head corrected data were rotated so that the origin of the spatial coordinates corresponds to the occlusal plane at the front teeth.

2.3 Data processing and analyses

The acoustic signals were manually segmented at the word and segmental level in PRAAT²² using standard criteria based on waveform and the spectrogram characteristics by two research assistants naive to the purpose of the study (see Fig. 4 below). The segmentation was later double-checked by the first author. The acoustic segmentation was used as a starting point for articulatory landmarking of target consonants. The only acoustic boundaries used in the analyses involve the acoustic duration of the target consonants and preceding and following vowels. Given the wide variety of vocalic contexts and the impossibility to consistently identify vocalic landmarks from the kinematics alone, we adopt acoustic landmarking as events that reasonably correlate with kinematic events, especially target achievement, following previous work (e.g., Ref. 23).

The articulatory analyses reported in this paper focus on apical/laminal consonants; thus, the main sensor of interest is the tongue tip, henceforth TT. The wide variety of vocalic contexts in which target consonants appears has a strong influence on horizontal tongue position; thus, our landmarking was based on the dimension that we found to be most reliable to identify consonantal articulation, namely the TT vertical movement (henceforth TTy) (cf. Ref. 24 for a similar choice). Support in favor of the decision of using TTy as the main landmarking dimension also comes from a principal component analysis (PCA). We entered all three dimensions of TT movement in a PCA and found that the direction of maximum movement, the 1st principal component (PC) (explaining 95.7% of variance), correlates almost perfectly with vertical movement (median $r = 0.97$ over all tokens). These results suggest that tongue movement during target consonants production takes place robustly in TTy and that this dimension is reliable to track articulation.

All articulatory signals were smoothed and interpolated using the algorithm of Ref. 25. For landmarking, TTy trajectories were upsampled to 1000 Hz for the NDI Wave sessions and kept at 1250 Hz for the AG501 sessions. All trajectories were smoothed using a Savitzky–Golay finite filter of polynomial order 4 and frame length of 21. Consonantal gestures were automatically landmarked using a custom MATLAB routine based on peak velocity thresholding. Acoustic boundaries were used to locate the consonant midpoint, which defined a symmetric window spanning twice the segment's acoustic duration. Within this window, velocity extrema were identified for upward closure and downward release movements in TTy. Gestural onset and target were defined as the first and last timepoints surpassing 20% of peak closure velocity, while release and offset were similarly defined for release velocity. Landmarks were visually inspected and adjusted for ~50 tokens as needed. Thirty tokens without clear gestural boundaries were excluded, along with their paired singleton or geminate, resulting in the exclusion of 60 tokens (~2% of the data). Excluding the paired token was necessary because some of our analyses require pairing singletons and geminates obtained from the same trial, as described below.

From the acoustic and articulatory signals, we derived a set of intragestural timing, kinematic, and intergestural timing measures. For intragestural timing (Fig. 1), we investigated the following: (i) The closure phase duration, defined as the lag between gestural onset and target; (ii) the plateau duration, defined as the duration of the lag between gestural target and release; (iii) the release phase duration, defined as the lag between gestural release and offset.

We also investigated the relationship between plateau duration and acoustic consonant duration using Pearson correlation.

Additionally, we also developed holistic analyses of the kinematic trajectories that rely on dynamic time warping (DTW). Specifically, we took each singleton/geminate pair in a trial and used DTW to derive a pairwise warping function that identifies which portions of a singleton articulatory trajectory need to be stretched to “derive” a geminate. We inspected the shape of the warping functions and obtained localized average warps over normalized duration, following Ref. 26.

For kinematic properties (Fig. 2), we investigated the following: (iv) Maximum vertical and horizontal tongue position, as a proxy for constriction target and location; (v) peak velocity during closure; (vi) movement amplitude from onset to maximum constriction during closure; (vii) kinematic stiffness, defined as the ratio of peak velocity to movement amplitude, providing a normalized measure of movement speed independent of amplitude. Higher stiffness indicates shorter times to reach peak velocity and target attainment.

