UiT The Arctic University of Norway

Faculty of Humanities, Social Sciences and Education
Department of Language and Culture

The weight and representation of Ryukyuan Miyako onsets

Initial geminate moraicity, markedness, and sonority

Matias Myhre

Master's thesis in Theoretical Linguistics – LIN3990 – September 2021

Table of Contents

CHAPTER 1: Introduction

It is generally accepted that intervocalic geminates and codas may add to the weight of a syllable, much as the nucleus does. Meanwhile, the onset has typically been considered obligatorily non-moraic. The Ryukyuan Miyako languages appear to contradict this notion, as they offer evidence for a moraic representation of geminate onsets. To determine if this is the case, the mechanisms that determine which segments may form Miyako initial geminates must be described. The claim made in this thesis is that these geminates in fact match the predictions of modern iterations of Moraic Theory, but that language-specific constraint rankings and prominence effects causes the range of coerced moraic onsets to deviate from what is expected due to sonority. A major part of the analysis of the Miyako onsets herein is done with constraint-based phonology, which to my knowledge has only been done for these languages to a minor extent previously. In this thesis, I will therefore account for the representation of Miyako onset phonotactics in the Irabu and Ikema varieties of Miyako, which will entail an Optimality Theoretic (OT) account of these. The analysis herein draws on the transcriptions and analyses of other authors, though it is maintained that neither variety features any complex margins, contradicting a number of claims in the literature. This is argued on the basis of syllable structure constraints and economy. This thesis also proposes that certain initial geminates and apparent partial geminates in Irabu and Ikema are better understood as disyllabic clusters. As will be demonstrated in the following chapters, Miyako offers strong evidence for the Moraic Theory account of geminate representation and indicates that voicing plays a role in restricting the moraicity of onset segments.

In the present chapter, I will first provide background on the Miyako languages in terms of geographic distribution, number of speakers, and their status as endangered languages. Next, I will provide a general overview of the linguistic features of the Miyako languages as a whole. Finally, I will describe the transcription styles used herein, before outlining the structure of this thesis.

1.1 **Background of the Miyako language context**

Situated in the Pacific Ocean, the Miyako islands are typically associated with either Okinawa or Japan in terms of politics and culture. It is therefore important to provide some context for this perception. From Hirara, the most populated and urbanized region in the Miyako islands, the prefectural capital of Okinawa (Naha) sits approximately 280 kilometers to the northeast. From Naha, there is another distance of roughly 600 km to Cape Seta, the southernmost point of mainland Japan. Considering these distances, it is perhaps unsurprising that the languages of these islands are both different in terms of linguistic features and mutually unintelligible with their Japonic cousins, the Japanesic and Okinawan languages. Along with Yaeyama, Miyako is categorized as belonging to the Southern Ryukyuan subdivision of Ryukyuan, while Northern Ryukyuan broadly includes the Amami and Okinawan languages (Pellard 2015).

Within Japan, the Ryukyuan languages have traditionally been analyzed as dialects of Japanese (Shibatani 1990: 191), but they are now generally considered to belong to the Japonic family as a separate group of languages. Although the Ryukyus have undergone a campaign of assimilation starting with their annexation by Japan in 1897, the local languages were still dominant until in the late 1920s, with Japanese only becoming the established throughout Okinawa in the early 1940s (Yoshimura 2014). This assimilation was intensified in part due to the political conditions leading up to and surrounding the reversion of the Ryukyus to Japanese sovereignty in the 1970s. More broadly within this history, the primary language contact in Miyako has been with Japanese, with linguistic evidence indicating that the majority of contact prior to the annexation was likely between Proto-Ryukyuan varieties, rather than with other geographically adjacent languages from Indonesia and the Philippines (Shimoji 2010: 4). Language loss and language shift towards Japanese has greatly reduced both the number of spoken varieties of Ryukyuan languages and the number of fluent speakers of any variety.

UNESCO recognizes six endangered languages in the greater Ryukyu area. This categorization may inadvertently conceal the fact that considerably more languages are endangered or near extinction in this region. Shibatani (1990: 194) notes that Shuri Okinawan served as a Ryukyuan lingua franca. This is presumably the case in Northern Ryukyu, as a standard variety of any of the Southern Ryukyuan languages does not exist and likely never has (Heinrich et al. 2015: 1). Adding to this difficulty, "no Ryukyuan language was ever popularly employed for writing" (ibid.: 2). This means that while these languages have historically been transcribed using Japanese syllabary scripts, there has never been a general Ryukyuan lingua franca in writing. Fluent speakers of the Ryukyuan languages are today generally older speakers, with most research conducted using participants in their 60s or older. Seemingly without exception, the authors of the research that has been used as reference for this thesis have stated that their participants are bilingual speakers whose L1 is Japanese and whose L2 is a Ryukyuan language.

Jarosz (2014) estimates that there are roughly 10,000 – 15,000 native speakers of Miyako, with some of the most-described varieties (Ikema, Irabu, Ōgami) having only between 2000 and 150 fluent speakers. Heinrich et al. (2015) furthermore note that significant efforts are not currently being made to preserve the Miyako language as a native language in Miyako, and that the measures that do exist for this purpose are insufficient. The same conclusion is reached in Anderson (2014), who finds that the lack of intergenerational mother-tongue language transmission is the biggest obstacle to Ryukyuan language maintenance. Even so, Takubo (2021: 65-66) notes that recent studies "have found that people in their late thirties can understand the language and can be considered to be speakers with passive knowledge of Ikema". This suggests that there remains a base of native Miyako linguistic intuition, though the lack of younger speakers is a serious concern for the longevity of the languages.

That the Miyako languages are in such a precarious position adds some urgency to their research and analysis. Krauss (1992: 10) concluded an essay on the topic of language death as follows: “Obviously we must do some serious rethinking of our priorities, lest linguistics go down in history as the only science that presided obliviously over the disappearance of 90% of the very field to which it is dedicated”. Although a discussion of language revitalization or preservation is outside the scope of this paper, it must be acknowledged that each language is irreplaceable for both more and less pragmatic reasons. For this reason, the goal of this thesis is twofold. First, I will use the descriptive data concerning the Miyako languages to assess the underlying structure of initial geminates. Second, it is my hope that this attempt to do so will encourage further interest in, and study of, the Ryukyuan languages.

1.2 Overview of the Miyako languages

There are far more Ryukyuan languages than have been sketched, let alone formally described, and the same applies to Miyako. As of this writing, the Miyako varieties with the largest bodies of descriptive research in the literature are Irabu, Ikema, Ōgami, and Tarama. For reasons of practicality and the availability of data, these are therefore also the varieties that will be examined in this thesis. There are several Miyako traits that are cross-linguistically rare, not to mention highly unusual among the Japonic languages, such as the partially fricative vowel /ɿ/ and the peculiar syllabic consonant behavior found in Ōgami. The central focus of this thesis will however be the Miyako initial geminates with particular attention given to how these geminates are represented in the context of Miyako onset phonotactics.

Geminate onsets are also found in the Southern Ryukyuan Yaeyama languages, such as Tedumuni (Shinohara & Fujimoto 2011, Ogino & Harada 2015) and Hateruma (Aso 2015). Among the Northern Ryukyuan Okinawan languages to the north, they appear to be less common, though this perception may prove more or less accurate as more varieties are documented and formally described. The Okinawan variety Tsuken does feature a labialized geminate onset given as [kkw] in Matayoshi (2010), but this seems to be the only occurrence of an initial geminate in this language. Shuri Okinawan also permits a few initial geminates, with Shimoji (2012: 352) attesting the examples /kkwa/ [kkʷa] ‘child’ and /ccu/ [ttɕu] ‘person’. While geminates are moraic and contrastive in Japanese phonotactics, they do not appear to occur word-initially in any variety of Japanese.

While language data on the Ryukyuan languages has been compiled by a number of researchers, many of these efforts have understandably been focused on documenting and recording these languages afforded to their lexicons and syntactic behavior. Perhaps because of this, the cross-linguistic implications of the patterns found in Miyako have not been thoroughly examined. Furthermore, little work has been done on Miyako within constraint-based phonology, with a few OT analyses made for Ikema in Celik & Takubo (2014) and less theory-specific constraints suggested in

Takubo (2021). Another angle to be explored here will therefore be the implications of Miyako onsets for the Moraic Theory analysis of initial geminates, and to model the behavior of these onsets in an Optimality Theoretic grammar.

1.2.1 Structure of the Miyako languages

I will begin this section with a brief account of the phonemic inventories of the Miyako varieties. These languages largely share most of their vowel and consonant inventories. Ōgami is a significant outlier here and must be examined in greater detail later. Hayashi & Pellard (2012/2019: 29) describe the vowel inventories of the Miyako varieties as belonging to one of three classes:

- Four-vowel system: /a, i, u, ɿ/
  Ikema
- Five-vowel system: /a, i, u, o, ɿ/
  Shimajiri, Irabu, Uruka, Boru, Nobaru
- Six-vowel system: /a, i, e, u, o, ɿ/
  Kurima, Kugai, Karimata, Ōura, Yonaha

To this set, we may add Ōgami, which features a five-vowel inventory of /ɑ, i, ɛ, ɯ, u/[1] (Pellard 2010). Absent in Ōgami is the "fricative vowel" /ɿ/, which surfaces as the "somewhat fronted close central vowel [ɨ] ~ unrounded close back vowel [ɯ]" and is accompanied by alveolar friction, producing "voiceless [s] when it is preceded by a voiceless onset consonant (…) and voiced [z] when it is preceded by a voiced onset or no onset" (Hayashi & Pellard 2012/2019: 21). /ɿ/ accordingly has been analyzed as either a vowel with associated frication or as a syllabic consonant with a vocalic component. Hayashi & Pellard (ibid: 22) describe this sound as "a phoneme that has both a consonantal and a vocal quality", noting that proponents of /ɿ/ as a vowel "still agree that it is accompanied by a friction noise", while proponents of /ɿ/ as a consonant "still recognize its vowel-like ability to occupy a syllable nucleus". This is further problematized by the fact that certain Miyako varieties, such as Irabu, permit broad ranges of syllabic consonants. While further phonetic and acoustic study of /ɿ/ is needed, it is apparent that this segment is underlyingly moraic and features a vocalic component at the very least. While literature on this segment differs on whether this vocalic component is a central vowel or an apical vowel (see Pellard & Hayashi 2012/2019: 21 and references therein), it appears that likely that /ɿ/ indeed serves as a vowel in these languages. In the present thesis, this will therefore be assumed to be the case as we proceed to the consonants. The following general Miyako consonant inventory is modified from a list of consonants given in Pellard & Hayashi (2012/2019: 30):

[1] It should be noted here that Pellard & Hayashi (2012) state that /a/ and /e/ occur as [a~ɑ] and [e~ɛ] in the Miyako languages.

(1) Miyako consonant inventory template

	Labial	Alveolar	Palatal	Velar	Glottal
Plosives	p b	t d		k g	
Fricatives	f v	s z			h
Affricates		ts (dz)			
Nasals	m	(n̥) n			
Taps/flaps		ɾ			
Approximants	(w)		j		

Among the Miyako dialects discussed in this thesis, the segments that appear in parentheses in the template above are generally less common or occur in only one variety, while those without are generally present in all. All the dialects discussed here (except Ōgami) feature a voicing contrast in both fricatives and stops, with labial, alveolar and velar places of articulation as the main contrasts. Alternatively, these places of articulation may be analyzed as a contrast between labial, coronal, and dorsal consonants, respectively (Aoi 2015: 407). Additionally, each language examined here features a place contrast in the nasals between the labial /m/ and the alveolar /n/, but only Ikema appears to feature a contrast between the voiced /n/ and voiceless /n̥/. Ōgami deviates considerably from this template, as it lacks the affricate as well as any voicing contrast, thus featuring only the voiceless stop series /p t k/ and the voiceless fricatives /f s/ (Pellard 2010).

Non-contrastive palatalization of obstruents is also fairly common among the Japonic languages and is not directly represented in this table. However, descriptions of Miyako varieties such as Irabu (Shimoji 2008) and Ikema (Hayashi 2010) show the approximants /j w/ fusing with onset segments, which is analyzed therein as an underlying glide in either the onset or the nucleus. As underlying approximants such as /j/ and /w/ do occur as onsets where no other onset segment exists, they appear to belong to the phonemic consonant inventory. As the majority of dialects and variants of Miyako lack formal phonological or in some cases phonetic descriptions, the table above should not be considered an exhaustive index of the phonemes present across the spectrum of Miyako languages. I will therefore refer to this table as a template, noting (where relevant) which segments are specific to the dialect in question, as well as which segments are notably absent.

1.2.2 Miyako syllable structure

It can generally be stated that the Miyako dialects permit initial geminates. Pellard & Hayashi (2012/2019: 52) provide the following canonical syllable structures for Miyako:

(2)

a. $(C_1)(C_2)(j)V(V)(C_3)$

b. $(C_4)\ C_{,5}(C_6)$

In (2a) above, C_1 and C_2 must be a geminate "fricative or a resonant /s, z, f, v, m, n, ɾ/ or a partial geminate with /v, m/ as C_1" (ibid.: 53). The term "partial geminate" refers here to a consonant cluster in which C_1 shares, or assimilates to, the place of articulation of C_2. For example, in Ikema this means that although the C_1 and C_2 positions in the onset (the initial geminate) must generally be filled by a geminate consonant such as [tː] (assuming that this segment occupies two timing nodes), forms such as /nta/ [nta] 'mud', /nkyaan/ [ŋkjaːɴ] 'past times' still appear to be permitted (Hayashi 2010: 170). The exact structure of these words may not be monosyllabic, however. As is discussed in section 3.4 of this thesis, these words are likely syllabified as a syllabic singleton consonant followed by a simple CV syllable. This is supported by the fact that [nta], like the minimal word /nːa/ [nːa], are permitted under the bimoraic minimality constraint. In each of the Miyako languages examined here, all initial geminates appear to occur moraic, as evidenced by the above.

Pellard & Hayashi (2012/2019) rightly note that further research is needed to determine whether the syllable is a meaningful prosodic entity in Miyako. Very little has been written about this question in this particular language, though a similar discussion can be found in the literature for the adjacent language of Japanese. Labrune (2012) cites the Japanese linguistic tradition in arguing for Japanese being a purely mora-timed language, suggesting that there is no empirical reason to assume that the syllable exists in Japanese (and thus also that the syllable is not universal). Kawahara (2016) rejects this claim, citing phonetic and psycholinguistic evidence, as well as deeper theoretical issues that would arise if the syllable were simply dismissed from Japanese phonology[2].

While it is beyond the scope of this thesis to discuss whether the syllable exists as a linguistic universal, it has been established in the literature that Miyako shares similarities with the prosody of other Japonic languages in that it is primarily mora timed. There does not appear to be any particular reason to reject the existence of the syllable in Miyako. If Miyako prosody were to be analyzed as solely being active at the mora level with no significant prosodic effects at the syllable level, any investigation of constraints on initial geminates (such as in an Optimality Theoretic account) must necessarily regard the mora and the foot as its most salient domains. On its own, this is more or less as expected, as the Miyako bimoraic minimality constraint will be argued here to be the strongest evidence for moraic onsets in these languages. As is the case for Japanese, however, there are significant reasons to consider the syllable important to Miyako. Shimoji (2008, 2011), Pellard (2010), and Takubo (2021) identify syllable-specific processes in Irabu, Ōgami, and Ikema Miyako, respectively, indicating that a mora-only analysis would be quite difficult to support. Given the weight of evidence for the syllable as a linguistic universal, as well as the attested significance of this prosodic unit in Miyako, it will therefore be assumed that the syllable is a significant unit in Miyako.

[2] See also Kubozono (2003), Starr & Shih (2017) concerning the syllable as a unit in Japanese.

In the final section of this introduction, I will outline and justify the transcription style used in this thesis.

1.3 Notes on transcription

In this thesis, underlying representations are given in slashes // and surface representations are given in square brackets []. Where transliterated words are written, these occur in italics and occur before any other representation of the words in question. Where a gloss is provided, this will appear inside single quotes to the right of the underlying and/or surface form of the word. Examples used here will therefore comply with the following format, shown here using an example from Irabu (Shimoji 2008: 42):

(1) *mta* /m.ta/ [m̩.ta] 'mud'

In the above example, *mta* is a transliteration, /m.ta/ is the underlying representation, [m̩.ta] is the surface representation, and 'mud' is the meaning of the word in English. It bears mentioning that while many words featuring initial geminates are transliterated here using a double consonant, e.g. Ikema *tta* /tːa/ 'tongue', there is no romanization standard for Miyako. Ryukyuan languages are generally written using Japanese hiragana or katakana syllabary script, in which an initial geminate is represented by a sokuon (っ or ッ) preceding a consonant-vowel sequence, thus *tta* is either った or ッタ in hiragana and katakana, respectively. As orthography is not examined in any detail in this thesis, IPA is used throughout to refer to Miyako words.

Segment length will be transcribed using the IPA "hourglass" length marker [ː] appended to consonant segments to show gemination, i.e., /CːV/ [CːV]. This is for two main reasons. It is argued in this thesis that the Miyako initial geminates provide evidence supporting the hypothesis that geminates are long monopositional segments whose length is either a correlate of, or a consequence of their moraic status. This assumption proves to be necessary to capture the nature of the Miyako geminates. The choice to use [Cː] is also made for the purposes of simplicity, as it is immediately clear to the reader where I am referring to a geminate segment instead of a sequence of singleton segments. In cases where an example is directly quoted from another author, or it is believed that separate consonant segments are present (i.e., a non-geminate), multiple consonant segments will be written instead, i.e., /CCV/ [CCV]. Two consonants are also used in transcriptions of the Ikema partial geminates featuring the voiceless nasal /ɴ̥/, e.g. [m̥m-] [n̥n-]. This is done because the exact identity of these geminates remains unclear, and because transcribing this geminate with two segments illustrates its partial lack of voice.

Some mention should be made of the traditional transcription convention used in Japanese and Japonic linguistics for a geminate segment, which is Q. In the literature, Q has historically been

used to indicate the presence of an unpronounced consonant-lengthening phoneme or *sokuon* in Japanese (Kubozono 2013) and has previously been suggested as one of several "empty" phonemes in Japanese. Proposals for other empty phonemes in Japanese can be found in the literature, such as in Labrune (2014), where it is argued that Japanese /r/ is an empty consonant (see also Pellard 2016 in response). The Japanese nasal coda /ɴ/ is also frequently described as placeless, invariably assimilating to the place of articulation of a following consonant or surfacing as uvular [ɴ], though some research has indicated that this segment may in fact have a target (Yamane 2013).

These discussions are mentioned here to underline the fact that this thesis is primarily concerned with the grammars of Miyako. The absence of the Q analysis in this thesis should therefore not be interpreted as a statement on the underlying nature of Japanese geminates. Rather, as the Ryukyuan languages must be understood as a related yet distinct group of Japonic languages, it may not be necessary to assume that Q or other concepts from the tradition of Japanese linguistics must necessarily find application in Miyako. While this segment has also been used in Ryukyuan linguistics, such as in Sakiyama (2003) for the Tarama and Minna varieties of Miyako (in which it is argued to represent a syllabic glottal stop), there are good reasons not to use it here. The Q segment in Japanese linguistics is, like the geminate in moraic theory, assumed to be an inherently moraic segment that duplicates (i.e., lengthens) the following consonant. This segment appears superfluous when operating under the theoretical frameworks used here. There is no particular need for an empty phoneme to explain why a singleton-geminate contrast exists, as this is taken to be either lexically specified or the result of lengthening. We also do not require a Q phoneme to explain why this geminate is moraic, as geminates are assumed to inherently feature weight in Moraic Theory.

It therefore seems more parsimonious to suggest that a word occurring as [fːa] may underlyingly be /fːa/ (in which case moraicity is lexically specified) or /fa/ (in which case the onset undergoes lengthening), rather than assuming that a segment Q may be underlyingly present word-initially or may be added to comply with bimoraic minimality. While this would functionally provide the same surface representation, the Q analysis would predict a set of positional and segmental restrictions specific to Q-consonant sequences that could be explained in a more satisfactory manner by the Miyako initial geminates simply being segments linked to moras. In the following section, I will summarize this introductory chapter and provide an outline of the present thesis.

1.4 Summary and outline

To summarize the preceding section, it has been established that the Miyako languages feature largely similar consonant and vowel inventories, and that initial geminates occur in all the Miyako dialects discussed here. The range of syllabic consonants and initial geminate segments varies greatly from variety to variety, however. Irabu features a large inventory of possible syllabic consonants, allowing any continuant or nasal to fill this role, as does Ōgami, while the other Miyako languages tend to only

allow nasals to be syllabic consonants. A similar distribution can be found for the coda position, as Irabu again permits any continuant or nasal to form a coda, while Ōgami restricts coda position to a nasal or the approximant [ʋ]. Once again, Ikema and other Miyako dialects tend to only permit nasal codas. While it is stated in Shimoji (2008, 2011) that affricates may be syllabic and fill coda position in Irabu, this is argued not to be the case in chapter 3. The distribution of initial geminates will be explored in greater detail in chapter 3, but the simplified description above shows that Irabu and Ōgami initial geminates and syllabic consonants are nearly the same set of segments. Ikema, however, permits initial geminate stops, affricates, and fricatives in addition to its syllabic nasals.

In the following chapters, these facts will be explored in greater detail. In chapter 2, the theoretical assumptions that form the framework of this thesis are explained, starting with general phonological concepts and proceeding to a discussion of the controversial issue of geminate representation. This is followed by a discussion of Moraic Theory and specific iterations of this theory that allow for moraic onsets. The chapter concludes with a discussion of Optimality Theory and the data used in this thesis. Chapter 3 provides a detailed description of Miyako onset phonotactics, starting with Irabu and its initial geminates, followed by an excursus on the nature of the Irabu affricates, as well as an exploration of the possibility that the Irabu initial geminate voiced sonorant-fricative sequences are not, in fact, initial geminates. This section is followed by a description of Ikema's onset phonotactics, in which the claim that Ikema features onset clusters is militated against for reasons of syllable structure and sonority. The partially voiceless Ikema nasal geminates are also examined in some detail here. The final language examined in this chapter is Ōgami, which is analyzed as featuring fully tautosyllabic geminates, despite previous analyses given in the literature. Optimality Theory analyses are integrated into chapter 3 at the end of the sections on Irabu, Ikema, and Ōgami. Chapter 4 then consists of a summary and the conclusion of the thesis, as well as suggestions for future research topics and unresolved issues.

CHAPTER 2: Theoretical assumptions

In order to describe the Miyako patterns that are of interest to this thesis, it is first necessary to establish the theoretical assumptions that form the argumentation herein. As such, it will also be essential to clearly define what is meant by each of the questions discussed. It is assumed here that the reader is familiar with the most fundamental concepts and transcription standards in phonological theory. A greatly simplified summary nonetheless follows.

2.1 General assumptions

Speech sounds (phones) are categorized by how they are produced in the vocal tract, and are generally named as a combination of a place of articulation (e.g., the soft palate or alveolum) and a manner of articulation, e.g. a full occlusion of the vocal tract followed by release, abbreviated as a stop or

plosive. These speech sounds are understood to be stored in the language faculty of the brain as underlying representations (phonemes), with sounds produced at the output level as surface representations (allophones) conditioned by the phonotactics of the language in question and the environments in which they occur.

Segments are linked to higher prosodic units, which in turn determine the segmental structure of words and utterances in a given language. The two prosodic units that are of interest to this thesis are the syllable and the mora. The syllable provides a structure for each beat of an utterance and consists of an onset, a nucleus (or peak) and a coda. The mora (μ) is a timing unit taken to exist at a level between the syllable and the segments and links to heavier segments, which tend to be more sonorous or more prominent. Syllables with more morae are considered heavy, while syllables with fewer morae are light. The unit that most often bears moraic weight is the nucleus, which is typically a vowel. Syllables are generally at least monomoraic for this reason. Codas are also moraic in many languages, with coda weight often being determined by the sonority of the coda (Zec 1988).

In many cases, geminate moraic identity is argued for through prosodic behaviors such as stress, such as in Pattani Malay (Hajek & Goedemans 2003) or Rural Jordanian Arabic (Al-Deaibes 2021), or through word minimality, such as in Trukese (Davis 2017) or Moroccan Arabic (Noamane 2018). Stress, predictable or lexical, does not appear to play a role in Japonic phonology, while lexical pitch accent does not support an analysis in favor of, or against, the moraicity of Miyako geminate onsets. Instead, moraic identity among the Miyako languages has been determined through application of the strict requirement that all words must contain at least two morae, a phenomenon that is both cross-linguistically common (Hayes 1995) and very well-described in Japanese (see for instance Itô 1990). This requirement will hereafter be referred to as bimoraic minimality.

As stated previously, the specific behavior that is being studied in this thesis centers around the geminate onsets found in Miyako. These are of interest for a few key reasons. First, gemination in any position other than intervocalically (i.e., word-medially) is rare. Second, as will be described further, these onsets appear to support the proposition in Topintzi (2008) and Davis (2011) that true geminates are inherently moraic and are represented as single units at the segmental level, as this provides the most parsimonious explanation for the Miyako bimoraic C:V syllables and disallowance of complex margins. This also holds implications for the nature of moraic consonants, which have previously been argued to only occur as nuclei or codas (Zec 1988, Morén 1999). Thirdly, there appears to be a relation between the syllabicity of certain consonant segments in Miyako and their tendency to occur as geminate onsets. With this in mind, it will be instructive to describe exactly what is meant by *geminate* here.

2.2 Geminate representations and moraicity

2.2.1 Phonetic correlates of geminate segments

To account for the patterning of geminate segments in the Ryukyuan languages, it will be useful here to first outline what segments are to be understood as geminates. Geminates are often described as consonants with a "long or 'doubled' consonantal sound" (Davis 2011: 1). Kubozono (2017: 2) notes that phonetically, consonant length is "primarily signalled by consonant duration", going on to point out that while it is known that other phonetic features are involved, it is "not well known how different languages employ these other phonetic features and how much they differ in this respect". Geminate duration varies cross-linguistically and may feature more phonetic correlates in addition to, or instead of, length. Durational differences also occur within languages between different types of long segments. For example, Sato (1998) gives the ratio of singleton to geminate stop duration in Japanese as 1:2.03 - 1:2.44, with geminate fricatives featuring a lesser durational increase with a singleton-geminate duration ratio of 1:1.79 - 1:1.82. As duration is somewhat variable, it is understandable that geminates may also be identified as segments that occur "stronger or more intense i.e. fortis" (Thurgood 1993: 129), as well as changes to preceding and following vowels (e.g., shortening of the previous vowel), measures of voice quality, and pitch patterns (Idemaru & Guion 2008). Like duration, secondary cues may also be stronger or weaker depending on the segment in question (Mitterer 2018). In Miyako, however, the primary cue for continuant and nasal geminate identity is duration, while the primary cue for geminate stops is the duration of closure (Matsuura 2012/2019).

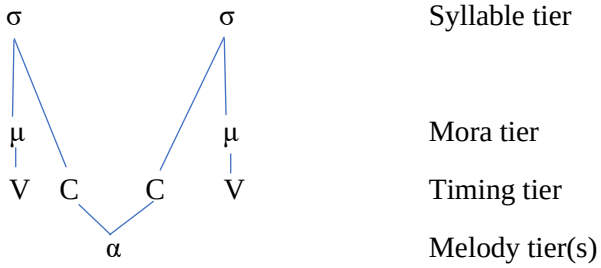
2.2.2 Phonological representation of geminate segments

This brings us to the phonological representation of geminate segments. Cross-linguistically, geminates are most often found intervocalically and thus word-medially (Thurgood 1993), (Ladefoged & Maddieson 1996). Intervocalic geminates are often easily analyzed as ambisyllabic, that is, as linking to the coda position of one syllable and the onset of the next. The exact representation of this linking has been somewhat controversial and varies depending on the theory being applied. As discussed in Davis (2011), geminates were represented in SPE phonology as [+long] segments (Chomsky & Halle 1968) and were later represented as segments linked to two C-tiers (McCarthy 1979), two X-tiers (Levin 1985) or two root nodes (Selkirk 1990). Davis goes on to discuss the representation of geminates in Moraic Theory (Hayes 1989), in which the geminate is a single segment that is underlyingly associated with a mora. To this list, we may also add the representation given in Ringen & Vago (2011), which essentially restates the CV-tier representation of the geminate without reference to a prosodic unit, linking a segment to two C-nodes. To keep this section relatively brief, this discussion can be summarized by stating that there are two popular methodologies used to describe geminates. One follows segmental length representation and stipulates

that the geminate is a segment linked to two timing nodes, while the other suggests that the geminate is a segment specified for moraic weight. The figure below illustrates the latter approach:

(3)

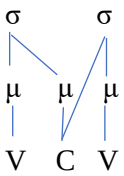
a. Ringen & Vago (2011: 156)



As is apparent in the above tree structure, this approach is similar to that of Selkirk (1990). As the only distinction between a geminate and a singleton is located at the melody tier (where phonetic material is located), this approach appears to predict that geminates should pattern with consonant clusters and not with single segments. Ringen & Vago (2011: 167-168) note however that while they claim that "with respect to quantity sensitive processes, all descriptions of single C geminates known to us are reanalyzable into CC representations", they are not claiming that geminates and consonant clusters necessarily pattern together, as the two are still structurally distinguishable. Conversely, they argue that geminates are unlikely to pattern with singletons. Following this representation, there is also no expectation that geminates should be underlyingly heavy, as there is no a priori assumption that moraicity and geminate identity are linked. A similar point of view can be found in Muller (2001: 75), who argues that there is "extensive empirical evidence indicating that prosodic weight is not a universal, uniform characteristic of geminates or coda consonants", highlighting languages such as Yawelmani in which there are apparent "inconsistencies" in mora assignment.

These analyses can be contrasted quite clearly with the syllable weight analysis of geminates. The representation given below is the "flopped" structure proposed by Davis (1999, 2011), in which the mora linked to the geminate is associated with the coda of the preceding syllable but not with the onset of the following syllable. Therefore, only the coda "portion" of the geminate is moraic:

(4)



Intervocalic geminates can thus be analyzed as consisting of a sequence of consonantal timing units syllabified as in (3), a segment linked to a mora unit at the coda position of the preceding syllable. It is of course necessary to note that neither theory exactly rules out the other's representation entirely.

Ringen & Vago (2011) make no attempt to argue that geminates may not surface as heavy or moraic, as is convincingly argued to be the case for Trukese (Davis 1999). Likewise, in Moraic Theory, it is not argued that two homorganic units never form a non-moraic sequence, but rather that this would be an example of a "fake" (underlyingly non-moraic) geminate contra a "true" (underlyingly moraic) geminate. False geminates would then be expected to pattern similarly to consonant clusters, as this is their structure. It should be mentioned here that geminate segments are generally assumed to not permit the insertion of a vowel, and this should be particularly unfeasible for a true geminate in Moraic Theory, as there should be no position to which a vowel may be inserted. False geminates, functioning essentially like consonant clusters, should allow for such phonological processes.

2.2.3 Initial geminate representations

Were all geminates medial, then speculations concerning the moraicity of geminates would likely be limited to the differences between geminates and medial coda segments. However, geminates are attested as occurring in both word-initial and in word-final position cross-linguistically. As such, initial gemination as seen in languages such as Cypriot Greek (Muller 2002), Trukese (Davis 2011), or Miyako (Kibe 2012/2019) serves as something of a challenge. Although word-initial geminates may be viewed as the exception rather than the rule, they are not as uncommon as this description would suggest. Kraehenmann (2011) notes that the cross-linguistic survey in Muller (2001) provides a non-exhaustive list of at least 29 languages known to contain initial geminates. Within this set, five languages (Leti, Ngada, Pattani Malay, Yapese, and Nyaheun) appear to *only* allow geminates to surface in the initial position, and not medially nor finally. Kraehenmann (ibid.: 4) also suggests that Sa'ban may also be included in this list, as initial geminates in this language greatly outnumber medial geminates. Twenty of the languages listed in Muller's survey feature both initial and medial geminates, while four contain geminates initially, medially and finally. As Kraehenmann points out, final geminates seem to imply the existence of medial and initial geminates, and no languages featuring only final geminates, or only initial and final geminates, appear to be attested. From this, we can tentatively state that while medial geminates are evidently the least marked cross-linguistically, it is nonetheless less marked to feature a geminate in word-initial position compared to word-final position.

This data presents us with an interesting problem. Moraic Theory as described in Hayes (1989) states both that the distinction between singletons and geminates can be understood as a distinction of underlying moraicity versus non-moraicity, but also that onsets cannot bear moraic weight. Final geminates do not appear to challenge this stipulation, as codas are generally recognized as potential sites of syllable weight. However, the initial geminates render this description of onsets a somewhat challenging position to maintain. Conversely, if geminates are not held to be underlyingly moraic, we should not expect to find languages like Miyako, in which geminates clearly satisfy a bimoraic minimality constraint.

Davis & Topintzi (2017) suggest that onset geminates may correctly be analyzed as moraic, and that the non-moraicity of onsets is instead due to the non-moraicity of singleton onsets, which are contrasted with geminate (and therefore moraic) onsets. In cases where geminates appear to surface without being linked to a mora, such as in the languages Selkup and Nglakgan, Davis (2003) proposes that Nglakgan geminates can be understood to be underlyingly moraic segments that occur as non-moraic in the surface representation due to specific constraint ranking issues in this language. In Selkup, where CVV is preferred for stress and CVC and CVC: syllables are ignored, Davis (2011) argues that "an independent constraint restricts pitch realization to vocalic elements", and that there would thus be no expectation for the geminate-closed syllable to attract stress. In other words, what Davis suggests in these arguments is that all geminates, including onset geminates, are underlyingly moraic, but that this moraicity may be absent or obscured in the output due to markedness. Topintzi (2008) builds on Davis's approach and suggests that weightless geminates are in fact represented as sequences of homorganic consonants in the underlying representation. In other words, a moraic onset following Topintzi is a consonantal segment linked to a μ node, while a weightless and therefore "fake" geminate consists of two homorganic segments with no link to a mora node.

Morén (1999) argues that weight can be present in long segments underlyingly (distinctive weight) or that segments may be non-moraic in the underlying representation and phonetically lengthened (coerced weight) due to factors such as (weight-sensitive) minimal word constraints, weight by position, constraints against word-final long vowels, and other phonological processes. This implies that underlyingly moraic geminates and codas surface as non-moraic in languages where moraicity is prohibited, while underlyingly non-moraic segments may also become moraic depending on language-specific constraint rankings. In a cross-linguistic survey, Gordon (2006) argues that weight-sensitive processes within languages do not show uniform sensitivity to segments, and that weight distribution differs by the process involved. For example, it is mentioned that languages in which vowel quality affects weight are problematic for moraic and skeletal slot models of weight, as there is no projection for the additional unit of weight (2006: 2). Gordon also cites languages with weight hierarchies exceeding the two expected categories, i.e., light and heavy. To explain the weight mismatches observed within languages, Gordon notes that:

> "The language specific choice in weight criteria for a given weight phenomenon is linked to language specific phonetic differences. Many, but not all, of these language specific phonetic differences can in turn be attributed to differences between languages in other aspects of the phonological system" (Gordon 2006: 244).

