

Voiceless nasals in the Ikema dialect of Miyako Ryukyuan

Catherine Ford 

University of Alberta
cford1@ualberta.ca

Benjamin V. Tucker

University of Alberta
benjamin.tucker@ualberta.ca

Tsuyoshi Ono

University of Alberta
tono@ualberta.ca

Voiceless nasal consonants are typologically rare in the world's languages. The present study investigates the acoustic realization of reported voiceless nasals in the Miyako Ryukyuan dialect Ikema. Voiceless nasals in Ikema occur word-initially and word-medially as part of a geminate or consonant cluster, and are phonemically distinct from modal voiced nasals. Initial observation of collected recordings revealed many instances of the voiceless phoneme with voicing throughout, leading to a re-evaluation of previous claims about its phonetic implementation. We hypothesized that word-medial and phrase-medial voiceless nasals surface as breathy voiced nasals. We analyzed the acoustic characteristics of nasal components of target words, focusing on duration, phonation state, and cepstral peak prominence (CPP), to determine whether reported voiceless nasal phonetic components with voicing are acoustically distinct from modal voiced nasal consonants. We find that voiceless nasals are produced with a voiceless component followed by a modal voiced component. Voiceless components and breathy components are found to be significantly shorter than modal components. We also find a significant difference between modal nasal, breathy nasal and voiceless nasal components' CPP values. The results confirm the observation that Ikema voiceless nasals are phonemically distinct from modal nasal consonants, and likely allophonically vary with breathy voiced nasals word-medially and phrase-medially. These findings align with the hypothesis that voiceless nasals require some voicing to be audible for perception, and are consistent with cross-linguistic findings, contributing to the typological understanding of the acoustics of voiceless nasals.

1 Introduction

Voiceless nasals are typologically rare phones employed contrastively in a handful of the world's languages (Ladefoged & Maddieson 2004). Acoustic characteristics of voiceless nasals have been analyzed in typologically diverse languages such as Burmese (Dantsuji 1986), Angami (Bhaskararao & Ladefoged 1991), Icelandic (Jessen & Pétursson 1998), and Romanian (Tucker & Warner 2010). One dialect of the under-documented language Miyako Ryukyuan, Ikema, has been reported to contain phonemic voiceless nasals, distinguishing it from other Ryukyuan languages and dialects of Miyako (Hayashi 2013). While

‘devoiced’ nasals have been described in the Miyako dialect Ōgami (Pellard 2009), their relationship appears to be allophonic with voiced nasals. Thus, studying Ikema provides a unique opportunity to observe the phonemic nature of voiceless nasals in the Ryukyuan language family.

Early observations of Ikema suggested that voiceless nasals may, in fact, be articulated with breathy voice. Thus, the present investigation seeks to determine (i) what acoustic characteristics distinguish voiceless nasals from voiced modal nasals in Ikema; and (ii) to what extent voiceless nasals are fully voiceless and how this may vary based on phonetic context. Secondary to these goals, as every voiceless nasal in Ikema appears to be obligatorily followed by a homorganic modal nasal, this study will also attempt to provide evidence regarding whether voiceless nasals form a geminate or consonant cluster with these following homorganic modal nasals. Investigating voiceless nasal durations and comparing those to modal voiced geminate nasals in the language may provide evidence to support or reject this hypothesis.

The paper is structured as follows. We provide a basic description and background of Miyako and Ikema, followed by a discussion of relevant literature involving voiceless nasals and breathy nasals. The elicitation methods and segmentation of data are then described, followed by analyses of duration and cepstral peak prominence. We conclude that Ikema voiceless nasals contain both a voiceless and modal voiced portion, such as those found in Burmese and Romanian. The voiceless portion is found to be significantly shorter than the modal portion. In addition, we propose that voiceless nasals and breathy nasals are allophones of the same phoneme in Ikema. These findings contribute to the broader hypothesis that voiceless nasals require some voicing in order to maintain sufficient amplitude to be perceptually distinguishable by listeners (Ohala & Ohala 1993).

2 Miyako Ryukyuan and Ikema

Miyako is a Ryukyuan language spoken on remote islands near Taiwan in the Ryukyu Archipelago, Japan. These islands have been politically part of Japan since the late 1800s, thus the modern population also speaks Japanese. Miyako is made up of a number of dialects although their mutual intelligibility is uncertain. The Ikema dialect of Miyako is spoken on Ikema Island, in the community of Nishihara on the northern part of Miyako Island, and in the community of Sarahama on Irabu Island (Figure 1). These communities are within relatively close proximity, having recently gained inter-island road access. At present it is unclear how this may impact Ikema subdialects, although communities are not known to frequently intermix according to our Ikema consultants. In total, an estimated 2000 people speak the Ikema dialect (Hayashi 2010), although this number is steadily declining as the majority of fluent speakers are over the age of 70; the younger generations are largely monolingual Japanese speakers. Accordingly, Miyako is considered ‘definitely endangered’ by the UNESCO *Atlas of the World’s Languages* (Iwasaki & Ono 2009). The present study thus provides a critical opportunity to document the phonetic realisation of voiceless nasals in this dialect. As Ikema continues to decline in use, documentation and analysis of such phenomena are vital as a means of preserving pieces of the language for heritage speakers. Additionally, exploratory analyses such as this provides a starting point for future researchers to continue phonetic documentation of Ikema and to contribute to our typological understanding of voiceless nasals.

While phonetic research on Ikema is limited, Hayashi’s (2013) grammatical sketch of Ikema includes an auditory description of the phonemic inventory, providing a strong foundation for the present study. It should be noted, however, that most of Hayashi’s speakers are from Nishihara, and thus her description may differ from the phonetic qualities observed in other Ikema communities. Although all three island communities speak the Ikema dialect, minor differences between communities have developed over time, further striating the

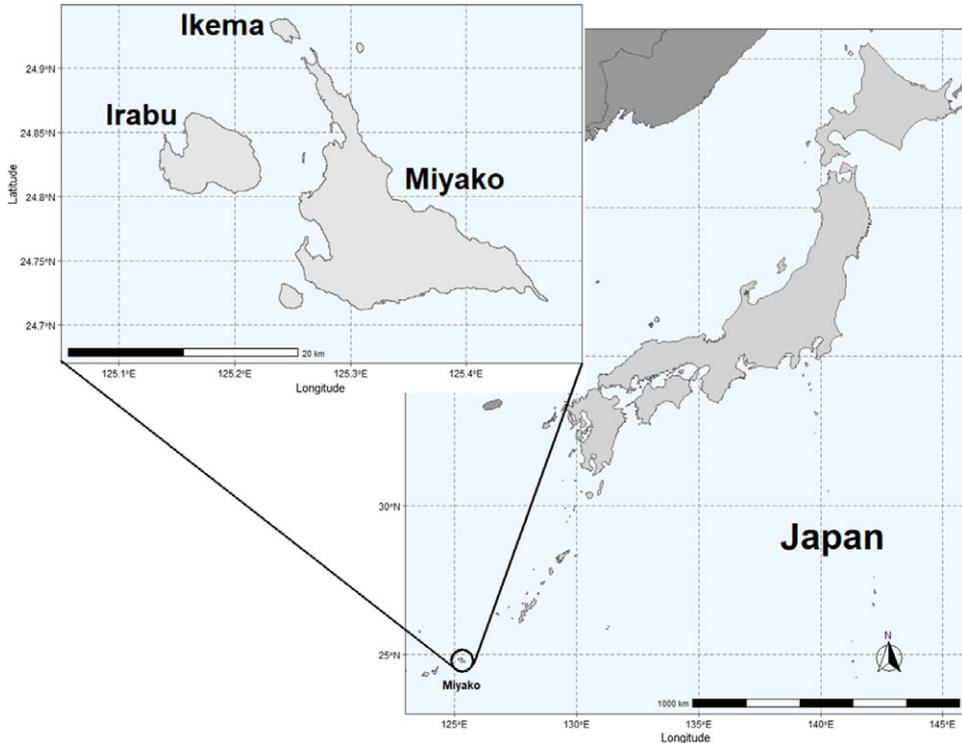


Figure 1 (Colour online) Map of Japan and the Ryukyu Archipelago, situated south of Kyushu, Japan; enlargement of Ikema, Irabu and Miyako islands. Created with *maps* package (Becker & Wilks 2018) in R (R Core Team 2016).

