Robbeets, M. (2017). The Japanese inflectional paradigm in a Transeurasian perspective. In M. Robbeets (Ed.), *Stability and borrowability* (pp. 130-164). London [u.a.]: Routledge.

The Japanese inflectional paradigm in a Transeurasian perspective

Martine Robbeets

Although the genealogical relationship between Japanese and the Transeurasian languages has been a source of contention for nearly two centuries, scholars seem to agree that paradigmatic morphology could substantially help to prove relatedness. Starting from this consensus, this contribution examines whether the correlations in verb inflections between Japanese and these languages can be characterized as "paradigmatic" and whether they are more likely to result from chance or borrowing than from inheritance. For this purpose, this paper advances Transeurasian cognates for the five basic inflected forms of Japanese grammar as well as one derived stem. Taking into account internal cohesion between ordered sets of cognate forms, shared idiosyncrasies and extended relationships of grammatical patterning, the paper concludes that the correlations in verb inflections are indeed paradigmatic and more likely to be inherited than to be coincidental or borrowed.

Keywords: genealogical relationship; Transeurasian languages; Japanese; verb inflection; paradigm

1. Introduction

The question of whether the languages here referred to as Transeurasian, namely the Japonic, Koreanic, Tungusic, Mongolic and Turkic languages, constitute a genealogical grouping remains one of the most disputed issues in historical comparative linguistics. The controversial classification has been on the table for nearly two centuries, but in spite of recent claims from both supporters and critics that the controversy has been resolved (Starostin et al. 2003: 7; Vovin 2005: 71), the debate is not so easily settled.[1] Nevertheless, irrespective of whether they favor or doubt Transeurasian relatedness, scholars seem to agree on at least this one point, i.e. that paradigmatic morphology could substantially help unravel the question. Vovin (2005: 73) begins his critique of Starostin et al. (2003) with the postulation that "The best way … is to prove a suggested genetic relationship on the basis of *paradigmatic* morphology", and in their rebuttal of this critique, Dybo & Starostin (2008: 125), supporters of relatedness, agree that "regular paradigmatic correspondences in morphology are necessarily indicative of genetic relationship." This relative consensus provides an interesting starting point for this paper because it raises two issues: the first is a theoretical question concerning the weight that should be given to paradigmatic evidence, the second is an empirical one, concerning the existence and the historical explanation of paradigmatic correlations between the

[1] For an overview of the history of the debate, I refer to Robbeets (2005: 18-29).

Transeurasian languages.

Given the relative resistance to code-copying of verb morphology vis-à-vis nominal morphology and of inflectional morphology vis-à-vis derivational morphology (Weinrich 1953: 35; Moravcsik 1978; Thomason & Kaufman 1988: 74-75; Wilkins 1996; Matras 2009: 153-165), it can be expected that inflectional verb morphology will provide rather reliable evidence to demonstrate common ancestorship. Therefore, this paper will focus on paradigmatic correspondences in inflectional verb morphology across the Transeurasian languages. The following Section deals with the importance of paradigms in establishing relatedness, including some guidelines that allow us to eliminate code-copying or chance as a motivation for paradigmatic correlations. Section 3 proposes Transeurasian cognates for the paradigm of Japanese basic inflected forms. Section 4 reflects upon the paradigmatic organization of the evidence, paying special attention to the likelihood of paradigmatic copying. By way of conclusion, Section 5 summarizes the main arguments for the availability of common paradigmatic morphology in support of the genealogical relationship of the Transeurasian languages.