For intergestural timing properties (see Fig. 4 below), we investigated the following: (viii) The duration of the vowel preceding the target consonants estimated from acoustics (V_1 Dur.); (ix) the duration of the lag between the preceding vowel acoustic onset and the consonantal gesture onset (V_1 Ons. – C Ons.); (x) the duration of the lag between the preceding vowel acoustic onset and the consonantal gesture target (V_1 Ons. – C Targ.); (xi) the duration of the trans-consonantal lag between the preceding and following vowels acoustic onsets (V_1 Ons – V_2 Ons.).

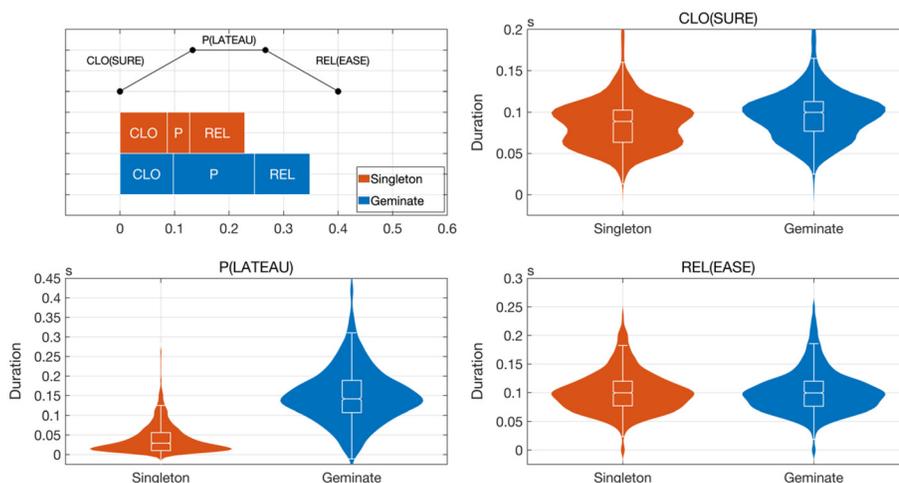


Fig. 1. (Top left) Schematic illustration of intragestural phases and their mean durations. (Top right) Box plot with superimposed Gaussian kernel density estimate (kde) of closure (CLO) duration. (Bottom left) Box plot of plateau (P) duration. (Bottom right) Box plot of release (REL) duration.