dialect. The differences are salient enough that Ikema Island community members can pinpoint subtle variants to their place of origin. For example the lexical item for ‘hiccup’ varies between communities: /saf:abi/ on Ikema Island and /s:abi/ in Nishihara. For the purposes of this study, all data was collected on Ikema Island, and may differ from Hayashi’s descriptions.

2.1 Ikema nasal consonants

Ikema is a moraic language with a CV structure, much like Japanese. Nasals occurring word-finally constitute a full mora (VN, e.g. [iN] ‘ocean’ is a two-mora word). Unlike Japanese, word-initial nasals can also occur before a homorganic stop consonant. In these cases the nasal is counted as a full mora (NCV, e.g. [ŋgi:] ‘to pull out’ is a three-mora word) (Hayashi 2013). While confirmation of moraic vs syllabic structure in Ikema is limited, an investigation of voiceless nasals by Shinohara & Fujimoto (2018) found mora count to correlate with word duration more reliably than syllable count, providing reasonable evidence for moraic timing in Ikema.

The phonemic nasals in Ikema include the bilabial /m/, alveolar /n/, and their voiceless counterparts /m̥/ and /n̥/. The alveolar and bilabial nasals occur before vowels and homorganic stop consonants, providing substantial evidence that these are distinct phonemes, as demonstrated in (1a–d). Example (2) illustrates minimal pairs of the voiceless nasals and their voiced counterparts. Word-medial examples of voiceless nasals can be found in (3). The velar nasal [ŋ] is also present within the language but only occurs preceding velar stop consonants [k] and [g], as illustrated in (4).

- (1) a. /mi-i/ → [mi:]
see-INF ‘to see’
b. /ni-i/ → [ni:]
boil-INF ‘to boil’
c. /mbu-munu/ → [mbumunu]
heavy-thing ‘heavy object’
d. /nta/ → [nta]
ground ‘ground’
- (2) a. /ᵐmi-i/ → [ᵐmi:]
step on-INF ‘to step on’
b. /m:i-i/ → [m:i:]
ripen-INF ‘to ripen’
c. /ᵐna:/ → [ᵐna:]
rope ‘rope’
d. /n:a/ → [n:a]
snail ‘snail’
- (3) a. /s:am̩mi-i/ → [s:am̩mi:]
go numb-INF ‘go numb’
b. /muzi-ᵐm/ → [muziᵐᵐ]
husks-step on ‘step on husks’ (process in harvesting wheat)
- (4) /ŋgjamasi/ → [ŋgjamasi]
noise ‘noisy’

The transcription of voiceless nasals followed by modal nasals throughout this paper is both for continuity with the Ryukyuan linguistics community (e.g. Hayashi 2013; discussions with Ikema speakers) and based on what is observed in spectrograms. It seems the voiceless nasal is obligatorily followed by a homorganic modal nasal, as no voiceless nasal has been observed without a significant voiced nasal portion following it; described as ‘half voiced geminates’ by Shinohara & Fujimoto (2018), which is discussed in greater detail in the analysis and results. At present, very little is known about Ikema voiceless nasals, and voiceless nasals typologically, and therefore it is unclear whether the modal nasal is a separate phoneme or a portion of the preceding voiceless nasal. Rhythm analysis by Shinohara & Fujimoto (2018) assumes voiceless nasals form a geminate with the following voiced modal nasal; however the authors reach the conclusion that both /n:/ and /ᵐn/ represent two mora, each nasal accounting for one mora. Under this analysis it would seem the nasals best represent a consonant cluster not uncommon with nasals in Ikema where clusters can be formed with voiced and voiceless stops as well. The current convention leans towards calling these

nasal constructions ‘geminates’ despite little confirmation of this categorization. For the purposes of continuity, this paper will also refer to the construction as a ‘geminate’, although this will be re-evaluated in the discussion.

In the absence of minimal pairs, [ŋ] is best judged as an allophone of /n/ (Hayashi 2013). Based on discussions with Ikema consultants, there is also likely the uvular nasal [ɴ] appearing as a word-final allophone, although this has yet to be confirmed with articulatory or acoustic data. Example (3b) presents a voiceless nasal realized as [ɴɴ] based on observations by an Ikema speaker.

2.2 Voiceless nasals

The phonetic characteristics of voiceless nasals have been widely discussed in the phonetic literature. Researchers have disputed whether they are completely voiceless and if the acoustic and aerodynamic evidence available allows interlocutors to perceive potential variants phonemically. Analyses of voiceless nasals in Romanian and Burmese reveal that voiceless nasals are comprised of both voiceless and voiced portions (Dantsuji 1986, Tucker & Warner 2010). Ohala & Ohala (1993) argue that without this voiced portion, place of articulation is likely difficult to portray acoustically due to aerodynamic principles of the nasal cavity. Unlike the oral cavity, the nasal cavity is lined with a dynamic layer of mucus which dampens acoustic signals, and thus may not enable productions with high enough amplitude for voiceless nasal place of articulation perception. Fully voiceless nasals have been described in languages such as Icelandic; however it is unclear whether these nasals are phonemic or allophones of voiced nasals (Jessen & Pétursson 1998, Hoole & Bombien 2010). Additionally, in southern dialects the voiceless nasal seems to be necessarily followed by a homorganic voiceless stop consonant, which may be providing much of the place cue for listeners (Jessen & Pétursson 1998).

Initial observations suggest that voiceless nasals in Ikema are much like those documented in Burmese, which Dantsuji (1986) described as having an initial voiceless ‘friction’ portion followed by a voiced ‘nasal murmur’ leading into a vowel, likely allowing for easier perception of sounds differing by place of articulation, as has been claimed by Ohala & Ohala (1993). These two components allow for phonemic variation, the former distinguishing sounds from their voiced counterparts, and the latter making each place of articulation perceivable (Bhaskararao & Ladefoged 1991).

However, Bhaskararao & Ladefoged (1991) suggest that voicing need not be present for place of articulation to be perceived. Analyses of the three voiceless nasals in the Tibeto-Burman language Angami shows all productions are entirely voiceless while phonemically distinguishable by speakers. This contradicts Ohala & Ohala (1993), who claim that voicing necessarily occurs in these sounds because of the relatively low level of turbulence created through the nasal cavity, making purely voiceless nasals difficult to perceive. Aerodynamic analysis of Angami voiceless nasals, however, shows a combination of oral and nasal airflow throughout. With joint oral-nasal airflow, speakers are better able to vary the degree of friction created than may be possible with nasal airflow alone, potentially allowing for accurate place of articulation perception by listeners. It is also possible perception is aided by the following vocalic context present in words tested, as each target nasal was followed by a vowel into which nasal airflow persisted. Regardless, the Angami findings suggest that there are multiple ways voiceless nasals may be articulated and perceived cross-linguistically.

