# Morpho-lexical evidence for Proto-Korean-Japanese

**Alexander Francis-Ratté**
**The Ohio State University**

**Abstract**
The hypothesis that Japanese and Korean share a common linguistic origin remains highly controversial, with detractors such as Vovin (2010) criticizing a lack of shared functional morphemes and phonological problems in proposed cognate vocabulary. Expanding on the analysis in Ratte (2015), this paper proposes concrete correspondences in the grammar and vocabulary of Old Japanese to Middle Korean. By analyzing the form and function of verbs in Old Japanese texts (*Man'yōshū*), this paper proposes a functional element \*(w)o- that reveals striking correspondences in Japanese and Korean noun-modifying structures, and a verb \*pə- that shows that identical verb-compounding structures exist in both languages. This paper also provides revisions of four important vocabulary items in previous literature whose correspondences have proven problematic, and improves the strength of the common origin hypothesis by proposing new cognate pairs. In so doing, this paper seeks to address deficiencies in the theory of Korean-Japanese common linguistic origin in order to demonstrate that the theory of proto-Korean-Japanese is stronger than has previously been assumed.

# 1. Introduction

Critics such as Vovin (2010) point to several problems in the theory that Korean and Japanese share a common linguistic origin. The first problem is that there is an insufficient number of convincing correspondences between Japanese and Korean paradigmatic morphology. Although there are many structural parallels between Old Japanese (OJ) and Middle Korean[1] (MK) grammar, formal correspondences in grammatical morphemes have been found wanting by scholars such as Vovin (2010). By contrast, some of the strongest evidence for the common origin of Indo-European languages such as Latin and Sanskrit can be found in close and interlocking correspondences in inflectional morphology, such as shared 3rd person singular endings in -t- and 3rd person plural endings in –nt-.

The second problem is that some of the Korean-Japanese lexical etymologies proposed by Martin (1966) and Whitman (1985) are semantically close but phonologically irregular. Unlike semantic changes, which are often culture-specific, sound changes are widely held to be conditioned by phonological environments alone and thus apply regularly throughout a language (Hale 2003). As a consequence, sound changes give rise diachronically to regular sound correspondences between related languages. The most important criterion therefore in evaluating the soundness of Japanese-Korean etymologies is not whether words correspond in meaning, but whether words correspond regularly in the sounds of their words. Building on the work in Ratte (2015), this paper addresses weaknesses in the hypothesis of Korean-Japanese common origin pointed out by Vovin (2010), and proposes new comparative evidence of lexical and morphological correspondences in support of the proto-Korean-Japanese (pKJ) hypothesis.

## 1.1. Previous Accounts and Literature

The two major works that have exercised the greatest influence on the Korean-Japanese relationship are Martin (1966) and Whitman (1985). Martin (1966) is widely acknowledged as a major work focused solely on establishing a common origin for the Japanese and Korean languages, coming up with a list of over 300 systematic lexical correspondences that he considers valid. Whitman (1985) builds on Martin (1966) by providing a more streamlined and reasonable set of Korean-Japanese phonological correspondences, as well as providing more potential cognates, such as 'sky' (MK *hanolh* 'sky' ~ OJ *swora* 'id.') and 'eye' (MK *meli* 'head' ~ OJ *mey* 'eye'). However, our understanding of the proto-Japanese vowel and consonant systems has changed considerably since 1985, changes which have rendered some of the correspondences and cognates in Martin (1966) and Whitman (1985) obsolete. Some cognates could be misidentified loans. Vovin (2010) once favored a Korean-Japanese relationship but has

---

[1] This paper compares vocabulary and functional elements in the earliest Japanese and Korean texts from which reliable phonological information can be obtained, namely the Old Japanese period (8th century C.E.) and the Late Middle Korean period (1446 – 1592 C.E.). The invention of Hangul, the Korean alphabet, marks the beginning of the Late Middle Korean period, and the use of the new alphabet means that texts from this period are phonologically transparent. Late Middle Korean is preceded chronologically by Early Middle Korean (10th century – 1446 C.E.) and Old Korean (7th – 10th centuries C.E.), and texts are extant from these periods (e.g. the Hyangga or 'Country Songs'). However, pre-Hangul texts are written with Chinese characters in a style that makes them difficult to decipher, and there is agreement among scholars that such texts do not provide the level of phonological detail necessary for accurate reconstruction (Nam 2012). Although there is a chronological difference between Old Japanese and Late Middle Korean, these languages are compared in relation to their hypothetical common ancestor, proto-Korean-Japanese, so chronological differences are not methodologically problematic; note that the differences are much greater in the comparison of Indo-European languages, such as Gothic (mid-first millennium C.E.) and Vedic Sanskrit (2nd millennium B.C.E.).

now become its most staunch opponent. It has leveled a great deal of criticism at Whitman's comparisons, and concludes that most of the lexical correspondences are early loans into Japanese. Unger (2009), on the other hand, disagrees with Vovin (2010) about the strength of the hypothesis. However, it nevertheless admits that a Korean-Japanese relationship requires better evidence than we currently have to be fully convincing. Thus even proponents of a Korean-Japanese relationship agree that more work needs to be done before any relationship can be seen as established.

To address weaknesses in the theory of common origin, this paper proposes new Korean-Japanese etymologies, including two specific, interlocking correspondences in important grammatical morphemes in Old Japanese (8th century C.E., e.g. *Man'yōshū*) and Late Middle Korean (15th & 16th centuries C.E.). In what follows, Section 2 proposes a correspondence of MK *pwó* - 'sees' to OJ *pe-* 'time passes' based on an analysis of the form and function of *pe-* in Old Japanese texts, which reveals an interlocking set of cognates built from shared proto-Korean-Japanese verb compounds. Section 3 proposes a correspondence of MK *wó* - 'comes' to OJ *wi-* 'comes to a stop' and *wor-* 'is present,' from which I reconstruct a pKJ source verb *(w)o- meaning 'comes'. This verb *(w)o- is employed in Korean to mark semantically active or 'modulated' clauses, and from this I argue that the proto-Japanese adnominal suffix *-or, which displays active syntactic alignment, may originally incorporate this verb *(w)o- 'comes' as the morphological source of its active semantics and syntax. This analysis thus reveals a striking correspondence between noun modification structures in Japanese and Korean. Section 4 proposes four revisions of important but problematic vocabulary in Martin (1966) and Whitman (1985) and shows that a stronger theory of Korean-Japanese lexical comparisons is within reach. Section 5 summarizes the main points of this paper, and concludes with a discussion of the Korean-Japanese question.