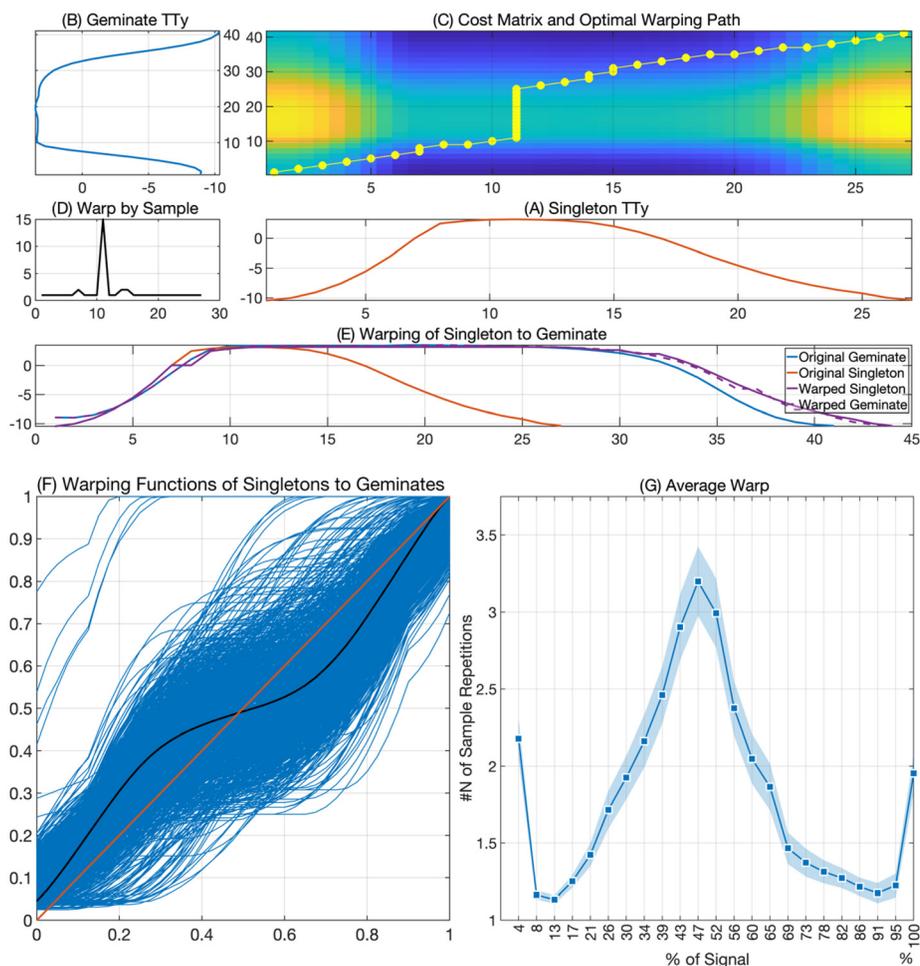


Fig. 2. (A) Example of TTy movement for singletons. (B) Example of TTy movement for geminates. (C) Cost matrix and optimal warping path for singleton to geminate alignment. (D) Warp function showing the number of repetitions each sample undergoes to stretch a singleton to a geminate. (E) Alignment of singleton and geminate TTy trajectories. (F) Warping function of singletons to geminates (blue lines), with average warping functions (solid black line) of singletons to geminates showing a strong distortion of linear time (orange line) in the plateau region (0.2–0.6). Time is normalized between 0 and 1. (G) Average warp at different percentages of the trajectory.

27 February 2025 14:35:41

We also investigated the relationship between consonantal and preceding vowel duration using Pearson correlation, following Ref. 11.

All properties (i)–(xi) were analyzed using linear mixed-effects regression models in MATLAB's *fitlme()* function. The effect of gemination was assessed via log-likelihood ratio tests, comparing a baseline model (without the geminate effect) to an alternative model including a fixed effect for gemination. Both models featured the same random effect structure, with by-subject and by-item random intercepts and a random slope for gemination, while the alternative model included a fixed effect for gemination (reference coded as singleton). Outliers were identified as values with *z*-scores below -4 or above 4 , resulting in the exclusion of 0% to 1.46% of the data across variables. For correlation analyses, values below the 1st percentile and above the 99th percentile were excluded to prevent extreme points from inflating or deflating observed correlations.

All data and scripts necessary to replicate our analyses and figures and to run separate analyses on the datasets are publicly available in an OSF repository at https://osf.io/27nyz/?view_only=4f93c383e24642e48d027c58fd945a27.

3. Results

3.1 Intra-gestural timing

We found a small but statistically reliable difference in closure phase duration [$\chi^2_{(1)} = 8.73$, $p = 0.003$]. The singleton closure phase is 86 ms (95% confidence interval [CI] [73–100] ms), while the geminate closure phase is +11 ms longer (95% CI [6–17] ms) (Fig. 1, top right).

A more robust difference was observed for plateau duration [$\chi^2_{(1)} = 12.61$, $p = 0.0004$]. The singleton plateau duration was 41 ms (95% CI [26–56] ms), while the geminate plateau duration was +108 ms longer (95% CI [72–144] ms) (Fig. 1, bottom left). No reliable differences were found in terms of the release phase duration [$\chi^2_{(1)} = 0.04$, $p = 0.84$] (Fig. 1, bottom right).