2. The importance of paradigms in establishing relatedness

2.1. Paradigm and paradigmaticity

In the introduction to this volume, Robbeets and Bisang (this volume: xx) characterize a paradigm as "an organized set of derivationally or inflectionally

related items that derive a particular semantic or morphosyntactic category from a common base or root". As such, they favor the view that derivation, like inflection, is regulated by paradigmatic principles and that what really matters in both cases is a certain internal organization within a coherent whole. Nevertheless, there are some differences between derivational and inflectional paradigms. In contrast with a derivational paradigm, for instance, the choices in an inflectional paradigm are mutually exclusive; German *stellen* 'to put' combines with two actional suffixes to form *wieder-her-stellen* 'to reconstruct', but the past and present of this verb cannot be simultaneously derived. Moreover, the semantic relationships among the cells of an inflectional paradigm remain constant from one lexeme to another; German *kratzen* 'to scratch' vs. *kratzte* 'scratched' expresses the same difference in meaning as *herstellen* 'to produce' vs. *stellte her* 'produced', whereas *kratzen* vs. *Kratzer* 'claw mark' and *herstellen* 'to produce' vs. *hersteller* 'producer' does not. The observation that derivational paradigms are often marked by semantic irregularity opens alternative opportunities for genealogical linguists in search of so-called "quirks", i.e. shared idiosyncrasies that are particularly telling for the establishment of linguistic relatedness.

The way in which the term "paradigm" is defined in the introduction to this volume is the full set of forms, inflectional and derivational, that a root enters into. Such a set is not small and not inherently closed, since every root enters into a different array of derivations and not every root has the full set of inflections. This characterization is in line with the definition of "paradigm" in morphological theory, but it is more general than the notion of "paradigmaticity" that has been advanced as diagnostic in historical comparison. As pointed out in

Johanna Nichols' internal review of this contribution, "paradigmaticity" in the comparative context refers to a closed set of form slots with positions defined by intersections of category dimensions, or the like that are definable independently of the forms that fill them. Since the set of forms is closed, it is not large, as illustrated, by the six forms filling the person-number paradigm of the copula in Indo-European in Table 1 below.

2.2. Inherited paradigms

Lexical comparison has two dimensions: form and meaning. The Latin root *es-* 'to be', for instance, can be compared to the copula Sanskrit *as-*, Greek *es-* and Gothic *is-*, reflecting a common form **h₁es-* and a common meaning 'to be'. Similarly, the Latin ending of the present indicative third singular *-t* in *es-t* 'he is' can be compared in form and function to Sanskrit *-ti*, Ancient Greek *-ti(n)* and Gothic *-t*, but here, as illustrated in Table 1, the patterning of person and number oppositions in the Latin present indicative paradigm can be compared to the oppositions in the other languages as well. In every language in the table, the copula fills a closed set of form slots with positions defined by intersections of the dimensions person and number agreement. The languages under comparison display correlations in grammatical patterning among ordered sets of disjunct forms, known as "multidimensional paradigmaticity" (Nichols 1996: 46), a correspondence, which adds a third dimension to the comparison.

This dimension can further be expanded by a shared irregularity in the formation of the copula root in the plural paradigm. Indeed, all paradigms except Greek, show a full-grade form with a vowel in the singular in opposition to a

zero-grade form with a vowel missing in the plural, e.g. Sanskrit *as-* vs. *s-*, Latin *es-* vs. *s-*, Gothic *is-* vs. *s-.* As such, it is not just the endings that match systematically, but there is also a systematic linkage between the roots, that would be hard to explain without recourse to a common ancestor.

Table 1. The present indicative paradigm based on the Indo-European copula **h₁es-* 'to be' (Beekes 1995: 13-14)

	Sanskrit	Homeric Greek	Latin	Gothic	pIE
1SG	*as-*	*es-*	*es-*	*is-*	**h₁es-*
2SG	*as-mi*	*ei-mi*	*s-um*	*i-m*	**h₁es-mi*
3SG	*as-i*	*es-si*	*es*	*is*	**h₁es-si*
1PL	*as-ti*	*es-ti(n)*	*es-t*	*is-t*	**h₁es-ti*
2PL	*s-mas*	*ei-men*	*s-umus*	*s-ijum*	**h₁s-més*
3PL	*s-tha*	*es-te*	*es-tis*	*s-ijuþ*	**h₁s-th₁é*
1SG	*s-anti*	*ei-si(n)*	*s-unt*	*s-ind*	**h₁s-énti*

Moreover, we find similar paradigms for the athematic present indicative, for the imperfect indicative and for non-indicative moods such as subjunctive, optative and imperative. Joseph (this volume: xx) introduces the notion of "extended paradigmaticity" to refer to external relationships of grammatical patterning among different paradigms, some of which may be the result of grammaticalization processes, such as the systematic linkage between personal pronouns and verb agreement markers. These external relationships of grammatical patterning among different systems make up the fifth dimension of paradigmatic evidence.