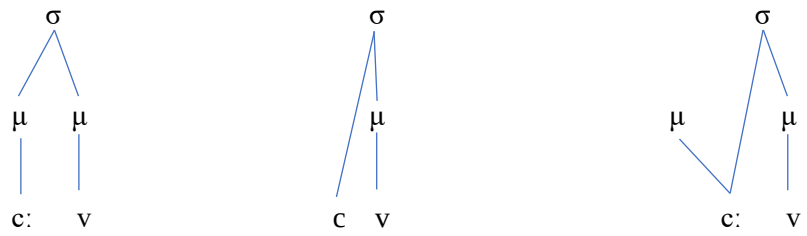
The same prominence-driven approach is given in Gordon (2005), where sensitivity to onsets is suggested to emerge from adaptation and recovery. Put very briefly, it is argued therein that onsets improve the perceptibility of the nucleus by providing a low-sonority, low-energy recovery phase

prior to the rime. In other words, onset weight may be assigned where the onset allows for a greater perceptual energy, given in Gordon (2005) as loudness or more broadly as phonetic "goodness" in Gordon (2006). Topintzi (2006: 28-31) directly critiques the prominence account of onset weight, noting most crucially that the perceptual energy-driven account of such weight does not account for the perceptual energy of the onset segment itself, and does not account for why the onset should behave differently from the coda with respect to recovery. As will become clear, the moraic account of onset weight appears to be quite applicable for the Miyako initial geminates.

To summarize, it has been established that geminates may occur initially, medially or finally, though they are most commonly found medially. The representation of geminates remains controversial, with the fundamental question being whether geminates are underlyingly moraic, and whether they are represented as a segment linked to multiple timing nodes or as a segment linked to a mora node. Per the approach suggested in Davis & Topintzi (2017), initial geminates may be analyzed as follows: in the initial position, a long (geminate) consonant C: simply links to a mora node μ. In the below figures, (5a) and (5b) illustrate a contrast between initial geminate and initial singleton identity. An alternate suggestion in Davis (2009), reiterated in Davis (2011), is that the edge geminate links to an extrasyllabic μ node. This is represented in (5c). C and V refer here to consonant and vowel segments, respectively:

(5)

(a) (b) (c)



As noted in Davis (2011), Topintzi's account of geminate weight as in (5a) allows Marshallese word-internal geminates to be analyzed as fully tautosyllabic moraic onset segments, as is proposed in Topintzi (2008). This suggests that geminates may be found to behave more similarly to singleton segments than previously believed. The distinction between (5a) and (5b) will be assumed in this thesis as the representation of Miyako onsets, as this best captures the near-minimal patterning of monosyllabic words with structures such as [CV:] and [C:V], without making additional stipulations about the extrasyllabic mora in (5c).

2.2.4 Geminate segment preferences

To determine whether any predictions can be made as to which segments in a language may form geminates, we must first account for which segments are attested as occurring as such. In a cross-linguistic survey of geminates, Thurgood (1993) notes that cross-linguistically, stops and nasals are

generally preferred over non-nasal approximants for gemination. Aoyama & Reid (2006) furthermore draw from this analysis to suggest that the alveolar place of articulation may be preferred for geminate-singleton contrast due to productions in this position being perceptually preferred in terms of phonetic quality. It should be noted, however, that Thurgood's data pertains to geminates of any type. As our interest here is in initial geminates, I will now briefly examine cross-linguistic data for initial geminates from Muller (2001), as analyzed in Kraehenmann (2011). There are a few key points to highlight here, which will be presented below. The following data, adapted from Kraehenmann (2011: 6-7), indicates how common each geminate is, cross-linguistically. Items on the left are more common, while items on the right are less common:

(6)

a. nn > tt > mm, kk > ss > pp > bb > ll > dd > ff, rr
b. stops > fricatives > nasals > liquids > glides > affricates
c. coronal > labial > dorsal > glottal > pharyngeal
d. voiceless stops > voiceless fricatives > voiced stops > voiced fricatives

Firstly, it should be noted that these are not implicational hierarchies, nor do they express any predicted universal ranking. Rather, these rankings are determined by the number of languages in the sample of 29 languages with initial geminates that feature specific initial geminate segments (6a), geminate segments by manner (6b), active place of articulation (6c), and finally voicing (4d). It must also be mentioned that the segments described in (6a) are intended to be understood as underlying forms, so other factors may apply to surface forms. As Kraehenmann (2011: 6-7) notes, these orders comply with what is expected from markedness: sonorants are coronal and voiced, while stops are voiceless. The presence of geminate fricatives implies geminate stops, while geminate affricates are rare and occur only in grammars with geminate stops and fricatives.

Based on cross-linguistic data, there does not appear to be any firm constraint against (or for) any particular geminate segment, but rather a strong tendency towards certain preferred candidates and implicational relationships. For example, Japanese in many ways aligns with hierarchies in (6) above, featuring long versions of voiceless stops at all places of articulation, as well as coronal voiceless fricatives and nasals. However, the nasals are not preferred as geminates in the native vocabulary, and there is evidence that nasal gemination is not a productive process in Japanese (Kawahara & Pangilinan 2017). It is also worth mentioning that while geminate voiced obstruents are highly marked and do not occur in the native phonology of Tokyo Japanese, they do occur in certain southern Japanese dialects such as Amakusa Japanese (Matsuura 2020). It must therefore be asked what processes govern the assignment of moraicity to consonant segments.

2.2.5 Distinctive and coerced geminate weight

More sonorous segments are expected to be permitted to occupy the nucleus position of a syllable and are more likely to be moraic when forming a coda. Zec (1988, 1995) finds that this is an implicational hierarchy, meaning that the least sonorous segment permitted to form a syllable nucleus or moraic coda implies that all more-sonorous segments must also be permitted to do so. Morén (1999, 2003) argues that this holds true for weight that is not specified in the input (coerced weight), but also that segments that are lexically specified as moraic (distinctive weight) may occur regardless of sonority. Distinctive moraicity may therefore cause less (or more) sonorous segments to be moraic in the rime on an arbitrary basis. In other words, distinctively moraic segments should be unpredictable and are not expected to form a natural class. While neither author considers the onset as a potential site of moraicity (both authors explicitly exclude the onset position) it is worth considering whether the sonority relation presented therein offers a possible explanation for the status of the initial geminate segments discussed here.

Naturally, it should not be assumed that restrictions on moraicity in onset position should mirror constraints on intervocalic moraicity. What we must then establish is which onsets are permitted to receive coerced weight, and whether the distinctive moraicity of onsets is similarly unbound by sonority constraints. Topintzi (2006, 2008, 2010) links moraicity to syllable well-formedness by way of sonority, noting that coerced moraic onsets are optimally less sonorous. The implicational hierarchy predicted in Topintzi is thus as follows (Topintzi 2006: 35):

(7)

Nucleus moraicity: *μ/PEAK/p,t,k >> … >> *μ/PEAK/e,o >> *μ/PEAK/a

Coda moraicity: *μ/CODA/p,t,k >> … >> *μ/CODA/e,o >> *μ/CODA/a

Onset moraicity: *μ/ONS/a >> … >> *μ/ONS/e,o >> *μ/ONS/p,t,k

In other words, the analysis here assumes that moraic onsets are subject to the exact opposite implicational hierarchy of that for moraicity in the syllable rime. The most marked moraic onsets should then be vowels, while the least marked should be the least sonorous, i.e., voiceless obstruents. If this is the case, it may follow that Morén's (1999, 2003) generalization that distinctive weight in the syllable rime does not correlate with sonority also applies to moraic onsets. Topintzi (2006: 32) argues that distinctive moraic onsets indeed occur regardless of sonority in languages such as Trukese, Pattani Malay, and others, while coerced (and thus sonority-driven) onset weight occurs in Pirahã, Arabela and Karo. According to this analysis, it must be predicted that coerced initial geminates in Miyako should be subject to a markedness hierarchy in which segments of higher sonority are more marked than segments of lower sonority. As opposed to overall sonority being the deciding factor,

however, Topintzi identifies the feature [voice] as the relevant distinction between the more marked and less marked initial geminates, summarized as follows (Topintzi 2006: 331):

> "the pitch perturbation caused by the voicing of consonants conditions the type of attested moraic onsets, always favouring voiceless over the voiced ones, unless onsets are underlyingly moraic, i.e. are geminates, in which case no such restrictions apply"

As such, it is primarily expected that voiceless segments should pattern together as the least marked candidates for coerced onset weight, just as the voiced segments are expected to pattern as the most marked. Sonorants behave somewhat unexpectedly under this analysis: They are assumed to either be specified as [voice], in which case they pattern with voiced obstruents, or they may be unspecified for voice, in which case they may pattern with the voiceless obstruents (Topintzi 2006: 102). Therefore, the constraint *μ/ONS/[voice] may or may not refer to sonorant segments on a language-specific basis. For Miyako, this raises some important questions. First, assuming that coercion occurs in Miyako initial geminates, are the Miyako moraic onsets optimally less sonorous, or is the meaningful distinction between voicing and voicelessness? Second, how do the Miyako sonorants behave in initial geminates? In the following section, I will provide a brief description of the constraint-based analytical framework that will be used to describe the Miyako geminates.

2.3 Optimality Theory (OT)

There has been a trend in phonological study since the inception of Optimality Theory (Prince & Smolensky 2004) to emphasize the emergent nature of grammars and their phonological phenomena. Moraic theory furthers this goal by establishing gemination as fundamentally a matter of weight, with non-moraic geminates emerging through constraint interaction. Consider, for instance, the proposed constraint *μ/ə for Piuma Paiwan (Shih 2018), which mandates non-moraic schwa in certain cases and not in others, even as the schwa is assumed to be underlyingly moraic.

When applied to phonology, Optimality Theory (OT) as a linguistic model suggests that phonological rules within a language are emergent from the rankings of universal constraints in that language. Barring certain pathologies, all speakers share the same set of mechanisms in OT. The Generator (GEN) produces input and output candidates. The Constraints (CON) are universally present (but may be differently ranked) in all speakers and are utilized by the Evaluator (EVAL) to determine the optimal output. In other words, OT follows from generative grammars in some key respects. The language faculty to which GEN, CON and EVAL belong is assumed to exist in all speakers as part of Universal Grammar (UG). It is also assumed that all linguistic output exists on an underlying and a surface level, and that underlying data (input) must in some way be parsed to occur as output.

What differentiates OT from, for instance, SPE phonology (Chomsky & Halle 1968), is its fundamental claim that the parsing of input to output is determined by constraints in competitive ranking relationships, and that that the process of parsing input to output is expected to occur on a single stratum. Specifically, this means that there are no sequential derivations or transformations in "standard" OT as described in Prince & Smolensky (2004) or McCarthy & Prince (1995). While some versions of OT, such as Stratal OT (Kiparsky 2000) utilize a model that posits a stratified evaluation of candidates on different lexical levels to deal with otherwise opaque processes, there is a strict emphasis on avoiding excessive derivation. In what is generally taken to be a parallel process, candidates are generated in GEN and selection is done in EVAL.

One of the advantages OT has over derivational models of output is that all input can be processed instantaneously. For instance, in languages where voiced stops are banned from word-final position, a derivational approach to an underlying voiced stop in this position, then the input must undergo derivation to get the devoiced or codaless output. This entails that more complex rules and more deviant inputs would necessitate longer and more complicated derivations, which should imply more cognitive effort to process. OT makes no such assumption, as simple inputs and complex inputs are effectively identical. All output candidates violate constraints, but OT attempts to find the "least bad" outcome of any parsing.

In OT, the underlying linguistic data (the input) is assumed to be completely unconstrained. This is known as *richness of the base* (Prince & Smolensky 2004) and is a fundamental assumption of OT. As such, only output candidates as these are produced by GEN can be weighed against CON, and not the input string. The implication of this is that an OT grammar should be able to generate candidates from any kind of input without selecting optimal outputs (optima) that are illegal in the actual language. To provide an example, this means that an OT grammar for a language that prohibits word-initial geminates encountering an input that contains a word-initial geminate must be able to select an optimal output that lacks this geminate. Likewise, we must expect grammars with a bimoraic minimality requirement to encounter monomoraic inputs and output optimal bimoraic candidates.

Concretely, OT posits that all candidate selection can be described through the mutual rankings of constraints. If constraint C1 outranks C2 (C1 >> C2), and C1 is undominated, then the optimal candidate may not violate C1, but may violate C2 if the candidate does not violate C1. The best possible candidate thus violates no constraints, while the worst possible candidate violates all constraints. All candidates should however be expected to violate some constraints while not violating others, as the opposition between the markedness of, and faithfulness to, the input is essentially what is considered to produce differences between languages. Constraints are typically formulated as either markedness or faithfulness constraints. Markedness constraints are violated by the presence or absence of a specific element in the output candidate. They include constraints against complex

syllable margins (*Complex, see Prince & Smolensky 2004) or against onsetless syllables (ONS). Markedness thus refers to both universal markedness, as becomes apparent form universal implicational hierarchies, and to language-specific markedness through language-specific rankings of markedness constraints. Faithfulness constraints are violated by differences between elements in the input and their corresponding elements in the output. Faithfulness can therefore be stated to be a correspondence relationship between the input and the output (McCarthy & Prince 1995). A typical example of a faithfulness constraint is the following from McCarthy & Prince (1995: 16):

(8) Maximality Input-Output (Max-IO):

Every segment of the input has a correspondent in the output.
(No phonological deletion.)

Max-IO is violated whenever a segment of the input has no correspondent in the output. This means that deletion incurs violation of Max-IO. An instantiation of a maximality constraint tends to be implied in OT grammars, as some explanation must be given for why marked elements or sequences are not simply omitted.

To conclude this discussion of OT, some mention must be made on the representation of OT that will be used here. Constraint rankings are typically represented in tableaux showing the input, the output candidates, and the relevant constraints. For reasons of practicality and space, each tableau does not contain all the constraints assumed to be present in the grammar. Similarly, only the most relevant output candidates are listed. The below table includes the constraint Uniformity (Prince & McCarthy 1995: 123), which restricts the coalescence of two or more input segments into one output segment. The following example uses the Irabu word /macja/ [matʃa] ‘little bird’ from Shimoji (2008: 39). /c/ here corresponds to the voiceless alveolar affricate /ts/:

(9) *Complex >> Max-IO >> Uniformity

/macja/	*Complex	Max-IO	Uniformity
ma.tsa		*!	
ma.tsja	*!		
☞ma.tʃa			*

In the above tableau, violations are indicated using an asterisk * and are followed by an exclamation point ! when this violation is fatal (disqualifying). The optimal candidate is indicated using a pointing hand ☞. The first candidate is disqualified due to violations of the constraint against deletion, while the second is ruled out due to its violation of the constraint against complex margins, here because of

a complex onset. The last candidate, [ma.tʃa], is optimal as the palatal approximant /j/ is realized in the output as palatalization of the preceding onset. While this candidate violates Uniformity, this constraint is dominated by the other two constraints in the tableau. Therefore, violations of Uniformity are permitted as long as higher-ranked constraints are not violated. Note that the table in (9) does not actually show a ranking argument for *Complex and Max-IO

While a wide range of extensions and revisions of OT have been made, such as Stratal OT (Kiparsky 2004), and probabilistic OT approaches such as partially ordered constraints (Anttila 2007), and Stochastic Optimality Theory (Boersma & Hayes 2001), the patterns explored in this thesis appear to be adequately accounted for using what could be considered "standard" OT. Or more accurately, what is examined here is not intra-speaker variation, but rather variations between different groups of speakers. Therefore, there is little need to posit probabilistic or serial constraint rankings. For reasons of space, I will not be examining whether these models may find application here except in the case of Ikema /z/ [z~dz]. One limitation of "standard" OT with exhaustive and fixed constraint ranking is that it is not well-suited to handling free variation, as a fixed constraint ranking should in theory generate the same output from the same input in all cases. Some modification is therefore necessary to account for Ikema /z/, which is in free variation between [z ~ dz].

To summarize this section, this thesis follows Topintzi & Davis (2017) in assuming that moraic onsets occur, and that the distinction between a moraic onset and a non-moraic onset is synonymous with the distinction between a geminate onset and a singleton onset, respectively. Furthermore, following Morén (1999, 2003) and Topintzi (2008) it is expected that the behavior of these geminate onsets must have some relation to either sonority or to voicing. This raises the question of what relation may be found in the actual data from the Miyako languages. Thus, the specific question this thesis asks is whether Miyako supports the proposal that initial geminates are moraic, whether these geminate onsets are best described by moraic theory, and whether the set of segments that may occur in these geminates pattern by sonority or voicing.

Furthermore, an OT approach will be utilized to describe the behavior of the geminate onsets in Miyako. To my knowledge, very little work has been done on OT analyses of the Miyako languages as of this writing. Takubo (2021) suggests a change in approach from derivational, rule-based analysis to constraint-based description of Ikema phonology, though this is not concretely given in terms of markedness or faithfulness or constraint rankings. Celik & Takubo (2014) represents a step towards identifying constraint rankings in Ikema that may account for some phonotactic patterns, but there is no discussion therein of onset moraicity. For this reason, an OT analysis of the Miyako onsets will provide new insights into the mechanisms of this group of languages, while also providing more data to understand the nature of geminates cross-linguistically. In the following section, I will briefly explain and justify the transcription styles that have been applied herein.

2.4 On the sources used

To my knowledge, no book has ever been published on an individual Miyako language in English or Japanese, with the exception of Hirayama (1967), as well as an Ōgami dictionary published by Hōsei University in 1977. There has been an increase in formal linguistic study conducted on the Miyako languages starting around the early 2000s, often in the form of bachelor's or master's theses (see for instance Ford 2016, Takei 2016), dissertations, and chapters in books about Ryukyu linguistics and the Ryukyuan languages more broadly. More recent works have also focused on language preservation and revitalization in the region, such as Heinrich & Anderson (2014).

For Irabu, extensive reference has been made to the research by Michinori Shimoji (2006, 2008, 2011, 2018), particularly Shimoji's 2008 Ph.D. dissertation on the language. Ikema has most thoroughly been described by Yuka Hayashi, starting with discourse data in Hayashi (2009), a description in Hayashi (2010) and subsequent dissertation in Hayashi (2013). Yukinori Takubo has also produced further analysis of Ikema morphophonemics (Takubo 2021). As of this writing, an online Japanese-Ikema-English dictionary has also been partially compiled (Kindred 2019), indexing some Ikema lexical items with phonetic transcription. There has also been some interest in the phonetics and acoustics of Ikema, primarily through the work of Shinohara & Fujimoto (2018, 2021), which further informs the analyses in this thesis. Ōgami has been described extensively by Pellard in his Ph.D. dissertation (2009), much of which is repeated in Pellard (2010). Tarama/Minna was sketched as early as Sakiyama (2003), though the differences between this and the description given in K. Shimoji (2004) and Aoi (2015) are significant enough that it is unclear if the earlier source is entirely reliable.

A few works collecting general research on the Ryukyuan languages have been published in more recent years. One of the earlier of these is the *Introduction to Ryukyuan Languages* (Shimoji & Pellard 2010), which alongside the *Handbook of the Ryukyuan Languages* (Heinrich, Miyara & Shimoji 2015), is a collection of grammar sketches and discussions of phenomena in the Ryukyuan family as well as specific Ryukyuan languages. Another major source of data for this thesis has been the research report *General Study for Research and Conservation of Endangered Dialects in Japan – Research Report on Miyako Ryukyuan* (Kibe 2012/2019), which contains a large amount of field data on multiple Miyako languages collected by researchers working in various locations across the Miyako islands. Perhaps just as importantly, the report also contains several articles discussing research on Miyako phonetics, syntax, and sociolinguistics. Some limitations must be acknowledged for this data set. Pellard & Hayashi (2012/2019: 14) note that the phonetic transcriptions collected in the research report are "phonetic transcriptions of utterances mostly obtained from a single speaker in a single survey", making it necessary to view some of the transcriptions there with a degree of caution. Nonetheless, while the data contained therein cannot provide an exhaustive and definite description of Miyako, it serves as a useful resource both for identifying patterns and for

corroborating patterns described elsewhere. While all efforts have been made to corroborate all attested forms, it may unfortunately be the case that some forms are of a dubious or uncertain nature. Where this is the case, it will be explicitly stated. Any errors or misinterpretations of the research discussed in this section are my own.

CHAPTER 3: Analysis of Miyako initial geminates and onset phonotactics

This chapter provides descriptions of the onset phonotactics and geminate structures in specific Miyako languages, starting with Irabu, then proceeding to Ikema, and finally to Ōgami. In order to account for the initial geminates, some description is also given for general phonotactic restrictions in Miyako, such as the undominated constraints against monomoraic words and complex margins. It is also argued here that the Irabu affricates /c z/ are not phonologically analyzed as fricatives, as is claimed in Shimoji (2008, 2011) and that both the initial geminates [ʊv] and [ʐz] and the apparent nasal-obstruent partial geminates in Ikema (Hayashi 2010, Pellard & Hayashi 2012/2019) may be better understood as heterosyllabic sequences. Each section begins with a brief phonological description of the language in question.

3.1 Irabu

Irabu is spoken on the island of Irabu and in the Sarahama region of Ikema island and was estimated by Shimoji (2008: 26) to have "approximately 2.000 to 2,500" speakers. This figure was determined in part by age, fieldwork having shown that fluent speakers of the Irabu dialects are "almost all over sixty years old" and census data from 2004 showing that one third of the island's population was at this age or older at that time.

In this section, I will begin by providing a basic description of Irabu, its syllabic consonant inventory and its initial geminate inventory. I will then proceed to outline several points of interest in the Irabu dialects. The first of these is the behavior of the segments /v/ and /ž/, which occur almost exclusively as approximants, but also appear to form partial geminates in onset position. These geminates are suggested here to be sequences of syllabic consonants and simple CV syllables due to markedness effects and the patterning of syllabic consonants in Irabu. The second point of interest is the behavior of the segments /c/ [ts] and /z/ [dz], which are proposed in Shimoji (2008, 2011) to be analyzed as fricatives. It is argued here that this is unlikely to be the case due to cross-linguistic evidence against this hypothesis from the other Miyako languages, the prohibition against initial [dzː], as well as the lack of alternation between the fricatives and the affricates.

3.1.1 Overview of Irabu phonology

(10) Irabu consonant inventory, adapted from Shimoji (2011: 79)

	Labial	Labiodental	Alveolar	Palatal	Velar	Glottal
Plosives	p b		t d		k g	
Fricatives		f v	s z [dz] c [ts] ž			(h)
Affricates						
Nasals	m		n			
Liquids			r			
Approximants	(w)			j		

Irabu is relatively close to the general Miyako syllable structure and consonant inventory given previously, lacking only the voiceless nasal /n̥/ and only marginally featuring the approximant /w/ and the glottal fricative /h/. The segments /z/ and /c/ consistently surface as [dz] and [ts] (unless phonetically palatalized as [dʑ] and [tʃ]), respectively, but are suggested to be underlyingly analyzed as fricatives in Shimoji (2008, 2011). It should also be noted here that /r/ surfaces as a flap [ɾ] only when found as a singleton onset segment but occurs as a voiced retroflex lateral approximant [ɭ] in all other contexts, including when it occurs as a geminate onset. The segments /v/ and /ž/ alternate primarily with the voiced fricatives [v z] and approximants [ʊ ʐ], with the fricative segments occurring in intervocalic geminates and the approximants occurring in all other contexts. Interestingly, neither segment may occur as a singleton onset. /v/ and /ž/ thus show some interesting behaviors: The phonemic representation of /v/ is given as a labiodental approximant [ʊ] in Shimoji (2008) and as a voiced fricative with the labiodental approximant as an allophone of the voiced labiodental /v/ phoneme in Shimoji (2011). /v/ appears to surface as [ʊ] in multiple contexts, just as /ž/ "includes allophones ranging from [z] with less friction to an approximant version of [z], [ʐ]" (Shimoji 2011: 79-80). The following contexts are taken from Shimoji (2008: 63):

(11)	#NN#	#(N)N.C	#Initial gem.	Medial gem.	Final coda#
/v/	[ʊ̥ː]	[ʊ̥(ː)]	[ʊv]	[vv~ʊC]	[ʊ]
/ž/	[ʐ̥ː]	[ʐ̥(ː)]	[ʐ̥z]	[ʒʒ~zz~ʐ̥C]	[ʐ̥]

In the above table, N stands for nucleus, showing that the nucleus-only syllable may either stand alone or precede another syllable. In both cases, the approximant allophone is preferred. I have followed Shimoji (2011) in assuming that the underlying phoneme for [ʊ] and [v] is the labiodental /v/. As will become clear in the following sections, this assumption is linked to the assumption that /v/ is underlyingly moraic in Irabu, and possibly also elsewhere among the Miyako languages.

3.1.2 Irabu syllable structure

Irabu syllables may feature singleton or geminate onsets, but no tautosyllabic cluster may occur at either edge. Onsets and codas are otherwise optional in word-initial syllables, but onsets are required in any non-initial syllables:

(12) Irabu syllable structure (Shimoji 2011: 80-81):
$((O_i)O_i)N_1(N_2)(C)$

As noted previously, per Shimoji's analysis, only a geminate may fill both onset slots. Consonant-glide (CG) sequences (in which G may be /j/ or /w/) may be underlyingly present in the onset, but the glide will surface as palatalization (or for /w/, labialization) of the initial consonant when an onset segment is present, e.g., /kakja/ [kakʲa] 'writer', /kwaas/ [kʷaːs] 'snack' (Shimoji 2011: 80). Heterorganic CC sequences in the onset are suggested to be syllabified as C.CV, as in /sma/ → [s.ma]. The absence of edge clusters is intriguing in the context of the typological prediction in Davis & Topintzi (2017), in which the following universals are proposed:

> "(I) if edge clusters do not exist, then edge geminates are moraic; (II) if edge geminates are non-moraic, then edge clusters are non-moraic" (Davis & Topintzi 2017: 262).

In the above statement, "'edge' refers to the same edge, either right or left edge, respectively" (ibid.). The reasoning behind these predictions is fairly straightforward: If geminates are analyzed as moraic consonant segments, it follows that non-moraic geminates are not geminates, and such geminates should instead be analyzed as sequences of homorganic consonants. Thus, a non-moraic onset geminate would in fact be a complex CC- onset, which does not occur in the Miyako languages examined here. In other words, if non-moraic onset clusters occur in the language, then syllables featuring these clusters should pattern with those featuring singleton onsets in terms of moraicity, as the initial cluster should have no impact on the weight of the syllable. Clearly, however, Irabu syllables do not permit edge clusters, while the initial geminates satisfy the bimoraic minimality constraint. This appears to provide some support for typological claim made in Davis & Topintzi (2017) cited above. Geminates in Irabu occur both word-initially and intervocalically and are moraic in both positions. While the moraicity of the intervocalic geminates could be explained by coda moraicity under a CC representation of the geminate, the word-initial geminate does not have access to this option. While nearly all the onset geminates appear to occur word-initially, it is noted in Shimoji (2008: 81) that medial onset geminates occur "only occasionally, when preceded by a syllabic consonant, as in /v.cca/ [ʊttsa] 'quail'".

Shimoji (2008, 2011) also identifies Irabu syllables that consist only of a long consonant as being phonologically analyzed as sequences of syllabic consonants filling two nucleus slots (NN) in the syllable. These bimoraic consonantal syllables are referred to as presyllables in Shimoji (2008). In the

present section, I will be referring to these syllables as NN syllables, and words consisting of such a syllable as an NN word. This is to avoid confusion with concept of minor syllables, which are also known by this term. It may initially appear tempting to posit that words consisting of an NN or N syllable and a CV syllable are sesquisyllabic. However, presyllables under this definition cannot themselves be well-formed words (Butler 2015: 448). In addition to occurring word-initially, Irabu NN syllables may serve as the sole constituent of a monosyllabic word, as in the below examples:

(13) (NN) monosyllables (Shimoji 2008: 50, 63-64):

/mm/	[mː]	‘potato’
/nn/	[ŋː]	‘yes’
/vv/	[ʊː]	‘sell’
/žž/	[ʐ̩ː]	‘rice ball’
/rr/	[l̩ː]	‘enter’

As such, it is stated in the above examples and further specified in Shimoji (2008: 63) that none of the obstruents may occur in the nucleus position of these NN syllables. Shimoji (2011: 82) contradicts this, proposing that these NN monosyllables may consist of obstruents. Specifically, it is noted that both voiceless fricatives and the voiced and voiceless affricates may form these words. Fieldwork in Kibe (2012/2019) shows even more differences in transcription. The table below shows some of the differences between these references. The page number for each reference is shown in the “p” column. Two of the examples are not recorded in Kibe for the Irabu dialects, and are marked as NR.

(14) NN obstruent monosyllables in Shimoji (2008, 2010)

Lexical item	Shimoji (2008)	p.	Shimoji (2011)	p.	Kibe (2012/2019)	p.
‘breast’	/cïï/ [tsɨː]	71	/cː/ [tsːʐ]	82	tsɻː	190
‘letter’	/zïï/ [dzɨː]	71	/zː/ [dzːʐ]	82	NR	-
‘come’	/fïï/ [fɯː]	71	/fː/ [fːᵛ]	82	NR	-
‘nest’	/sïï/ [sïː]	55	/sː/ [sːʐ]	82	sɻː	189

It appears that the fricative-like vowel /ɻ/ may account for some of the differences between the two analyses in Shimoji (2008, 2011). Recalling that /ɻ/ is characterized by both vocalic and fricative phonetic qualities, it is therefore unclear whether the words in (14) are bimoraic syllabic consonants or simply CVː syllables with some degree of frication in the nucleus. While /ɻ/ is not mentioned in Shimoji (2008, 2011), /ɻ/ appears to be widely attested in Kibe (2012/2019). The presence of off-glide vocoids indicated in Shimoji (2011: 82) therefore appears to simply be a consonant analysis of /ɻ/. We are therefore left with two possible conclusions: Either Irabu permits long, syllabic obstruents to form NN monosyllables, or there is a set of words consisting of a fricative or strident segment followed by

/ɳ/. It is also worth noting here that the approximant allophones of /v/ and /ž/, [ʋ] and [ʐ̩] respectively, are selected in coda and nucleus position, making the preferred output as sonorous as possible. Given that syllabic [v] and [z] do not appear to occur in these contexts, it should be unexpected for [f] and [s], let alone [ts] and [dz], to do so. As such, the analysis in Kibe (2012/2019) appears more feasible.

In the following sections, it will therefore be assumed that words such as /sː/ 'nest' actually feature the less marked CVː structure /sɳː/. Sequences such as the purported /sta/ [s$^{(z)}$.ta] 'tongue' (Shimoji 2011: 82) should then also be analyzed as /sɳta/ [sɳta]. This brings the set of NN monosyllables down to the sonorants, liquids, and nasals listed in (13). In other words, the set of syllabic segments is therefore the same as the set of singleton coda segments. When a stop or fricative appears medially, it must either form an onset or an intervocalic geminate. Only the more sonorous segments in (13) may occur without gemination in this context. Note that, as established previously, Shimoji (2011: 82) identifies the affricate series /c/ [ts], /z/ [dz] as fricatives due to their patterning with the fricative series. Part of the argument for this is the suggestion that the affricates may occur in nucleus position when preceding a heterorganic consonant. Also like the fricatives, they may be followed by an off-glide vocoid when preceding another consonant. These vocoids are described as voiced approximants in the same place of articulation as the preceding consonant and are described as being "predictable and phonologically invisible" (ibid.):

(15)

/ckara/ [ts$^{(z)}$.ka.ra] 'power'
/ftai/ [f$^{(ʋ)}$.tai] 'forehead'
/sma/ [s$^{(z)}$.ma] 'island'

As was discussed above, this will here be considered evidence of frication from the presence of the vowel /ɳ/. Per this analysis, the initial syllables of the words in (15) above are in fact plain CV syllables of the template [Cɳ]. This would not appear to impact the example of /v.cca/ [ʋttsa] 'quail' cited above, however, as [ʋ] here does not indicate any frication. We return now to the initial geminates.

3.1.3 Irabu initial geminate contrasts

With regard to contrastive length, the bilabial nasal /m/ and the alveolar nasal /n/ appear to be the only Irabu consonants that feature a length contrast in root-initial position. In derived contexts, however, /f/, /s/ and the affricate /c/ [ts] also show contrastive length word-initially:

(16) Initial singleton and geminate onset pairs, as given in Shimoji (2008: 68):

/fau/ [fau] 'eat' CVV[3] /ffau/ [ffau] 'child' (accusative) CCVV

[3] This syllable structure is given as CV in Shimoji (2008). I have assumed this to be a typographical error.

/sa.gi/ [sagi] ‘k.o.bird’ CV.CV	/ssa.gi/ [ssagi] ‘bridal’ CCV.CV
/ci.bi/ [tʃibi] ‘hip’ CV.CV	/ccir/ [ˀttʃil̩] ‘pipe’ CCVC
/maa.su/ [maːsu] ‘salt’ CVV.CV	/mmaa/ [mmaː] ‘No’ CCVV
/na.ma/ [nama] ‘raw’ CV.CV	/nna.ma/ [nnama] ’now’ CCV.CV

The labiodental fricative /v/, the voiced alveolar fricative /ž/, and the flap /r/ (as [l̩]) occur as long syllabic segments in monosyllabic roots due to word minimality constraints (Shimoji 2008: 68-69). /v/ and /ž/ are furthermore obligatorily long when occurring in onset position. Note also that /ffau/ receives its initial geminate as a result of word minimality, which in Irabu applies to the “grammatical word rather than a word-plus” (Shimoji 2011: 66). In other words, the affixation of the accusative case clitic /=u/ to the monomoraic root /fa/ does not prevent repair of the root through onset gemination, even though the minimal word in Irabu is obligatorily bimoraic and does not require three morae.

Long /t/ also does occur as an initial geminate in Irabu, but this is an extremely marginal case (Shimoji 2011: 55-56). We may therefore state that the full set of contrastive or non-contrastive initial geminate segments occurring in Irabu surface forms is as found in (16b), and that the subset of initial geminates that are contrastive root-initially is as in (16c). (16a) shows the set of segments permitted to geminate intervocalically. When occurring in intervocalic position, the stops, fricatives and affricates are obligatorily geminate, as only nasals, sonorants and [l̩] may form a coda. (16b) and (16c) are subsets of (16a). In the below example, the geminates [tː] and [dzː] are in parentheses to indicate that initial [tː] is marginal and intervocalic [dzː] appears to be in the process of neutralizing towards [tsː].

