



**HAL**  
open science

## Articulation of geminate obstruents in the Ikema dialect of Miyako Ryukyuan: A real-time MRI analysis

Masako Fujimoto, Shigeo Shinohara, Daichi Mochihashi

► **To cite this version:**

Masako Fujimoto, Shigeo Shinohara, Daichi Mochihashi. Articulation of geminate obstruents in the Ikema dialect of Miyako Ryukyuan: A real-time MRI analysis. *Journal of the International Phonetic Association*, 2021, pp.1-25. hal-03096154

**HAL Id: hal-03096154**

**<https://hal.science/hal-03096154>**

Submitted on 4 Jan 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

(Final draft ver. Dec 2020 to appear in *Journal of International Phonetic Association*,  
Journal xx, page number xx-xx, and DOI xxxxx)

## **Articulation of geminate obstruents in the Ikema dialect of Miyako**

### **Ryukyuan: A real-time MRI analysis**

Masako Fujimoto<sup>1\*</sup>, Shigeko Shinohara<sup>2</sup>, Daichi Mochihashi<sup>3</sup>

<sup>1</sup>Advanced Research Center for Human Sciences, Waseda University

2-579-15, Mikajima, Tokorozawa, Saitama, 359-1192, Japan

<sup>2</sup>Laboratoire de Phonétique et Phonologie (UMR7018, Centre National de la  
Recherche Scientifique-Sorbonne Nouvelle), 19 rue des Bernardins, Paris,  
75005, France

<sup>3</sup>The Institute of Statistical Mathematics, Department of Statistical Inference  
and Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo, Japan

\*Corresponding Author:

Email: mfuji@viola.ocn.ne.jp, m.fujimoto5@kurenai.waseda.jp

Phone: +81 (0429)49-6848

Address: 2-579-15, Mikajima, Tokorozawa, Saitama, 359-1192, Japan

### **Abstract**

The Ikema dialect of Miyako Island in Okinawa, Japan, has typologically rare word-initial and voiced geminate obstruents (e.g., /vva/ ‘you’, /ffa/ ‘child’, /tta/ ‘tongue’, /badda/ ‘side’). These sounds are marked in two ways: Voicing through geminate obstruents is hard to produce and initial voiceless plosives seem to be difficult to perceive. This study investigated real-time magnetic resonance imaging (rt-MRI) to examine the

articulatory settings underlying contrasts between singleton and geminate obstruents. Our analyses of two male speakers' utterances showed the following characteristics. 1) Geminate obstruents in Ikema have longer duration of articulatory constrictions regardless of position and consonant types. 2) The voiced alveolar plosive geminate /dd/ is articulated with a larger linguopalatal contact than its singleton counterpart but such difference depends on the speaker for the voiceless plosive pair /tt/-/t/ and the fricative pairs /ss/-/s/ and /zz/-/z/. 3) Alveolar voiceless plosives /t/ and /tt/ have a greater degree of linguopalatal contact than their voiced counterparts /d/ and /dd/, respectively, but fricatives show inter-speaker variation. 4) Fricatives do not show any systematic difference in degree of (midsagittal) linguopalatal contact between geminates and singletons, or between voiceless and voiced consonants. 5) Voiced geminate obstruents are accompanied by pharyngeal expansion for both speakers and by lowering the larynx for one speaker, and never by lowering of the velum. We also observed that voiced fricatives tend to realize as affricates, which we interpret as part of the articulatory adjustments for (full) voicing of phonologically voiced geminate fricatives.

### **Keywords**

Southern Ryukyuan, real-time MRI, voiced geminates, initial geminates, articulatory settings

### **1. Introduction**

Miyako Ryukyuan is one of the Ryukyuan languages spoken in the Ryukyu Islands and belongs to the Japonic family, along with Japanese. Dialects of Miyako Island in southern Ryukyu were designated by UNESCO as endangered (Moseley ed. 2009). Along with other Miyako Ryukyuan dialects, the Ikema dialect has typologically rare word-initial and voiced geminate obstruents. This study examined articulatory strategies employed by speakers of the dialect to maintain the geminate-singleton contrast in voiced obstruents (e.g., /bada/ 'gutter' vs. /badda/ 'side', /zI/ 'soil' vs. /zzu/ 'fish'). We also

addressed the question of how, if at all, articulatory settings differ between word-initial voiceless singletons and their geminate counterparts (e.g., /ta/ ‘rice field’ vs. /tta/ ‘tongue’). The paper reports qualitative and quantitative analyses of the utterances of two male speakers of the Ikema dialect of Miyako Ryukyuan, using real-time magnetic resonance imaging (rt-MRI).

Geminate consonants are phonologically longer consonants, and their phonological length is acoustically manifested by a longer duration when compared with singletons (e.g., Ladefoged & Maddieson 1996, Ham 2001). Initial geminates are cross-linguistically less common than word-medial geminates (Thurgood 1993, Topintzi & Davis 2017) and have been an object of phonological (e.g., Hume, Muller & van Engelenhoven 1997, Davis 1999, Topintzi 2008) and phonetic studies (e.g., Abramson 1986 et seq., Arvaniti & Tserdanelis 2000, Kraehenmann 2001, Muller 2001). Our previous acoustic study involving five speakers of the Ikema dialect observed longer frication noise for geminates than for singletons in word-initial position, as well as longer closure durations of word-medial /dd/ and /tts/ (Shinohara & Fujimoto 2018). However, closure durations of word-initial plosives could not be measured in phrase-initial position, which served as one of the motivations for this articulatory study. The acoustic cues of the length difference of voiceless geminates, particularly voiceless plosive ones, may be difficult to perceive even by native speakers (Kraehenmann 2001, Ridouane & Hallé 2017).

In the Ikema dialect, initial geminates occur in words such as /tta/ ‘tongue,’ /ssa/ ‘grass,’ /zza/ ‘father,’ /ffa/ ‘child,’ and /vva/ ‘you.’ The length of voiceless geminates, such as /tta/, might barely be recognized as long when uttered phrase-initially. Secondary cues may also be present as in Pattani Malay, wherein multiple perceptual cues were identified

(Abramson 1991; 1999b; 2003). Notably, the voice onset time (VOT) of the geminate /tt/ was observed to be shorter than that of its singleton counterpart /t/ in Ikema (Shinohara & Fujimoto 2018). This difference may be perceived as a cue to the singleton–geminate distinction for /t/. Although this particular aspect of articulatory events occurring between an onset consonant and the following vowel is outside the scope of this study, articulatory durations and degrees of linguopalatal contact are compared between singleton and geminate consonants by using rt-MRI (see the remainder of the Introduction and Section 3.3 for the advantages and limitations of rt-MRI).

Regarding the voicing, place, and manner of consonants, there are universal tendencies wherein certain types occur as geminates more frequently than others. In this regard, Jaeger (1978) predicted an implicational relation between voiceless and voiced obstruent geminates, whereby only languages with voiceless geminate obstruents could have voiced geminate obstruents. A survey of the Stanford Phonology Archive comprising 221 languages supports his prediction. Jaeger reported that nasals are more prone to be geminates (67 out of the 72 languages with some sort of geminates) and that, among obstruents, voiced geminates are less widespread than voiceless ones. The study provided a list of 16 languages that had voiceless obstruent geminates and ones that were not voiced, whereas only two languages (Somali and Island Carib) were observed to have a skew toward voiced obstruent geminates. Maddieson (1984) also reported an asymmetry among 16 languages possessing geminate obstruents. The languages that had voiceless geminates were twice as many as those that had voiced ones. Finnish, Maranungku, and Cypriot Greek are examples of languages that have only voiceless obstruents or sonorant geminates. This type of asymmetry led Kirchner (2000) to view the voiced geminate obstruents as marked when compared with voiceless ones. Jaeger

provided the following articulatory reason for the asymmetry:

a stop closure of long duration will allow air pressure in the oral cavity enough time to equalize with sub-glottal pressure and cause voicing to stop; this is also true of the narrow constriction for fricatives, but not true of nasals (Jaeger 1978: 320).

This quote implies that greater effort is required to produce voiced obstruents than that required to produce voiceless ones. Consequently, voiced obstruents are likely to become more or less devoiced (Jaeger 1978, Ohala 1983; 1997), and geminates are even more likely to become devoiced than singletons (Ohala 1997). Hayes and Steriade (2004) also claimed that such aerodynamic constraints are reflected in phonological patterns. They established a markedness scale concerning place and length of voiced plosives (\*gg >> \*dd >> \*bb >> \*g >> \*d >> \*b). Regarding the dimension of the manner of articulation, fricatives tend to be more devoiced than stops (Ohala 1997). This phenomenon may support the additional markedness by Kirchner (2000) for fricative geminates when compared with stop and affricate geminates, which was proposed on the basis of the typological data in Maddieson (1984).

Focusing on geminate obstruents, how are the aforementioned universal trends reflected in Japonic languages? In standard Japanese, voiced obstruent geminates are observed only in recent borrowings such as /beddo/ ‘bed’ or /baggu/ ‘bag’ and in onomatopoeia. Acoustic analyses clearly indicated that those voiced geminates are partially or fully devoiced (Kawahara 2006, Matsuura 2012, Fujimoto & Funatsu 2018, Hussain & Shinohara 2019). However, acoustic analyses of voiced geminates in the dialects of the Kumamoto and Amakusa areas of Kyushu, Japan, showed that the amount of voicing during the closure phase varied, depending on words, speakers, sub-dialects,

and tokens, indicating the instability of voiced geminates in those dialects (Matsuura 2016, Takada 2018). Voiced geminates are used in everyday words in the Ikema dialect both word-medially, such as in /badda/ ‘one’s side’ and /kuvva/ ‘calf,’ and word-initially, such as in /zza/ ‘father’ and /vva/ ‘you’<sup>i</sup>. Acoustic investigations of the dialect have indicated that voicing was maintained throughout word-medial voiced geminates (Matsuura 2012, Fujimoto & Shinohara 2013, Shinohara & Fujimoto 2018) and word-initial geminates were accompanied by a long-voiced period (Fujimoto & Shinohara 2013, Shinohara & Fujimoto 2018).

We investigated how voicing is maintained throughout the geminates by using rt-MRI, which is a noninvasive technique to examine the settings of articulators and articulators’ transitions in the vocal tract, including the laryngeal area, observations of which are not possible by other methods such as electromagnetic articulography (EMA) or electropalatography (EPG). As rt-MRI allows dynamic recordings of many words and repetitions within a short recording time, it is ideal for analyzing co-articulation in running speech, something that cannot be done with static MRI (Engwall 2006). Studies using static and real-time MRI include qualitative observations of the overall vocal tract (e.g., Ramanarayanan et al. 2013) as well as quantitative analyses of specific regions of interest (e.g., Lammert et al. 2013) and contour tracking (e.g., Bresch et al. 2008, Bresch & Narayanan 2009) (see Ramanarayanan et al. [2018] for a detailed review of rt-MRI studies). Linguistic investigations using MRI include (Kim, Honda & Maeda 2005, Kim Maeda & Honda 2010) on Korean obstruents; Martins et al. (2008) on the vowels and consonants of European Portuguese; Proctor, Shadle & Iskarous (2010) on American English; Iribar et al. (2018) on Basque sonorants; and Fujimoto & Shinohara (2018a) on geminate plosives of the Japanese Kumamoto dialect.

In our preliminary observation of rt-MRI data of Ikema Ryukyuan, 8 words uttered by a single speaker (Fujimoto & Shinohara 2015) and 17 words for voicing contrast with two speakers (one being common) (Fujimoto & Shinohara 2017) were analyzed solely by visual inspection. In this paper, we used (near-) minimal pair words uttered by two speakers to conduct systematic comparisons, more detailed observations, and quantitative analyses of two points of interest, namely, the size of linguopalatal constriction length and pharyngeal aperture on the midsagittal plane, and the frame count for the comparison of consonant duration.

In the following sections, we first describe relevant articulatory studies to make predictions in Section 2. Methods are described in Section 3, Section 4 presents the results, and Section 5 provides the discussion.

## **2. Articulatory studies of obstruents and predictions**

In our previous study of Ikema, acoustic signals of geminate plosives in word-medial position and geminate fricatives in the initial and medial positions presented a longer duration (Shinohara & Fujimoto 2018). Articulatory studies in standard Japanese have observed that oral constriction is longer in duration for (word-medial) voiceless geminates than for their singleton counterparts (Takada 1985, Löfqvist 2007, Kochetov & Kang 2017, Kawahara & Matsui 2017). The same tendency was observed in Italian word-medial geminates (Gili-Fivela et al. 2007) and in both word-initial and word-medial geminates in Swiss German (Kraehenmann & Lahiri 2008) and Tashlhiyt Berber (Ridouane 2010, Ridouane & Hallé 2017). Thus, we predicted that geminates in the Ikema dialect would have a longer duration in oral constriction both word-initially and word-medially.



