

3 EVIDENCE FOR SEVEN VOWELS IN PROTO-JAPANESE

BJARKE FRELLESVIG AND JOHN WHITMAN

University of Oxford and Cornell University

1. Introduction

In this paper we present evidence to support the reconstruction of seven short vowels for Proto-Japanese in (1) originally proposed and preliminarily presented in Frellesvig and Whitman (2004, forthcoming).¹ The seven vowel system for Japanese is reconstructed on Japanese-internal grounds (primarily internal reconstruction and dialect comparison), but it turns out to find further support from Japanese/Korean comparative evidence.

- (1) **i* **ɨ* **u*
 **e* **ɛ* **o*
 **a*

OJ had the following eight distinct vowels and postconsonantal diphthongs, shown in (2) in the notation used in this book (see Table 2 in §0.1.2). These are the eight entities which earlier were thought to be unitary vowel phonemes and referred to as the ‘eight vowels’ of OJ. Since Lange (1973) and Matsumoto (1974) a diphthongal interpretation has become current.

- (2) *i* *u* -*wi*
 e *o* -*ye* -*wo*
 a

¹ Vowel length has been reconstructed for pJ, based mainly on interpreting low pitch in EMJ as reflecting pJ long vowels, supplemented with Ryukyuan evidence in the form of what seem to be primary long vowels. Vovin (1993) offers additional external evidence. The precise role of this feature in changes between pJ and OJ is far from clear. Vowel length has been proposed to have been a conditioning environment for certain sound changes; for example, vowel raising only applying to short vowels (Hayata 1998), or loss of **m* and **r* only taking place after (some) short vowels (Whitman 1985). The hypothesis we set forth here deals only with short vowels.

The correspondences between the pJ vowels we reconstruct and the vowels and postconsonantal diphthongs of OJ are shown in (3).

(3) Seven vowel hypothesis (7VH)

pJ	OJ
*i	i
*e	i, ye
*a	a
*o	u, wo
*u	u
*ɨ	o
*ə	o

Pre-OJ	OJ
*ui	wi
*ɨi	wi
*əi	e
*ai	e

*iɨ	ye
*iə	ye
*ia	ye
*uɨ	wo
*uə	wo
*ua	wo

Until recently, a four vowel hypothesis (4VH) has been dominant, holding that pJ had only four primary vowels (e.g. Miller 1967, Matsumoto 1974, 1984, 1995, Ohno 1977, Whitman 1985), /*i, *a, *u, *ə/, reflected as OJ /i, a, u, o/.

(4)

*i	*u
*ə	
*a	

The correspondences between the pJ four vowel system and OJ are given in (5).

(5) Four vowel hypothesis (4VH)

pJ	OJ
----	----

**i*

i

**a*

a

**u*

u

**ə*

o

Pre-OJ	OJ
* <i>ui</i>	<i>wi</i>
* <i>ɔi</i>	<i>wi</i>
* <i>ai</i>	<i>e</i>
* <i>ia</i>	<i>ye</i>
* <i>iə</i>	<i>ye</i>
* <i>ua</i>	<i>wo</i>
* <i>uə</i>	<i>wo</i>

Comparing (3) and (5), it will be seen that our 7VH offers a more differentiated and complex account of the origin of OJ *-i*, *-u*, *-o*, *-e*, *-ye*, *-wo*, *-wi* than does the 4VH. Thus, the 4VH has single sources for *-i*, *-e*, *-o*, *-u* and it regards *all* OJ *-ye* and *-wo* as secondary, resulting from contractions of pre-OJ vowel sequences. As opposed to this, we argue (§3) that some *-i*, *-ye*, *-u*, *-wo* reflect primary mid vowels, **e* and **o*, which in main island varieties in most positions rose to merge with **i* and **u*, but in some positions gave *-ye* and *-wo*. We also argue (§4) that OJ *-o* had *two* sources, a mid and a high central vowel, **ɔ* and **ɔ̄*, which in the central dialect merged to give OJ **o*, outside of contraction with **-i* (or *-y*) where the outcomes were **ɔi* > *-e* and **ɔ̄i* > *-wi*, respectively, thus also positing **two** sources for OJ *-e* (viz. **ai*, **ɔi*).

2. *-e*, *-ye*, *-wo*, *-wi*

The proposals we present in this paper are mainly based on different interpretations of the origin of OJ *-e*, *-ye*, *-wo*, *-wi* from previous scholars. While *-e*, *-ye*, *-wo*, *-wi*, which are all regarded as secondary in the 4VH, are similar in being relatively infrequent and mostly occurring in morpheme- or root-final position, we argue that this is in fact coincidental and that their distributions have different sources: *-wi* and *-e* are truly secondary, but most *-ye* and *-wo* directly reflect primary vowels.

2.1 *Frequency and distribution.*

A count of the text occurrence of the different orthographically distinct syllable types in the Man'yōshū shows the following distribution (Ohno 1980:151ff), (6). Although a frequency count in running text gives no real picture of lexical distribution, these figures do give some indication that the main primary vowels were those in the syllables *Ci₍₁₎*, *Ca*, *Co₍₂₎*, *Cu*. Likewise, the fact that whereas simple morphemes of the structure *CiCi*, *CaCa*, *CoCo*,

CuCu all are well represented in the lexicon, *CwoCwo* is very rare² and there are no simple morphemes of the type *CyeCye*, *CeCe*, or *CwiCwi*, shows that *Cwo*, *Cye*, *Ce*, *Cwi* are somehow not on a par with *Ci₍₁₎*, *Ca*, *Co₍₂₎*, *Cu*.





(6)

<i>Ci₁</i>	3,160	<i>Ca</i>	12,120
<i>Ci</i>	6,103		
<i>Ci₂</i>	370	<i>Co₁</i>	1,030
		<i>Co</i>	3,631
<i>Ce₁</i>	686	<i>Co₂</i>	5,280
<i>Ce</i>	2,299		
<i>Ce₂</i>	853	<i>Cu</i>	6,415

In terms of lexical distribution, *-wi*, *-e*, *-ye*, *-wo* are similar in being relatively infrequent and mostly occurring in morpheme- or root-final position (cf. JLTT, Matsumoto 1995). In nonfinal position, *-wi*, *-e*, *-ye*, *-wo* are all rare. In simple forms, nonfinal *-wi* is found only in:³ *pwiwe-* ‘to scrape, slice thin’, *kwisi* ‘shore’, *kwiri* ‘fog’.⁴ Truly nonfinal *-wo* is rare, although a number of instances are in a morpheme or root final syllable and not in absolute word final position (cf. §§3.1, 3.3 below); JLTT:60-2 gives a list of OJ words with *-wo*. Also *-e* and *-ye* are infrequent in nonfinal position, although not exceptional; some cases may be positively

² *Mwomwo* ‘hundred’ and *mwomwo* ‘thigh’ look like reduplications; early OJ *mwokwo* (~ later OJ *mukwo*) ‘partner, bridegroom’ may be a compound with *kwo* ‘child’.

³ The word *mwina* ‘all’ is most likely a compound: *mwi-na* ‘body-person’. Sometimes *kokwida* ‘this much’ is also cited as an example of nonfinal *-wi*, but that is at least bimorphemic: *kokwi-da*, cf. the following OJ forms built on *koko-/kokwi-* ‘this much’: *kokoda*, *kokwida*, *kokodaku*, *kokwidaku*, *kokoba*, *kokobaku*, *kokwibaku*. *koko-/kokwi-* is itself analyzable at least into *ko* ‘proximal’ + *-kV* ‘degree, quantity’, cf. *i-ku-* ‘how much?’ and *so-kV-* ‘that much’, both with a number of further derived forms.

⁴ *-wi* in *kwisi* and *kwiri* (and possibly *pwiwe-*, see further below) may in fact be rare examples of sporadic ‘umlaut’ under influence of *i* in the following syllable, going back to **kusi/k si* (cf. MK *k*   < **k*   ‘bank, shore’) and **kuri* (cf. Hachijōjima *kuri* ‘fog’, MK *kurum* ‘cloud’). Other potential examples of sporadic umlaut include *seki* (~ *saki*) ‘barrier’, *neri* ‘paste’, *semi* ‘cicada’, *seri* ‘parsley’. In the following Nidan verbs, the Nidan theme (cf. §2.3 below) may have induced umlaut in the penultimate syllable (in addition to having been contracted with the root final vowel): *nedi-* ‘twist’ (?~ *yodi-*), *negwi-* ‘appease’ (~ *nagu-* ~ *nagwo-*), *nige-* ‘flee’ (~ EMJ *nogare-*), *pwiwe-* (see Whitman 1985:133 for a different etymology), *wosipe-* ‘teach’ (~ NJ *osowar-* ‘learn’ < OJ **wosopar-*).

identified as loanwords,⁵ and others may in fact be loanwords (see further §5.3 below about loanwords).