(17)

a. [mː] [nː] [fː] [vː] [sː] [žː~zː] [l̩ː] [tsː] ([dzː]) [pː] [tː] [kː]
b. [mː] [nː] [fː] [vː] [sː] [žː~zː] [l̩ː] [tsː] ([tː])
c. [mː] [nː]

As can be seen in (17b), it is clear that the Irabu initial geminate does not pattern with the frequency hierarchy found in Kraehenmann (2011) except in a few specific ways: the contrastive root-initial nasals [nː] and [mː] are cross-linguistically common, as are the voiceless fricatives. It is furthermore interesting if not necessarily significant that the most frequently occurring places of articulation, the coronals and labials, are quite prominently featured here. However, with [tː] as a highly marginal initial geminate segment, the near-total absence of the stop series is directly contradictory to what would be expected cross-linguistically. Of course, frequency data alone does not provide an explanatory mechanism for the initial geminate segments found in Irabu.

To deal with these questions, let us review which segments may form moraic onsets in Irabu. Shimoji (2011: 81) notes that the “geminate onset must be fricative or resonant”, except for [tː], which

is “only found in a limited number of roots (…) or in morpheme boundaries”. Note that this includes the entire set of moraic onsets, not just a subset of coerced or distinctive moraic onsets. Returning to the initial geminates in (17), there is no particular reason to restrict the lexical specification of these initial geminate segments. Instead, the issue here is that if we assume that [vː], [ʐː], [fː], and [sː] are the only coerced forms, Irabu appears to contradict the expected implicational hierarchy from sonority regardless of whether voicing is taken to be the aspect of sonority that controls coerced onset moraicity. Problematically, while [fː] and [sː] occurring coerced is exactly as anticipated, we would expect to find coerced [tː] and [pː], while [ʊː] and [ʐː] should only be available to distinctive moraicity. As has been mentioned, the occurrence of [tː] is also perplexing from this perspective. Due to its extremely limited environment, it appears that only distinctive [tː] may occur in Irabu. As such, the question becomes whether the apparently coerced moraicity of [ʊː] and [ʐː] poses a challenge to the moraic account of onset geminates, and whether it truly is the case that Irabu allows voiceless fricatives to be coerced while banning voiceless stops from the same. However, as will be argued below, the affricate segments [ts] and [dz] may be better understood as stops.

3.1.4 Irabu affricates as stops

In this section, it will be proposed that Irabu /c z/ be analyzed as stops in light of both the cross-linguistic behavior of affricate segments, the markedness of initial and intervocalic [dzː], and the presence of /ɳ/ in Irabu suggested above. For the sake of clarity, I will refer to the phonemes given as /c/ and /z/ in Shimoji (2008, 2010) as /ts/ and /dz/. This is done to avoid confusion between /z/ [dz] and /ž/ [ʐ, z], and to clearly state the argument being made here, which is that Irabu features the affricate phonemes /ts/ and /dz/. Being affricates, it is suggested here that these segments are analyzed as stops.

3.1.4.1 Irabu affricate patterning

The ways in which the Irabu affricates are argued to form a class with the fricatives can be summarized as in (20) below:

(18)

i. Fricatives (including the affricates) are syllabic when followed by a consonant segment, e.g. /s.ma/ vs /mas/ (Shimoji 2011: 82).

ii. Long sonorants or fricatives, including the affricates, may fill the nucleus slots of an (NN) syllable, e.g. /mm/ [mː] ‘potato’, /ff/ [fː] ‘come’, /cː/ [tsᶻː] ‘breast’, /zː/ [dzᶻː] ‘letter’ (ibid.).

iii. /ï/ and /ïː/ can only take a fricative (or affricate) onset (Shimoji 2008: 39).

iv. Root morphemes that underlyingly consist only of a single fricative or affricate segment receive epenthetic /ïï/ (Shimoji 2008: 66).

Firstly, it should be noted that /h/ is not included in any of the stipulations in (20). As this segment clearly deviates from any of the fricative behaviors discussed above, it may be tempting to speculate that this is evidence for the hypothesis presented here. However, there does not appear to be much support for either analysis. As Shimoji (2008: 38) states, /h/ is "mostly restricted to non-native words" and is "the only phoneme whose place of articulation does not form a natural class with other phonemes". Furthermore, /h/ should also generally expected to pattern differently from other fricatives cross-linguistically. Due to the phonetics of [h], the question of whether this segment can actually be considered a fricative has been a matter of some controversy, see for instance Ladefoged (1990).

Per the discussion in 3.1.2 above, it is argued here that both the affricates and the fricatives are non-syllabic and non-moraic in cases such as those in (18-i) and (18-ii). For (18-iii) and (18-iv), we must first determine whether these two generalizations can be collapsed into one. Shimoji (2018: 6-8) states that /ɨ/ (which is assumed here to correspond to /ɪ/) can be identified as an epenthetic vowel that repairs morphophonemically derived consonant clusters and illegal codas. It is also stated in Shimoji (2008: 73) that the same vowel is "best treated as being underlyingly absent, where the surface /Cɨ/ is underlyingly //C//". This pattern appears to apply to stems consisting of //f//, //s//, and //c//, but not //v//, //ž//, //r//, //m// or //n// (Shimoji 2008: 50, 282). //f s c// are thus suggested to be the only underlyingly monomoraic obstruent roots. As no such roots may consist of plain stops, it cannot be determined whether the same epenthesis would apply to these segments. It therefore seems likely that the morphophonemic rules in (18-iii) and (18-iv) can be collapsed into a single process specifying that the fricatives (and the affricates) require the insertion of an epenthetic high vowel when found word-finally. However, it appears that this epenthetic vowel cannot follow a geminate fricative. It is apparent that where an underlyingly long fricative would otherwise be followed by the epenthesized vowel, the long segment is shortened instead of being syllabified as an initial geminate. The vowel must then also be long to satisfy bimoraic minimality. In the below example, /ɨ/ is used by Shimoji (2018), though it is argued here that this vowel may be /ɳ/:

(19) /ɨ/ epenthesis and lengthening (Shimoji 2018: 8, translation mine):

	'read'	'lend'	'know'
Affix /-a/ (volitional)	jum-a	karas-a	ss-a
Affix /-i/ (imperative)	jum-i	karas-i	ss-i
No affix (non-past)	jum	*karas → karasɨ	*ss → *ssɨ → sɨ̈

Epenthesis occurs in the above examples because [s] cannot surface as a singleton coda (as in **karas*), and also cannot form the nucleus of a syllable, as in (*ss). This raises the question of why, given that initial geminate fricatives and affricates are permitted, we do not see a geminate onset and a singleton

epenthesized segment [ï] or [ɿ]. Shimoji (2008: 73) suggests that the prohibition against the geminate fricative/affricate + /ɨ/ sequence may be due to phonetic factors, such as the avoidance of low sonority. Interestingly, there is a major difference in transcription between Shimoji (2008) and Pellard & Hayashi (2012/2019) with regard to the structure of certain lexical items. (20-a) below shows Shimoji's underlying morphophonemic unit and phonemic derivation of the word *cïcï* 'the moon' (Shimoji 2008: 73), while (20-c) shows the phonemic and phonetic forms of the word *sïïsï* 'meat' (ibid.). (20-b) and (20-d) show the phonetic transcriptions of these words given in Pellard & Hayashi (2012/2019: 26). The syllable structure notations have been added for illustrative purposes. /ɿ/ has been indicated as a vowel segment below (V):

(20)	a. //cc// 'the moon'	>	/cïcï/	'the moon'	CV.CV
	b. -		[tsɿttsu]	'the moon'	CVC:V
	c. /sïïsï/ 'meat'	>	[sɨːsɨ]	'meat'	CV.CV
	d. -		[ᶻɿzu]	'meat'	$^{(C)}$ɿCV

This concretely displays the disputed nature of /ɿ/ between different researchers, indicating two possibilities. Irabu [ï] may be analyzed as an epenthetic /ï/, or it is the (underlying or epenthetic) segment /ɿ/ [ɿ], as is assumed in Pellard & Hayashi (2012/2019). Shimoji (2011: 114) transcribes the word given as /cïcï/ in (20) as *cc=nu* 'moon=NOM' in discourse data, indicating a third possibility, which is that /ts/ may form a licit NN monosyllable. As was argued in 3.1.2, this does not appear to be the case, and the /ɿ/ analysis is assumed here. Thus, words such as 'grass', given as *fïsa* in Shimoji (2008: 111) and as *fsa* in Shimoji (2011: 122), are argued here to feature the structure /fɿsa/. Because the segment /ɿ/ is between two voiceless segments, it is devoiced and may resemble a fricative while still being analyzed as a vowel. This claim receives some support in Kibe (2012/2019: 230), where the same Irabu word is recorded as [fu̥sa]. This raises the question of why /ɿ/ only follows these segments. This is not necessarily the case, however, as this segment appears to occur following stops, vowels, as well as word-initially (Pellard & Hayashi 2012/2019: 23). This distributional restriction thus does not appear to apply, and the behavior of /ɿ/ furthermore does not indicate that /ts/ and /dz/ pattern as fricatives.

However, the ban on geminate fricative + /ɿ/ sequences such as *[sːɿ appears to be valid. If /ɿ/ is assumed to be the vowel in words such as *sïï* 'know', the motivation for the prohibition against these sequences appears even more likely to be conditioned by the avoidance of low sonority, as was suggested by Shimoji (2008: 73). Because the vowel /ɿ/ is suggested to feature a kind of fricative co-articulation, these sequences may result an illegally small sonority change from onset to nucleus. Alternatively, the specific ban on long alveolar fricatives preceding /ɿ/ may be conditioned by the markedness of the "fricative vowel" following a moraic fricative segment. Given that singleton fricatives appear to form onsets to syllables in which /ɿ/ occupies nucleus position, it appears highly

likely that the geminate identity of the fricative segments cause sequences such as *[sːɳ] to be ill-formed. More research is needed on /ɳ/ and its exact acoustic and phonetic qualities to draw firm conclusions on this matter, however.

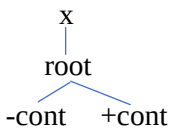
3.1.4.2 Affricate representations

One could consider whether the Irabu affricates are merely allophones of the fricative series. However, there is clearly a contrast between the fricatives and the affricates in most contexts, and it is not suggested that /ts/ and /dz/ feature fricative allophones in any context (Shimoji 2008: 63). Furthermore, there is no proposed pattern of phonetic affrication in any environment for either the stops or the fricatives. Intriguingly, while /dz/ consistently surfaces as [dz], the phoneme /ž/ does feature a fricative allophone [z] in intervocalic geminates, where it occurs as [zː] or as [ʒː] when palatalized. It is thus in contrast with intervocalic /dz/, which surfaces as [dzː] or variably as [tsː] in this context. Lastly, if /ts/ and /dz/ were fricatives in the underlying representation, the assumption would then be that the optimal output of these input segments is their respective affricates [ts], [dz]. However, this would imply that the inputs /s/ and /z/ (or indeed any fricative input) are sub-optimal options in the surface representation compared to the affricates in certain contexts. This would in turn suggest that these segments should be found in complementary distribution to some extent. Because this is not the case, and because there is no apparent evidence of contrast neutralization between the affricates, the plain stops, and the fricatives, the classification of the affricates as fricatives is worthy of questioning. As such, we must consider how an affricate should be represented, and how they could occur as syllabic or otherwise pattern with fricatives (if this indeed were the case).

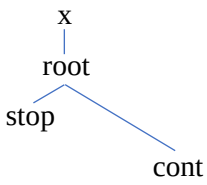
Kehrein (2002) describes a series of approaches taken to representing affricate segments; as [strident] stops (Jakobson et al. 1951), as phonemes specified as [-cont] [+delayed release] (Chomsky & Halle 1968), or as segments with specifications for both [stop] and [cont]. This then raises the question of whether there is in fact a feature that could cause the affricate to belong to this group of segments. The challenges of this former analysis hinge on how affricates are represented in the underlying structure. For instance, the representations given in Sagey (1986: 81) and Lombardi (1990) appear to provide the features needed to ensure that the affricates pattern with fricatives:

(21)

a. Sagey (1986)



b. Lombardi (1990)



Sagey's representation in (21a) describes the affricate as a segment that contains "a sequence of articulations on a single timing unit" (ibid.), or an underlying series of specifications of values of the feature [continuant], i.e. [-continuant][+continuant]. From this perspective, it can be argued that the Irabu constraints on syllabic and moraic segments simply require the presence of a segment marked for the value [+continuant] or the unary feature [continuant], allowing the affricates to pattern with the fricative series. Recalling that voiced fricatives are permitted to serve as initial geminate, it would then have to be assumed that the constraint prohibiting initial geminate [dz] is in fact a constraint against moraic voiced [-continuant] segments. This would also apply if, as in (21b), it were assumed that the underlying structure is [stop][continuant] with these features specified at different tiers, and thus without an underlyingly ordered [+continuant][-continuant] sequence.

However, there are a few crucial problems involved in analyzing the Irabu affricates as such. Kehrein (2002) examines cross-linguistic affricate patterning and demonstrates that centrally released non-strident affricates such as [p^{ϕ} t^{θ} k^{x}] do not appear to contrast with homorganic stops. Were the structure of the affricates represented as in (21a) or (21b), this would perhaps allow us to account for the Irabu pattern. Furthermore, it would explain the inclusion of the phoneme /ž/ with the segments in (19), as this segment can variably be either a fricative or approximant (and therefore [continuant] in either case). However, we would also expect to find phonemic affricates such as /pf/ and /qχ/ in cross-linguistic data, which does not seem be the case. Kehrein (ibid.) also highlights a theoretical issue in that the feature [continuant] by definition entails that there is no complete closure in the oral cavity. The specification of affricates as [continuant] thus runs contrary to the fact that nasal affricates (and indeed nasals in general) cannot feature a [+continuant] or [continuant] feature on either node, as the feature [nasal] entails oral closure. Kehrein's own solution is to suggest that affricates are not a distinct phonemic category, and that they should instead be understood to form a class with stops at the phonological level. The distinction between plain stops and affricates would then only be whether the stop segment is specified for one of the manner features [strident], [lateral], or [nasal]. [ts, dz] are thus differentiated from plain [t, d] as they are specified as [strident].

While Irabu does not feature nasal affricates, the nasals /m n/ are suggested to belong to the same class of consonants as those in (18). In other words, as the Irabu nasals must be understood to be [-continuant], it appears that [continuant] is unable to explain their apparent patterning alongside the fricatives and affricates in Irabu. Finally, it should be noted that there does not appear to be any

context in Irabu in which there is neutralization between the affricates and the fricatives, nor is there any deaffrication or spirantization in which the affricates are reduced to simple fricatives. Finally, one could ask whether the affricates pattern alongside fricatives (or fricatives and sonorants) in Irabu due to a feature such as [strident], [coronal], or whether there is another factor responsible for this patterning. These possibilities will now be examined in turn.

Without the [continuant] affricate representation, it is difficult to identify any single distinguishing feature that would include the affricates with the entire set of fricatives in Irabu. The [strident] feature accurately captures the segments [ts dz s z] but although the labiodental fricative /f/ has been identified as [strident] in languages such as Ewe (Utman & Blumstein 1994), there is to my knowledge no empirical basis, nor has it been suggested in the literature, that this should be the case for the Irabu labiodentals. Clearly, the same issue applies if the place feature [coronal] is used instead. Once again, the correct alveolar segments are selected while the non-coronal [f] fails to be selected. Additionally, it is a significant problem that neither solution accounts for the fact that sonorants may pattern in the same manner as the affricates and fricatives in (18-i) and (18-ii). An alternate solution would then be to posit a two-predicate constraint *C[-continuant, -strident]/nucleus, specifying that segments occupying this position must be at least either [continuant] or [strident].

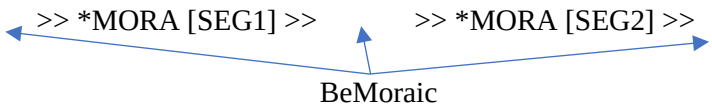
This solution does not appear to be particularly convincing, as it appears less complex to suggest that the occurrence of /ɳ/ interacts in some way with the strident segments /s, z, ts, dz/. As previously mentioned, the use of the feature [continuant] is problematic for the nasals, which are attested as occurring alongside the fricatives and affricates. The nasals are clearly neither [continuant] nor [strident], and while /n/ is [coronal], /m/ is not. This would then predict that the nasals should be unable to surface in this context. It would be possible here to specify further, stipulating that the segments in question cannot feature negative values for all three features [continuant], [strident], and [nasal]. At this point, however, a more parsimonious solution is desired. The inability of the approach discussed here to capture the seemingly syllabic Irabu affricates thus lends some support to the proposal that they do not pattern with the Irabu fricatives, and in fact serve as onsets to the vowel /ɳ/ in NN monosyllables proposed to consist of affricate segments.

3.1.4.3 Affricate sonority and markedness

It therefore appears that the distribution of syllabic consonant segments may be better understood as being driven by sonority. The set of consonants that may be syllabic neatly aligns with sonority-driven predictions concerning the implicational hierarchies of nucleus and moraic segments in the literature[4]. Morén (1999, 2003) provides a model for describing the ranking relationships for coerced weight within a universal markedness hierarchy as follows (Morén 2003: 290):

[4] See for instance Zec (1988), Morén (1999, 2003).

(22) Coerced weight (simplified)



In the above model, BeMoraic stands in for any constraint that coerces moraicity of a specific segment in a specific context. If BeMoraic is ranked between the two markedness constraints *MORA[SEG1] and *MORA[SEG2], it follows that the constraint against segment [SEG1] surfacing as moraic outranks the constraint coercing weight onto the segment and will not be moraic. Segment [SEG2], however, will surface as moraic. If it is such that the affricate stops distinguish themselves from regular stops in terms of sonority, then it should be possible to represent this in terms of a constraint ranking such as in (22). Morén (1999, 2003) follows Selkirk (1984) in describing the minimally necessary sonority ranking of obstruents as follows: continuants >> voiced stops >> plain stops >> aspirated stops.

However, if it is assumed that the affricates are permitted to be syllabic due to sonority, further problems emerge. First, if weight in this context is coerced, it is expected that the sonority of the affricates must be higher than that of the plain stops. Although there is no phonetic or acoustic data to specifically support this possibility in Irabu, it is not a priori impossible. Parker (2002: 71) notes that in the literature, affricates have variously been proposed to be more sonorous than plain stops, less sonorous than plain stops, or equally as sonorous as plain stops. Accordingly, acoustic research has suggested that the sonority of certain disputed sonority contrasts, in this case affricates vs. stops, may be language specific (Jany et al. 2007). This implies that there should be no cross-linguistically fixed ranking for the sonority of these contrasts. However, this is not itself proof that a potential difference in sonority between the affricates and the plain stops is significant or insignificant in Irabu.

The larger issue is that if these affricate stops can be syllabified in nucleus position, i.e., that they may receive coerced weight in the nucleus, and this coercion is licensed by sonority, it should be expected that the more-sonorous voiced stops and voiced fricatives should be capable of receiving coerced weight in the nucleus. However, while the sonorant allophones of the voiced fricative series may be syllabic, the voiced fricatives themselves never occur as syllable nuclei in the surface representation. Likewise, it does not appear that any voiced stops may occur as intervocalic geminates. The only exception to this in Irabu is the affricate /dz/, which maintains a contrast with /ts/ when geminated intervocalically. Note however that this contrast is also undergoing neutralization in the direction of the voiceless affricate [ts]. Phonetically, [ts] and [dz] occur in free variation when surfacing as allophones of intervocalic geminate /dz/, while the underlyingly voiceless affricate features no alternation with the voiced affricate in the same context. Shimoji (2008: 56) notes that the occurrence of intervocalic [dz:] is present in “some very old speakers”, while others exclusively use

intervocalic [tsː]. When /z/ [dz] occurs as a singleton onset, no such neutralization occurs (the following examples are taken from Shimoji 2008: 41, 56-57, 72):

(23) /fiz.za/ [fɯddza ~ fɯttsa] ‘whale’ /fɨc.ca/ [fɯttsa] ‘mouth’
/az.za/ [addza ~ attsa] ‘taste’ /ac.ca/ [attsa] ‘side’
/aza/ [adza] ‘elder brother’ /aca/ [atsa] ‘tomorrow’

Some caution is necessary here, however. That the voiceless affricate should be preferred intervocalically in spite of its relative sonority is not entirely unexpected. Voiced sibilant affricates such as /dz/ have been argued to be more cross-linguistically marked than voiceless ones due to their phonetic qualities (Zygis et al. 2012), and so the absence of [dz] where voiced stops are prohibited may be phonetically motivated, rather than constituting clear support for the hypothesis that the Irabu affricates should be considered stops. Voiced geminate obstruents are generally found to require more effort to produce than voiceless geminate stops (e.g., Ohala 1983 and Kirchner 2000). This may provide part of the motivation for the diachronic trend towards neutralization of the intervocalic affricate voicing contrast. Phonetic evidence of this nature is relevant, as there is no disagreement as to whether /ts/ and /dz/ are affricates in the phonetic output. It is worth noting that the preference for the voiceless affricate in the intervocalic geminate may be motivated by factors outside of the phonology.

It should briefly be noted that it is not necessarily the case that a /ts/ geminate is exclusively the result of coercion in every context. There are minimal pairs indicating that this segment may be distinctively long intervocalically as well:

(24) /acca/ [attsa] ‘side’ (Shimoji 2008: 72)
/aca/ [atsa] ‘tomorrow’ (Shimoji 2008: 272)

In the above example, there is no reason to assume that weight is coerced through the bimoraic minimality constraint or other morphophonemic processes, as both words feature at least two morae and are not monomoraic at any level of derivation. As discussed previously, Shimoji (2008: 69-70) describes a consonant lengthening rule (referred to as Geminate Copy Insertion) in which a geminate results from the adjacency of moraic //C// and //(G)V// in a word-plus. For /c/, this appears to occur for the verb /cca/ ‘wear’ (volitional), which is stated to consist of the stem //c-// and the volitional suffix /-a/. It is again argued here that this is not a process of epenthesis, but rather that either the underlying /ts/ onset is moraic (and therefore geminate), or that there is a word with an underlying non-moraic /ts/ onset, which receives moraicity by coercion. There are few attested examples of intervocalic /V[ts].[ts]V/ geminates. The only examples in Shimoji (2008, 2011, 2018) appear to be /acca/ [attsa] ‘side’, and /fic.ca/ [fɯttsa] ‘mouth (topic)’. This may of course be due to an accidental gap, or due to the fact that no broad, transcribed corpus has been compiled for this language as of this

writing. Regardless, as obstruents in medial codas are required to geminate, it can be established that the prohibition against intervocalic voiced obstruent geminates outranks sonority-driven markedness for these segments. Here and elsewhere, /dz/ clearly patterns with the voiced stops in terms of markedness.

3.1.4.4 Markedness of geminate /dz/

The claim that Irabu /ts/ and /dz/ form a class with the stops is supported somewhat by the fact that the voiced affricate /dz/ cannot form an initial geminate (Shimoji 2008: 39). As discussed above, this is exactly the case for the stop series as well, as /b d g/ are prohibited in this context. In general, it appears that voiced stops are not permitted to form initial geminates in any Miyako, Ishigaki or Okinawan language. A caveat to this claim is that, as will be discussed in the section on Ikema below, Ikema /z/ does appear to be in free variation with [dz] also when this segment forms an initial geminate. The distinction that must be drawn between these cases is that Irabu /dz/ and /ž/ do not show variation between fricative and affricate identity. Contrast is thus maintained between /dz/ and /ž/ [ʐ, z]. Pellard & Hayashi (2012/2019: 40-41) state that historical proto-Miyako *z has produced contemporary Miyako [z ~ dz] in most varieties, and the data therein shows that Irabu /dz/ may occur in free variation as [dz ~ dʑ ~ d] in initial and medial onsets. Crucially, Pellard & Hayashi (ibid.) note that "In Irabu, Kuninaka, Kurima, and Ikema, *gɻ has changed into /dzɻ/".

It thus appears to be the case that Irabu /dz/ derives both from proto-Miyako *z and *gɻ, and that *z has therefore been reanalyzed as two separate phonemes, the affricate /dz/ and the fricative or sonorant /ž/. In codas and initial geminates, [ʐ] appears to be cognate with Miyako /z/. The Irabu examples below are from Shimoji (2008: 47, 172) and Shimoji (2010: 123), while the Ikema examples are from Takubo (2021: 67) and Kibe (2012/2019: 223). Note that for the below examples, Irabu /ž/ is stated not to occur as a singleton onset, while Ikema only permits nasal codas and nasal NN words. Ōgami examples from Pellard (2010: 77, 187, 302) are included, as this is the only other Miyako language examined here that has been shown to clearly feature non-nasal syllabic consonants:

(25) Irabu – Ikema – Ōgami cognates containing /z/, /ž/

	Lexical item	Irabu	Ikema	Ōgami
Initial geminate	'father'	/žža/ [ʐza]	/zza/	/ɯa/
	'fish'	/žžu/ [ʐzu]	/zzu/	/ɯu/
Nucleus	'rice ball'	/žž/ [ʐː]	[maii]	/muks/
Coda	'rice'	/maž/ [maʐ]	[mai]	/maɯ/

Examining the examples in (25), Irabu /ž/ appears to correspond to Ikema /z/ and Ōgami /ɯ/, except in cases where /z/ would be illegal in Ikema (i.e., coda or nucleus position). While the lexical items for ‘rice ball’ and ‘rice’ may be structurally too different to compare for this purpose, we may compare these with other recorded Miyako dialect data in Kibe (2012/2019: 223-224):

(26)

Lexical item	Irabu	Kurima	Kugai
‘rice ball’	/žž/ [ʐ̩ː]	[zz] / [nnaɾi]	[nnaz]
‘rice’	/maž/ [maʐ̩]	[maz]	[maz]

As the data for Kurima and Kugai in (26) are provided as phonetic transcriptions in Kibe (2012/2019), their use here is not intended here to indicate the phonemic structure of the words in question. As is demonstrated in (25) and (26), it would appear that Miyako /z/ has been diachronically reanalyzed in Irabu as either the segment /ž/, whose allophones range from “[z] with less friction to approximant version of [z] ([ʐ̩])” (Shimoji 2010: 79-80) or as the voiced affricate /dz/ [dz ~ dz̞ ~ d]. The claim made here is that there appears to be little reason to suggest that the latter segment is analyzed as a fricative, as a fricative exists in this place of articulation with a contrast between the segments /ž/ and /dz/. Rather, it should be anticipated that the [-continuant] specification of the affricates should be the single distinctive feature separating them from the fricatives, which would entail that their status as stops is entirely necessary in perceptual terms.

While the absence of [dz] in the initial geminate does suggest that this segment is a stop, it does not explain its presence in intervocalic geminates, such as in /fïz.za/ [fɯddza]. There does not seem to be an undominated constraint against voiced fricatives in the initial geminate, nor in the intervocalic geminate. However, Irabu intervocalic geminates do appear to avoid voiced stops (Shimoji 2008: 56). As will be examined in further detail later, for younger speakers of Irabu, the contrast between [ts] and [dz] does appear to be undergoing neutralization in the direction of [tsː] in intervocalic geminates (Shimoji 2008: 57). Assuming that this neutralization indicates that intervocalic [dzː] is marked, it is suggested here that the affricate pair [ts], [dz] behave similarly to stops. As such, this indicates that Irabu does in fact feature the predicted range of anticipated coerced moraic onsets, satisfying a markedness hierarchy of *sonorants >> *voiced fricatives >> *voiceless fricatives >> *voiceless stops.

To summarize this section, it has been argued that the Irabu segments /c z/ [ts dz] form a natural class with the stops in the language. The fact that they feature a contrast with the fricatives and the stops, and that this contrast never neutralizes except where /dz/ variably surfaces as [d] suggests that they are not analyzed as fricatives. It has also been suggested that NN words that seemingly

consist of a long, bimoraic affricate actually may consist of an affricate followed by the "fricative vowel" /ɿ/. In turn, this likely indicates a sequencing condition of consonants and /ɿ/. It has also been argued that licit sequences of /ɿ/ and non-fricative (and non-affricate) segments occur, and that featural specifications fail to capture the apparent patterning of the affricates and fricatives. While it has been noted that that the Irabu affricates /ts dz/ may be more sonorous than their plain stop counterparts /t d/, the assumption that the affricates may occupy nucleus position is viewed as problematic with regard to sonority. Further support for the analysis of /ts dz/ as stops has been found in the prohibition against initial [dzː] and a tendency towards neutralization of the /ts dz/ voicing contrast in intervocalic geminates. One option that has not been examined here is whether place of articulation could account for the patterning of the affricate and continuant segments discussed above. This possibility appears largely fruitless, however, as only the velar and glottal places of articulation are blocked from being syllabic or from forming initial geminates. We are nonetheless left with the question of why the initial geminate must be either the affricate /ts/, a fricative, or a sonorant.

3.1.5 Coerced and distinctive initial geminates in Irabu

Therefore, we must account for exactly which consonants may be coercively or distinctively moraic. In Morén (2003), distinctiveness for consonant moraicity pertains to underived geminates. In other words, any initial geminate resulting from the application of a phonological rule (or constraint, assuming OT) should be considered a case of coercion. As has been established, initial /v/ and /ž/ are obligatorily long in onset position and the underlying morphemes //va// '2SG' and //ža// 'father' are lengthened to /vva/ [ʊva] and /žža/ [ʐza] respectively. As noted above, /f s c/ belong to this category as well. The exception to this rule appears to be the nasals /m n/, which are not coerced into moraicity in the initial geminate. Rather, long [mː nː] appear to be lexical in some words such as /mmaa/ [mmaː] 'No', /nna.ma/ [nna.ma] 'now'. Elsewhere, they appear to be syllabified as long nuclei either as monosyllables or as a long syllabic consonant preceding another syllable. /r/, which occurs as [l̥] except in medial singleton onsets, appears to pattern with the nasals in this matter as well (note that /r/ does not occur as a singleton word-initially):

(27) Geminate and non-geminate contrast (Shimoji 2008: 68):

/mmaa/ [mmaː] 'No'	/maa.su/ [maː.su] 'salt'
/nna.ma/ [nna.ma] 'now'	/na.ma/ [na.ma] 'raw'
/rra/ [l̥a] 'placenta'	/ku.ri/ [ku.ɾi]

Monosyllables (ibid.: 49)	Disyllables (ibid.: 49)
/mm/ [mː] 'potato'	/mm.ta/ [mː.ta] 'a kind of tree'
/nn/ [ŋː] 'yes'	/nn.sa/ [nː.sa] 'dumb'
/rr/ [l̥ː] '(the sun) sets' (participle stem)	/m.ta/ [m.ta] 'mud'

/n.fi/ [ŋ.fi] ’warm’
/prr.ma/ [plː.ma] ‘daytime’

While the underlying nasal or rhotic may receive coerced weight in the surface representation, this appears to lead it to be reanalyzed as a nucleus instead of forming an initial geminate. This suggests that the true initial geminate forms of these segments are distinctively moraic and therefore not subject to a markedness hierarchy. We are thus left with the following set of coerced and distinctive moraic onsets:

(28)

Distinctive onset weight:		Coerced onset weight:	
labial	alveolar	labial	alveolar
mː	nː l̪ː (tː)	fː (ʊv)	sː (ʐz) tsː

For the above table, it must be acknowledged that there may be words in which the initial geminate segments listed as coerced can occur as distinctively moraic. For instance, the onset in /ssam/ [ssam] ‘louse’ is not compelled by any phonotactic requirement to lengthen, as it already satisfies minimality in the underlying representation. Were the underlying representation assumed to be /sam/, this would still be bimoraic in Irabu. On the other hand, there are no contexts in which the segments listed as distinctive surface due to coercion. In cases where these segments form the onset of a monomoraic root, or where a sequence of homorganic segments may be present in onset position, resyllabification or nucleus lengthening occur instead. This allows us to form some new generalizations for the Irabu initial geminate: (I) sonorants and /t/ are distinctively moraic in onset position, and (II) initial geminates are limited to the labial/labiodental and alveolar places of articulation. Proceeding with the assumption that /ts/ is indeed a stop, we can also observe that the predicted range of sonority is revealed in the coerced moraic onsets, namely: *voiced fricatives >> *voiceless fricatives >> *stops. In the table in (29), the geminate forms of the sonorant segments [ʊ] and [ʐ] are placed in brackets because their status as initial geminates is problematic, as will be discussed later in this section. First, however, we must consider why /p t k/ do not receive coerced weight in the onset.

Itô et al. (2017) show that the feature [+anterior] can be used to differentiate between the types of voiced obstruents that are permitted to geminate in Japanese loanwords, following Chomsky & Halle (1968: 304) in defining this feature as “sounds produced with an obstruction that is located in front of the palate-alveolar region of the mouth”. [+anterior] thus accurately describes the set of coerced moraic onsets. Assuming further that a strict ban on voiced initial geminate stops is

undominated in the grammar, we can also stipulate that /d/ and /dz/ must not occur here. This leaves /p t/ with an unexplained absence in the initial geminate. Intriguingly, neither distinctive nor coercive initial [pː] are attested in any of the Miyako languages to my knowledge. Among the Southern Ryukyuan languages, [pː] appears to be exclusive to the Yaeyama language Tedumuni (Shinohara & Fujimoto 2011), and it is absent from neighboring Yaeyama languages such as Hatoma (Lawrence 2012) and Yonaguni (Izuyama 2012). What this indicates is that onset gemination of the labial stops is in fact highly marked in Miyako and perhaps also in the Japonic languages in general. Therefore, it seems that the constraints on the initial geminate favor the alveolar place of articulation for stops. Recalling that the frequency data in Kraehenmann (2011) indicates coronal stops as the most cross-linguistically common consonants in initial geminates, this seems to be in line with empirical observation as well.

An alternative hypothesis here would be to suggest that there is a contrast between the bilabial and labiodental places of articulation, allowing for a ban on the bilabials /p b/ without affecting the labiodentals /f v/. As there are no known phonemic labiodental stops, and since the Irabu alveolar segments are not attested as being dental, this would leave only the marginal nature of initial [tː] as an unresolved issue. However, this notion would be somewhat difficult to support. There is no phonemic contrast between the bilabial and the labiodental segments, and no other part of the Irabu grammar seems to distinguish between these places of articulation in any other way.