The fact that singletons and geminates differ primarily in terms of their plateau duration is also evident from DTW analyses. When each mimetic singleton is stretched to its geminate counterpart, a strong distortion of time is observed around the midpoint of the consonant, 0.2 to 0.6 of its proportional duration (Fig. 2F). Each singleton sample between 33% and 66% of the singleton trajectory is repeated between two and almost 3.5 times to derive a geminate, indicating a stretching of the plateau region (Fig. 2G).

The observed longer acoustic duration for geminates is closely related to the longer plateau, as indicated by a robust correlation between target consonant acoustic duration and plateau duration ($r = 0.84$, $p < 1 \times 10^{-215}$) (see Fig. 6 in the supplementary material).

3.2 Kinematic properties

We found a higher TTy position for geminates than singletons [$\chi^2_{(1)} = 8.76$, $p = 0.003$]. Compared to singletons (6.6 mm, 95% CI [1.6–11.6]), geminates are produced with higher/more constricted targets, +0.80 mm (95% CI [0.39–1.19] mm) (Fig. 3, mid left). In the horizontal position, the difference was not significant [$\chi^2_{(1)} = 3.25$, $p = 0.07$]; however, geminates tend to be more fronted, +0.23 mm (95% CI [0–0.47] mm).

The difference in movement amplitude was significant [$\chi^2_{(1)} = 8.5$, $p = 0.003$]. Compared to singletons (8.16 mm, 95% CI [6.7–9.7]), geminates are produced with greater closure movement amplitude, +0.99 mm (95% CI [0.45–1.53] mm) (Fig. 3, mid right). We found no significant differences in closure peak velocity [$\chi^2_{(1)} = 1.45$, $p = 0.22$] (Fig. 3, bottom left).

Finally, we found a significantly lower stiffness for geminates ($\chi^2_{(1)} = 5.13$, $p = 0.02$). Compared to singletons (19.92 s^{-1} , 95% CI [16.17–23.67 s^{-1}]), geminates are produced with lower stiffness, -2.61 s^{-1} (95% CI [−4.54 to −0.69] s^{-1}), indicating a slower movement and time to target (Fig. 3, bottom right).

3.3 Intergestural timing properties

We found that vowels preceding geminates had longer duration than vowels preceding singletons [$\chi^2_{(1)} = 7.57$, $p = 0.006$]. Specifically, compared to vowels preceding singletons (70 ms, 95% CI [60–79]), vowels preceding geminates were +13 ms longer (95% CI [6–20] ms) (Fig. 4, mid left panel).

The durations of the lag between vowel acoustic onset and consonantal gesture onset (V1 Ons. – C Ons.) were significantly different, but very similar between geminates and singletons [$\chi^2_{(1)} = 4.12$, $p = 0.04$] (Fig. 4, mid right panel). Compared to the lag of vowels preceding singletons (−4 ms, 95% CI [−1–7]), the lag of vowels preceding geminates is slightly longer by +7 ms (95% CI [1–13] ms). Note, however, that the 95% CIs are very close to an overlap with 0. Moreover, this finding only emerges by pooling our datasets together, suggesting that the difference between singleton and geminates in V1 Ons. – C Ons. is very small and variable. Future work should further investigate the robustness of this finding.

On the other hand, the duration of the lag between vowel acoustic onset and consonantal gesture target (V1 Ons. – C Targ.) was clearly longer in geminates than in singletons [$\chi^2_{(1)} = 12.6$, $p = 0.0004$]. Compared to the singleton lag (83 ms, 95% CI [69–96] ms), the geminate V1 Ons. – C Targ lag was +20 ms longer (95% CI [13–27] ms) (Fig. 4, bottom left panel).

