Title

An Alternative Account of the Japanese-Ryukyuan Split Tonal Correspondences

Journal

Japanese/Korean Linguistics, 32(1)

Author

Igarashi, Yosuke

Publication Date

2026-07-31

DOI

10.5070/J7.66707

An Alternative Account of the Japanese-Ryukyuan Split Tonal Correspondences*

YOSUKE IGARASHI
NINJAL/Graduate University for Advanced Studies, SOKENDAI

1 Introduction

This study proposes an alternative explanation for the long-standing problem in the historical linguistics of Japonic languages, first noted by Hattori (1958), namely the split tonal correspondences between Ryukyuan and Japanese. This issue is of central importance, as its interpretation fundamentally reshapes the tonal system reconstructed for proto-Japonic (pJ), the common ancestor of Ryukyuan and Japanese.

Kindaichi (1974) reconstructed tonal classes for 11th-century Middle Japanese (MJ) based on both modern Japanese varieties and historical textual sources. His reconstruction identified three classes for monosyllabic nouns (1.1–1.3), five for disyllabic nouns (2.1–2.5), and six for trisyllabic nouns (3.1–3.2, 3.4–3.7, with 3.3 absent). Because Kindaichi did not sufficiently incorporate data from Ryukyuan—Japonic varieties spoken on the Ryukyu Islands, his tonal classes cannot be regarded as pJ classes. Accordingly, this study refers to them as MJ tonal classes, although, for convenience, the same class labels are also used to describe earlier stages.

* I am grateful to Kenan Celik, Kohei Nakazawa, and Simão Gonçalves for fruitful discussions. This study is partially based on an earlier oral presentation (Igarashi 2022b), but has been revised with new discussion and empirical data. This research was supported by JSPS KAKENHI Grant Number JP24K00070 and by the NINJAL project "Empirical Study on the Intonational Diversity in Japanese and Ryukyuan Dialects."

Japanese/Korean Linguistics 32.
Edited by Youngdong Cho, Young-Hoon Kim, and John Whitman.

MJ classes	pR classes	Sub-classes		Kyoto	Oita	Yoron	Nakijin
2.1	A		‘nose’ ‘water’	háná mízú	haná-gá midú-gá	pana miʑi	pʰanáá midʑíi
2.2	A		‘sound’ ‘snow’	óto júki	otó-gá jukí-gá	utu juki	ɸutˀúu jutʑˀíi
2.3	B		‘mountain’ ‘year’	jáma tóɕi	jamá-ga toɕí-ga	jamá tuɕí	jamaá tʰuɕii
2.4	C	2.4a	‘middle’ ‘breath’	naká ikí	náka-ga íki-ga	náá íkí	náhaa ʔítʑˀi
	B	2.4b	‘shoulder’ ‘barley’	katá mugí	káta-ga múgi-ga	hatá mugí	hatˀaá mudʑii
2.5	C	2.5a	‘pot’ ‘monkey’	nabê sarû	nábe-ga sáru-ga	nábí sárú	nábi sáru
	B	2.5b	‘rain’ ‘sweat’	amê asê	áme-ga áse-ga	amí aɕí	amii aɕii

Table 1. Tonal correspondences between Japanese varieties (Kyoto and Oita; Hirayama et al. eds. 1992–93) and Ryukyuan varieties (Yoron; Kiku and Takahashi 2005, and Nakijin; Nakasone 1983). Acute accents indicate high pitch, whereas circumflex accents indicate falling pitch.

Ryukyuan varieties form a distinct branch composed exclusively of Japonic varieties spoken on the Ryukyu Islands, and their most recent common ancestor is termed proto-Ryukyuan (pR) (Pellard 2015). Studies on Ryukyuan have shown that three tonal classes (A, B, and C) can be reconstructed for pR (Hattori 1958, 1978–79; Matsumori 2012). Notably, as first demonstrated by Hattori (1958) and confirmed by subsequent research (Matsumori 1996, 2009, 2012; Shimabukuro 2007; de Boer 2010; Igarashi 2022a), MJ Classes 2.4–2.5 exhibit split correspondences with pR tonal classes.

Table 1 summarizes the tonal correspondences between two Japanese varieties (Kyoto and Oita) and two Ryukyuan varieties (Yoron and Nakijin). Word forms are presented in broad phonetic transcription: glottalized consonants are indicated with the diacritic ˀ, and long vowels are represented as a sequence of two identical vowels. Acute and circumflex accents indicate high and falling pitches, respectively.

MJ Classes 2.1 and 2.2 correspond to pR Class A, and MJ Class 2.3 corresponds to pR Class B. By contrast, MJ Classes 2.4 and 2.5 exhibit a split correspondence: approximately half of the items correspond to pR Class C, and the remaining half to pR Class B. To account for this pattern, scholars have posited additional subclasses—Classes 2.4a and 2.5a (corresponding to pR Class C) and Classes 2.4b and 2.5b (corresponding to pR Class B).

Hattori (1958, 1978–79) also identified split correspondences among Class 2.3 nouns, and Matsumori (2000) argued for similar splits in Classes 3.4 and 3.5. Nevertheless, most Class 2.3 and Class 3.4 nouns correspond to pR Class B (Igarashi 2022a: fn. 1). Therefore, these irregular correspondences can be treated as sporadic exceptions, or at least as distinct from the systematic split of Classes 2.4–2.5, where nouns are roughly evenly divided between pR Classes B and C. Regarding Class 3.5, irregular correspondences are also attested within Japanese itself (Matsumori 1997), and it remains debated whether this class is derived from a single pJ source (Uwano 2006).

Thus, the split correspondences in Classes 2.3, 3.4, and 3.5 likely stem from different principles and will not be discussed further here.

If the split correspondence in MJ Classes 2.4–2.5 is not an innovation within pR, it must reflect a distinction inherited from pJ. The straightforward interpretation, then, is that pJ has more than five tonal classes of disyllabic nouns. Such a reconstruction would yield a tonal system with a large inventory and few restrictions on tone distribution. However, some scholars argue that split correspondences arose instead through a conditioned merger in pR (Kindaichi 1960; Hirayama et al. 1966; Tokugawa 1990). This account, the *conditioned merger hypothesis* (CMH), has proven untenable as it entails too many exceptions (Uwano 1996; de Boer 2010). In the absence of a clear conditioning factor, scholars have therefore posited additional tonal contrasts in pJ. The reconstruction of four subclasses—2.4a, 2.5a, 2.4b, and 2.5b—in addition to 2.1–2.3 is now widely accepted (Hattori 1978–79; Vovin 2008; Shimabukuro 2008; Pellard 2009).

Contrary to this prevailing view, this study advances an alternative account: the *revised conditioned merger hypothesis* (RCMH). The RCMH explains the split correspondence as the result of a tonal merger conditioned by two factors: a word-final high vowel and an original morpheme boundary. Section 2 reviews the CMH and identifies its limitations. Section 3 outlines the RCMH. Section 4 presents evidence that a now-lost word-internal morpheme boundary can be recovered through the internal reconstruction of Old Japanese (OJ)—the earliest attested stage of the language dating back to the eighth century—as well as through unexpected heavy syllables in modern Japanese varieties and complementary etymological analysis. Section 5 discusses exceptions to the RCMH, and Section 6 concludes the paper.

Our investigation is restricted to Class 2.4–2.5 nouns that satisfy the following criteria: (i) they are listed in Kindaichi (1974); (ii) their distribution across all Ryukyuan varieties permits an unambiguous reconstruction at the level of pR; and (iii) they display a high degree of regularity in their sound correspondences. In total, 75 nouns meet these criteria (Igarashi 2022a). The MJ tonal classes were taken from Kindaichi (1974), the pR classes from Igarashi (2022a)—a revision and extension of Hattori (1978–79), Matsumori (2009, 2012), Shimabukuro (2007), and de Boer (2010)—and the OJ forms from JJ (1967) and ONCOJ (2020).

2 Previous CMH

According to the earlier version of the CMH (Kindaichi 1960; Hirayama et al. 1966; Tokugawa 1990), Class 2.4–2.5 nouns ending in a non-high vowel merged into Class 2.3 (corresponding to pR Class B), whereas those ending in a high vowel remained in a distinct tonal class (pR Class C).

For the tonal system of pR prior to the merger, the CMH assumes essentially the same system as that of the so-called Gairin type (Kindaichi 1974), found in Japanese varieties such as Oita in northeastern Kyushu. Specifically, it posits σσ (unaccented) for Class 2.1–2.2 nouns, σσ́ (final accent) for Class 2.3 nouns, and σ́σ (initial/penultimate accent) for Class 2.4–2.5 nouns, where σ indicates a syllable and the acute accent marks the accent, phonetically realized as a pitch peak (1).

(1) Tonal patterns in the early stage of pR proposed in the CMH

a.	Class 2.1–2.2	σσ	e.g.,	*kaze 2.1 ‘wind’, *kawa 2.2 ‘well (noun)’
b.	Class 2.3	σσ́	e.g.,	*jamá 2.3 ‘mountain’
c.	Class 2.4–2.5	σ́σ	e.g.,	*íta 2.4 ‘board’, *íki 2.4 ‘breath’
				*áme 2.5 ‘rain’, *sáru 2.5 ‘monkey’

Group	No. of nouns	Final vowel	Tonal class	
Group 1	15	non-high	2.4a	***ato** ‘trace’, ***ito** ‘thread’, ***jado** ‘shelter’, ***kesa** ‘short while ago’, ***kuda** ‘tube’, ***naka** ‘middle’, ***pera** ‘spatula’, ***pune** ‘boat’, ***sora** ‘treetop’
			2.5a	***kage** ‘shadow’, ***koe** ‘voice’, ***mae** ‘front’, ***moko** ‘bridegroom’, ***nabe** ‘pot’, ***woke** ‘bucket’
Group 2	21	High	2.4a	***beni** ‘rouge’, ***iki** ‘breath’, ***kami** ‘above’, ***keu** ‘today’, ***matu** ‘pine’, ***nau** ‘what’, ***nusi** ‘owner’, ***obi** ‘belt’, ***omi** ‘sea’, ***Usu** ‘mortar’, ***pari** ‘needle’, ***pasi** ‘chopstick’, ***suzi** ‘sinew’, ***tubu** ‘grain’, ***tumi** ‘sin’, ***tuti** ‘hammer’, ***wanu** ‘I’
			2.5a	***kobu** ‘spider’, ***saru** ‘monkey’, ***tabi** ‘socks’, ***tuju** ‘dew’
Group 3	29	non-high	2.4b	***awa** ‘millet’, ***ine** ‘rice plant’, ***ita** ‘board’, ***kado** ‘corner’, ***kama** ‘sickle’, ***kasa** ‘umbrella’, ***kata** ‘shoulder’, ***keta** ‘rafter’, ***mino** ‘raincoat’, ***miso** ‘bean paste’, ***nae** seedling, ***pada** ‘skin’, ***poka** ‘outside’, ***saja** sheath, ***soba** ‘side’, ***soto** ‘outside’, ***tane** ‘seed’, ***toga** ‘sin’, ***wana** ‘trap’, ***wara** ‘straw’
			2.5b	***ame** ‘rain’, ***ase** ‘sweat’, ***koto** ‘zither’, ***mado** ‘spare time’, ***majo** ‘cocoon’, ***majo** ‘eyebrow’, ***momo** ‘thigh’, ***puna** ‘gibel’, ***tate** ‘upright’
Group 4	10	High	2.4b	***kasu** ‘dregs’, ***kinu** ‘garment’, ***mogi** ‘barley’, ***nomi** ‘chisel’, ***ori** ‘melon’, ***siru** ‘soup’, ***zeni** ‘money’
			2.5b	***ai** ‘indigo’, ***kimi** ‘millet’
			2.4b~2.5b	***joru** ‘night’

Table 2. Classification of Class 2.4–2.5 nouns by their tonal classes in pR and the height of the word-final vowel (Igarashi 2022a).
A capital *U is used where it is unclear whether *o or *u should be reconstructed.