In order to make sense of this seemingly opaque system of segmental prohibitions, it is possible that a more feasible solution will be to view this pattern as an emergent property of the grammar, restricting moraicity in onset position to certain classes of consonants based on sonority due to voicing, as well as constraints against specific segment types. These arguments will be presented in the following section, in which an Optimality Theory approach is applied to the Irabu initial geminates.

3.1.6 OT analysis of Irabu initial geminates

Topintzi (2006: 33) refers to the constraint Moraic Onset as “the equivalent of Weight-by-Position for onsets”. Here, it may serve the role of the placeholder constraint BeMoraic given by Morén (1999). In other words, it stands in for other constraints that compel moraicity in onsets:

(29) Moraic Onset: Onsets are moraic (Topintzi 2006: 45).

In Irabu, the constraint in question is likely to be related to the constraint on minimal word size. This preference for bimoraic words is widely attested cross-linguistically (see for instance Hayes 1995) and is argued in McCarthy & Prince (1999) to primarily be composed of markedness constraints on the prosodic structure of the language. Of particular interest to the discussion here are the constraints Headedness/PrWd and Foot Binarity:

(30) a. Headedness/PrWd
Every prosodic word contains a foot
b. Foot Binarity (FtBin)
Feet are binary under syllabic or moraic analysis

Put very simply, these constraints ensure that all prosodic words contain at least one foot consisting of two syllables or two morae, depending on which unit the language is sensitive to. Further constraints then specify that all syllables or morae are parsed into feet, then aligned with a position in the word. Like most Japonic languages, Irabu is mora-timed, and syllable count does not appear to be relevant to word minimality, as both monosyllabic and disyllabic words may satisfy word minimality if two morae are present. As mentioned previously, Irabu onsets appear to receive weight even when the derived word does not require it. Therefore, it may be argued that the root receives status as the prosodic word by way of a morphological category (MCat) constraint as given in Prince & Smolensky (2004: 51):

(31) Lx ≈ Pr (MCat)
A member of the morphological category MCat corresponds to a PrWd.

The assumption then is that the root must be minimally bimoraic. Regardless of the theoretical assumptions made for how the bimoraic minimality constraint arises, however, it is quite clearly active in Irabu. We may therefore collapse this set of constraints into the constraint Word Minimality (WdMin). This is preferred over the Weight By Position-type constraint Moraic Onset, as a constraint ranking of *μ/ONS[SEG1] >> Moraic Onset >> *μ/ONS[SEG2] would imply that [SEG1] never receives coerced weight and [SEG2] always does. While many segments never occur as initial geminates, coerced onset weight does not appear to occur where the bimoraic minimality constraint is met:

(32) /fau/ [fau] ‘eat’
/ci.bi/ [tʃibi] ‘hip’
/sa.gi/ [sagi] ‘k.o. bird’

Conversely, coerced weight does appear to occur in a manner predicted by Moraic Onset for the initial geminates [ʊv] and [ʐz], as these segments are obligatorily moraic in the onset. As described previously, the surface forms of /v/ and /ž/ alternate between voiced fricatives and approximants. In the initial geminate, they are analyzed as an approximant-fricative sequence:

(33) /vː/ [ʊv]
/žː/ [ʐz]

It is therefore unclear exactly how these geminates should be analyzed. By definition, geminates are essentially “long” stops, though gemination may be articulated in ways other than duration (or

duration of closure for stops). A contour of manner features such as in (34) seems less consistent with a geminate and more akin to a consonant cluster. However, if analyzed as a cluster, [ʋv] and [ʐz] appear to be ill-formed onsets in terms of sonority. The Sonority Sequencing Principle (Selkirk 1984) states that, in principle, sonority within a syllable should rise from the onset to the nucleus and drop from the nucleus to the coda. Consider the following (examples taken from Shimoji (2008: 47):

(34) a.

/vva/ 'you' [ʋva]: ʋ v a

b.

/žža/ 'father' [ʐza]: ʐ z a

Sonority indicated here for illustration.

In [ʋva] and [ʐza], sonority starts high in the approximants and drops to the voiced fricative, before rising again to the low vowel [a]. There are famous exceptions to the SSP in the literature, with perhaps the most well-examined of these being the /s/ in /st-/ or /sp-/ onset clusters in various Indo-European languages. Based on a strict reading of the SSP, the more sonorous /s/ is not expected to precede a voiceless stop in the onset. However, such clusters are cross-linguistically quite common, suggesting that the /s/ is either ignored in the syllable (i.e. it is extrametrical), or that some other feature allows it to be present in this onset cluster. Extrametricality is clearly not applicable here, as this would leave the word monomoraic, and thus illegal. Morelli (1999) argues that such sequences are essentially independent of sonority and that they are governed by specific markedness for obstruent clusters. This analysis is however contingent on both members of the cluster belonging to a group of obstruents, which is not the case for [ʋ] and [ʐ].

Sonorant-obstruent onset clusters are predicted to be occur in Jay & Parker's (2020) typological approach, but this cluster type (referred to as glide-obstruent clusters) is the most highly marked candidate in an implicational hierarchy of possible onset clusters. Following this approach, we should expect to see each of the more harmonic onset clusters emerge in the language, meaning that onset clusters featuring rising sonority should necessarily also occur. This is clearly not the case in Irabu, however. Clements (1990: 288) indexes a large number of examples of violations of the SSP, such as Yateé Zapotec [wbey] 'hoe'. This specific example is of interest as it features an even larger drop in sonority within the onset, and because it features a labial sonorant-obstruent sequence akin to what is seen in [ʋva]. In other words, this violation of the SSP does not necessarily indicate that the onset cluster is ill-formed. Once again, however, the existence of a sonorant-obstruent onset cluster appears to predict the existence of onset clusters with more appropriate sonority sequences. Because Irabu (and indeed Miyako in general) does not permit CC- onsets of any kind, the cluster interpretation seems ill-advised.

It is therefore unclear whether these sequences should in fact be analyzed as geminate onsets or if they are better understood as syllabic approximants followed by a CV syllable, i.e. [ʐ.za], [ʋ.va].

If they were instead to be considered onset clusters, they would still have to be specified as moraic to satisfy minimality in words such as [ʊva] ‘you’ and [ʐza] ‘father’ (Shimoji 2008: 47). A third option is to view these as syllables with peaks that are not determined by sonority, as Clements (ibid.) suggests is the case in Berber. The syllable [wn̩] is here expected to feature a syllabic [w] due to the relative sonority of approximants and nasals. To summarize, we have three possible analyses for the voiced fricatives: (a) these sequences behave similarly enough to geminates that they should simply be analyzed as a moraic initial geminate, (b) they are heterosyllabic sequences of a syllabic sonorant followed by a singleton fricative onset and therefore do not violate any constraints on onsets, or (c) they are moraic onset clusters, and thus also the only edge clusters in Irabu.

A fourth option would be that the onset cluster in this case is non-moraic but that words such as [ʊva] are lexically specified as being exempt from the minimal word constraint. This appears to occur to a limited extent in Japanese, where certain lexical items are monomoraic. Unlike Irabu, however, word minimality in Japanese is enforced only on derived forms (Itô 1990), meaning that all productive word formation processes invariably generate bimoraic words, while Irabu enforces word minimality on underived forms as well (Shimoji 2008: 66). Additionally, if words such as [ʊva] were analyzed as monomoraic words with non-moraic onset clusters, we should also expect to find monomoraic words with a CV structure, such as /ka/ ‘mosquito’ or /hi/ ‘day’ in Japanese. However, this occurs nowhere in the Miyako languages. It is therefore apparent that the [ʊv] and [ʐz] sequences must contain at least one mora in words that would be monomoraic without them.

In the following section, it will be argued that the heterosyllabic analysis is the most feasible approach. The apparently coerced initial geminate voiced fricatives are in fact coerced into moraic weight in nucleus position, which should be preferred in terms of sonority. The interpretation of [ʊv] and [ʐz] as occupying a nucleus and singleton onset also implies that the voicing distinction for coerced moraic onsets used in Topintzi (2006) is supported by the Irabu data. The absence of [ʊ] and [ʐ] from the initial geminate would then suggest that no separate mechanism is needed to explain why voiced geminate fricatives may be moraic while unvoiced stops are not. Additionally, because the only voiced initial geminates would be distinctive, it would yield a simpler analysis in which all instances of [voice] in the initial geminate. The distinctive moraicity of the sonorant initial geminates ([nː], [mː], and [lː]) would then be preserved by faithfulness, as has been suggested. For this reason, (28-b) will be considered below as well.

3.1.6.1 Coerced initial [fː] and [sː]

In Irabu, bimoraic minimality is primarily expected to give weight to the voiceless fricatives and the affricate. I will begin by discussing the voiceless segments and will start by using the minimal word constraint WdMin. Using WdMin and the markedness constraints *μ/ONS[continuant], we can successfully select the coerced initial geminates in Irabu:

(35) WdMin >> *μ/ONS[continuant]

/fa/	WdMin	*μ/ONS[continuant]
[fa]	*!	
☞ [f:a]		*

Note however that this simplified approach fails to explain why a lengthened vowel is not preferred over the lengthened onset:

(36) WdMin >> *μ/ONS[continuant]

/fa/	WdMin	* μ/ONS[continuant]
☹ [f:a]		*!
💣 [fa:]		

Because the lengthening of the vowel satisfies bimoraic minimality, [fa:] is erroneously selected as the optimal candidate. Although the underlying input could be specified as /f:a/, this would simply describe distinctive weight, rather than coercion. A valid constraint ranking for Irabu must be able to derive the correct outputs for both /fa/ and /f:a/, but as the suggestion here is that underlying /fa/ will result in [f:a], only the monomoraic input is needed. Therefore, we must identify the other constraints that are acting on the output. One option here is to include a constraint such as Dep-IO (definition here from Kager 1999: 68):

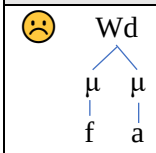
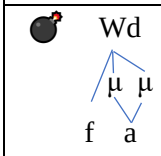
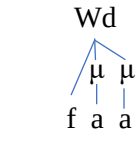
(37) Dependence Input-Output (Dep-IO)
Output segments must have input correspondents.
('No epenthesis')

Of course, the problem with using this constraint here is that it is not necessarily the case that vowel epenthesis. More accurately stated, the tableau must account for potential output candidates that feature epenthesized vowels (or codas) or added moraic weight in the nucleus. Constraints restricting epenthesis and moraic lengthening of the vowel must therefore dominate constraints on the moraic lengthening of the onset. To prevent [fa:] from being optimal, a constraint limiting the insertion of additional morae must be involved, i.e., Dep-μ:

(38) Dependence-Mora (Dep-μ)
Assign a violation mark for each mora present in the output with no corresponding mora in the input.

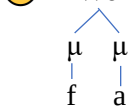
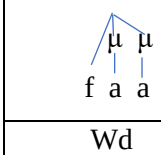
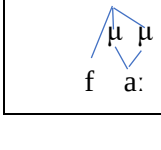
However, this constraint blocks added weight on the onset and nucleus equally. Comparing the Dep-IO and Dep-μ approaches in (35) and (36), respectively, it is clear that other constraints are needed:

(39) WdMin >> Dep-IO >> *μ/ONS[continuant]

/fa/	WdMin	Dep-IO	* μ/ONS[continuant]
☹ 			*!
💣 			
		*!	

While epenthesis is successfully blocked in (39), the lengthened vowel is optimal compared with the correct output, [fːa]. In (40) below, Dep-μ proves overly restrictive:

(40) WdMin >> Dep-μ >> * μ/ONS[continuant]

/fa/	WdMin	Dep-μ	* μ/ONS[continuant]
☹ 		*!	*
		*!	
		*!	

Here, no optimal candidate can be selected. Violating lower-ranked constraints does not disqualify [fːa] unless additional assumptions are made about the relevant constraint interactions (e.g. constraint conjunction), but the correct candidate cannot be inferred to be optimal from this tableau. To resolve this unfortunate situation, we should instead stipulate a restriction on where a mora may be inserted. For this purpose, we may resort to positional dependence. Positional mora licensing has been argued for in Bermudez-Otero (2001) and put to a similar use in Topintzi (2010). The definitions of positional μ-licensing and P-Dep-μ below are taken from Topintzi (2010: 111):

(41) Positional μ-licensing: A segment α is positionally μ-licensed by a mora iff μ is the only prosodic unit directly dominating α

(42) P-Dep-μ: A non-positional μ-licenser mora in the output has a correspondent in the input

In other words, the precise outcome blocked by P-Dep-μ is the lengthening of an underlyingly moraic segment, whose moraic status is licensed by its syllable position. Segments in nucleus position are assumed to be moraic by position, as are codas. As there is no indication that codas may surface as bimoraic in Irabu, the more relevant constraint here is against underlyingly monomoraic nuclei surfacing as bimoraic. This becomes clear if we include P-Dep-μ in the table in (39). Note that WdMin must outrank P-Dep-μ, as monomoraic words never occur, while vowels may be lengthened. The minimality constraint is therefore not included in the tableau below. Both of these constraints must also outrank *μ/ONS[continuant], to allow /f/ to occur as [fː]. Furthermore, some version of Dep-IO should remain as well, to prevent epenthesis of either an additional vowel or a coda. Once again, it is not yet clear if any ranking can be determined between Dep-IO and P-Dep-μ given the present input. WdMin remains undominated and is not included here:

(43) P-Dep-μ >> Dep-IO >> *μ/ONS[continuant]

/fa/	P-Dep-μ	Dep-IO	* μ/ONS[continuant]
☞ Wd (μ μ; f a)			*
Wd (μ μ; f a)	*!		
Wd (μ μ; f a r)		*!	

In (43), the epenthesized coda [r] in [far] may be assigned a positionally licensed mora. This does not violate WdMin or P-Dep-μ. Its presence here is nonetheless penalized, as this segment is absent in the input. There does not appear to be any pattern of coda epenthesis in Irabu, so this outcome is expected. Vowel epenthesis appears to be limited to the insertion of short /ï/ [ɨ] to break up illegal consonant clusters (Shimoji 2008: 40), but this may also be interpreted as an underlying [ɿ] per the discussion in 3.1.2. [faː] violates P-Dep-μ, whereas [fːa] violates only the lowest-ranked constraint and is thus optimal. All the voiceless segments that appear to occur as coerced initial geminates (/f s/) may form either singleton or geminate onsets, depending on whether WdMin is satisfied:

(44)	Bimoraic without initial geminate	Bimoraic with initial geminate
	/fau/ [fau] ‘eat’	/ffa/ [fːa] ‘child’
	/sau/ [sau] ‘pole’	/ssu/ [sːu] ‘white’
	/tsuu/ [tsuː] ‘strong’	-
	/tsi.bi/ [tʃĩbi] ‘hip’	-

The distribution of initial geminate [tsː] appears to be very limited in the onset, with the only attested examples being the word ‘(tobacco) pipe’ cited in Kibe (2012/2019: 51) as [ttɛ(ɻ)z] and in Shimoji (2008: 47) as /ccir/ [ˀttʃĩl], and the word /ttsir/ [tʃːil] ‘wear’ (Shimoji 2006: 37). Both words appear to share the same lexical representation. As initial geminates, vowels and codas are moraic in Irabu, this essentially means that these are thus far the only attested examples of bimoraic words with initial [tsː] onsets in these sources. To make matters worse, there appear to be more examples of medial onset geminates featuring [tsː] than initial geminates:

(45) /tur.ccjaa.ki/ ‘while taking’ (Shimoji 2008: 52)
/v.cca/ [ṿ.ttsa] ‘quail’ (ibid.: 59)
/m.cca/ [m̩.ttsa] ‘road:TOP’ (ibid.)

In the examples above, the morpheme /-ccjaaki/ is a suffix that is productively added to verbs to indicate that two actions are taking place simultaneously. This morpheme appears to belong to a highly limited set of suffixes that violate the general rule that all medial clusters should form intervocalic geminates in the language. Recalling that the domain of WdMin in Irabu is argued to be the word, and not the word-plus, it is clear here too that the suffix is far from violating bimoraic minimality. Thus, it seems to be the case that [tsː] initial geminate is not coerced at all, and that it should instead be considered distinctive.

This poses something of a problem. The expected implicational markedness hierarchy of coerced weight in onsets has thus far been suggested to be *sonorants >> *voiced fricatives >> *voiceless fricatives>> *voiced stops >> *voiceless stops, with the possibility of inserting affricates between the voiceless stops and the voiced stops. Here, it appears that voiceless fricatives may be the least marked candidates for initial geminate in Irabu. In other words, the most strongly preferred candidate is neither the most nor the least sonorous. This lends some support to the suggestion that the distinction between consonants that may receive coerced weight in onset position and those that cannot is the feature [voice]. While none of the voiced obstruents may occur as moraic in the onset, it is not necessarily the case that there is a universal ranking between the coerced, unvoiced obstruents. The pattern may therefore be that sonority explains the markedness of initial geminate voiced obstruents, while the markedness of the unvoiced stops may be due to other mechanisms.

As such, the motivation for the behavior observed in Miyako may be found in saliency. While it has been found that initial [sː] is less salient than intervocalic [sː] (Pajak 2013), this merely suggests

a motivation for the general markedness of initial geminates. However, if saliency is weaker in this position, it may be the case that contrasts involving geminates with weaker perceptibility should be more marked. Kawahara (2007) finds this to be the case for intervocalic geminates, as intervocalic sonorant geminates appear less salient than obstruent geminates and are therefore more marked cross-linguistically. Dmitrieva (2017) finds the inverse result in a perception study in which "listeners were better at identifying intervocalic sonorants (liquids and nasals) than for intervocalic obstruents, voiceless stops in particular", noting also that the voiceless alveolar fricative /s/ patterned with the sonorants in this context (39-40). The argument raised therein is, among other things, that fricatives, sonorants and nasals may be more perceptible due to their status as "filled intervals" as opposed to "empty intervals" (ibid.). In the same study, listeners showed a greater ability to discern initial geminate fricatives over final ones. This suggests that the preference for [fː] and [sː] in the Irabu initial geminate is driven by perception.

As far as I am aware, there are hitherto no studies comparing the saliency of initial stop geminates and initial fricative geminates beyond the alveolar segments, let alone ones that compare voiced and voiceless versions of these. Pajak (2013: 17) tentatively finds higher perceptibility in initial [zː] compared to initial [sː], but as the assumption here is that coerced [zː] is highly marked, this finding is of less importance here. Therefore, we may consider what acoustic data has been collected for the Irabu initial geminates. As may be expected, initial geminates in Miyako are clearly shorter in duration compared to intervocalic ones. Matsuura (2012) shows that the durational ratios of singletons to geminates are as follows (156):

(46) a. Intervocalic

Consonants	Ratio
[t] and [tt]	1:2.02 – 2.86 (Irabu)
[ts] and [tts]	1:2.47 (Kugai)
[z] and [zz]	1:2.33 (Kugai)

b. Word-initial

Consonants	Ratio
[n] and [nn]	1:2.24 (Irabu)
[f] and [ff]	1:1.45 (Irabu) 1:1.32 (Kugai)
[s] and [ss]	1:1.51 (Kugai)
[v] and [vv]	1:1.70 (Kugai)

While further phonetic and acoustic research is needed to determine whether it is likely that perception can explain the relative markedness of the fricatives and stops in Irabu initial geminates, the data in (49) suggests that the initial geminates are durationally shorter than intervocalic geminates. Although the nasals appear to be considerably lengthened, the fricative initial geminate is relatively short. Unfortunately, no direct comparisons are made in this data between intervocalic and initial geminate versions of the same segments, such as comparing initial and intervocalic geminate [fː]. As such, this remains a future subject of research. In addition to duration, it must be determined what other phonetic correlates may be present in Irabu, and whether these are more prominent for the fricatives than they are for the other obstruents.

Furthermore, it is worth considering whether the contextual occurrences of the initial geminates are relevant to their saliency. Because the fricatives are generally not permitted in coda position in Irabu, they are likely to be preceded either by sonorants or vowels (or silence). This may further aid listeners in perceiving length. Regardless of the underlying cause, it is apparent that /f s/ are the least marked initial geminates in Irabu. We must therefore assume that a markedness constraint, tentatively given here as *μ/ONS[continuant] is ranked low enough to coerce moraic voiceless fricative onsets. Because it is proposed here that no voiced segments may receive coerced weight in Irabu onsets, the constraint *μ/ONS[voice] will be used instead of specifying a separate markedness constraint for each segment. Although it is possible that constraints may be given for each individual segment in the grammar, it will be shown here that the same outcome can be generated using a limited set of constraints. We may further specify this to be ranked as *μ/ONS[voice] >> *μ/ONS[continuant], only permitting unvoiced fricatives to be coerced in onset position. This ranking, and indeed *μ/ONS[voice] itself, is furthermore justified per the claim in Topintzi (2006) that voiced segments do not receive coerced weight cross-linguistically. We should therefore expect either *μ/ONS[voice] or a specification of *μ/ONS for each voiced segment to outrank all other relevant place or manner feature specifications of *μ/ONS, as a voiceless obstruent should always be preferable to a voiced one in a moraic onset. These constraints and their rankings will now be examined.

As has been established, none of the sonorants occur in the initial geminate through coercion. As will be argued later in this section, this stipulation also applies to the sonorant allophones of the voiced fricatives. Because the voiced fricatives do not occur in this context, it appears that it is unnecessary to propose multiple feature specifications such as [+continuant, -voice] in the lowest-ranked constraint. As long as a constraint against voiced moraic onsets outranks the markedness of moraic continuant onsets, the only fricatives that will be long due to this coercion will be voiceless. However, it must still be expected that monomoraic inputs such as /ba/ should be repaired. Therefore, the most feasible solution appears to be to rank P-Dep-μ and Dep-IO between WdMin and initial geminate continuant markedness. This ensures that /fa/ is repaired to [fːa] and that /ba/ may still

surface as bimoraic, e.g. [baa]. I therefore propose the following rankings, illustrated in the tableaux below. Note that the dependency constraint used below is P-Dep-μ, as it appears that nucleus lengthening is preferred over the epenthesis of a heterorganic vowel. Dep-IO therefore likely outranks P-Dep-μ:

(47) WdMin >> P-Dep-μ >> *μ/ONS[continuant]

/fa/	WdMin	P-Dep-μ	*μ/ONS [continuant]
fa	*!		
☞f:a			*
fa:		*!	

In (47), bimoraic minimality enforces the coercion of [f:] and [s:] in the onset. The monomoraic candidate is highly marked, and accordingly, no words of this type occur in Irabu. P-Dep-μ constrains the lengthening of moraic [a], causing the candidate with bimoraic [a:] to be less optimal than the candidate with the moraic onset, [f:a].

(48) WdMin >> *μ/ONS[voice] >> P-Dep-μ

/gi/	WdMin	*μ/ONS[voice]	P-Dep-μ
gi	*!		
g:a		*!	
☞gi:			*

In (48), a constraint on voicing in the initial geminate outranks P-Dep-μ. This appears to be a fairly safe assumption, as voiced stops never receive coerced weight in the onset position in Irabu. The relevance of [voice] here finds some support in the suggestion that voiced consonants never receive coerced weight in the onset (Topintzi (2006), as well as the general Japonic preference to avoid voiced geminates (in any position).

(49) WdMin >> *μ/ONS[stop] >> P-Dep-μ

/ka/	WdMin	*μ/ONS[stop]	Dep-IO/P-μ
ka		*!	
k:a	*!		
☞ka:			*

Above, (49) shows that stops are marked in the initial geminate, ensuring that they do not receive

coerced weight in this position. The winning candidates given in the tableaux (47, 48, 49) all also appear as words in Irabu. The following examples use the transcription found in Shimoji (2008):

(50) Irabu minimal words

a. /fːa/ [fːa] ‘child’ (Shimoji 2008: 72)

b. /giː/ [giː] ‘tree’ (Shimoji 2008: 114)

c. /kaː/ [kaː] ‘skin’ (Shimoji 2008: 73)

The three rankings in (47, 48, 49) show that WdMin must necessarily outrank all of the other constraints being discussed here. Furthermore, *μ/ONS[continuant] must be outranked by the dependency constraints to ensure that onset coercion is preferred over epenthesis or nucleus lengthening. The outcome of these three rankings is therefore that in Irabu, only the voiceless continuants /f/ and /s/ may be coerced in onset position by WdMin. The ranking of *μ/ONS[stop] and *μ/ONS[voice] will be explained in the next section.

3.1.6.2 Initial [ʊv] and [ʐz]

This brings us to the obligatorily long segments discussed previously. Following from the previous discussion regarding /ts/, it is assumed here that only /v ž/ are obligatorily moraic in syllable-initial position. Thus, there is a clear difference between these segments and /f s/. However, it does not appear that the length of these segments is coerced by the bimoraic minimality constraint, as they occur as moraic regardless of the environment. It may therefore be the case that these segments are obligatorily moraic in general. As seen in (51), these segments are always moraic in the nucleus or coda (final or intervocalic geminate):

(51)	/vva/	[ʊva] ‘2SG’	$C\text{ː}^{\mu}V^{\mu}$	(Shimoji 2008: 69)
	/žza/	[ʐza] ‘father’	$C\text{ː}^{\mu}V^{\mu}$	(ibid.)
	/pav/	[paʊ] ‘snake’	$CV^{\mu}C^{\mu}$	(Shimoji 2008: 41)
	/paž/	[paʐ] ‘man’	$CV^{\mu}C^{\mu}$	(ibid.)
	/aʊ.ʊa/	[avva] ‘oil’	$V^{\mu}C\text{ː}^{\mu}V^{\mu}$	(Shimoji 2008: 56)
	/taž.žasɨ/	[tazaasɨ] ‘bind’	$CV^{\mu}C\text{ː}^{\mu}V^{\mu}CV^{\mu}$	(ibid.)

The syllable structure in (51) is provided to indicate the morae assigned in each item. In (51), it is clear that these segments also only occur as voiced fricatives when forming an intervocalic geminate. Therefore, it appears equally likely that these segments must be specified as moraic in the grammar. One option then is to propose a constraint similar to Vowel-Mora (Rosenthall 1994: 36):

(52) Vowel-Mora (V-Mora)

For every vocalic root node rt_i, there is a mora μ_i.

The above constraint does not technically imply that vowel segments are inherently or underlyingly moraic. Instead, it is violated by vowels that surface without a link to a mora node. The notion would then be to suggest something along the lines of the following constraint for /v/ and /ž/:

(53) v/ž-Mora (VZ-Mora)
For every v/ž root node rt_i, there is a mora μ_i.

While the literature differs on whether vowels should always be considered moraic in the input, vowels are generally assumed to be inherently moraic, though rare counterexamples do occur, such as the non-moraic schwa in languages like Piuma Paiwan (Shih 2018). It must therefore be asked whether the modified constraint in (53) finds any support in cross-linguistic evidence. McKeever (2014) suggests that glides in the Fur language are moraic in all positions, as they cannot close syllables that contain long vowels (only non-moraic codas are permitted). Denzer-King (2012) argues that Blackfoot /s/ is inherently moraic, indicating that /s/ occurs as extrasyllabic where it would otherwise form an illegal onset cluster. A similar approach is taken in Lin (1997), where it is argued that Piro features moraic extrasyllabic consonants as well. Deletion of the Piro extrasyllabic consonants triggers compensatory lengthening, but the consonants are otherwise ignored by all other phonological rules. In Irabu, the foot is suggested to be binary or ternary (Shimoji 2008: 90). As such, extrasyllabic, moraic /ʋ/ could be licensed if, as Lin suggests for Piro, "an unsyllabified mora is in principle licensed by the foot unless it is adjacent to unfooted syllables" (Lin 1997: 415). Problematically, however, the foot itself is still minimally bimoraic. While the approach in Lin (1997) may license a third mora in a binary foot, it does not describe patterns such as bimoraic [ʋva], which cannot be more than bimoraic. Both morae must therefore be linked to a foot, making it difficult to interpret this as a case of extrasyllabicity.

Although Miyako codas always occur with moraic weight, the relation between segmental identity and weight follows the expected sonority relation (Zec 1988, 1995). Miyako singleton codas are invariably of a high sonority, generally nasals or other sonorants. In Irabu, this is similarly restricted to nasals, liquids, or sonorant allophones of the voiced fricatives. The same sonority relation applies to the occurrence of the segments [ʋ] and [ʐ̩] in nucleus position. In both these positions, moraic weight is exactly what is expected, and the occurrence of these specific segments is predictable. Only these segments' seemingly obligatory weight in onset position then stands out as peculiar, suggesting that it may be preferable to avoid stipulating that [ʋ] and [ʐ̩] are underlyingly moraic, or that they are coerced into moraicity through a constraint such as the one suggested in (53) above.

Notably, however, Shimoji (2011: 84) suggests that words such as /vːa/ and /žːa/ are "underlyingly /va/ and /ža/ (moraic C + V), and this impermissible phonotactics is fixed by GCI (e.g. /va/ > /vva/)". GCI in this context refers to Geminate Copy Insertion, which as discussed previously is

the term Shimoji uses to refer to the lengthening of singleton segments into geminates. This is required if it is assumed that onsets are non-moraic and geminates occupy two C-slots, as it would then be suggested that the CC geminate sequence is linked to two syllables. Following the Moraic Theory representation of geminates, however, this could be stated more simply as underlying /$v_{\mu}a_{\mu}$/ and therefore as [vːa] in the surface representation. The moraicity of /v/ in this context is therefore distinctive, not coerced.

Singleton initial /v/ does appear to be quite rare across the Miyako languages, occasionally occurring as a medial onset, such as in Uruka Miyako [tuvi] ‘to jump’ (Kibe 2012/2019: 286) or word-initially instead of /w/ in Kugai Miyako [vaa] ‘pig’ (Matsuura 2012/2019: 152). This would appear to lend some support to the possibility that /v/ is inherently moraic in languages where [ʊv] or [vː] occurs, and there is no length contrast between this segment and singleton [v]. However, this does not explain the phonetics of the Irabu /v/ and /ž/ initial geminates, which are given as [ʊv] and [ʐ̩z] in Shimoji (2008, 2011). If these sequences are represented as a segment linked to a single timing node, such as the mora, then the geminate should not be expected to feature two specifications for manner of articulation over the course of its duration. Assuming that this sequence is indeed analyzed as a geminate, we must also expect that the sequences /CV/, /CːV/ or /CCV/ may be present in the input, depending on whether the geminate is coerced or distinctive. If the geminate corresponds to two segments in the input with two respective correspondents in the output, /C_1C_2V/ → [C_1C_2V], then it does not appear reasonable to analyze it as a true geminate, but rather as a consonant cluster. The same applies if underlying CːV is output as a sequence of C_1C_2V, unless a CV-tier representation of the geminate is assumed. For [ʊv], [ʐ̩z] sequences to be analyzed as initial geminates, we would then have to assume that while the onset is one geminate segment in the surface representation, phonetic or acoustic qualities cause it to be produced or perceived as featuring two manners (ostensibly [+sonorant], [-sonorant]) in the same place of articulation. Another approach would be to assume that these sequences are heterosyllabic, similar to the Irabu words with initial syllabic nasal + obstruent clusters. The latter option is more compatible with the Irabu phonotactics given in Shimoji (2008, 2011) and will be explored in the following section.

3.1.6.3 Heterosyllabic [ʊv] and [ʐ̩z]

The heterosyllabic approach can be summarized as follows:

(54) #ʊv(V) and #ʐ̩z(V) are syllabified as #Ç̩[sonorant].C[obstruent]
(V) here stands for a vowel segment.

This would entail that coerced onset weight does not occur with these segments, but rather that there is coercion in nucleus position. This would simplify the constraint hierarchy assumed to be responsible for Irabu initial geminates, as it would no longer be assumed that any voiced segments receive coerced weight in onset position. As predicted, then, the only voiced segments that may be

geminate in initial position are distinctively voiced. Consequently, it should be expected that repair occurs through resyllabification with the approximants [ʋ] and [ɻ] as a nucleus followed by a simple CV syllable. It is therefore suggested that the underlying representations of these words may follow those given in Shimoji (2008, 2011), namely /vva/ and /žža/, represented as CCV rather than C:V. Because *Complex rules out non-geminate /vv/ and /žž/ in the onset, the initial fricative is syllabified as a nucleus.

Some support for this is found in the structure of the NN monosyllables, which are restricted to the nasals [n:] and [m:] and otherwise the approximants [ʋ], [ɻ], or [l]. According to (54), the moraic structure of the word would be unchanged, as geminate onsets and nuclei are both taken to be moraic. This is also supported by the fact that while /ž/ corresponds to /z/ in most other Miyako variants, the sequence given as onset *žž-* in Irabu, e.g., *žžu* 'fish' Shimoji (2008: 472) has elsewhere been phonetically transcribed as a VC sequence [ɭz] in fieldwork transcriptions of Irabu. In other words, this sequence has been recorded as heterorganic, with the first segment being vocalic. It furthermore appears that another cognate of Miyako initial /z:/ replaces the initial geminate with a vocalic component. Below I compare Irabu and Ikema:

(55) Attested Irabu initial [ɭz-] sequences:

	'fish'	'scythe'	'father'
Irabu	ᶻɭzu	ɭzaɾa	ʊja
Ikema	zzu	zzaɾa ~ ᵈzaɾa	zza ~ ᵈza

In (55), the transcriptions for 'fish' (Pellard & Hayashi 2012/2019: 23), 'scythe' (Pellard & Hayashi 2012/2019: 48), and 'father' (Kibe 2012/2019: 207, 210) indicate either an interspeaker variation, or a difference in transcription between the researchers in question. If it is indeed the case that /#žž-/ in fact yields [#ɭz-], then this implies that Irabu [ɻz] is not syllabified as an onset. Due to *Complex, it should be expected that it would be less costly to syllabify a highly sonorous segment such as [ɭ] as a nucleus rather than producing a complex onset. This applies in particular as the sequence [ᶻɭz] in [ᶻɭzu] should not be syllabified as a complex onset, but rather as a heterosyllabic sequence of a nucleus and onset. However, if the underlying representation of these initial sequences is taken to be an onset cluster, then the presence of [ɭ] is no longer required. While it is entirely possible that either transcription is accurate, or indeed that both are, [ɭz] and [ɻz] should both yield heterosyllabic outputs when preceding a vowel. The question then becomes what motivates this syllabification in the grammar. Regardless of whether the transcription in Pellard & Hayashi (2012/2019) or Shimoji (2008, 2011) is assumed, the markedness constraint Onset (ONS) must be present to prevent sequences such as [Ɍ.V]. ONS is justified in Irabu as /j/ does not occur in the surface representation as a separate segment except to satisfy ONS in syllables where no other onset is possible. ONS has been proposed

as an undominated constraint in Ikema (Celik & Takubo 2014), lending further support to its application here. In (56-a) below, the constraint *μ/ONS[voice] is vacuously satisfied as no moraic onset occurs.