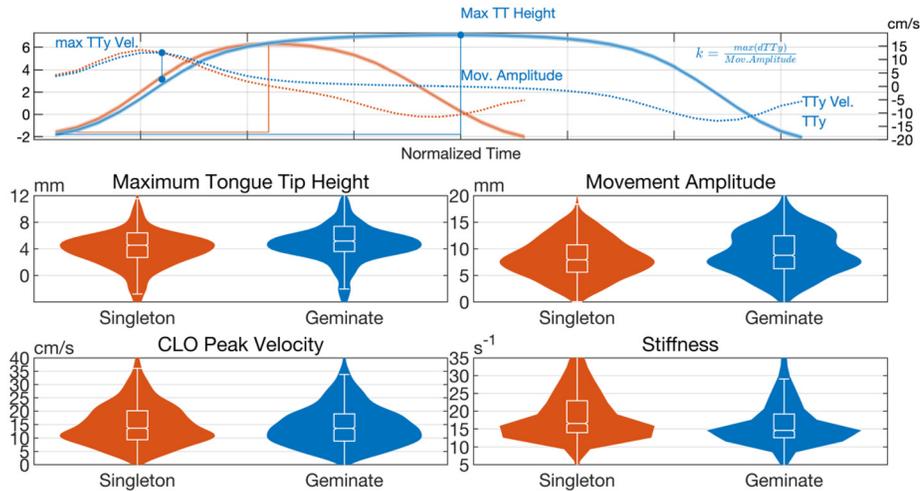


Fig. 3. (Top panel) Averaged vertical movement trajectories for singleton and geminate production (solid lines), together with their first derivatives (dashed lines) and overlaid markers for TT height and vertical velocity. Vertical lines from maximum TT height to onset illustrate movement amplitude. The kinematic stiffness (k) calculation is also displayed. (Mid left) Box plot with superimposed Gaussian kde of maximum TTy values. (Mid right) Box plot of movement amplitude values. (Bottom left) Box plot of CLO peak velocity values. (Bottom right) Box plot of stiffness values.

Finally, the trans-consonantal lag between the preceding vowel acoustic onset and following vowel acoustic onset ($V_1 \text{ Ons} - V_2 \text{ Ons.}$) was much longer in geminate than in singleton production [$\chi^2_{(1)} = 19.43, p < 0.0001$]. Compared to the lag in singleton production (218 ms, 95% CI [158–279] ms), the $V_1 \text{ Ons.} - V_2 \text{ Ons.}$ lag in geminate production was +134 ms longer (95% CI [108–160] ms) (Fig. 4, bottom right panel).

Given the findings presented above, we hypothesized that the longer duration of vowels preceding geminates may be due to a longer time to target compared to singletons, effectively allowing for a longer steady-state period for vowel production. Evidence in favor of this hypothesis is offered by a strong correlation ($r = 0.70, p < 1 \times 10^{-215}$) observed between time to target and preceding vowel duration (see Fig. 7 in the supplementary material).

4. Discussion

Returning to our research questions, our analyses revealed robust differences between Japanese geminates and singletons in terms of intragestural timing, kinematic properties, and intergestural timing organization.