Linguopalatal contact has also been reported to be larger in area in geminates than in singletons in Japanese (Kochetov & Kang 2017, Kawahara & Matsui 2017); Italian (Payne 2006, Hagedorn, Proctor & Goldstein 2011); and Tashlhiyt Berber (Ridouane 2010, Ridouane & Hallé 2017). The larger contact area is viewed as “stronger” articulation in geminates. This stronger linguopalatal contact in geminates is reported for phrase-initial voiceless plosives in Swiss German (Kraehenmann & Lahiri 2008) and for utterance-initial voiceless plosives in Tashlhiyt Berber (Ridouane 2007; 2010, Ridouane & Hallé 2017)<sup>ii</sup>. Hence, a larger size of oral constriction could also appear in geminates of Ikema speakers both word-initially and word-medially.

Few articulatory studies on voiced geminates in standard Japanese have been conducted because they only occur in recent loanwords (and onomatopoeia). Kawahara and Matsui’s (2017) EPG study using mimetics is an exception and showed that linguopalatal contact is greater in (word-medial) /dd/ than in /d/ and so is in /tt/ than in /t/. Another study observed that the oral airflow patterns are similar between voiced and voiceless geminates and that this may result in voiced geminates being partially devoiced in standard Japanese (Fujimoto & Funatsu 2018). According to our review of the literature, no articulatory studies have focused on (initial and medial) voiced geminates of the Miyako dialects (except for our preliminary reports [Fujimoto & Shinohara 2015; 2017]). Thus, we investigated whether voiceless and voiced geminate obstruents are articulated in a distinct fashion in Ikema.

Within the singleton category, the size of the linguopalatal constriction area was smaller in voiced obstruents than in voiceless ones in German (Fuchs et al. 2006) and Berber (Ridouane 2007) although Dagenais et al. (1994) observed the opposite pattern in American English, i.e., the linguopalatal contact is greater for voiced stops than for their

voiceless counterparts. Studies of standard Japanese have generally agreed that the linguopalatal constriction area is smaller in voiced singleton plosives than in their voiceless counterparts (National Institute for Japanese Language and Linguistics 1990, Kitamura & Hatano 2012<sup>iii</sup>, Kochetov 2014). Hence, one can speculate that the size of linguopalatal constriction is smaller in voiced consonants than in voiceless ones in Ikema speakers.

Regarding linguopalatal constriction, an asymmetry has been found between plosives and fricatives. Within singleton obstruents in German, the target of the closing gesture is lateral in the case of the fricative /z/, but it is on the midsagittal plane in that of /t/ (Fuchs et al. 2006). In Japanese, linguopalatal contact is greater in voiceless plosives than in voiced ones, and it is greater in voiced fricatives than in voiceless ones (Kochetov 2014). Additionally, linguopalatal constriction has been observed to be greater for geminates than for singletons in plosives, but fricatives do not share this difference in standard Japanese (Kawahara & Matsui 2017). This phenomenon may occur because the size of lateral constriction, or the central groove, plays a critical role for fricatives when compared with midsagittal contact in stops. If so, the asymmetry between plosives and fricatives might also appear in Ikema speakers' utterances.

In addition to weaker linguopalatal constriction, pharyngeal cavity expansion, nasal leakage, and larynx lowering have been reported as articulatory correlates for voicing. Rothenberg (1986) argued for the necessity of other mechanisms that absorb the glottal airflow to maintain vocal fold vibration when the vocal cavity is obstructed, i.e., 1) a passive, pressure-actuated expansion of one or more of the walls of the supraglottal cavity; 2) a muscularly activated enlargement of the supraglottal cavity; and 3) some nasal airflow through an incomplete velopharyngeal closure. Westbury (1983) showed that in

American English, the tongue root is advanced more during the production of voiced stops when compared with their voiceless counterparts, suggesting an increase of pharyngeal volume. Similarly, pharyngeal expansion has been reported to occur during singleton voiced obstruents (Kent & Moll 1969, Perkell 1969, Ohala & Riordan 1979). In the Kumamoto dialect of Japanese, pharyngeal distance was observed to be longer during (word-medial) voiced geminate plosives when compared with their voiceless counterparts (Fujimoto & Shinohara 2018a). Hence, pharyngeal expansion during production of voiced obstruents seems to be a cross-linguistically common maneuver to maintain the voicing during geminate plosives. Pharyngeal expansion may also be observed during voiced consonants in Ikema speakers' utterances.

Larynx lowering has also been reported during the production of voiced consonants in Kent and Moll (1969), Perkell (1969), and Ewan and Kronen (1974) in American English. This gesture can also enlarge the volume of the pharyngeal cavity. Larynx lowering may occur during voiced consonants in Ikema speakers' utterances. Nasal leakage during voiced obstruents has been reported in Spanish, English, and French (Solé, Sprouse & Ohala 2011, Solé 2018), and Russian (Kharlamov 2018). It was also observed in voiced plosives of non-words used by speakers of the Kansai dialect of Japanese (Hirata et al. 2002). If a large amount of nasal leakage occurs, velum lowering may appear during voiced consonant production in Ikema speakers.

Based on the results in the literature, this study investigated whether any of the maneuvers facilitating length and voicing distinctions in articulation could be observed in the rt-MRI data of Ikema consonants. We predicted that 1) oral constriction would be longer in duration for geminates than it would be for singletons regardless of consonants; 2) the size of the oral constriction would be larger in geminates than that in singletons; 3)

the size of oral constriction would be larger for voiceless consonants than it would be for voiced ones; 4) larger constriction in geminates than that in singletons and/or in voiceless consonants than that in voiced ones, if at all, would appear more significantly in plosives than it would in fricatives; and 5) the size of the pharyngeal aperture would be larger for voiced obstruents than that for voiceless counterparts, and the difference is more evident for geminates. We also examined whether a lowering of the larynx and/or velum occurs.

### 3. Methods

#### 3.1 Speakers

Two male speakers of Ikema dialect (M1 and M2) participated in this study. They were raised in Nishihara District, Miyako Island. M1 was born and grew up on Miyako Island; studied (18–22 yrs) and worked (24–57 yrs) in Naha on the main Okinawa Island; and then returned to Miyako Island. M2 spent his entire life on Miyako Island. M1 and M2 were aged 67 and 61 years, respectively, at rt-MRI recording (in 2014).

#### 3.2 Test words

According to Pellard and Hayashi (2012), the consonant inventory of Ikema is the following: /p, b, t, d, k, g, ts, s, z, f, v, h, m, n, r, j, w, ŋ/. Among the obstruents, the geminates /tt, tts, ff, ss, vv, zz/ appear in initial and medial positions (/tta/ ‘tongue,’ /uttu/ ‘husband,’ /ttsutsI/ [ttɛutsI] (see below for the quality of /I/) ‘cicada,’ /attsɑ/ [attea], /ffa/ ‘child,’ /maffa/ ‘pillow,’ /ssa/ ‘grass,’ /ussa/ ‘cow.TOP,’ /vva/ ‘you,’ /kuvva/ ‘calf,’ /zzu/ ‘fish,’ /tuzza/ ‘wife.TOP’), whereas the geminate /dd/ occurs only in word-medial position (/baddɑ/ ‘side’). In addition to /a, i, u/, another unrounded close (often referred to as central) vowel is used in Ikema. This vowel has been variably transcribed in the literature

as /i/ /ɨ/, or /i/ (see references in Fujimoto & Shinohara [2018b]). An rt-MRI study of Fujimoto and Shinohara (2018b) reported that its primary constriction is at the alveolar ridge with the front of the tongue; acoustically, the second formant (F2) lies roughly in the area between the F2 of [i] and that of [u]. In this paper, we use /I/ or [I] (with a normal size uppercase letter) to denote the vowel. The vowels have long and short length contrasts, except for the underlyingly monomoraic words. As in many Ryukyuan dialects, Ikema has a word minimality constraint wherein a word must have at least two-mora length on the surface. Due to this constraint, monomoraic words, such as /ta/ ‘rice field’ or /zI/ ‘soil,’ are pronounced with lengthened vowels [taa] or [zII], respectively (double segment symbols represent a two-mora length here). As aforementioned, vowel length also contrasts elsewhere. Thus, pairs such as /ti/ [tii] ‘hand’ and /tii/ [ttii] ‘pipe’ are possible; however, there is no contrast such as in [tta] vs. [ta] or in [ta] vs. [taa]. We selected as much as possible singleton–geminate pairs with a similar vocalic environment, such as [tta]-[taa] and [ssa]-[sata].

Table 1 presents the test words used in this study. Glosses are provided between single quotations. The test words are all existing words and consist of those involving voiced geminates /dd, zz, vv/ and those starting with /tt, ss, zz, ff, vv/.<sup>iv</sup> Their singleton counterparts are /t, d, s, z, f/. Note that there is no word containing a singleton /v/ or beginning with /dd/. Words were selected so that they formed singleton–geminate minimal pairs; however, near-minimal pairs were used when not possible.

**Table 1** Test words and glosses sorted by singleton and geminate consonants

Consonants	Singletons	Geminates
t	ta ‘rice field’	tta ‘tongue’
d	nada ‘tear’	badda ‘side’
s	sata ‘sugar’	ssa ‘grass’

z	zI 'soil'	zza 'father'
f	fau 'to eat'	ffa 'child'
v	N/A	vva 'you'

### 3.3 Method of real-time MRI

Real-time MRI acquisitions were performed at the Brain Activity Imaging Center at ATR-promotions (ATR-BAIC), Kyoto, Japan, with a 3 Tesla Siemens MRI System (MAGNETOM Verio) equipped with three types of coils. These three coils, a small flex coil, head coil (posterior), and neck coil (posterior), were combined with a specially made coil holder. To obtain 50 sec dynamic images, a 500 times continuous, fast low-angle shot sequence was employed. The equipment had the following characteristics: field of view = 256 mm × 256 mm; size of acquisition matrix = 87 × 128, interpolated to 256 × 256; pixel size (after interpolation) = 1.0 mm × 1.0 mm; slice thickness = 10 mm; number of slices = 1; TR = 3.4 ms; TE = 1.38 ms; flip angle = 8°; and parallel acquisition factor = 3. Therefore, each frame acquisition time was 98.6 ms, and the frame rate was 10.14 fps.

The midsagittal plane of the speaker's head and neck was determined for the plane to cut the center of the vocal tract observed in the coronal plane and to cut the center of the tongue in the transverse plane of the 3D pilot scanning. This process was performed for each speaker. Next, midsagittal images were sampled continuously. Recall that this frame rate does not allow us to determine the exact timing of very short events, such as a release in the oral cavity. Consequently, the timing of the maximum constriction of consonants may not have always been captured within an image. Another limitation of MRI is that it cannot image the organs that do not contain water such as teeth and bones.

Thus, some oral contact information for /t, d/ may be missing because the portions of linguodental contact were not shown in the images. In addition, the degree of lateral constriction for fricatives could not be observed by a midsagittal slice. Nevertheless, the articulators' overall shapes during the production of phones, along with their dynamic movements, could be observed when several images for the same consonants were referred to.

We prepared 16 sets of word lists for a recording session, comprising 3–6 words. These lists included the words used for other analyses as well as the 11 words used in this analysis shown in Table 1. During the recording time of 50 seconds, speakers uttered the words in a list repeatedly at a self-regulated, comfortable tempo. Each test word was recorded 6–12 times, depending on the number of words in a list, word length, and speech tempo. We discarded the first trials of M1 because the speaker sometimes hesitated at the beginning. One token of /tta/ by M2 was discarded because the speaker put a long pause (for about 1 second) before the release of /tt/. Six tokens for each word per speaker were used in this study.

The speech sound was simultaneously recorded by using an Optoacoustics surveillance optical microphone. A spectral subtraction method was applied to remove the MR's scanning noise as follows. First, fast Fourier transform (FFT) was performed for the period including only scanning noise (FFT1). Second, FFT was performed for the period of the recorded experimental session that included the speaker's voice and scanning noise (FFT2). Third, spectral subtraction (FFT2-FFT1) and inverse FFT were performed. (Sampling frequency = 44.1 kHz, Window size of FFT = 10000, and Window shift = 5000). The rt-MR images and the speech sound were integrated into movies. The speech sound somewhat deteriorated because of the noise-canceling process; nevertheless,

it had enough quality for reliable spectral analyses of the sounds because the (noise-canceling) treatment did not affect any specific frequency ranges of the speech sound. The sound treatment and synchronization were also conducted by ATR-BAIC.