⁵ E.g., *kyesa* ‘priest’s robe’ (袈裟, cf. LMC **kja*: ʔa:, EMC **ka* ❖ ʔa ❖; Skt. *kaś”ya* ‘the (yellow) robe of Buddhist clergy’), *pyera* ‘spatula, pallet’ (cf. MK *pyet* ‘moldboard’), *tera* ‘temple’ (cf. MK *tyel* ‘temple’, 刹 EMC **tʰa* ❖t, OC **tshraat/ksraat*, Skt. *kāśetra* ‘place’). EMC and LMC reconstructions follow Pulleyblank (1984); OC follows Miyake (1997 and p.c.).

2.2 Lexical contraction

Common to any account of pre-OJ phonology is the observation that OJ *-wi*, *-e*, *-wo*, *-ye* in some words result from contraction of vowel sequences, e.g. (7),⁶ where juxtapositions with a vowel initial morpheme as the second part were univerted.

- (7) a. *wi* < *u+i* *wakwiratukwo* < *waku+iratukwo* ‘young+honored male’
 wi < *o+i* *opwisi* Proper name < *opo+isi* ‘big+stone’
 e < *o+i* *toneri* ‘palace servant’ < *tono+iri* ‘palace+enter’
 e < *a+i* *taketi* Proper name < *taka+iti* ‘high+market’
- b. *wo* < *u+o* *sitwori* ‘type of cloth’ ~ *situ-ori* ‘native.weaving-weave’
 wo < *u+a* *kazwope-* ‘count’ < *kazu+ape-* ‘number+join’
 ye < *i+a* *sakyeri* ‘is blooming’ < *saki+ari* ‘bloom+is’
 ye < *i+o* *pyeki* Proper name < *pi+oki* ‘sun-put’

Although these contractions have, rightly, attracted much attention it must be noted that examples of lexical contractions of this sort in fact are quite few. The only solid cases of *-wi* and *-wo* are those listed in (7) and there is a small handful of examples with *-e*. Only *-ye* is found in a sizeable number of forms, which all are fairly transparent contractions from **i-a* involving grammatical morphemes, including: *-yer-* Stative < **-i-ar-* ‘Infinitive-be’,⁷ *-kyeku* Past Nominal < **-ki-aku* ‘Past-Nominalizer’, *-kyeku* Adjectival Nominal < **-ki-aku* ‘Adjectival.Adnominal-Nominalizer’, *-kyem-* Past Conjectural < **-ki-am-* ‘Past-Conjectural’, *-kyer-* Modal Past < **-ki-ar-* ‘Past-be/come-be’, *myes-* ‘see.Honorific’ < **mi-as-* ‘see-Honorific’, *kyes-* ‘wear.Honorific’ < **ki-as-* ‘wear-Honorific’, *kyesa* ‘this morning’ (cf. §3.4 below) < **ki-asa* ‘this-morning’.

2.3 Apophony

⁶ The notation in (7) is shorthand for saying, e.g., that *wakwiratukwo* is thought to reflect a contraction of two pre-OJ forms which are reflected as OJ *waku* and *iratukwo*, exemplifying OJ *-wi* originating in a contraction of two pre-OJ vowels which are reflected as OJ *u* and *i*.

⁷ This is the regular morphological Stative, exemplified in (7) by *sakyeri*; included here are also some lexicalized forms: *kyer-* ‘be wearing’ < **-ki-ar-* ‘wear-be’, *kyer-* ‘have come’ < **-ki-ar-* ‘come-be’, *ser-* ‘have done’ < (**syer-* <) **si-ar-*.

A number of nouns have alternating shapes with variation in the final syllable; these are often referred to as ‘apophonic’ stems. One variant, the free form (known as *roshutsukei* ‘exposed form’), occurs in word final position, while the other, the compound form (*hifukukei* ‘covert form’), usually occurs in compounds or derived forms. Traditionally the compound form is thought to represent the original shape of the root.

(8)		<i>roshutsukei</i>	<i>hifukukei</i>
wi ~ u	‘body’	<i>mwi</i>	~ <i>mu-kapari</i> (‘substitute’) ‘hostage’
	‘yellow’	<i>kwi</i>	~ <i>ku-gane</i> (‘metal’) ‘gold’
	‘moon’	<i>tukwi</i>	~ <i>tuku-ywo</i> (‘night’) ‘moonlit night’
wi ~ o	‘tree’	<i>kwi</i>	~ <i>ko-dati</i> (‘stand’) ‘grove’
	‘Hades’	<i>yomwi</i>	~ <i>yomo-tu-kuni</i> (‘Gen-land’) ‘id.’
e ~ o	‘back’	<i>se</i>	~ <i>so-muku</i> ‘turn’
e ~ a	‘eye’	<i>me</i>	~ <i>ma-pye</i> (‘side, direction’) ‘front’
	‘ <i>saké</i> ’	<i>sake</i>	~ <i>saka-duki</i> (‘cup’) ‘ <i>saké</i> cup’

One view of these cases of nominal apophony is that the *roshutsukei* originates from the contraction of the original root-final vowel with a morpheme of the shape **i* (Yoshitake 1930), as shown in (9a). Phonologically this conforms to the contractions noted in §2.2. The function and origin of this **i* is subject to some debate, but one view relates it to the subject marker *i* vestigially attested in OJ. A further development of this derives the apophonic nouns from consonant final shapes, with the final consonant being lost before **i* (Whitman 1985), see (9b).

(9)			(a)	(b)	
	‘body’	<i>mwi</i>	< <i>*muy</i>	< <i>*mu-i</i>	< <i>*mum-i</i>
	‘yellow’	<i>kwi</i>	< <i>*kuy</i>	< <i>*ku-i</i>	< <i>*kur-i</i>
	‘eye’	<i>me</i>	< <i>*may</i>	< <i>*ma-i</i>	< <i>*mar-i</i>

Yet a third view (Murayama 1962) of the apophonic nouns reconstructs a specific final consonant, **-r*, which weakened to a yod (yodization) in phrase final position, (10).

(10)	‘yellow’	<i>kwi</i>	< <i>*kuy</i>	< <i>*kur</i>
	‘eye’	<i>me</i>	< <i>*may</i>	< <i>*mar</i>

On both Whitman’s and Murayama’s views, the final consonant of the reconstructed forms is deleted in composition before another consonant. The evidence for the identity of the reconstructed final consonants is in all cases comparative, from Korean.

Thus, nonfinal *-wi*, *-e*, *-wo*, *-ye* are all quite rare. Some instances may positively be traced back to lexical contraction or borrowing (and possibly to other events, such as sporadic umlaut, suggested in §2.1, note 4) and we believe that all cases of truly nonfinal *-wi*, *-e*, *-wo*, *-ye* have such sources (apart from those which may originate in the type of irregular, singular, word individual changes found in any language).

In final position *-wi*, *-e*, *-wo*, *-ye* are not frequent, but nor are they rare. Their distribution is markedly and significantly different. Almost all final *-wi*, *-e* (and thus the vast majority of all *-wi*, *-e*) are found in apophonic environments, either in nominal apophony or as the stem vowel of Nidan verbs (whether or not with an attested *hifukukei*), both of which are thought to originate in contractions of a primary vowel with **i* or **y*. As opposed to this, final *-wo* and *-ye* do not take part in apophonic alternations, nor can they generally be etymologized as contractions.⁸ Contra the four vowel hypothesis, the evidence that OJ *-wo* and *-ye* are diachronically secondary is thus on closer inspection far less persuasive than for *-wi* and *-e*.

3. Mid vowel raising: pJ **e*; **o* > OJ *i*, *ye*; *u*, *wo*

The fact that *-wo* and *-ye* generally cannot be etymologized as secondary contractions is one motivation for the hypothesis, originally due to Hattori, that *-ye* and *-wo*, as opposed to *-wi* and *-e*, reflect two primary pJ vowels, reconstructed as **o* and **e*. See Hattori (1976), NSNT, Hayata (1998, 2000), Serafim (1999), and Hino (2003) for proposals based on internal and comparative dialect evidence for reconstruction of pJ **e*, **o*. Miyake (2003) provides philological evidence for this proposal. All versions of this hypothesis imply a sound change of mid vowel raising (MVR) in the course of which **e* and **o* raised in all or certain environments in pre-OJ. In the following (§§3.1, 3.2) we review the case for reconstruction of **e* and **o*. We propose, as shown in (15) (see further §3.3), that pJ **e*, **o* raised in all environments, but that the outcome was conditioned by whether MVR took place in ‘final position’: > *wo*, *ye*, ‘partial raising’ (reflexes of pJ **i* vs. **e* and **u* vs. **o* remaining distinct); or non-final position: > *i*, *u*, ‘full raising’ (merger of pJ pJ **i* vs. **e*, and **u* vs. **o*, respectively). On our proposal, dialects/varieties differed in the domain (word, morpheme, root) with reference to which ‘final’ was defined.

⁸ Some cases of final *-wo*, *-ye* may go back to forms of the shape **CVCV*, > **CVV* through consonant loss and further > *Cwo*, *Cye* through contraction; but it is in fact difficult to come up with plausible candidates for such developments motivated by anything other than the presence of final *-wo*, *-ye*. Whitman (1985) compares OJ *kwo* with MK *koma* ‘child’, seemingly providing an external motivation for the reconstruction of the *-wo* as originating in contraction, but Lee (1991:246) points out that the latter is a loan from Mongolian.