## 2. Morpho-Lexical Correspondence: pKJ *pə-

This section proposes a comparison of MK *pwó-* 'sees' with OJ *pe-* 'time passes' to reconstruct a common pKJ verb *pə- 'sees'. Not only is 'sees' an important item in core vocabulary, this paper will show that in pKJ *pə- is also present in the formation of two additional verbs that show perfect Korean-Japanese correspondences, 'sends out' and 'wishes'. The reconstruction of 'sees' thus provides an example of interlocking compound morphology inherited into both languages, which is highly unlikely to be due to chance or borrowing.

### 2.1. Comparison of 'sees'

OJ *pe-* 'time passes' < pJ *pə- 'sees' ~ MK *pwó-* 'sees it'. This etymology is proposed in Ratte (2015) but with a vowel reconstruction that is problematic and resolved in the analysis presented here. OJ pe- 'time passes' is a lower bigrade verb. Most verbs in the lower bigrade conjugation are in all likelihood derived with a 'transitivity switch' formant *-e- or *-(C)i- (Unger 1993; Whitman 2008). Thus, OJ *pe-* '(time) passes, (time) is passed' probably incorporates the bigrade formant *-e-, which implies a proto-Japanese verb root *pə- whose transitivity is opposite to that of OJ *pe-*. OJ *pe-* has an intransitive meaning '(time) passes, (time) elapses,' so a plausible reconstruction for its transitive root *pə- is '(subj.) passes time, experiences a time'. As argued in Ratte (2015), evidence from its interaction with aspect militates towards *pe-* originally deriving from a verb denoting an instantaneous action like 'see' rather than 'pass'. For instance, in *Man'yōshū*, OJ *pe-* is mostly attested with the perfective auxiliary *-nu* (here in its infinitive form *-ni*):

(1) a. *tosi pa pey-ni-tutu*  'the years, they have kept passing by' (*Man'yōshū* 1080)
  b. *tukwi zo pey-ni-kyeru*  'the moons have gone by' (*Man'yōshū* 1464, 2093).

Why should *pe-* be consistently attested with the perfective auxiliary? If the root of *pe-* meant 'sees' or 'experiences,' then we can explain the preference for perfective marking by noting that 'experiencing' an event is logically prior to the description of a resulting state as 'passing by'. I therefore hypothesize that OJ *pe-* comes from pJ *pə- 'sees, experiences it'. Ratte (2015) reconstructs *pwə- for this verb, but given that there is no evidence for the distinctiveness of *Cwə versus *Cə in proto-Japanese, there is no Japanese-internal evidence for reconstructing a lost labial in the Japanese form.

MK *pwó-* is an accented stem that patterns with MK *wó-* 'comes'; given that *wó-* functions as root affix indicating active semantics ('volitive / modulator,' see Section 3), a plausible explanation for the accentual class of MK *pwó-* 'sees' is that it incorporates the modulator verb *wó-*. Hence, MK *pwó-* < *po-wó- 'see-volitive'. Since the root vowel is entirely erased, it is likely to have been minimal *o /ə/. PKJ *pə- 'sees, experiences it'.

## 2.2. Comparison of 'is sent out'

MK *pwonáy-* 'releases, sends it' ~ OJ *panas-* / *panat-* 'releases it,' MJ *fanare-* 'is separated,' pKJ *pə-na-* 'see-go out; is sent out'. Ratte (2015) points out that both MK *pwonáy-* 'sends it, releases it' and MJ *fanare-*/ OJ *panas-* are likely derivations from the same root: *pə- 'sees' (in Section 2.1) + *na- 'goes out' (Whitman 1985). The difference in vowel quality of OJ *panas-* with MK *pwonáy-* is explained by the sound co-occurrence rules known as Arisaka's Laws, which have caused a shift of *pəna- > *pana- in proto-Japanese. PKJ *pə-na- is a pre-existing compound of *pə- 'sees it' + *na- 'goes out' to mean 'is sent out, released'.

This cognate is not likely to represent a loanword correspondence. Assuming that pre-OJ distinguished *po /pə/ ≠ *pwo /po/[2], the above reconstruction of pKJ *pə-na- for OJ *panas-* violates Arisaka's Third Law of vowel co-occurrence, which bars B-type *o /ə/ from co-occurring in the same word with the vowel *a[3]. Since no branch of Japanese or Ryukyuan seems to preserve forms that violate Arisaka's Third Law, we can conclude that Arisaka's Laws describe the vocalic co-occurrence patterns of *ə in proto-Japanese, the common ancestor of all Japonic varieties. Therefore, any sound changes that gave rise Arisaka's Laws of the non-cooccurrence must predate the differentiation of proto-Japanese into Japanese and Ryukyuan branches. This means that Korean-Japanese cognates posited on the basis of a shift *ə > *a explaining Arisaka's Third Law are dependent on a sound change in proto-Japanese, and thus cannot be late borrowings into Old Japanese. The comparison of OJ *panas-* / *panat-* < *pana- to Korean

---

[2] These two syllable types are no longer distinguished by the Old Japanese period, but it is quite likely that such a distinction once existed in preceding periods. One piece of evidence for distinct *po ≠ *pwo is the difference in the allomorphy of OJ *pwi / po-* 'fire' and *puney* 'boat' / *pey* 'boat?,' (*peysaki* 'prow'). OJ *pey* can only come from original *poy* or *pay*, and its allomorph *pu(+ney* 'root'*)* points to *poy* with raising in the compound *puney*. OJ *pwi / po-* 'fire' must similarly come from *poy*, but mid-vowel raising does not take place, even in compounds (e.g. OJ *ponopo* 'flame,' not **punopo*). Given that mid-vowel raising is thought to have affected pJ *o (not *ə), reconstructing 'boat' as *poj and 'fire' as *pəj is the logical conclusion, indicating distinct *po ≠ *pwo in pre-OJ.
[3] Arisaka's Third Law states that *otsu-rui* or B-type *o* syllables /ə/ are almost always found in words that contain other examples of B-type *o* /ə/ or A-type *i* /i/. B-type *o* syllables almost never co-occur word-internally with *a*, *u*, or A-type *o* syllables. In other words, B-type *o* /ə/ co-occurs only with other /ə/ or /i/, and almost never with /a, u, o/. Francis-Ratte (2016, to appear) hypothesizes that a sound change of *ə > *a explains the co-occurrence patterns of Arisaka's Third Law.