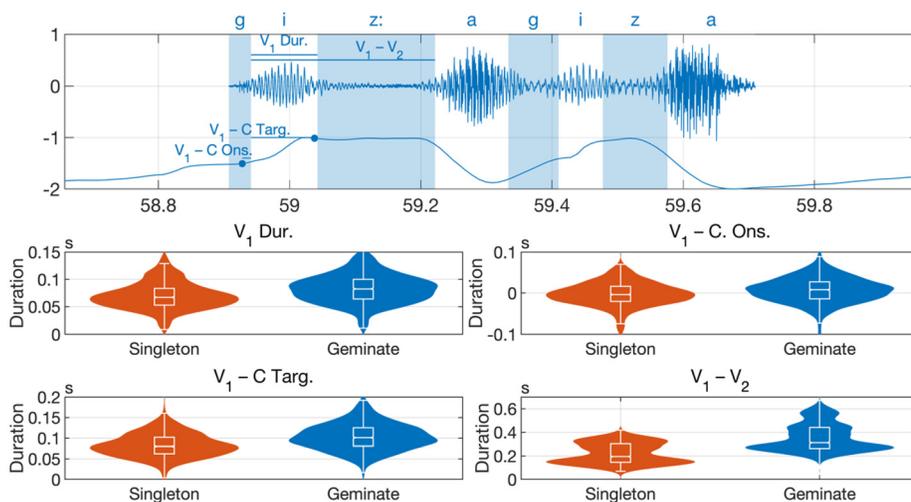


Fig. 4. (Top panel) Illustration of waveform and TTy trajectories, with shaded areas marking segmental boundaries for production of [gizaagiza]. Circular markers on TTy trajectory mark onset and target of tongue tip gesture for [z]. Labeled blue segments mark lag durations: V_1 duration, $V_1 - V_2$, $V_1 - C \text{ Ons.}$, and $V_1 - C \text{ Targ.}$ (Mid left) Box plot with superimposed Gaussian kde of V_1 duration values. (Mid right) Box plot of $V_1 - C \text{ Ons.}$ values. (Bottom left) Box plot of $V_1 - C \text{ Targ.}$ values. (Bottom right) Box plot of $V_1 - V_2$ values.

With respect to intragestural timing, geminates were produced with slightly longer closure phases and much longer gestural plateaus, a finding confirmed by DTW analyses. No differences were observed for release duration, in line with the previously-reported acoustic analyses, which reported no difference in Voice Onset Time.¹³ We also found a very strong correlation between consonantal acoustic duration and TTy plateau duration. Taken together, our findings align with previous work,^{5,9} indicating that Japanese speakers produce geminates with slightly longer closure movements and much longer gestural plateaus. Findings of longer closures and plateaus for geminates are not unique to Japanese. They have also been reported in other languages, like Italian.¹⁵

Turning to kinematic features, we found that Japanese speakers produce (lingual) geminates with a higher TT, i.e., with a more constricted posture, slightly larger movements, similar peak velocity, and lower stiffness, in line with previous reports.⁷ Taken together, the differences in kinematic features suggest that Japanese geminates are not just longer versions of singleton consonants. Their articulation is characterized by a different set of kinematic parameters. In this respect, our findings support the view that geminates differ from singletons not only in terms of durational properties but also in terms of more general articulatory strategies, an idea proposed for Tashlhiyt Berber² and Italian^{2,15} and that also been entertained for Japanese^{3,13} and other Japonic languages.²⁷

Finally, in terms of intragestural timing, we found that Japanese speakers produced slightly longer vowels before geminate consonants. Geminates also start roughly at the same time as singletons with respect to the preceding vowel, but their targets are reached later. Longer trans-consonantal lags between preceding and following vowels across geminates are also observed. These findings suggest that Japanese speakers produce geminates and singletons with relatively similar timing organization with respect to the preceding vowel. This has been noted in previous work^{5,6} and is a feature that sets apart Japanese geminate production from languages like Italian, where geminate closure robustly “intrudes” in the preceding vowel resulting in shortening of pre-geminate vowels.^{5,6,15}

Additionally, unlike some previous work,¹¹ our findings also suggest a robust positive correlation between the V_1 – C Targ. lag and V_1 duration. This pattern provides a possible kinematic basis for the longer vowel durations preceding Japanese geminates. Longer times to target allow the vocalic gesture to be acoustically exposed for a longer period of time before the acoustic consequences of consonantal articulation “kick in” during the consonantal plateau, which, as we have demonstrated, is the main articulatory correlate of consonantal acoustic duration. The longer time to target also reinforces other potential bases for longer vocalic acoustic duration. Namely, the slower and longer tongue body movements associated with vowels during geminate production^{5,6,9} can contribute to longer acoustic vowel durations in the presence of delayed consonantal target achievements for geminates and their associated acoustic consequences. Such slower tongue body movements are also observed in our data. This is illustrated in Fig. 8 in the supplementary material, where we present DTW analyses to show that there is a generalized need to warp time. Specifically, to obtain the combined vertical and horizontal tongue movement observed during pre- and post-geminate vowel production from the tongue movement observed during pre- and post-singleton vowel production time needs to be slowed down, especially while the consonant is being produced.