(14) **e* > -*ye* in final position (partial raising)
 i elsewhere (full raising)

**o* > -*wo* in final position (partial raising)
 u elsewhere (full raising)

It should be noted that the effect of *partial raising*, because of the merger of *Ce* and *Cye*, and *Co* and *Cwo*, respectively, was neutralized in OJ for **e* after coronals and glides, for **o* after labials, and eventually in EMJ after all consonants, see (15).

(15)

pJ	pre-OJ	OJ	EMJ
	/ C = {p, b, m, k, g}	<i>Cye</i> >	<i>Ce</i>
* <i>Ce</i> > * <i>Cye</i> >	/ C = {t, d, s, z, n, r, w (, y)}	<i>Ce</i> >	<i>Ce</i>
	/ C = {t, d, k, g, s, z, n, r, y}	<i>Cwo</i> >	<i>Co</i>
* <i>Co</i> > * <i>Cwo</i> >	/ C = {p, b, (m,) w}	<i>Co</i> >	<i>Co</i>

Finally in this section we show how the hypothesis of MVR helps us understand the correspondences between the Japanese and Korean demonstratives (§3.4).

3.1 *pJ* **o* > *OJ* u, wo

Internal evidence for raising of **o* may be found in variation between *-u* and *-wo* within the OJ lexicon (16) and across dialects (17). See Hayata (1998), Matsumoto (1995:79ff, 132f) for *-wo*~*-u* forms, also Thorpe (1983).

(16) **yo(-)ri* *ywo(-)ri* ‘from’ ~ *yuri* ‘behind; from’
**mo(-)ko* *mwokwo* ‘partner, bridegroom’ ~ *mukwo* (> EMJ *muko*)

(17) Central OJ *-wo* Eastern OJ *-u*
**yo* *ywo* ‘night’ :: *yu-* (*yutoko* ‘night bed’)
**no* *nwo* ‘field’ (?~ *numa* ‘marsh’): *nu-* (*nu-no* ‘field-Gen.’)

Note that final position is implicated in the alternations above: in (16), *ywo*~*yuri* is usually analyzed as involving a morpheme boundary between *ywo*~*yu* ‘from’ and the ending *-ri*, while *mwokwo* > *mukwo* ‘partner, bridegroom’ is arguably a compound involving a morpheme **mwo* + *kwo* ‘child, man’. In each case, **o* > *u* can be associated with loss of transparency in the derivational or compound boundary. The facts are similar in (17); *-wo* appears finally as in *ywo* ‘night’ and *nwo* ‘field’, but *u* appears word-internally. The forms with *u* are attested in Eastern OJ material, but what is at issue here is not dialectal correspondences but different

responses to the compound boundary in different varieties.⁹ The crucial point in these examples is that **o* is reflected as *-wo* in final position in one variety, but raised fully to *-u* before an internal morpheme boundary in another.

While the pattern of Eastern OJ *u* for COJ *wo* reflects varying interpretations of compound boundaries, or varying degrees of lexicalization, a second pattern of contrast between EOJ and COJ may reflect more systematic variation in the environment for raising **o*.¹⁰ EOJ *wo* for COJ *u* is attested throughout the Eastern OJ region. In the westernmost part of the region, this is limited to a single morpheme, the finite primary verb ending¹¹ *-wo* corresponding to COJ *-u*. Further to the east, **o* is reflected as *-wo* before derivational suffix boundaries in some inflecting stems.

(18)	Central OJ <i>u</i>		Eastern OJ <i>wo</i>
	<i>*-o</i>	<i>-u</i> finite verb ending	:: <i>-u</i> ~ <i>-wo</i>
	<i>*su/ogo-su</i>	<i>sugusu</i> ‘pass it’	:: <i>sugwosu</i>
	<i>*popo-mu</i>	<i>pupumu</i> ‘to bud’	:: <i>popomu</i>

These facts suggest that at least some EOJ varieties - those most distant from the central region - had a more restricted environment for raising **o*, although the exact nature of what counted as a ‘final’ environment for the relevant EOJ varieties - perhaps before transparent derivational suffixes - is difficult to recover from the surviving data.

Examples like those in (18) allow us to recover pJ **o* in contexts where in COJ it raised and merged with **u*. Comparative Ryukyuan data, here due to Thorpe (1983), reveals further examples of this sort. Ryukyuan in isolation provides no basis for distinguishing vowels corresponding to OJ *-wo* (= *-o₁*) and *-o* (= *-o₂*), and in most varieties pR **o* raises and merges with **u*. But in some Amami varieties (a subset of those which distinguish reflexes of **i* and **e*; cf. Thorpe 1983:32), *o* is preserved. The examples in (19) are cases where this vowel corresponds to Central OJ *u*.

⁹ It is instructive here to note that the Central OJ attestation of ‘night bed’ as *ywodoko* in *Nihon shoki* (NS 47) has rendaku after the partially raised *-wo*, showing that the compound (or word) boundary was more transparent: *ywo-doko*.

¹⁰ Fukuda (1965: 301) observes that the pattern of EOJ *u* corresponding to COJ *wo* is not found in the four easternmost *kuni* (Shimotsuke, Kozuke, Kazusa, and Mutsu); on the other hand EOJ *wo* :: COJ *u*, as in (18), is found across the EOJ region. The vast majority of examples of the latter type involve the ending of finite verbs.

¹¹ In traditional terminology, the *rentaikei* or Adnominal ending in consonant base verbs. See, however, Frellesvig, this volume (fn. 16).

(19)	Central OJ - <i>u</i>	proto-Ryukyuan * <i>o</i>
	* <i>siro(-)si</i> <i>sirusi</i> ‘mark’	:: * <i>sirosi</i>
	* <i>mo(-)ko</i> <i>mwokwo</i> ~ <i>mukwo</i> (see (17) above)	:: * <i>moko</i> ‘son in law’
	* <i>tuko-yo</i> <i>tukuywo</i> ‘moonlit night, moon’	:: * <i>tukoyo</i> ‘moon’
	< * <i>tuko</i> ‘moon’ + * <i>yo</i> ‘night’	

As with the EOJ data, pR gives no evidence for word nonfinal pJ **o* except before possible morpheme boundaries (the argument that *sirus-* ‘write, mark down’ originates from a morphologically complex form **sirV-s* is weaker than in the case of the other two compound forms, but is not to be rejected out of hand). This raises the possibility that raising of **o* was already underway at the pre-OJ stage at which pR split off.

3.2 pJ **e* > OJ *i*, *ye*

Comparative Ryukyuan-Japanese evidence gives the primary impetus for reconstructing an additional pJ front vowel, reconstructed by Hattori (1976), NSNT and Thorpe (1983) as pJ **e*. This evidence involves cases where pR **e* corresponds to OJ *i*. Hattori and Thorpe reconstruct pR **e* for words where North Amami dialects give a central vowel *i* corresponding to *i* elsewhere in the Ryûkyûs as well as OJ. Examples of this sort are listed in (20);¹² the pR reconstructions follow Thorpe (1983). See Serafim (this volume) for further examples and discussion of this correspondence.

(20)	OJ <i>i</i>	proto-Ryukyuan * <i>e</i>
	<i>midu</i> ‘water’	:: * <i>medu</i>
	<i>oyobi</i> ‘finger’	:: * <i>UyUbe</i> (* <i>U</i> represents pR * <i>u</i> or * <i>o</i>)
	<i>piru</i> ‘garlic’	:: * <i>peru</i>
	<i>kizu</i> ‘wound’	:: * <i>kezu</i>

¹² Whitman (1985:57), working within the 4VH, proposes instead that these correspondences reflect pJ **ə*, and that an umlaut-like process raised and fronted **ə* before **C_{son}(C)i*, where *C_{son}* represents a sonorant. He suggests that in the same environment, **ə* was fronted to **e* in pR, where it fell together with secondary **e* from the sources discussed below. pJ **ə* > OJ *o* elsewhere and the chief argument for this hypothesis is a distributional gap: OJ *o* is rare before *C_{voi}i*. However, Miyake (2003) cites Serafim (1999) for a number of forms that counter-exemplify Whitman’s umlaut hypothesis and thus strengthen the Hattori/Thorpe argument for pJ **e*. These include OJ *sima* ‘island’ :: pR **sema* and OJ *sita* ‘tongue’ :: pR **seta* (Serafim’s reconstructions). These reconstructions, too, if correct, indicate that some instances of pR **e* :: OJ *i* must reflect a pJ vowel other than **i* or **ə*.

idu ‘wh-’ :: **edu*
pidi ‘elbow’ :: **pedi*

OJ internal evidence for **e* may be found in variation between *-ye* and *-i* within OJ (21).