Ohala & Ohala (1993) claim that differences in perception are further highlighted by studies in other areas of linguistics, including historical linguistics, child language acquisition and phonetic documentation. They further claim that there is a close relationship with voiceless nasals and fricatives, and that voiceless nasals can be classified as [–sonorant] with their voiced variants largely accepted as [+sonorant]. This phonological proposition centers on the typical environment in which voiceless nasals are found, often clustered with the voiceless fricative /s/. An example of this can be found in Romanian, where nasals are

allophonically devoiced after /s/, as in /bas_m/ ‘fairy-tale’ (Tucker & Warner 2010). Children learning English have been reported to replace /sn/ and /sm/ consonant clusters with voiceless nasals of the same place of articulation (Greenlee 1973), suggesting a strong perception of the voiceless turbulent airflow from the /s/ segment in these clusters. Based on these findings, it would appear that turbulent airflow is a perceptually salient property of voiceless nasals that may be realized by speakers through a variety of mechanisms. It should also be mentioned, however, that studies discussing the perception of voiceless nasals such as those by Ohala & Ohala (1993) used articulatory and acoustic data as opposed to perceptual data. To our knowledge, no true perception study has been conducted on voiceless nasals, leading to the conclusion that little is known about how speakers perceive these phonemes, and what is acoustically required for accurate perception.

Our preliminary observations of spectrographic characteristics which inspired the present study suggested that word-medial voiceless nasals are produced with voicing throughout, and word-initial voiceless nasals are often produced with voicing phrase-medially. The nasals in question appear to be only truly voiceless when produced phrase-initially. Despite this, there is reason to believe that these nasals differ from modal nasals present in the language. Firstly, speakers are aware of their differing articulation from the modal nasal, one speaker describing the sound as ‘not quite /n/, but not quite /ŋ/ either’. Spectrograms also reveal differences between modal nasals and the target sound; the target sound is produced with considerable noise, as will be demonstrated in the Methods section. Because of the observed voicing throughout these productions, however, this sound cannot be considered a voiceless nasal. One other possibility is that speakers may be manipulating phonation state and producing breathy nasals in these contexts.

Cross-linguistically, previous investigations suggest that breathy nasal and voiceless nasal airflow is quite similar. The Tibeto-Burman languages Sumi and Angami have phonemic nasal contrasts for phonation state (Harris 2010) and voicing (Bhaskararao & Ladefoged 1991), respectively. Airflow analysis of Sumi breathy nasals and Angami voiceless nasals are both marked by increased oral and nasal airflow compared to modal nasals in either language. Based on these findings, it seems likely that the only major difference between voiceless and breathy nasal articulation is vocal fold vibration while maintaining higher airflow than a modal nasal. Thus, theoretically these two nasal phones could easily be in allophonic variation. We investigate the possibility of this variation within Ikema.

3 Method

3.1 Participants

Six speakers (five male, one female) over the age of 56 participated in this study. One male speaker had to be excluded from the analysis, as he seems to have lost the voiceless nasal phonemic distinction, likely from Japanese influences having lived away from the community for over ten years. The remaining four male speakers were born on Ikema Island and grew up speaking Ikema. Their ages range from mid-50s to mid-70s. The female speaker was in her 60s and was raised by her grandmother from Nishihara and may have some dialectal variation in her pronunciation as a result.

3.2 Procedure

Recordings were conducted on Ikema Island in December 2015, either at the local community centre or in the speaker’s home depending on their physical mobility. When able, recordings were made using a Countryman Associates Inc. E6 Isomax head-mounted microphone. Two speakers found the head-mounted microphone uncomfortable, in which case a Sennheiser System K6 ME66 microphone was used with a tabletop microphone stand. While the use

of two microphone setups may have influenced acoustic data collection, speaker comfort was determined to be the priority in the collection of this data. In order to control for possible variation in microphone frequency response patterns which may impact measures of interest, microphone type was considered in the analysis. A Marantz PMD660 recorder was used to make digital recordings (44.1 kHz and 16 bit) of the speech. To capture more natural productions, speakers were asked to produce sentences using the target word. Following the sentence elicitation, speakers were then asked to produce each word in isolation. Elicitation sessions were conducted in Japanese, asking speakers to translate from Japanese to Ikema in order to avoid influence from the researcher's attempted pronunciation of Ikema. In total, 26 target words were elicited from each speaker (13 voiceless nasal words, 13 minimal/near-minimal pair words). The word list can be found in the Appendix. Occasionally speakers were unable to remember a word or produced a different word than expected. Thus, not all 26 words were collected from each speaker. Speakers were encouraged to produce multiple productions of targets as naturally as possible. Thus, in order to avoid list intonation or unnatural speech rates, no specific repetition guide such as 'repeat the word ten times' was given. This did lead to unbalanced numbers of each target word but ultimately allowed for additional data to be gathered. In total, 285 words with voiceless nasals and 248 words with modal nasals were collected for analysis.

3.3 Acoustic analysis

Utterances were segmented by hand using Praat (Boersma & Weenink 2017). Nasal segments were separated from the following vowel based on where antiformant structure ended and vowel formant structure began. The waveform was used to verify this segmentation as nasals generally have a simpler sinusoidal waveform than vowels. Nasals were then further segmented based on voicing; segmentation differed based on the nasal's word position. Word-initial voiceless nasal portions were segmented from the beginning of the nasal formant, which is typically visible just above the voicing bar (e.g. Figure 2A), to the start of voicing. Examples using /m̥mi/, /m̥mu/ 'to wear (shoes)' and /m̥:ui/ 'ripen' in both phrase-initial (in isolation) and phrase-medial positions can be found in Figure 2A and 2B respectively. While segmenting recordings, it was discovered that almost all word-medial voiceless nasals appeared to have voicing throughout, examples of which can be found in Figure 3. This was also the case for word-initial voiceless nasals produced within a phrase (e.g. Figure 2B). However, a significant decrease in amplitude and increase in noise compared to modal nasals, much like voiceless nasals produced in isolation or phrase-initially, was still apparent in expected regions (i.e. the initial portion of the nasal). These target phonetic components with voicing throughout were presumed to be breathy, and were measured as the section of the nasal where amplitude significantly drops based on observations of the waveforms. In the few cases where the voiceless and following modal portion could not be distinguished, the two portions were segmented together. Based on this process, three categories of nasals were created for analysis: voiceless nasals, breathy nasals (which are phonologically voiceless), and modal nasals (which are phonologically modal).

These segmentations were used to determine: (i) whether voiceless nasals as a phoneme are fully or partially voiceless, (ii) phonation state differences between nasal segments, (iii) whether phrase position influences the phonation state of nasal segments, and (iv) whether the durations of nasal segments differed based on phonation state. The terms 'segment' and 'portion' are used throughout this paper interchangeably to refer to phonetic components of utterances and should not be interpreted as an application of phonology. It should be noted that, as voiceless nasals seem to be obligatorily followed by a modal nasal there are no word-final or phrase-final productions.