To account for the merger, the CMH postulates a rightward accent shift in Class 2.4–2.5 nouns. For nouns ending in a non-high vowel, the accent shifted from the initial syllable to the second syllable, thereby merging with Class 2.3 (2b–c). By contrast, the rightward shift did not occur in nouns ending in a high vowel, as the final high vowel blocked the merger (2d–e).

(2) Rightward accent shift postulated in the CMH
   a. *jamá 2.3 ‘mountain’ > pR *jamá B (no shift)
   b. *íta 2.4b ‘board’ > pR *itá B
   c. *áme 2.5b ‘rain’ > pR *amé B
   d. *íki 2.4a ‘breath’ > pR *íki C (no shift)
   e. *sáru 2.5a ‘monkey’ > pR *sáru C (no shift)

However, the CMH faces too many exceptions to be maintained in its original form. Table 2 classifies Class 2.4–2.5 nouns by tonal correspondence and height of the word-final vowel into four groups (Groups 1–4). This classification shows not only that a small number of Class 2.4b/2.5b nouns ending in a high vowel occur (Group 4) but, more importantly, that Class 2.4a/2.5a nouns are roughly balanced between those ending in non-high vowels (Group 1) and those ending in high vowels (Group 2). Based on this evidence, the CMH has been rejected by many scholars (Hattori 1978–79; Shimabukuro 2007, 2008; Pellard 2009).

However, it is noteworthy that relatively few Class 2.4b/2.5b nouns end in a high vowel (Group 4). As de Boer (2010: 232) observes, "it is not impossible that the presence of a close vowel in the second syllable played some role in preventing the merger of members of class 2.4/5 with class 2.3." Therefore, it may be premature to dismiss the possibility of a conditioned merger altogether. In fact, the skewed distribution suggests the existence of an additional conditioning factor overlooked in previous studies, which this study identifies as the original word-internal morpheme boundary.

3 The RCMH

As with the CMH, the RCMH assumes the Gairin-type system (1) as the pre-pR tonal system—that is, the system before the merger—and takes a rightward accent shift as the mechanism responsible for the merger. Crucially, any revised hypothesis must also account for exceptions to the CMH (Groups 1 and 4 in Table 2). The RCMH succeeds in explaining most of the nouns in Group 1, whereas Group 4 remains an exception and is considered separately in Section 5. Section 4 presents evidence that many Group 1 nouns were polymorphemic in pJ. This section outlines the RCMH.

The RCMH can be summarized as shown in (3). For clarity, we distinguish between two stages of pR during the relevant period: an earlier stage (pre-pR), which preceded the merger, and a later stage (pR), which followed it.

(3) The RCMH
- a. The split correspondence between Ryukyuan and Japanese results from a conditioned rightward accent shift in pre-pR that merged Class 2.4–2.5 nouns into Class 2.3.
- b. The rightward accent shift was blocked by two factors: a word-final high vowel and reflexes of original morpheme boundaries inherited from pJ.
- c. Not all Class 2.4–2.5 nouns of polymorphemic origin preserved their morpheme boundaries until pre-pR. Some became lexicalized as simplex nouns, losing their original boundary without a phonological trace prior to pre-pR, and thus merged into Class 2.3.
- d. The original polymorphemic structure manifested in pre-pR as either an actual word-internal morpheme boundary or an underlying heavy syllable, with the latter reflecting a lost boundary.

To illustrate the processes in (3a–b), consider the examples in (4).

(4) Examples of mergers under the RCMH
- a. pJ *jama 2.3 'mountain' > pre-pR *jamá > pR *jamá B (no shift)
- b. pJ *ita 2.4 'board' > pre-pR *íta > pR *itá B
- c. pJ *amai 2.5 'rain' > pre-pR *áme > pR *amé B

d. pJ *iki 2.4 ‘breath’ > pre-pR *íki > pR *íki C (no shift)
e. pJ *saru 2.5 ‘monkey’ > pre-pR *sáru > pR *sáru C (no shift)
f. pJ *na-ka 2.4 ‘middle’ > pre-pR *ná-ka > pR *náka C (no shift)
g. pJ *wo-kai 2.5 ‘bucket’ > pre-pR *wó-ke > pR *wóke C (no shift)

The Class 2.4–2.5 nouns in (4b–c) have word-final non-high vowels; they undergo a rightward accent shift and merge with Class 2.3. These belong to Class 2.4b/2.5b, which corresponds to pR Class B. The nouns in (4d–e) end in high vowels, preventing the rightward shift, thereby avoiding mergers. These constitute Class 2.4a/2.5a, corresponding to pR Class C. Finally, the nouns in (4f–g) end in non-high vowels; nonetheless, they resist the rightward shift. This is because they retained a morpheme boundary inherited from pJ to pre-pR, which blocked the shift. Section 4 discusses evidence that these words can be reconstructed with morpheme boundaries in pJ. Consequently, they also belong to Class 2.4a/2.5a, which corresponds to pR Class C.

Thus, by assuming that, in addition to word-final vowel height, word-internal morpheme boundaries blocked the merger, the RCMH succeeds in explaining most of the Group 1 nouns that had previously been treated as exceptions under the CMH.

To make the RCMH viable, additional refinement is required, as formulated in (3c–d). Specifically, morpheme boundaries present at the pJ stage are not assumed to have been consistently maintained in the pre-pR stage. Regarding (3c), some polymorphemic nouns in pJ are hypothesized to have undergone lexicalization and lost their internal boundaries by the pre-pR stage. A representative example is pR *madó B ‘spare time’, which Igarashi (2022a) argues is cognate with OJ *mado*$_1$ 2.5 ‘window’ (the subscript “1” denotes the *kōrui* syllable). It has been analyzed etymologically as a compound of *ma, the covert form (*hifukukei*) of *mai ‘eye’, and *to ‘door’ (Ono et al., eds. 1974: 1196; Martin 1987: 469). The compound consisting of *ma and *to in pJ later developed into *ma-do through voicing of the initial obstruent of the second component, a process conventionally known as *rendaku*. Obstruent voicing is assumed to occur when a nasal and a voiceless obstruent come in succession. When such a nasal cannot be interpreted as part of either root in a compound with *rendaku*, this study reconstructs an intervening *n, typically interpreted as a genitive marker. Accordingly, pJ *ma-n-to is reconstructed. The resulting form, *ma-do, is then assumed to have lexicalized as a simplex noun before the pre-pR stage, yielding pre-pR *mádo ‘spare time’. Because the morpheme boundary had already been lost by that stage and the word ended in a non-high vowel, it subsequently underwent the rightward accent shift, as shown in (5), merging with Class 2.3.

(5) Earlier lexicalization of a pJ polymorphemic noun
pJ *ma-n-to 2.5 ‘window’ > *ma-do > pre-pR *mádo ‘spare time’ > pR *madó B

In the case of pR *madó ‘spare time’, the motivation for lexicalization can be demonstrated relatively clearly. This noun originated as a compound of ‘eye’ and ‘door’, and its original meaning, ‘window’, retained the compositional meaning of ‘eye-door’, or ‘space for seeing’. However, with the subsequent semantic shift from ‘window’ to ‘spare time’, the noun lost its semantic compositionality and, with it, the transparency of its morpheme boundary.

Notably, the loss of morpheme boundary transparency is not a regular sound change. It can occur idiosyncratically in individual words. Consequently, when a sound change interacts with a

morpheme boundary, the sound correspondence across varieties can become irregular. Therefore, the presence of the original morpheme boundaries in some Group 4 nouns does not constitute an exception to the RCMH. What count as genuine exceptions are Group 1 nouns in which morpheme boundaries cannot be reconstructed at the pJ stage.

We now turn to the complication described in (3d). Similar to (3c), (3d) involves the opacification of an originally polymorphemic structure; however, in this case, traces of the original boundary persisted until the pre-pR stage. Specifically, the boundary was reflected as the initial heavy syllable. This development is hypothesized to have resulted from the contraction of words with more than two syllables into disyllabic forms. Because the contraction likely occurred at a relatively early stage, the early loss of morphological transparency complicates etymological analysis.

A typical example is pR *képu 'today', cognate with OJ *ke₁pu*. Many researchers agree that the noun has a polymorphemic origin (Matsumoto 1975; Martin 1987; Unger 2010). As discussed in Section 4, the presence of the non-final *e₁* strongly suggests that the word was originally polymorphemic. However, there is no consensus regarding the identities of the constituent morphemes. Martin (1987: 464) proposes an etymology deriving from a compound of 'this' and 'day'; however, his analysis does not account for the irregular vowel correspondences (Unger 2010). Despite the agreement on the polymorphemic origin, the lack of consensus on etymology suggests that contraction occurred early enough to obscure evidence recoverable through either comparative or internal reconstruction.

As an experiment, let us adopt Unger's (2010) etymology, according to which the noun originates from a compound containing *keru 'be coming' and *pu 'juncture, evening'. The component *keru, in turn, derives from *ki (the adverbial form of *kə- 'to come') combined with the adnominal form of *ar- 'to be'. Following Hattori (1978–79), the adnominal verbal suffix in pJ can be reconstructed as *-o, yielding (pre-)pJ *ki-ar-o pu 'today, the coming evening'. The sporadic loss of the vowel *-o likely produced the contracted form *kiarpu, which subsequently developed into a form with an initial heavy syllable, *kjaapu or *kjappu. This transitional heavy syllable can be hypothesized as the source of the non-final *e₁* in OJ *ke₁pu* 'today' and as the factor that blocked the rightward accent shift at the pre-pR stage. Modern Ryukyuan data are consistent with the reconstruction of *keppu in pre-pR.[1] Therefore, the development in (6) can be posited.

(6) Earlier contraction of a (pre-)pJ polymorphemic noun
(pre-)pJ *ki-ar-o pu 2.4 'today' > *kjappu > pre-pR *képpu > pR *képu C

[1] The reconstruction of pre-pR *képpu 'today', with geminate *pp, involves the controversial correspondence of intervocalic OJ *p* to pR. As Thorpe (1983) notes, intervocalic *p* in OJ generally does not correspond to *p reflexes in Ryukyuan varieties, except in a few cases—mostly adjectives. On this basis, Thorpe (1983) does not reconstruct an intervocalic *p for pR. For words in which some Ryukyuan varieties exhibit intervocalic *p*, he reconstructs *pp ("*Qp" in his notation), suggesting that it originated from emphatic expression, as in pR *koppa- 'hard' (OJ *ko₁pa-* 'hard'). However, words exhibiting intervocalic *p* in Ryukyuan are not limited to adjectives; they also include the cognate of OJ *ke₁pu* 'today'. Specifically, Ashikebu *kɨbu* 'today' (Uwano 1996) shows an intervocalic labial stop, which underwent sporadic voicing (*p > *b*). Therefore, this study proposes that, for Ryukyuan words in which the labial stop corresponds to OJ intervocalic *p*—including the noun for 'today'—*pp should be reconstructed for pre-pR, as in *koppa- 'hard' and *képpu 'today'. This geminate subsequently degeminated in the pR stage, giving rise to a small set of pR words with intervocalic *p, such as *kopa- 'hard' and *képu 'today'.