*pwonáy-* therefore cannot be attributed to late borrowing, and must either be due to chance or common inheritance.

### 2.3. Comparison of 'wishes'

MK *pólá-* 'wishes it' ~ OJ *por-* 'wishes it', pKJ *pə-(a)ra-* 'see-have'. The comparison of MK *pólá-* 'wishes it' and OJ *por-* 'wishes it' is well-known, but Whitman's (1985) account for the correspondence does not explain final the root-final vowel *á* in Korean. In addition, neither Martin (1966: #62, DESIRE) nor Whitman (1985: #10) elucidates the the internal relationship of MK *pólá-* 'wishes it' to MK *pwó-* 'sees'.

I propose that both OJ *por-* 'wishes it' and MK *pólá-* are lexicalizations of *pə-* 'sees' (Section 2.1) + a pKJ verb *ara-* 'has'. MK *pólá-* and MK *pwó-* 'sees' appear to be related, and the fact that *pólá* displays an initial vowel ó strongly suggests that the root of MK *pwó-* 'sees' is in fact *po-* as argued in Section 2.1. Root-final *á* in Korean *pólá-* can be explained by reconstructing a vowel-final verb suffix *ara-* (see Francis-Ratte, 2016 to appear). The fact that both the Japanese and Korean forms can be analyzed as morphological derivations from a shared verb root suggests that the correspondence is not one of a loanword.

### 3. Morpho-Lexical Correspondence: pKJ *(w)o

The similarity of the Japanese adnominal suffix *-u* / *-uru* / *-ru* (proto-Japanese *-or; see Whitman 2004) to the Korean prospective adnominal suffix *-(o/u)l*[4] is often cited as evidence of possible Korean-Japanese common origin (Whitman 1985; Robbeets 2005). However, Vovin (2010) criticizes the comparison on the basis that the vowels are incongruous. To resolve this incongruence, I will argue in Section 3 that the correct correspondence of the proto-Japonic adnominal marker *-or is not directly to MK *-(o/u)l* 'adnominal' < *-r, but rather to the Middle Korean 'modulated' adnominal *-wol* / *-wul* < pK *-o-r, which incorporates what Martin (1992) calls the 'modulator' and what Lee and Ramsey (2011) call the 'volitive' suffix.

### 3.1. Comparison of 'comes'

The key to understanding the correspondence of adnominal morphemes in Korean and Japanese begins with a diachronic analysis of two important verbs, MK *wó-* 'comes' and OJ *wi-* 'comes to a stop'. Ratte (2015) compares OJ *wor-* 'is present' and OJ *wi-* 'sits down' to MK *wó-* 'comes,' which is a phonologically strong correspondence, and reconstructs pKJ *(w)o- 'comes'. This reconstruction takes into account the arguments by Kinsui (1983) that the root of OJ *wi-* 'sits down' and *wor-* 'is present' must be a verb of motion meaning *'comes to a stop'. I will further argue in Section 3.2 that the comparison of pJ *wo- to MK *wó-* 'comes' and the reconstruction of a motion verb *(w)o- in proto-Korean-Japanese also has implications for the compositionality of noun-modifying structures in both languages.

Ratte (2015) further postulates that the Middle Korean 'volitive' morpheme *-wó-* is etymologically identical to MK *wó-* 'comes' and derives from this same root pKJ *(w)o- 'comes'. The volitive (Lee and Ramsey 2011) or modulator (Martin 1992) suffix *-wó-* (and dark harmony alternant *-wú-*) is a verbal affix that attaches directly to the verb root in Middle Korean. Although the exact meaning of the volitive *-wó-*/*-wú-* is unclear, it appears to have been employed when the predicate expressed a "subjective will or intent" (Lee and Ramsey, 2011:

---

[4] The adnominal suffix in MK takes the form *-l* following vowel-final roots, *-ol* following consonant-final roots exhibiting light vowel harmony, and *-ul* elsewhere.

206). The volitive or modulator morpheme is obligatorily suffixed when an adnominalized or nominalized predicate expresses an action done upon the syntactic object of the verb (Lee and Ramsey 2011: 206). I will argue in Section 3.2 that the volitive or modulator morpheme may have originally functioned as a marker of active semantics.

## 3.2. Noun-modification in comparative perspective

Morphological comparisons of Japanese to Korean such as Whitman (1985) compare the OJ adnominal suffix *-u* (quadrigrade) / *-uru* (bigrade and irregular verbs) / *-ru* (monograde verbs) directly to the MK adnominal *-(o/u)l*. In light of Whitman's (2004) reconstruction of the proto-Japanese adnominal suffix as *-or, both Japanese and Korean adnominal suffixes do exhibit similarities in that both possess a final liquid consonant.