Once segmented, a Praat script was used to extract acoustic measures on duration and phonation state. We extracted segment duration to compare voiceless, breathy and modal

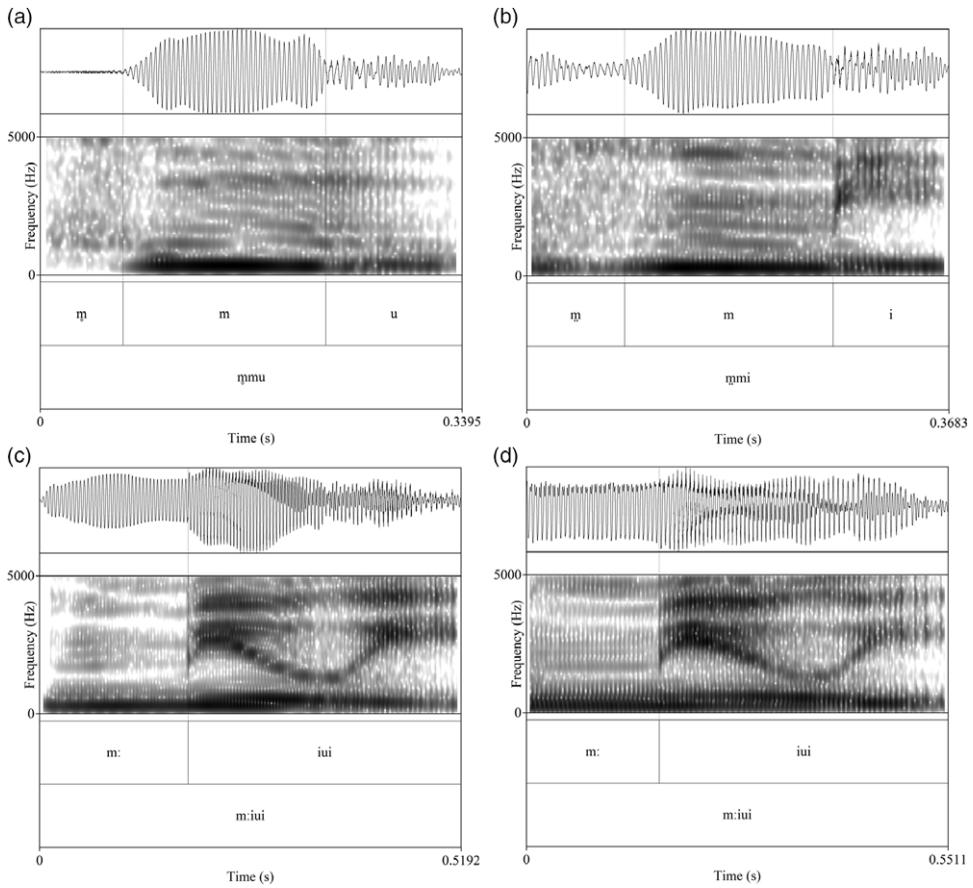


Figure 2 Spectrograms from a single speaker of /ŋmu/ (A) produced in isolation and /ŋmi/ (B) produced phrase-medially, both meaning 'to wear (shoes)', /ŋmi/ as the command form of /ŋmu/. Spectrograms (C) and (D) are of /m:iui/ 'ripen' in continuative form produced in isolation (C) and phrase-medially (D).

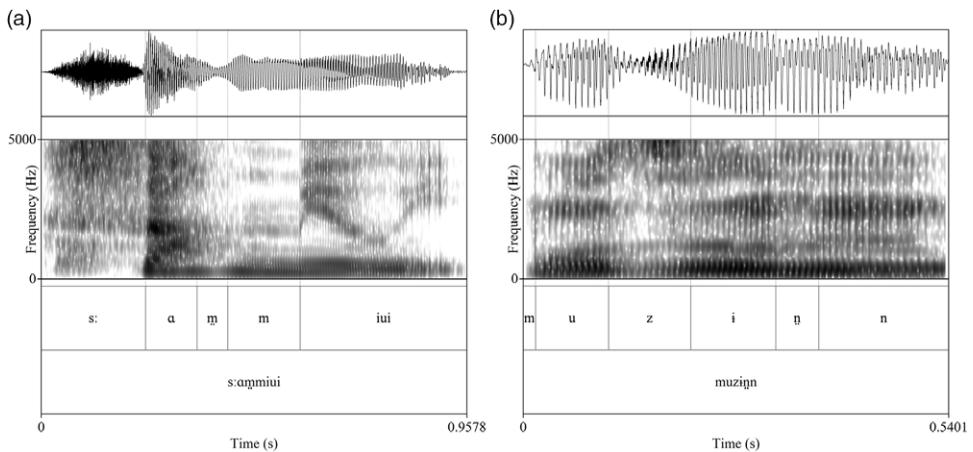


Figure 3 Spectrograms of /s:ŋmiui/ 'to fall asleep (limb)' (A) in continuative form, and /muziŋn/ 'step on (wheat)' (B), both produced in isolation.

nasals. Although speaker variation is likely present in the data, raw duration, rather than normalized duration, was used due to varying prosodic structures between target words. Speaker variation is thus accounted for during the analysis, and is described in detail in the Results section.

The amplitude measure cepstral peak prominence (CPP) was used to measure differences between phonation states (Hillenbrand, Cleveland & Erickson 1994). Breathy voice is marked by an increase in glottalic airspace during voicing vibration cycles, which causes an increase in aspiration noise. CPP assesses the regularity of harmonic peaks; the more regular and high amplitude harmonics are, the higher the CPP value. Thus, modal voicing should have a higher CPP than breathy voicing, and breathy voicing should have a higher CPP than voicelessness. The CPP measure has been shown to be an acoustic correlate of breathy voice (Hillenbrand et al. 1994, Hillenbrand & Houde 1996) and was demonstrated by Samlan & Story (2011) to mark perceptually different phonation states between breathy and modal vowels. Using the kinematic model ‘Tube Talker’ (Story 2005), Samlan & Story (2011) demonstrated vocal fold shapes and kinematics that cause the perception of breathy voice, and used CPP as their acoustic measure. CPP decreased as the separation between vocal folds increased, aligning with the above hypothesis regarding modal, breathy and voiceless phones.

4 Results

Voiceless and breathy nasal data is analyzed to confirm the validity of the methodological approach to segmentation and phoneme categorization. Following this, a series of linear mixed-effects logistic regression models are used to determine whether duration and cepstral peak prominence can distinguish voiceless, breathy and modal nasals, and whether voiceless nasals are best judged as geminates consisting of a voiceless portion followed by a modal portion.

4.1 Categorization of voiceless and breathy nasals

Over the course of the analysis thus far, an assumption was made that all phonemically voiceless nasals that did not have a significant voiceless portion, based on visual analysis of spectrograms, should be categorized as breathy. In order to avoid the potential for false positives in the results, a statistical analysis was conducted comparing voiceless nasals based on phrase position, regardless of how they were produced. It was originally hypothesized that a phonemically voiceless nasal would be realized as breathy phrase-medially, and voiceless phrase-initially. This preliminary analysis investigates the validity of this hypothesis, categorizing nasals based on phrase position alone. Table 1 gives a breakdown of nasals based on phrase position (occurring initially or medially) compared to the original breathy and voiceless categorization. While the division isn’t entirely categorical, the nasals tend to be realized as voiceless phrase-initially and breathy phrase-medially. This is proportionally more the case for voiceless nasals, where only 4.5% of voiceless nasals occur phrase-medially.