Section 4 addresses the empirical question of how now-lost morphemic boundaries can be identified and reconstructed.

4 Evidence for Polymorphemic Origin

4.1 Non-Final OJ *e₁* and *o₁*

The first piece of evidence in support of the RCMH is the strikingly regular correspondence between OJ nouns with a non-final *e₁* or *o₁* and pR Class C nouns—a pattern that has gone unnoticed in previous studies.

OJ distinguished between the so-called *kō-rui* and *otsu-rui* syllables, which merged in the transition to Early MJ. This distinction is indicated here by the subscripts "1" and "2." Of particular importance are *e₁* and *o₁*. Non-final *e₁* and *o₁* are exceedingly rare, occurring almost exclusively in the final position, as in OJ *jome₁* 'bride', *kabe₁* 'wall', *ipe₁* 'house', *suso₁* 'skirt hem', *simo₁* 'below', and *kuso₁* 'dung'. However, non-final *e₁* and *o₁* are attested in six Class 2.4–2.5 nouns, and notably, all but one correspond to pR Class C (7).[2]

(7) OJ Class 2.4–2.5 nouns with non-final *e₁* and *o₁*

a.	OJ *mo₁ko₁* 2.5 'bridegroom'	::	pR *móko C 'bridegroom'
b.	OJ *ke₁pu* 2.4 'today'	::	pR *képu C 'today'
c.	OJ *ke₁sa* 2.4 'this morning'	::	pR *késa C 'a short while ago'
d.	OJ *pe₁ra* 2.4 'spatula'	::	pR *péra C 'spatula'
e.	OJ *so₁ra* 2.4 'sky'	::	pR *sóra C 'treetop'
f.	OJ *jo₁ru* 2.4~2.5 'night'	::	pR *jorú B 'night'

To account for the restricted distribution of OJ *e₁*/*o₁*, two types of changes from pJ to OJ have been proposed: lexical contraction (Matsumoto 1975; Ono 1977) and conditioned vowel raising (Frellesvig and Whitman 2008). The former derives OJ *e₁* and *o₁* from the contraction of heteromorphemic vowel sequences, whereas the latter holds that pJ vowels *e and *o merged into *i₁* and *u* in the non-final position but were preserved in the final position. Crucially, both changes presuppose the presence of word-internal morpheme boundaries. In other words, as discussed below, the occurrence of non-final *e₁* and *o₁* can be taken as evidence that the word was originally polymorphemic.

This lexical contraction is illustrated by the OJ nouns in (8), which are widely accepted to be of compound origin.

(8) OJ *e₁* and *o₁* arising from lexical contraction (Ono 1977; Frellesvig and Whitman 2008)

a.	pJ *ua > OJ	*o₁*	pJ *kazu-apai 'number-match'	>	OJ *kazo₁pe₂* 'to count'

[2] Following Arisaka (1932), interpreting *mo₁mo₁nagani* in OJ as containing a Class 2.5 noun meaning 'thigh', cognate with pR *momó B 'thigh', would add another example of a Class 2.4–2.5 noun with non-final *o₁*. However, *mo₁mo₁* is often interpreted as 'hundred, many' (Kurano 1963) and is therefore excluded from (7). If the 'thigh' interpretation is correct, pJ *mo-mo may be reconstructed, and its Class B reflex in pR can be attributed to the opacification of polymorphemic structure by the pre-pR stage. Similarly, since the vowel quality of the first syllable in OJ *keta* 2.4 'rafter' (whether *e₁* or *e₂*) is uncertain, it is also omitted from (7). Martin (1987: 448) analyzes OJ *keta* as a compound. If so, its Class B reflex in pR *ketá 'rafter' implies that the polymorphemic structure had already become opaque before pre-pR.

b.	pJ *uə > OJ *o₁*	pJ *situ-əri ‘native-weave’	>	OJ *sito₁ri* ‘native weave’
c.	pJ *ia > OJ *e₁*	pJ *saki-ari ‘bloom-is’	>	OJ *sake₁ri* ‘is blooming’
d.	pJ *iə > OJ *e₁*	pJ *pi-əki ‘sun-put’	>	OJ *pe₁ki* ‘proper name’

In these pJ compounds, a vowel sequence occurs across the morpheme boundary. According to the lexical contraction account, such heteromorphemic vowel sequences were reduced to yield the OJ vowels transcribed as *e₁* and *o₁*.

The lexical contraction account is valid only for those OJ nouns with non-final *e₁*/*o₁* that can be etymologized as contracted forms. However, with the progress of comparative research between Ryukyuan and Japanese, many scholars now agree that OJ *e₁* and *o₁* reflect not only heteromorphemic vowel sequences but also primary vowels *e and *o of pJ. Researchers adopting this view posit a sound change in OJ, first proposed by Hattori (1978–79) and later termed Mid Vowel Raising (MVR) (Frellesvig and Whitman 2008), which affected pJ *e and *o (9).

(9) MVR
pJ *e > OJ *i₁* pJ *o > OJ *u*

MVR raised pJ *e and *o to OJ *i₁* and *u*, respectively, so that pJ *e and *i merged as OJ *i₁*, and pJ *o and *u merged as OJ *u* (Hattori 1978–79; Frellesvig and Whitman 2008; Pellard 2013). Importantly, MVR did not apply to pR; pJ *e and *o are preserved as pR *e and *o (Hattori 1978–79; Pellard 2013). Thus, MVR can be regarded as an innovation that defines the Japanese branch.

However, not all pJ *e and *o were raised in OJ. Many OJ words retain unraised *e₁* (< *e) and *o₁* (< *o). Although their frequent occurrence in the word-final position suggests that word boundaries conditioned the application of MVR, there are also rare examples of word-internal unraised *e₁*/*o₁* that cannot easily be explained as contracted forms, such as OJ *so₁de* ‘sleeve’ and *me₁pi* ‘niece’. Consequently, word boundaries cannot be the sole blocking factor for MVR.

To account for these exceptions, Hattori (1978–79) reconstructed a contrastive vowel length in pJ, *e and *o vs. *ee and *oo, and proposed the changes pJ *ee, *oo > OJ *e₁, o₁*. In this view, short vowels undergo MVR, whereas long vowels resist it. However, this account faces serious problems: typologically, long vowels are more likely than short vowels to undergo raising (Pellard 2013), and more importantly, there is no independent evidence for reconstructing long vowels in pJ (Frellesvig and Whitman 2008).

Frellesvig and Whitman (2008) offered an alternative: domain-final vowels fall outside the scope of MVR, with the relevant domain (word, morpheme, or root) differing across dialects. This conditioned MVR account is motivated by the existence of OJ doublets, in which one form shows the raising of *e/*o while the other does not—for example, OJ *jo₁ri* ~ *juri* ‘from’ and OJ *mo₁ko₁* ~ *muko₁* ‘bridegroom’. Of particular note are the former variants, where *e₁*/*o₁* occur in non-word-final position but do not undergo MVR. Frellesvig and Whitman interpret such differences as dialectal variation: they reconstruct a morpheme boundary between the first and second syllables in forms such as pJ *jo-ri ‘from’ and pJ *mo-ko ‘bridegroom’, and argue that the transparency of morpheme boundaries was preserved in some dialects but lost in others. Once transparency was lost, first-syllable *e/*o ceased to be domain-final and thus became subject to MVR, producing the latter variants *juri* and *muko₁*. Indeed, OJ *so₁de* ‘sleeve’ and *me₁pi* ‘niece’ are neatly accounted for by this analysis: in both cases, the unraised non-final *e₁*/*o₁* is consistent with established etymological analyses that posit a morpheme boundary immediately following the vowel—*so*

'cloth' plus *te* 'hand' (Martin 1987: 529) and *me₁* 'female' plus *pi* 'niece, nephew' (cf. OJ *wopi₂* 'nephew' and *wo* 'male') (Martin 1987: 475). Vowels directly preceding a transparent morpheme boundary were exempt from MVR; therefore, these forms precisely fit the predictions of Frellesvig and Whitman's analysis.

Thus, by interpreting the origin of unraised non-final *e₁*/*o₁* in OJ either as the result of lexical contraction or as the non-application of MVR to domain-final *e and *o, we may conclude that OJ nouns with non-final *e₁*/*o₁* were polymorphemic at the pJ stage. Therefore, the RCMH explains why most Class 2.4/2.5 nouns with these vowels are reflected as pR Class C.

In the case of OJ *mo₁ko₁* 'bridegroom' (7a), the morpheme boundary is assumed to have persisted until the pre-pR stage. Frellesvig and Whitman (2008) analyze it as a compound with pJ *ko 'child, man' (> OJ *ko₁* 'id.'), although the first component, *mo, remains without etymology. This boundary likely survived beyond the pR stage. Alongside regular reflexes of pR *móko—such as Nakijin *muɸu* (Nakasone 1983), Tarama *muku* (Tokuyama and Celik 2020), and Yonaguni *mugu* (Yonaguni Board of Education 2019)—we also find forms like Asama *mukk ʔwa* (Uwano 2017) and Yoron *muukwaa* (Kiku and Takahashi 2005). The final syllable in these latter forms is not a direct reflex of pJ *ko, but of pR *kua 'child' (Igarashi 2022a). This suggests that the final syllable of pR *móko was reanalyzed as *kua, indicating that speakers continued to perceive the word as polymorphemic, even after the diversification of Ryukyuan. Thus, we may convincingly assume that although *mo eventually became a cranberry morpheme, the morpheme boundary persisted, at least to the pre-pR stage, thereby blocking the rightward accent shift, as shown in (10).

(10) pJ *mo-ko 2.5 'bridegroom' > pre-pR *mó-ko > pR *móko C

As discussed in Section 3, early lexical contraction can be assumed for OJ *ke₁pu* 'today' (7b). A similar process can be posited for OJ *ke₁sa* 'this morning' (7c), which is cognate with pR *kesa 'short while ago' (Igarashi 2022a). Unger (2010) analyzes OJ *ke₁sa* as a compound consisting of *asa* 'morning' as the second component and the same first component as in *ke₁pu*. Following Unger's etymology, this paper reconstructs pre-pJ *ki-ar-o asa 'this morning, the coming morning'. This phrase is thought to have undergone vowel loss, yielding forms such as *kjaasa or *kjassa. As with pR *kepu 'this morning' discussed in Section 3, we reconstruct pre-pR *kessa 'short while ago' with an initial heavy syllable—specifically, a syllable closed by a coda consonant—and regard this heavy syllable as blocking the rightward accent shift, as shown in (11).