### 3.4 Methods of analyses

We identified images corresponding to target phones by visual inspection. Figure 1 presents four consecutive frames of the word /basa/ [basa] ‘Japanese banana plant’ produced by M1. In this example, one frame corresponds to one phone. However, the number of frames corresponding to one phone differs depending on the words, tokens, and speakers.

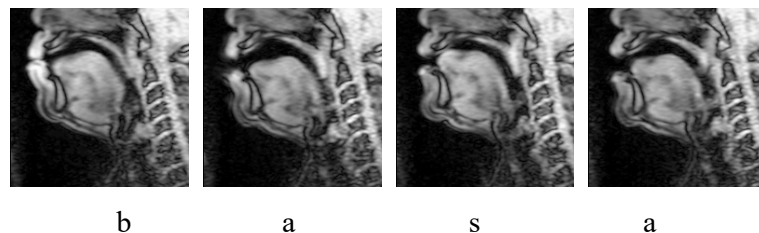


Figure 1 Example of representative frames with phones of Speaker M1 uttering /basa/.

To explain the durational difference in the articulation of singleton and geminate obstruents, we counted the number of frames of the target consonants. The relatively slow frame rate allows only a coarse comparison. Although observing the difference in milliseconds between the two is impossible, it can be an efficient means to confirm a duration difference in the silent closure duration of initial plosive segments, which is not possible by performing an acoustic analysis. In this evaluation, frames for each consonant were counted from the first visible contact through the release to a following vowel for plosives and from the first visible beginning of the constriction with voicing or frication



noise through release to a following vowel for fricatives.

As part of our qualitative analyses, we observed shapes of articulators and the size of constriction between articulators during the production of each phone with special reference to singleton and geminate and voiced and voiceless contrasts. In our preliminary studies (Fujimoto and Shinohara 2015; 2017), we observed characteristic differences in the size of the linguopalatal constricted area and the pharyngeal aperture between these contrasts. This finding called for measuring the size of the linguopalatal constriction and pharyngeal aperture. Thus, we conducted the following two quantitative analyses of the images.

Linguopalatal constriction was measured for coronal obstruents, /*(t)t*/, /*(d)d*/, /*(s)s*/, and /*(z)z*/ (parenthesized segments within slashes indicate optionality, e.g., /*(t)t*/ reads as “either /*t*/ or /*tt*/”). We measured the length of a straight line drawn on the flat part of the tongue as shown in the left panel in Figure 2. As mentioned in Section 3.3, the size of the linguopalatal contact may be underestimated for plosives because the (presumably existing) invisible linguodental contact was not included. Some studies such as those by Kim et al. (2005) and Labrunie et al. (2018) have estimated the shape and location of incisors from the remainder of the (MRI) information. However, such estimation could not be achieved in this study, mainly because of the limited resolution of the images.

Pharyngeal aperture was measured for all consonants (/ *(t)t*/, /*(d)d*/, /*(s)s*/, /*(z)z*/, and /*vv*/). In the studies on a similar effect of voicing, variability has been observed regarding the level of the pharynx measured. In Martins et al. (2008), the upper part of the pharynx was observed to be more expanded in (singleton) voiced consonants when compared with their voiceless counterparts. In Westbury’s (1983) study, on the other hand,

the lower part of the pharynx expanded more than the upper part. In our observation of geminates, the lower part of the pharynx showed the greatest displacement. Therefore, we measured in the lower part of the pharynx, more precisely, on the horizontal line crossing the bottom of the third cervical spine, in the area between the two vertical lines in the right panel of Figure 2. The measurements of linguopalatal constriction and the pharyngeal aperture were conducted manually by counting the pixels where 1 pixel corresponds to 1 mm. Premier Elements 10, Photoshop 10 by Adobe, and Image J 1.51 (Schneider, Rasband & Eliceiri 2012) were used for these analyses.

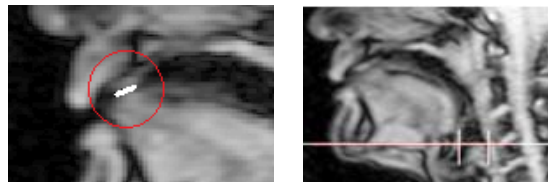


Figure 2 Representative place in the measurements of linguopalatal constriction (left figure) and of the pharyngeal aperture (right figure). The dotted line in the circle on the left figure indicates the extent of linguopalatal constriction. The distance between the two vertical lines in the right figure indicates the size of the pharyngeal aperture on the horizontal line.

For the statistical evaluation of frame counts between singletons and geminates, we performed a generalized linear model (GLM) using Poisson distribution (Crawley 2014). This method was used because the frame counts were nonparametric and expected to be a small integer number with non-normal distribution. For the evaluation of the linguopalatal constriction and pharyngeal aperture, we performed a three-way analysis of variance (ANOVA), with Geminant, Voice, and Manner as independent variables. We used ANOVA due to the limited number of speakers, tokens, and factor levels. Speakers were treated separately to observe individual variation. Given that GLM and ANOVA were performed with a single speaker, the possibility of type I error increased because

individual tokens were used as an experimental unit. To compensate potential overestimation of the results, we set the alpha value at 0.01 (i.e., significance level  $p = .01$ ), smaller than the conventional 0.05 (Grafen & Hails 2002). SPSS version 22 (IBM) was used for the statistical analyses. For the supplementary acoustic materials presented in Section 4.2.3, Praat 6.0.21 (Boersma & Weenink 2016) was employed.

## 4. Results

### 4.1 Length distinction of singletons versus geminates

Table 2 shows the number of frames of each word uttered by each speaker averaged over six tokens, confirming that the frame count was greater in geminates than in singletons regardless of consonants and speakers. No differences were observed in the two speakers' patterns of frame count ( $t(130) = .112, p = .911$ ). Table 3 shows the result of GLM using Poisson distribution for 60 tokens (10 consonants (5 singleton + 5 geminate)  $\times$  6 tokens) separated by the speaker, with Frame Count as the dependent variable, and with Gemination (singleton or geminate), Voice (voiced or voiceless), and Manner (plosive or fricative) as independent variables. The consonant /vv/ was excluded from the analysis as the language lacks singleton counterpart /v/ in its phonotactics. In this analysis, only Gemination was marginally significant ( $p = .018$ ) for M1 and significant ( $p < .01$ ) for M2. Voice, Manner, and all the interactions were not. Frame Count averaged over consonants was greater for geminates (2.30 [standard deviation (SD) .466] for M1 and 2.33 [SD .711] for M2) than for singletons (1.40 [SD .498] for M1 and 1.27 [SD .450] for M2). These results indicate that geminate articulation takes longer time than singleton articulation. Thus, our first hypothesis "oral constriction is longer in duration for geminates than for singletons" is substantiated. To conclude, geminates'

articulation is longer in duration in word-initial position in the Ikema dialect. Notably, the difference in duration of the silent closure phase of utterance-initial /tt/ and /t/, which was difficult to see acoustically, became clear.

Table 2 Frame counts and SD of the consonants for two speakers averaged over six tokens. “C” denotes singletons and “CC” denotes geminates.<sup>v</sup>

Consonant	M1		M2	
	C	CC	C	CC
t	1.5 (.55)	2.7 (.52)	1.8(.41)	3.3 (.52)
d	1.0 (0.0)	2.0 (0.0)	1.0 (0.0)	2.8 (.41)
s	1.7 (.52)	2.8 (.41)	1.5 (.55)	2.3 (.52)
z	1.7 (.52)	2.0 (.00)	1.0 (.00)	2.2 (.41)
f	1.2 (.41)	2.0 (0.0)	1.0 (.00)	1.7 (.52)
v	---	2.5 (.55)	---	3.2 (.41)

Table 3 Results of GLM using the Poisson distribution of Frame Count as the dependent variable and Gemination, Voice, and Manner as independent variables. Data of the two speakers were treated separately. Asterisk denotes the significance codes (.01 \*).

	M1			M2		
	Walt Chi Square	df	Sig.	Walt Chi Square	df	Sig.
(Intercept)	29.100	1	.000	24.791	1	.000
Gemination	5.626	1	.018	9.280	1	.002 *
Voice	.740	1	.390	1.085	1	.298
Manner	.180	1	.671	1.891	1	.169
Gemination × Voice	.078	1	.780	.022	1	.882
Gemination × Manner	.634	1	.426	1.085	1	.298
Voice × Manner	.435	1	.509	.311	1	.577
Gemination × Voice × Manner	.315	1	.575	.022	1	.882

## 4.2. Qualitative observation of voiced and voiceless geminate obstruents

The articulatory variability of the target consonants was observed to be small among the tokens of the same test words within the speakers. Hence, we show images of the representative token of each word for illustration purposes. Unless explained otherwise, we present one frame for singletons and two successive frames for geminates that immediately precede a vowel.

### 4.2.1 Plosives: /t, tt, d, dd/

Figure 3 shows the frames of /t/ in /ta/ ([taa]) and /tt/ in /tta/ [tta]. For M1, tongue shape and the size of the linguopalatal contact area clearly differed between /t/ and /tt/. Notably, while only the tongue tip was raised toward the palate during /t/, the tongue body was also raised during /tt/. Of the two frames of /tt/, linguopalatal contact length was often greater in the second frame. For M2, the size of the linguopalatal contact was similar between /t/ and two frames of /tt/. Figure 4 shows the frames of /d/ in /nada/ and those of /dd/ in /badda/. For both speakers, the size of the linguopalatal contact was larger during /dd/ than during /d/. Often, contact size was greater in the second frame of /dd/. Because these consonants were both preceded and followed by the same vowel /a/, the difference must be a result of the singleton versus geminate contrast. When we compared the voiceless and voiced pair, /t/-/d/ and /tt/-/dd/, the size (of linguopalatal contact) was larger for /t/ and /tt/ than for /d/ and /dd/, respectively, for both speakers. Thus, linguopalatal contact was observed to be larger in voiceless plosives than in the voiced counterparts.

When we compared images of geminates /tt/ with /dd/, a difference was noticed in the pharyngeal area. The pharyngeal distance was observed to be greater for /dd/ than for /tt/, especially in the lower part around the tongue root; that is to say, the pharynx expanded during the voiced /dd/. In the two frames of /dd/, pharyngeal distance was

observed to be generally greater in the second frame. Hence, pharynx expansion must have occurred after linguopalatal contact. By contrast, pharyngeal expansion was not observed during /tt/ for either speaker. Thus, we posit that the articulation of /dd/ is accompanied by significant pharyngeal expansion. The larynx was slightly lower during /dd/ than during /tt/ for M1, but this was not the case for M2. Thus, during /dd/, pharyngeal expansion and laryngeal lowering were observed for M1, but only pharyngeal expansion was observed for M2. Velum lowering, another possible articulatory correlate for voicing, was not observed during /(d)d/ for either speaker.

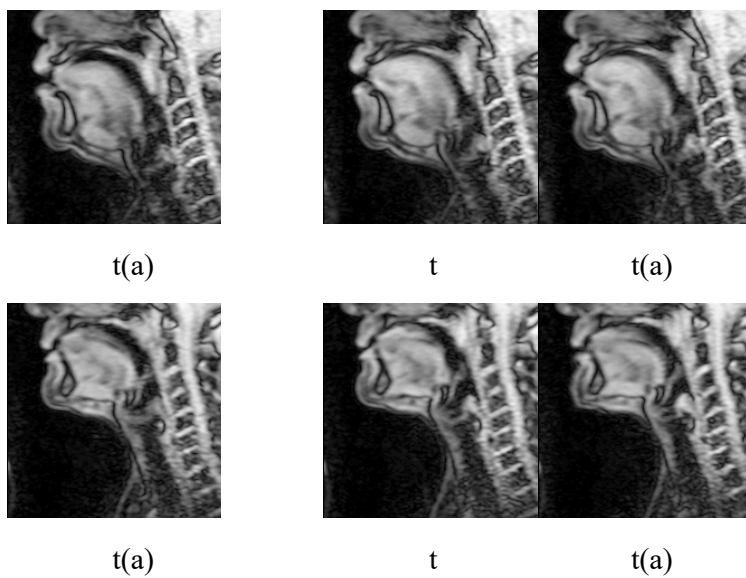
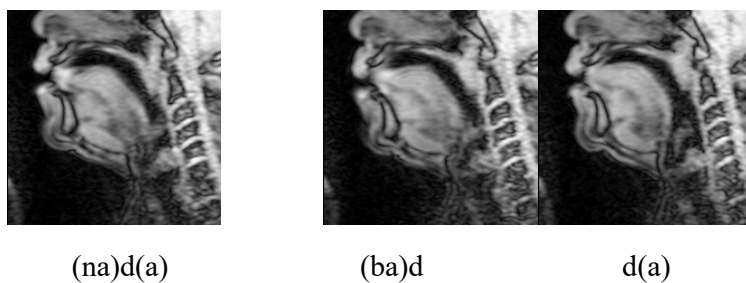


Figure 3 Frames of /t/ in /ta/ (left) and /tt/ in /tta/ (middle and right). M1 (upper), M2 (lower).