To confirm whether the previous nasal categorization based on phrase position aligns with the phonation characteristics of the recorded nasals, the CPP values of target nasals

Table 1 Subset of voiceless nasal productions, categorized based on phrase position and observed phonation state in spectrograms.

	Phrase-initial	Phrase-medial
Breathy	17	68
Voiceless	126	6

in different phrase positions were compared. Due to the amount of data in this subset analysis (using only those nasals labelled as ‘voiceless’ or ‘breathy’), a *t*-test was chosen, as any other statistical analysis would be unreasonable for the size of the sample. Results indicate significantly higher CPP values, or more regular harmonic peaks indicative of breathy voice, for phrase-medial nasals ($M = 19.39$, $SD = 3.15$), than phrase-initial nasals ($M = 18.06$, $SD = 3.10$), $t(146) = -2.96$, $p = .0036$). These initial findings validate the visual categorization used in the segmentation process, supporting the data analysis that follows.

4.2 Duration

Duration data for the three types of nasals (modal, breathy, and voiceless) are represented in Figure 4. Modal nasal segments are longer than voiceless or breathy segments, with a mean duration of 143.6 ms, almost double that of both the breathy and voiceless segments (73.9 ms and 63.8 ms respectively). The geminate constructions have a mean duration of 225.3 ms and voiced modal geminates have a mean duration of 190.9 ms.

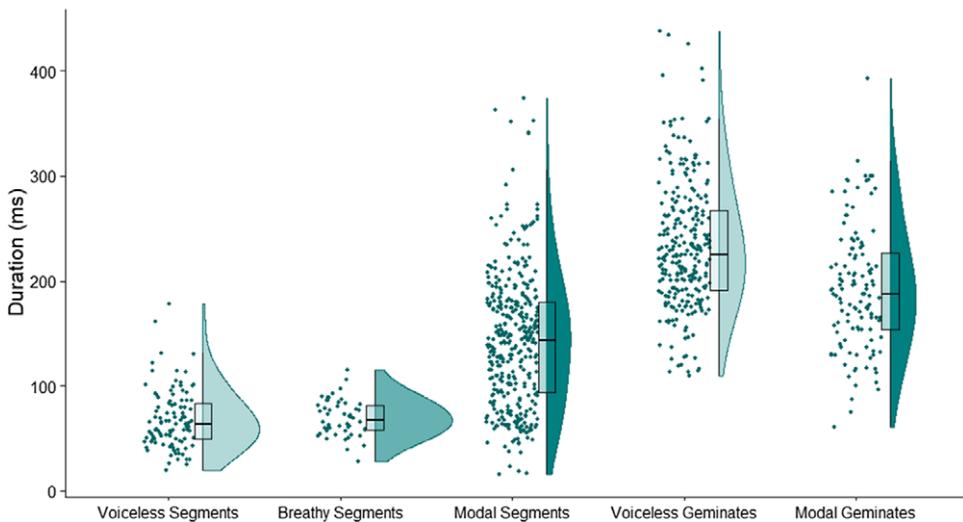


Figure 4 (Colour online) Raincloud plot of duration of voiceless segments / $\text{ŋ} \text{ŋ}/$, breathy segments / $\text{ŋ} \text{ŋ}^h/$, modal segments / $\text{m} \text{ n}/$, ‘voiceless geminates’ which are the voiceless/breathy segments followed by a modal segment / $\text{ŋm} \text{ ŋn}/$, and modal geminates / $\text{m} \text{ n}:/$.

A linear mixed-effects logistic regression model calculated with the *lme4* package (Bates et al. 2017) in R (R Core Team 2016), was used to determine whether observed duration could be explained by different phonation states of nasal segments. Duration was the dependent variable; Phonation State (voiceless segment, breathy segment, voiceless/breathy geminate, modal segment, and modal geminate) was the independent variable of interest and Mora Count was included as an additional fixed effect to control for possible rhythm timing influences on segment duration. Speaker and Word were included as random effects. There was not enough data in the analysis to support random slopes in the models. The *emmeans* package (Lenth 2020) was used to verify effects with a Bonferroni correction and no effects changed. The model, with voiceless segment as the intercept, indicated that voiceless segments are significantly shorter in duration than all other phonation states analyzed. Mora count did not significantly vary with duration. In a second model with breathy segment used as the intercept, breathy segments were found to be significantly shorter than modal segments

Table 2 Estimated coefficients, their standard errors, and *t*-statistics according to linear mixed-models fitted to Duration, with Phonation State as main predictor and Mora as a controlled fixed effect. An asterisk indicates a significant comparison.

	Estimate	Standard Error	<i>t</i> -statistic
Intercept (voiceless segment)	32.202	23.971	1.343
Breathy segment	17.652	7.662	2.304 *
Modal segment	113.840	5.134	22.171 *
Voiceless/Breathy geminate	185.936	5.163	36.012 *
Modal geminate	174.130	13.532	12.868 *
Mora	-0.318	4.989	-0.064
Intercept (breathy segment)	49.855	24.433	2.040
Voiceless segment	-17.652	7.662	-2.304 *
Modal segment	96.188	6.749	14.252 *
Voiceless/breathy geminate	168.284	6.811	24.709 *
Modal geminate	156.478	14.159	11.051 *
Mora	-0.318	4.989	-0.064
Intercept (voiceless/breathy geminate)	218.139	23.710	9.200 *
Voiceless segment	-185.936	5.163	-36.012 *
Breathy segment	-168.284	6.811	-24.709 *
Modal segment	-72.097	3.614	-19.947 *
Modal geminate	-11.806	12.969	-0.910
Mora	-0.318	4.989	-0.064

as well. When the model was relevelled to use voiceless/breathy geminates as the intercept, all phonation states except modal geminates were found to be significantly shorter. There was no significant difference between the duration of voiceless/breathy geminates and modal geminates. Coefficients for the model can be found in Table 2. We note that Bates et al. (2015: 34) indicate that any method for approximating degrees of freedom for linear mixed-effects regression is ‘at best *ad hoc*’ and thus we have not included *p*-values.

4.3 Cepstral peak prominence

The CPP data are plotted in Figure 5. Following our prediction, we see that CPP values are highest for modal nasals ($M = 22.5$). By comparison, CPP is lower for breathy nasal segments ($M = 19.3$), and lowest for voiceless nasal segments ($M = 17.3$). Breathly segments have higher overall CPP values than voiceless segments, suggesting the phonation state of these two groupings does indeed differ.