(11) a. (pre-)pJ *ki-ar-o pu 2.4 'today' > pre-pR *képpu
> pR *képu C
b. (pre-)pJ *ki-ar-o asa 2.4 'this morning' > pre-pR *késsa 'short while ago'
> pR *késa C

Reconstructing an original morpheme boundary in OJ *pe₁ra* 'spatula' (7d) and *so₁ra* 'sky' (7e) is more challenging. One hypothesis identifies them as possible Korean loans, based on a comparison with Middle Korean *pyet* 'moldboard' and *hanolh* 'sky', respectively (Whitman 1985). If these Korean words had been borrowed into OJ after MVR, as forms with *e₁* and *o₁* on the OJ side, the presence of unraised vowels would be straightforwardly explained without recourse to a morpheme boundary. However, the fact that their pR forms can be reconstructed with mid

vowels—as in *péra and *sóra based on regular correspondences among the Ryukyuan languages—casts doubt on the post-MVR borrowing hypothesis, since pR branched off before MVR in OJ. If the borrowing hypothesis is maintained, it must be assumed that these words entered pJ before MVR in forms with vowels not subject to raising, such as *piara and *saura. If the diphthong in the initial syllable was interpreted as producing a heavy syllable and this heavy syllable persisted into pre-pR, it would have blocked the rightward accent shift and yielded reflexes in pR Class C. In this case, the RCMH should be revised to include heavy syllables introduced by borrowing as a factor capable of blocking mergers.

However, this study does not adopt the borrowing hypothesis. Instead, it considers the possibility that OJ *pe₁ra* 'spatula' and *so₁ra* 'sky' were of polymorphemic origin. OJ *pe₁ra* can plausibly be analyzed as containing a root related to the OJ lower bigrade (*shimonidan*) verb *pi₂we-* 'to scrape'. This verb is one of only four exceptional words exhibiting a non-final *i₂* (< pJ *əi, *ui or *oi); the others are OJ *mi₂na* 'all', *ki₂si* 'shore', and *ki₂ri* 'fog' (Frellesvig and Whitman 2004, 2008). OJ *mi₂na* 'all' has previously been analyzed as originally polymorphemic (Hattori 1978–79; Frellesvig and Whitman 2004: 283). Frellesvig and Whitman (2008: 18) suggest that OJ *ki₂si* 'shore' and *ki₂ri* 'fog' are rare cases of sporadic "umlaut" triggered by *i in the following syllable, going back to earlier forms: *kusi and *kuri.

However, in the case of OJ *pi₂we-*, such a sporadic umlaut cannot be assumed. Since the stem-final vowel of lower bigrade verbs is reconstructed as pJ *-ai (Hattori 1978–79), the vowel in the second syllable of *pi₂we-* could not be *i. Therefore, an alternative explanation is to posit that the vowel in the initial syllable of OJ *pi₂we-* was originally *i, which was later labialized through sporadic assimilation to the following *w. Accordingly, the pJ form of OJ *pi₂we-* can be reconstructed as *piw-ai-. Given the attested pairs OJ *mure-* 'to flock' ~ *mura* 'village' (< pJ *mur-ai ~ *mur-a) and OJ *kare-* 'to wither' ~ *kara* 'husk, empty' (< pJ *kar-ai ~ *kar-a), it is reasonable to posit a similar pairing of *piw-ai- with the noun *piw-a. From this, OJ *pe₁ra* 'spatula' may be interpreted as a polymorphemic noun containing this root.

The origin of the final syllable *-ra* in OJ *pe₁ra* remains uncertain. However, it may be related to the diminutive suffix *-ra*, as in OJ *wagi₁pe₁-ra* 'my.house-diminutive'. If so, a pJ form *piw-a-ra 'spatula' may be reconstructed. Through the sporadic loss of *w, the *piwara sequence may have contracted to *piara. This, in turn, may have resulted in glide formation and compensatory lengthening, producing *pjaara. Such a development would yield the pre-pR form *péera with a heavy initial syllable, which, in turn, would block the rightward accent shift, resulting in pR Class C, as shown in (12).

(12) pJ *piw-ar-a 2.4 'spatula' > *pjaara > pre-pR *péera > pR *péra C

OJ *so₁ra* is likely to have contained a root related to OJ *ura* ~ *ure* 'tip of a branch'. This interpretation is supported by the semantic range of *sora* itself. Notably, *sora* is attested in Japanese literature not only with the meaning 'sky' but also with reference to the upper part of an object, such as 'roof', 'ceiling', 'upper seat', and 'treetop' (NKD 2001: 507). In many modern Japanese varieties as well, *sora* denotes 'upper place, high place', 'mountaintop', or 'ceiling' (SKJ 1989: 1340). In the modern Ryukyuan languages, reflexes of pR *sora, which Igarashi (2022a) regards as cognate with OJ *so₁ra*, mean 'treetop', while 'sky' is instead expressed by reflexes of pR *ten 'sky'. Therefore, it is reasonable to assume that the pJ word reflected by OJ *so₁ra* originally

referred to something like ‘the uppermost part or space of an object’ rather than specifically to ‘sky’.

Cognates of OJ *ura* survive in modern Japanese varieties in the forms *ura* or *ora*, meaning ‘tip, end’, ‘tip of a leaf’, ‘tip of a stem’, ‘treetop’, ‘mountain top’, ‘upstream of a river’, or ‘the mountain peak at a river’s headwaters’ (SKJ 1989: 299), suggesting that the word originally denoted ‘the extremity of an object that typically lies at the top’. Thus, given their formal and semantic similarity, OJ *so₁ra* ‘sky’ can reasonably be assumed to have consisted of *ura/ora* as its second component.

The initial component is more difficult to etymologize, although one possibility is *su (> OJ *su* ‘plain, natural, unadorned’ [Martin 1987: 531], as in OJ *sunapo* ‘straight, naïve’). If the initial component was *su, then, to yield the final form *sora*, the initial vowel of the second component must have been a non-high vowel. Therefore, the second component should be reconstructed as *ora rather than *ura, allowing us to posit a form such as *su-ora. Its meaning can be reconstructed as ‘the empty or open space at the top’. This heteromorphemic vowel sequence *uo would then have undergone lexical contraction, producing pre-pR *sóora ‘treetop’, where the initial heavy syllable blocked the rightward accent shift, resulting in pR *sóra, as shown in (13).

(13) pJ *su-ora 2.4 > pre-pR *sóora ‘treetop’ > pR *sóra C

The RCMH accounts for the fact that, except for one case, all six Class 2.4–2.5 nouns containing a non-final OJ *e₁* or *o₁* are reflected as Class C in pR—the sole exception being OJ *jo₁ru* 2.4~2.5 ‘night’ (7f). This word is assumed to have been originally polymorphemic, containing a cognate of OJ *jo₁* ‘night’ as its first component (Martin 1987: 557). At first glance, its Class B reflex in pR can be explained by assuming that it underwent early lexicalization, losing its morpheme boundary before the pre-pR stage. However, its word-final high vowel should have blocked the rightward accent shift, yielding a Class C reflex in pR. Accordingly, OJ *jo₁ru* 2.4~2.5 must be regarded as an exception to the RCMH, which will be addressed separately in Section 5.

In summary, OJ nouns exhibiting rare non-final *e₁* or *o₁* can be considered polymorphemic at the pJ stage, and the fact that Class 2.4–2.5 words with these vowels belong to pR Class C can be accounted for by the RCMH. This hypothesis posits that word-internal morpheme boundaries, or the heavy initial syllables arising from them, prevent rightward accent shifts at the pre-pR stage. The plausibility of the individual etymologies proposed in this section may vary; however, their validity should be evaluated independently of the RCMH. The correlation of non-final *e₁/o₁* with pR Class C appears to be too systematic to be accidental; therefore, it is reasonable to assume that OJ nouns with non-final *e₁/o₁* ultimately reflect polymorphemic origins.

4.2 Violation of Arisaka’s Law

The second piece of evidence comes from apparent violations of Arisaka’s Law (Arisaka 1932), which formulates phonotactic constraints on vowel co-occurrence within OJ morphemes. Arisaka’s law states that in OJ, *o₂* does not co-occur with *a*, *o₁*, or *u* within a single morpheme. However, several Class 2.4–2.5 nouns violate this constraint, as shown in (14).

(14) Class 2.4–2.5 nouns that violate Arisaka’s Law
  a. OJ *ato₁ ~ ato₂* 2.4a ‘trace’ :: *pR *áto C ‘trace’

b. OJ *ko₂we* 2.5a ‘voice’ :: *pR *kóe C ‘voice’
c. OJ *to₂ga* 2.4b ‘sin’ :: *pR *togá B ‘sin’
d. OJ *poka* 2.4b ‘outside’ :: *pR *poká B ‘outside’
e. MJ *soba* 2.4b ‘side’ :: *pR *sobá B ‘side’

The occurrence of vowel combinations that are otherwise prohibited within a morpheme is most plausibly explained by assuming that these words were originally polymorphemic. Accordingly, this study argues that the four items in (14) are of polymorphemic origin.

Phonograms of the second vowel in OJ *ato* 2.4a ‘trace’ (14a) fluctuate between *o₁* and *o₂* (JJ 1967: 32). OJ *ato* has generally been analyzed as an original compound, whose second component can be etymologized as *to₂* ‘trace’ (Ono et al., eds. 1974: 39; Matsumoto 1975; Martin 1987). Assuming that this analysis is valid, the original form must have been *ato₂*, which violates Arisaka’s Law. The variant *ato₁* likely represents a dialectal form in which the transparency of the original morpheme boundary was lost, leading to assimilation of the second vowel to the first, thereby restoring conformity with Arisaka’s law.

Martin (1987: 387) analyzes this word as consisting of OJ *asi* 2.3 ‘foot’ and *to₂* ‘trace’. The first component may alternatively be a bound root *a-* ‘foot’, attested in forms such as OJ *abumi₁* ‘stirrup’ and *ajupi₁* ‘cord attached to one’s skirt’. The pJ form of ‘trace’ can therefore be reconstructed as pJ *asi-tə or pJ *a-tə. The sporadic vowel loss in the second syllable of the former could have produced a lexical contraction, yielding pre-pR *átto with an initial heavy syllable that would block the rightward accent shift. Alternatively, in the latter case, the morpheme boundary itself may have persisted until the pre-pR stage, thus preventing the shift. These two possible scenarios are illustrated in (15).

(15) a. pJ *asi-tə 2.4 ‘trace’ > pre-pR *átto > pR *áto C
b. pJ *a-tə 2.4 ‘trace’ > pre-pR *á-to > pR *áto C

OJ *ko₂we* ‘voice’ (14b) is not a straightforward case of violation of Arisaka’s Law. In Japanese, this noun alternates between an exposed form (*roshutsu-kei*) *kowe* and a covert form (*hifuku-kei*) *kowa-*, from which pJ *kəwai can be reconstructed. The reconstruction of pJ *kəwai predicts an OJ covert form *ko₂wa-, which would indeed violate Arisaka’s Law. However, this form is not attested in OJ; the covert form *kowa-* is found only from MJ onward.