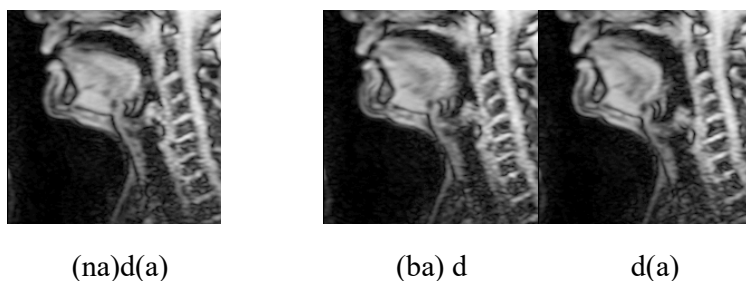


Figure 4 Frames of /d/ in /nada/ (left) and two frames of /dd/ in /badda/ (middle and right). M1 (upper), M2 (lower).

#### 4.2.2 Fricatives: /s, ss, z, zz, f, ff, vv/

Figure 5 shows the frames of /s/ in /sata/ and /ss/ in /ssa/.<sup>vi</sup> The size of the linguopalatal constriction area and the shape of the vocal tract were similar for /s/ and /ss/, except that M2 had greater constriction than M1. Figure 6 presents the frames of /z/ in /zI/ and /zz/ in /zza/. The size of the linguopalatal constriction area was again similar across /z/ and two frames of /zz/ for both speakers. When we compared the voiceless and voiced pairs, /s/-/z/ and /ss/-/zz/, the size of the linguopalatal constriction area along the palate was similar between the two pairs. However, the degree of aperture between the tongue and the palate differed considerably between the voiceless and the voiced fricatives. Figure 7 compares the enlarged images of the alveolar area of the first frame of /ss/ and /zz/ and reveals that complete linguopalatal closure was formed at the midsagittal plane for /zz/ but not for /ss/ (relevant areas are circled in the figures). This finding suggests that /zz/ might have been realized as an affricate, as reported in the acoustic analysis (Fujimoto & Shinohara 2017).

The pharyngeal distance was greater during /zz/ than during /ss/. Thus, the pharynx expanded during /zz/. The degree of expansion was often greater in the second frame of /zz/, and this observation holds true for both speakers. The larynx was slightly

lower during /zz/ than during /ss/ for M1 but not for M2. Thus, as in the case of alveolar plosives, during /zz/, pharyngeal expansion and laryngeal lowering were observed in M1, but only pharyngeal expansion was observed in M2. Velum lowering during /(z)z/ was not observed in either speaker.

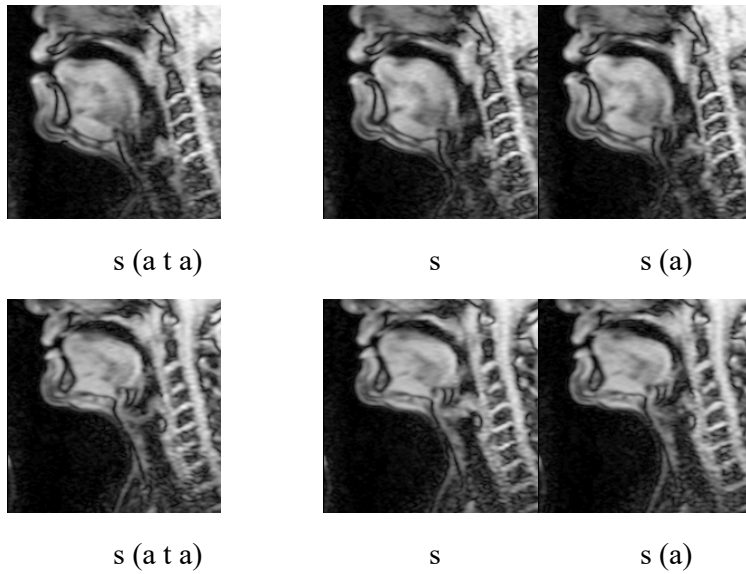


Figure 5 Frames of /s/ in /sata/ (left) and /ss/ in /ssa/ (middle and right). M1 (upper), M2 (lower).

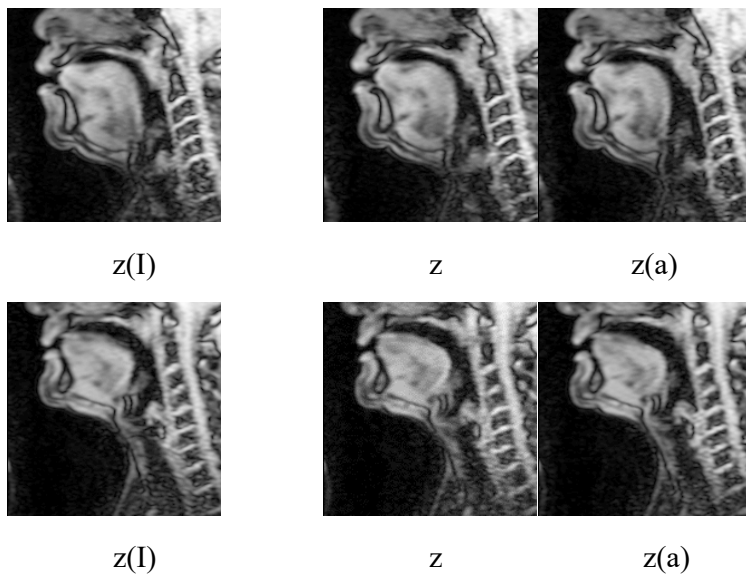


Figure 6 Frames of /z/ in /zI/ (left) and /zz/ in /zza/ (middle and right). M1 (upper), M2 (lower).



(lower).

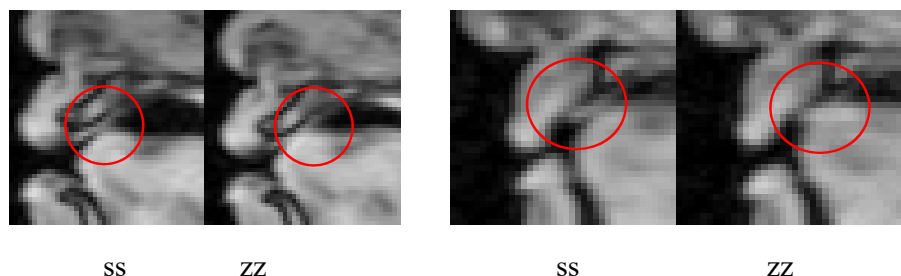


Figure 7 Comparisons of the first frames of /ss/ and /zz/. M1 (left two frames), M2 (right two frames).

Figure 8 shows the frames of /f/ in /fau/ and /ff/ in /ffa/. The shapes of the articulators for /f/ and /ff/ were similar between the two speakers. The degree of lip protrusion was small for both speakers. M1's lower lip showed some dip, suggesting the presence of labiodental contact; however, this was not seen for M2. Thus, the place of articulation for /f/ was observed to differ between the two speakers: M1 may be labiodental, while M2 may be bilabial. However, because teeth cannot be observed in the images, we cannot empirically assess labiodental contact. Figure 9 shows frames of /vv/ in /vva/. Recall that no word contains a singleton /v/. The lower lip showed some dip for M1, suggesting labiodental articulation (for /vv/). M2's upper surface of the lower lip was flatter in /vv/ than the same part in /f/ in Figure 8. This finding suggests that the upper part of the lower lip was pressed by the upper teeth and flattened. Although unclear without an image of the teeth, a closure may have been formed during /vv/, and the place of articulation of M2's /vv/ may have also been labiodental.

When we compared /ff/ and /vv/, the pharyngeal distance was greater for /vv/, indicating pharyngeal expansion, which tended to be greater in the second frame of /vv/. The larynx was slightly lower during /vv/ than during /ff/ for M1, but this was not the

case with M2. Thus, again, both pharyngeal expansion and laryngeal lowering were observed in M1, but only pharyngeal expansion was observed in M2. We did not identify velum lowering during /vv/ in either speaker.

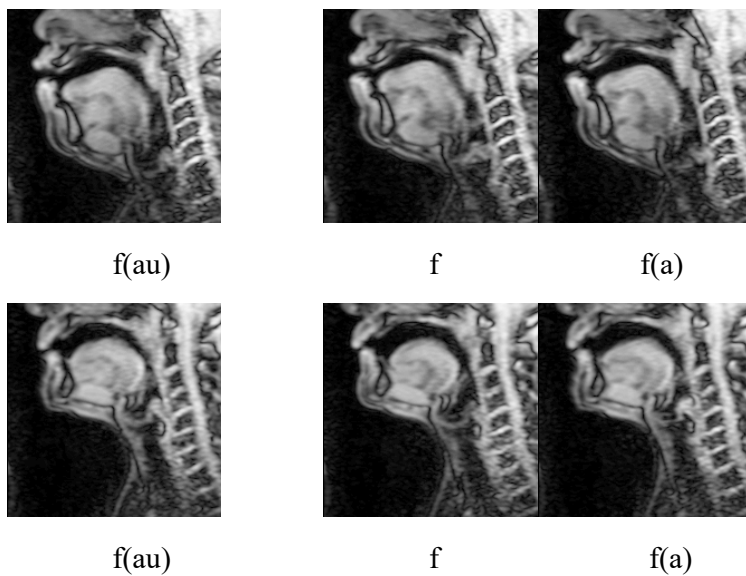


Figure 8 Frames of /f/ in /fau/ (left) and /ff/ in /ffa/ (middle and right). M1 (upper), M2 (lower).

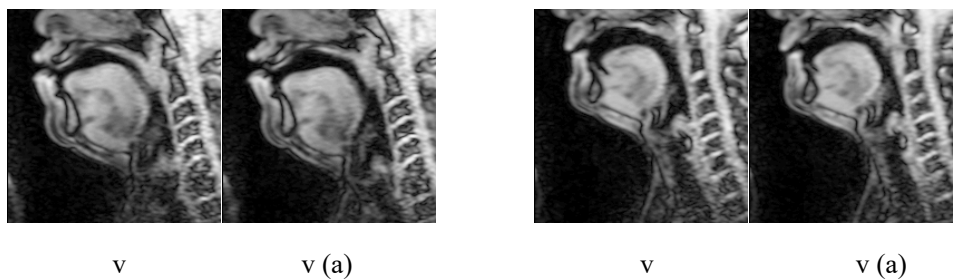


Figure 9 Frames of /vv/ in /vva/. M1 (left two), M2 (right two).

#### 4.2.3 Measurement of linguopalatal contact

To examine quantitative differences among coronal segments of singletons vs. geminates, as well as voiceless vs. voiced consonants, we measured the size of the

linguopalatal contact of /t/, /d/, and /z/ and that of the (narrow) constriction of /s/ for 96 frames (8 (4 singleton + 4 geminate) consonants × 6 tokens × 2 speakers) (Figure 2). Measurements of linguopalatal contact /f/ and /v/ were not possible because we had no information about the teeth. For geminates, the second frame was used. Figure 10 shows the result separated by the speakers. The length was on average more in geminates than in singletons in—not all but—many pairs (/t/, /d/, and /s/ for M1 and /d/ and /z/ for M2), and it was longer in voiceless consonants than in the voiced counterparts in many pairs (/t-/d/ and /tt-/dd/ for M1 and M2, and /s-/z/ for M2). Additionally, it was longer for plosives than for fricatives except that singleton /d/ was shorter than /z/ for both M1 and M2. The average duration of linguopalatal constriction significantly differed between the two subjects ( $t(94) = -3.109, p = .002$ ). Table 4 shows the result of ANOVA, with the size of Linguopalatal Constriction as the dependent variable and with Gemination, Voice, and Manner as independent variables. The results of M1 showed that all the main effects and the interactions of Gemination and Manner, and those of Voice and Manner were significant. The results of M2 showed that the main effects except for Gemination were significant and all the interactions except for that of Gemination, Voice, and Manner were significant.