(56) a. ONS, *μ/ONS[voice]

/ɳzu/	ONS	*μ/ONS[voice]
ɳz.u	**!	
☞ɳ.zu	*	

b. *μ/ONS[voice] >> ONS >> Ident(voice) >> Ident(manner)

/zzu/	*μ/ONS[voice]	Ident(voice)	ONS	Ident(manner)
a. z̩zu	*!			*
b. sːu		*!		
c. ☞z̩.zu			*	*
d. z̩z.u			**!	*

Candidate (a) in (56-b) should be understood here as a [CːV] syllable with a moraic onset. As this is highly marked, it cannot be coerced from the underlyingly monomoraic /zzu/ (the same would apply to underlying /zu/). [sːu] violates the faithfulness constraint Ident(voice), whose presence is justified here as devoicing does not appear to occur in Irabu. Voiceless segments do appear to become voiced in an intervocalic voicing context, but this is assumed to involve voicing agreement. Because no words may end in a voiceless segment, there is no context at the word or phrase level in which an onset would be motivated to devoice. While both candidates in (56-a), as well as (c) and (d) in (56-b) violate ONS, one of the two candidates incurs two violations of this constraint as it features two onsetless syllables. Not included in this table is the coda condition that applies in Irabu, as [z], like [v], is not permitted to form an Irabu singleton coda, and a voiced fricative coda segment would have to surface as a sonorant to be permitted. This would still incur the same violations of ONS, rendering the candidate impermissible.

The analysis in (56-a) proceeds on the assumption that the initial segment is a vowel, while (56-b) shows that the initial segment may simply be a fricative that surfaces as a syllabic sonorant. The latter solution appears to be more appropriate for [ʊv] than the former. Alveolar [s] or [z]-like frication is expected from the vowel [ɿ], but labiodental [ʋ] is phonetically quite dissimilar. Accordingly, Pellard & Hayashi (2012/2019: 43) record some form of geminate voiced labiodental for the word ‘you’ for at least twelve Miyako varieties, seven of which are [vva], three given as [ʊʊa], and one other dialect (Karimata) featuring [ʊva]. In other words, none of the researchers involved in the

NINJAL research report (Kibe 2012/2019) suggest that [ɽ] is present in any of these words. Interestingly, the Irabu word is given as [jaː], ostensibly then homophonous with *jaa* 'house' (Shimoji 2008: 169). There is little other data on this word, however, as there appear to be very few Irabu words that show word-initial /v/. 'sell' is given as /vv/ [ṿː] (Shimoji 2008: 50), where [ʋ] appears as part of an initial geminate when a vowel is suffixed, e.g., /vvi/ [ʋvi] 'sell-IMP' (Shimoji 2008: 69). Due to the limited distribution of these onsets, it is not unfeasible to suggest that because initial /v/ is rare in the Miyako languages, the occurrence of initial geminate /v/ is distinctive.

To summarize this section, it has been proposed that [ʋv] and [ʐz] are not geminate onsets where they precede a vowel, but rather that they are sequences of a syllabic sonorant and an obstruent onset. Evidence from the other Miyako dialects suggests that [ʐz] may be the VC sequence [ɽz], but a similar claim cannot be made for [ʋv]. Due to the fricative properties of the vowel [ɽ], it is assumed here that the [ɽz] sequence transcribed in Pellard & Hayashi (2012/2019) may be interpreted as either indicating the presence of a vowel or the presence of a syllabic fricative. Both interpretations support the analysis proposed here, however, as the salient point is that [ʐː] and [ʋː], not to mention initial [zː] and [vː], do not occur as Irabu onsets. If we fully assume the transcriptions of these sequences in Shimoji (2008, 2011), the assumption may either be that [ʋv] and [ʐz] are genuine, distinctive geminates, or that they form Ɛ.CV words. In the following section, I will examine the former possibility, as well as the occurrence of distinctively moraic onsets in Irabu.

3.1.6.4 Distinctive onset weight

For the distinctively moraic onsets, the following faithfulness constraints from Morén (2003: 294) may apply. The descriptions below have been simplified for reasons of space:

(57) a. MaximalityLink-Mora[SEGMENT] (MaxLink-Mora[SEG])
Assign one violation mark for each mora link that is not associated with a segment in the output and that is associated with a corresponding segment in the input.

b. DependenceLink-Mora[SEGMENT] (DepLink-Mora[SEG])
Assign one violation mark for each mora link that is associated with a segment in the output that is not associated with the corresponding segment in the input.

As coerced weight implies that weight is absent in the underlying representation, the maximality constraint does not apply to the coerced examples in discussed in the previous section. What must be specified next are the identities of the segments to which these constraints apply. To review, the segments that have been argued here to be distinctively moraic in the initial geminate are as follows:

(58) Sonorants: /m n r/
Stops: /t ts/

Distinctive weight suggests that the (non-)moraicity of these segments should not be altered in the onset of the output word. An OT analysis of this issue must therefore be able to account for the following: First, the underlyingly moraic segments must not lose their morae, and second, where these segments are underlyingly non-moraic, they must not receive coerced weight. Thirdly, because all constraint rankings are assumed to exist in parallel, the constraint ranking arguments made here cannot cause a sub-optimal candidate to be selected in other areas of the grammar.

3.1.6.5 Distinctively moraic sonorant onsets

To account for the sonorant initial geminates, the content of MaxLink-Mora[SEG] is specified here as MaxLink-Mora[sonorant]. This means that any sonorant that is associated with a mora unit in the input must remain as such in the output. For this ranking argument, we do not yet need to stipulate a ranking of DepLink-Mora, as the markedness of the moraic sonorant onset effectively also penalizes coercion. It must therefore be assumed that if distinctively moraic sonorant onsets occur, MaxLink-Mora[sonorant] must outrank *μ/ONS[sonorant].

(59) MaxLink-Mora[sonorant] >> *μ/ONS[sonorant]

μ μ m a	MaxLink-Mora [sonorant]	* μ/ONS [sonorant]	DepLink-Mora [sonorant]
☞ a. μ μ m a		*	
b. μ μ m a	*!		*

In (59), only the ranking of MaxLink-Mora is relevant. The ranking of *μ/ONS[sonorant] with regard to DepLink-Mora only becomes relevant where the input is monomoraic, as the markedness constraint blocks coercion:

(60) MaxLink-Mora[sonorant] >> *μ/ONS[sonorant] >> DepLink-Mora[sonorant]

μ m a	WdMin	MaxLink-Mora [sonorant]	* μ/ONS [sonorant]	DepLink-Mora [sonorant]
a. μ μ m a			*!	*
☞ b. μ μ m a				*
c. μ m a	*!			

In (60), lengthening the nucleus avoids violating the minimal word requirement and *μ/ONS[sonorant]. The fully faithful monomoraic output [ma] is ruled out by the minimal word requirement, while the coerced nasal onset is blocked by markedness. It may initially appear to be a problem that [sonorant] is understood to include the approximants in the language, as the glides /j/ and /w/ never surface as moraic onsets. These segments are expected to pattern differently from the other sonorants, as they are the most sonorous in the consonant inventory. Nevins & Chitoran (2008) suggest that what separates glides and their respective vowel counterparts are the feature specifications [-consonantal] and [-vocalic]. In the context of onset position, however, it appears it is unnecessary to specify markedness for *μ/ONS[-vocalic], as [-consonantal] captures a full range of prohibited sonorant moraic onset segments, i.e., glides. This ranking also prohibits the occurrence of moraic vowels in onset position, but this is to be expected.

(61) *μ/ONS[-consonantal] >> MaxLink-Mora[sonorant] >> *μ/ONS [sonorant]

	* μ/ONS [-consonantal]	MaxLink-Mora [sonorant]	*μ/ONS [sonorant]
μ μ μ j a m a			
a. μ μ μ j a m a	*!		*
☞ b. μ μ j a m a		*	

In the above tableau, candidate (61-a) violates both *μ/ONS [-consonantal] and *μ/ONS [sonorant], as /j/ is both [-consonantal] and [sonorant]. However, only the higher-ranked constraint is relevant here. As such, distinctive weight is blocked for the approximants /j w/.

The use of [-consonantal] may seem somewhat arbitrary. After all, the same outcome may be expected if the markedness constraint were specified as [approximant]. However, [l] very clearly patterns with the nasals in Irabu, and is also clearly attested as an initial geminate. Thus, while it is entirely possible to specify these constraints even further to *μ/ONS[j] and *μ/ONS[w], using the feature *[-consonantal] allows for a more parsimonious analysis. With regard to the specifications of MaxLink-Mora, it is argued in both Morén (1999) and Topintzi (2006) that the segmental restrictions in distinctive moraicity is largely arbitrary. The specifications of MaxLink-Mora given here might therefore successfully be defined for each individual segment that is present in the inventory, i.e., MaxLink-Mora[n], MaxLink-Mora[m], and so forth. As is shown in the following section, the same outcome can be derived through a smaller set of constraints through reference to manner and place specifications.

3.1.6.6 Distinctively moraic obstruent onsets

Next, we must account for initial moraic /t/ and /ts/. It has been noted that [tː] onsets are very rare in Irabu, and only a few are described by Shimoji (2008: 44, 56):

(62) i. /ttjaa/ [ʔtʲaː] ‘then’
ii. /tti.gaa/ [ʔtti.gaː] ‘then’
iii. /ttar/ [ʔtta̠] ‘came’

However, rare does not mean non-occurring. The grammar must therefore be able to generate the attested outputs in (62). Similarly, [tsː] onsets do not occur frequently, with only a few attested examples:

(63) i. /ccir/ [ttʃi̠] ‘pipe’
ii. /-ccjaaki/ simultaneous converb suffix (medial onset geminate)
iii. [tstsu ~ tsttsu] ? ‘the moon’ (Pellard & Hayashi 2012/2019: 35)

(63-i) above provides a full morphemic word with an initial [ts] geminate, while (62-ii) provides an example of a medial [tsː] geminate Shimoji (2008: 52) indicates that this suffix is syllabified as featuring a geminate onset when appended to a verb: “/jum-/ ‘read’ + /-ccjaaki/ > /jum.ccjaa.ki/ ‘while reading’ (GVC.CCGVV.CV)”. Example (63-iii) is a phonetic transcription from field research as described in Kibe (2012/2019). The question mark next to the transcription is present in the source and is assumed here to indicate uncertainty on the exact phonetic form of this output. The segments therein appear to indicate another medial onset geminate. MaxLink-Mora[stop] >> *μ/ONS[stop] at first appears to provide the expected pattern in (64-a) below. However, in (64-b), it does not accurately distinguish between the velar stops and the alveolar stops:

(64) (a)

μ μ μ t a r /tːar/ ‘came’	MaxLink-Mora [stop]	*μ/ONS [stop]
☞ a. μ μ μ t a ɭ		*
b. μ μ t a ɭ	*!	

(b)

μ μ μ k a r /k:ar/ (not attested)	MaxLink-Mora [stop]	*μ/ONS [stop]
💣 a. μ μ μ k a ɭ		*
☹ b. μ μ k a ɭ	*!	

As /k/ is erroneously permitted to be distinctively moraic in (64-b), it must be assumed that the specification in MaxLink-Mora cannot simply be given as [stop]. Morén (2003: 295) does not provide a more specific iteration of this constraint than [obstruent] and [sonorant]. To distinguish between the places of articulation involved here, i.e., the alveolars and velars, as well as the labials, we must evidently provide either a more specified MaxLink-Mora constraint or a more specified *μ/ONS constraint. The former will be explored below, as it requires fewer assumptions concerning *μ/ONS:

(65) (a) MaxLink-Mora[coronal] >> *μ/ONS[stop]

/t:ar/	MaxLink-Mora [coronal]	*μ/ONS [stop]
☞t:aɭ		*
ta:ɭ	*!	

As indicated in (65-a) above, one option is to suggest that MaxLink-Mora constraint be specified as [coronal]. In addition to correctly predicting that underlying [t:] should surface faithfully in onset position, this also implicitly predicts that the alveolar fricative /s/ and the nasal /n/ may feature distinctive weight in Irabu. While the latter has been assumed thus far, the former has not been analyzed as such. However, given that the segments that are subject to coerced weight are by definition unmarked in this context, it should not be unexpected that segments that may feature coerced weight may also be distinctively moraic. Therefore, as Irabu shows cases of /s/ and /f/ geminate onsets that are not clearly motivated by bimoraic minimality, it must be anticipated that weight may be lexically specified for these onsets as well. Thus, [coronal] resolves the patterning of /t/ and /ts/ onset geminates, but a featural specification for /f/ is still needed.

The voiceless stop /p/ must still be blocked, so it is assumed here that *μ/ONS[stop] outranks MaxLink-Mora[labial]. Note however that in order to prevent the voiced fricatives from being

distinctive in the onset, *μ/ONS[voice] must also outrank MaxLink-Mora[labial], as /v/ and /ž/ are not analyzed here as receiving onset weight in initial position. Furthermore, *μ/ONS[voice] must outrank MaxLink-Mora[coronal] or underlying initial /dː/ would be permitted. Here and elsewhere in this section, it must be assumed that WdMin remains undominated. This constraint will therefore not be included in the following tableaux:

(66) *μ/ONS[voice] >> MaxLink-Mora [coronal]

/dːa/	*μ/ONS[voice]	Ident(voice)	MaxLink-Mora [coronal]
dːa	*!		
tːa		*!	
☞daa			*

As discussed above, [dːa] is banned for being voiced in the initial geminate. Because devoicing may not occur to repair an illegal onset, Ident(voice) must outrank MaxLink-Mora[coronal] in the above ranking. Below, in (56), *μ/ONS[stop] prevents the surface occurrence of underlying initial geminate /p/, though /fːa/ [fːa] may still occur.

(67) *μ/ONS[stop] >> MaxLink-Mora [labial]

/pːa/	*μ/ONS[stop]	MaxLink-Mora [labial]
☞paa		*
ppa	*!	

In addition to the rankings in (54-56), it is expected that WdMin must also outrank the markedness constraints against, and MaxLink-Mora constraints concerning, the sonorants. It does not appear to be necessary for WdMin to be dominated by the constraint against vowels in the onset, although this is unattested in any language. Provided *μ/ONS[-consonantal] is ranked above either dependence constraint, the grammar will always favor epenthesis or vowel lengthening over permitting an illegal vowel onset. I therefore propose the following rankings:

(68)

a. WdMin >> *μ/ONS[-consonantal]

b. *μ/ONS[-consonantal] >> MaxLink-Mora [sonorant]

c. MaxLink-Mora [sonorant] >> *μ/ONS[sonorant]

Although vowels and glides are [sonorant] by definition, their presence in the initial geminate is prevented by *μ/ONS[-consonantal]. It is also clear that while approximant initial geminates will

violate *μ/ONS[continuant], this constraint is taken to be the least dominant in the hierarchy proposed here.

3.1.7.1 Summary of Irabu initial geminates

The following markedness hierarchy of initial geminates has been proposed identified in Irabu:

(69) a. *[-consonantal] >> *[sonorant] >> *[voice] >> *[stop] >> *[continuant]
b. *vowels/glides >> *approximants >> *voiced segments >> *stops >> fricatives

(69-a) above describes the features given thus far, while (69-b) indicates the specific classes that are restricted by the relevant constraint. The Irabu initial geminate markedness hierarchy thus contradicts the expectation that the least sonorous segments (voiceless stops) should be preferred, but offers strong support for the hypothesis that voicing is the distinction between segments that may or may not receive coerced weight:

(70) *sonorants >> *voiced obstruents >> *voiceless stops >> *voiceless fricatives

It must therefore be assumed that the ability of the voiceless fricatives to receive coerced weight is conditioned by some other factor, such as saliency. Support is found for the typological distinction of discussed in Topintzi (2006), as coercion seems to exclusively apply to voiceless segments in the grammar. Rankings for distinctive weight-related constraints are not expected to pattern in accordance with any clear pattern, except that some segments may be more cross-linguistically common as initial geminates than others (e.g., Kraehenmann 2011). MaxLink-Mora is assumed to refer to both place and manner of articulation in the grammar. This is supported by the fact that no other feature specifications capture the distinction between both the sonorants and the obstruents in the language and the labial, coronal, and velar places of articulation. The rankings of the relevant initial geminate constraints in Irabu are thus as follows:

(71) i. Rankings of *μ/ONS:

*μ/ONS[-consonantal] >> *μ/ONS[voice] >> *μ/ONS[sonorant] >> *μ/ONS[stop] >> *μ/ONS[cont]

ii. Rankings of MaxLink-Mora:

MaxLink-Mora[sonorant] >> MaxLink-Mora[coronal] >> MaxLink-Mora [labial]

iii. Full constraint ranking:

*μ/ONS[-consonantal] >> MaxLink-Mora[sonorant] >> *μ/ONS[sonorant] >>
*μ/ONS[voice] >> Ident(voice) >> MaxLink-Mora[coronal] >> *μ/ONS[stop] >>
MaxLink-Mora [labial] >> *μ/ONS[continuant]

In the following sections, I will compare the Irabu initial geminates with those of other Miyako languages.

3.1.7.2 Comparison of initial geminate voiced fricatives in Irabu, Ikema, and Tarama/Minna

Somewhat problematically, one or both of the initial geminates [vː] or [zː] are also found in Ikema (Hayashi 2010) and Tarama/Minna (Aoi 2015). The following data is taken from Pellard & Hayashi (2012/2019: 43, 48):

(72) Miyako initial geminate /v/, /z/

Only [vː]	[vː] and [zː]	
Karimata (ʊva)	Ikema	Kurima
Ōura (ʊː)	Tarama/Minna	Uruka
	Yonaha (ʊː)	Bora
	Kugai	
	Uechi	
	Kuninaka (ʊː)	
	Shimajiri	

None of the languages feature initial [zː] unless they also feature some form of initial [vː], including [ʊː] or the Irabu-like [ʊva]. Additionally, where initial /vː/ occurs and /zː/ does not, the only geminate voiced labiodentals appear to be [ʊː] or the sequence [ʊva]. The data in Pellard & Hayashi (2012/2019: 43, 48) records the Irabu cognate of Miyako [vːa] ‘you’ as [jaː] and [zːaɾa] ‘sickle; scythe’ as [ɳzaɾa]. While this instance of [jaː] may be due to regional or interspeaker variation, it should be noted that the data does not show Irabu featuring initial geminate [v] or [z]. Irabu is therefore not included in this table. The islands of Tarama and Minna are geographically more remote from the Miyako islands, and accordingly are but whose languages also clearly belong to the Miyako language group, appear to feature an even broader range of initial geminates in Sakiyama (2003: 153), where it is stated that they permit initial geminate /b/, /t/, /d/, /s/, /c/, /z/, /š/, /č/, /ž/ and /h/. Sakiyama analyzes these strings as featuring a word-initial /Q/ phoneme (suggested to be a glottal stop) that triggers gemination. K. Shimoji (2004) and Aoi (2015) contradict this view, in part by proposing a distinctly different consonant inventory, which is closer that of Irabu:

(73) Proposed consonant inventories of Tarama & Minna

a. Sakiyama (2003: 150)

/p, b; t, d; k, g; c, z; š, č, ž; m, n; r; h, ‘/

b. K. Shimoji (2004: 95), Aoi (2015: 407):

	Labial	Coronal	Dorsal	
Plosive	p b	t d	k g	
Affricate		c [ts]		
Fricative	f v	s z		(h)
Nasals	m	n		
Liquids		r		
Laterals		(l)		

K. Shimoji (2004) and Aoi (2015) essentially provide the same consonant inventory, except that K. Shimoji includes the lateral /l/ as part of the underlying phonology. Both also include /w/ [w~ʊ] and /j/ as semi-vowels in Tarama. It should be noted that the majority of data available on Tarama and Minna solely describes Tarama, as the former island has essentially been depopulated since the 1960s. Tarama does not feature the voiced affricate /dz/ [dz], and features /z/, which corresponds to Irabu /ž/, such as in the words *zzu* 'fish' (Tarama) and *žžu* 'fish' (Irabu). The syllable structure of Tarama is also fairly similar to Irabu, as seen below:

(74) Aoi (2015: 407): Tarama syllable template: (C0)(C1)(C2)V1(V2)(C3)
Shimoji (2011: 80): Irabu syllable template: $((O_i)O_i)N_1(N_2)(C)$

Like Irabu, the syllable structure of Tarama permits an initial geminate in C0 and the nucleus V1(V2) is generally filled by vowels but may optionally consist of a syllabic, bimoraic consonant, in which case no onset or coda is permitted. Aoi suggests that the C1 position is occupied either by a singleton onset segment or the second part of a geminate segment. C2 may then only be occupied by a glide. Any sonorant may fill the coda slot C3 word-medially or finally, but any obstruents in C3 must be intervocalic geminates. One important difference between Irabu and Tarama/Minna is that only the nasals may form consonantal nuclei. The latter may perhaps be explained by the specification of this segment as a semi-vowel as opposed to an approximant. Given that /w j/ are provided as semivowels in Aoi (2015: 407), and as they do not appear to occur in the onset position C1, they are perhaps best understood as non-moraic components of the nucleus or as a quality of the onset segment in C1, implying that true onset clusters do not occur in Tarama.

The nasal-only syllables are of particular interest here due to the potential interaction between the syllabification of consonants as nuclei compared with their syllabification as moraic onsets. It should be assumed that if markedness for nuclei, i.e. *μ/NUC[SEG1] outranks *μ/ONS[SEG1], as well as the presumed general markedness constraint against moraic onsets (e.g. *μ/ONS) it will be preferable to syllabify a moraic onset [SEG1] as a geminate onset. The inverse ranking will then enable syllabic [SEG1] to surface as a nucleus. Unfortunately, little published data on the specific initial geminate segments that are permitted in Tarama appears to exist beyond that given in the sources mentioned thus far, and much of what does exist is largely only available in Japanese. A few

examples are given in the grammar sketch in Aoi (2015), all but one of which (Tarama *nna* ‘rope’) are identical to their Irabu equivalents. The phonological description in K. Shimoji (2004) provides more examples, such as *mmaga* ‘grandchild’ and *nnabïkal* ‘lightning’ below. In the below examples, the gloss for Tarama *ssam* originally given in Aoi (2015) is ‘loose’. This is likely a typographical error, and Tarama *ssam* ‘louse’ is corroborated in Igarashi (2012/2019).

(75) Irabu – Tarama cognates

Lexical item	Irabu	Tarama
‘child’	*ffa*	*ffa*
‘louse’	*ssam*	*ssam*
‘grandchild’	*mmaga*	*mmaga*
‘white’	*ssu*	*ssu*
‘rope’	*tsɨna*	*nna*
‘lightning’	*(mnapskaɣ)*	*nnabïkal*

The transcription of Irabu *mnapskaɣ* ‘lightning’ in (75) is taken from fieldwork in Kibe (2012/2019: 193). Per the discussion of the fricative vowel /ɣ/ in 1.1, it is quite possible that an underlying /ɣ/ occurs with more frication for this speaker, though further data collection is needed to determine whether this is the case. If so, the structure of the Irabu word may be more similar to the Tarama example than it appears at first.

As the analysis presented in Aoi (2015) is assumed here for this language, I will proceed to another Miyako language that shows similar behavior with regard to nasal-only syllables, namely Ikema. In the next section, I will provide a brief description of Ikema phonology before discussing the issues relevant to the occurrence of initial geminates in this language.

3.2 Ikema

3.2.1 Overview of Ikema phonology

I will begin this section by examining the consonant inventory of Ikema:

(76) Consonantal inventory of Ikema (Hayashi 2010: 169, Takubo 2021: 67)

	Labial	Alveolar	Palatal	Velar	Glottal
Stops	(p) b	t d		k g	
Affricates			c [ts ~ tɕ] ([ts])		
Fricatives	f v	s z [s ~ ɕ] [z ~ ʑ] ([z ~ dz])			h [hʷ ~ ç ~ h]
Nasals	m	n̥ n (ɴ̥) (ɴ)			
Flaps		r [ɾ]			
Approximants	w		y[j]		

The phonemic inventory of Ikema is largely the same as that given in 1.2 for Miyako in general, with the notable addition of the cross-linguistically rare voiceless alveolar nasal /n̥/. Although not noted specifically by either author, /w/ appears to have a very limited distribution in Ikema. Examining the data in Kibe (2012), Hayashi (2010, 2013), Takubo (2021) and Igarashi et al. (2016), the only instances of surface [w] that seem to have been recorded are the words *waa* 'pig' and *waigawaiti* 'with utmost effort', the latter of which is recorded only in Hayashi (2013). Takubo (2021: 77) discusses the possibility of there being an underlying verb stem-final /w/ that is deleted in all surface occurrences, though this does not appear to be necessary to explain Ikema morphophonemics. Pellard & Hayashi (2012/2019: 49) note that historically, Proto-Japonic *w changed to Proto-Miyako *b, and /w/ is in complementary distribution with /v/ in contemporary Miyako.

Hayashi (2010) and Takubo (2021) differ somewhat in their analysis of the Ikema inventory. In the table in (69), the differences given in Takubo (2021) are presented in parentheses. Of particular note here is the allophony of Ikema /z/, which corresponds to Irabu /ž/ and /z/ in other Miyako languages. In Ikema, it has been analyzed as surfacing in free variation as [z ~ ʑ] (Hayashi 2010) or as [z ~ dz] (Takubo 2021). The latter pattern is similar to Japanese, in which /z/ may occur in free variation as [z ~ dz] or [ʑ ~ dʑ] in certain environments (Labrune 2012: 64).

In terms of phoneme inventory, the biggest difference between the accounts in Hayashi (2010) and Takubo (2021) here is found in the nasal series. In Takubo (2021), in addition to the labial and alveolar nasals /m n/, there are also the uvular (assimilating or "placeless") nasals /ɴ/ and /ɴ̥/, the latter of which takes the place of /n̥/ in Hayashi (2010). Shinohara & Fujimoto (2018) assume the same consonant inventory as Takubo. /p/ is given as a contrastive segment by all these authors, though it is noted that /p/ often neutralizes to /h/. In the following section, I will be following Takubo (2021) in transcribing the voiceless nasal as /ɴ̥/ [n̥, m̥]. As such, I will also be distinguishing between the voiced

nasal segments /m, n, ɴ/. While there is neutralization between /m, n, ɴ/ when preceding an obstruent, they are contrastive when geminated or preceding a vowel (examples below modified from Takubo (2021: 68):

(77) **Contrast**

/nna/ ‘turban shell’, /nnna/ ‘all’ vs. /mma/ ’mother’, /mmma/ ‘the head of female priests’

Assimilation

/nta/ [nta~mta] ‘earth’, /nbu/ [mbu] ‘heavy’, /nba/ [mba] ‘disagree’

/nndi/ [nːdi~mːdi] ‘yes’

Note that in the above examples, Takubo opts to transcribe /ɴ/ as /n/ in contexts where /ɴ/ is only distinguishable from /n/ by its distribution, i.e. the apparent nasal-initial clusters, such as those listed in (77). Like Irabu, Ikema features the central high vowel [ɨ], which is always preceded by one of the fricatives /s, z, f/ or the affricate /c/ [ts]. This may either be analyzed as part of the phoneme inventory or as an epenthetic vowel. As minimal pairs exist between /i/ and /ɨ/, it would appear that there is a contrast in at least some environments (Takubo 2021: 3):

(78)

/siba/	‘worry’	/sɨba/	‘lip’
/muːs-ɨ/	‘burn-CONCLUSIVE’	/muːs-i/	‘burn-IMPERATIVE’

The syllable structure of Ikema is similar to that of Tarama. As such, onsets include singleton consonants, initial geminate, or C + /j/, as well as an apparent set of nasal-obstruent onset clusters. Like in Japanese, codas are optional and consist either of a nasal (medially or word-finally) or a geminate (medially). Nuclei may be short (V) or long (VV). Otherwise, the syllable may also consist of a long, syllabic nasal, in which case no onset or coda is permitted. Like in most Ryukyuan languages, the voiced stops are not permitted to be geminate in the onset. Unlike Irabu, however, the velar stop /k/ and the alveolar stop /t/ appear to be fully licit as initial geminates, while /b, d/ are permitted in intervocalic geminates. There are some differences between Hayashi (2010) and Takubo (2021) with regard to permitted initial geminate segments. In both analyses, /ŋ̊/ (or /ɴ̊/) does not form a regular geminate such as *[ŋ̊ː], but instead precedes a voiced nasal in the onset, as in /ɴ̊ɴa/ [ŋ̊na] ‘rope’ or /ɴ̊mu/ [m̥mu] ‘cloud’.

(79) Ikema initial geminates:

/tː, cː, (kː), fː, vː, sː, zː, mː, nː, (ɴ̊)/

Ikema intervocalic geminates:

/tː, dː, cː, kː, pː, bː, fː, vː, sː, zː, rː, mː, nː/

As is clear from (79), initial geminates in Ikema may consist of any voiceless stop, voiced or voiceless fricative, a nasal, or the flap. As such, only the glottal fricative /h/, the sonorants /w, j/, and the voiced stops /b, d, g/ are prohibited from being moraic in the onset. However, although reported in Hayashi (2010) and Takubo (2021), initial [kː] appears to be very rare. The only word attested to feature this onset is /kːunuc̚i/ '9 pieces' (Takubo 2021: 68). Discussing the field data in Kibe (2012/2019), Pellard & Hayashi (2012/2019: 51) state that although /kːunuc̚i/ "appears as 'kukunutsi̚' in the reported data, the variant 'kkunutsi̚' is also encountered". Therefore, it would seem that either initial [kː] is in variation with [kuku] in Ikema, or there is some variation among the different local Ikema variants. It is perhaps for this reason that Shinohara & Fujimoto (2018) make no mention of initial [kː] in their acoustic study of Ikema geminates. If it is indeed the case that the underlying representation of the word reported as either [kːunutsi̚] or [kukunutsi̚] features an underlying geminate, then this suggests that this geminate may be split in the output. Traditionally, geminates are not expected to be susceptible to being broken up by epenthesis (see for instance Kenstowicz & Pyle 1973, Hayes 1986). This rule has been observed cross-linguistically and is generally referred to as Integrity, which is defined as follows in Hayes (1986: 321):

> Integrity: Insofar as they constitute two segments, long segments can not be split by rules of epenthesis.

The premise of Integrity aligns with the analysis of geminates assumed here thus far, as there should be no reasonable way for epenthesis to occur within a long [Cː] segment. As has been the case for many assumptions in rule-based linguistics, the inviolability of this rule has been questioned within the framework of constraint-based phonology. Noamane (2018) reviews Integrity and the version of Integrity proposed by Benhallam (1980), incorporating the latter into an OT account of geminates in Moroccan Arabic. As the geminates in Moroccan Arabic may in fact be broken up by schwa epenthesis, geminate integrity in that language appears to be violable as long as the violation of this constraint is done to comply with a higher-ranked morphophonemic constraint. This may suggest that the underlying representation of this word in fact features a non-geminate /kk-/ onset cluster. Unless this sequence surfaces as a monopositional geminate, it should violate *Complex, which would motivate vowel epenthesis in this context. It is nonetheless unclear why the instances of [kː] that have been recorded may occur as such.

Because of the lack of data on speakers producing [kː], and because of the theoretical predictions this involves, more research will be needed to determine whether this geminate is a matter of free variation, or if /kːunuc̚i/ has undergone reanalysis diachronically, or that the [kː]/[kuku] split is a matter of regional variation. While it does raise some intriguing possibilities, the data on long [kː] is scarce and inadequate for the purpose of analysis. It will therefore be assumed herein that initial [kː]

does not occur consistently in Ikema. This brings us to the question of coercive and distinctive moraicity in Ikema.

3.2.2 Ikema initial nasal-obstruent sequences and nasal sonority

One significant phonotactic difference between Ikema and the other Miyako languages discussed thus far is that onset clusters starting with /n/ may seem to be permitted in words such as /nta/ [nta~mta] ‘mud’ or ’earth’, and /nbu/ [mbu] ‘heavy’. These clusters are described as partial geminates (Hayashi 2010), with the nasal assimilating to the place of articulation of the following consonant. This appears superficially similar to intervocalic geminates in native Japanese phonology, which produce heterosyllabic nasal-obstruent clusters. Kawahara (2006: 549) describes the suffix [-ri] as containing a floating mora μ. When affixed to a word, this causes the second consonant in a mimetic root to geminate, i.e. /tapu+μ+ri/→[tappuri] ‘a lot of’. If the second consonant is a voiced obstruent, this instead results in a nasal coda: /zabu+μ+ri/→[zamburi] ‘splashing’. These forms are not in variation, as *[tampuri] and *[zabburi] are both illegal in Japanese. In terms of constraint ranking, Kawahara explains this pattern using the markedness of the voiced obstruents *VoiObs over a faithfulness constraint on the nasal identity of the coda $\text{IDENT(nas)}_{\text{coda}}$, thus *VoiObs >> $\text{IDENT(nas)}_{\text{coda}}$. The assumption therein is that the voiced obstruent geminate violates *VoiObs twice, making a faithful voiced geminate obstruent candidate less optimal than one in which one segment is made nasal.

It is clear that this specific constraint ranking is inadequate for the Ikema initial geminates, as the voiced obstruents /z/ and /v/ appear to occur as moraic onsets, while /b/ and /d/ are also permitted in intervocalic geminates. If *VoiObs were included in this ranking, this could conceivably affect the initial [mb] sequences (if these were interpreted as underlying /bb/), but the same cannot be argued for [tː]. There is no indication that this geminate is restricted in Ikema, and thus there is no motivation for any constraint ranking that changes initial /tː/ to [nt].

Instead, if /nt-/ is underlying in /nta/, then its surface occurrence may be explained if we interpret these sequences like the Irabu nasal + obstruent clusters in the previous section. Takubo (2021: 67) notes that the contrast between the nasals is neutralized when they “appear as an independent syllable or appear before an obstruent”. This may suggest that the two contexts for this neutralization is in fact a single context, i.e. the contrast between /m/ and /n/ is neutralized when the nasal is syllabic. This supports the analysis of the nasal segment as being placeless in this context, meaning that it is unable to bear a place feature. This approach is identical to that of Itô & Mester (1993), in which the placeless nasal is given as the only singleton segment permitted by the Japanese coda filter.