Matsumoto (1975) and Frellesvig and Whitman (2008) regard the alternation between *kowe* and *kowa-* as a secondary analogical development that occurred in MJ. They further relate OJ *ko₂we* to the onomatopoetic expression *ko₂woro₂ko₂woro₂* ‘rumbling noise’ in OJ, and based on this, Frellesvig and Whitman reconstruct pJ *kəwər > *kəwəi ‘voice’. If their reconstruction is correct, OJ *ko₂we* does not exemplify a violation of Arisaka’s Law and thus provides no evidence for a polymorphemic origin. At the same time, this interpretation rests on a broader reconstruction that remains open to discussion. Frellesvig and Whitman’s analysis is embedded in their seven-vowel hypothesis, in which *əi is reconstructed for forms showing the *o₂* ~ *e₂* alternation, in contrast to *ɨi, which is reconstructed for forms showing the *o₂* ~ *i₂* alternation. Pellard (2013) notes that this hypothesis is based mainly on OJ-internal evidence and does not receive clear independent support from Ryukyuan, Eastern Old Japanese, or Hachijo. It may thus be possible to account for OJ *ko₂we* as a reflex of pJ *kəwəi, but this analysis does not by itself rule out an alternative morphological explanation.

For the purposes of the present analysis, OJ *ko₂we* is provisionally treated as a violation of Arisaka's Law and as having a polymorphemic origin, although its precise etymology remains uncertain. As a working hypothesis, we may tentatively reconstruct pJ *kətə-wai 'tip of word, voice'. The first component, *kətə, is reflected in OJ *ko₁to₁* 'thing, word', and the second, *wai, may be cognate with the second component of OJ *suwe* 'tip, end, edge'. Sporadic loss of intervocalic *t would have produced the lexically contracted form, *kəəwai, with a long vowel in the initial syllable. The regular sound changes pJ *ə > *o and *ai > *e would yield pre-pR *kóowe, where the initial heavy syllable blocked the rightward accent shift, resulting in pR *kówe, as shown in (16).

(16) pJ *kətə-wai 2.5 'voice' > *kəə-wai > pre-pR *kóowe > pR *kóe C

OJ *to₂ga* 'sin' (14c) clearly violates Arisaka's Law, indicating that it was polymorphemic in pJ, although its precise etymology remains uncertain. Even without a detailed analysis, pJ *tə-ga (< *təN-ka, *tə-Nka, or *tə-n-ka) can be reconstructed. Because this noun belongs to pR Class B, the original morpheme boundary is assumed to have been lost before the pre-pR stage, thereby allowing the rightward accent shift, as shown in (17).

(17) pJ *tə-ga 2.4 'sin' > pre-pR *tóga > pR *togá B

The quality of the vowel in the first syllable of OJ *poka* 'outside' (14d) and MJ *soba* 'side' (14e)—whether *o₁* or *o₂*—remains uncertain. Nevertheless, the morpheme boundary between the two syllables must be reconstructed for pJ in both cases. If the vowel was *o₁* (< pJ *o), it would have been raised to *u* in a monomorphemic form by MVR. Therefore, the absence of the expected †puka and †suba points to the presence of a morpheme boundary. Conversely, if the vowel was *o₂* (< pJ *ə), the form would have violated Arisaka's Law, again suggesting a polymorphemic origin.

Accordingly, the pJ form can be reconstructed as either *po-ka or *pə-ka for (14d), and as *so-ba or *sə-ba for (14e). The second component of MJ *soba* 'side' (14e) is likely *ba 'place', which Vovin (2014) interprets as a contracted form of *nipa* 'place for conducting affairs'. Consequently, pJ *so-nipa or *sə-nipa can ultimately be reconstructed. Because the nouns for 'outside' and 'side' both belong to pR Class B, they must have lost their internal boundary before the pre-pR stage, thereby undergoing the rightward accent shift, as illustrated in (18) and (19).

(18) a. pJ *po-ka 2.4 'outside' > pre-pR *póka > pR *poká B
b. pJ *pə-ka 2.4 'outside' > pre-pR *póka > pR *poká B

(19) a. pJ *so-nipa 2.4 'side' > *so-ba > pre-pR *sóba > pR *sobá B
b. pJ *sə-nipa 2.4 'side' > *sə-ba > pre-pR *sóba > pR *sobá B

Overall, these apparent violations of Arisaka's Law provide further empirical support for the RCMH, reinforcing the view that preserved morpheme boundaries—or their relics—played a decisive role in conditioning the tonal merger.

4.3 Unexpected Initial Heavy Syllables in Modern Varieties

The third piece of evidence concerns unexpected heavy syllables in certain modern Japanese varieties. Here, *unexpected* refers to instances in which, relative to regular correspondences, a noun exhibits an unpredicted syllable weight that cannot be accounted for by any known variety-internal conditioning. Some apparent cases may ultimately reflect sporadic changes in individual varieties. However, among Class 2.4–2.5 nouns, such unexpected heaviness occurs almost exclusively in Class 2.4a/2.5a (pR Class C). Within the framework of the RCMH, this skewed distribution is best interpreted as relict morphology—residual traces of the original word-internal morpheme boundaries, manifesting as increased syllable weight.

In Shiramine Japanese, spoken in Ishikawa Prefecture, six disyllabic nouns exhibit an unexpected heavy first syllable: *ooke* 2.5 'bucket', *naabe* 2.5 'pot', *kaŋge* 2.5 'shadow', *noodo* 'throat', *soore* 'sleigh', and *njoozi* 'rainbow' (Nitta 2010). The first three belong to Class 2.4–2.5, and notably, all correspond to pR Class C—*wóke, *nábe, and káge.

The view that Shiramine's unexpected heavy syllables are remnants of their polymorphemic origin can be supported by *noodo* 'throat' in this variety. Crucially, the form cognate to Modern Japanese *nodo* 'throat' is attested in OJ as a trisyllabic compound, *no₂mi₁to₂* ~ *no₂mi₁do₁*. The variation between *to₂* and *do₁* is not straightforward to interpret; however, if the established etymology—that this OJ noun consists of OJ *no₂mi₁* 'drinking' and *to₁* 'door' (Martin 1987: 499)—is correct, the final vowel should be *o₁*. The variant *no₂mi₁to₂*, which conforms to Arisaka's Law, suggests that lexicalization was already underway at the OJ stage. The pJ form can thus be reconstructed as *nəmi-to. Its reflex *nodo* in multiple modern Japanese varieties, including Standard Japanese, is evidently a contracted outcome. The sporadic loss of the second-syllable vowel would have yielded an intermediate *nomto, subsequently giving Modern Japanese *nodo*. Indeed, an intermediate stage with a nasal coda, *nondo* 'throat', is attested in Japanese sources up to the nineteenth century (NKD 2001: 828, 894). Shiramine *noodo* likely developed from *nomto* through voicing assimilation and lenition-denasalization of *m; that is, *nomto > *nomdo > *noodo*.

Importantly, the tonal correspondence of *nodo* among modern Japanese varieties is equivalent to Class 2.4 (Martin 1987: 499). Furthermore, pR *nódo C 'throat' can be reconstructed, as evidenced by Ikema *nudu* (Nakama et al. 2022), Ishigaki *nudu* (Miyagi 2003), and Yonaguni *nudu* (Yonaguni Board of Education 2019). Thus, the Class C reflex in pR can be accounted for under the RCMH by assuming that the initial heavy syllable blocked the rightward accent shift in the pre-pR stage, as illustrated in (20).

(20) pJ *nəmi-to 'throat' > *nomto > pre-pR *nóndo > pR *nódo C

An unexpected heavy syllable in Shiramine *naabe* 'pot' can also be explained by positing an original morpheme boundary. This noun is cognate with OJ *nabe₂* 2.5, which is generally analyzed as consisting of OJ *na* 'side-dish, greens' and *pe₂* 'pot' (JJ 1967: 531; Matsumoto 1975: 99; Martin 1987: 490). Accordingly, the pJ form of 'pot' can be reconstructed as *na-n-pai, with *n functioning as a genitive marker and giving rise to a subsequent form with *rendaku*, *na-be. The long vowels attested in the Shiramine form can be attributed to Foot Binarity (McCarthy and Prince 2001), which requires prosodic feet to be binary: compounding would have blocked the formation of prosodic feet across morpheme boundaries, thereby triggering vowel lengthening. Under this account, monomoraic morphemes within compounds surface as bi-moraic forms; for instance, *na-

be could yield *naa-bee, which was later lexicalized with the shortening of the final syllable, resulting in *naabe*.

In the pR reflex, by contrast, no intermediate vowel lengthening needs to be assumed. It is sufficient to posit that the morpheme boundary was retained until the pre-pR stage. Consequently, pre-pR *ná-be can be reconstructed, with the preserved boundary blocking the rightward accent shift, yielding pR *nábe Class C, as shown in (21).

(21) pJ *na-n-pai 2.5 'pot' > pre-pR *ná-be > pR *nábe C

An unexpected long vowel in the initial syllable of 'bucket' is attested not only in Shiramine but also in other Japanese varieties—in the form *ooke*—in the Shimabara Peninsula of Nagasaki Prefecture (SDJKS 1975: 95) as well as in Gifu and Mie Prefectures (NKD 2001: 1085). The view that this word has a polymorphemic origin is supported by established etymological analyses (JJ 1967: 831; Martin 1987: 505). This noun is cognate with OJ *woke₂* 2.5 'bucket', which is generally analyzed as consisting of OJ *wo* 'hemp' and *ke₂* 'container'. Accordingly, the pJ form can be reconstructed as *wo-kai 'bucket'. The long vowels observed in these varieties can be accounted for by assuming the same type of vowel lengthening as that proposed for 'pot'. However, no intermediate lengthening needs to be assumed for the pR reflex. pJ *wo-kai developed into pre-pR *wó-ke, with the retained morpheme boundary blocking the rightward accent shift, yielding pR *wóke C, as shown in (22).

(22) pJ *wo-kai 2.5 'bucket' > pre-pR *wó-ke > pR *wóke C

An unexpected nasal coda in the initial syllable of cognates of OJ *kage₂* 2.5 'shadow' is attested not only in Shiramine Japanese but also in Yamagata Japanese, as in *kaŋŋe* 'backside' (SKJ 1989: 518). Similarly, Shimabara Japanese exhibits an unexpected long vowel in this noun, as in *kaage* 'shade' (SDJKS 1975: 32). Under the RCMH, OJ *kage₂* is assumed to have been polymorphemic. However, its etymology remains highly controversial. Martin (1987: 432) proposed that OJ *kage₂* was originally polymorphemic, derived from *ka ni ka, whose first and last components are related to OJ *ka* 'day', as seen in *putuka* 'two days', *nanuka* 'seven days', *ika* 'fifty days', and *ikuka* 'how many days'. Yet Vovin (2020: 383–393) reconstructs the proto-form corresponding to OJ *ka* as *uka, not *ka, rendering Martin's proposal difficult to maintain.

OJ *kage₂* is polysemous, denoting not only 'shadow' but also its opposite, 'light', along with related senses such as 'shade', 'figure', 'reflection', and 'appearance' (JJ 1967: 182). Its meanings range from tangible phenomena, such as sunlight or the dim area cast by an object, to more abstract notions, such as a remembered figure or an image arising from imagination. In short, OJ *kage₂* encompassed both physical and perceptual manifestations of light and darkness. Two OJ verbs, *kagajo₁p-* 'to twinkle' and *kagi₁r-* 'to glimmer', likewise express the alternation between light and darkness and are formally reminiscent of OJ *kage₂*. Both verbs are clearly compounds: the former may incorporate OJ *kajo₁p-* 'to go back and forth', and the latter OJ *ki₁r-* 'to make fire by friction'. It follows that the shared component among *kage₂*, *kagajo₁p-*, and *kagi₁r-* should be reconstructed as pJ *ka or *kaN 'light and darkness', rather than *kaNk (> *kag*), since *k belongs to the second component.