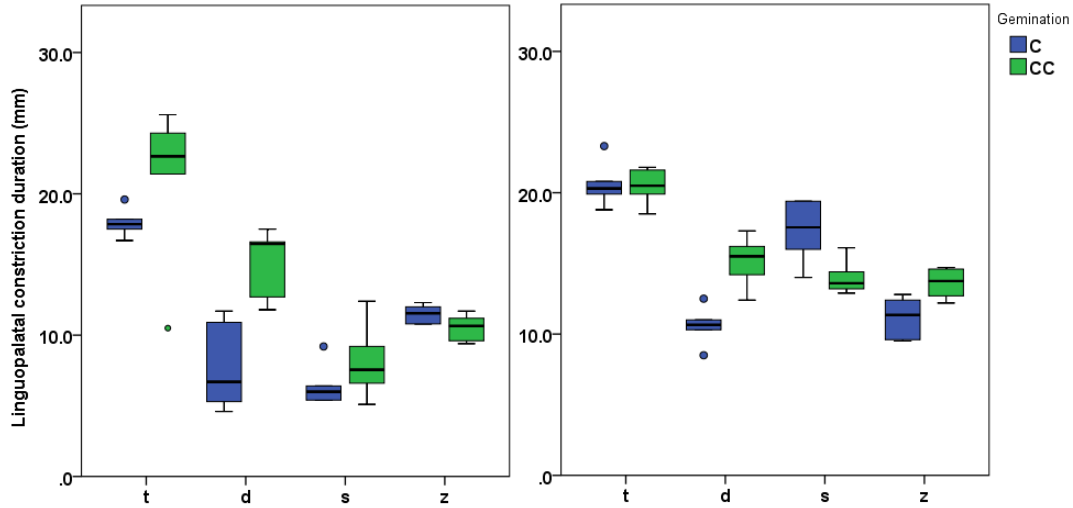


Figure 10 Size of linguopalatal constriction. M1 (left) and M2 (right). (For geminates, second frames were used.) “C” and “CC” denote singletons and geminates, respectively.

Table 4 Results of the three-way ANOVA, with the size of Linguopalatal Constriction as the dependent variable, and with Gemination, Voice, and Manner as independent variables. Asterisk denotes the significance codes (.01\*; .001\*\*).

		M1		M2	
	Df	F value	Sig.	F value	Sig.
Gemination	1	14.424	.000 **	4.460	.041
Voice	1	8.145	.007 *	164.299	.000 **
Manner	1	70.695	.000 **	40.136	.000 **
Gemination × Voice	1	.326	.571	38.169	.000 **
Gemination × Manner	1	11.135	.002 *	10.065	.003 *
Voice × Manner	1	61.423	.000 **	26.591	.000 **
Gemination × Voice × Manner	1	5.313	.026	.433	.514
Residuals	40				

Figure 11 shows the interaction of Gemination, Voice, and Manner separated by the speakers. In Figure 11, the longer constriction in geminates than in singletons more

clearly appears in the plosive pairs, especially in /d/-/dd/, than in the fricative pairs. Analysis by Tukey's multiple comparisons of means showed that (for a 95% family-wise confidence level) the size was significantly larger in /dd/ than /d/ ( $p < .001$ ) slightly larger in /tt/ than /t/ ( $p = .039$ ) but was similar between /ss/ and /s/ ( $p = .279$ ) and /zz/ and /z/ ( $p = .528$ ) for M1: and for M2, the size was significantly larger in /dd/ than /d/ ( $p < .001$ ) and in /zz/ than /z/ ( $p = .006$ ) but similar between /tt/ and /t/ ( $p = .907$ ) and significantly smaller in /ss/ than /s/ ( $p < .001$ ). Thus, our second hypothesis that "the size of linguopalatal constriction is larger in geminates than in singletons" is robustly substantiated for /d/, marginally substantiated for /t/ and /z/, but not for /s/.

Similarly, from Figure 11, the longer constriction in voiceless consonants than voiced ones were clearer in the plosive pairs than in fricative pairs. Tukey's multiple comparisons of means showed that the size was larger in /t/ than /d/ ( $p < .001$ ) and in /tt/ than /dd/ ( $p < .001$ ) but was similar between /ss/ and /zz/ ( $p = .112$ ) and smaller in /s/ than /z/ ( $p = .002$ ) for M1; for M2, the size was larger in /t/ than /d/ ( $p < .001$ ), in /tt/ than /dd/ ( $p < .001$ ), and in /s/ than /z/ ( $p < .001$ ) but was similar between /ss/ and /zz/ ( $p = .682$ ). Thus, the third hypothesis, "the size of linguopalatal constriction is larger for voiceless consonants than for voiced ones," is substantiated for the plosives for two speakers and only partly for the fricatives in one of the speakers. Therefore, fricatives demonstrated the tendency observed in plosives to a lesser degree for both Gemination and Voice. Hence, the fourth hypothesis that "larger constriction in geminates than that in singletons and/or in voiceless consonants than that in voiced ones, if at all, would appear more significantly in plosives than it would in fricatives" is substantiated.

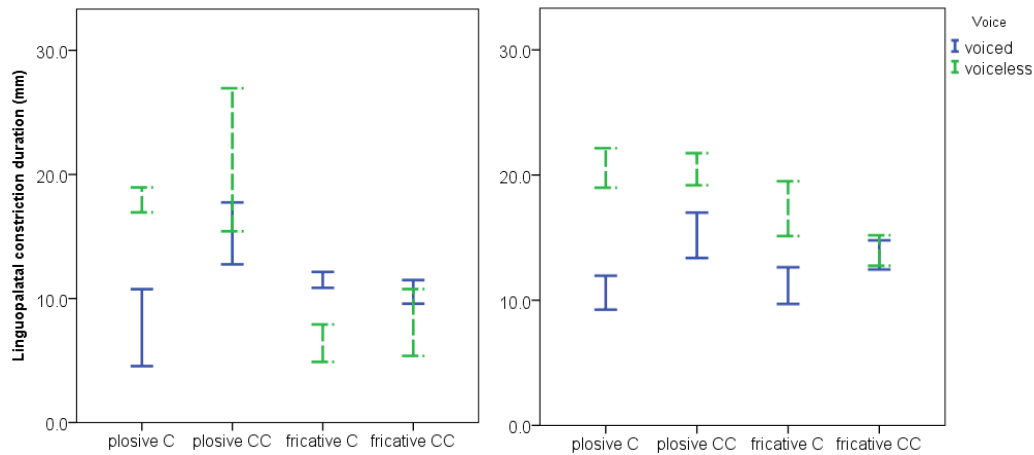


Figure 11 Size of linguopalatal constriction according to Manner, Gemination, and Voice. M1 (left) and M2 (right). Solid (blue) lines indicate voiced consonants and dotted (green) lines, voiceless ones. “C” and “CC” denote singletons and geminates, respectively. The error bar shows 95 % confidence intervals.

#### 4.2.4 Measurement of pharyngeal aperture

During the voiced geminate obstruents, expansion was observed in the pharynx. Figure 12 shows the result of the measurement of the lower part of the pharynx, separated by the speakers and averaging over six tokens for 132 frames (11 consonants (5 singleton + 6 geminate) × 6 tokens × 2 speakers). In Figure 12, pharyngeal distance was longer in geminates than in singletons for voiced consonants /d/ and /z/ for both speakers, regardless of the manner of the consonants. However, for voiceless consonants /t/, /s/, and /f/, the two speakers demonstrated opposite tendencies regarding increasing or decreasing the pharyngeal distance due to gemination. The average duration of pharyngeal distance significantly differed between two subjects ( $t(130) = 4.802, p < .001$ ). Table 5 shows the result of a three-way ANOVA, with Pharyngeal Distance as dependent variable, and with Gemination, Voice, and Manner as independent variables. The consonants /f/, /ff/, and /vv/ were excluded from the analysis as the language lacks singleton /v/ in its phonotactics. The result of M1 showed that the main effects of Gemination and Voice were significant,

and all the interactions were significant. The results of M2' showed that the main effects of Gemination and Voice were significant and all the interactions except for those of Gemination and Manner were significant.

Figure 13 shows the interaction among Gemination, Voice, and Manner. According to the figure, pharyngeal distance was larger for voiced consonants than their voiceless counterparts except for the singleton plosive pair, which showed the opposite tendency. Analysis by Tukey's multiple comparisons of means showed that (for a 95% family-wise confidence level) the size was significantly larger in /dd/ than /tt/ ( $p = .001$ ), /zz/ than /ss/ ( $p < .001$ ) but similar between /s/ and /z/ ( $p = .469$ ) and smaller in /d/ than /t/ ( $p < .001$ ) for M1, and for M2, it was significantly larger in /dd/ than /tt/ ( $p = .001$ ), /z/ than /s/ ( $p < .01$ ) and /zz/ than /ss/ ( $p < .001$ ) but it was smaller in /d/ than /t/ ( $p < .001$ ). Thus, our fifth hypothesis that “the size of the pharyngeal aperture is larger for voiced obstruents than for their voiceless counterparts, and the differences are more evident for the geminates” is generally confirmed while it is not true for the singleton plosive /t/-/d/ pair. Note that the similar results were obtained when /f/ and /ff/ were included in the analysis.

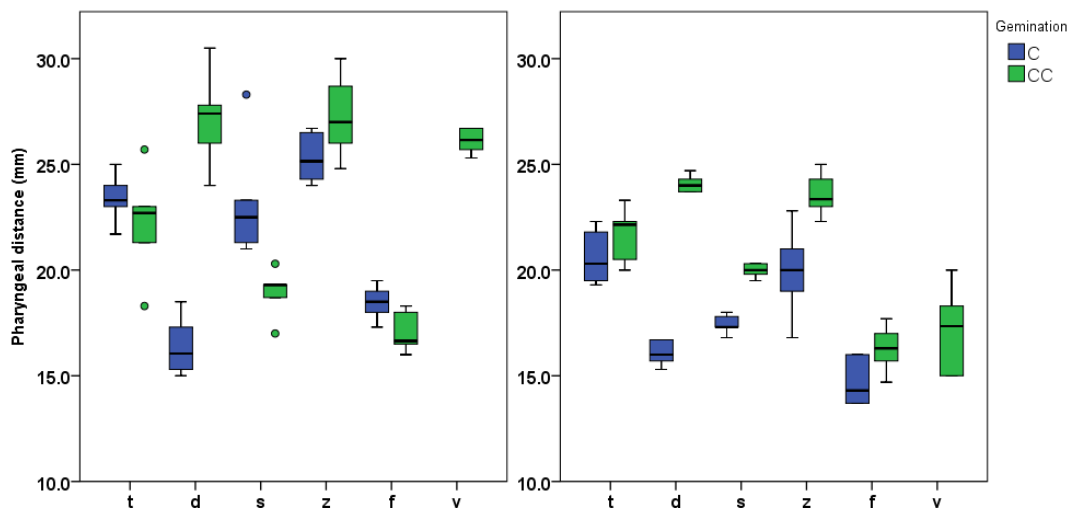


Figure 12 Pharyngeal Distance. M1 (left) and M2 (right). (For the geminates, second frames were used)

Table 5 Results of three-way ANOVA, with the size of Pharyngeal Distance as dependent variable, and with Gemination, Voice, and Manner as independent variables. Asterisk denotes the significance codes (.01\*; .001\*\*)

		M1		M2	
	Df	F value	Sig.	F value	Sig.
Gemination	1	12.673	.001 *	157.323	.000 **
Voice	1	15.521	.000 **	10.178	.003 *
Manner	1	6.733	.013	1.642	.207
Gemination × Voice	1	73.267	.000 **	41.761	.000 **
Gemination × Manner	1	32.084	.000 **	5.889	.020
Voice × Manner	1	35.391	.000 **	45.728	.000 **
Gemination × Voice × Manner	1	7.579	.009 ‡	22.510	.000 **
Residuals	40				

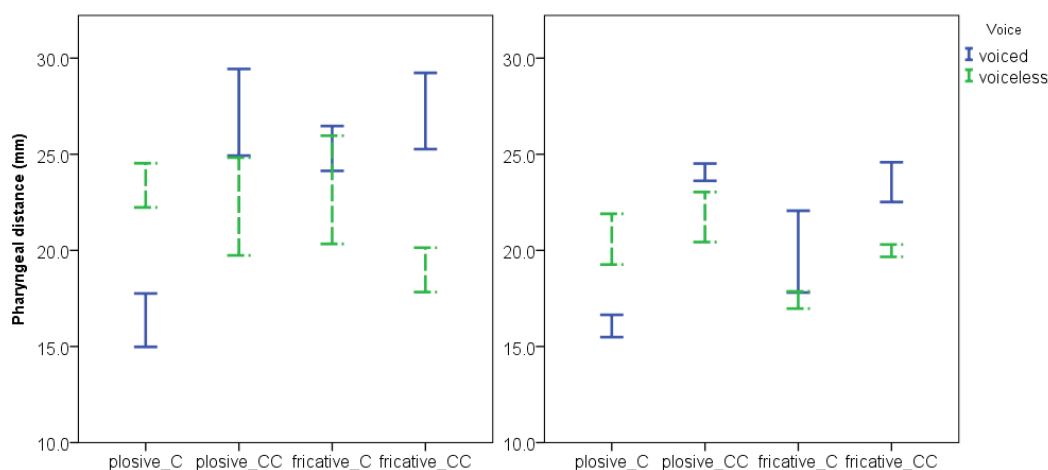


Figure 13 Comparison of Pharyngeal Distance according to Gemination, Voice and Manner. M1 (left) and M2 (right). Solid (blue) lines indicate voiced consonants and dotted (green) lines, voiceless ones. “C” and “CC” denote singletons and geminates, respectively. The error bar shows 95% confidence intervals.