Krämer & Zec (2020) argue against nasal placelessness, instead suggesting that nasals occupy two spots in the sonority hierarchy. Nasals are then divided into high-sonority nasals and low-sonority nasals, the latter of which are specified for a value of [±continuant]. It is shown therein that in some

languages, nasals "are specified for this feature across the board, and in others only in certain prosodic positions, or not at all" (ibid.: 59). The criteria for identifying high-sonority nasals outlined therein generally involve the relative sonority of nasals and liquids within a given language. If nasals are permitted in the coda while liquids and obstruents are not, then "there is reason to suspect that the language has high-sonority nasals, at least in postvocalic position" (ibid.: 39). Similarly, if "a language with liquids has syllabic nasals, but not syllabic liquids, the nasals must be of the high-sonority type, at least when they are syllabic" (ibid.). As mentioned previously, only nasals may form syllabic segments or singleton codas in Ikema. Because contrast exists in onset position but not when the nasal occurs as syllabic or in coda position, we may analyze Ikema as featuring two contrastive low-sonority nasals /m/ and /n/ in Ikema, in addition to the high-sonority nasal, which is given as placeless /ɴ/ in Takubo (2021: 67). This would make the typology of Ikema nasals similar to the languages Wan and Sentani, which also feature complementary distribution between high- and low-sonority nasals (Krämer & Zec 2020: 47-49).

It should be noted that it is not argued here that the transcriptions given in Hayashi (2010) or Takubo (2021) are inaccurate. Rather, as the Miyako languages are generally considered to be mora-timed and as fewer prosodic processes take place at the syllable level, descriptions of these languages tend not to discuss or indicate syllabification for many lexical items. Takubo briefly mentions a hypothesis related to what is being discussed here in a footnote (Takubo 2021: 67), wherein it is suggested that words such as /nnna/ 'all' and /mmma/ 'the head of female priests' may be syllabified as [nn.na] and [mm.ma], respectively. The nasal sequence in the word-initial syllable here is assumed to assimilate to the onset of the word-final syllable. What I propose is that this analysis also be applied to apparent consonant clusters such as [nta], which are therefore argued to be /ɴta/ [ɳ.ta], in which /ɴ/ represents a high-sonority nasal that in the output is syllabic and moraic. There does not appear to be any contrast between the high- and low-sonority nasals, as low-sonority nasals in this position should be expected to either surface with high sonority (i.e., neutralization), or must otherwise feature some other form of repair to avoid violating *Complex.

It must also be noted that all the Miyako languages discussed thus far are subject to a strict bimoraic minimality constraint. In Ikema, this applies at the morpheme level. As it is held to be the case that no free morphemes may be monomoraic (Takubo 2021: 68-69), we may question where the mora would otherwise be placed in words such as the following:

(80) /nta/ [nta] 'mud' /nba/ [mba] 'disagree'
/nsu/ [nsu] 'miso' /nbu/ [mbu] 'heavy'

It is clear that without the initial nasal, the resulting word is illegal due to minimality, e.g. *[ta], *[bu]. The alternative option would then be to analyze these clusters as being moraic, but this raises the question of why nasal-obstruent clusters should occur in the language when sequences that are more

harmonious in terms of sonority do not. Furthermore, it would be peculiar that these clusters should be moraic. The only other edge clusters that have been proposed for Ikema are C+/j/ clusters, though it is assumed here that these clusters are not analyzed as onset clusters. They also do not appear to carry weight. The /j/ is also fully fused with a preceding consonant where possible, as seen in the following examples:

(81) C+/j/ clusters

/kuː/ [kuː] 'suffer' /kjuː/ [kjuː] 'today'

/cɨː/ [tsɨː] 'breast milk' /cjuː/ [tɕuː] 'dew'

As indeed appears to be the case for all the C+/j/ clusters, minimality seems to be satisfied by the length of the vowel. That C+/j/ is non-moraic is further supported in Takubo (2021: 72-73), where it is argued that there is a constraint specifying that the "number of morae must be the same in the input and output". There are a few cases in which the mora count of the input is reduced, such as when a vowel is changed into a glide. One instance of this is when three consecutive vowels occur in the input, in which case the third vowel changes to a glide. The same occurs when rising diphthongs occur in the language, i.e., *ia→ja, iu→ju*. To compensate for this mora removal, the remaining rightmost vowel is lengthened. Accordingly, we do not see words like *[kju] in Ikema. This would of course not be necessary if C+/j/, or /j/ alone, were moraic. Thus, it appears that /j/ does not in fact serve as a full consonant in these cases, but rather serves as palatalization of the onset or as an on-glide to the nucleus. This being the case, we would be left with only the nasal-obstruent sequence as a proposed onset cluster. As such, we must ask whether nasal-obstruent sequences are in fact tautosyllabic.

Some attention has been given in the literature to the theoretical question of whether partial geminates behave like "full" geminates. Davis (1999) argues that partial (or place) geminates may be lexically specified as moraic, but are not inherently moraic as a rule, citing moraic homorganic nasal clusters in the Bantu language Jita (Downing 1990), as well as evidence against the moraicity of Korean partial geminates. As mentioned previously, moraicity is also observed in Japanese partial geminates (Kawahara 2006). These examples are all of intervocalic partial geminates, however. Japanese codas are assumed to be moraic regardless, and thus provide little evidence for the nature of partial geminates. Topintzi & Davis (2017) cite two examples of languages with moraic initial (non-geminate) clusters: Cypriot Greek and Ponapean. Armosti (2011) shows that Cypriot Greek initial geminates appear to pattern with onset clusters with regard to certain phonological processes. Ponapean is of particular interest here, as it shows the same pattern as has been proposed here for Ikema, namely that of word-initial clusters whose first member is a syllabic nasal and whose second member is a stop (Kennedy 2003). The below examples are taken from Davis & Topintzi (2017: 274):

(82) Ponapean: initial geminates and NC clusters

c. [m̪met] ‘full’
[ŋŋet] ‘to pant’

d. [m̪pek] ‘to look for lice’

As is implied in (82-b) above, the sequence [m̪p] is heterosyllabic, i.e., the word is syllabified as [m̪.pek], which is the analysis that has been presented for the Ikema clusters discussed above. Another point of consideration is word duration. At the word level, the durations of [nta] and [nada] are identical. Unlike the Ikema singleton and geminate onsets, the durations of the initial nasal onsets in [nta] and [nada], i.e. the [n] segments, were not different to a significant degree. While a correlation between duration and moraicity may be present in Ikema, it should be emphasized that duration can generally be considered the primary correlate of gemination (Ehrenhofer et al. 2017: 205). It is suggested here that this is also the case for Ikema. Thus, there is good reason to expect that the initial geminate must show greater duration, and it can reasonably be argued that the same should not necessarily be expected by a syllabic (and therefore moraic) singleton nasal in word-initial position to the same extent. Considering the relative sonority of /n/ and /t/ (as well as /m/ and /b/), the fact that both [nada] and [nta] feature the same duration, that [nta:] is not required, and the fact that there is contrast between [n:ta:] and [nta], it would appear highly likely that /nta/ is bimoraic and disyllabic [n.ta].

Considering all available options, it would appear more parsimonious to assume an analysis in which all Ikema initial nasal-obstruent clusters are analyzed as a syllabic (and thus moraic) nasal followed by an obstruent onset. This makes the initial clusters symmetrical with medial syllable boundaries, as only the nasal may occur as a moraic singleton coda, while other segments may only occur as geminates intervocalically. In other words, heterosyllabic nasal-obstruent sequences may surface both word-initially and word-medially, and the nasal will be moraic in both cases. This is in line with the Syllable Contact Law (Vennemann 1988), which predicts falling sonority from the coda of one syllable and the onset of the next, as is clearly the case for [n.t].

How, then, is this represented in terms of constraint rankings? Firstly, it is clear that there is some constraint interaction involving the markedness of either complex margins in general or complex onsets specifically:

(83) No Complex Onset (*ComplOns) (Prince & Smolensky 1993)
Add one violation mark for each complex onset.

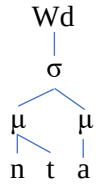
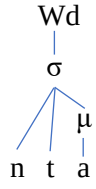
(84) No Complex Margins (*Complex) (Prince & Smolensky 1993)
Add one violation mark for each complex onset or coda.

Violations of *ComplOns are incurred by onsets consisting of more than one segment. It is well-attested that Miyako avoids onset and coda clusters, which provides some empirical support for the

application of these constraints. The coda shows no behavior similar to that of the nasal-obstruent clusters, except in the sense that a syllable-final nasal coda may be followed by an obstruent onset. However, there is no equivalent of the C+/j/ clusters, i.e., *jC]. Therefore, it appears highly likely that initial C + /j/ sequences do not violate *Complex, or that [CjV] sequences do violate *Complex but surface faithfully, implying that it is relevant to rank *ComplOns separately from *Complex. If it is truly the case that complex onsets occur in this language, *Complex would not appear to be sufficient to account for this behavior. Conversely, if both (or neither) margins were obligatorily simple, no specification beyond *Complex would be needed. As in Irabu, however, /j w/ may only be found in onset position where no other onset segment is present. It is therefore assumed here and throughout that any consonant-glide sequence in the onset may only be a cluster in the underlying representation, while the output obligatorily simplifies this to a simple onset. While not directly addressed, the same assumption is made in Celik & Takubo (2014), as all instances of surface Cj are given as C^j.

The first ranking argument here is between word minimality, *Complex, and *μ/NUC[nasal], showing that WdMin and *Complex must outrank the constraint against nasal syllabicity. No ranking can be determined between WdMin and *Complex, as it does not appear to be the case that monomoraic words or complex margins surface in any context.

(85) WdMin, *Complex >> *μ/NUC[nasal]

/ɴta/	WdMin	*Complex	*μ/NUC[nasal]
☞ a. 			*
b. 		*!	
c. 	*!	*	

As is indicated in (86), the initial segment bears a mora in the output. This tableau makes no theoretical assumption about the moraicity of the nasal in the input. There is no indication in the literature that any Miyako nucleus is non-moraic. Therefore, the moraicity of the initial nasal may either be distinctive (/$ɴ^{μ}ta^{μ}$/), or it will be coerced to avoid violating bimoraic minimality (/$ɴta^{μ}$/ →

[nᵘ.taᵘ]). The moraic edge cluster candidate (86-b) is disfavored due to the markedness of the complex onset. In either ranking of WdMin or *Complex, the complex onset candidate (86-b) and the monomoraic output candidate (86-c) are disqualified. (86-a) is therefore selected as optimal. It must be assumed here that a ranking of WdMin, *ComplOns >> *μ/NUC[nasal] is present, as nucleus position is preferred over onset position for the nasal.

A low ranking of *μ/NUC[nasal] is expected here for two reasons: Firstly, syllabic nasals are cross-linguistically common and are predicted to be among the least marked syllabic consonants according to sonority (Zec 1988). Second, like many other Miyako languages, Ikema only permits nasals to be syllabic. As discussed previously, the nasal preceding the obstruent is assumed here to be a high-sonority nasal segment whose place of articulation either assimilates to a following segment or is in free variation between [n ~ m]. We can therefore derive the following rule for the initial nasal-obstruent sequences:

(86) n, m → Ṇ[place$_1$]/_C[place$_1$]V
/NCV/ → [Ṇ.CV]

This can be accounted for in an OT grammar as an emergent property of markedness constraint rankings. In the below tableau, it is assumed that both candidates feature a mora in the onset, whether this is in the syllabic segment or the cluster:

(87) *Complex >> *μ/NUC[nasal]

/nbu/	*Complex	*μ/NUC[nasal]
mb.u	*!	
☞ ṃ.bu		*

In (87), the violation of *μ/NUC[nasal] is permitted as long as *Complex is not violated. Constraints such as *μ/NUC[obstruent] and *μ/NUC[approximant] must necessarily outrank *μ/NUC[nasal]. As obstruent and approximant nuclei never occur in Ikema, so the former constraints are assumed here to be undominated and will not be examined further here. With regard to the nasal-only words, we similarly need to specify that WdMin >> *μ/NUC[nasal], as is implied by (85). What has not yet been explained in (87) is why [ṃ.bu] should be preferred over [ṇ.bu]. This appears to be a fairly classic case of nasal place assimilation. The clearest generalization for this is that the nasals must feature agreement with the place of articulation of the following consonant.

One potential explanation would be to follow Takubo (2021) in assuming that Ikema features /N̥, ɴ/ as placeless phonemes. would be assumed that Ikema, like Japanese, features a nasal that is underlyingly unspecified for a place feature (Itô 1987). This would then allow for feature spreading under the assumption that all segments must have a place feature in the output. Alternative approaches

include analyzing nasal assimilation as a ranking of markedness constraints for specific sequences of /n/ and non-nasal stops within a foot or monomorpheme over a constraint against feature spreading (Kang 1996). Another option would be to propose that a form of the constraint AGREE (see for instance Lombardi 1995) compels the nasal to acquire the place of articulation of the following segment. Because only other nasals or vowels may directly precede nasals in Ikema, place feature spread may only occur in the sense of a nasal receiving a place feature from another segment. For reasons of space, the exact mechanism of place assimilation will not be explored here.

To summarize this section, the following arguments have been made:

(88) i. Ikema surface representations do not feature complex margins.
ii. Apparent nasal-obstruent onset clusters /NC/ are syllabified as heterosyllabic [N̩.C].
iii. Ikema features both low- and high-sonority nasals in complementary distribution, with low-sonority nasals occurring only as (singleton or geminate) onsets.

It follows from (88) that Ikema initial geminates cannot be analyzed as clusters, and thus pattern with singleton onsets except in the sense that they are moraic. Due to the contrast between low-sonority /n/ and /m/, and the lack of contrast between the high-sonority nasals in [nː ~ mː], it is also argued that the Ikema initial geminate nasals cannot be analyzed as disyllabic, e.g. [n.na], [m.ma]. One aspect of the Ikema inventory that has not been explored yet is the Ikema voiceless nasal /n̥/. In the following sections, it is argued that this segment may best be understood as a preaspirated nasal, and its associated geminate segment [n̥n-]/[m̥m-] as a "true", monopositional geminate.

3.2.3 Identity of the Ikema voiceless nasal segment

As mentioned in 3.2.1, Ikema features a voiceless nasal, tentatively given here as /n̥/, which occurs only in initial geminate sequences. Takubo (2021: 68) gives the context of the voiceless nasal as "before /ɴ/, /n/, /m/", and interprets the segment as "a marker for devoicing nasals", which is represented as a segment *h* preceding the nasal (i.e., [n̥na]→[hna]). While generally rare, voiceless nasals occur in several languages, e.g., Oma Longh Kenyan (Blust 2007), Comaltepec Chinantec (Silverman 1996), and particularly among the Tibeto-Burman languages (Chirkova et al. 2018). Research on the latter group in Chirkova et al. (2018) supports the proposal in Bhaskararao & Ladefoged (1991) that there are two types of voiceless nasals; preaspirated nasals, which begin "with a voiceless period characterized by nasal and oral airflow and ending in a voiced period characterized only by nasal airflow", and voiceless aspirated nasals, which begin "with a voiceless period characterized only by nasal airflow and ending in a partially voiced period characterized by simultaneous nasal and oral airflow" (Chirkova et al. 2018: 17-18). Shinohara & Fujimoto (2018) describe the Ikema [n̥na] sequence as a "half voiceless nasal geminate", in which a voiceless nasal component of the geminate is followed by a voiced one.

However, this does not mean that voicing is evenly distributed from the beginning to the end of this geminate. Shinohara & Fujimoto (2018: 264-265) show that the actual voicelessness present in [n̥n] is very short, noting that the "proportion of the voiceless nasal to the entire geminate is remarkably short if we consider it as a moraic segment", while noting that it is nonetheless enough to signal the voicelessness of the geminate segment, as there may be other cues of voicelessness such as pitch and intensity. Chirkova et al. (2018: 19) find that in Burmese voiceless nasals, "the duration of the voiced period amounts to approximately one third of the total duration of the nasal", and that "the voiced closure period is located at the nasal-vowel boundary". In other words, these voiceless nasals feature early voicelessness and become voiced near the vowel. The Ikema [n̥n] features far more voicing, as the initial voicelessness in the Ikema segments is considerably shorter. One immediate distinction to be drawn here is that the Burmese voiceless nasals examined in Chirkova et al. are singleton onsets, rather than geminates of any kind.

Given that voicing is expected to occur in the latter part of the articulation of a voiceless nasal single, we may assume that the same holds true for a voiceless nasal geminate. However, if the voiceless geminate merely occurs as a long version of this segment, it is unclear to what extent this would affect the duration of the voiceless and voiced periods of the nasal. Conversely, if the geminate were represented as two separate segments (or two separate segmental timing units), one voiceless C̥ and one voiced C, then it is unclear what accounts for the reduction of the voiceless unit. Tentatively, the best option appears to be to consider the Ikema voiceless nasal a preaspirated nasal as in Northern Otomi (Palancar 2013), Burmese (Chirkova et al. 2018), or as has been suggested for Proto-Hlai (Norquest 2015). Further research on the acoustics and articulation of the voiceless nasal is needed to gain empirical support for either analysis. As noted above, the /n̥n-/ [n̥n-] and /m̥m-/ [m̥m-] sequences are generally analyzed as a cluster of some sort, with the voiceless nasal preceding [n] or [m] in word-initial position. In the following section, I argue that based on Ikema mora assignment rules and syllable phonotactics, these sequences are likely best understood as geminates.

3.2.4 Geminate identity of the Ikema [n̥n-] onset

It may be questioned here whether the semi-voiceless geminates in words such as [n̥na] are analyzed as geminates. As other Miyako geminates appear to be fully homorganic, and no suprasegmental feature has been proposed to explain the occurrence of the geminate voiceless nasal, it is necessary to examine whether sequences such as [n̥n] may instead be analyzed as clusters. Shinohara & Fujimoto (2018) argue that at least in terms of phonetics, they clearly pattern with geminates in duration and moraicity. The reasoning therein is relatively straightforward: bimoraic minimality does not compel lengthening of the vowel in words such as /n̥na/ to produce outputs such as [n̥naː]. [n̥na] is therefore most likely bimoraic, which entails that both the nucleus and at least some component of the onset must be moraic. In terms of duration, [nna] and [n̥na] are essentially identical, with the onsets [nn] and [n̥n] being nearly twice as long as the onset [n] in /nada/ [nada] 'tears'. Shinohara & Fujimoto

argue that this indicates that duration is a correlate of moraicity in Ikema. However, they also note that past research has established that in Tokyo Japanese, "each moraic element is not isochronous due to incremental segmental duration," but that word duration is "adjusted by all segments at the word level" (Shinohara & Fujimoto 2018: 262). As noted previously, this supports the suggestion that, in Ikema, duration indicates the geminate identity of a particular segment, while word duration indicates the presence of moraic segments in that word.

One consequence of the analyses proposed in this and the preceding sections is that there must be a relationship between the relative markedness of geminate nasal onsets and word-initial nasal syllabics. The input /nːa/ must be expected to surface faithfully as [nːa] and not as [n.a]. Likewise, /n.ta/ must faithfully be output as [n.ta] and not as [nːta]. This may be explained through markedness, as onsetless syllables are cross-linguistically more marked and there would be a preference to syllabify the nasal as the onset to the vowel. This could be viewed as a preference for generating geminates over generating syllabic consonants. Bearing in mind that the only consonants permitted to be syllabic in Ikema are nasals, gemination is clearly the preferred option for repairing phonotactically illegal monomoraic words. It is therefore likely that all specifications of the constraint *μ/NUC[C], with the most relevant constraint here being *μ/NUC[nasal], must outrank all specifications of *μ/ONS[C]. This has no bearing on [n.ta], as monosyllabic [nta] is disqualified by the undominated constraints *Complex and WdMin.

It follows that the constraint on complex onsets must remain undominated here as well, allowing for words such as [n.ta]. As can be inferred from the above, this leaves no reason for the nasal in /nta/ to lengthen, as bimoraic minimality is achieved through the syllabification of the initial cluster. Finally, across Irabu, Tarama and Ikema, it is clear that labial stops, voiced stops and non-nasal sonorants are largely avoided for geminate onsets, as are dorsal segments. Thus, the key distinction between Ikema and Irabu is that [vː], [zː] and [tː] are licit as initial geminates. The question this raises is how this distinction can be accounted for in OT.

First, we must establish which of the Ikema initial onsets are distinctively moraic and which may be coercively moraic. This task primarily involves determining which onsets may not be coerced into moraicity, as any occurrence of these segments as initial geminates must then be distinctive. Following Morén (1999) and Topintzi (2006), it is again expected that distinctively moraic segments may not necessarily form any kind of natural class (though they may appear to do so) and may be arbitrary with regard to sonority and featural specifications. By contrast, the coerced members are predicted to pattern with voiced segments being highly marked and voiceless segments being less marked. Recall that for Irabu, which features the most similar range of permitted initial geminates to Ikema, a distinction could be drawn between geminate onsets and long syllabic consonants. As was concluded in that section, the distinctively moraic onsets were best identified as those that did not

participate in any gemination (and thus coercion) process in onset position, which tended to be the segments that were available to the grammar as nuclei.

Note that the initial geminates that must be distinctively long are clearly the most sonorant in the set, whereas the segments that may undergo coercion are less sonorant. Geminate identity is preferred over syllabicity in situations such as those where a word with an underlying CV monomoraic structure begins with a singleton consonant segment, such as /f/. The only exception to this is the long alveolar stop [tː], which is marginal and also does not occur as syllabic in any cases. In this respect, there is a clear and somewhat drastic difference between Irabu and Ikema, as Ikema only permits nasals to be syllabic, yet it also permits a wider range of initial geminates. The avoidance of initial [pː] in all the languages surveyed thus far is perhaps less surprising in Ikema as /p/ itself appears to be undergoing a neutralization towards /h/ "in many Ikema (regional and generational) varieties" (Hussain & Shinohara, 2020). It is worth mentioning that among the Ryukyuan languages (and indeed the Japonic languages), initial [pː] appears to only occur in the Yaeyama language Tedumuni (Shinohara & Fujimoto 2011), in which it appears along with [sː], [kː], [tː], and [mː].

Returning to Ikema, I will now compare the distinctive and coerced sets of Ikema geminates. Once again, one of the primary challenges in this regard is that no Ikema words may be monomoraic, while words featuring geminate onsets tend to be less complex. As such, a true minimal pair such as [ffa] – [fa] cannot be found in this language. Examples (i, ii, iii, vi) below are taken from Takubo (2021) and Kindred (2019), while (iv, v, vii, viii) are taken from Shinohara & Fujimoto (2021). Romanizations have been added where none were present in the source.

(89) Ikema initial geminates and singletons, monosyllabics

i.	[n]	*nna*	[nːa] 'turban shell'	*naa*	[naː] 'name'
ii.	[m]	*mma*	[mːa]'mother'	*maa*	[maː] 'trace'
iii.	[f]	*ffa*	[fːa] 'child'	*fau*	[fau] 'eat'
iv.	[s]	*ssa*	[sːa] 'grass'	*suu*	[suː] 'tide'
v.	[t]	*tta*	[tːa] 'tongue'	*taa*	[taː] 'rice field'
vi.	[ts]	*ccya*	[tsːa] 'come'	*cjuu*	[tɕuː] 'dew'
vii.	[z]	*zza*	[zːa] 'father'	*zii*	[zɻː] 'soil'
viii.	[v]	*vva*	[vːa] 'you'	-	

Based on the table in (95), it is clear that Ikema features a near-minimal contrast between geminates and singletons within the domain of word minimality (i.e., in the unaffixed word). Like in Irabu, /v/ also has no singleton counterpart in word-initial position. However, this segment is permitted to occur as [vː] initially, and does not surface as [ʊv]. Recalling that a geminate cannot co-occur with a complex nucleus within a syllable, the pairs in (89) provide some evidence of a contrast. This is reinforced by morphemes with more syllables. Disyllabic words are used below because the vast

majority of Ikema free morphemes are either monosyllabic or disyllabic. Longer words usually result from affixation or other morphological processes, thus taking them out of the domain of the bimoraic minimality constraint. The examples here are taken from Kibe (2012/2019), Hayashi (2009, 2010), Takubo (2021), and Kindred (2019):

(90) Ikema initial geminates and singletons, disyllabic monomorphemes

i.	[n]	*nnagu* [nːagu] ‘sand’	*nasi* [nasɨ] ‘pear’
ii.	[m]	*mmara* [mːara] ‘similar’	*mata* [mata] ‘also’
iii.	[f]	*ffaci* [fːatsɻ] ‘hoe’	*fusa* [fu̯sa] ‘grass’
iv.	[s]	*ssabi* [sːabi] ‘remove dirt’	*saki* [saki] ‘liquor/alcohol’
v.	[t]	-	*taja* [taja] ‘strength’
vi.	[ts]	*ccutsi* [ttɕuːtsɻ] ‘sago palm’	*cigusi* [tsɨgusɨ] ‘knee’
vii.	[z]	*zzara* [zːaɾa] ‘sickle; scythe’	*zɻki* [zɻkɨ] ‘kin (unit)’
viii.	[v]	*vvadu* [vːaduː] ‘2PL’	-

The picture becomes a little less clear in (90), as multiple gaps emerge. Some of these may be accidental, owing either to coincidence or to a general lack of data. Problematically, productive initial gemination appears to be less common in Ikema. In the following section, I will account for the distinctive and coerced moraic onsets in Ikema.

3.2.5 Coerced and distinctive onset weight in Ikema

While no disyllabic (or longer) free morphemes were found with initial [tː] in the sources examined for (90), this geminate is otherwise quite common in word-initial position in monosyllables. The affricate /ts/ has a wider distribution, appearing either as long [tsː] in monosyllables or as short [ts] in longer words. The only clear example of a polysyllabic word with initial [tsː] in Ikema appears to be *ccutsi*, which is recorded as [ttɕuːtsɻ] in Kibe (2012/2019: 195). While this may indicate that /t/ receives coerced weight in monosyllabic words, it is unclear whether the same holds for /ts/. Takubo (2021: 76) suggests that the underlying root forms of verbs such as [sːan] ‘know=NEG’ and [tsːan] ‘wear=NEG’ are *ss*-, and *cc*-. While no underlying syllabification need be assumed for these words, the implication is clearly that /s/ and /ts/ is underlyingly moraic in these cases.

Faithfulness to an underlying mora, identified as involving MaxLink-μ for Irabu and Tarama above, implies that underlyingly long segments should occur also where they are not compelled to do so by bimoraic minimality. There is some support for this in the case of [tː], and somewhat weaker evidence for [tsː] in this regard. We may therefore assume that some degree of distinctive moraicity must be present. This brings us to the voiceless fricatives, /f s/. In terms of theoretical approach, neither Hayashi (2010) nor Takubo (2021) analyze Ikema roots as being underlyingly monomoraic. It is therefore difficult to determine whether onset lengthening occurs in Ikema. Although there are no true minimal pairs in (90) above, the near-minimal pairs that are present may tentatively suggest that

each of the fricatives may be lexically specified as moraic in onset position. As noted in Takubo (2021), Hayashi (2013) assumes a set of verbs to feature underlying singleton onsets that are lengthened in the output when a stem-final vowel (or vowel sequence) is deleted when the negation suffix *-an* is added. However, this analysis is contested in Takubo (2021: 75-76), who instead posits underlyingly moraic consonants:

Lexical item	Hayashi (2013)		Takubo (2021)	
'know'	*sɨɨ* + *an*	*ssan*	*ss-* + *an*	*ssan*
'wear'	*cɨɨ* + *an*	*ccan*	*cc-* + *an*	*ccan*

Going by the analysis in Takubo (2021), there appears to be little definite evidence for coerced onset moraicity in Ikema in general. The strongest evidence thus far for any form of coerced moraicity appears to be the absence of initial [tː] in underived polysyllabic lexical items, as well as the relative absence of initial [tsː] in longer forms, though the latter may simply be due to a lack of data. It will tentatively be assumed here that [t] indeed receives coerced moraicity in initial position, though further research is needed to determine the underlying morphology that may trigger such coercion.

3.2.6 Distinctive moraicity of initial geminate nasals

Following the discussion of the Irabu initial geminates, it is expected here that in principle, all Ikema nasals should permit distinctive moraicity in onset position. In Irabu, in which it is also licit for the nasals to form a monosyllabic word consisting of only a long syllabic consonant (hereafter NN), this was argued to be evidence for the distinctiveness of the nasal moraic onsets. There are some issues that must be resolved if we are to adopt the same hypothesis here, however. There is a distinct difference between the nasals present in the moraic onsets and the nasals present in NN. As has been argued here, the syllabic nasals and nasal codas are likely high-sonority nasals as described in Krämer & Zec (2020).

The assumption then is that the high-sonority nasal /ɴ/ assimilates in place to a following consonant but is in free variation between [m ~ n] when occurring as the only part of an NN monosyllable. The only underlyingly syllabic consonant in Ikema can therefore be stated as /ɴ/, though place assimilation may result in phonetic [n] or [m] as a syllabic consonant (where preceding an obstruent or geminate onset) or as a coda (between two morphemes). Where the nasal is not followed by any onset that would permit assimilation, the nasal is in free variation [n ~ m]. Considering that the contrast between [n] and [m] is otherwise retained in other Miyako varieties, this neutralization implies that the nasals preceding the obstruent in words such as [nta] or [mbu] are analyzed in the underlying representation as high-sonority /ɴ/, permitting them to be syllabic in this environment. Interestingly, the high-sonority nasal does not itself occur as a geminate onset, e.g., *[ɴːa], but the voiceless nasal [ɴ̥] does, as the following examples show (Takubo 2021: 68):

(91) a. /ɴː/ [nː] ‘sweet potato’
b. /ɴ̥ɴ/ [n̥n] ‘step on, scoop (water)’
c. /ɴ̥ɴa/ [n̥na] ‘rope’
d. /ɴ̥mu/ [m̥mu] ‘cloud’

As /ɴ̥/ is suggested in (91) to only occur before another nasal, it is interpreted in Takubo (2021) as a devoicing instruction and is orthographically represented there as *h*, e.g., *hna* /ɴ̥ɴa/. Due to its limited distribution, it is indeed tempting to analyze this segment as a suprasegmental feature or allophone of /ɴ/. However, [n̥] and [m̥] appear to contrast with [n] and [m], respectively, in geminates. There also appears to be a contrast between [n̥] and [m̥]:

(92) a. /n̥na/ [n̥na] ‘rope’ /nːa/ [nːa] ‘turban shell’
b. /ɴ̥mu/ [m̥mu] ‘cloud’ /mːa/ [mːa] ‘mother’
c. /ɴ̥mu/ [m̥mu] ‘cloud’ /ɴ̥nu/ [n̥nu] ‘yesterday’

As far as the representation of geminate initial onsets is concerned, it appears that the Ikema nasals must receive distinctive weight. As in Irabu, illegal onset clusters are repaired through syllabification of the nasal, which in Ikema entails assimilation to the following consonant (and therefore neutralization of place contrasts). Similarly, where a nasal is located after a vowel in the input string, it is syllabified as a coda, also leading to neutralization. Neutralization notably does not occur where a nasal is followed by a vowel and is parsed in onset position, whether it forms a singleton or geminate onset. This suggests that the difference between Ikema geminate nasal onsets contra nasal syllabics and codas is specifically one of distinctiveness (moraic onset) versus coercion (moraic coda or nucleus). As explored previously, it is suggested here that the voiceless nasal is better described as a preaspirated nasal geminate. As a geminate nasal is primarily characterized by continuous airflow through the nasal passage accompanied by voicing, there should be no particular expectation that the voiceless component of this nasal geminate should be voiceless throughout. Instead, it should only require this quality early in its articulation.

A consequence of this analysis is that the voiceless component of the “half-voiceless” geminates is not considered a separate segment from the voiced component. Instead, the geminate is considered one long segment with an aspirated (and therefore voiceless) period at the beginning of its articulation, accounting for the brief nature of the voicelessness described in Shinohara & Fujimoto (2018). As such, the contrast between the nasals in [n̥nu] ‘horn, yesterday’ and [m̥mu] ‘cloud’ (Takubo 2021: 69) is taken to be phonemic, indicating that Ikema features a phonemic contrast between /n̥/ and /m̥/ in addition to /n/ and /m/. The word /ɴ̥ɴ/ [n̥n] ‘scoop (water)’ (Takubo 2021: 68) seems problematic in this regard, as it has thus far been assumed that the syllabic nasal is limited to a high-sonority nasal that is either phonemic or an allophone of the nasal series. However, it may be appropriate here to suggest that the preaspiration found in /n̥/ and /m̥/ is distinctive for the high-

sonority nasal as well, accounting for the contrast between [n̥n] ‘scoop (water)’ and [nː] ‘sweet potato’.

Another possibility would be that the partially voiceless nasal geminates are in fact specified for breathy voice, the quality in which the vocal folds vibrate “without appreciable contact” and with a “higher rate of airflow than in modal voice” (Ladefoged & Maddieson 1996: 48). A relationship between nasality and breathy voice is recognized, as it has been claimed that breathiness may arise in nasality due to listener misperception or phonetic enhancement, and that breathiness and nasality may co-occur as part of a diachronic sound change (Garellek et al. 2016). Breathiness as a phonemic contrast would explain why partial devoicing appears to occur while /N̥ː-/ remains moraic in onset position. The contrast between [m̥], [n̥], and [m], [n] could then be understood as a contrast between specification for breathy voice versus no such specification. Further phonetic and acoustic study of these onsets is needed to determine airflow, vocal fold contact and arytenoid cartilage distance.

3.2.7 Distinctive moraicity of initial geminate /z/ and /v/

There are two major differences between Ikema and Irabu with regard to the voiced fricatives. First, initial geminate /v/ and /z/ surface [vː] and [zː] with no indication of these segments occurring as heterorganic sequences. The other major distinction is that initial /zː/ is in free variation between [zː ~ dzː] in Ikema. Like in Irabu, however, /v/ never occurs as a singleton onset. Instead, this segment only appears as a geminate onset or intervocalic geminate. This appears to support the hypothesis that this /v/ is underlyingly specified as moraic among the Miyako languages. As mentioned previously, non-moraic /v/ appears to be quite rare across the Miyako dialect spectrum. Pellard & Hayashi (2012/2019: 43) show that this segment or /ʋ/ occur as moraic in most Miyako dialects, but length contrasts involving /v/ are generally rare and usually result from a change from Proto-Miyako *w > v in words like **waa* > *vaa* ‘pig’ (Pellard & Hayashi 2012/2019: 50).

Because /v/ may only surface with moraic weight, it is assumed here to be underlyingly moraic in Ikema. This is supported by the complete absence of singleton /v/ onsets in this dialect, as well as the general tendency to only feature long /v/ in Miyako. Referring back to the lists in (89) and (90), the distribution of /z/ is distinctly different from that of Irabu, as initial singleton and geminate /z/ both appear to occur regularly. Other near-minimal pairs suggest that distinctive moraicity may occur in the case of /z/ as well, such as *zyau* [ʑau] ‘good’, and *zza* [zːa] ‘father’ (Hayashi 2010: 175, 182). Words such as *zzaku* ‘oar’ (Igarashi et al. 2016: 54) also do not require coercion to form an initial geminate.