Linear mixed effects regression models were calculated to determine whether observed CPP could be explained by different phonation state was the dependent variable, Phonation State (Voiceless, Breathly, Modal) was the main predictor, and Microphone Type (Countryman Associates Inc. E6 Isomax or Sennheiser System K6 ME66) was included as a fixed effect to control for any variation due to microphone differences. Speaker and Word were included as random effects. As before, there was not enough data in the analysis to support random slopes in the models. The results of the analysis are reported in Table 3. The *emmeans* package was used to verify effects with a Bonferroni correction and no effects changed. The model with voiceless segment as the intercept indicated that voiceless segments have significantly lower CPP values than modal or breathly segments. In a second model, breathly segments had significantly lower CPP values than modal segments, but had significantly higher CPP values than voiceless segments. A higher CPP indicates more regular high amplitude harmonics. Thus, it fits the prediction that modal nasals have the highest CPP values while voiceless

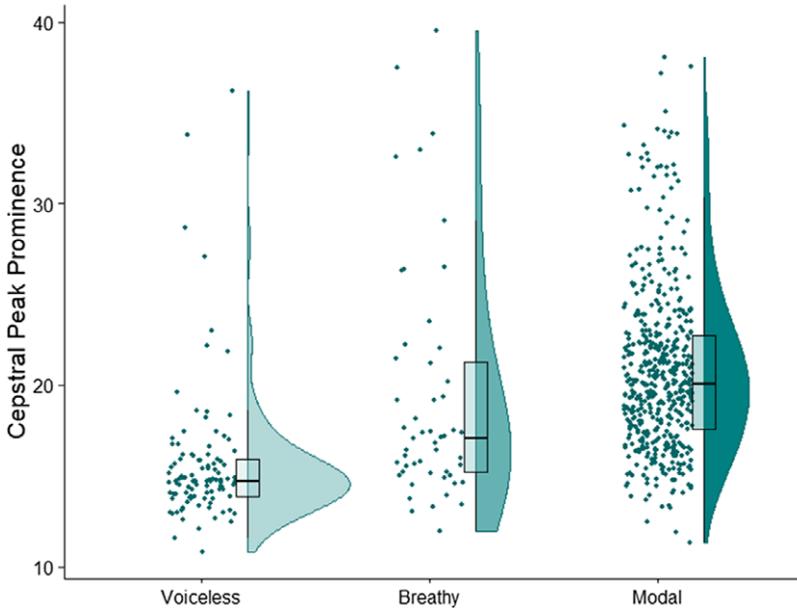


Figure 5 (Colour online) Raincloud plot of CPP data for voiceless nasal segments, breathy nasal segments, and modal nasals.

Table 3 Estimated coefficients, their standard errors and associated *t*-statistics according to linear mixed-models fitted to CPP, with Phonation State as main predictor. An asterisk indicates a significant comparison.

	Estimate	Standard Error	<i>t</i> -statistic
Intercept (voiceless)	15.430	0.806	19.137 *
Breathy	4.408	0.082	5.495 *
Modal	6.037	0.529	11.403 *
Microphone	-1.302	1.023	-1.273
Intercept (breathy)	19.838	0.950	20.891 *
Voiceless	-4.408	0.802	-5.495 *
Modal	3.268	0.415	7.883 *
Microphone	-1.302	1.022	-1.273

nasals have the lowest CPP values; breathy nasals have CPP values that fall between these two groups. Microphone was not significant in either model, suggesting that the microphone used in each recording did not impact CPP measures.

5 Discussion

The present study investigated the acoustic characteristics of voiceless nasals in Ikema and how these may differ from modal nasals in the language. The findings in this study contribute to the typological study of nasals cross-linguistically, as voiceless nasal productions in Ikema share characteristics with voiceless nasals found in other languages, such as containing a voiceless segment followed by a voiced segment. Table 4 provides a comparison of voiceless nasal productions cross-linguistically. Of note, Angami is the only language with confirmed phonemic voiceless nasals that are realized as fully voiceless. As Ikema voiceless nasals always occur in voiceless + modal nasal constructions, we conclude that the modal portion

Table 4 Realizations of voiceless nasals cross-linguistically.

Nasal production	Languages	Phonological status	References
Voiceless portion + voiced portion	Burmese	Phonemic	Dantsuji 1986, Chirkova, Basset & Amelot 2019
	Ikema dialect of Miyako	Phonemic	Shinohara & Fujimoto 2018
	Kham Tibetan	Phonemic	Chirkova et al. 2019
	Romanian	Allophonic with modal nasal	Tucker & Warner 2010
	Xumi	Phonemic	Chirkova et al. 2019
Fully voiceless	Angami	Phonemic	Bhaskararao & Ladefoged 1991
	Icelandic	Unclear	Jessen & Pétursson 1998
	Ogami dialect of Miyako	Allophonic with modal nasal	Pellard 2009

is likely used to aid perception, aligning with the hypotheses presented by Ohala & Ohala (1993).

Further, the present study also reveals that the Ikema voiceless nasal phoneme is generally voiceless in word-initial position and breathy voiced in word-medial position. Visual inspection of spectrograms shows a clear decrease in amplitude and an increase in noise, typical of voicelessness and breathy voicing (Gordon & Ladefoged 2001). This is also reflected in the CPP values, exhibiting regular high amplitude harmonic peaks for modal nasals, and significantly lower amplitude and lower regularity of harmonic peaks for voiceless nasals. Voiceless nasals produced phrase-medially have a significantly higher CPP than those produced phrase-initially, providing strong evidence for a breathy nasal which occurs phrase-medially.

At present Miyako is the only language we are aware of to have been reported with this type of variation, although it seems possible for Tibeto-Burman languages where either the breathy or voiceless nasal have prevalence. Overall, these results provide evidence that one of the major production cues of voiceless nasals is simply increased airflow, a characteristic predominant in breathy phonation state as well. Confirmation of this hypothesis, however, would require oronasal airflow analysis. We also suspect that breathy realization of these word-medially is a possible solution for languages with nasal voicing distinctions without sacrificing perceptibility, along the lines of Ohala & Ohala (1993). Thus, Ikema breathy nasals may in part be used to aid listeners, as truly voiceless nasals are likely difficult to perceive due to their decreased amplitude. However, as a reviewer pointed out, based on this argument we would expect the breathy allophone to occur word and phrase-initially as well, although the results suggest this happens infrequently. Arguably a more reasonable explanation centers around ease of production. Based on Lindblom's (1990) theory of speech production and perception, speakers balance discriminatory needs of the listener with economical hypoarticulations during productions. Ceasing voicing intervocalically is a more effortful articulatory movement for the speaker than maintaining voicing. Thus, allowing voicing to continue by producing a breathy nasal in continuous speech, even if that nasal is in word-initial position, eases articulation while maintaining sufficient distinguishability for listeners. This is supported by our results, where only 4.5% of voiceless nasals occurred phrase-medially, indicating that the preferred allophone intervocalically is the breathy nasal where voicing is maintained. Intervocalic voicing of typically voiceless consonants isomonly seen in other languages, particularly in spontaneous speech where phonetic reduction occurs more frequently (e.g. Hualde, Simonet & Nadeu 2011, Torreira & Ernestus 2011). As a result the breathy nasal is useful from both production and perception standpoints in Ikema, and is phonetically distinct from the modal nasal.

Similarly, voiceless nasals are likely obligatorily followed by modal nasals in Ikema to improve perceptibility. This hypothesis is supported by findings in a variety of languages, where voiceless nasals contain a voiced portion (see Table 4). It has been proposed that this partial voicing helps distinguish place of articulation and overall perception of the nasal itself (Ohala & Ohala 1993). This also resembles what has been found with breathy nasals in Sumi, which are reported as containing a breathy portion and a modal portion (Harris 2010). Although production data seems to abound, this hypothesis has yet to be confirmed by perception data. While aerodynamic and articulation theories provide a logical basis, without speech perception data we will not fully understand how listeners perceive voiceless nasals and what cues are relevant for successful perception.