#### 4.2.5 Affricate realization of voiced fricatives

As mentioned in Section 4.2.2, voiced fricative /zz/ and /vv/ were observed to be produced as affricates. Although an affricate variation of /z/ has been reported in Miyako dialects (Pellard & Hayashi 2012), no such allophonic variation has been claimed for /v/. We briefly examine this point by acoustically analyzing the speech sound. Figure 14 shows waveforms and spectrograms of representative tokens of /vva/. A voice bar without components of higher frequency precedes the frication noise during /vv/. In M2's spectrogram, a spike-like burst was observed. These examples confirm that both speakers produced /vv/ as an affricate. In addition, the spectrogram of the same word /vva/ in our previous acoustic study showed a very similar pattern (Shinohara & Fujimoto 2018). These findings strongly suggest the presence of complete closure at the onset of /vv/ for both speakers although this phenomenon was not evident from the MRI images.

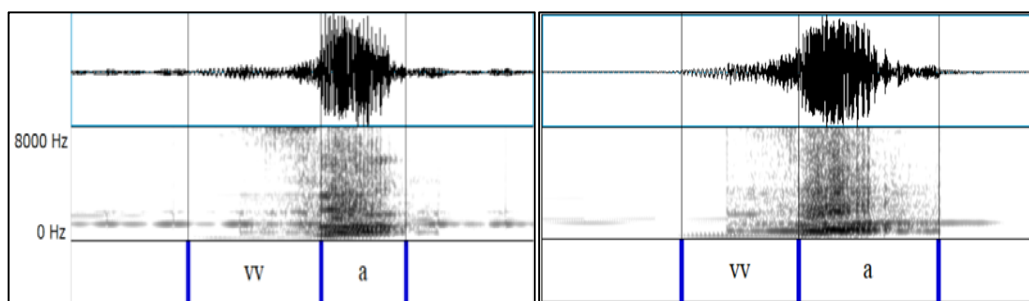


Figure 14 Wave forms and spectrograms (700 ms) of /vva/. M1 (left) and M2 (right).

## 5. Discussion

This paper examined articulatory characteristics of geminate obstruents of the Ikema dialect of Miyako Ryukyuan by using rt-MRI. We reported our results of a length comparison of singleton and geminate obstruents based on frame counts, qualitative

observations of articulatory settings of obstruent consonants, and measurements of linguopalatal contact and pharyngeal distances from the MRI images. On the basis of previous literature, we predicted 1) longer articulatory duration in geminate obstruents than that in their singleton counterparts; 2) stronger articulation in geminates; 3) stronger articulation in voiceless consonants than that in their voiced counterparts; 4) asymmetry between plosives and fricatives; and 5) existence of some articulatory adjustments for voicing initiation and for the maintenance of voicing in voiced geminate obstruents, such as pharyngeal expansion and lowering of the larynx and/or velum. By examining articulatory images, we identified five characteristics. First, geminate obstruents in Ikema had a longer duration of articulatory constrictions. Second, the voiced alveolar plosive geminate /dd/ was articulated with a larger (midsagittal) contact area than that of its singleton counterpart /d/, implying “stronger” articulation; however, the voiceless plosive pair, /tt/-/t/, and the fricative pairs, /ss/-/s/ and /zz/-/z/, did not show such difference in a systematic way. Third, the alveolar voiceless plosives /(t)t/ had a larger linguopalatal contact area than their voiced counterparts /(d)d/ did, but the fricative pairs, /(s)s/-/(z)z/, did not show any difference. Fourth, fricatives did not show any systematic difference in the size of (midsagittal) linguopalatal constriction between geminates and singletons or between voiceless and voiced consonants. Fifth, voiced geminate obstruents were accompanied by pharyngeal expansion for both speakers; also, laryngeal lowering was seen for one speaker but neither speaker showed velum lowering. We also observed that voiced fricatives tend to be produced as affricates.

### **5.1 Articulatory duration and strength**

As predicted, the result of the rt-MRI frame count showed that the labiodental or

alveolar constriction in the geminate obstruents was significantly longer in duration than that in their singleton counterparts. This finding clearly indicates that speakers of the Ikema dialect differentiate between singletons and geminates in word-initial plosives and fricatives and the word-medial plosive (/d/ vs. /dd/). Although the frame rate of the rt-MRI system that we used is relatively low, we could observe the articulatory duration of inaudible closure phase of the voiceless geminate plosives in phrase-initial position by using this method. The result of the rt-MRI supports our previous acoustical analysis with five speakers of the Ikema dialect (including the two analyzed in this paper), showing that word-medial geminate consonants and word-initial geminate fricatives are significantly longer than their singleton counterparts (Shinohara & Fujimoto 2018). The clear length difference conforms to word-medial geminates in standard Japanese (Homma 1981, Beckman 1982, Han 1962; 1994, Idemaru & Guion 2008, Hussain & Shinohara 2019, among others) and other languages that have a singleton–geminate contrast (e.g., Ham 2001, Ridouane 2007, Hussain 2015). The distinction also held for the closure period of initial voiceless stop (/t/ vs. /tt/), as in the cases of Pattani Malay (Abramson 1986 et seq.) and Tashlhiyt Berber (Ridouane 2010, Ridouane & Hallé 2017).

Our measurements of the images revealed that the size of the linguopalatal contact area of the alveolar plosives was significantly larger in word-medial geminate /dd/ than that in singleton /d/, but no difference was statistically confirmed for the word-initial pair, /tt/-/t/; however, a more raised tongue body and longer contact duration was recognized in the images of /tt/ than in those of /t/ for one speaker. EPG studies have indicated that geminate plosives have a wider tongue contact than singleton ones in both voiceless and voiced plosives, which has been interpreted as geminates having “stronger” articulation (Payne 2006, Ridouane 2007, Kochetov & Kang 2017, Kawahara & Matsui

2017). Our results for plosives agree with their findings for /d/ but not for /t/. We suggest that the reason for not finding greater linguopalatal contact for geminate /tt/ in the measurement across speakers is the position of /t/ in the word. In our recording corpus, the positions of /t/-/tt/ and /d/-/dd/ are uneven: /t/-/tt/ contrast in /\_a/ context was found only in word-initial position whereas /d/-/dd/ contrast occurs only in word-medial position in Ikema. Articulatory strengthening is perhaps more implementable word-medially than word-initially. If so, this result may be one of the grounds across languages for the favoring of a word-medial position for geminates over word-initial position. Another possibility is that Ikema has a different articulatory patterning in terms of the singleton–geminate contrast for plosives from other languages. Although an acoustic study showed that VOT was longer in singleton /t/ than in /tt/ (Shinohara & Fujimoto 2018), our MRI study did not allow evaluation of the effect of VOT duration. Further research is necessary to clarify this point.

In addition, for plosives, the size of linguopalatal constriction was larger in voiceless consonants than in their voiced counterparts. The result agrees with the studies of standard Japanese (National Institute for Japanese Language and Linguistics 1990, Kitamura & Hatano 2012, Kochetov 2014), suggesting that the Ikema dialect is similar to standard Japanese with respect to linguopalatal constriction. Moreover, this phenomenon may be a universal articulatory tendency.

For alveolar fricatives, the size of linguopalatal constriction varied depending on the voicing, gemination, and speakers. This finding might partly be because lateral constriction plays a greater role than mid-sagittal constriction in fricatives, as shown by Fuchs et al. (2006). If so, for fricatives, the difference between singletons vs. geminates and voiced vs. voiceless consonants cannot be determined solely from midsagittal plane

observations. The analysis of cross-sectional plane is necessary to assess these assumptions in a future study.

## 5.2 Pharynx expansion

As predicted, voiced geminates were accompanied by significant pharyngeal expansion in plosive /dd/ and fricative /zz/ ([ddz], Section 5.3 presents a discussion on affrication). There was a clear difference of the size of the pharynx between voiceless and voiced geminate obstruents. Pharyngeal expansion may facilitate vocal fold vibration during voiced obstruent production. Thus, we regard pharyngeal expansion as an articulatory adjustment to initiate and/or maintain vocal fold vibration during voiced obstruents, especially for geminates with elongated durations. A large pharyngeal expansion may not be essential for word-medial voiced singleton obstruents because the vocal folds may easily continue to vibrate from the preceding vowel. Nevertheless, the difference of the size was found in the singleton pair, /z/-/s/, but less than in the geminate pair, /zz/-/ss/. As for the singleton plosive pair, pharyngeal distance was even smaller for /d/ as compared to /t/. However, as mentioned in Section 5.1, the voicing contrast of the plosive pairs appear in different positions. Thus, the result for the /t/-/d/ comparison might be due to their position rather than the voicing difference. However, the size difference in the corresponding geminate pair, /tt/-/dd/, must have been large enough to overcome the positional difference.

Kent and Moll (1969: 1552) reported that voiced stops ‘are associated with an increase in cavity size during the closure period.’ By contrast, ‘the voiceless stops have smaller increases, no increases at all, or a slight reduction in cavity size during the period of closure.’ Our results are in line with this description although the languages studied are

different. In our previous rt-MRI study of a speaker of the Kumamoto dialect of Japanese, acoustic data showed half-devoicing of word-medial voiced geminate /dd/ and /gg/, which are phonologically contrastive with their voiceless counterparts. Additionally, moderate pharyngeal expansion was observed during the voiced geminates (Fujimoto & Shinohara 2018a). Although the pharyngeal distance was not measured in the study, the size in geminates was observed to be much smaller than that of the speakers of the Ikema dialect in this study. To confirm the general assumption that the amount of expansion is proportional to the degree of voicing in the geminate obstruents, further evaluation by comparing voiced and voiceless obstruents across dialects is essential.

A slight lowering of the larynx was observed for one speaker but not for the other. We did not recognize velum lowering for either speaker. Hence, we assume that our speakers of the Ikema dialect use pharyngeal expansion as a main strategy to initiate and maintain voicing during voiced geminates.

### **5.3 Affricate realization of voiced fricatives**

Voiced fricatives /z/ and /zz/ were produced as affricates in our audio recordings. This paper supplied articulatory evidence for the description of alveolar affricates by Pellard and Hayashi (2012). At the onset of /vv/, another voiced fricative in Ikema, a dip in the lower lip was observed in one speaker (M1) and a flat shape at the top of the lower lip in the other speaker (M2). These observations suggest a closure between the upper teeth and the lower lip. Hence, a labiodental closure was likely formed. Acoustic analysis supports this observation (Figure 14). However, we have not observed any description of the affrication of /vv/ in the Ikema dialect in the literature. More articulatory and acoustic examinations are required to evaluate this analysis.

From phonological perspectives, the voiced fricatives are typologically disfavored (Maddieson 1984, Kirchner 2000). In the case of Ikema, the constraint against voiced geminate fricatives seems to be respected by means of affrication at the phonetic level. Zygis, Fuchs & Koenig (2012) claimed, by contrast, that voiced (sibilant) affricates are typologically more marked than voiced plosives and fricatives because of their complex oral articulation, in addition to their demanding voicing maintenance during oral constriction. Their claim may seem counter to the affrication process in Ikema at first glance. However, Zygis et al. (2012) define the phonetic foundation for markedness of voiced affricates in the stricter timing coordination of constrictions for voiced (singleton) affricates when compared with voiced simplex obstruents (i.e., plosives and fricatives) and not along the dimension of voicing. Given the long duration imposed on the geminates, it might be that the voicing of long voiced fricative segments is aided by a complete closure.