The identity of the second component in OJ *kage₂* remains uncertain, although it is likely derived from a form going back to pJ *kai. A plausible candidate is OJ *ke₂* ‘sign, appearance’. Accordingly, the pJ precursor of OJ *kage₂* can be reconstructed as a polymorphemic form, either *kaN-kai or *ka-n-kai. The former contains an earlier contracted form of *kanV or *kamV as the first root, and the latter incorporates a genitive marker *n between the two roots. The former is particularly noteworthy, as it corresponds to OJ *kamu-* ‘divine, sacred’, the covert form of *kami₂* ‘god’. Therefore, it is conceivable, albeit highly speculative, that OJ *kage₂* ultimately reflects *kamu-kai ‘sign of light and darkness’. The root *kamu- can then be interpreted as originally denoting the unified concept of light and darkness—a perception of the two as inseparable aspects of a single phenomenon—which was later extended metaphorically to denote what is ‘divine’ or ‘sacred’. This interpretation accords with the fact that the exposed form pJ *kamui (> OJ *kami₂*) denotes ‘god’ in the Japanese religious tradition—an ambivalent being embodying both beneficent and malevolent aspects.

Regardless of the validity of this etymology, the preceding discussion suggests that OJ *kage₂* was originally polymorphemic. After an early contraction that obscured its internal structure, the nasal element may have persisted into the pre-pR stage as *kánge ‘shadow’, with the initial heavy syllable blocking the rightward accent shift, yielding pR *káge, as shown in (23).

(23) pJ *kaN-kai 2.5 (< *kamu-kai?) ‘shadow’ > pre-pR *kánge > pR *káge C

Shimabara Japanese, already discussed several times in this subsection, exhibits a relatively large number of disyllabic nouns with a long vowel in the initial syllable. However, most of these result from synchronic phonological rules rather than from remnants of polymorphemic structure inherited from pJ. An analysis of SDJKS (1975), a valuable collection of early lexical data based on surveys conducted across the Shimabara Peninsula before 1931, reveals two recurrent patterns. The first, shared by several varieties in southwestern Kyushu (Okumura 1958; Kamimura 1983; Matsumori 2017), is that a word-final high vowel triggers lengthening of the vowel in the penultimate syllable, as in *maatsu* ‘pine’. The second is that when both the final and penultimate vowels are /a/, the penultimate vowel lengthens, as in *saaka* ‘slope’ and *kaasa* ‘umbrella’. Nevertheless, a few nouns display a long vowel that cannot be attributed to either of these patterns; such lengthening must be regarded as underlyingly specified.

Among these cases are several Class 2.4–2.5 nouns with an unexpected long vowel in the initial syllable—namely, *ooke* 2.5 ‘bucket’ and *kaage* 2.5 ‘shadow’, already discussed, as well as *heera* 2.4 ‘spatula’ (SDJKS 1975: 114) and *miino* 2.4 ‘raincoat’ (SDJKS 1975: 111). Shimabara *heera* ‘spatula’ corresponds to OJ *pe₁ra* 2.4 and pR *péra C. As discussed in Section 4.1, this word can be reconstructed as the polymorphemic pJ *piwa-ra. Thus, the form *heera*, similar to pre-pR *peera, is interpreted as retaining traces of its original polymorphemic structure. Shimabara *miino* ‘raincoat’ corresponds to OJ *mi₁no₂* 2.4 and pR *minó B. Martin (1987: 480) suggests that this noun derives from *mina ‘water’ plus OJ *so₂* ‘cloth’. Regardless of the validity of this particular etymology, *miino* can be regarded as originally polymorphemic within the RCMH framework, which interprets unexpected heavy syllables as vestiges of earlier morpheme boundaries. Since pR *minó belongs to Class B, the morpheme boundary must have been lost by the pre-pR stage, while Shimabara *miino* preserves its trace in the form of a long vowel.

Oita Japanese has *suura* 2.4 ‘falsehood’, with an unexpected long vowel in the initial syllable (SKJ 1989: 1340). This word can reasonably be regarded as cognate with OJ *so₁ra* 2.4 ‘sky’. Notably, OJ *so₁ra* also carries the figurative meaning ‘a state of mental emptiness’ (JJ 1967: 407), and Standard Japanese *sora* likewise extends to the meaning ‘falsehood’. These meanings are plausibly semantic extensions of the notion of ‘emptiness’ implied in the proto-form reconstructed in this study as *su-ora ‘the empty or open space at the top’, representing a metaphorical shift from spatial emptiness to psychological emptiness and ultimately to insubstantiality or falsehood. This, in turn, supports the interpretation that the first component is *su ‘plain, natural, unadorned’.

The Oita form *suura*, similar to pre-pR *sóora, is therefore interpreted as retaining traces of its original polymorphemic structure. Since the word for ‘sky’ in Oita Japanese is *sora* or *tɕen* (Hirayama et al. eds. 1992–93: 2850), the form *suura* may be understood as an archaic reflex inherited from pJ, which survived only in this derived meaning.

As described above, the polymorphic origin of the six words belonging to Class 2.4a/2.5a is also supported by unexpected heavy syllables in Modern Japanese varieties.

4.4 Etymological Analysis

To substantiate the RCMH, it is necessary to provide evidence that the fifteen nouns in Group 1 of Table 1 are derived from polymorphemic sources. In the preceding subsections, independent arguments for a polymorphemic origin have already been presented for nine of these fifteen items. The remaining task is to provide similar evidence for the six nouns that have not yet been discussed.

Three of these nouns have been independently analyzed as polymorphemic by multiple scholars, and such analyses are fully compatible with the assumptions of the RCMH. pR *jádo C ‘shelter’ is cognate with OJ *jado₁* 2.4 ‘shelter’, which has been analyzed as consisting of OJ *ja* ‘house’ and OJ *to₁* ‘door’ (Ono et al., eds. 1974: 1311; Martin 1987: 570). Matsumoto (1975: 121, 135) instead interprets the second component as OJ *to₁* ‘place’. Therefore, the pJ form can be reconstructed as pJ *ja-n-to, with the morpheme boundary preserved until the pre-pR stage, thereby blocking the rightward accent shift, as shown in (24).

(24) pJ *ja-n-to 2.4 ‘shelter’ > pre-pR *já-do > pR *jádo C

pR *náka C ‘middle’ is cognate with OJ *naka* 2.4 ‘middle’, which has been etymologized as a compound of OJ *na-* ‘inside’ and OJ *-ka* ‘place’ (Ono et al., eds. 1974: 940; Martin 1987: 491). OJ *na-* ‘inside’ in *naka* is attested only in proper nouns, where the phonogram 中 ‘inside’ represents the syllable /na/, allowing the reconstruction of the bound form OJ *na-* ‘inside’ (JJ 1967: 512). The second component *-ka* is likewise a bound morpheme, attested, for example, in OJ *sumi₁ka* ‘residence’ and OJ *woka* ‘hill’. Thus, the pJ form can be reconstructed as pJ *na-ka, whose morpheme boundary persisted until the pre-pR stage and prevented the rightward accent shift, as shown in (25).

(25) pJ *na-ka 2.4 ‘middle’ > pre-pR *ná-ka > pR *náka C

pR *máe C ‘front’ corresponds to OJ *mape₁* 2.5 ‘front’, analyzed as consisting of OJ *ma-* ‘eye’—the covert form of *me₂* ‘eye’—and OJ *pe₁* ‘vicinity’ (Ono et al., eds. 1974: 1201; Martin 1987: 469). Accordingly, the pJ form can be reconstructed as pJ *ma-pe, with the morpheme

boundary again retained until the pre-pR stage, thereby blocking the rightward accent shift, as shown in (26).

(26) pJ *ma-pe 2.5 ‘front’ > pre-pR *má-pe > pR *máe C

The etymology of pR *púne C ‘boat’ is more controversial. It is cognate with OJ *pune* 2.4 ‘boat’, whose covert form is *puna*. Whitman (2012) regards the word as polymorphemic, containing *nai ‘root’ (> OJ *ne* ‘root’) as the second component. Whitman (2012), drawing on comparison with Korean, further proposes that the first component is cognate with OJ *pe₂* ‘prow’. The alternation between *u* and *e₂* in OJ is highly unusual, prompting Whitman (2012) to propose the following account. He reconstructs pJ *pəi ‘boat’, whose vowel later developed into *u* in OJ *pune* and *e* in OJ *pe₂* ‘prow’. In his analysis, when combined with *nai ‘root’, *i of *pəi was lost first, yielding the covert form *pə, after which *ə rounded to *o following the labial consonant, producing *po-nai. Once the morphemic boundary became opaque, *o underwent MVR (*o > *u), resulting in *punai > OJ *pune*. For OJ *pe₂* ‘prow’, *əi in *pəi is assumed to have monophthongized to *e₂*, directly giving OJ *pe₂*.

However, several issues remain unresolved in the reconstruction of pJ *pə-nai ‘boat’. First, if *po-nai is reconstructed as an intermediate stage, the predicted pR reflex should be †póne, since pR does not undergo MVR. In reality, however, the form reconstructed for pR is *púne. Alternatively, if pR had branched off before the supposed labialization (*pə-nai > *po-nai), the expected outcome would again be pR †póne, which contradicts the comparative evidence. Instead, one could assume that pR borrowed the post-MVR form *pune* from Japanese, which would remove the problem of irregular correspondences. In this case, the morphemic boundary must have been lost already by the time of borrowing, and to remain consistent with the RCMH, where the loss of transparency triggers the rightward accent shift, borrowing must have occurred after the pre-pR stage.

Second, it remains uncertain whether the proposed rounding of *ə to *o after a labial is supported by additional evidence. Under this hypothesis, explaining why words such as OJ *poko₂* ‘spear’ are not realized as *puko* requires positing a morpheme boundary between the two syllables that blocks MVR, although, to the author’s knowledge, independent evidence for such a boundary has not been demonstrated.

Therefore, even if the second component of ‘boat’ is taken to be *nai ‘root’, the precise etymology of the first component remains to be established and must be left for future research. The simplest explanation would be to regard *pune ‘boat’ as a cultural borrowing from Japanese after the pre-pR stage. However, in the present study, within the framework of the RCMH, I adopt Whitman’s (2012) view that the second component is *nai and reconstruct pJ *pu-nai ‘boat’, leaving the first component presently unetymologized, as shown in (27).

(27) pJ *pu-nai 2.4 ‘boat’ > pre-pR *pú-ne > pR *púne C

No established etymological theory has been proposed for the remaining two items in Group 1, namely *íto C ‘thread’ and *kúda C ‘tube’. To the author’s knowledge, pR *íto ‘thread’, cognate with OJ *ito* 2.4 ‘thread’, has not been analyzed as polymorphemic in previous studies. However, Japanese preserves a root *i* ‘thread’ in fossilized expressions such as *kumo-no i* ‘spider’s thread’,

attested not only in classical poetry but also in modern varieties (SKJ 1989: 137). Similar usage appears in *renkon-no i* ‘sticky string drawn from lotus root’ in Yamaguchi Prefecture (SKJ 1989: 137). This *i* could represent the first component of OJ *ito*, although the absence of independent attestation in OJ suggests that it is a later innovation rather than a pJ inheritance.