(21) **me* *mye* ‘woman’ ~ *-mi-* (*womina* ‘young woman’)
**e* *ye* ‘placenta’ ~ *i-* (*iro* ‘of same mother’)

Examples of this sort are not as numerous as the examples we saw in (17) showing final **o* reflected as *-u* word-internally, but the pattern is the same. Cross-dialectal correspondences within OJ show some of the variation parallel to the correspondences between pR and OJ. These include well known cases of Central OJ *-i* : Eastern OJ *-e* (23), where the Eastern varieties pattern like pR. Thus, for example, the adjectival adnominal ending Eastern *-ke* (:: Central *-ki*) may be considered to reflect a pre-raising stage (cf. also Hino 2003); note there is also a single example of Central OJ *-kye* for *-ki* of this auxiliary in an early poem (K 32).

(22)	Central OJ		Eastern OJ
	<i>*te</i>	(<i>mi</i>) <i>ti</i> ‘road’	:: <i>te</i> (<i>miti-no-naga-te</i>)
	<i>*ke</i>	<i>-ki</i> adj.Adn.	:: <i>-ke</i>

There are also, however, examples of COJ *-ye* :: EOJ *-i* . These examples highlight the fact that different dialects had different criteria for full and partial raising, as we saw in the case of **o* > OJ *u, wo*.

(23)	Central OJ		Eastern OJ
	<i>*ipe</i>	<i>ipy</i> e ‘house’	:: <i>ipi</i>
	<i>*kaper-</i>	<i>kapyer-</i> ‘return’	:: <i>kapir-</i>

We propose that pJ **e* raised in the same environments as **o*, merging with **i* in all positions except final, where it became OJ *-ye*. Thus pJ **e* is reflected as COJ *-ye* in **me* > *mye* ‘woman’, but nonfinally this morpheme surfaces as *-mi-*, as in *womina* ‘young woman’ (*wo-* ‘small’, *na* ‘person’) or *omina* ‘old woman’ (*o-* prefix, cf. also *okina* ‘old man’, probably related to *oi-* ‘age, get old’).

3.3 *The environment for mid vowel raising*

Several proposals have been made about the environments and conditions for MVR. Serafim (1999) and Miyake (2003) advance the view that pJ **e*, **o* raised in pre-OJ and merged unconditionally with **i*, **u* respectively, giving OJ *i*, *u*, but on this view, all instances of OJ *-ye* and *-wo* are secondary, resulting from contraction, and as we noted above (§2.4) virtually none of these instances have clear etymologies involving contraction. We therefore adopt the view that the outcome of MVR was variable and conditioned, as proposed by Hattori (1976), NSNT (and Hayata 1998, 2000 for **o*). Hattori and Hayata propose that length was the crucial conditioning factor for raising. Under this view, short **e* and **o* merged with **i*, **u*, but long **ee* and **oo* gave OJ *ye* and *wo*.¹³ There are a number of problems with this view. First, there seems to be no independent evidence for the long vowel in virtually all of the relevant cases. Second, there is no explanation for why we find no *CyeCV* < **CeeCV* or *CwoCV* < **CooCV*. Finally, the condition seems to imply that all word final vowels were lengthened; although allophonic lengthening is plausible for monomoraic words in phrase-final position, as in modern Kansai dialects, it is difficult to find supporting evidence for lengthening of all final vowels. Without invoking vowel lengthening or long vowels it seems likely, and simpler to assume, that raising followed the course of a gradual phonetic diphthongization, eventually giving a high vowel, except in certain environments where the diphthongal realization was phonemicized as a diphthong, see (24), thus distinguishing between ‘partial’ (**o*, **e* > *wo*, *ye*) and ‘full’ (**o*, **e* > *u*, *i*) raising as detailed in §3.

(24) /**o*/ > [^w*o* > *u*] > /*wo*, *u*/
 /**e*/ > [^j*e* > *i*] > /*ye*, *i*/

As outlined in §3 and as noted throughout §3.1 and §3.2, we propose that ‘final position’ is the core environment for partial raising (i.e., the environment in diphthongal realization was phonemicized as a diphthong), although the details are not yet clear. As we saw in §§3.1-2, there seems to have been variation between (and possibly also within) dialects with regard to the domains with reference to which ‘final’ was defined, especially for inflecting word classes. In COJ ‘word final’ seems to have been the main environment for partial raising, but for example, the well-known contrast between COJ *sugus-* ‘pass it’ and EOJ *sugwos-* ‘id.’ indicates that in the latter dialects root- or stem-final, rather than word-final, position was the environment for **o* > *-wo*. More generally, examples of non-final partial raising are found before fairly transparent boundaries, e.g., COJ *sukwo-si* ‘a little’ (attested with *-wo* in the *Shinsen jikyô* 892) (~ *sukuna-* ‘few’) < **su/oko* + **-si* ‘adverbial

¹³ This is simply tantamount to proposing that the first half of a long vowel raised (**e* > *i*, **o* > *u*, but **ee* > *ie*, **oo* > *uo*), although not described in that way by Hattori and Hayata.

formant’, or EOJ *te-kwo/te-gwo* ‘baby’ (~ *ti-go* ‘id.’, attested in EMJ) < **te* ‘breast, suckle’ + **ko* ‘child’. Such forms exhibit partial raising either because the word structure remained transparent until after MVR was complete and the variety in question had morpheme-, or lexical morpheme, final as the environment for partial raising (e.g., **te-ko* > **tye-kwo* > *te-kwo* > *tekwo*), or because the derivative or compound itself was only formed after the completion of MVR (e.g., **te* > **tye* > *te*, then *te+kwo* => *tekwo*). It is finally possible that some dialects had unconditional full raising, i.e., merger of **i* and **e* as *i*, and **u* and **o* as *u*, or at least very narrowly defined environments for partial raising. Forms such as *ipi* and *kapir-* cited in (24) above may represent such dialects.¹⁴

3.4 Mid Vowel Raising and Demonstratives.

As is well known, the Japanese and Korean systems of demonstrative/interrogative pronouns are structurally identical, with a three way distinction among the demonstratives (25). There is also a good form fit between MK *k*↗, *tyə*, *ə-* ~ *ənu* and EMJ *ko*, *so*, *i-* ~ *idu-*. However, the semantics do not match in two of the sets: MK mesial *k*↗ corresponds to EMJ proximal *ko*, and MK distal *tyə* to EMJ mesial *so*. In this section we show how raising of pJ mid vowels offers support for a scenario of changes which has led to this situation.

(25)	proximal	mesial	distal	interrogative
MK	<i>i</i>	<i>k</i> ↗	<i>tyə</i>	<i>ə-</i> ~ <i>ənu</i>
EMJ	<i>ko</i>	<i>so</i>	<i>ka</i>	<i>i-</i> ~ <i>idu-</i>

Previous researchers have pointed out that there is some evidence for an earlier additional demonstrative *i-*, retained as the OJ personal pronoun *i* (second person), and also in *ima* ‘now’ < *i* ‘this’ + *ma* ‘interval’. If *i* does reflect an earlier demonstrative, its function would be proximal. This raises the possibility that OJ *i* descends from an earlier proximal pronoun, which was displaced by the ancestor of OJ *ko*.

There are internal and typological arguments for this view. First of all, it is less well known that the EMJ system is an innovation and that OJ in fact, as demonstrated in detail by Hashimoto (1966, 1982), only has a two term plus interrogative system : *ko* (speaker), *so* (non-speaker), *i-* ~ *idu-* (interrogative) (26). Most accounts posit a three term ‘proximal - mesial - distal’ system for OJ, built on *ko* - *so* - *ka*. However, there is no evidence within OJ of *ka*

¹⁴ Other apparent examples of final full raising may have a different origin, namely as having been extracted from non-final position in compounds after completion of full raising. A likely example is *ti* ‘breasts’ which in OJ mainly occurs as the first member of a compound, but which from EMJ onwards is found as an independent noun; see §5.2.

being a productive member of the demonstrative system. Two forms are attested in OJ: long *kare* is found once (MYS 18.4045); what may be taken to be short *ka*, as distinct from the adverb *ka* ‘this way’, is attested at most twice, both in dialect poems (MYS 14.3565, 20.4384). While these forms most likely represent the budding of the distal demonstrative which is so frequent in EMJ, they clearly did not form a productive integrated part of the OJ system of demonstratives. Other *ka-* based forms often cited are actually attested only from EMJ. Hashimoto’s study is the first to consider the OJ system on its own merits, rather than in terms of the EMJ system. He shows that the *ko-* versus *so-* system is entirely speaker based, with no primary reference to the hearer. ‘Speaker’, *ko-*, refers to what is within the speaker’s domain of direct sensory perception, or experience. ‘Non-speaker’, *so-*, refers to what is outside of the speaker’s domain of direct experience.