Hence, the strength of paradigmatic evidence as an indicator of genealogical

relatedness is in its combining multiple dimensions of comparison. In addition to formal and functional criteria, the correspondences can be assessed in terms of paradigmatic behavior such as internal cohesion between the ordered slots of a set of forms, shared formational irregularity in specific parts of the system and external relationships of grammatical patterning among different systems.

2.3. Copied paradigms

Although paradigmatic correspondences in morphology provide solid evidence for genealogical relationship, this statement cannot be taken to imply that every paradigm shared between two languages must necessarily be the result of inheritance. While many contact linguists stress that anything, including paradigmatic inflectional morphology, can be copied (e.g. Thomason & Kaufman 1988: 19-20; Thomason 2001: 65), some even go as far as to suggest that paradigmatic copying is actually favored over copying individual morphemes (e.g. Kossmann 2010; Seifart 2012). Even if this assumption seems plausible, it may still be possible to distinguish between the effects of contact and inheritance in shared paradigmatic morphology (Robbeets 2012). This is because copied paradigms usually give themselves away by reflecting some typical restrictions of lexical, morphological, semantic, morphosyntactic or distributional nature.

First, the most common type of paradigmatic borrowing cross-linguistically, is when languages copy morphological paradigms together with foreign lexicon (Kossmann 2010). Ajia Varvara Romani, a Romani dialect spoken in a suburb of Athens, for instance, copied the entire Turkish nonfocal present paradigm, but all copied morphemes are hosted by verbs copied from Turkish (Friedman 2009:

112). Therefore, shared paradigms in which the common inflections are restricted to common lexemes are likely to be copied.

Second, in copied paradigms the endings are often morphologically complex in the model language, but simplex in the recipient language. This is for instance the case for the copying of all Sakha paradigms copied in Uchur Evenki or Lamunkhin Even, discussed by Pakendorf (2009, and this volume: xx). The copied inflections consist of a Sakha ending along with the entire paradigm of Sakha person markers, which do not occur as simplex morphemes in Uchur Evenki, Lamunkhin Even or elsewhere in Tungusic. Therefore, shared paradigms in which the endings are morphologically complex in one language, but not in the others, are likely to be copied.

Furthermore, in Pakendorf's (this volume: xx) example, the Sakha assertive-presumptive suffix in *-Tax* also occurs in the temporal-conditional paradigm, from which it originates. However, the copy is restricted to the assertive-presumptive paradigm in Uchur Evenki, and to the assertive paradigm in Lamunkhin Even. Therefore, paradigms in which only secondary semantics are shared are likely to be copied.

Fourth, the copying of paradigms in Uchur Evenki and Lamunkhin Even clusters in a very specific morphosyntactic subsystem, notably mood (assertive-presumptive, necessative and hypothetical), while paradigms in other parts of grammar are left unaffected. In most documented cases of paradigmatic borrowing, such as in Resígaro (Seifart 2012), Michif (Thomason & Kaufman 1988: 228-233; Bakker 1997: 97-102; Comrie 2008: 21-22) or Copper Island Aleut (Thomason & Kaufman 1988: 233-238; Sekerina 1994; Thomason 1997; Comrie 2008: 24-31; Comrie 2010: 28-30), we find a similar imbalance whereby

certain morphological subsystems have been entirely copied, while others have hardly been influenced at all. Therefore, shared paradigmaticity that is restricted to specific morphosyntactic subsystems is indicative of borrowing.