5. Conclusion

To conclude, our analyses have revealed that geminates, as produced in Japanese mimetic words, exhibit longer closure phases, extended gestural plateaus, higher TT positions, and more constricted postures. These articulatory profiles accord well with the acoustic properties of Japanese geminates, like their longer closure durations and lengthened preceding vowels. Additionally, our analyses also situated Japanese geminate production in a wider cross-linguistic context. Japanese geminates seem to be primarily produced by lengthening gestural plateaus compared to singletons, as demonstrated by our DTW analyses. However, even so, they are not simply extended versions of singletons: Some of their kinematic parameters are also different. These considerations lend plausibility to the proposal that even “canonical” geminates like those of Japanese and Italian are actually implemented by speakers using dimensions beyond duration, such as tighter constrictions and generally different kinematic profiles that have larger movements and lower stiffness. Finally, our results show that languages can differ substantially in the timing of geminates.⁵ Japanese geminates and singletons start around the same time with respect to the preceding vowels, yet geminates reach their targets later, allowing for longer acoustic vowel durations. This is unlike other languages where geminate production starts earlier with respect to the preceding vowel, in effect, shortening it.¹⁵

Supplementary Material

See the supplementary material for additional analyses and figures.

Acknowledgments

We thank Piyapath Srisomyos and Teerawee Sukanchanon for help with data annotation, as well as Jeff Moore, Lia Bučar Shigemori, and Ulrike Rupprecht for help with data acquisition. Data collection was supported by Japan Society for the Promotion of Science (JSPS) Grant Nos. 15F15715 to the second and third authors and 26770147 and 26284059 to the second author.

Author Declarations

Conflict of Interest

The authors have no conflicts to disclose.

Ethics Approval

The authors obtained ethical approval from the institutional review boards (IRBs) of Western Sydney University and Keio University (Protocol No. HREC 9482) and from the IRB of the University of Munich. Informed consent was obtained from all participants.

Data Availability

The data that support the findings of this study are openly available in an OSF repository at https://osf.io/27nyz/?view_only=4f93c383e24642e48d027c58fd945a27.