One conceivable scenario is that pJ *i-tu-wo, consisting of *i ‘thread’, the genitive *tu, and *wo ‘cord’, developed into pre-pR *ítto through sporadic loss of the second vowel; the resulting heavy syllable would have blocked the rightward accent shift. However, this hypothesis remains highly speculative. Alternatively, considering the ritual use of thread in ancient Japan (JJ 1967: 85), pR *íto could derive from pJ *itu-wo ‘sacred cord’, with *itu- corresponding to OJ *itu-* ‘sacred’, used in religious compounds such as OJ *itu-kasi* ‘sacred oak’, and *wo ‘cord’. Nevertheless, such reconstructions lack independent motivation beyond maintaining consistency with the predictions of the RCMH. Therefore, it may be most plausible to regard pR *íto ‘thread’ as a cultural borrowing from Japanese, introduced after the pre-pR period.

The reflexes of pR *kúda C ‘tube’ are primarily used in the meaning ‘shaft around which the weft thread is wound and placed into the shuttle’, rather than ‘tube’ in general. Given this function, one might posit a compound origin involving the cognate of OJ *kum-* ‘to be entangled’. However, this explanation is difficult to maintain, since OJ *kuda* 2.4 already denotes ‘flute made of bamboo or horn’ (JJ 1967: 260), while its meaning, ‘spinning tool’, is attested only from the late ninth century onward, suggesting that the original meaning was simply ‘tube’ or ‘hollow cylinder’. As pR *kúda refers to a technical implement for weaving or spinning, it may be safest to interpret it as a cultural borrowing from Japanese introduced after the pre-pR stage.

In summary, the historical and phonological evidence discussed in Sections 4.1–4.3 indicates that many of the fifteen Class 2.4–2.5 nouns that appear as pR Class C and end in a non-high vowel without undergoing a rightward accent shift (Group 1 in Table 1) have a polymorphemic origin. For the six items lacking direct phonological evidence, as shown in this subsection, most have likewise been independently identified by multiple scholars as polymorphemic. Combined, these findings reinforce the RCMH, which posits that both word-final high vowels and internal morpheme boundaries inhibited the rightward accent shift.

5 Exceptions to the RCMH

As noted in Section 3, ten nouns in Group 4 listed in Table 2 constitute exceptions to the RCMH and, therefore, require alternative explanations.

One plausible explanation is analogical change. For instance, pR *jorú B ‘night’ (cognate with OJ *jo$_1$ru* 2.4~2.5) may have undergone an analogical tonal shift from Class C to Class B, influenced by the synonymous noun pR *jo B ‘night’ (cognate with OJ *jo$_1$*), which belongs to Class B.

Similarly, pR *aí B ‘indigo’ (cognate with OJ *awi* 2.5) may have changed to Class B by analogy with pR adjectives denoting colors—particularly the etymologically related *ao- B ‘blue’ (cognate with OJ *awo* 2.5). All the basic color terms in pR belong to Class B—*ao- B ‘blue’, *kuro- B ‘black’, and *siro- B ‘white’—except *aka- A ‘red’ (Igarashi 2022a).

Analogical change can likewise be posited for pR *mogí B ‘barley’ (cognate with OJ *mugi$_1$* 2.4) and *kimí B ‘millet’ (cognate with OJ *ki$_1$mi$_1$* 2.5), both denoting types of grain. Many disyllabic nouns referring to cereal crops in pR belong to Class B—for example, *komé B ‘rice’, *awá B ‘foxtail millet’, *iné B ‘rice plant’, and *naé B ‘seedling’. *mamé B ‘bean’ can also be

added to this group. It is therefore plausible that pR *mogí and *kimí underwent an analogical shift from Class C to Class B within pR.

Another plausible explanation is borrowing. pR *zení B ‘money’ is best regarded as a loanword from Japanese. Since the word ultimately derives from Chinese 錢 ‘money’, it cannot be traced back to pJ. The word-initial voiced obstruent likewise points to a non-native origin. Although it is difficult to rule out the possibility of borrowing before the separation of pR and Japanese, the unattestation of the word in OJ supports the view that it was borrowed after the separation. At the same time, the regular correspondences among Ryukyuan varieties that allow the reconstruction of pR *zení B indicate that the borrowing must have occurred by the pR stage or earlier. Therefore, it is reasonable to assume that *zení was borrowed sometime between the separation of pR and Japanese and the subsequent diversification of the Ryukyuan languages. The close formal similarity between pR *zení B and MJ *zeni* 2.4 further suggests that Japanese, rather than Chinese, served as the donor language. If this Sino-Japanese loanword in the donor variety had a form comparable to the final-accented *zení, its reflex as Class B in pR can be naturally explained.

If the remaining five words—pR *kasú B ‘dregs’, *kinú B ‘garment’, *nomí B ‘chisel’, *orí B ‘melon’, and *sirú B ‘soup’—were borrowed from donor varieties with forms similar to the final-accented patterns reconstructed for (pre-)pR, their Class B reflexes could in principle be explained as borrowings. However, evidence supporting this hypothesis is insufficient. Furthermore, as noted in Section 4.4, while borrowing was considered possible for pR *kúda C ‘tube’, *íto C ‘thread’, and perhaps *púne C ‘boat’, those nouns would have to have been borrowed in penultimate-accented form at the pR stage.

Therefore, the borrowing hypothesis raises unresolved questions—for instance, whether the donor variety for pR *zení B ‘money’ differed from that for pR *kúda C ‘tube’, or whether the same donor variety underwent a change in its accentual system between the two borrowing events, or alternatively, whether the loanword phonology of (pre-)pR itself shifted during that period. At present, these questions cannot be answered based on available evidence.

In summary, while analogy and borrowing may plausibly account for some exceptional cases, several nouns remain unexplained under the RCMH. These unresolved items delineate the current empirical limitations of the hypothesis and point to areas where further evidence is required.

6 Conclusion

This study proposed the RCMH as an alternative to the traditional view that split tonal correspondences between Japanese and Ryukyuan directly reflect the multiple tonal contrasts already present in pJ. Rather than positing such inherited contrasts, the RCMH explains the split as the outcome of a merger conditioned by two factors: the presence of a word-final high vowel and the retention of an original word-internal morpheme boundary inherited from pJ.

The evidence from multiple domains supports this hypothesis. The correlation between non-final e_1/o_1 in OJ and pR Class C demonstrates that words containing relics of internal boundaries behaved differently from those lacking them. Additional corroboration comes from apparent violations of Arisaka’s Law and from unexpected heavy syllables in several modern Japanese varieties, both independently pointing to earlier polymorphemic structure. These convergent observations substantiate the central claim that the retention of an original word-internal morpheme boundary in pre-pR blocked the rightward accent shift that produced the Japanese–Ryukyuan split tonal correspondences.

Although the RCMH accounts for most of the observed correspondences, several exceptional cases remain, revealing the empirical limits of the hypothesis and indicating areas in which additional evidence is required. Nevertheless, because the RCMH successfully explains most of the instances that earlier innovation-based hypotheses could not, it remains a proposal worthy of further investigation and refinement.

If the RCMH holds, it eliminates the need to posit more tonal classes in pJ disyllabic nouns than those assumed for MJ by Kindaichi (1974). Previous studies dealing with split tonal correspondences in Japonic have typically adopted one of two approaches: disregarding them altogether (Kindaichi 1960; Hirayama et al. 1966; Tokugawa 1990) or increasing the number of tonal categories in pJ to account for them (Uwano 2006). While the latter approach is methodologically legitimate within the comparative linguistic framework, it should be recognized as a working hypothesis, rather than a definitive reconstruction. Even when no phonological conditioning factors have been identified, tonal innovations may arise from other sources.

The present study has highlighted one such source: word formation, specifically the opacification of polymorphemic structures and possible lexical contraction. These processes have received little systematic attention in previous scholarship on Japanese historical linguistics. Traditional reconstructions of pJ tone have generally assumed that word forms attested in historical stages remain constant over time, reconstructing tonal changes while effectively leaving segmental and morphological structures unexamined. However, the word forms themselves must have changed, and such changes could, in turn, have influenced tonal development.

By applying comparative and internal reconstruction methods rigorously and incorporating the view that opacification of word-internal morphological structure can occur independently across words and varieties, apparent irregular sound correspondences can be systematically explained. This approach does not entail an unnecessary expansion of the tonal inventory; rather, it enables the reconstruction of unattested patterns of word formation.

In this respect, this study departs from the long-standing tendency in Japanese linguistics to posit even more tonal distinctions the further back one reconstructs. Instead, it advances a more constrained and explanatory model—one that reduces rather than inflates the pJ tonal inventory—and may offer a preliminary step toward future inquiry into the pre-pJ origins of tonogenesis—a still-uncharted domain whose exploration may one day echo the breakthrough initiated by Haudricourt (1954), which revealed non-tonal ancestry in Vietnamese and inspired subsequent accounts of Chinese tonogenesis.

References

Arisaka, H. 1932. Kojiki ni Okeru *mo* no Kana no Yōhō ni Tsuite [On the Use of the Kana *mo* in *Kojiki*]. *Kokugo to Kokubungaku* 9(11): 74–93.

de Boer, E. 2010. *The Historical Development of Japanese Tone*. Wiesbaden: Harrassowitz.

Frellesvig, B. and J. Whitman. 2004. The Vowels of Proto-Japanese. *Japanese Language and Literature* 38: 281–299.

Frellesvig, B. and J. Whitman. 2008. Evidence for Seven Vowels in Proto-Japanese. *Proto-Japanese: Issues and Prospects*, eds. B. Frellesvig and J. Whitman, 15–41. Amsterdam: John Benjamins.

Hattori, S. 1958. Amami Guntō no Shohōgen ni Tsuite: Okinawa, Sakishima Shohōgen to no Hikaku [On the Dialects of the Amami Islands: A Comparison with the Okinawa and Sakishima Dialects]. *Jinrui Kagaku* 11: 79–99.

Hattori, S. 1978–79. Nihon Sogo ni Tsuite (1)–(22) [On Proto-Japanese (1)–(22)]. *Gekkan Gengo* 7(1)–8(12). Reprinted in S. Hattori. 2018. *Nihonsogo no Saiken* [Reconstruction of Proto-Japanese], 87–401. Tokyo: Iwanami.

Haudricourt, A.-G. 1954. De l'Origine des Tons en Vietnamien [The Origin of Tones in Vietnamese]. *Journal Asiatique* 242: 69–82. Translated into English by M. Brunelle, 2018. Available at: https://shs.hal.science/halshs-01678018 (accessed 15 October 2025).

Hirayama, T., I. Oshima, and M. Nakamoto. 1966. *Ryūkyū Hōgen no Sōgōteki Kenkyū* [A Comprehensive Study of the Ryukyuan Dialects]. Tokyo: Meiji Shoin.

Hirayama, T., I. Oshima, M. Ono, M. Kuno, M. Kuno, and T. Sugimura, eds. 1992–93. *Gendai Nihongo Hōgen Daijiten* [The Comprehensive Dictionary of Modern Japanese Dialects]. Tokyo: Meiji Shoin.