However, one distinct difference between the forms is that the pJ adnominal *-or must be reconstructed with a rounded back vowel *o. This reconstruction is necessary to explain why quadrigrade verbs in Eastern Old Japanese have an adnominal ending *-o* where Western Old Japanese quadrigrade verbs have an adnominal ending *-u*, for which Whitman (2004) posits mid-vowel raising of pJ *or > *-u* in Western Old Japanese. However, the proto-Korean adnominal suffix can be reconstructed as simply *-r[5]; it appears not to have had a back vowel *o as the proto-Japanese adnominal suffix does. This mismatch in vowel quality in part leads Vovin (2010) to reject the comparison and to claim that noun-modifying structures in Japanese and Korean are unrelated. I agree with Vovin (2010) that a direct comparison of noun-modifying structures is problematic. Instead, the proper cognate to the proto-Japanese adnominal suffix *-or is the modulated adnominal suffix *-wól*, whose reconstructed proto-Korean form *-o-r is segmentally identical to the proto-Japanese adnominal suffix.

According to Yanagida and Whitman (2009), Old Japanese can be characterized as following a split-active system of morphosyntactic alignment. Non-nominalized clauses, such as finite (*shūshikei*) clauses, follow the nominative-accusative alignment characteristic of Middle and Modern Japanese, with subjects unmarked and objects marked with *-wo*. By contrast, nominalized and adnominalized clauses in Old Japanese (e.g. *rentaikei* clauses) seem to show a very different relationship between verbs and their arguments, instead displaying an active or active-stative morphosyntactic alignment. In such clauses, active *agents* become marked with the postposition *-ga* (which is normally a marker of genitive case for animate nouns), and nouns marked with *-wo* often precede the agent marked with *-ga*. Under the theory proposed by Yanagida and Whitman (2009), active alignment originates from the fact that such clauses are syntactically nominal in origin. Yanagida and Whitman (2009: 134) provide some reasons for thinking that the development of active-stative alignment in syntactically nominalized clauses might be a cross-linguistic and universal development, though they note that the lack of clearly defined genetic affiliations for Japonic render a diachronic explanation difficult.

In this paper, I put forward that a diachronic explanation of active alignment in OJ nominalized clauses may be found in a morphological comparison of the proto-Japanese adnominal (*rentaikei*) suffix to the Middle Korean volitive adnominal suffix *-wól*. Sections 3.1 and 3.2 have set forth four premises, which can be summarized as follows:

---

[5] The MK adnominal suffix *-l* is obligatorily preceded by a vowel, and a minimal vowel *o/u* is inserted when it attaches to consonant stems. Because the inserted vowel is minimal, it is highly likely to be epenthetic in origin, which is to say that the proto-Korean adnominal suffix is likely simply *-r with underlying no vowel.

(3) a. The MK modulator morpheme *-wó-* marking agentive predicates likely represents a grammaticalization of pK *(w)o- 'comes' as a root-affix.

    b. Proto-Korean *(w)o- 'comes' is shared with Japanese, cognate with proto-Japanese *(w)o- 'comes (to a stop),' the root of OJ *wi-* 'sits down' and OJ *wor-* 'is

    c. The proto-Korean modulator morpheme *-o- combines with the plain adnominal *-r to form an active or agentive nominalized clause in *-o-r.

    d. The reconstructed adnominal morpheme *-or in proto-Japanese exhibits an active syntactic alignment.

The key to uncovering both a Korean cognate to pJ *-or and the possible origin of active morphosyntactic alignment in OJ is in a proper analysis of the proto-Korean system for adnominalization. Often overlooked in comparative treatments of Korean to Japanese is the fact that Late Middle Korean shows an important morphological distinction between unmarked narrative clauses and clauses marked with the modulator or volitive morpheme *-wó-* (with dark harmonic variant *-wu-*). As noted above, the usage of the volitive or modulator morpheme in MK appears to correlate with agency in the clausal predicate. The use of the LMK modulator morpheme *-wo-* with the base adnominal suffix *-l* creates a complex adnominal form *-wol* (proto-Korean *-o-r) that is used with adnominalized predicates expressing agentive semantics, and *-wo-* is obligatory in adnominalized clauses when the predicate describes an action done upon a syntactic object. The use of the volitive or modulator is thus motivated by the semantics of the clausal predicate, that is, whether the predicate describes an agent-driven action. This distribution is strikingly reminiscent of active marking in Old Japanese, which is semantically conditioned by the relationship of the verb to its arguments. According to Yanagida and Whitman (2009), nominalized clauses that express active agency gain special marking with *-ga*, indicating that agentive semantics are an important variable in OJ morphosyntactic alignment. On semantic and morphosyntactic grounds, there is ample reason for thinking that the MK volitive or modulator morpheme may be connected to the peculiar phenomenon of active marking in OJ nominalized clauses.

The observation that both proto-Korean modulated adnominal *-o-r and proto-Japanese adnominal *-or both exhibit active or agentive properties (as in 3c and 3d) suggests that pK *-o-r and pJ *-or might be cognates. Given that proto-Korean and proto-Japanese share the verb root *(w)o- 'comes' as likely cognates (as in 3c), the grammaticalization of *(w)o- as an active marker may be a feature of proto-Korean-Japanese. I therefore propose that pJ adnominal *-or corresponds directly in form and function to the pK modulated adnominal *-o-r. The vowel *o common to both pJ adnominal *-or and pK agentive adnominal *-o-r was originally a verbal suffix marking a predicate with active or agentive semantics, and is a grammaticalization of pKJ *(w)o- 'comes' (as in 3b). Like the volitive or modulator morpheme (pK *-o-), the pKJ active verb marker *-(w)o- originally functioned as a means of marking active or agentive semantics in its clause.

This analysis points to a proto-Korean-Japanese morphological system that distinguished between a simple adnominal suffix *-r used for non-active, non-agentive adnominalizations, and an active adnominal suffix *-o-r used for adnominalized predicates expressing active agency. In the Korean lineage, the use of modulator *-wo-* with clauses involving agentivity reflects this original active morphosyntax triggered by the presence of root-affixed *-o-, and Middle Korean inherits both narrative (non-agentive) adnominalizations in *-l* as well as active adnominalizations

in *-wól*[6]. In the Japanese lineage, this system underwent a shift whereby speakers generalized the use of the active adnominal *-o-r to all adnominal clauses (agentive and non-agentive) and ceased using the plain adnominal *-r. This shift must have taken place in a period prior to the differentiation of the Japonic family (i.e. in pre-proto-Japanese before the differentiation of Japanese and Ryukyuan), as I am unaware of any evidence from varieties of Japonic for an adnominalizating morpheme *-r.