In general, oral closure would stop the existing vocal fold vibration (Ohala & Riordan 1979). However, Rothenberg (1986) noted possible mechanisms to sustain glottal oscillations with complete articulatory closure and include passive or active enlargement of the supraglottal cavity and nasal venting. While our speakers did not demonstrate any sign of velopharyngeal opening, we did observe pharyngeal expansion. Arguably, pharyngeal expansion might become easier with oral closure (i.e., plosives/affricates) than without (i.e., fricatives), which would explain the considerable amount of pharyngeal expansion in the voiced geminates in our Ikema speakers (especially remarkable in the second frame). In the case of Ikema, because there are no contrastive voiced affricates, affricates can be used as allophonic variants of fricatives. Oral closure at the beginning of the voiced fricatives /*(z)z*/, and possibly /*vv*/, may be an articulatory adjustment to

expand the pharynx, particularly for the geminate.

Affrication of voiced fricatives in the Ikema dialect may be related to the observation that the voiced fricative /z/ can be realized as an affricate in many Japanese dialects, including standard Japanese (Shibatani 1990, Maekawa 2010)<sup>vii</sup>, especially word-initially (Kawakami 1977, Vance 1987). The voiced fricative /z/ does not have a contrastive voiced affricate counterpart in Japanese (except at the surface level in Shikoku dialects), and this also applies to /z/, /zz/, and /vv/ in the Ikema dialect. In Japanese loanwords, /z/ is likely to be realized as an affricate when geminated, as observed in /bazzi/ [bad.dʒi]<sup>viii</sup> ‘badge’ (Kawagoe 2015). Cross dialectal studies on the affrication of voiced fricatives may add an insight to the phonetics and phonology of Japanese and Ryukyuan languages.

Although the segmental environments and number of speakers are limited, our analyses provide additional insights on the articulatory maneuver of initiating and maintaining voicing throughout the long duration of geminate obstruents. For voiceless word-initial singletons vs. their geminate counterparts, the size of linguopalatal constriction did not show statistical difference. It may be that producing geminates in the initial position might be more articulatorily challenging, which accords with the cross-linguistic rarity of initial geminate consonants. To evaluate our findings, additional analyses using finer techniques are essential. Further research could include more words with more speakers. In addition, a comparison of our results with those of Tokyo speakers, whose voiced geminates tend to be devoiced, would be desirable.

### **Acknowledgement**

This paper is a follow-up analysis to our studies presented at GemCon2015 and the Spring



Meeting of the Acoustical Society of Japan 2016. We are grateful to our participants for their willingness and patience, to Yuriko Yokoe-Matsumoto for helping data management and to Chakir Zeroual, Hyunsoon Kim, Lise Crevier-Buchman, and Pierre Badin for helpful comments and information on MRI analyses. We thank the editors and three anonymous reviewers of JIPA and Donna Erickson for useful comments and constructive criticism. Special thanks are due to Tatsuya Kitamura for his help on articulatory measurements and Kikuo Maekawa, Mafuyu Kitahara, and Keiichi Tajima for the useful advice on statistical analyses. Shortcomings belong solely to the authors. The study was supported by research grants from SOLIFIC for the project ‘Acoustic and Articulatory Study of Geminates and their Phonological Structures’, Sophia University Phonetics Laboratory and JSPS Kakenhi No. 26370470 and No. 17K02707 attributed to the first author.

## References

- Abramson, Arthur S. 1986. The perception of word-initial consonant length: Pattani Malay, *Journal of the International Phonetic Association* 16, 8–16.
- Abramson, Arthur S. 1987. Word-initial consonant length in Pattani Malay. *Proceedings of the XIth International Congress of Phonetic Sciences*, Tallinn, Academy of Sciences of the Estonian USSR. 68–70.
- Abramson, Arthur S. 1991. Amplitude as a cue to word-initial consonant length: Pattani Malay. *Proceedings of the XIIth International Congress of Phonetic Sciences*, Aix-en-Provence, Centre National de la Recherche Scientifique, 98–101.
- Abramson, Arthur S. 1999a. The complex acoustic output of a single articulatory gesture: Pattani Malay word initial consonant length. In Udom Warotamasikkhadit & Thanyarat Panakul (eds.), *Papers from the Fourth Annual Meeting of the Southeast Asian Linguistics Society 1994*, Tempe, Arizona State University, 1–20.
- Abramson, Arthur S. 1999b. Fundamental frequency as a cue to word-initial length: Pattani Malay. *Proceedings of the XIVth International Congress of Phonetic Sciences*, San Francisco, USA, 591–594.
- Abramson, Arthur S. 2003. Acoustic cues to word-initial stop length in Pattani Malay. *Proceedings of the XVth International Congress of Phonetic Sciences*, Barcelona,

- Spain, 387–390.
- Arvaniti, Amalia & Georgios Tserdanelis. 2000. On the phonetics of geminates: evidence from Cypriot Greek. In Baozong Yuan, Taiyi Huang & Xiaofang Tang (eds.), *Proceedings of International Conference on Spoken Language Processing ICSLP 2000*, Beijing, China, 559–62.
- Beckman, Mary E. 1982. Segment duration and the ‘Mora’ in Japanese, *Phonetica* 39. 113–135.
- Boersma, Paul & David Weenink. 2016. Praat: doing phonetics by computer [Computer program]. Version 6.0.22, retrieved 15 November 2016 from <http://www.praat.org>.
- Brain Activity Imaging Center at ATR-promotions (ATR-BAIC): <http://www.baic.jp/reference/imageMRI.html>.
- Bresch, Eric, Yoon-Chul Kim, Krishna Nayak, Dani Byrd & Shrikanth Narayanan. 2008. Seeing speech: Capturing vocal tract shaping using real-time magnetic resonance imaging [Exploratory DSP], *IEEE Signal Processing Magazine* 25(3), 123–132. doi: 10.1109/MSP.2008.918034.
- Bresch, Eric & Shrikanth Narayanan. 2009. Region Segmentation in the Frequency Domain Applied to Upper Airway Real-Time Magnetic Resonance Images, *IEEE Transactions on Medical Imaging* 28(3), 323–338.
- Crawley, Michael J. 2014. Count Data, *Statistics: An Introduction Using R*, 2nd edition. 234–255. New York: Wiley.
- Dagenais, Paul A. Leah C. Lorendo & Martic J. McCutcheon. 1994. A study of voicing and context effects upon consonant linguapalatal contact patterns. *Journal of Phonetics* 22, 225–238.
- Davis, Stuart. 1999. On the representation of initial geminates. *Phonology* 16, 93–104.
- Engwall, Olov. 2006. Assessing Magnetic Resonance Imaging measurements: Effects of sustenation, gravitation, and coarticulation. In Jonathan Harrington & Marija Tabain (eds.), *Speech Production: Models, Phonetic Processes, and techniques*, 301–314. New York: Psychology Press.
- Ewan, William G. & Robert Krones. 1974. Measuring larynx movement using the thyroumbrometer, *Journal of Phonetics* 2, 323–335.
- Fuchs, Susanne, Pascal Perrier, Christian Geng & Cristine Mooshammer. 2006. What role

- does the palate play in speech motor control? Insight from tongue kinematics for German Alveolar obstruents, In Jonathan Harrington & Marija Tabain (eds.), *Speech production: Models, Phonetic Processes, and Techniques*, 149–164. New York: Psychology Press.
- Fujimoto, Masako & Seiya Funatsu. 2018. Airflow patterns of voiced geminate stops in Japanese, *Journal of the Acoustical Science & Technology* 39(2), 66–74.
- Fujimoto, Masako & Shigeko Shinohara. 2013. Acoustic characteristics of geminate consonants in the Ikema variety of Miyakojima dialect of Southern Ryukyuan, *Proceedings of autumn meeting of Acoustical Society of Japan*. 431–432. (in Japanese).
- Fujimoto, Masako & Shigeko Shinohara. 2015. Articulation of voiced geminate and word-initial geminate obstruents in the Miyakojima Ikema dialect of Ryukyuan: A Real-time MRI analysis, *Paper presented at ICPhS 2015 Satellite Workshop on 'Geminate Consonants across the World,'* Glasgow, England. 27–28.
- Fujimoto, Masako & Shigeko Shinohara. 2017. Articulatory manifestations for voiced geminates and initial geminates in Ikema Ryukyuan, *Abstract/Talk presented at Methods in Dialectology XVI*, National Institute for Japanese Language and Linguistics, Tokyo, Japan, 1 Page.
- Fujimoto, Masako & Shigeko Shinohara. 2018a. Articulatory Characteristics of Geminate Plosives of Kumamoto Dialect: A Case Study Using MRI, *Journal of the Phonetic Society of Japan* 22(2), 95–108. (In Japanese).
- Fujimoto, Masako & Shigeko Shinohara. 2018b. Miyakojima no tyuuzetsu boin no tyoon [Articulation of central vowel in Miyakojima], *Proceedings of Phonetic Society of Japan*, Okinawa, Japan, 120–126. (In Japanese).
- Gili-Fivela, Barbara, Claudio Zmarich, Pascal Perrier, Christophe Savariaux, & Graziano Tisato. 2007. Acoustic and Kinematic correlates of phonological length contrast in Italian consonants. *Proceedings of 16<sup>th</sup> International Conference of Phonetic Sciences*, Saarbrücken, Germany, 469–472.
- Grafen Alan & Rosie Hails. 2002. *Modern Statistics for the Life Sciences*, Oxford: Oxford University Press.
- Hagedorn, Christina, Michael Proctor & Louis Goldstein. 2011. Automatic analysis of singleton and geminate consonant articulation using real-time magnetic resonance

- imaging. *Proceedings of the 12th Annual Conference of the International Speech Communication Association*, 409–412.
- Ham, William H. 2001. *Phonetic & phonological aspects of geminate timing*. New York: Routledge.
- Han, Mieko S. 1962. The feature of duration in Japanese, *Onsei no Kenkyuu* 10, 65–80.
- Han, Mieko S. 1994. Acoustic manifestations of mora timing in Japanese. *Journal of the Acoustical Society of America* 96, 73–82.
- Hayes, Bruce & Donca Steriade. 2004. The phonetic bases of phonological Markedness. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically-Based Phonology*, 313–345. Cambridge: Cambridge University Press.
- Hirata, Soichiro, Takeshi Wada, Takashi Tachimura, Hisanaga Hara, Kanji Nohara & Koichi Satoh. 2002. Kansai hogen washa ni okeru nazomeetaa kensa deno nihongo hikenbun to bikuu heisa kinoo huzen no hyooka [Nasalance of speech samples in kansai (the mid-western part of Japan) dialect speakers and its availability as an evaluation tool for identifying patients with velopharyngeal incompetence following cleft palate surgery], *Journal of Japanese Cleft Palate Association* 27, 14–23. (In Japanese).
- Homma, Yayoi. 1981. Durational relationships between Japanese stops and vowels. *Journal of Phonetics* 9 (3), 273–281.
- Hume, Elizabeth, Jennifer Muller & Aone van Engelenhoven. 1997. Non-Moraic Geminate in Leti. *Phonology*, 14 (3), 371–402.
- Hussain, Qandeel. 2015. Temporal characteristics of Punjabi word-medial singletons & geminates, *The Journal of the Acoustical Society of America* 138, EL388–EL391. doi: 10.1121/1.4931832.
- Hussain, Qandeel & Shigeeko Shinohara. 2019. Partial devoicing of voiced geminate stops in Tokyo Japanese, *The Journal of the Acoustical Society of America* 145 (1), 149–163.
- IBM Corp. Released. 2013. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.
- Idemaru, Kaori & Susan G. Guion. 2008. Acoustic covariants of length contrast in Japanese stops. *Journal of the International Phonetic Association* 38, 167–186.