(26) OJ demonstratives	speaker	nonspeaker	interrogative
	<i>ko</i>	<i>so</i>	<i>i- ~ idu-</i>

We believe on the basis of comparative Ryûkûan evidence that the pJ shape of the interrogative was **e-*, and we have mentioned above evidence for an earlier proximal demonstrative **i-*. We propose that the OJ system evolved from the pJ system in (28). What upset the pJ system, we suggest, was raising of pre-OJ */*e/*, resulting in homonymy between pre-OJ proximal **i* (< pJ **i*) and interrogative **i* (< pJ **e*), a stage we represent as pre-OJ(a).¹⁵ We suggest that this ‘pernicious homonymy’ was resolved by proximal **i* being discarded as a productive member of the system and mesial **k*↗ and distal **s*↗ being reanalyzed as speech event participant and nonparticipant respectively: pre-OJ(b). This was subsequently reanalyzed as speaker and nonspeaker, respectively, the system exhibited by OJ. The OJ system was later augmented to include a distal (*ka*), resulting in the EMJ and later system.

(27) pJ	proximal	mesial	distal	interrogative
	<i>*i</i>	<i>*k</i> ↗	<i>*s</i> ↗	<i>*e-</i>
pre-OJ(a)	proximal	mesial	distal	interrogative
	<i>*i</i>	<i>*k</i> ↗	<i>*s</i> ↗	<i>*i-</i>
pre-OJ(b)	participant	nonparticipant	interrogative	
	<i>*k</i> ↗	<i>*s</i> ↗	<i>*i-</i>	

¹⁵ Note that interrogative **e-*, like OJ *i-*, was not used in isolation, but always with following morphological material, and that **e-* therefore raised fully to > *i-*.

OJ	speaker		nonspeaker		interrogative
	<i>ko</i>		<i>so</i>		<i>i- ~ idu-</i>
EMJ	proximal	mesial	distal		interrogative
	<i>ko</i>	<i>so</i>	<i>ka</i>		<i>i- ~ idu-</i>

The pJ demonstrative system we reconstruct is a good fit, phonologically and semantically, with the MK system (28):

(28)		proximal	mesial	distal	interrogative
	MK	<i>i</i>	<i>k</i> ↗	<i>tyə</i>	<i>ə</i>
	pJ	* <i>i</i>	* <i>k</i> ↗	* <i>s</i> ↗	* <i>e</i>

Internal Japanese support for the reconstruction of OJ proximal *ko-* as reflecting an earlier mesial term is provided by a number of deictic OJ time words with initial *k-* (including both *ko-* and a variant *ki-*) (30), in three sets: (a) proximal, (b) ambiguous, and (c) non-proximal. The deictic difference between (a) and (c) and the ambiguity within (b) is easily understood on the hypothesis of a deictic shift in pre-OJ from mesial to proximal of the ancestor of OJ *ko-*, with the (b) and (c) forms reflecting the pre-shift stage.

- (29) a. *kyepu* ‘today’, *kyesa* ‘this morning’;
 b. *kozo* ‘tonight; last night’,¹⁶ *koyopi* ‘tonight; last night’;
 c. *kozo* ‘last year’, *kiso/kizo* ‘last night’, *kinopu* ‘yesterday’.

More generally, the shift of the mesial in a three term demonstrative system to replace the original proximal is attested for example in the development of the Latin demonstrative system. The Classical Latin three term system had proximal *hic* (*haec, hoc*) ‘this’ (close to speaker), mesial *iste* (*ista, istud*) ‘that’ (close to hearer), and distal *ille* (*illa, illud*) ‘that yonder’ (distant from speaker and hearer). In the transition to medieval Latin, the original mesial term *iste* came to replace proximal *hic* in its function (Väänänen 1963:128-9).

4. Two sources for o: pJ *-ə, *-ǝ > OJ -o

¹⁶ This word is in Zdb translated as ‘tonight’, noting that it may also mean ‘last night’, whereas Ohno 1974 translates it simply as ‘last night’; these differences provide a good illustration of the ambiguity involved.

A main feature of our seven vowel hypothesis is that we propose that OJ *o* is the reflex of two pJ vowels, which we reconstruct as * \varnothing and * ∂ , a high and a mid central vowel respectively. The basic observation underlying this proposal is evident in (7a) and (8) above: OJ *o* takes part in two alternations, $-o \sim -(w)i$ and $-o \sim -e$,¹⁷ reflecting two outcomes of contraction with **i* or **y*. Of these, $-o \sim -e$ has received little serious attention, but has usually been thought to be an irregular and/or nonstandard alternative to $-o \sim -(w)i$. There is, however, little support for this view and in this section we argue on the contrary that $-o \sim -e$ simply reflects a different source for OJ $-o$ than that reflected in $-o \sim -(w)i$; and more generally that each of the alternations in (7a) and (8) above reflect a different pJ vowel, as shown in (30). Pre-OJ * \varnothing and * ∂ merged after these contractions took place, eventually giving OJ $-o$.

- (30) a. * $-u > -u \sim -(w)i < *-ui$
 b. * $\varnothing > -o \sim -(w)i < *-\varnothing i$
 c. * $\partial > -o \sim -e < *-\partial i$
 d. * $-a > -a \sim -e < *-\partial i$

Below we review relevant cases from lexical contraction (§4.1), nominal apophony (§4.2), and verbal apophony (§4.3), of the two alternations $o \sim e$ and $o \sim wi$ in support of our interpretation.

4.1 Lexical contraction

The examples in (31) seem to be the only solid cases of lexical contraction of an ancestor of OJ *o* with **i*. They do not show $-wi$ to be the regular and $-e$ the irregular outcome: there is only one case of the former versus two of the latter. We interpret this to mean that OJ *o* had two pJ sources which gave different outcomes when contracted with **i*, as shown in (32). On this proposal, pre-OJ * \varnothing and * ∂ merged after these contractions had taken place and eventually gave OJ *o*, hence OJ *opo-*, *tono*, *wo-*.

- (31) a. *opwisi* Proper name < *opo+isi* ‘big+stone’
 b. *toneri* ‘palace servant’ < *tono+iri* ‘palace+enter’
wenu ‘puppy’ < *wo+inu* ‘small+dog’

- (32) a. * $\varnothing p - > opo- \sim opwisi < * \varnothing p \varnothing -isi$

¹⁷ $-wi$ is neutralized as $-i$ after coronals (*t, d, s, z, n, r*), hence the notation $-(w)i$.

- b. **tənə* > *tono* ~ *toneri* < **tənə-iri*
 **wə-* > *wə-* ~ *wenu* < **wə-inu*

4.2 Nominal Apophony

As mentioned, apophonic alternations such as those in (8) are thought to originate in contractions of an older shape with *-i or *-y (possibly going back further to a consonant). OJ monosyllabic apophonic nominal stems are listed in (33).

(33) a. *Cu* ~ *C(w)i*

- ku-* ‘fortress’ ~ *kwi*
- mu-* ‘body’ ~ *mwi*
- nu* ‘bead’ ~ *ni*
- tu* ‘miscantus reed’ ~ *ti* ‘chigaya’
- turi* ‘hook’ ~ *ti*

b. *Co* ~ *C(w)i*

- ko-* ‘tree’ ~ *kwi*
- no-* ‘load’ ~ *ni*
- po-* ‘fire’ ~ *pwi*
- so* nonspeaker demonstrative ~ *si* 3rd sg. personal pronoun

c. *Co* ~ *Ce*

- mo* ‘algae; ?seaweed’ ~ *me* ‘(edible) seaweed’
- so-* ‘back’ ~ *se*
- yo* ‘branch’ ~ *ye*
- yo-* ‘good’ ~ *ye-*

d. *Ca* ~ *Ce*

- ka-* ‘hair’ ~ *ke*
- ka-* ‘day(s)’ ~ *ke*
- ma-* ‘eye’ ~ *me*
- sa-* ‘narrow’ ~ *se-*
- ta-* ‘hand’ ~ *te*

Not all these cases of apophony involve variation between a compound and a free variant, but it is clear that also here OJ *o* takes part in two alternations: *-o- ~ -(w)i* and *-o- ~ -e*. Again, the distribution is not persuasive for *-o ~ -(w)i* being regular and *-o ~ -e* being irregular. Instead, we interpret this to support our proposal that OJ *o* has two sources, $*\varnothing$ and $*\partial$, and that the forms above should be reconstructed and understood as (34), with $*\varnothing$ and $*\partial$ merging (eventually to give OJ *o*) at some point after the contractions had taken place.

(34)

b.	‘tree’	$*k\varnothing$	$>$	<i>ko-</i>	\sim	<i>kwi</i>	$<$	$*k\partial i$
	‘load’	$*n\varnothing$	$>$	<i>no-</i>	\sim	<i>ni</i>	$<$	$*n\partial i$
	‘fire’	$*p\varnothing$	$>$	<i>po-</i>	\sim	<i>pwi</i>	$<$	$*p\partial i$
	‘dem.’	$*s\varnothing$	$>$	<i>so-</i>	\sim	<i>si</i>	$<$	$*s\partial i$
c.	‘seaweed’	$*m\partial$	$>$	<i>mo</i>	\sim	<i>me</i>	$<$	$*m\partial i$
	‘back’	$*s\partial$	$>$	<i>so</i>	\sim	<i>se</i>	$<$	$*s\partial i$
	‘branch’	$*y\partial$	$>$	<i>yo</i>	\sim	<i>ye</i>	$<$	$*y\partial i$
	‘good’	$*y\partial$	$>$	<i>yo-</i>	\sim	<i>ye-</i>	$<$	$*y\partial i$

The picture is different if we look at polysyllabic stems: Among polysyllabic stems we only find the three alternations OJ *CVCu- ~ CVC(w)i*, *CVCo- ~ CVC(w)i* and *CVCa- ~ CVCe*; but not *CVCo- ~ CVCe* (the alternation which supports our proposal of two pJ sources for OJ *o*).