As a secondary goal, the present study contributes evidence to help determine whether voiceless nasals form a geminate with their following homorganic modal nasal. In all productions the first portion of the construction is voiceless or breathy and the second portion is modal. Thus, voiceless/breathy segments do not seem to be realized without modal voicing in Ikema, aligning with other languages with nasals, such as Burmese (Dantsuji 1986). Duration analysis indicated that voiceless/breathy + modal constructions are not significantly different from modal geminates. However, analysis by Shinohara & Fujimoto (2018) concluded both voiceless and modal geminates (/ŋŋ/ and /nn/) constitute two mora, one per nasal segment. These results would suggest the two nasals form a consonant cluster. Under a cluster construction, it is likely that a portion of the modal nasal is part of the voiceless nasal due to theoretical rhythmic requirements. Mora timing assumes that each mora is approximately equal in length (Warner & Arai 2001). Based on our duration analysis, where voiceless and breathy portions alone are almost two times shorter than following modal segments, voiceless/breathy segments are likely too short to be considered a full mora on their own and could possibly violate the rhythmic requirements of Ikema, unless we assume voiceless nasals in Ikema invariably contain a modal component. Yet, an investigation into moraic timing in Ikema has yet to be conducted, and moraic timing in Japanese has been shown to break down in spontaneous speech (Warner & Arai 2001). Therefore, it is unclear how timing may play a role in these productions and durations of voiceless/breathy and modal components of these ‘clusters’.

In order to find stronger evidence of whether Ikema voiceless nasals form a geminate or consonant cluster with the following modal segment, additional rhythmic and phonological investigations would need to be conducted. One possibility could be a mora counting study where Ikema speakers count out mora in predetermined words and show how they counted. If voiceless nasals are truly geminates, a word such as [muziŋŋ] ‘the process in harvesting wheat of stepping on husks’ would be counted as three mora, the nasals counting as one mora together. Something less direct such as haiku formation, where the three lines of the poem must coincide to five-seven-five mora, could be used to gain a similar understanding of the voiceless nasal timing unit. A preliminary investigation suggests the voiceless nasal and modal nasal are counted as two separate mora. Haiku, however, are strictly Japanese poems, and thus an Ikema language game may be more appropriate.

It is also possible that historically there may have been a voiceless nasal singleton that progressed over time into a mixed-voice construction. Under this explanation, speakers may have recognized the two portions of the voiceless nasal, and slowly lengthened the modal portion leading to a reanalysis of the segment as a geminate or consonant cluster. While not historical evidence, speakers do display awareness of the voiceless and modal portions of their productions orthographically. Although Miyako and its dialects are traditionally oral languages, some speakers have begun to use Japanese kana to represent Ikema. During field work when speakers were asked to write the voiceless nasal, many said there is not an appropriate kana for the sound. This has led speakers to use multiple syllabaries to represent the voiceless nasal or to develop a new kana outside of the Japanese writing system. This voiceless representation is always followed by a typical Japanese kana for the modal nasal, reflecting the possible modern lengthening of voiceless segment-modal segment productions.

For example, while the primary Japanese syllabary, hiragana, has been used by speakers to represent Ikema sounds, the voiceless nasal is sometimes written in katakana, the syllabary typically used to represent foreign words (e.g. [ŋna:] ‘rope’ written as ンナー). Without historical data, however, this hypothesis is impossible to confirm. Additionally, the authors are unaware of similar historical sources for modern voiceless nasals. However, a study with historical data along the lines of Pellard & Hayashi (2012) may be illuminating. It appears that voiceless nasals are at risk of being lost from the language in the future, as speakers continue to modify their productions, or lack there of, of voiceless nasals.

As in many projects concerning endangered languages, there are limitations of the research process that are difficult to avoid. Namely, the need to use the dominant language, Japanese, in elicitation sessions could impact speakers. Japanese is known to have an influence on the way in which speakers use Miyako, as in a natural discourse setting speakers often mix the two languages (Nakayama & Ono 2013). Additionally, as a foreigner asking for productions of these words, speakers may have over-articulated to accommodate to their non-native listener. There is also a chance that speakers may avoid articulating the voiceless nasal phoneme because they are aware it is difficult for non-native speakers to perceive. To avoid this, speakers were asked to produce sentences with the target words spontaneously in this study, hoping that rapid speech would minimize these effects. It is unclear to what degree this methodological strategy aided natural productions. While it is assumed having a researcher present during any speech production task has an impact on speakers, it is difficult to determine how this presence may specifically impact indigenous language production and what may mitigate possible artificialities.

Future directions include targeted investigations of perception and how these sounds surface in spontaneous speech. To confirm the validity of hypotheses regarding the perceptibility of voiceless nasals and whether breathy nasals are truly perceived as allophones, a dedicated speech perception analysis is needed. Studying the target sound in spontaneous speech would also give more insight into the nature of the allophonic variation between breathy and voiceless nasals, specifically whether the voiceless nasal surfaces regularly in any context other than in careful speech phrase-initially. Furthermore, analysis of spontaneous speech could give a greater understanding of how voiceless nasal constructions with modal nasals manifest in the language. Finally, a study comparing voiceless nasal productions between Ikema subdialects of Ikema Island, Nishihara and Sarahama may provide additional evidence with regards to the questions addressed in this study.

6 Conclusion

The present study investigates voiceless nasals in Ikema and furthers our understanding of voiceless nasals cross-linguistically. We found that voiceless nasals are only partially voiceless in Ikema, similar to those in other languages. Ikema’s voiceless nasal is articulated as a mixed-voicing geminate or consonant cluster, with a voiceless portion followed by a significantly longer modal portion. Modal voicing may be used to aid perception of place of articulation for these phones, which may have led to the voiceless nasal becoming a mixed-voicing construction in Ikema. Additionally, we found that the voiceless nasal appears as both voiceless and breathy in fairly predictable contexts. The voiceless nasal surfaces as a breathy nasal intervocalically, leading to the conclusion that the voiceless nasal has an allophonic variant. The breathy allophone likely surfaces as a coarticulatory effect while still maintaining more noise and airflow than modal nasals in Ikema. The results presented here deepen our understanding of acoustic properties of voiceless nasals in speech, and how endangered language contexts may contribute to further acoustic variation.

Acknowledgements

We would like to thank the community on Ikema Island for their hospitality during our time on the island, their endless patience answering questions, and contributing to the exploration of their fascinating language. We would also like to thank members of the Alberta Phonetics Laboratory for giving feedback at various stages of this project, as well as the reviewers for providing commentary that improved the quality of the manuscript as a whole.