Reconstructing *-(w)o- as a marker of active or agentive semantics in the predicate also harmonizes with the independently formulated reconstruction of pKJ *(w)o- as a verb of motion 'comes' as in Section 3.1. It is cross-linguistically common for verbs of motion to be grammaticalized into verbal auxiliaries expressing intentionality; for example, the subject of the English phrase *be going to* was originally restricted to an active agent going to a location for some purpose, and the verb following this phrase was restricted to a dynamic action (Bybee 2005: 605). It is entirely reasonable to think that a motion verb *(w)o- 'comes' in pKJ may be the etymological source for the putative marker of active / agentive action *-(w)o- that shares its segmental shape. Thus, the existence of just such a plausible internal etymology for active *-(w)o- provides further circumstantial support for reconstructing *-(w)o- as an active marker.

This proposal explains for the vowel discrepancy between the pJ adnominal *-or and the pK non-modulated adnominal *-r (rightly pointed out as problematic by Vovin), and also provides a diachronic explanation for why nominalized and adnominalized clauses in Old Japanese display active syntactic alignment as noted by Yanagida and Whitman (2009). Under this analysis, the diachronic origin of active alignment in OJ nominalized clauses is the fact that the pJ adnominal *-or incorporates an active marker *o from proto-Korean-Japanese. A consequence of this analysis is that all OJ morphology that incorporates and builds on pJ adnominal *-or (*rentaikei*, *izenkei*, and verbal nominal *-aku) should also inherit active morphosyntactic properties, just as Yanagida and Whitman (2009) observe.

## 3.3. Potential problems with active *-(w)o-

One potential problem for the hypothesis that active syntax and semantics are properties associated originally with *-(w)o- is the fact that *mizenkei* (irrealis) clauses in OJ, in particular the *mizenkei* conditional in *-(a)-ba*, also seem to show features of active morphosyntactic alignment. Yanagida and Whitman (2009) note that OJ *mizenkei* (irrealis) clauses also show agent-marking with *-ga* just as *rentaikei* and *rentaikei*-derived clauses do[7]. Yanagida and Whitman (2009: 132) believe that *mizenkei* clauses may be derived from a verbal nominalization in *-a, which leads them to posit that active morphosyntactic alignment is a property of OJ nominalized clauses more generally. The absence of the adnominal formant *-ur-* in *mizenkei* forms of OJ verbs (but present in the *izenkei* and verbal nominal *aku) strongly suggests that the

---

[6] Some scholars have attempted to show that Korean once had non-accusative syntactic alignment by focusing on the nominative marker *-i*; however, Whitman and Yanagida (2012) have recently re-assessed the evidence for non-accusative alignment with *-i* in pre-modern Korean and shown conclusively that *-i* could not have been an ergative marker in pre-modern Korean. Analyzing *-(w)o- as a pKJ active marker recognizes non-accusative features of pre-modern Korean morphosyntax without reconstructing ergative morphosyntactic alignment in proto-Korean.

[7] The *izenkei* (realis) stem in *-e(y)* / *-ure* and the verbal nominalization in *aku are widely held to be morphologically derived from the *rentaikei* base *-or due to the paradigmatic similarities that appear across the conjugational classes of Old Japanese (Whitman 2004). For example, vowel-stem verbs in Old Japanese (e.g. *ake-* 'brightens') all exhibit the formant *ur* in their *izenkei* form (*akure*), *rentaikei* form (*akuru*) and verbal nominalization form (*akuraku*), but do not exhibit this formant in other conjugational stems (e.g. *mizenkei* form *ake*).

*mizenkei* stem is morphologically unrelated to the *rentaikei* stem, and thus not derived from the reconstructed active marker *-(w)o-. This observation may represent a problem for the hypothesis proposed here that active syntax and semantics are properties associated originally with the *-(w)o- fused within the *rentaikei*. However, this too may have a plausible diachronic explanation.

There is no active predicate marker *-(w)o- that is productive in Old Japanese, nor is such a productive marker found in any dialects of Japanese or Ryukyuan. It is reasonable to conclude therefore that if such a morpheme existed, the active marker *-(w)o- lost productivity before the differentiation of the proto-Japanese into Japanese and Ryukyuan. The early loss of an active marker *-(w)o- from proto-Korean-Japanese into proto-Japanese can help to explain why *rentaikei*-based clause types and *mizenkei* clauses both display active syntactic alignment.

Let us now imagine this situation from the perspective of a speaker of proto-Japanese, in which an active marker *-(w)o- has lost productivity, but constructions built from this morpheme (e.g. *rentaikei* *-o-r) continue to retain the properties of active syntactic alignment originally associated with the presence of *-(w)o-[8]. Following the loss of *-(w)o- from their morphological inventory, speakers of proto-Japanese would have been presented with an odd, illogical morphosyntactic distribution: nominalized and adnominalized clauses with *-or in their language would have shown active morphosyntactic alignment, speakers would know of no structural justification for their being so. I believe that speakers that were faced with this distribution would naturally have been driven to reanalyze active syntactic alignment from a property of the embedded *-(w)o- in the adnominal *-or to a property of nominal forms of verbs more generally. This reanalysis allows speakers to form an internally consistent understanding of active syntactic alignment, and possesses a greater logic than the arbitrary stipulation that only nominalizations of the *rentaikei* type display active properties. This discussion provides a plausible historical pathway for explaining why *mizenkei* conditional clauses pattern together with *rentaikei* clauses in Old Japanese morphosyntax.

### 3.4. Summary: pKJ *-(w)o-

Sections 3.1 to 3.3 have provided arguments for a novel theory that the adnominal verb suffix of Old Japanese is directly related to the modulated adnominal in Middle Korean. To account for the correspondence, I have reconstructed an active marker *-(w)o- in proto-Korean-Japanese that explains the diachronic origins of active morphosyntactic alignment in Old Japanese[9], as well as the volitive or modulator morpheme in Middle Korean. An explanatory correspondence of adnominal forms in Korean and Japanese is particularly significant as several important constructions in OJ are built from the *rentaikei*, such as the *izenkei* (realis) stem and the OJ verbal nominalization in *-aku*, which suggests the possibility that more morphological correspondences between Japanese and Korean remain to be discovered.