3.2.8 Summary

To summarize this section, the nasals and voiced fricatives have been argued to permit distinctive moraicity in onset position in Ikema. Unlike the voiced fricatives, the nasals may be coerced into moraicity in coda position, or as syllabic consonants to syllabify initial nasal-obstruent clusters. The

voiced fricatives, however, never occur in coda or nucleus position. The obstruents may only be moraic in the onset or as part of an intervocalic geminate. It has also been suggested, tentatively, that /t/ and /ts/ may receive coerced weight in onset position. In the following section, I will provide an OT analysis of the Ikema initial geminate patterns, contrasting this with Irabu where relevant.

3.2.9 Ikema initial geminates in OT

3.2.9.1 Overview of Ikema initial geminates

I will begin this section by comparing the initial geminate segments permitted in the Miyako languages discussed thus far:

(93) Miyako initial geminates

Manner	Irabu			Tarama			Ikema		
Place	Labial	Coronal	Dorsal	Labial	Coronal	Dorsal	Labial	Coronal	Dorsal
Stop		t: ts:						t: ts:	(k:)
Voiceless Fric.	f:	s:		f:	s:		f:	s:	
Voiced Fric.							v:	z:	
Nasal	m:	n:		m:	n:		m: m̥:	n: n̥:	
Approximant		l̥:							

Ikema appears to feature largely the same initial geminate segments as Irabu, lacking only the long approximant [l̥:] and featuring the additional voiceless nasals [m̥:] and [n̥:]. The long coronal stop [t:] occurs more frequently in Ikema, and a marginal occurrence of [k:] is also attested. [k:] is of particular interest here, as it is the only dorsal segment that appears to be permitted to form an initial geminate. However, as stated previously, initial geminate [k:] is attested only in a single word and may possibly be in variation with [kuku]. More data is needed to determine the exact identity of Ikema [k:], and this segment will therefore be ignored for the following discussion.

Unlike the prior discussion for Irabu, [v:] and [z:] are not analyzed here as sequences of a syllabic segment and a non-syllabic segment. They evidently may not occur as syllabic in Ikema, nor do any allophones of these segments surface as syllabic. Additionally, the geminate onsets [v:] and [z:] are structurally consistent with geminates, unlike the sonorant-obstruent sequences found in Irabu. Only the nasals form syllables of the type [C.CV], as the nasals are the only syllabic consonants in Ikema. Consequently, it will be more parsimonious to assume that Ikema simply permits more types of segments to form its initial geminates, as it allows /v/ and /z/ to behave as the other obstruents do in onset position. This is further supported by the fact that a wide range of voiced stops are permitted medially. Takubo (2021: 68) states that all Ikema consonants except /j, w, h, n̥/ may be geminated

intervocalically, which implicitly includes the entire stop series with the sole exception of the voiceless/preaspirated nasal geminate.

By contrast, Irabu permits only /pː, tː, kː/ in terms of intervocalic geminate stops (Shimoji 2008: 56), though we may expand this to /pː, tː, tsː, kː/ per the discussion in a previous chapter. This suggests that the prohibition against [pː, bː, dː] in word-initial position is not necessarily a constraint against geminate voiced obstruents, but rather at least one constraint against voiced stops, as well as a constraint specifically restricting labial stops. Adopting a highly ranked markedness constraint against voiced initial stops should be unproblematic here as in Irabu. While any constraint against initial geminate voiced stops will implicitly restrict the nasals /m̥ː, mː, n̥ː, nː/, these are taken to be distinctively moraic and thus not subject to sonority-driven markedness. A further issue here is that /z/ appears to show some degree of variability between [z] and [dz]. As discussed in the section on Irabu, it is assumed here that, following Kehrein (2002), affricates form a class with stops and not with fricatives. Affricates such as [ts], [dz] are understood here to be [strident] and [-continuant]. Essentially, this suggests that if initial [dzː] surfaces, then this segment is the sole exception to the Miyako ban on initial geminate voiced stops. The affricate variant appears to have been recorded in word-initial position in Ikema *ddʑitai* ‘to.get-PAST’ (Karimata 2012/2019: 100), also recorded as [ddʑittai] in Kibe (2012/2019: 315). The tendency of /z/ to affricate in Ikema is further corroborated in MRI research by Shinohara & Fujimoto (2021). Interestingly, the latter authors also note that their speakers appeared to form a labiodental occlusion when producing initial [vː] (ibid.: 37), which does not appear to have been described in the literature on Ikema previously.

Though this closure was evidently not viewed as being phonetically prominent enough to result in a transcription of [vː] as [bv], this finding implies some support for the preference for lower sonority segments in geminate onsets, as the long, phonetically affricate segments [dzː] and [bvː] are expected to be less sonorant than their long voiced fricative counterparts [zː] and [vː]. There is another key distinction to be made between [zː] and [vː], in that [vː] is obligatorily moraic in onset position, while [zː] shows a contrast with short [z] in disyllabic words. Another distinction to consider with regard to [vː] and [zː] is featural: like /ts/ and unlike /v/ (or [bv]), [dz] is [strident]. As [ts] does not alternate with [t] in any context, it could be posited that an identity constraint enforces the strident identity of the output segment corresponding to the strident segment in the input. Because [dz] does not occur phonemically in Ikema, instead being in free variation as an allophone of /z/, it is not difficult to conceive of why /s/ does not show the same behavior. /ts/ contrasts phonemically with /s/ and there does not appear to be any neutralization in the direction of either segment in their allophonic distributions.

The relevant constraint ranking for [zː ~ dzː] thus appears to be to posit a constraint MaxLink-Mora (strident), ensuring that underlying /zː/ does not become non-moraic in the output. This

constraint may be highly ranked, outranking *μ/ONS[voice], as the only Ikema strident segments are /s/, /ts/, /z/, and [dz], all of which occur as word-initial geminates and have been argued to occur with distinctive moraicity. It is assumed here that this constraint ranking is not what triggers free variation between [z ~ dz], as this variation applies for all instances of /z/, whether it is a singleton onset, an intervocalic geminate, or an initial geminate. For reasons of space, I will therefore not be discussing what motivates this variation further.

To summarize the above discussion, the below segments have been identified as occurring with distinctive and coerced weight in initial position. As noted previously, the identification of a segment as coerced in the table below does not mean it cannot bear distinctive weight, but rather that the distinctively moraic segments cannot be coerced:

(94) Ikema geminate onsets, distinctive and coerced weight

Manner	Distinctive		Coerced	
	labial	alveolar	labial	alveolar
Stop				t: ts:
Fricative	v:	z:	f:	s
Nasal	m̥m m:	n̥n n:		

Recalling that the proposed universal markedness for coerced morae in onset position goes from the most sonorant/most marked to the least sonorant/least marked, we can assume that *μ/ONS[stop] is ranked low. The only highly marked onset geminate stops will here be the voiced plosives /b/, /d/, /g/, and the voiceless labial /p/. All of these occur in singleton onsets and may even be geminate intervocalically but are prohibited from forming a geminate onset.

3.2.9.2 Voiced initial geminate [z: ~ dz:]

As discussed, the only voiced plosive segment that may variably be geminate in the onset is [dz], as an allophone of [z]. Owing to the largely strict and unviolated constraint against moraic voiced stops in the onset in the other Miyakoan and Ryukyuan languages, it must be asked why this segment is permitted to surface as a voiced stop in this position. Simply put, it is unclear why there should be an exception for [dz:]. Assuming that some constraint enforces the strident identity of the onset in the output with regard to its corresponding segment in the input, and that voicing be preserved at the same time, the optimal output may in fact be the affricate. In terms of sonority, this is actually the outcome predicted by *μ/ONS[fricative] >> *μ/ONS[stop], as [dz] is [-continuant] and thus optimal by comparison:

(95) WdMin >> *μ/ONS[fricative] >> *μ/ONS[stop]

/za/	WdMin	*μ/ONS[fricative]	*μ/ONS[stop]
a. za	*!		
b. z:a		*!	
c. ☞ dz:a		*!	

For this to be the case, we must also assume that (a), there are two constraints, one that may be violated by a change from fricative to stop, and one that is violated by any change to the featural specification [strident], respectively, or (b) that a single constraint is violated by both changes, causing the most faithful output candidate to be optimal where this constraint dominates. A good example of such a constraint would be Ident-Manner (Lombardi 2003: 237), "a constraint violated by any change of Stop, Cont or Strident". Note that an approach focusing on CV linkage constraints (Itô & Mester 1995) cannot apply here, as [tu] is licit in Ikema (unlike Japanese). Lombardi suggests that Ident-Manner may be exploded into Ident Stop and Ident Cont, as well as Ident Strident, in languages where this specification is necessary, such as in languages that feature many fricative/stop alternations. If we assume a ranking of Ident Stop >> Ident Strident >> Ident Cont, all else being equal, three conclusions follow: All stops in the input must remain stops in the output, strident segments must remain strident except where this violates Ident Stop, and fricatives may become stops to avoid violating another constraint provided they do not violate Ident Strident. Ident Strident, or Ident-IO (strident) as in Kawai (2004), conversely also prohibits the assignment of a [strident] feature to segments that are not underlyingly specified as such (Lombardi 2003: 236). It may not be necessary to explode Ident-IO(manner), however. Assuming that these onsets are moraic, we can derive the following constraint ranking. Once again, WdMin is undominated and is not included here:

(96) *μ/ONS[fricative] >> Ident(Manner) >> *μ/ONS[stop]

za	*μ/ONS[fricative]	Ident(Manner)	*μ/ONS[stop]
a. z:a	*!		
b. ☞ dz:a		*	
c. d:a		**!	*

In the above tableau, [d:a] is disqualified because it contains two violations, as opposed to the one violation featured by [dz:a]. [z:a] is ruled out here due to markedness. Of course, this is not a desirable outcome in and of itself. Generating free variation such as Ikema [z~dz] in OT requires the involvement of probabilistic extensions to this model, such as partially ranked constraints (Anttila 2007) or stochastic constraint rankings (Boersma & Hayes 2001). The premise of partially ranked

constraints is that, rather than all constraints featuring a fixed ranking, certain constraints or sets of constraints may be probabilistically selected in EVAL. If two constraints are partially ranked, the probability of either constraint being selected as dominant is 50%. With three constraints, each individual order of constraints is approximately 16.6% likely to be selected. No data on the exact probability of the free variation of [z ~ dz] in Ikema has been collected, so further study is needed to determine whether either option is more common. Assuming a 50% distribution of [z ~ dz], however, free variation between the constraints *μ/ONS[fricative] and Ident Manner may be given as follows:

(97) {*μ/ONS[fricative], Ident Manner}

za	*μ/ONS[fricative]	Ident Manner
a. z:a	*!	
☞ b. dz:a		*
c. d:a		**!
za	Ident Manner	*μ/ONS[fricative]
☞ a. z:a		*
b. dz:a	*!	
c. d:a	**!	

While this resolves the variability of [z: ~ dz:], it has the unfortunate effect of implying that the same variation should apply to [s]. Furthermore, as noted previously, /z/ [z~dz] is in variation in all environments, and not simply when it occurs as a geminate onset. If [stop] or [continuant] were left unspecified for /z/, then Ident Manner would have no application, as /z/ would incur violations regardless of whether it occurred as a stop or as a continuant in the output. It therefore seems equally likely that this variation can instead be modelled through the markedness of the individual segments [z] and [dz], ranked as {*z, *dz}. If this is the case, the partially ordered set of {*z, *dz} must necessarily outrank all constraints pertaining to the moraic onsets [z:], [dz:]. Provided that the winner constraint of {*z, *dz} dominates the constraint against voiced obstruents in the geminate onset, then the candidate that violates the lower-ranked constraint output will be selected in spite of any constraints against this segment in the geminate onset.

It should in other words be clear here that we are assuming the same markedness hierarchy as was suggested for Irabu, with only a small amendment. First, it is no longer necessary to explore the possibility of any intermediate step in terms of sonority between the Ikema stops and fricatives. The affricates pattern identically to the stops, which unlike the Irabu stops, appear to be less marked than the fricatives in a geminate onset. This produces the following hierarchy:

(98) *μ/ONS[sonorant] >> *μ/ONS[fricative] >> *μ/ONS[stop]

With this constraint ranking, we see a pattern similar to that of Irabu emerge. Ikema initial [tː] and [tsː] behave similarly to Irabu initial [tsː]. For the distinctively moraic nasals, we again find application for *μ/ONS[-consonantal] and MaxLink-Mora [sonorant]. Unlike Irabu, Ikema has no moraic liquids in the onset, and as such, it is not necessary or possible to formulate a constraint ranking that preserves these. Instead, it will be more parsimonious to specify MaxLink-Mora as selecting the distinctive feature [nasal]. Because there is no expectation for distinctive weight to comply with sonority, the MaxLink-Mora[nasal] constraint need not fall anywhere within a universal hierarchy. Another advantage to specifying MaxLink-Mora[nasal] is that there is no indication of non-moraic coda nasals in Ikema. As all codas are nasals, and all codas are moraic, mora linking remains consistent. This allows for a simpler constraint ranking for distinctive moraicity:

(99) MaxLink-Mora[nasal] >> *μ/ONS[sonorant]

/n$^{\mu}$a$^{\mu}$/	MaxLink-Mora [nasal]	* μ/ONS [nasal]
☞ n$^{\mu}$a$^{\mu}$		*
na$^{\mu\mu}$	*!	

In (99) above, the violation is incurred by the word featuring the bimoraic vowel, as it entails the deletion of the link between this mora and the nasal. As this constraint outranks *μ/ONS[nasal], moraic nasals in onset or any other position must remain moraic. However, nasals will not be coerced into moraicity in onset position, as *μ/ONS[nasal] still outranks P-Dep-μ, which inhibits vowel lengthening as a strategy to repair illegal monomoraic words. More crucially, *μ/ONS[nasal] outranks *μ/NUC[nasal] as well. As such, it is expected that nasals are more likely to be coerced into serving as nuclei than they are to be coerced into forming moraic onsets. A consequence of MaxLink-Mora[nasal] >> *μ/ONS[nasal] >> *μ/NUC[nasal] is that all nasal nuclei will be preserved, and that in all cases of repair where a moraic onset is not underlying, the preference will be for the nasal to be syllabic over forming a moraic onset.

3.2.9.3 Voiced initial geminate [vː]

In the section on Irabu, it was suggested that /v/ is simply specified as obligatorily moraic in that variety. That is, the segment itself is subject to a constraint such as v/μ, causing the segment to surface as a nucleus (the allophone [ʊ]) in Irabu. As mentioned, no allophone of /v/ is capable of forming a consonantal nucleus in Ikema, as is the case for all the other obstruents in the language. All specifications of *μ/NUC[obstruent] must therefore outrank all specifications of *μ/ONS[obstruent].

The voiced obstruents are thus also banned from forming a nucleus, while *μ/ONS[obstruent] is less relevant than the constraint *μ/ONS[voice].

Assuming, then, that the underlying form of [vː] is underlyingly moraic /vᵘ/, the matter of initial /vː/ is one of faithfulness vs. markedness. In order to surface as a moraic onset, we may posit that MaxLink-Mora(continuant) outranks the high-ranked *μ/ONS[voice]:

(100) MaxLink-Mora(continuant) >> *μ/ONS[voice]

/vːa/	MaxLink-Mora(continuant)	*μ/ONS[voice]
vːa		*
vaː	*!	

This ranking is necessary to account for the underlying moraicity of Ikema /vː/ and /zː/. Although MaxLink-Mora(continuant) is satisfied by /v/ surfacing as a syllabic consonant, it can be assumed that *μ/NUC[obstruent] is undominated due to the absence of obstruent syllabic consonants.

To summarize this section, it has been argued that the ranges of segments that are permitted to form initial geminates and syllabic consonants in Irabu and Ikema onsets can be accounted for by the ranking of markedness and faithfulness constraints. The patterns in Irabu and Ikema are generally consistent with the predictions of Topintzi (2006), as all initial geminates are moraic and distinguish between distinctive and coerced moraic onsets. Coerced moraic onsets pattern with sonority, as voiced segments are banned from moraicity in onset position. However, it has also been found that the markedness ranking of moraic initial fricatives versus moraic initial stops is not fixed between the Miyako languages. This implies that sonority is not the only relevant condition determining the markedness of specific segments and types of segments in moraic onsets. The difference between fricatives and stops in this position may therefore be motivated by other factors, such as prominence or saliency. It has also been posited that constraints involving place of articulation may be relevant, but only for distinctively moraic onsets. The ranking of these constraints may be arbitrary, or may relate to issues in perception and production of geminate segments. All the Miyako languages appear to avoid initial geminate velars, as well as initial geminate labial stops. These constraints should not be understood as sonority-driven, but rather as a markedness relationship within Miyako. In the final section of this chapter, comparison will be made between Irabu, Ikema and Ōgami. The latter is the only Miyako language that has been argued to feature no phonemically voiced stops (Pellard 2009, 2010).

3.3 Ōgami

3.3.1 Overview of Ōgami phonology

The last variety that will be examined in this thesis is Ōgami. As the phoneme inventory and phonotactic behavior of Ogami deviates considerably from the other varieties of Miyako discussed

here, I will begin this section with a brief summary of this language's consonantal inventory as given in Pellard (2010: 116):

(101) Ōgami consonant inventory

	Labial	Labiodental	Dental/alveolar	Velar
Plosives	p		t	k
Fricatives		f	s	
Nasals	m		n	
Taps/flaps			ɾ	
Approximants		ʋ		

The Ōgami inventory is notably sparse among the Ryukyuans, featuring only nine consonants. The stop series /p t k/ features no phonemic voicing contrast but stops "can be optionally voiced between vowels". All instances of distinctively voiced stops are either Japanese loanwords or loans from other Miyako varieties (Pellard 2010: 116-117). Voicing in such cases is unstable, however, attesting to the lack of voice as a contrastive feature. It should also be noted here that there is no phonemic glottal fricative /h/ and no affricate /ts/, though [ts] and [tɕ] do occur in borrowings. /ʋ/ is also not consistently moraic as it is in Irabu, or as /v/ is in Ikema, and there is a contrast between /ʋaː/ 'pig' and /ʋːa/ 'you' (Pellard 2009: 58). While /ʋ/ generally surfaces as the labiodental approximant [ʋ], it "can be a fricative when [v] when geminated/long" (Pellard 2010: 117). In terms of phonotactics, the pattern is even more peculiar. In addition to featuring length distinctions for the consonants, all the continuants except /ɾ/ may be syllabic. As is noted in Pellard (2010: 119), this is particularly unexpected, as the existence of syllabic obstruents should entail the existence of syllabic liquids (Zec 2007). Pellard argues that the Miyako apical vowel *ɿ and the rounded back vowel *u completely assimilated to voiceless fricatives in Ōgami, resulting in a broad range of voiceless fricative syllable nuclei. These syllables may have the following structure (Pellard 2010: 120):

(102) (O)Nu(Nu)(Co)

In the above template, O represents the onset, Nu represents a nucleus segment and Co represents a coda. Onsets and codas are optional, and vocalic and consonantal nuclei may be simple, long or complex (diphthongal). Like the other Miyako languages, complex onsets and codas are not permitted. Also like the other Miyakoans, there is a bimoraic minimal word constraint, with deletion repaired through compensatory lengthening. The syllabic consonants behave more similarly to vowels in Ōgami than they do in Irabu and Ikema, as certain syllabics may take onsets:

(103) Ōgami syllabic consonants (Pellard 2010: 119-120):

e. Syllabic consonant-only words:

ix. /mm/ ‘yam’
/ʊʊ/ ‘sell’
/ss/ ‘dust’, ‘rub’
/ff/ ‘comb’, ‘bite’. ‘fall (rain)’

f. Syllabic /s/ with /p-/ or /k-/ onset:
x. /kss/ ‘breast’, ‘fish-hook’, ‘to fish’, ‘come’
xi. /ksks/ ‘month’, ‘listen’
xii. /fks/ ‘mouth’, ‘build’
xiii. /psks/ ‘pull’

g. Syllabic /f/ with /k-/ onset
xiv. /kff/ ‘make’

As is clear from the above examples, only the voiceless fricatives /f/ and /s/ may take onsets when syllabic. Voiceless syllables with onsets comply with sonority sequencing, as the nucleus (here the voiceless fricative) must be more sonorous than the onset (the voiceless stop). Tautosyllabic stop-fricative sequences do not occur, as the only attested examples of stop-fricative sequences involve a long fricative segment and are analyzed as heterosyllabic (Pellard 2009: 81). No obstruent-sonorant sequences of any kind occur. Like the other Miyako varieties, words consisting only of one long syllabic consonant are still subject to bimoraic minimality. Sequences such as /mm/ are therefore analyzed as bimoraic as well, in the manner of a long vowel or diphthongal sequence of vowels. The labial nasal [m] and labiodental approximant [ʋ] pattern with /f/ and /s/ in forming long homorganic monosyllables, as well as serving as syllabic segments much as was suggested for Ikema in the previous section (Pellard 2010: 119, 2009: 59):

(104) a. /nta/ [n̩ta] ‘where?’
b. /pstu/ [ps̩tu] ‘person’
c. /ftai/ [f̩tai] ‘forehead’
d. /mna/ ‘seashell’
e. /mta/ ‘soil’
f. /sta/ ‘below’
g. /fta/ ‘lid’

In the examples in (101), Ōgami /nta/, /mta/, /sta/, /fta/ appear to behave identically to Ikema /nta/ ‘soil’, with a syllabic nasal followed by a simple CV syllable. [ftai] is phonetically similar to the surface representation of Ikema /fusa/, which has been recorded with a devoiced /u/, e.g., [fu̥sa ~ fʷsa] (Kibe 2012/2019: 39). Ikema /ftai/ ‘forehead’ is in fact recorded as [ftai] in Kibe (2012/2019: 42), though this is likely a matter of vowel devoicing, as is seen in Irabu [fʊ̥tai], Bora [fʊ̥tai], Kuninaka [fu̥tai], and other Miyako varieties (ibid.). Pellard argues that while this may historically have applied

in Ōgami, there is no evidence that synchronic vowel devoicing occurs in the latter language, clearly making the fricatives in (101-b) and (101-c) the peaks of their respective syllables. Perplexingly, /ɯ/ and syllabic /s/ "seem to be in the process of merging to /ɯ/ after /k-/ in word-final position" (Pellard 2010: 123), as there are attested examples of both contrasts between word-final /-ks/ and /-kɯ/, there is often free variation between the two and neutralization of this contrast.

Pellard also notes a three-way length distinction for consonants, in which a consonant segment may be short, geminate, or extra-long, similar to what has been described as a suprasegmental feature for languages such as Estonian (Prince 1980) or Lule Saami (Fangel-Gustavson, et al. 2014). This three-way distinction occurs with the voiceless fricatives /f/ and /s/, as well the bilabial nasal as /m/. We may compare this with the Ikema example of the same. In the below examples, the underlying and surface transcriptions are copied directly from their respective authors:

(105) Ōgami three-way length distinction (Pellard 2010: 188)
a. /faa/ [faː] 'child'
b. /f.fa/ [fːa] 'grass'
c. /ff.fa/ [fː.a] 'comb=TOP'

Ikema three-way length distinction (Takubo 2021: 67-69)
d. /maa/ [maː] 'trace'
e. /mma/ [mːa] 'mother'
f. /mmma/ [mː.ma] 'the head of female priests'

The Ikema examples in (102-d,e,f) are also similar to Ōgami /mɑɑɯ/ [mɑːɯ] 'turn', /mmɑ/ [mmɑ] 'mother', /mmmɑ/ [mːmɑ] 'potato=TOPIC' (Pellard 2009: 59), with /mmɑ/ being cognate in the two languages. Takubo and Pellard both appear to agree on the syllabification of their respective onset lengths in (102), as both (102-c) and (102-f) are suggested to feature a boundary at the right edge of the long syllabic [fː] and [mː] in Ōgami and Ikema, respectively. Both authors therefore also argue that this three-way distinction is not phonemic, arising only from the adjacency of the two syllables in question. Following the descriptions given above for Ikema and Irabu, this is more or less what is expected with regard to syllable structure. Length distinctions in all the Miyako languages appear to be limited to a singleton/geminate or short/long distinction, the latter of which has been used to refer here to codas and onsets, and the latter of which has been used to refer to nuclei.

Some mention must be made here of theoretical assumptions, specifically with regard to geminates. Pellard (2010: 118) transcribes the underlying structure of [fːa] 'grass' as /f.fa/, suggesting an interpretation of initial geminates as consisting of two underlying segments spread across two syllables. Moraicity in Ōgami according to Pellard (2009: 69) is assigned to (nucleic) vowels, syllabic consonants, codas and "the first half of a geminate". This is essentially in keeping with a skeletal CV

or X-tier representation of geminates, but as has been argued here, an approach rooted in Moraic Theory appears to adequately explain these patterns in a satisfactory manner. Applying the analysis that the geminate onset is fundamentally just an onset segment linked to a mora node, we derive a slight modification of the Ōgami syllable template given in Pellard:

(106) Revised Ōgami syllable template
(O)(ː)Nu(Nu)(Co)

Because complex margins and trimoraic nuclei are still prohibited, the only difference is that the onset may be defined as long or short. This shows that despite the seemingly odd nature of Ōgami outputs, the overall pattern largely complies with that of the other Miyako languages. We once again see a pattern of less-to-more-sonorant consonants that may bear moraic weight in onset position, more sonorant coda consonants, and a mix of syllabic consonants that form a "presyllable" structure of the type described in Shimoji (2008) or fully consonantal words. The major distinction between Ōgami and Irabu, Ikema, and Tarama in this respect is the range of consonants that may serve as nuclei, which is both broader and stranger with regard to sonority, allowing for fully voiceless words. With this in mind, I will proceed now to describe the Ōgami initial geminate pattern in OT, paying particular attention to the relationship between syllabic consonants and onset geminates.

3.3.2 Ōgami in OT

The first thing to note in discussing the differences between Ōgami and the other Miyako languages is that it is expected that more specifications of *μ/ONS[C] will outrank more specifications of *μ/NUC[C], where C refers to any consonant segment or distinctive feature of consonantal segments. As may be clear from the description of Ōgami provided above, the only undominated specification of *μ/NUC[C] should be *μ/NUC[ɾ], or stated more generally, *μ/NUC[liquid]. While cross-linguistically odd, considering the hierarchies in both Morén (1999, 2003) and Zec (1988, 2007), *μ/NUC[liquid] must outrank *μ/NUC[obstruent]. Alternatively, it will be sufficient to simply give *MORA[liquid] (Morén 2003) as an undominated constraint, as the rhotic /ɾ/ also does not occur in coda position, and therefore also does not occur as moraic.

Another notable fact about Ōgami here is that (almost) none of the stops are permitted to be geminate in onset position. With the exception of [tː], which like in Irabu is present in only one lexical item, the hearsay marker /tːa/ (Pellard 2009: 58), the only initial geminate obstruents are the voiceless fricatives /f/ and /s/. We must therefore assume that *μ/ONS[stop] is not dominated by its faithfulness counterpart, and in fact dominates the constraint MaxLink-Mora[stop]. This reveals the pattern that only segments that are either [nasal] or [continuant] may surface as initial geminates, not unlike Tarama. The fact that this is the same set of segments that may occur as syllabic is significant, as this is the opposite of the case in Ikema and Tarama, where moraic onsets considerably outnumber syllabics.

It nonetheless does not appear to be useful to consider the initial geminates to be syllabic segments. While seemingly phonotactically licit, there is no motivation to output /ffa/ as [fː.a], unless the interpretation of geminates used is [f.fa]. Under a moraic geminate analysis, it would be particularly unexpected in cross-linguistic terms for a string such as /fːa/ to output a syllabic segment followed by hiatus and a vowel. The position of the following analysis will therefore be that these are long segments syllabified to onset position.

The question that immediately emerges here is why the stop series is restricted from occurring as initial geminates, assuming the implicational hierarchy given in Topintzi (2006). Given that the range of initial geminates runs from highly sonorous [mː], [nː], to less sonorous [vː], to least sonorant [fː], [sː], it would appear that either the hierarchy simply does not properly predict this pattern, or that we are again observing a distinction between coerced and distinctive moraicity. Recalling again that all of these segments may occur as singletons, as moraic onsets or as syllabic consonants, it seems likely that distinctive weight applies. In the following set of constraint rankings, I will explore the possibility of distinctive weight for [nː], [mː]. To account for the possibility of /tsːa/ outputting [tsː.a], the constraint ONSET (Prince & Smolensky 2004) will be used here:

(107) ONSET (Ons): Syllables must have onsets.

ONSET (ONS) is one of the earliest OT constraints, and is supported by a robust base of cross-linguistic data. Onset is violated by any syllable that does not feature an onset, e.g., syllables of a format such as V, VV, VVC, and so forth. This includes monosyllabics such as Ōgami [vː] ‘sell’. Because such words are completely licit, it stands to reason that ONS must be outranked by a constraint restricting epenthesis. The presence of ONS in Ōgami is further supported by the fact that an underlying /i/ in syllables where this /i/ precedes [ɑ] or [u] with no consonant onset, the vowel becomes an approximant onset [j] (Pellard 2010: 117). Onset epenthesis is not attested in Ōgami, so this suggests that a CV syllable [jɑ] is optimal compared to a diphthongal syllable such as [iɑ]. ONS must outrank the other two constraints in the below ranking to prevent [fː.a] from being the optimal output of /fːa/:

(108) ONS >> MaxLink-Mora[nasal] >> *μ/ONS[nasal]

/nːᵘaᵘ/ *nna*	ONS	MaxLink-Mora [nasal]	* μ/ONS [nasal]
☞ nːᵘaᵘ			*
naːᵘᵘ		*!	
nː.a	*!		

In the above tableau, the underlyingly moraic segment [nː] maintains its moraicity due to MaxLink-Mora, which rules out the bimoraic [naː]. [nː.aː] is similarly ruled out due to the constraint ensuring that an onset is present. Here, the violation of *μ/ONS[nasal] is irrelevant, as it is the lowest ranked constraint in the ranking. Because WdMin is held to be undominated here as well, the same ranking rules out the lengthening of an underlyingly non-moraic nasal onset:

(109) WdMin >> ONS >> MaxLink-Mora[nasal] >> *μ/ONS[nasal]

/naᵘ/	WdMin	ONS	MaxLink-Mora [nasal]	* μ/ONS [nasal]
naᵘ	*!			*
☞naːᵘᵘ				
n.a		*!		
nᵘaːᵘ				*!

In the above tableau, a completely faithful output [na] with a non-moraic onset and monomoraic nucleus is ruled out by the bimoraic minimality constraint. This again fits neatly in with the constraint rankings suggested for the other Miyako varieties, in which this constraint remains undominated.

Considering the description in Pellard (2009, 2010), it could be questioned why the analysis of Ōgami initial geminates should be that they are fully tautosyllabic [CːV], as opposed to heterosyllabic [C.CV]. However, if the assumption is that, following the moraic geminate analysis given in Davis & Topintzi (2017), the latter syllable structure should be both the least marked, yielding bimoraic [CːV]. However, it is clearly also the case that Ōgami, like Irabu and Ikema, features heterosyllabic sequences of a short [Ç] or long syllabic consonant [Çː] followed by a simple CV syllable. Like in the other Miyako varieties, it is also clear that the relationship between the geminate onsets and their respective syllables cannot be purely contextual, as contrast occurs here as well:

(110) Ōgami geminate-singleton contrast (Pellard 2009: 58-59):

Geminate	Singleton
/mmɑ/ [mmɑ] ‘mother’	/mɑɑɯ/ [mɑːɯ] ‘turn’
/nnɑmɑ/ ‘now’	/nɑmɑ/ ‘raw’
/ffɑ/ [ffɑ] ‘grass’	/fɑɑ/ [fɑː] ‘child’
/ssu/ [ssu] ‘white’	/suu/ [suː] ‘vegetable’
/ʊʊɑ/ ‘you’	/ʊɑɑ/ ‘pig’

It appears likely in (107) that the nasals are once again distinctively moraic in onset position,

suggesting MaxLink-Mora[nasal] >> *μ/ONS[nasal]. Because the nasals may occur as moraic in any syllable position, it is feasible to posit that a broader MaxLink-Mora constraint is undominated, except by *μ/ONS[liquid]. This appears to function well for the other syllabic Ōgami consonants as well, as the set of segments that may be geminate onsets, may be syllabic, or may be codas are all the same consonants. We are therefore unable to posit any kind of distinction between coerced and distinctive moraicity in Ōgami, as all segments may feature distinctive moraicity.

As with the other Miyako varieties, /p/ and /k/ are notably absent from initial geminate position. This implies one of two things. Either this is an accidental gap, and there is simply no moraic /pː/ or /kː/ input that would test whether underlying initial /pː/ or /kː/ could surface as moraic onsets, or MaxLink-Mora must be expanded to include place of articulation specifications as has been proposed for Irabu and Ikema. We would then find the same set of MaxLink-Mora constraints as in Irabu and Ikema to be relevant, with MaxLink-Mora(velar) and (labial) ranked below all specifications of *μ/ONS.

3.3.3 Summary

In this section, a brief comparison has been made between Ōgami, Irabu and Ikema, finding that Ōgami features a more faithful input-output relation in terms of moraicity, as only /p/ and /k/ may not surface as syllabic or as initial geminates. /t/ also shows extremely limited distribution in terms of initial geminates, occurring in only one lexical item. This, along with the near-minimal pairs in (110), indicates that all initial geminates in Ōgami may be distinctively moraic. The sets of consonants that may be syllabic, and those that may form initial geminates, overlap almost completely. Broadly speaking (and setting initial /tː/ aside for the moment), this implies that all moraic faithfulness is undominated for the continuants and the nasals.

CHAPTER 4: Conclusion

4.1 Summary

In this thesis, Miyako geminate onsets and onset phonotactics have been examined through the Miyako varieties of Irabu, Ikema and Ōgami, and an attempt has been made at accounting for both the patterning of the onset geminates and for the syllabic consonants that occur in each variety. It has been assessed whether Moraic Theory or CV/X-tier representations provide a more satisfactory analysis of the Miyako geminate patterns, and it has been concluded that the Moraic Theory representation given in Davis & Topintzi (2017) most accurately describes the Miyako pattern. Like the other Japonic languages, Miyako is strongly mora-driven, and it is apparent that the analysis of initial geminates as moraic onsets involves the fewest assumptions about the language.