Appendix. Wordlist

Voiceless nasal	Meaning	Japanese	Minimal pair	Meaning	Japanese
/ᵐmi:/	draw (e.g. water) put on (shoes) step on (it)	<i>kuminasai</i> (<i>kutsu o hakinasai</i>) <i>fuminasai</i>	/m:i:/	to ripen	<i>jukusuru</i>
/ᵐmu/	cloud to draw (water) to put on (shoes) to step on	<i>kumo</i> <i>kumu</i> (<i>kutsu o haku</i>) <i>fumu</i>	/mu:/	seaweed	
/ᵐna:/	rope	<i>tsuna</i>	/n:a/	snail	<i>sazae</i>
/ᵐnagi:/	to attach to	<i>tsunagu</i>	/n:a:gja:/	not yet	<i>mada</i>
/ᵐnu/	horn yesterday	<i>tsuno</i> <i>kinou</i>	/n:uzi:/	type of octopus	<i>iidako</i>
/ᵐnta:gu:ta:/	to march	<i>hageshiku fumu</i>	/n:ta/	nut	<i>kinomi</i>
/ᵐndi/ ¹	yes	<i>hai, soudesu</i>	/n:di/ ¹	yes	<i>hai, soudesu</i>
/ᵐmpadʒ:a/ or /ᵐnhadʒ:a/	to step wrong, miss footing	<i>fumihazusu</i>	/m:pa/	no	<i>iie</i>
/ᵐmbiki:/	step on (it)	<i>fumu</i>	/mbjai:/	endure	<i>taete</i>
/s:aᵐmi:/	go numb	<i>shibireru</i>	/sanmin/	understanding	<i>rikai</i>
/f:ᵐmu/	storm cloud	<i>kurokumo</i>	/nfumunu/	warm	<i>atatakai</i>
/muziᵐm/	process in harvesting wheat of stepping on husks	<i>mugifumi</i>	/jamaziᵐm/	form into piles	<i>sanseki</i>
/ʃ:iᵐm/	Clamming	<i>shiohigari</i>	/baʃ:in/	do not forget	<i>wasurenai</i>

¹Both /ᵐndi/ and /n:di/ are accepted productions by Ikema Island speakers for 'yes'.

References

- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using *lme4*. *Journal of Statistical Software* 67(1), 1–48.
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2017. Package *lme4*. R package version, 1.1-13.
- Becker, Richard A. & Allan R. Wilks. 2018. Package *maps*. R package version, 3.3.0.
- Bhaskararao, Peri & Peter Ladefoged. 1991. Two types of voiceless nasals. *Journal of the International Phonetic Association* 21, 80–88.
- Boersma, Paul & David Weenink. 2017. Praat: Doing phonetics by computer (version 6.0.28) <http://www.praat.org/>.
- Chirkova, Katia, Patricia Basset & Angélique Amelot. 2019. Voiceless nasal sounds in three Tibeto-Burman languages. *Journal of the International Phonetic Association* 49(1), 1–32.

- Dantsuji, Masatake. 1986. Some acoustic observations on the distinction of place of articulation for voiceless nasals in Burmese. *Studia phonologica* 20, 1–11.
- Gordon, Matthew & Peter Ladefoged. 2001. Phonation types: A cross-linguistic overview. *Journal of Phonetics* 29(4), 383–406.
- Greenlee, Mel. 1973. Some observations on English initial consonant clusters in a child two to three years old. *Papers and Reports on Child Language Development, Stanford* 6, 97–106.
- Harris, Tom. 2010. Phonation in Sumi nasals. *Australian Speech, Science, and Technology Association conference proceedings*, 14–16.
- Hayashi, Yuka. 2010. Ikema (Miyako Ryukyuan). In Michinori Shimoji & Thomas Pellard (eds.), *An introduction to Ryukyuan languages*, 167–188. Tokyo: Research Institute for Languages and Cultures of Asia and Africa.
- Hayashi, Yuka. 2013. *A grammar of Southern Ryukyuan Miyako dialect Ikema*. Ph.D. dissertation, Kyoto University at Kyoto, Japan. [林由華 (2013) 「南琉球宮古語池間方言の文法」 京都大学大学院文学研究科博士論文.]
- Hillenbrand, James, Ronald A. Cleveland & Robert L. Erickson. 1994. Acoustic correlates of breathy vocal quality. *Journal of Speech and Hearing Research* 37, 769–778.
- Hillenbrand, James & Robert A. Houde. 1996. Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *Journal of Speech and Hearing Research* 39, 311–321.
- Hoole, Philip & Lasse Bombien. 2010. Velar and glottal activity in Icelandic. In Susanne Fuchs, Phil Hoole, Christine Mooshammer & Marzena Żygis (eds.), *Between the regular and the particular in speech and language*, 171–204. New York: Peter Lang.
- Hualde, José I., Miquel Simonet & Marianna Nadeu. 2011. Consonant lenition and phonological recategorization. *Laboratory Phonology* 2(2), 301–329.
- Iwasaki, Shoichi & Tsuyoshi Ono. 2009. Ikema Ryukyuan: Investigating past experience and the current state through life narratives. *Japanese/Korean Linguistics* 19, 351–364.
- Jessen, Michael & Magnús Pétursson. 1998. Voiceless nasal phonemes in Icelandic. *Journal of the International Phonetic Association* 28, 43–53.
- Ladefoged, Peter & Ian Maddieson. 2004. *The sounds of the world's languages*, 2nd edn. Malden, MA: Blackwell.
- Lenth, Russel V. 2020. *emmeans*: Estimated Marginal Means, aka Least-Squares Means. R package version, 1.4.8. <https://CRAN.R-project.org/package=emmeans>.
- Lindblom, Björn. 1990. Explaining phonetic variation: A sketch of the H&H Theory. In William J. Hardcastle & Alain Marchal (eds.), *Speech production and speech modelling*, 403–439. Dordrecht: Kluwer.
- Nakayama, Toshihide & Tsuyoshi Ono. 2013. Having a shinshii/shiishii ‘master’ around makes you speak Japanese! Inadvertent contextualization in gathering Ikema data. In Elena Mihas, Bernard Perley, Gabriel Rei-Doval & Kathleen Wheatley (eds.), *Studies in language companion series*, vol. 142, 141–156. Amsterdam: John Benjamins.
- Ohala, John & Manjari Ohala. 1993. The phonetics of nasal phonology: Theorems and data. *Phonetics and Phonology* 5, 225–249.
- Pellard, Thomas. 2009. *Ōgami: Éléments de Description d'un Parler du Sud des Ryūkyū*. Docteur de l'ehess, École des Hautes Études en Sciences Sociales, Paris, France.
- Pellard, Thomas & Yuka Hayashi. 2012. The phonology of the Miyako dialects: Phonological systems and comparisons. In Nobuko Kibe (ed.), *General study for research and conservation of endangered dialects in Japan: Research report on Miyako Ryukyuan*, 13–55. Tokyo: National Institute for Japanese Language and Linguistics.
- R Core Team. 2016. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Samlan, Robin A. & Brad H. Story. 2011. Relation of structural and vibratory kinematics of the vocal folds to two acoustic measures of breathy voice based on computational modeling. *Journal of Speech, Language, and Hearing Research* 54, 1267–1283.

- Shinohara, Shigeko & Masako Fujimoto. 2018. Acoustic characteristics of the obstruent and nasal geminates in the Ikema dialect of Miyako Ryukyuan. In Elena Babsouli (ed.), *Crosslinguistic research in monolingual and bilingual speech*, 253–270. Chania, Greece: ISMBS.
- Story, Brad H. 2005. A parametric model of the vocal tract area function for vowel and consonant simulation. *The Journal of the Acoustical Society of America* 117(5), 3231–3254.
- Torreira, Francisco & Mirjam Ernestus. 2011. Realization of voiceless stops and vowels in conversational French and Spanish. *Laboratory Phonology* 2(2), 331–353.
- Tucker, Benjamin V. & Natasha Warner. 2010. What it means to be phonetic or phonological: The case of Romanian devoiced nasals. *Phonology* 27, 289–324.
- Warner, Natasha & Takayuki Arai. 2001. The role of the mora in the timing of spontaneous Japanese speech. *The Journal of the Acoustical Society of America* 109(3), 1144–1156.