(35) a. *CVCu- ~ CVC(w)i*

kamu- ~ kamwi ‘god’
kuku- ~ kukwi ‘stem’
kuru- ~ kuri ‘chestnut’
kutu- ~ kuti ‘mouth’
satu- ~ sati ‘hunting device (or its spiritual power)’
tuku- ~ tukwi ‘moon’
tuku- ~ tukwi ‘zelkova (tree)’

b. *CVCo- ~ CVC(w)i*

yomo- ~ yomwi ‘Hades’
woto- ~ woti ‘distant place’

c. *CVCa- ~ CVCe*

ama- ~ *ame* ‘rain’
ama- ~ *ame* ‘sweet’
kana- ~ *kane* ‘metal’
kaza- ~ *kaze* ‘wind’
muna- (not phonographically attested in OJ) ~ *mune* ‘ridge’
muna- ~ *mune* ‘breast’
mura- ~ *mure* (not as noun in OJ; <= S2 *mure-* ‘to flock’) ‘crowd’
napa- ~ *nape* ‘seed’
puna- ~ *pune* ‘boat’
suga- ~ *suge* ‘sedge’
saka- ~ *sake* ‘wine’
taka- (not phonographically attested in OJ) ~ *take* ‘bamboo’
tana- ~ *tane* ‘seed’
tuma- ~ *tume* ‘claw’
upa- ~ *upe* ‘top’

It is clear that *CVCa-* ~ *CVCe* is fairly frequent (the listed forms are taken JLTT:48),¹⁸ but examples of both *CVCu-* ~ *CVC(w)i* and in particular *CVCo-* ~ *CVC(w)i* are few. The absence of examples of nominal apophony of the shape *CVCo-* ~ *CVCe* may therefore not be significant for our argument.

4.3 *Verb(al) Apophony*

Nidan verbs (see §2.3) do not regularly take part in the alternation *-o-* ~ *-e-*, but it is possible to identify a few pairs exhibiting this alternation, see (36). Of these, only *komor-* ~

¹⁸ Martin cites one more pair, *kowa-* ~ *kowe* ‘voice, sound’. Note, however, (a) that *kowa-* is attested only from EMJ, and then in variation with *kowe-* in compounds (e.g. *kowe-daka*, *kowa-daka* ‘loud’), whereas only *kowe* ‘voice’ is found in OJ, and (b) that OJ **kowa-* would have violated Arisaka’s Law (see §6.1 below). This suggests that *kowa-* is not an original *hifukukei*, but in fact an ‘analogical’ EMJ formation, based on the alternation among the forms of the older apophonic nouns. In that case, *kowe* could be from **kəwe/kʷwe* with mid vowel raising (> **kəwe/kʷwe* > *kowe*), but it is more likely that it goes back to a consonant final form: **kəwər* > (yodization, cf. §2.3 above) **kəwəy* > (contraction) **kəwe* > *kowe*; **kəwər* would then also, with an epenthetic echo vowel conforming to phonotactic surface constraints, be reflected in the reduplicated mimetic OJ *koworo-koworo* ‘rolling, rumbling noise’.

kome- and *yosor-* ~ *yose-* are true lexical verbs.¹⁹ *Se-* is the (derived) basic stem of ‘to do’; the OJ reflex of its root, *so*, is used in the prohibitive construction *na-...-so* as a fossilized imperative. *-Kose-* is the (derived vowel) base of a defective optative auxiliary verb, the OJ reflex of its root, *koso*, being used as its imperative of that verb; the root is also most likely reflected in the focus particle *koso*. *To* is the copula infinitive and *-te* the derived adverbializer and gerund formant. (See Frellesvig 2001:12-16 about *kose*, *koso*, *te*, *to*.)

- (36) *komor-* ‘hide’ (INTR) ~ *kome-* ‘hide, enclose’
yosor- ‘be drawn near’ ~ *yose-* ‘bring close’
so ~ *se-* ‘do’
-koso ~ *-kose-* ‘do for me’
to copula ~ *te* adverbializer.

These forms seem to be relevant examples of *o* ~ *e* in verb(like) derivation and we suggest that they should be reconstructed as in (37).

- | | | | | | |
|------|-------------|-------------------|---------------|---|--------------------------------|
| (37) | ‘hide’ | <i>*kəmə-r-</i> > | <i>komor-</i> | ~ | <i>kome-</i> < <i>*kəmə-Y</i> |
| | ‘close’ | <i>*yəśə-r-</i> > | <i>yosor-</i> | ~ | <i>yose-</i> < <i>*yəśə-Y</i> |
| | ‘do for me’ | <i>*kəsə-</i> > | <i>-koso</i> | ~ | <i>-kose-</i> < <i>*kəsə-Y</i> |
| | ‘do’ | <i>*sə</i> > | <i>so</i> | ~ | <i>se-</i> < <i>*sə-Y</i> |
| | ‘copula’ | <i>*tə</i> > | <i>to</i> | ~ | <i>te</i> < <i>*tə-Y</i> |

4.4 Summary

In this subsection we have presented evidence internal to OJ for two pJ sources of OJ *o*. Overall this evidence consists in the existence of alternations between OJ *o* ~ *e* in addition to the more widely recognized *u* ~ (*w*)*i*, *o* ~ (*w*)*i* and *a* ~ *e*. Although *o* ~ *e* has a more limited distribution, especially numerically, than the other three, it is found in the same three general contexts: lexical contraction (§4.1), nominal apophony (§4.2), and verbal apophony (§4.3). We thus suggest that these four alternations within OJ reflect contractions of four different pJ vowels with a following **i*, as in (30) above, repeated as (38).

- (38) **u* > *u* ~ (*w*)*i* < **ui*

¹⁹ Note that *kome-* has a Kami Nidan variant which is attested only once, in the compound *tuma-gomwi-* ‘spouse-hiding, spouse-enclosing’ (Kojiki, song 1). Also, the relationship of the verb *yor-* ‘approach’ to *yosor-*, *yose-* and *yos-* ‘bring close’ is not clear.

*ɨ >	o	~	(w)i	<	*ɨi
*ə >	o	~	e	<	*əi
*a >	a	~	e	<	*ai

5. Supporting comparative evidence

The seven vowel system we have reconstructed for proto-Japanese has the same number of vowels as the system of Middle Korean. (39) shows the vowels of LMK (15th century), with transcription following Lee (1972a) and with Hankwul and Yale Romanization (following Martin 1992) in brackets.

(39)	i (◦ i)	ɨ (◦ u)	u (◦ wu)
		ə (◦ e)	o (◦ wo)
		· (◦ o)	
		a (◦ a)	

In this section we point out that many of the vowel correspondences proposed for Japanese and Korean in previous literature can be explained rather straightforwardly by the 7 vowel system we reconstruct for proto-Japanese. OJ *o* has been noted to correspond both to MK ❖ and ◦. This is explained by our hypothesis that OJ *o* results from the merger of pJ *❖ and *ə. MK *o* has been noted to correspond to OJ *u* and (more rarely) *w**o*. This is explained by the hypothesis of mid vowel raising. Matters are slightly more complex in the case of our reconstructed pJ mid front vowel **e*, but here too we show that the observed correspondences support the hypothesis of Japanese mid vowel raising.