---

[8] It is common for morphemes that become fossilized as to retain properties of their original function (Traugott 2003).

[9] This analysis does not preclude the possibility of a further diachronic connection between active morphosyntactic alignment and the OJ accusative marker *-wo* as Yanagida and Whitman (2009: 134) suggest, particularly in light my reconstruction of the pKJ active morpheme as *(w)o. However, I believe that care should be exercised in reconstructing uninflected verb roots as nominal postpositions, given that verb roots always require inflectional support in Japanese in all but a few cases of imperative mood.

## 4. Lexical Evidence for pKJ: Four Revised Etymologies

As discussed in Section 1, the existence of weak cognate pairs is a significant barrier to broader scholarly consideration of the pKJ hypothesis. This section discusses four problematic comparisons and identifies more convincing Korean cognates for each of the Old Japanese forms.

### 4.1. Revision of 'water'

Martin (1966) compares OJ *midu* 'water' directly to MK *múl* 'id.' and reconstructs pKJ *myaldu. However, OJ *midu* is likely to be morphologically complex, since *mi* alone is attested in compounds meaning 'water,' e.g. OJ *idumi* 'spring,' literally 'leaving waters' (OJ *ide-* 'leaves'). Also, Unger (2009) points out philological evidence that *tu* meant 'river' in pre-OJ, which suggests *midu* 'water' may be derived from a compound of pre-OJ *\*mi* 'water' and *\*tu* 'river'. Does pre-OJ *\*mi* 'water' correspond then to MK *múl* 'water'? Perhaps not; other examples of MK syllables ending in *-ul* generally correspond to OJ *-wi*, *-ey*, or *-o* (compare the correspondence of MK *pul* 'fire' to OJ *pwi* / *po-* 'fire'). The correspondence of MK *múl* 'water' to pre-OJ *\*mi* 'water' is phonologically irregular. Furthermore, there is distinct evidence that the proto-Japanese form of 'water' is not *mi but rather *me. As Vovin (2010: 198) correctly notes, Ryukyuan languages preserve traces of an original distinction between pJ *me and pJ *mi, syllables which have mostly merged in Old Japanese, and 'water' shows evidence of having been proto-Ryukyuan (and hence proto-Japanese) *me. Proto-Japanese *-e does not correspond regularly to MK *-ul*, so reconstructing proto-Japanese *me for 'water' renders the comparison to Korean *múl* entirely problematic.

An alternative to comparing MK *múl* to OJ *midu* is a comparison of MK *múl* to OJ *mopi* 'drinking water; gourd for drinking water'. The idea for the comparison is mentioned in Martin (1966), though he supports the traditional view that MK *múl* corresponds to OJ *midu*. Omodaka et al. (1967: 747) list 'jug' as the primary meaning of OJ *mopi* in *Man'yōshū* and 'usable, potable water' as secondary, but compounds like OJ *mopi-tori* 'drawing water for use' (*tori* 'taking') indicate that *mopi* also meant 'water'. The semantic specificity of *mopi* and lack of a transparent derivation suggest that *mopi* may be a lexicalization containing a morpheme meaning 'water'. I analyze OJ *mopi* 'jug for usable water' as a lexicalized compound with *op-i* 'carrying, bearing (OJ *op-* 'carries, bears'), originally *\*mo-opi* 'carrying *\*mo*'. I hypothesize that *\*mo* in this compound meant '(potable) water,' and the compound *\*mo-opi* meant 'bearing water'. Comparing this pre-OJ reconstructed form *\*mo* 'water' to MK *múl* 'water,' I reconstruct pKJ *\*mïr* '(drinking) water'. The OJ vowel *o* corresponds regularly to MK *u*, and the final consonant *\*r* shows no reflex in OJ due to the fact that *\*mo* is the initial element in a compound (compare how the original vowel *\*o* of OJ *pwi* 'fire' is exposed when 'fire' is the initial element in a compound, e.g. *ponopo* 'flame').

### 4.2. Revision of 'sky'

Ratte (2015) notes that the comparison OJ *swora* 'sky' to MK *hanólh* 'id.' in Whitman (1985) is difficult to accept, and that the stronger comparison for MK *hanólh* 'sky' is to OJ *kamwi* 'god'. This implies a pJ nominal *\*ka* 'godly,' from a pKJ root *\*ha-* 'great, celestial'. Whitman (2012) compares LMK *há-* 'is great' to OJ *kasa* 'a volume, a quantity' which he analyzes as *\*ka* 'great' + the adjective nominalizer *-sa*. OJ *kasa* 'volume' is usually analyzed as a single morpheme, but if Whitman (2012) is indeed correct, this would constitute further

evidence of pJ *ka 'great'[10]. The match to MK *há-* 'great,' *hanólh* 'sky' is semantically strong and phonologically perfect.

The origin of OJ *swora* 'sky' is more difficult to explain. Ratte (2015) compares OJ *swora* to Early Middle Korean *soy* 'dawn,' but this is evidently incorrect. Early Middle Korean *soy* is not attested in Late Middle Korean, and the fact that it is observed only after the merger of Late Middle Korean *o* with *a* means that *soy* 'dawn' is in all likelihood a hyper-corrective form of MK *say* 'id.' (e.g. MK *say-pay* 'id.'). Since MK *-a* does not correspond regularly to OJ *-wo*, OJ *swora* is probably not cognate with MK *say*. Instead, MK *swoynakí* 'brief rain shower' may be the proper cognate of OJ *swora* 'sky'. MK *swoynakí* 'a sudden rain shower' appears to be composed of an unknown element *\*swoy + naki*; the latter element can be identified as the nominalized form of *na-* 'goes, comes out of'. Internal analysis thus indicates that *swoynakí* 'a sudden rain shower' is a lexicalization from a phrase 'coming out of the \*swoy'. Hypothesizing *\*swoy* as 'sky' and *swoy-nakí* as 'coming out of the sky' ( > 'sudden rain shower') is a reasonable internal reconstruction. I reconstruct pre-MK *\*swoy* 'sky,' which is supported by the evidence discussed above that MK *hanólh* 'sky' is an innovation from *ha-* 'great'. PKJ *\*sorə >* pK *\*soj* by a hypothetical shift of *\*rə > y*; compare the shift of the medial liquid to a glide in MK *sye:(Wul)* 'capital city,' likely from Old Korean 斯盧 ／ 斯羅 ?*\*sirə*, supported by the Japanese cognate *siro* 'castle'. PKJ *\*sorə >* OJ *swora*, undergoing a shift of *\*ə > a* as per Arisaka's Third Law.