References

- ¹P. Ladefoged and I. Maddieson, *The Sounds of the World's Languages* (Blackwell, Oxford, UK, 1996), Vol. 1012.
- ²R. Ridouane, "Geminates at the junction of phonetics and phonology," in *Laboratory Phonology*, edited by C. Fougeron, B. Kuehnert, M. D'Imperio, and N. Vallee (De Gruyter Mouton, Berlin, 2010), Vol. 10, pp. 61–90.
- ³S. Kawahara, "The phonetics of sokuon, or geminate obstruents," in *Handbook of Japanese Phonetics and Phonology*, edited by H. Kobuzono (De Gruyter, Berlin, 2015), pp. 43–78.
- ⁴K. Idemaru and S. Guion-Anderson, "Relational timing in the production and perception of Japanese singleton and geminate stops," *Phonetica* **67**, 25–46 (2010).
- ⁵C. L. Smith, *The Timing of Vowel and Consonant Gestures* (Yale University, New Haven, CT, 1992).
- ⁶C. L. Smith, "Prosodic patterns in the coordination of vowel and consonant gestures," in *Phonology and Phonetic Evidence. Papers in Laboratory Phonology*, edited by B. Connell and A. Arvanati (Cambridge University Press, Cambridge, UK, 1995), Vol. IV, pp. 205–222.
- ⁷A. Löfqvist, "Lip kinematics in long and short stop and fricative consonants," *J. Acoust. Soc. Am.* **117**, 858–878 (2005).
- ⁸A. Löfqvist, "Articulatory coordination in long and short consonants: An effect of rhythm class?," in *The Phonetics and Phonology of Geminate Consonants*, edited by H. Kubozono (Oxford University Press, Oxford, UK, 2017), pp. 118–129.
- ⁹A. Löfqvist, "Tongue movement kinematics in long and short Japanese consonants," *J. Acoust. Soc. Am.* **122**, 512–518 (2007).
- ¹⁰M. Morimoto and T. Kitamura, "Articulation of geminated liquids in Japanese," in *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS 2019)*, Melbourne, Australia (International Phonetic Association, London, 2019), pp. 2811–2815.
- ¹¹M. Fujimoto, S. Funatsu, and P. Hoole, "Articulation of single and geminate consonants and its relation to the duration of the preceding vowel in Japanese," in *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*, Glasgow, Scotland (International Phonetic Association, London, 2015).
- ¹²A. Löfqvist, "Interarticulator programming: Effects of closure duration on lip and tongue coordination in Japanese," *J. Acoust. Soc. Am.* **120**, 2872–2883 (2006).
- ¹³S. Kawahara and F. M. Matsui, "Some aspects of Japanese consonant articulation: A preliminary EPG study," *ICU Work. Papers Linguist.* **2**, 9–20 (2017).
- ¹⁴A. Kochetov and Y. Kang, "Supralaryngeal implementation of length and laryngeal contrasts in Japanese and Korean," *Can. J. Linguist.* **62**, 18–55 (2017).
- ¹⁵F. Burroni, S. Maspong, N. Benker, P. Hoole, and J. Kirby, "Spatiotemporal features of bilabial geminate and singleton consonants in Italian," in *Proceedings of the 13th International Seminar on Speech Production (ISSP 2024)*, Autrans, France (ISCA, Stockholm, 2024).
- ¹⁶J. A. Shaw and S. Kawahara, "Effects of surprisal and entropy on vowel duration in Japanese," *Lang. Speech* **62**, 80–114 (2019).
- ¹⁷F. Tomaschek, B. V. Tucker, M. Wieling, and R. H. Baayen, "Vowel articulation affected by word frequency," in *Proceedings of the 10th International Seminar on Speech Production (ISSP)*, edited by S. Fuchs, M. Grice, A. Hermes, L. Lancia, and D. Mücke, Cologne, Germany (May 5–8, 2014), available at http://www.issp2014.uni-koeln.de/wp-content/uploads/2014/Proceedings_ISSP_revised.pdf.
- ¹⁸U. Cohen Priva, "Informativity affects consonant duration and deletion rates," *Lab. Phonol.* **6**, 243–278 (2015).
- ¹⁹F. Tomaschek, B. V. Tucker, M. Fasiolo, and R. H. Baayen, "Practice makes perfect: The consequences of lexical proficiency for articulation," *Linguist. Vanguard* **4**, 20170018 (2018).
- ²⁰S. Kawahara, "Emphatic gemination in Japanese mimetic words: A wug-test with auditory stimuli," *Lang. Sci.* **40**, 24–35 (2013).
- ²¹matlab, R2024B (MathWorks, Inc., Natick, MA, 2024).
- ²²P. Boersma and D. Weenink, "Praat: Doing phonetics by computer, URL (version 6.4.25) [computer program]," <https://www.fon.hum.uva.nl/praat/> (Last viewed 8/12/2024).
- ²³C. P. Browman and L. Goldstein, "Some notes on syllable structure in articulatory phonology," *Phonetica* **45**, 140–155 (1988).
- ²⁴J. A. Shaw and S. Kawahara, "The lingual articulation of devoiced /u/ in Tokyo Japanese," *J. Phonet.* **66**, 100–119 (2018).
- ²⁵D. Garcia, "Robust smoothing of gridded data in one and higher dimensions with missing values," *Comput. Stat. Data Anal.* **54**, 1167–1178 (2010).
- ²⁶F. Burroni and S. Tilsen, "The online effect of clash is durational lengthening, not prominence shift: Evidence from Italian," *J. Phonet.* **91**, 101124 (2022).
- ²⁷F. Burroni, R. Lau-Preechathammarach, and S. Maspong, "Unifying initial geminates and fortis consonants via laryngeal specification: Three case studies from Dunan, Pattani Malay, and Salentino," in *Proceedings of the Annual Meetings on Phonology (AMP 2021)*, Virtual Meeting (Linguistics Society of America, Washington, DC, 2021).