We propose the basic vowel correspondences in (40) (see also Frellesvig and Whitman 2006):

(40)	MK	::	OJ	<	pJ
	a. i		i		*i
	b. ❖		o		*❖
	c. u		u		*u
	d. yə < *e		i, ye		*e
	e. ə		a, o		*a, *ə
	f. o		u, wo		*o
	g. ·		o		*ə
	h. a		a		*a

5.1 Correspondences of OJ *o*.

Previous researchers have observed that OJ *o* shows correspondences with Middle Korean ❖ (40b), and also with MK \cdot (Kôno 1967/1979:561, Martin 1966:220) (40g). The two correspondences are exemplified in (41-2):

(41)	MK ❖	::	OJ <i>o</i>	pJ *❖
a.	$k\text{❖}$ ‘that’ (mesial)	::	<i>ko</i> <	$\text{*}k\text{❖}$ (mesial, cf. §3.4)
b.	$m\text{❖}s\text{❖}$ - ‘wh(at)’	::	<i>mosi</i> <	$\text{*}m\text{❖}si$ (conjunctural adverb)
c.	$t\text{❖}m\text{❖}r$ - ‘rare’	::	<i>tomo-si</i> <	$\text{*}t\text{❖}m\text{❖}$ - ‘scarce’
d.	$t\text{❖}r$ - < $\text{*}t\text{❖}r\text{❖}$ - ‘hold, lift’	::	<i>tor</i> - <	$\text{*}t\text{❖}r$ - ‘take’
e.	$m\text{❖}r\text{❖}$ - <i>t</i> ‘all’	::	<i>moro</i> - <	$\text{*}m\text{❖}r\text{❖}$ ‘all’

(42)	MK \cdot	::	OJ <i>o</i>	pJ ə
a.	$t\text{·}rk$ ‘chicken’	::	<i>tori</i> <	$\text{*}t\text{ə}r(-i)$ ‘bird’
b.	$k\text{·}th$ - ‘alike’	::	<i>koto-si</i> <	$\text{*}k\text{ə}t\text{ə}$ - ‘resemble’
c.	$m\text{·}t$ ‘eldest, chief (of kin)’	::	<i>moto</i> <	$\text{*}m\text{ə}t\text{ə}$ ‘base, origin’
d.	$n\text{·}r$ - ‘fly’	::	<i>nor</i> - <	$\text{*}n\text{ə}r$ ‘ride’
e.	$k\text{·}\beta\text{·}r$ ‘county’	::	<i>kopori</i> <	$\text{*}k\text{ə}\beta\text{ə}r(-i)$ ‘county’ (loan)

These two correspondence sets show us that two pJ vowels were involved: OJ *o* < pJ *❖ :: MK ❖ , and OJ *o* < pJ *ə :: MK \cdot , providing strong comparative support for our reconstruction of two central pJ vowels.

The same two sets of correspondences are observable in OJ correspondences for MK words in final /r/ or /y/:

(43)	MK $\text{❖}r$		OJ <i>o</i> -, <i>wi</i>	pJ $\text{*❖}y$
a.	$p\text{❖}r$ ‘fire’	::	<i>po</i> -, <i>pwi</i> <	$\text{*}p\text{❖}y$ ‘fire’
b.	$m\text{❖}r\text{❖}$ - < $\text{*}m\text{❖}r\text{❖}r$ - ‘withdraw’	::	<i>mo</i> , <i>mwi</i> - <	$\text{*}m\text{❖}y$ - ‘turn about’
c.	$k\text{❖}r\text{❖}h$ ‘stump, counter for trees’	::	<i>ko</i> -, <i>kwi</i> <	$\text{*}k\text{❖}y$ ‘tree’

(44)	MK $\text{·}r$		OJ <i>o</i> , <i>e</i>	pJ $\text{*ə}y$
a.	$m\text{·}r$ ‘seaweed’	::	<i>mo</i> , <i>me</i> <	$\text{*}m\text{ə}(y)$ ‘seaweed’
b.	$p\text{·}yam$ ‘snake’	::	<i>pemi</i> <	$\text{*}p\text{ə}y\text{m}(-i)$ ‘snake’
c.	$p\text{·}y$ ‘boat’	::	<i>pe</i> <	$\text{*}p\text{ə}y$ ‘boat (prow)’

The two sets of correspondences in (43-4) involve apophonic nouns in the pattern discussed earlier in (34): although the *hifukukei* (where observable) is OJ *o* in both (43) and (44), there are two distinct *roshutsukei* or free forms, *wi* and *e*. In (38) we derived this distinction from pre-OJ diphthongs with two distinct nuclei: * $\text{ɰ}y$ and * $\text{ə}y$. The Korean comparanda confirm the internally motivated reconstructions of these two distinct nuclei. Thus the OJ apophonic nouns ‘fire’ and ‘tree’ in (43) behave exactly as predicted on internal grounds by the hypothesis that their nuclear vowel reflects pJ * ɰ , and this is confirmed by the Korean comparisons. Likewise, the vowel in (44a) alternates between *e* and *o*, suggesting on internal grounds the pJ shape * $m\text{ə}(y)$ confirmed by the Korean comparison. Note that the comparisons in (43) and (44a) also provide comparative confirmation for Murayama’s hypothesis that in some instances the final segment accounting for the *roshutsukei* in apophonic nouns may have been **r*. Again, these correspondences provide strong support for reconstructing two central pJ vowels. And conversely, our overall proposal makes possible a simple, systematic, and coherent account of a number of correspondences which hitherto have appeared contradictory and competing.

5.2 Correspondences for the pJ mid vowels **e* and **o*

Under the mid vowel raising hypothesis, MK *o* is expected to correspond to OJ *u*, except in final position. This is borne out, as follows:

(45) MK <i>o</i>		OJ <i>-u-</i>		pJ <i>*o</i>
a. <i>kokori</i> ‘stem’	::	<i>kuku-</i> , <i>kukwi</i> , ‘stem’	?<	<i>*ku/okoy</i>
b. <i>oy</i> ‘melon’	::	<i>uri</i> ‘melon’	?<	<i>*u/ori</i>
c. <i>tok</i> ‘jug’	::	<i>tuki</i> ‘jug’	?<	<i>*tu/oki</i>
d. <i>kom</i> < <i>koma</i> ‘bear’	::	<i>kuma</i> ‘bear’	?<	<i>*ku/oma</i>
e. <i>mom</i> ‘body’	::	<i>mu-</i> , <i>mwi</i> ‘body’	?<	<i>*mu/o(C)(i)</i>
f. <i>moyh</i> ‘mountain’	::	<i>mure</i> ‘mountain’	?<	<i>*mu/ore</i>
g. <i>kop-/koβ</i> < <i>kop-</i> ‘beautiful’	::	<i>kupasi-</i> ‘beautiful’	<	<i>*kopV-</i>

MK <i>o</i>		OJ <i>-wo</i>		pJ <i>*o</i>
h. <i>moh</i> ‘corner, side’	::	<i>mwo</i> ‘direction’	<	<i>*mo</i>
i. <i>-to</i> ‘also, too’	::	<i>two</i> ‘outside, other’	<	<i>*to</i>

In (45 a-f) there is no Japanese-internal evidence that non-final /u/ results from pJ **o*; the OJ and pR vowel shape, *u* in a non-final syllable, is consistent with either **u* or **o*. In (g), there is evidence for OJ *u* < **o*: the original vowel is reflected as partially raised *-wo* in the Kami Nidan verb *kwopwi-* ‘love, long for’; it is fully raised in stem-nonfinal position in the adjective *kupasi-* ‘fine, beautiful’. In (h-i), pJ **o* is, as expected, reflected as *wo* in final position.

Matters are more complex with correspondences for OJ *i/ye* < pJ **e*. As shown in (39), the MK inventory lacks a non-high front vowel. We adopt a proposal due to Leon Serafim, brought to our attention by Marc Miyake, that pre-MK had a mid front vowel **e*. This vowel was centralized to merge with *ə* in absolute initial position, but in other positions it diphthongized and left a trace in the form of palatalization of the preceding consonant, resulting in syllables of the shape *Cyə* in MK. This hypothesis accounts for the inordinately large number of syllables of shape *Cyə* in Korean, in comparison with other syllables with a palatalized onset. The predicted OJ correspondence for this vowel is *i* in nonfinal (46a-e), *ye* in final position (f-g):

(46) MK <i>yə</i>		OJ <i>i</i>		pJ <i>*e</i>
a. <i>syəm</i> ‘island’	::	<i>sima</i> ‘island’	?<	<i>*syema</i>
b. <i>syəy-</i> ‘whiten’ cf. <i>h y-/h ❖y-</i> ‘white’	::	<i>sirwo</i> ‘white’	?<	<i>*sero</i>
c. <i>nyəy-</i> ‘go’	::	<i>nige-</i> ‘flee’	?<	<i>*nenk(V)-</i>
d. <i>pyərok</i> ‘flea’	::	<i>piru</i> ‘leech’	<	<i>*peru</i>
e. <i>ə-</i> ‘wh’	::	<i>i-</i> ‘wh’	<	<i>*e-</i> (cf. §3.4)
f. <i>myən ri</i> ‘wife’	::	<i>mye, mi-</i> ‘woman’	<	<i>*me</i>
g. <i>cyəs</i> ‘breasts’	::	<i>ti, titi</i> ‘breasts’	?<	<i>*te</i>

‘Woman’ in (f) gives the clearest evidence of an alternation between fully raised nonfinal OJ *i* and partially raised final *-ye* (cf. §3.2). ‘Breasts’ in (g) is less clear, because central dialect OJ does not attest the predicted **te* in final position. However OJ *ti* ‘breasts’ is attested primarily as the first member of a compound; in examples where it occurs in combination with a verb, it is always immediately preverbal, and this is also in a possible compound environment, making it likely that the independent form *ti* was extracted from non-final position in compounds after full raising: **te-.. > ti-.. > ti*. Japanese-internal support for *ti* < **te* is provided by the Azuma compound *tekwo, tegwo* ‘baby’ which is most plausibly analyzed as a compound of *te-* ‘breast, suckle’ and *kwo* ‘child’, where *te* reflects the original mid vowel as *te* (**tye* < **te*, neutralized partial raising), for either of the possible reasons given in §3.3.