## 4.3. Revision of 'eye'

Vovin (2010: 191) is correct that Whitman's (1985) comparison of OJ *mey* 'eye' and MK *melí* 'head' is problematic, given that Early Middle Korean transcriptions of MK *melí* 'head' indicate that the liquid of *melí* is probably the result of lenition of *\*t*, from pre-MK ?*\*mati*. In addition, the meanings 'eye' and 'head' are not particularly close. Given that both a light harmony form (*mali*) and a dark harmony form (*melí*) of 'head' exist in Late Middle Korean, the pre-MK form of 'head' may have exhibited a root-internal harmonic mismatch that was resolved by speakers leveling the harmony of the vowels in both light and dark directions. The vowel *i* is harmonically neutral, so the only way for root-internal harmonic mismatch to have existed in the pre-MK form of 'head' is if the attested vowel *i* of the second syllable is a monophthongization of an original diphthong *\*uy*[11]. Therefore, I reconstruct pre-MK *\*matúy* for 'head,' where the harmonic mismatch of initial syllable *a* (light harmony) and second syllable *úy* (dark harmony) leads some speakers to innovate MK *melí* via harmonic leveling of *\*a > e*, and other speakers to innovate MK *mali* via monophthongization of *\*uy > i*. Reconstructing *\*matúy* 'head' suggests a further internal relationship of 'head' to MK *motoy* 'joint, knuckle' (note also MK *mátoy* 'directly above,' which may be related as well). I reconstruct proto-Korean *\*mətaj* or *\*mataj*[12] 'place where body parts are joined, come together,' and I compare this form to OJ *mata* 'fork, bend, crotch'. The comparison points pKJ *\*mətaj* or *\*mataj* 'place where body parts are joined'.

## 4.4. Revision of 'blue'

Whitman (1985) compares MK *pholol-*, *phulul-* 'blue / green' tp OJ *awo* 'blue,' but as

---

[10] Though OJ *kasa* may be a better cognate with MK *kes* 'thing'.

[11] The MK syllabic nuclei *uy* and *i* are virtually never distinctive in polysyllabic words.

[12] Based on internal evidence, the vowel of the initial syllable is ambiguous; it may be *\*a* with minimalization in *motoy*, or it may be *\*ə* with strengthening to *a* in pre-MK *\*matuy*.

Vovin (2010) points out, the initial aspirate *ph* of MK *pholol-* must go back to *\*pVk* or *\*kVp*, which invalidates the match to OJ *awo*. Instead, MK *pholol-*, *phulul-* 'blue / green' and the related nominal MK *phul* 'grass, shoots' can be compared to OJ *waka* 'young,' from pKJ *\*wakə* 'young' (+ pK continuative suffix *\*-ol-*).

The adjective 'blue, green' in Korean belongs to a small class of irregular inflecting stems displaying an excrescent *-l-* in their infinitival form, a class that also includes *nwoló(l)-* 'yellow'. Their morphological irregularity likely indicates a derived class, and the excrescent *-l-* may reflect part of a derivational suffix that is suppressed in most environments (i.e. before a consonant) but expressed before a vowel (infinitive *-a/e*). Hence, MK *pholo(l)-* 'blue' < pre-MK *\*pVkol-ol-*. Lee and Ramsey (2011: 74) cite a reconstructed form *\*pwok* or *\*pwak* as the Old Korean word for 'youth,' based on its phonographic representations 福, 卜, 巴, and 伏; this *\*pwak* 'youth' appears to be the root in MK *pholo-*/*phulú-* 'blue/green' < *\*pVkol-ol-* and MK *phúl* 'grass, shoot' < *\*pVkul*.

I reconstruct MK *phúl* 'grass, shoots' as a deverbal expression from this putative adjective root *\*pak-* 'young' with the adnominal suffix *\*-(o/u)l* meaning 'the young one' in a development similar to the use of English *green* to mean 'naive, young, inexperienced'[13]. MK *phúl* usually refers to grass, but can refer to any number of short-stemmed annuals that grow quickly; moreover, one of the glosses for 筍 (Sino-Korean *swun* 'shoot') is the word *phúl*. The MK adjective *pholo(l)-* 'blue, green' is built from this stem *\*pak-ər* 'the young one' by the addition of a 'continuative' verbalizing auxiliary *\*-(o/u)l-* [14] (Whitman 2012). The irregular morphological class of MK *pholo(l)-* 'blue, green' reflects its diachronic origin as a double derivation from *\*pak-ər-ər-*.

## 5. Conclusions

This paper has built on the work in Ratte (2015), which proposes new Korean-Japanese cognates but contains errors that are corrected herein. In so doing, this paper seeks to build a stronger theory of proto-Korean-Japanese that synthesizes the results of previous Korean-Japanese research with recent criticism and advances in our understanding through the in-depth internal work accomplished in the three decades since Whitman (1985). The reconstruction of an inflecting stem *\*(w)o-* in proto-Korean-Japanese not only connects MK *wo-* 'comes' to OJ *wi-* 'comes to a stop,' but also is key to understanding the origin of the proto-Japanese adnominal suffix *\*-or* and how it corresponds to the Korean volitive or modualtor morpheme *-wo-*. Reconstructing pKJ *\*pə-* 'sees' reveals a set of common compound structures that link two additional Korean and Japanese verbs, pKJ *\*pə-na-* and *\*pə-(a)ra-*; interlocking correspondences such as these are significant for being far more specific than any single lexical comparison. Finally, this paper has shown that weak comparisons in Martin (1966) and Whitman (1985) can be revised into strong cognates, demonstrating that the pKJ hypothesis has even greater explanatory power than is currently thought.