- Iribar, Alexander, Rosa M. Pagola, Itziar Turrez, José L. G. Arroyo, Begoña G. Zapirain & Ibon O. Ruiz. 2018. Parameters of tongue shape of /n/ and /l/ in Basque. *Journal of the International Phonetic Association*, 1–15.
- Jaeger, Jeri. 1978. Speech aerodynamics & phonological universals. *Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society*, 311–329.
- Kawagoe, Itsue. 2015. The phonology of sokuon, or geminate obstruents, In Haruo Kubozono (ed.), *The handbook of Japanese phonetics and phonology*. Berlin: De Gruyter Mouton, 79–119.
- Kawahara, Shigeto. 2006. A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese. *Language* 82(3), 536–574.
- Kawahara, Shigeto & Michinao Matsui F. 2017. Some aspects of Japanese consonant articulation: A preliminary EPG study. *ICU Working Papers in Linguistics* 2, 9–20.
- Kawakami Shin. 1977. *Nihongo onsei gaisetsu* [Introduction to Japanese Sounds], Tokyo: Ohfusya. (in Japanese).
- Kent, Raymond & Moll, Kenneth. 1969. Vocal-Tract Characteristics of the Stop Cognates, *The Journal of the Acoustical Society of America* 46, 1549–1555.
- Kharlamov, Viktor. 2018. Prevoicing and prenasalization in Russian initial plosives, In Taehong Cho, Gerry Docherty & Douglas H. Whalen (eds.), *Journal of Phonetics* 71, *Special Issue: Marking 50 Years of Research on Voice Onset Time*, 215–228.
- Kibe Nobuko (ed.) 2012. *General Study for Research and Conservation of Endangered Dialects in Japan: Research Report on Miyako Ryukyuan*, Tokyo: National Institute for Japanese Language and Linguistics. (in Japanese).
- Kim, Hyunsoon, Kiyoshi Honda & Shinji Maeda. 2005. Stroboscopic-cine MRI study of the phasing between the tongue and the larynx in the Korean three-way phonation contrast. *Journal of Phonetics* 33, 1–26.
- Kim, Hyunsoon, Shinji Maeda & Kiyoshi Honda. 2010. The laryngeal characterization of Korean fricatives: Stroboscopic cine-MRI data. *Journal of Phonetics* 39(4), 626–641.
- Kirchner, Robert. 2000. Geminate inalterability and lenition. *Language* 76(3), 509–545.
- Kitamura, Tatsuya & Hiroaki Hatano. 2012. Measurement of temporal change of vocal

- tract volume during production of plosive and fricative consonants, *IEICE Technical report* 112 (281), SP2012-75, 19–23. (in Japanese).
- Kochetov, Alexei. 2014. Voicing and Tongue-Palate Contact Differences in Japanese Obstruents, *Journal of the Phonetic Society of Japan* 18 (2), 63–76.
- Kochetov, Alexei & Yoonjung Kang. 2017. Supralaryngeal implementation of length and laryngeal contrasts in Japanese and Korean, *Canadian Journal of Linguistics*. 62(1), 18–55.
- Kraehenmann, Astrid. 2001. Swiss German Stops: Geminates all over the word, *Phonology* 18, 109–145.
- Kraehenmann, Astrid & Aditi Lahiri. 2008. Duration differences in the articulation and acoustics of Swiss German word-initial geminate and singleton stops, *Journal of the Acoustical Society of America*. 123 (6), 4446–4455.
- Labrunie, Mathieu, Pierre Badin, Dirk Voit, Arun Joseph, Jens Frahm, Laurent Lamalle, Coriandre Vilain & Louis-Jean Boë. 2018. Automatic segmentation of speech articulators from real-time midsagittal MRI based on supervised learning. *Speech Communication* 99, 27–46. doi:10.1016/j.specom.2018.02.004.
- Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell.
- Lammert, Adam, Louis Goldstein, Shrikanth Narayanan & Iskarous Khalil. 2013. Statistical Methods for Estimation of Direct and Differential Kinematics of the Vocal Tract. *Speech Communication* 55(1), 147–161.
- Löfqvist, Anders. 2007. Tongue movement kinematics in long and short Japanese consonants, *Journal of the Acoustical Society of America*, 122(1), 512–518.
- Maddieson, Ian. 1984. *Patterns of Sounds*, Cambridge: Cambridge University Press.
- Maekawa, Kikuo. 2010. Coarticulatory reinterpretation of allophonic variation: Corpus-based analysis of /z/ in spontaneous Japanese. *Journal of Phonetics* 38(3), 360–374.
- Martins, Paula, Inês Carbone, Alda Pinto, Augusto Silva & António Teixeira. 2008. European Portuguese MRI based speech production studies. *Speech communication* 50, 925–952.
- Matsuura, Toshio. 2012. Miyako shohogen no onsei jitsugen ni kansuru yobiteki kento

- [Preliminary study of voicing realization in Miyako dialects], In Kibe Nobuko (ed.), *General Study for Research and Conservation of Endangered Dialects in Japan: Research Report on Miyako Ryukyuan*, 111–126, Tokyo: National Institute for Japanese Language and Linguistics. (in Japanese).
- Matsuura, Toshio. 2016. Amakusa shohogen ni okeru yusei sokuon no oninron-teki onseigaku-teki kijyutu [Phonological and phonetic description of voiced geminates in Amakusa, Japanese], *National Institute for Japanese Language and Linguistics, Research Papers* 10, 159–177. (in Japanese).
- Moseley, Christopher (ed.). 2009. *Atlas of the World's Languages in Danger*, 3rd edition. Paris, UNESCO Publishing. Online version:  
<http://www.unesco.org/culture/en/endangeredlanguages/atlas>
- Muller, Jennifer. 2001. *The phonology and phonetics of word-initial geminates*.  
 Doctoral Dissertation, Ohio State University.  
[https://etd.ohiolink.edu/pg\\_10?0::NO:10:P10\\_ACCESSION\\_NUM:osu1364226371](https://etd.ohiolink.edu/pg_10?0::NO:10:P10_ACCESSION_NUM:osu1364226371).
- National Institute for Japanese Language and Linguistics (NINJAL). 1990. Nihongo no Boin, Shiin, Onsetsu: Cyoon undo no jikken onseigakuteki kenkyu [Japanese vowels, consonants, syllables: Experimental phonetics research of articulatory movements], *The National Language Research Institute Research Report* 100. (info:doi/10.15084/00001212)
- Ohala, John. 1983. The origin of sound pattern in vocal tract constraints. In Peter MacNeilage (ed.), *The Production of Speech*, 189–216. New York: Springer.
- Ohala, John. 1997. Aerodynamics of phonology, *Proceedings of 4th Seoul International Conference on Linguistics*, Seoul, Korea, 92–97.
- Ohala, John & Carol Riordan. 1979. Passive vocal tract enlargement during voiced stops, *Proceedings of the Acoustical Society of America annual meetings Spring* 65, Suppl.1, 923.
- Payne, Elinor M. 2006. Non-durational indices in Italian geminate consonants. *Journal*

- of the *International Phonetic Association* 36(1), 83–95.
- Pellard, Thomas & Yuka Hayashi. 2012. Miyako shohogen no onin: Taikei to hikaku [Phonology of Miyako dialects -system and comparison-]. In Kibe Nobuko (ed.), *General Study for Research and Conservation of Endangered Dialects in Japan: Research Report on Miyako Ryukyuan*, 13-51, Tokyo: National Institute for Japanese Language and Linguistics. (in Japanese).
- Perkell, Joseph. 1969. *Physiology of speech production: Results and implications of a quantitative cineradiographic study*. Cambridge MA: MIT Press.
- Proctor, Michael I., Christine H. Shadle & Khalil Iskarous. 2010. Pharyngeal articulation in the production of voiced and voiceless fricatives. *Journal of the Acoustical Society of America* 127(3), 1507–1518.
- Ramanarayanan, Vikram, Louis Goldstein, Dani Byrd & Shrikanth S. Narayanan. 2013. An investigation of articulatory setting using real-time magnetic resonance imaging, *Journal of the Acoustical Society of America* 134, 510–519.
- Ramanarayanan, Vikram, Sam Tilsen, Michael Proctor, Johannes Töger, Louis Goldstein, Krishna S. Nayak & Shrikanth Narayanan. 2018. Analysis of speech production real-time MRI. *Computer Speech & Language* 5, 1–22.
- Ridouane, Rachid. 2007. Gemination in Tashliyt Berber: an acoustic and articulatory study. *Journal of the International Phonetic Association* 37(2), 119–142.
- Ridouane, Rachid. 2010. Gemination at the junction of phonetics and phonology, *Laboratory Phonology X*, 61–90.
- Ridouane, Rachid & Pierre Hallé. 2017. Word-initial geminates from production to perception, In Haruo Kobozono (ed.), *Phonetics and phonology of geminate consonants*, 66–84. Oxford: Oxford University Press.
- Rothenberg, Martin. 1968 *The breath-stream dynamics of simple-released-plosive production*, *Biblioteca Phonetica* 6. Basel: Karger.
- Schneider, Caroline A. Wayne S. Rasband & Kevin W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of image analysis, *Nature methods* 9 (7), 671–675.
- Shibatani, Masayoshi. 1990. *The languages of Japan*. Cambridge: Cambridge University Press.
- Shinohara, Shigeo & Masako Fujimoto. 2018. Acoustic characteristics of the obstruent



- and nasal geminates in the Ikema dialect of Miyako Ryukyuan. In Elena Babatsouli (ed.), *Cross-linguistic research in monolingual and bilingual speech*, Institute of Monolingual and Bilingual Speech, Chania, Greece, 253–270.
- Solé, Maria-Josep. 2018. Articulatory adjustments in initial voiced stops in Spanish, French and English, *Journal of Phonetics* 66, 217–241.
- Solé, Maria-Josep, Ronald Sprouse & John J. Ohala. 2011. Voicing control and nasalization, *Proceedings of Laboratory Phonology* 11, Wellington, New Zealand. 127–128.
- Takada, Masaharu. 1985. Sokuon no choonjyo no tokucho ni tsuite [On some articulatory characteristics of the mora obstruent], *National Language Research Institute Research Report 83, Occasional Papers* 6. National Institute for Japanese Language and Linguistics, 17–40. (in Japanese).
- Takada, Mieko. 2018. The Phonetic Details of Voiced Geminates in the Kumamoto Dialect: The Effects of Generation and Gender. *Journal of the Phonetic Society of Japan* 22(2), 109–121. (in Japanese).
- Digital Museum for Endangered Languages and Cultures. "Languages and cultures"/ Miyako islands/ Miyako Nishihara quarter, <http://kikigengo.jp/library/miyako/nishihara/index.php> (Currently not accessible)
- Thurgood, Graham. 1993. Geminates: A cross-linguistic examination. In Joel Ashmore Nevis, Gerald McMenemy & Graham Thurgood (eds.), *Papers in Honor of Frederick H. Bregelman on the Occasion of the Twenty-Fifth Anniversary of the Department of Linguistics*, 129–139. Fresno: Department of Linguistics, California State University.
- Topintzi, Nina. 2008. On the Existence of Moraic Onset Geminates, *Natural Language & Linguistic Theory* 26(1), 147–184.
- Topintzi, Nina & Davis Stuart. 2017. On the weight of edge geminates, In Haruo Kobozono (ed.), *The phonetics and phonology of geminate consonants*, 260–282. Oxford: Oxford University Press
- Vance, Timothy J. 1987. *An introduction to Japanese phonology*. New York: University of New York Press.

Westbury, John. 1983. Enlargement of the supraglottal cavity and its relation to stop consonant voicing, *The Journal of the Acoustical Society of America* 73(4), 1322–1336.

Zygis, Marzena, Susanne Fuchs & Laura L. Koenig. 2012. Phonetic explanations for the infrequency of voiced sibilant affricates across languages, *Laboratory Phonology* 3, 299–336.

---

<sup>i</sup> This dialect also has voiced and voiceless nasal geminates in word-initial position; however, this paper does not focus on those (see Shinohara & Fujimoto 2018).

<sup>ii</sup> A plausible explanation is that a stronger articulation in geminates may provide cues to the length distinction of consonants, especially in utterance-initial position. Possible acoustic cues provided by a stronger articulation may include the intensity of the burst, rate of formant transition, fundamental frequency perturbations, relatively greater amplitude of the following vowel (Abramson 1986, 1991), and VOT (Ridouane & Hallé 2017).

<sup>iii</sup> We assumed that pharyngeal expansion was not observed because it was not mentioned.

<sup>iv</sup> To compile test words, we consulted Kibe (2012), Pellard and Hayashi (2012), and Digital Museum for Endangered Languages and Cultures, Miyako Islands, Miyako Nishihara area, Dictionary (in Japanese, not in service in September 2020). Words starting with a geminate /dd/ or words containing /gg/ or singleton /v/ were not found in any of these corpora. Words starting with /tts/ were too limited to make a reasonable minimal pair. /kk/ occurred in a very limited number of items (/gakkoo/ ‘school’).

<sup>v</sup> A zero standard deviation for some cells means that the number of the frames for these consonants was one, as for M1’s /d/, or two, as for M1’s /dd/, across the tokens.

<sup>vi</sup> Contrary to the description that /ssa/ is pronounced as [fwsa] (Pellard & Hayashi 2012), protrusion of the lips did not appear ahead of /s/. Thus, the phonetic realization of /ssa/ was [ssa] for the present speakers.

<sup>vii</sup> In standard Japanese, the opposition between /z/ and /d/ is neutralized before /u/ to [dz]~[z], and it is neutralized before /i/ and /j/ to [z]~[dz]. Recent loanwords, where the sequences /du/ [du] and /di/ [di] are employed, are exceptions to these rules.

<sup>viii</sup>[dʒ] is used for [dz] in the original transcription.