As part of this analysis, some assumptions made about the languages examined have also been questioned, and alternative analyses have been proposed. In Irabu, it has been argued that due to evidence from other transcriptions, phonetic qualities, and general markedness, long /ts/ and /dz/ do not occur. This argument applies to both geminate /ts/ and /dz/ in the onset and long /ts/ and /dz/ as syllabic segments in an NN monosyllable. Furthermore, it has been argued that /ts/ and /dz/ cannot be represented as fricatives underlyingly, as they appear to pattern with stops. It has also been argued that due to their phonetic qualities, the apparent partially geminate voiced fricatives [ʊv] and [ʐz] are best understood as heterosyllabic sequences of a syllabic consonant and a simple onset. Finally, it has been proposed that the voiceless fricatives /f/ and /s/ receive coerced weight in Irabu, while /t/ and /ts/ do not. This has been interpreted as support for the claim in Topintzi (2006) that voicing is the most relevant distinction with regard to the markedness of coerced moraic onsets, as all voiced segments are more marked than all voiceless segments, but voiceless fricatives are still preferred over the less-sonorous voiceless stops.

In Ikema, it has been argued that the proposed nasal-obstruent onset clusters are, in fact, heterosyllabic clusters on the basis of sonority sequencing, the syllabicity of nasals in Ikema, and the lack of contrast between nasal segments in this position. As all place contrasts are neutralized between the Ikema nasals in nucleus and coda position, neutralization in the first part of a nasal-obstruent cluster indicates syllabicity. Little conclusive evidence has been found for coerced moraicity in Ikema, though there is some indication that the voiceless obstruents /t/ and /ts/ may receive coerced weight in certain morphemes. If so, it would constitute further support for a potential distinction between Ikema and Irabu being the permitted re-ranking of constraints against onset moraicity for stops and fricatives, respectively. Further research is needed to determine whether this indication can actually stand up to scrutiny, however.

Finally, Ōgami has been compared with the other two languages in this thesis as an example of a Miyako language with a considerably broader set of syllabic consonants. Due to the ability of all continuant segments in Ōgami to be syllabic, it is argued that essentially all continuant and nasal consonants may feature a lexical length contrast in word-initial position. It therefore does not appear to be the case that onset coercion occurs in Ōgami, as the nucleus may be lengthened or initial segments in monomoraic roots may be syllabified as nuclei.

4.2 Future areas of research

4.2.1 Constraint-based typology of Miyako varieties

One issue that has not been examined in detail in this thesis has been the issue of constraint-based typology. The fundamental goal of Optimality Theory is to map out the range of constraints that exist in the language faculty of the human brain, and to account for which rankings are universal, which rankings may be different from language to language, and what the consequences of these rankings

are. As such, there have been some proposals to clearly define how an OT typology should look and what predictions it should make. One such approach is Property Theory (DelBusso 2018), which seeks to derive properties from winner-loser pairs of constraint sets. Constraints that function together are defined as constraint classes, with one example of this being footing constraints favoring iambic or trochaic feet, i.e. {Ia, Tr} (Alber et al. 2016). Footing is thus a property with two mutually exclusive values. This may find application in the description of onset moraicity, as we clearly see a family of constraints governing which consonant segments may form initial geminates, which may be syllabic, and which may not be moraic under any context. The question this raises is whether the patterning of Miyako consonant moraicity can be defined as typological properties. This would potentially then allow for a greater understanding of not only the typological specificities of the Miyako varieties, but also the broader nature of dialectal variation.

4.2.2 Acoustic research on voiceless nasal and fricative vowel

As mentioned in chapter 3, much work still remains in terms of acoustic and phonetic research. At present, significant efforts have been made to examine Ikema and Ōgami phonetics, but further study is needed to clearly identify both the voiceless nasal in Ikema and the Miyako fricative vowel. Only by describing these phenomena in detail will we be able to fully account for the geminate patterns found in the Miyako languages. Another area of focus for future research may be to determine whether the Ikema voiceless nasal occurs elsewhere in Miyako or the Southern Ryukyuan language group in general, or if it is truly exclusive to Ikema. Additionally, further data is needed on the voiced fricative /v/ to determine whether it is feasible to maintain that this is an underlyingly moraic segment as has been argued in the present thesis. Further documentation of the Miyako varieties will provide considerably more data to both determine what patterns can be found throughout the Miyako language family, and to determine what theoretical implications these patterns hold.

4.2.3 Further description of Miyako varieties

Finally, it is clear that further description of the various Ryukyuan dialects will be necessary if we are to gain a complete understanding of what segments are or are not permitted to form initial geminates. This thesis has revolved around a limited number of languages out of necessity, and firm statements about phenomena in all Miyako dialects run the risk of being skewed by the availability of data on a select few dialects over the many that are still largely unrecorded. It is therefore essential that descriptive data continue to be collected from across the Ryukyus, so that we may examine in detail patterns such as the initial geminate /k/ proposed for Shimajiri Miyako, or the typological differences between the broader language groups of Miyako and Yaeyama.

4.3 Conclusion

The key findings of this thesis are as follows: Miyako onsets (and codas) strictly adhere to a constraint against complex margins. Where edge clusters have been proposed in the literature, these have been

found to be better analyzed as heterosyllabic sequences. The initial geminates found in the Miyako languages invariably satisfy bimoraic minimality where they occur. As such, it appears that the Miyako initial geminates are best described as monopositional, moraic segments. Where geminate identity is coerced by bimoraic minimality, this always involves voiceless segments. Voiced segments in illegal phonotactic contexts may surface as syllabic, and singleton voiced onsets in monomoraic roots trigger lengthening of the vowel nucleus of the word. This supports the analysis of voicing being marked for moraic onsets.

However, it also appears that the ranking between voiceless fricatives and stops is not universal, as Irabu does not permit voiceless stops to receive coerced weight, while Ikema appears to not allow fricatives to do the same, while allowing (some) stops to be coerced. Finally, it has been found that the patterning of these Miyako onset phenomena can be fully captured by violable markedness and faithfulness constraints. An attempt has been made to describe these geminate and seemingly complex onsets through the constraint-based grammar of Optimality Theory. The findings of this analysis indicate that the rankings of the constraints in each language accounts for the smaller inter-Miyako variations that characterize the onset phonotactics of each dialect.

It is my hope that this attempt at examining the Miyako geminates will inspire further analysis and study of the Ryukyuan languages, both for the documentation of languages that thus far lack formal descriptive grammars, and for the further examination of the data that has been gathered regarding this group of languages. As mentioned in the introduction to this thesis, many of the Miyako or Ryukyuan languages may continue to progress towards extinction in the near future. It is nonetheless an encouraging sign that more awareness is being raised about the linguistic diversity of the Ryukyus, and that academic interest in these languages has been strengthened in recent years. Languages with unexpected patterns allow us to gain a greater understanding of the boundaries of what is possible within natural language, and we must be quick to study them while we still have the option to do so.

WORKS CITED

Alber, Birgit, et al. “From Intensional Properties to Universal Support.” Language, vol. 92, no. 2, 2016, pp. e88–e116., doi:10.1353/lan.2016.0029.

Al-Deaibes, Mutasim. “Gemination: Weight or Length? Evidence from Rural Jordanian Arabic.” *WORD*, vol. 67, no. 2, 2021, pp. 159–171., doi:10.1080/00437956.2021.1909841.

Anderson, Mark. “Revitalisation Attempts and Language Attitudes in the Ryukyus.” *Language Crisis in the Ryukyus*, edited by Mark Anderson and Patrick Heinrich, Cambridge Scholars Publishing, 2014, pp. 1–30.

Anderson, Mark, and Patrick Heinrich, editors. *Language Crisis in the Ryukyus*. Cambridge Scholars Publishing, 2014.

Anttila, Arto. “Variation and Optionality.” *The Cambridge Handbook of Phonology*, edited by Paul De Lacy, 519–36. Cambridge: Cambridge University Press, 2007.

Aoi, Hayato. “Tarama Miyako Grammar.” *Handbook of the Ryukyuan Languages: History, Structure, and Use*, edited by Patrick Heinrich et al., De Gruyter Mouton, 2015, pp. 405–422.

Aoyama, Katsura, and Lawrence A. Reid. “Cross-Linguistic Tendencies and Durational Contrasts in Geminate Consonants: An Examination of Guinaang Bontok Geminates.” *Journal of the International Phonetic Association*, vol. 36, no. 2, 2006, pp. 145–157., doi:10.1017/s0025100306002520.

Armosti, Spyros. “Fricative and sonorant super-geminates in Cypriot Greek: a perceptual study.” *Studies in Modern Greek Dialects and Linguistic Theory*. Research Centre of Kykkos Monastery, 2011, pp. 97-112.

Aso, Reiko. “Hateruma Yaeyama Grammar.” *Handbook of the Ryukyuan Languages: History, Structure, and Use*, edited by Thomas Pellard et al., De Gruyter Mouton, 2015, pp. 423–447.

Butler, Becky. “Approaching a Phonological Understanding of the Sesquisyllable with Phonetic Evidence from Khmer and Bunong.” *Languages of Mainland Southeast Asia: The State of the Art*, edited by N. J. Enfield and Bernard Comrie, De Gruyter Mouton, 2017, pp. 443–499.

Benhallam, Abderrafi. “Syllable structure and rule types in Arabic.” *University of Florida*, 1980, Ph.D. dissertation.

Bermúdez-Otero, Ricardo. “Underlyingly Nonmoraic Coda Consonants, Faithfulness, and Sympathy.” University of Manchester, Jan. 2001. Retrieved from bermudez-otero.com/DEP-mora.pdf.

Bhaskararao, Peri, and Peter Ladefoged. “Two Types of Voiceless Nasals.” *Journal of the International Phonetic Association*, vol. 21, no. 2, 1991, pp. 80–88., doi:10.1017/s0025100300004424.

Blust, Robert A. “Òma Lóngh Historical Phonology.” *Oceanic Linguistics*, vol. 46, no. 1, 2007, pp. 1–53., doi:10.1353/ol.2007.0016.

Boersma, Paul, and Bruce Hayes. “Empirical Tests of the Gradual Learning Algorithm.” *Linguistic Inquiry*, vol. 32, no. 1, 2001, pp. 45–86., doi:10.1162/002438901554586.

Celik, Kenan, and Yukinori Takubo. ”Ikema Topic and Accusative Marker, an OT Analysis.” FAJL poster sessions, NINJAL, 2014. Retrieved from iculinguistics.files.wordpress.com/2015/01/celik_fajl7.pdf

Chirkova, Katia, et al. “Voiceless Nasal Sounds in Three Tibeto-Burman Languages.” *Journal of the International Phonetic Association*, vol. 49, no. 1, 2018, pp. 1–32., doi:10.1017/s0025100317000615.

Chomsky, Noam, and Morris Halle. *The Sound Pattern of English*. Harper & Row, 1968.

Clements, George N. “The Role of the Sonority Cycle in Core Syllabification.” *Papers in Laboratory Phonology*, 1990, pp. 283–333., doi:10.1017/cbo9780511627736.017.

Davis, Stuart. “Geminates and Weight-Manipulating Phonology in Chuukese (Trukese).” *The Phonetics and Phonology of Geminate Consonants*, edited by Haruo Kubozono, Oxford University Press, 2017, pp. 230–259.

Davis, Stuart. “Geminates.” *The Blackwell Companion to Phonology*, Wiley-Blackwell, 2011.

Davis, Stuart. “On the Representation of Initial Geminates.” *Phonology*, vol. 16, no. 1, 1999, pp. 93–104., doi:10.1017/s0952675799003711.

DelBusso, Natalie R. “Typological Structure and Properties of Property Theory.” Rutgers, the State University of New Jersey, 2018, Ph.D. dissertation.

Denzer-King, Ryan E. “The Distribution of /s/ in Blackfoot: An Optimality Theory Account.” *The University of Montana*, 2009, M.A. thesis.

Dmitrieva, Olga. “The Role of Perception in the Typology of Geminate Consonants: Effects of Manner of Articulation, Segmental Environment, Position, and Stress.” *Language and Speech*, vol. 61, no. 1, 2017, pp. 43–70., doi:10.1177/0023830917696113.

Downing, Laura J. “Problems in Jita Tonology.” *University of Illinois*, 1990.

Ehrenhofer, Laura, et al. “Asymmetric processing of consonant duration in Swiss German.” *The Phonetics and Phonology of Geminate Consonants*, edited by Haruo Kubozono, Oxford University Press, 2017, pp. 205–229.

Palancar, Enrique L. “Preaspiration in Northern Otomi: A relic of an old fortis/lenis contrast.” Base articulatorie arrière. Backing and Backness, LINCOM Europa, LINCOM Studies in Phonology 01, 2013, hal-00920515.

Fangel-Gustavson, Nora, et al. “Quantity Contrast in Lule Saami: A Three-Way System.” Unpublished, 2014, doi:10.13140/2.1.5171.5846.

Ford, Catherine. “The Status of Voiceless Nasals in Ikema Ryukyuan”. University of Alberta, 2016, BA thesis.

Gordon, Matthew. *Syllable Weight: Phonetics, Phonology, Typology*. Routledge, 2006.

Gordon, Matthew. “A Perceptually-Driven Account of Onset-Sensitive Stress.” *Natural Language & Linguistic Theory*, vol. 23, no. 3, 2005, pp. 595–653., doi:10.1007/s11049-004-8874-9.

Hajek, John, and Rob Goedemans. “Word-Initial Geminates and Stress in Pattani Malay.” *The Linguistic Review*, vol. 20, no. 1, 2003, doi:10.1515/tlir.2003.003.

Hayashi, Yuka. “Ikema (Miyako Ryukyuan).” *An Introduction to Ryukyuan Languages*, edited by Michinori Shimoji and Thomas Pellard, Research Institute for Languages and Cultures of Asia and Africa, 2010, pp. 167–188.

Hayashi, Yuka. “Ryūkyū-go Miyako Ikema Hōgen-no danwa siryō.” [Ryukyuan Miyako Ikema dialect discourse data.], *Chikyū-ken gengō kijūtu ronsyū 1*, 2009, pp. 153-199.

Hayes, Bruce. *Metrical Stress Theory: Principles and Case Studies*. The University of Chicago Press, 1995.

Hayes, Bruce. “Compensatory Lengthening in Moraic Phonology.” *Linguistic Inquiry*, vol. 20, no. 2, 1989, pp. 253–306.

Hayes, Bruce. “Inalterability in CV Phonology.” *Language*, vol. 62, no. 2, 1986, p. 321., doi:10.2307/414676.

Heinrich, Patrick, et al., editors. *Handbook of the Ryukyuan Languages: History, Structure, and Use*. De Gruyter Mouton, 2015.

Hirayama, Teruo (editor). “Ryūkyū Sakishima-Hōgen-no Sōgōteki-Kenkyū.” [Comprehensive research on the Ryukyuan Sakishima dialects]. Meiji Shoin, 1967.

Hōsei daigaku Okinawa bunka kenkyūjo [Hōsei University Okinawa Cultural Research]. ”Ryūkyū-no hōgen: Miyako Ōgami-jima.” [Dialects of Ryukyu: Miyako Ōgami-jima]. Hōsei University Press, 1977.

Idemaru, Kaori, and Susan G. Guion. “Acoustic Covariants of Length Contrast in Japanese Stops.” *Journal of the International Phonetic Association*, vol. 38, no. 2, 2008, pp. 167–186., doi:10.1017/s0025100308003459.

Igarashi, Yosuke, et al. “The Ikema Dialect of Miyako Ryukyuan Has a Three-, Not Two-, Pattern Accent System.” *Journal of the Phonetic Society of Japan*, vol. 20, no. 1, Apr. 2016, pp. 47–60.

Igarashi, Yosuke. “The Nominal Accent System of the Yonaha Dialect of the Southern Ryukyuan Language of Miyako.”, *General Study for Research and Conservation of Endangered Dialects in Japan: Research Report on Miyako Ryukyuan*, edited by Nobuko Kibe, NINJAL Collaborative Research Project Reports 12-02, 2012, published in English in 2019, pp. 56-75.

Itô, Junko. “Prosodic Minimality in Japanese.” *Papers from the Parasession on the Syllable in Phonetics and Phonology*, CLS 26-II, 1990, pp. 213–239.

Itô, Junko, and Armin Mester. “Licensed Segments and Safe Paths.” *Canadian Journal of Linguistics/Revue Canadienne De Linguistique*, vol. 38, no. 2, 1993, pp. 197–213., doi:10.1017/s0008413100014742.

Itô, Junko, et al. “A Prosodic Account of Consonant Gemination in Japanese Loanwords.” *The Phonetics and Phonology of Geminate Consonants*, edited by Haruo Kubozono, Oxford University Press, 2017, pp. 283–320.

Itô, Junko. “Syllable Theory in Prosodic Phonology.” University of Massachusetts, University Microfilms International, 1987, Ph.D. dissertation.

Izuyama, Atsuko. “Yonaguni.” *The Languages of Japan and Korea*, edited by Nicolas Tranter, Routledge, 2012, pp. 412–457.

Jakobson, Roman, et al. *Preliminaries to Speech Analysis*. M.I.T. Press, 1951.

Jany, Carmen, et al. “How universal is the sonority hierarchy?: A cross-linguistic acoustic study.” (2007).

Jarosz, Aleksandra. “Reconstructing Miyako-Ryukyuan Of 1920s: a Tentative Grammar Sketch.” *Silva Iaponicarum*, 2014, pp. 25–46., doi:10.14746/sijp.2015.38/39.2.

Jay, Stephen, and Steve Parker. “Onset Cluster Typologies.” *Occasional Papers in Applied Linguistics*, Dallas International University, 2020.

Kager, René. *Optimality Theory*. Cambridge University Press, 1999.

Karimata, Shigehisa. “Verb Conjugation in the Miyako Language: Perfective, Negative, Past, and Continuative Forms.”, *General Study for Research and Conservation of Endangered Dialects in Japan: Research Report on Miyako Ryukyuan*, edited by Nobuko Kibe, NINJAL Collaborative Research Project Reports 12-02, 2012, published in English in 2019, pp. 76-140.

Kawahara, Shigeto. “Japanese Has Syllables: A Reply to Labrune.” *Phonology*, vol. 33, no. 1, 2016, pp. 169–194., doi:10.1017/s0952675716000063.

Kawahara, Shigeto, and Melanie Pangilinan. “Spectral Continuity, Amplitude Changes, and Perception of Length Contrasts.” *The Phonetics and Phonology of Geminate Consonants*, edited by Haruo Kubozono, Oxford University Press, 2017, pp. 13–33.

Kawahara, Shigeto. “Sonorancy and geminacy.” *University of Massachusetts Occasional Papers in Linguistics, 32*, 2007, pp. 145–186.

Kawai, Junji. “Phonological Analysis of Casual Japanese Speech in Optimality Theory.” *University of Canterbury*, 2004, Ph.D. dissertation.

Kehrein, Wolfgang. *Phonological Representation and Phonetic Phasing: Affricates and Laryngeals*. Niemeyer, 2002.

Kenstowicz, Michael, and Charles Pyle. “On the Phonological Integrity of Geminate Clusters.” *Issues in Phonological Theory*, 1973, pp. 27–43., doi:10.1515/9783110905137.27.

Kibe, Nobuko (editor). “General Study for Research and Conservation of Endangered Dialects in Japan: Research Report on Miyako Ryukyuan.” *NINJAL Collaborative Research Project Reports 12-02*, 2012, published in English in 2019.

Kindred, Shaun (editor). “Miyako-Ikema - Japanese - English Dictionary.” Compiled by Celik et al., 2019. www.webonary.org/miyako-ikema/

Kiparsky, Paul. “Opacity and Cyclicity.” *The Linguistic Review*, vol. 17, no. 2-4, 2000, doi:10.1515/tlir.2000.17.2-4.351.

Kirchner, Robert. “Geminate Inalterability and Lenition.” *Language*, vol. 76, no. 3, 2000, pp. 509–545., doi:10.2307/417134.

Kraehenmann, Astrid. “Initial Geminates.” *The Blackwell Companion to Phonology*, 2011, pp. 1–23., doi:10.1002/9781444335262.wbctp0047.

Krämer, Martin, and Draga Zec. “Nasal Consonants, Sonority and Syllable Phonotactics: The Dual Nasal Hypothesis.” *Phonology*, vol. 37, no. 1, 2020, pp. 27–63., doi:10.1017/s0952675720000032.

Krauss, Michael. “The World's Languages in Crisis.” *Language*, vol. 68, no. 1, 1992, pp. 4–10., doi:10.1353/lan.1992.0075.

Kubozono, Haruo. “The Syllable as a Unit of Prosodic Organization in Japanese.” *The Syllable in Optimality Theory*, edited by Féry Caroline and Ruben van de Vijver, Cambridge University Press, 2003, pp. 99–122.

Kubozono, Haruo. “Introduction to the Special Issue on Japanese Geminate Obstruents.” *Journal of East Asian Linguistics*, vol. 22, no. 4, 2013, pp. 303–306., doi:10.1007/s10831-013-9109-z.

Labrune, Laurence. “Questioning the Universality of the Syllable: Evidence from Japanese.” *Phonology*, vol. 29, no. 1, 2012, pp. 113–152., doi:10.1017/s095267571200005x.

Labrune, Laurence. “The Phonology of Japanese /r/: a Panchronic Account.” *Journal of East Asian Linguistics*, vol. 23, no. 1, 2013, pp. 1–25., doi:10.1007/s10831-013-9117-z.

Ladefoged, Peter, and Ian Maddieson. *The Sounds of the World's Languages*. Blackwell, 1996.

Lawrence, Wayne P. “Southern Ryukyuan.” *The Languages of Japan and Korea*, edited by Nicolas Tranter, Routledge, 2012, pp. 381–411.

Levin, Juliette. “A Metrical Theory of Syllabicity.” *Massachusetts Institute of Technology*, 1985, Ph.D. dissertation.

Lin, Yen-Hwei. “Syllabic and Moraic Structures in Piro.” *Phonology*, vol. 14, no. 3, 1997, pp. 403–436., doi:10.1017/s0952675798003455.

Lombardi, Linda. “The Nonlinear Organization of the Affricate.” *Natural Language and Linguistic Theory*, vol. 8, no. 3, 1990, pp. 375–425., doi:10.1007/bf00135619.

Matayoshi, Satomi. “Tsuken (Okinawan).” *An Introduction to Ryukyuan Languages*, edited by Michinori Shimoji and Thomas Pellard, Research Institute for Languages and Cultures of Asia and Africa, 2010, pp. 89–113.

Matsuura, Toshio. “Phonetic Realizations in the Miyako Dialects: A Preliminary Investigation.”, *General Study for Research and Conservation of Endangered Dialects in Japan: Research Report on Miyako Ryukyuan*, edited by Nobuko Kibe, NINJAL Collaborative Research Project Reports 12-02, 2012, published in English in 2019, pp. 141-158.

Matsuura, Toshio. “Voiced Geminate Obstruents in Sino-Japanese Words in the Fukami Dialect of Amakusa Japanese.” *GENGO KENKYU (Journal of the Linguistic Society of Japan)*, vol. 158, 2020, pp. 29–61., doi:https://doi.org/10.11435/gengo.158.0_29.

McCarthy, John Joseph. “Formal Problems in Semitic Phonology and Morphology.” *Massachusetts Institute of Technology*, 1979, Ph.D. dissertation.

McCarthy, John J., and Alan Prince. “Faithfulness and Reduplicative Identity.” *Papers in Optimality Theory, University of Massachusetts Occasional Papers*, vol. 18, 1995, pp. 249–384.

McKeever, Ashley L. “Sonority and Syllable Weight in Fur.” *University of Texas at Arlington*, 2014, Ph.D. dissertation.

Mitterer, Holger. “Not All Geminates Are Created Equal: Evidence from Maltese Glottal Consonants.” *Journal of Phonetics*, vol. 66, 2018, pp. 28–44., doi:10.1016/j.wocn.2017.09.003.

Morelli, Frida. “The Phonotactics and Phonology of Obstruent Clusters in Optimality Theory.” *University of Maryland*, 1999, Ph.D. dissertation.

Morén, Bruce. “Weight Typology: An Optimality Theoretic Approach.” *The Linguistic Review*, vol. 20, no. 2-4, 2003, doi:10.1515/tlir.2003.011.

Morén, Bruce. “Distinctiveness, Coercion and Sonority: A Unified Theory of Weight.” University of Maryland at College Park, 1999, Ph.D. dissertation.

Muller, Jennifer S. “On the Theoretical Implications of Cypriot Greek Initial Geminates.” *Journal of Greek Linguistics*, vol. 3, no. 1, 2002, pp. 115–137., doi:10.1075/jgl.3.06mul.

Muller, Jennifer S. “The Phonology and Phonetics of Word-Initial Geminates.” *Ohio State University*, 2001, Ph.D. dissertation.

Nevins, Andrew, and Ioana Chitoran. “Phonological Representations and the Variable Patterning of Glides.” *Lingua*, vol. 118, no. 12, 2008, pp. 1979–1997., doi:10.1016/j.lingua.2007.10.006.

Noamane, Ayoub. “On the Integrity of Geminates in Moroccan Arabic: An Optimality-Theoretic Account .” *Journal of Applied Language and Culture Studies*, 2018, pp. 133–164.

Norquest, Peter K. *A Phonological Reconstruction of Proto-Hlai*. Brill, 2015.

Ohala, John J. “The Origin of Sound Patterns in Vocal Tract Constraints.” *The Production of Speech*, 1983, pp. 189–216., doi:10.1007/978-1-4613-8202-7_9.

Pajak, Bozena. “Non-Intervocalic Geminates: Typology, Acoustics, Perceptibility.” *UC San Diego: San Diego Linguistics Papers*, 2013. Retrieved from https://escholarship.org/uc/item/2bb4829q

Parker, Stephen G. “Quantifying the Sonority Hierarchy.” *University of Massachusetts Amherst*, 2002, Ph.D. dissertation.

Pellard, Thomas. “The Linguistic Archeology of the Ryukyu Islands.” *Handbook of the Ryukyuan Languages: History, Structure, and Use*, edited by Patrick Heinrich et al., De Gruyter Mouton, 2015, pp. 13–37.

Pellard, Thomas. “Ōgami: Éléments De Description d’Un Parler Du Sud Des Ryūkyū.” *École des Hautes Études en Sciences Sociales (EHESS)*, 2009.

Pellard, Thomas. “Ōgami.” *An Introduction to Ryukyuan Languages*, edited by Michinori Shimoji and Thomas Pellard, Research Institute for Languages and Cultures of Asia and Africa, 2010, pp. 113–166.

Pellard, Thomas. “Why /r/ Is Not a Special, Empty Consonant in Japanese.” *Journal of East Asian Linguistics*, vol. 25, no. 4, 2016, pp. 351–383., doi:10.1007/s10831-016-9147-4.

Pellard, Thomas, and Yuka Hayashi. “The Phonology of the Miyako Dialects: Phonological Systems and Comparisons.”, *General Study for Research and Conservation of Endangered Dialects in Japan: Research Report on Miyako Ryukyuan*, edited by Nobuko Kibe, NINJAL Collaborative Research Project Reports 12-02, 2012, published in English in 2019, pp. 13-55.

Prince, Alan, and Paul Smolensky. “Optimality Theory: Constraint Interaction in Generative Grammar.” *Optimality Theory in Phonology: A Reader*, edited by John J. McCarthy, Media Production Services Unit, Manitoba Education, 2004, pp. 1–71.

Prince, Alan S. “A Metrical Theory for Estonian Quantity.” Linguistic Inquiry, vol. 11, no. 3, MIT Press, 1980, pp. 511–62, http://www.jstor.org/stable/4178178.

Ringen, Catherine O., and Robert M. Vago. “Geminates: Heavy or Long?” *Handbook of the Syllable*, edited by Charles E. Cairns and Eric Raimy, Brill, 2011, pp. 155–169.

Rosenthall, Samuel. “Vowel/Glide Alternation in a Theory of Constraint Interaction.” *University of Massachusetts*, 1994, Ph.D. dissertation.

Sagey, Elizabeth. “The Representation of Features and Relations in Non-Linear Phonology.” *Massachusetts Institute of Technology*, 1986.

Sakiyama, Osamu. *Studies of Minority Languages in the Western Pacific Rim*. Osaka Gakuin University, 2003.

Sato, Yumiko. “The Phonetic Reality of the Mora in Japanese: A Cross-Linguistic Study on Timing in Japanese, English, and Korean.” *University of Hawai'i at Manoa*, 1998, Ph.D. dissertation.

Selkirk, Elisabeth. “A Two-Root Theory of Length.” *University of Massachusetts Occasional Papers in Linguistics*, Vol. 16, Iss. 3, Article 7, 1990.

Selkirk, Elisabeth. “On the Major Class Features and Syllable Theory.” *Language Sound Structure: Studies in Phonology*, edited by Morris Halle et al., MIT Press, 1984, pp. 107–136.

Shibatani, Masayoshi. *The Languages of Japan*. Cambridge University Press, 2010.

Shih, Shu-Hao. “Non-Moraic Schwa: Phonology and Phonetics.” *Rutgers, The State University of New Jersey*, 2018, Ph.D. dissertation.

Shimoji, Michinori. “Minami-Ryūkyū Miyako Irabu-jima Hōgen-no ‘go’ ni-tsuite” [About the ‘words’ of the Southern Ryukyuan Irabu dialect.], 2018. Retrieved from www.academia.edu/38333439.

Shimoji, Michinori. “Irabu Ryukyuan.” *Grammatical Sketches from the Field*, edited by Yasuhiro Yamakoshi, Research Institute for Languages and Cultures of Asia and Africa, 2011, pp. 77–131.

Shimoji, Michinori. “Ryukyuan Languages: An Introduction.” *An Introduction to Ryukyuan Languages*, edited by Michinori Shimoji and Thomas Pellard, Research Institute for Languages and Cultures of Asia and Africa, 2010, pp. 1–13.

Shimoji, Michinori, and Thomas Pellard, editors. *An Introduction to Ryukyuan Languages*. Research Institute for Languages and Cultures of Asia and Africa, 2010.

Shimoji, Michinori. “A Grammar of Irabu, a Southern Ryukyuan Language.” *The Australian National University*, 2008, Ph.d. dissertation.

Shimoji, Michinori. “A general overview of Irabu Ryukyuan, a Southern Ryukyuan language of the Japonic language group.”, *Tokyo University of Foreign Studies Descriptive Linguistics Collection*, 2006.

Shimoji, Michinori. "Northern Ryukyuan.", in The Languages of Japan and Korea, edited by Nicolas Tranter, Abingdon: Routledge, 16 May 2012, Routledge Handbooks Online.

Shimoji, Kayoko. “Miyako Tarama Hōgen no On’in oyobi Sono Henka no Genshō.” [The Phonology and change phenomena of the Tarama Miyako Dialect.] *Ryūkyū no Hōgen* 28, 2003, pp. 93 – 113.

Shinohara, Shigeko, and Masako Fujimoto. “Acoustic characteristics of the obstruent and nasal geminates in the Ikema dialect of Miyako Ryukyuan.” Babatsouli, E. Crosslinguistic research in monolingual and bilingual speech, pp.253-270, 2018, 978-618-82351-2-0. halshs-01896658

Shinohara, Shigeko, and Masako Fujimoto. “Moraicity of Initial Gemiantes in the Tedumuni Dialect of Okinawa.” *17th International Congress of Phonetic Sciences*, Aug. 2011, pp. 1826–1829.

Shinohara, Shigeko, and Qandeel Hussain. “Ikema Miyako (Japan) – Language Snapshot.” *Language Documentation and Description*, 2021, EL Publishing, Language Documentation and Description, 19, pp. 240-248. (hal-03184734).

Silverman, Daniel. “Voiceless Nasals in Auditory Phonology.” *Annual Meeting of the Berkeley Linguistics Society*, vol. 22, no. 1, 1996, p. 364., doi:10.3765/bls.v22i1.1357.

Starr, Rebecca Lurie, and Stephanie S. Shih. “The Syllable as a Prosodic Unit in Japanese Lexical Strata: Evidence from Text-Setting.” *Glossa: a Journal of General Linguistics*, vol. 2, no. 1, 2017, p. 93., doi:10.5334/gjgl.355.

Takei, Honoka. “An Investigation of Direct Object Markers in the Ikema Dialect of Miyako.” University of Alberta, 2016, MA thesis.

Takubo, Yukinori. “Morphophonemics of Ikema Miyakoan.” *Studies in Asian Historical Linguistics, Philology and Beyond*, edited by John Kupchik et al., Brill, 2021, pp. 65–79.

Thurgood, Graham. “Geminates: A cross-linguistic examination.” *Papers in honor of Frederick H. Brengelman on the occasion of the twenty-fifth anniversary of the Department of Linguistics, CSU Fresno,* edited by Nevis et al., 1993, pp. 129-139.

Topintzi, Nina, and Stuart Davis. “On the Weight of Edge Geminates.” *The Phonetics and Phonology of Geminate Consonants,* edited by Haruo Kubozono, Oxford University Press, 2017, pp. 260–282.

Topintzi, Nina. “On the Existence of Moraic Onset Geminates.” *Natural Language & Linguistic Theory*, vol. 26, no. 1, 2008, pp. 147–184., doi:10.1007/s11049-008-9034-4.

Topintzi, Nina. *Moraic Onsets*. University of London, 2006.

Topintzi, Nina. *Onsets: Suprasegmental and Prosodic Behaviour*. Cambridge University Press, 2010.

Utman, Jennifer Aydelott, and Sheila E. Blumstein. “The Influence of Language on the Acoustic Properties of Phonetic Features: A Study of the Feature [STRIDENT] in Ewe and English.” *Phonetica*, vol. 51, no. 4, 1994, pp. 221–238., doi:10.1159/000261978.

Vennemann, Theo. *Preference Laws for Syllable Structure and the Explanation of Sound Change: With Special Reference to German, Germanic, Italian, and Latin*. Mouton De Gruyter, 1988.

Yamane, Noriko. “'Placeless' Consonants in Japanese: An Ultrasound Investigation.” University of British Columbia, 2013, Ph.D. dissertation.

Yoshimura, Sayaka. “Japanese Language Education in the Meiji Period.” *Language Crisis in the Ryukyus*, edited by Mark Anderson and Patrick Heinrich, Cambridge Scholars Publishing, 2014, pp. 31–53.

Zec, Draga. “Sonority Constraints on Prosodic Structure.” Stanford University, 1989, Ph.D. dissertation.

Zec, Draga. “Sonority Constraints on Syllable Structure.” *Phonology*, vol. 12, no. 1, 1995, pp. 85–129., doi:10.1017/s0952675700002396.

Zec, Draga. “The Syllable.” *The Cambridge Handbook of Phonology*, edited by Paul De Lacy, Cambridge University Press, 2007, pp. 161–194.

Żygis, Marzena, et al. “Phonetic Explanations for the Infrequency of Voiced Sibilant Affricates across Languages.” *Laboratory Phonology*, vol. 3, no. 2, 2012, doi:10.1515/lp-2012-0016.