Previous treatments of correspondences like those in (46) have posited primary **i* for both languages (Lee 1959, Whitman 1985). But (d-e) give Ryukyuan evidence for pJ **e* (see also §3.4 about (e)), while (f) gives Japanese-internal evidence for **e*, as it shows *ye*, not *i*, in the position predicted by the mid vowel raising hypothesis. Our proposal provides a more balanced vowel system for pre-MK, by setting up a non-high front vowel, and also makes sense of this correspondence set.

5.3 Pre-OJ Mid Vowel Raising in Japanese Loans from Korean

We know that Japanese absorbed lexical loans from Korean in the early historical period. We might expect some of these to have entered Japanese after mid vowel raising. The mid vowel raising hypothesis enables us to distinguish such later loans from material present in the language at the preraising stage:

(47) MK	J Preraising	Postraising
a. <i>syəm</i> ‘island’	:: <i>sima</i> ‘island’	<i>sema</i> (<i>Nihon shoki</i>)
b. <i>kyən</i> ‘silk’ (Sino-Korean)	:: <i>kinu</i> ‘silk’	<i>ken</i> (SJ)
c. <i>tyər</i> ‘Buddhist temple’	::	<i>tera</i> ‘Buddhist temple’
d. <i>puthə</i> ‘Buddha’ < <i>*put</i> ʔhyə ::		<i>potoke</i> ‘Buddha’
e. <i>Kokuryə</i> ‘Koguryō’ (SK)	:: <i>Kokuri</i> (MJ)	<i>Kure</i> ‘Koguryō’
f. <i>kom</i> < <i>koma</i> ‘bear’	:: <i>kuma</i> ‘bear’	<i>koma</i> Koguryō ethnonym
g. <i>mosi</i> ‘ramie’	:: <i>karamusi</i> ‘(Korean) ramie’	

Doublets such as (a), (b), (e), and (f) are of particular interest. The postraising forms *sema* (island in Paekche material in *Nihon shoki*), *Kure* (designator for Chinese or continental cultural objects), *Koma* ‘Koguryō’ have clear continental connections, and thus are good

candidates for loans into Japanese of a relatively late date, i.e. after raising. The postraising forms *tera* and *potoke* must have been borrowed after the introduction of Buddhism into Japan in the first half of the sixth century which thus provides a *terminus ante quem* for the completion of raising of **e*.

6. Concluding

Recapitulating, we propose the following seven vowel system for pJ.²⁰

- (48) *i *ɨ *u
 *e *ə *o
 *a

Three sets of sound changes applied to that system to give the OJ system in (2) above:

(49) 1. Contraction

	Pre-OJ	OJ
a.	*ui	wi
	*ɨi	wi
	*əi	e
	*ai	e
b.(i)	*i ɨ/*iə	ye
	*ia	ye
(ii)	*u ɨ/*uə	wo
	*ua	wo

2. Merger of central vowels and backing

pJ { *ɨ, *ə } > *ə > OJ o.

3. Mid vowel raising

- a. pJ *e > OJ i, ye
b. pJ *o > OJ u, wo

As for the dating and ordering of these changes, we are at the moment not in a position to offer a fully explicit account, but can say the following:

²⁰ It is worth noting that (48) is a well-attested system. Crothers (1978) reports that vowel system type 7:2 (seven vowels with two of the central vowels /ɨ/, /ə/, or /ü/) is the fifth most common in his sample of vowel systems, and the most common type among seven vowel systems, with 14 exemplars. Three of these are reported with the exact system in (48): Amharic, Sundanese, and Rumanian.

Within (49.1), sets (a), contraction of falling diphthongs, and (b), contraction of rising diphthongs, are very different and should be divorced temporally; as we showed above (§2.2), the (a) set is pervasive, whereas (b) is really very marginal, with only **ia > ye* being found in more than a single example. It seems likely that the (b) set arose as singular or rare phonemicizations, coincidental to MVR when, as a result of MVR, the relevant phonemic strings (/ye, wo/) became available in the system. We would therefore order (49.1.b.i) after (49.3.a) and (49.1.b.ii) after (49.3.b).

As for (49.2), the merger of pJ **ə* and **a* took place after (1.a); equally obviously, the backing of the result of the merger, here tentatively **a*, to give OJ *o* must have taken place after (49.3.b) (or at least after the first moment of that change: **o > [ʷo]*). It is likely that the two parts of MVR were not synchronous: In the central dialect a number of early forms such as *mwokwo* for *mukwo* or *pom-* ‘step’ for *pum-* suggest that raising of **o* was not entirely complete at the beginning of the 8th century. Raising of **e* in the central dialect, on the other hand, seems to have been completed earlier, and singular forms such as *-kye* for *-ki* may be regarded as archaic literary fossils, reflecting the pre-raising stage. As mentioned above, words such as *tera* or *potoke* were probably borrowed shortly after the introduction of Buddhism into Japan, indicating that raising of **e* was complete by the first half of the 6th century. If we are right about the dating of raising of **o*, the type (50.1.a) contractions must have taken place before that, but it is not possible, strictly speaking, to determine the ordering of contraction and raising of **e*.

Proto-Ryukyuan, on the other hand, suggests a contrasting chronology, with pR **e* retained later and in a larger range of environments than pR **o < pJ *o*, and pJ **ɔ̄* and **ə* merged prior to the pR stage. The point, however, is the same: raising of the front and back mid vowels did not necessarily occur synchronously.

The following more detailed ordering for OJ of the component parts of the changes above is consistent with the facts.

(50) 1. Contraction of falling diphthongs

Pre-OJ	OJ
<i>*ui</i>	<i>wi</i>
<i>*ɔ̄i</i>	<i>wi</i>
<i>*əi</i>	<i>e</i>
<i>*ai</i>	<i>e</i>

2. Merger of central vowels:

pJ { **ɔ̄, *ə* } > **ə*

3. Mid vowel raising:

pJ **e* > OJ *i*, *ye*

4. Contraction of rising diphthongs

pre-OJ **ia* > **ya* > *ye*

pre-OJ **iə* > *ye*

5. Mid vowel raising

pJ **o* > OJ *u, wo*

6. Backing of central vowel

pre-OJ **ə* > OJ *o*.

7. Contraction of rising diphthongs

pre-OJ **ua* > **wa* > *wo*

pre-OJ **uo* > *wo*

6.1 *Arisaka's Law and the 7VH.*

The well known 'Arisaka's Law' is the observation that in OJ the syllable *Co* (= *Co*₂) did not occur in a root morpheme with *Ca, Cwo, Cu*. That is to say, there were (almost) no apparently simple words of the structure *CoCa, CoCwo, CoCu* or *CaCo, CwoCo, CuCo*. This has been taken to indicate that Japanese earlier had some form of 'vowel harmony' (in the form of restrictions on the shape of root morphemes rather than the kind of vowel harmony which applies to the concatenation of stem and affix in 'vowel harmony languages' such as Turkish, for example). However, Arisaka's Law never made much sense under the 4VH: first, four vowels seem few for a harmony system of any kind; (51) is the harmony system which would have held if both the 4VH and Arisaka's Law were correct. Second, on the 4VH, OJ *-wo* is always a secondary vowel, < **-u+a* or **-u+ə*; there is thus no scope for OJ *-wo* to take part in Arisaka's law, and furthermore where *(-)Cwo(-)* was etymologized as < **(-)Cuə(-)* < **(-)CuCə(-)* (rather than **(-)CuCa(-)*), the source string itself violated Arisaka's Law. Third, on the 4VH, *-e* (= *-e*₂) is a secondary vowel, deriving exclusively from **-a-i*; thus the majority of OJ *CoCe-* forms are etymologized as < **CəCa(-)(C)i*, with the source string violating Arisaka's Law.

(51)	neutral * <i>i</i>	central * <i>ə</i>	back * <i>u</i> * <i>a</i>
------	-----------------------	-----------------------	--------------------------------------

On the seven vowel hypothesis, on the other hand, Arisaka's Law makes far better sense. It provides better balanced sets of vowels, with the 'central' and 'back' sets in opposition (52), and direct OJ reflexes of the pJ vowels which conform to the observed cooccurrence relations

(54). On our 7VH, OJ *-wo* simply in the overwhelming majority of cases reflects a member of the pJ ‘back’ set, accounting for its lack of cooccurrence with OJ *-o*; and OJ forms of the shape *CoCe* may be simply thought to be < **CəCə(-)i* or **C Cə(-)i*, with no violation of Arisaka’s Law.

(52)	neutral	central	back
	*i	*ɨ	*u
	*e	*ə	*o
			*a

(53)	neutral	A	B
	-i (< *i, *e)		-u (< *u, *o-)
	-ye (< *e)	-o (< * , *ə)	-wo (< *-o)
			-a (< *a)

In conjunction with this, note that the syllable */wo/* has had three distinct pJ sources under the seven vowel hypothesis: **/wə/*, **/wɨ/*; **/wo/*, thus representing a merger of three different vowels. This helps explain why */wo/* appears to escape the effects of Arisaka’s laws, combining both with */o/* (as in *to*