Although this paper has presented only a small number of new Korean-Japanese lexical and morphological correspondences, it is evident that they have broad implications for the ultimate

---

[13] The development in English is obviously opposite to the hypothesized development of proto-Korean *\*'young' → 'green,' but I have no reason to think that the development is unidirectional; note Japanese *wakatake-iro* 'blue' (lit. 'the color of young bamboo'), indicating the metaphor of 'young (plant)' → 'blue/green'.

[14] Though superficially similar in form to the MK adnominal suffix, the continuative is an inflecting extension (see the analysis of MK *pólá-* 'desires it' in Section 2.3).

structure of the pKJ reconstruction as a whole. This paper thus underscores the critical need to consider alternative etymologies and sound correspondences before adopting a radically skeptical hypothesis such as that of Vovin (2010).

**References**

Bybee, J. (2003) Mechanisms of Change in Grammaticalization. In B. Joseph & R. D. Janda (eds.), *The Handbook of Historical Linguistics*, 602-23. Malden: Blackwell.

Francis-Ratte, A. (2016., to appear). Proto-Korean-Japanese: a new reconstruction of the common origin of the Japanese and Korean languages. Doctoral dissertation, The Ohio State University.

Frellesvig, B. (2001) A common Korean and Japanese copula. *Journal of East Asian Linguistics* 10, 1–35.

Frellesvig, B. (2008) On the reconstruction of proto-Japanese and pre-Old Japanese verb inflection. In B. Frellesvig and J. Whitman (eds.), *Proto-Japanese: Issues and Prospect*, 175-192. Amsterdam: John Benjamins.

Frellesivig, B. and J. Whitman. (2008) Evidence for seven vowels in proto-Japanese. In B. Frellesvig and J. Whitman (eds.), *Proto-Japanese: Issues and Prospect*, 15-42. Amsterdam: John Benjamins.

Frellesvig, B. (2012) Chronological Layering in proto-Japanese and pre-Old Japanese Verb Inflection. In B. Frellesvig, J. Kiaer, and J. Wrona (eds.), *Studies in Japanese and Korean Linguistics*, 28-47. Munich: Lincom.

Kinsui, S. (1983) Jōdai, chūko no *wiru* to *wori* - jyōtaika keishiki no suii. *Ōsaka daigaku daigakuin bungaku kenkyūka kiyō* 42, 1-25.

Martin, S. (1966) Lexical Evidence Relating Japanese to Korean. *Language* 42.2, 185-251.

Martin, S. (1987) *The Japanese Language through Time.* New Haven: Yale University Press.

Martin, S. (1996) *Consonant Lenition in Korean and the Macro-Altaic Question.* Honolulu: University of Hawai'i Press.

Nam, K. (1997) *Kwo'e sacen*. Seoul: Donga Publishing:

Nihon Daijiten Kankōkai and Shōgakkan. (2000) *Nihon kokugo daijiten*. Tōkyō: Shōgakukan.

Ramsey, R. and K. M. Lee. (2011) *A History of the Korean Language*. Cambridge: Cambridge University Press.

Ratte, A. (2015) Towards a stronger theory of proto-Korean-Japanese. To appear in M. Kenstowicz, T. Levin and R. Masuda (eds.), *Japanese / Korean Linguistics 23*, Chicago: University of Chicago Press.

Robbeets, M. I. (2005) *Is Japanese related to Korean, Tungusic, Mongolic and Turkic?* Wiesbaden: Harrassowitz.

Traugott, E. (2003) Constructions in Grammaticalization. In B. Joseph and R. D. Janda (eds.), *The Handbook of Historical Linguistics*, 624-47. Malden: Blackwell.

Unger, J. M. (1993) *Studies in Early Japanese Morphophonemics.* Bloomington: IULC.

Unger, J. M. (2009) *The Role of Contact in the Origins of the Japanese and Korean Languages*. Honolulu: University of Hawai'i Press.

Vovin, A. (1993) Long Vowels in Proto-Japanese. *Journal of East Asian Linguistics* 2.2, 125-134.

Vovin, A. (2010) *Koreo-Japonica: A Re-Evaluation of a Common Genetic Origin*. Honolulu: University of Hawai'i Press.

Whitman, J. (1985) The Phonological Basis for the Comparison of Japanese and Korean. Doctoral dissertation, Harvard University.

Whitman, J. (1990) A rule of medial *-r- loss in pre-Old Japanese. In J. Baldi (ed.), *Trends in Linguistics 45*, 512-545.

Whitman, J. (2004). The Form and Function of the Rentaikei and Izenkei Suffixes in Proto-Japanese. Presented at the Second Oxford-Kobe Linguistics Seminar on the History and Structure of Japanese, 27 Sept. 2004.

Whitman, J. (2008) The source of the bigrade conjugation and stem shape in pre-Old Japanese. In B. Frellesvig and J. Whitman (eds.), *Proto-Japanese: Issues and Prospects*, 159-174. Amsterdam: John Benjamins.

Whitman, J. (2012) The Relationship of Japanese and Korean. In D. Tranter (ed.), *The Languages of Japan and Korea*, 24-38. London: Routledge.

Yanagida, Y. and J. Whitman (2009) Alignment and word order in Old Japanese. *Journal of East Asian Linguistics,* 18.2, 101-144.

Yanagida, Y. and J. Whitman. (2012) A Korean Grammatical Borrowing in Early Middle Japanese *Kunten* Texts and its Relation to the Syntactic Alignment of Earlier Korean and Japanese. In S. Nam, H. Ko, & J. Jun (eds.), *Japanese / Korean Linguistics 21*, 121-135. Stanford: CSLI.