Finally, the paradigmatic copying in Pakendorf's example reflects a triple setting, going from Sakha into Uchur Evenki and, in parallel, from Sakha into Lamunkhin Even. This situation is rather rare: paradigm copying typically goes from a model language into a recipient language, but there are few instances where the same paradigm is copied into yet another language. Paradigms shared in a multiple setting such as between Sanskrit, Old Greek, Gothic and Latin in Table 1 are likely to have a genealogical explanation.

In Section 4.2, these five indications of paradigmatic copying will be extrapolated to linguistic prehistory in order to determine the likelihood of copying *vis-à-vis* inheritance as a motivation for the paradigmatic correlations between the Transeurasian languages.

2.4. Coincidentally matching paradigms

As noted by Hyman (this volume: xx) and Janhunen (this volume: xx), an important problem that reduces the diagnostic value of morphological comparisons is that morphological elements are typically very short, often mono- or bisegmental. The disadvantage of this typical shortness is that similar forms tend to occur by coincidence even in unrelated languages. Nevertheless, the comparison of morphemes as pursued in this contribution, displays a number of other characteristics that counterbalance the statistical effect of shortness. Factors that reduce the probability that the proposed morphological

correlations are coincidental include (1) the number of proposed etymologies, (2) the number of branches in which the morphological item has a match, (3) the relatively small size of the inventory of verb morphemes in an average language, (4) the verification of sound correspondences in matching morphemes against regular correspondences previously established on the basis of lexical data, (5) shared polysemy, (6) shared irregularity and (7) the occurrence of paradigmatic correlations.

First, whereas the present contribution lists six etymologies, Robbeets (forthcoming c) proposes over twenty different etymologies for verb suffixes, well represented in the five branches of the family. An approximation of the number of etymologies needed to exclude chance as a plausible explanation is proposed in Robbeets (forthcoming c: Section 2.3.2), but one intuitively feels that twenty etymologies with matching members in up to five branches is too striking a number to be attributed to coincidence.

Second, chance probability decreases with the number of branches in which the morphemes are matched, provided that the match is simultaneously present in each and every branch It is much easier to find coincidental matches in a binary setting, such as perhaps — if not genealogically motivated — the resemblances between the Transeurasian and Uralic deverbal noun suffixes pTEA *-*mA* and pUr *-*mə* or pTEA *-*i* and pUr *-*i*/ -*y*, than it is to add more coincidental look-alikes that are simultaneously present in — say — Niger-Congo, Sinitic and Austronesian.

Third, the probability that a certain correspondence is due to coincidence decreases with the number of elements that are open to comparison, i.e. the number of trials we make. Whereas the average number of words in a language

exceeds several tens of thousands, the average number of verb morphemes remains below one hundred. By consequence, the probability that a certain correspondence in verb morphology is due to coincidence will be lower than that for a similar correspondence within the lexicon, because the body of elements open to comparison is much smaller.

Fourth, comparing the Transeurasian and Uralic deverbal noun suffixes pTEA *-*mA* and pUr *-*mə*, we are unable to test the assumed sound correspondences pTEA **m* :: pUr **m* and pTEA **A* :: pUr **ə* against regular sound correspondences established on the basis of lexical comparison. By contrast, the formal correspondences of morphological cognates across the Transeurasian languages can be confirmed on the basis of independently established set of phonological rules proposed in Robbeets (2005).

Fifth, comparing forms with diffuse or general meanings such as the common denominator 'deverbal noun suffix' for pTEA *-*mA* and pUr *-*mə* or pTEA *-*i* and pUr *-*i*/ -*y*, may enhance coincidental matches. However, shared semantic specialization such as the distinctive use of pTEA *-*mA* in color nouns or shared polysemy such as the secondary use of pTEA *-*i* in converbs decreases chance probability. Note that the Uralic deverbal noun suffixes share neither this semantic specialization nor this polyfunctionality with the Transeurasian languages.

Sixth, shared irregularity such as in the formation of the Indo-European copula root in Table 1 reduces the probability that the correlations are due to sheer chance. Shared irregularity in bound morphemes may include shared allomorphy conditioned by a specific phonological environment, such as the allomorphy in the reflexes of the converb marker pTEA *-*i* ~ *ø* below or, shared

functional irregularity conditioned by a specific semantic environment, such as the aspectual and temporal distinctions of the reflexes of pTEA *-*rA* below, which are conditioned by the telicity of the base verb.