Igarashi, Y. 2022a. Reconstruction of Ryukyuan Tone Classes of Middle Japanese Class 2.4 and 2.5 nouns. *Open Linguistics* 8(1): 232–257.

Igarashi, Y. 2022b. 2-onsetsu Meishi Dai-4/5-rui ni Taiō Suru Ryūkyūsogo B-rui wa Kaishin de Aru to Suru Kasetu [The Hypothesis that Proto-Ryukyuan Class B Corresponding to Disyllabic Nouns of Classes 4/5 Is an Innovation]. A talk presented at the Annual Meeting of the Linguistic Society of Japan.

JJ = Jodaigo Jiten Editorial Committee, ed. 1967. *Jidaibetsu Kokugo Daijiten: Jōdaihen* [The Historical Dictionary of the Japanese Language: The Old Japanese Period]. Tokyo: Sanseido.

Kamimura, K. 1983. Kyūshū Hōgen no Gaisetsu [An Outline of the Kyushu Dialects]. *Kōza Hōgengaku 9 Kyūshū Chihō no Hōgen* [A Course in Dialectology, Vol. 9: The Dialects of the Kyushu Region], eds. K. Iitoyo, S. Hino, and R. Sato, 1–28. Tokyo: Kokusho Kankokai.

Kiku, C. and T. Takahashi. 2005. *Yoron Hōgen Jiten* [The Dictionary of the Yoron Dialect]. Tokyo: Musashino Shoin.

Kindaichi, H. 1960. Akusento kara Mita Ryūkyūgo Shohōgen no Keitō [The Genetic Relationship of the Ryukyuan Dialects from the Perspective of Accent]. *Tōkyō Gaikokugo Daigaku Ronshū* 7: 59–80.

Kindaichi, H. 1974. *Kokugo Akusento no Shiteki Kenkyū: Genri to Hōhō* [A Historical Study of Japanese Accent: Principles and Methods]. Tokyo: Hanawa Shobo.

Kurano, K. 1963. *Kojiki*. Tokyo: Iwanami Shoten.

McCarthy, J. J. and A. Prince. 2001. *Prosodic Morphology: Constraint Interaction and Satisfaction* (RuCCS Technical Report #3). Rutgers University Center for Cognitive Science. https://doi.org/10.7282/T3B856GM.

Martin, S. E. 1987. *The Japanese Language through Time*. London: Yale University Press.

Matsumori, A. 1996. Ryūkyū ni Okeru 2-onsetsugo Dai 4-5 rui no Gotō Chōboin o Meguru Mondai: Kita Ryūkyū Sogo no Akusento Saiken ni Mukete [Problems Concerning Word-initial Long Vowels in Class 4–5 Disyllabic Words in the Ryukyus: Toward a Reconstruction of Proto-Northern Ryukyuan Accent]. *Nihongo Kenkyū no Shoryōiki no Shiten* [Perspectives on Various Fields of Japanese Linguistics], ed. Hirayama Teruo Hakase Beiju Kinenkai, 1130–1147. Tokyo: Meiji Shoin.

Matsumori, A. 1997. Tokushimaken Wakimachi, Mikamochō no Akusento to Hondo Sogo no Akusento Taikei [Accent in Wakimachi and Mikamocho Dialects in Tokushima and the Proto-accent System of Mainland Japanese]. *Kokugogaku* 189: 55–68.

Matsumori, A. 2000. Ryūkyū no Takei Akusento Taikei ni Tsuite no Ichi Kōsatsu: Ryūkyū Sogo ni Okeru Ruibetsu Goi 3-pakugo no Gōryū no Shikata [An Examination of So-called Multi-patterned Accent Systems in Ryukyuan Dialects Focusing on Three-syllable Words]. *Kokugogaku* 51(1): 93–108.

Matsumori, A. 2009. Okinawa Hontō Kin Hōgen no Akusento Gata to Sono Keiretsu: Ryūkyū Chōsayō Keiretsubetsu Boi no Kaihatsu ni Mukete [Accent Types and Classes of Nominals in the Kin Dialect

of Okinawa Island: Toward the Development of Lexical Lists for Ryukyuan Accent Research]. *Memoirs of the Japan Women's University, Faculty of Literature* 58: 97–122.

Matsumori, A. 2012. Ryūkyūgo Chōsayō Keiretsubetsu Goi no Soan [Toward a Categorized Vocabulary for Ryukyuan Field Research]. *Journal of Phonetic Society of Japan* 16(1): 30–40.

Matsumori, A. 2017. Kita-Ryūkyū ni Okeru C-keiretsu 2-onsetsu Meishi no Gotō Onsetsu no Chōonka: Sono Gen'in ni Tsuite Kangaeru [Analysis of the Cause of Vowel Lengthening in the Initial Syllables of the Class-C Disyllabic Nouns in Northern Ryukyuan Dialects]. *Nihongo no Kenkyū* 13(1): 1–17.

Matsumoto, K. 1975. Kodai Nihongo Boin Soshiki kō: Naiteki Saiken no Kokoromi [On the Vowel System of Ancient Japanese: An Attempt at Internal Reconstruction]. *Kanazawa Daigaku Hōbun Gakubu Ronshū* 22: 83–152.

Miyagi, S. 2003. *Ishigaki Hōgen Jiten* [The Dictionary of the Ishigaki Dialect]. Okinawa: The Okinawa Times.

Nakama, H., Y. Takubo, S. Iwasaki, Y. Igarashi, and N. Nakagawa. 2022. *Minami Ryūkyū Miyako-go Ikema Hōgen Jiten* [The Dictionary of the Ikema Dialect of Southern Ryukyuan Miyako]. Tokyo: NINJAL.

Nakasone, S. 1983. *Okinawa Nakijin Hōgen Jiten* [The Dictionary of the Nakijin Dialect of Okinawa]. Tokyo: Kadokawa Gakugei Shuppan.

Nitta, T. 2010. Hakusanroku Shiramine no Hōgen Tokuchō to Mukashibanashi ni Mirareru Hōgen no Gohō (2) [Dialectal Features of Shiramine in the Hakusan Foothills and the Usage of Dialect in Local Folktales (2)]. *Kanazawa Daigaku Rekishi Gengo Bunkagaku-kei Ronshū: Gengo, Bungaku-hen* 2: 61–92.

NKD = Nihon Kokugo Daijiten Editorial Committee, ed. (2001) *Nihon Kokugo Daijiten* [The Comprehensive Dictionary of the Japanese Language]. 2nd ed. Tokyo: Shogakukan.

Okumura, S. 1958. On'in to Akusento: Akusento Kenkyū no Igi [Phonology and Accent: The Significance of Accent Studies]. *Kokugo Kokubun*. 27(9): 1–16.

Ono, S. 1977. On'in no Hensen (1) [Phonological Change (1)]. *Iwanami Kōza Nihongo 5: On'in* [Iwanami Lectures on the Japanese Language 5: Phonology], ed. M. Hashimoto, 147–219. Tokyo: Iwanami.

Ono, S., A. Satake, and K. Maeda, eds. 1974. *Iwanami Kogo Jiten* [Iwanami Dictionary of Classical Japanese]. Tokyo: Iwanami.

ONCOJ. 2020. *Oxford NINJAL Corpus of Old Japanese*. https://oncoj.ninjal.ac.jp/ (accessed 22 February 2021).

Pellard, T. 2009. Once Again on Long Vowels and Register in Proto-Japonic. Presented at the International Workshop on "The History & Reconstruction of Japanese Accent". INALCO, Paris.

Pellard, T. 2013. Ryukyuan Perspectives on the Proto-Japonic Vowel System. *Japanese/Korean Linguistics* 20, eds. B. Frellesvig and P. Sells, 81–96. Stanford: CSLI Publishers.

Pellard, T. 2015. The Linguistic Archeology of the Ryukyu Islands. *Handbook of the Ryukyuan Languages: History, Structure, and Use*, eds. P. Heinrich, S. Miyara, and M. Shimoji, 13–37. Berlin: De Gruyter Mouton.

SDJKS = Shimabara Daiichi Jinjo Koto Shogakko, eds. 1975. *Shimabara Hantō Hōgen no Kenkyū* [A Study of the Dialects of the Shimabara Peninsula]. Tokyo: Kokusho Kankokai.

SKJ = Shogakkan Kokugo Jiten Editorial Committee, ed. 1989. *Nihon Hōgen Daijiten* [The Comprehensive Dictionary of Japanese Dialects]. Tokyo: Shogakkan.

Shimabukuro, M. 2007. *A Reconstruction of the Accentual History of the Japanese and Ryukyuan Languages*. Folkestone: Global Oriental.

Shimabukuro, M. 2008. A Reconstruction of Proto-Japanese Accent for Disyllabic Nouns: Focusing on the Problem of Subclasses. *Proto-Japanese: Issues and Prospects*, eds. B. Frellesvig and J. Whitman, 125–139. Amsterdam: John Benjamins.

Thorpe, M. L. 1983. Ryūkyūan Language History. Ph.D. thesis, University of Southern California.

Tokugawa, M. 1990. Nihon no Hōgen: Nihongo no Keisei to no Kakawari [Japanese Dialects: Their Relation to the Formation of the Japanese Language]. *Nihongo no Keisei* [The Formation of the Japanese Language], ed. O. Sakiyama, 248–272. Tokyo: Sanseido.

Tokuyama, S. and K. Celik. 2020. *Minami Ryūkyū Miyako-go Tarama Hōgen Jiten* [The Dictionary of the Tarama Dialect of Southern Ryukyuan Miyako]. Tokyo: NINJAL.

Unger, J. M. 2010. New Etymologies for Some Japanese Time-words. *Journal of the American Oriental Society* 130(1): 35–41.

Uwano, Z. 1996. Naze-shi Ashikebu, Arira Hōgen no Meishi no Akusento Taikei [The Nominal Accent Systems of the Ashikebu and Arira Dialects of Naze City]. *Tōkyō Daigaku Gengogaku Ronshū* (15): 3–68.

Uwano, Z. 2006. Nihongo Akusento no Saiken [On the Reconstruction of Japanese Accents]. *Gengo Kenkyū* 130: 1–42.

Uwano, Z. 2017. Tokunoshima Asama Hōgen no Akusento Shiryō (3) [Materials on the Accent of the Asama Dialect of Tokunoshima (3)]. *NINJAL Research Papers* 12. 139–161.

Vovin, A. 2008. Proto-Japanese beyond the Accent System. *Proto-Japanese: Issues and Prospects*, eds. B. Frellesvig and J. Whitman, 141–156. Amsterdam: John Benjamins.

Vovin, A. 2014. Out of Southern China? Some Linguistic and Philological Musings on the Possible Urheimat of the Japonic Language Family. Paper presented at the 27th Journées de Linguistique d'Asie Orientale, June 26–27, 2014.

Vovin, A. 2020. *A Descriptive and Comparative Grammar of Western Old Japanese*. Brill Academic Pub.

Whitman, J. 1985. *The Phonological Basis for the Comparison of Japanese and Korean*. Ph.D. thesis, Harvard University.

Whitman, J. 2012. The Relationship between Japanese and Korean. *The Languages of Japan and Korea*, ed. D. N. Tranter, 24–38. London: Routledge.

Yonaguni Board of Education. 2019. *Dunan Munui Jiten* [The Dictionary of the Yonaguni Language]. Okinawa: Nanzansha.