Finally, the sharing of an ordered paradigm of individual morphemes is more difficult to attribute to chance than the sharing of a list of randomly amassed morphemes. Examples of coincidentally matching paradigms are extremely difficult to find across the languages of the world. As illustrated in Table 2, Campbell & Poser (2008: 188) refers to the coincidences between a set of verb agreement endings in Proto-Eastern Miwok (Central California) and in Indo-European, but in this case the matches involve only five cells of the paradigm, not extending to the third person plural.

Table 2. Coincidences between Proto-Eastern Miwokan and Indo-European (Campbell & Poser 2008: 188)

	Proto Eastern Miwokan declarative suffixes	Indo-European active suffixes
1SG	**-ma· ~ -m*	**-m*
2SG	**-sY ~ -ṣ*	**-s*
3SG	**-Ø*	**-t < **-Ø*
1PL	**-maṣ·i ~ *-maṣ*	**-me(s)/-mo(s)*
2PL	**-to-k*	**-te*

3. The basic inflectional paradigm of Japanese verbs in a comparative perspective

The standard way of representing Japanese verb morphology within the Japanese tradition is to list six paradigmatically opposed verb forms for which any lexical verb inflects and on the basis of which one can derive any other verb form. This may have been inspired to some extent by the four main verb forms ("principal parts") in Latin grammar. If one knows *ferre*, *ferō*, *tulī*, *lātum*, the infinitive, present, perfect and perfect participle of the verb 'to carry', one can also derive any other verb form. The six "principal parts" of Japanese grammar are the *mizenkei*, which is a surface stem that diachronically reflects resegmentation of suffixes with initial *-*a*; the *rentaikei* or adnominal form in -*uru*; the *izenkei* or subjunctive form in -*ure*; the *renyōkei* or conjunctive form in -*i*; the *shūshikei* or finite indicative in -*u* and the *meireikei* or imperative, which originally goes back to a bare verb root.

A conceptual shortcoming of this system is that it presents the six basic inflections as equivalent morphological forms, but in reality they are not equivalent. Although five forms (adnominal, conjunctive, indicative, subjunctive, imperative) indeed are basic inflected word forms, the *mizenkei* is not a word form, but simply a derived stem. The five inflected forms, however, can be referred to as the basic paradigm for Japanese verbs (Frellesvig 2010: 118). It is a fixed set of verb forms that is closed not only by convention in that it is the accepted set of principal parts within Japanese tradition, but also inherently in that it exhausts the entire set of forms taking part in Japanese verb inflection.

In what follows, I will propose an internal analysis as well as an external etymology for the five basic inflected forms of Japanese grammar and the so-called *mizenkei* derived stem. The putative cognates in the Transeurasian

languages are required to match the Japanese comparanda in terms of form, function, combinational properties and typological characteristics including grammaticalization patterns. As such, I start from a fixed set of forms chosen on independent, non-circular grounds (i.e. not because they were known in advance to have many matches in an otherwise sparse field) and I use independent non-circular criteria to identify matches.

*3.1. The copula pTEA *a- as the origin of the mizenkei or surface-stem*

Some Old Japanese verb forms are seemingly formed by adding a suffix (-*zi* negative tentative, -*na* desiderative, -*ba* conditional, -*s*- honorific, -*sime*$_2$- causative, -*ye*- passive, -*re*- passive, -*n*- negative, -*zu* negative, -*m*- tentative, -*masi* subjunctive) to the base of vowel final stems, while an extra element -*a*- is added to the base of consonant final stems. The *a*- stem of consonant verbs is referred to as the *mizenkei* in traditional Japanese grammar, but diachronically it represents a secondary formation. Current linguistic scholarship (Whitman 1985: 244; Takeuchi 1999: 91; Unger 2000: 664; Vovin 2003: 168; Robbeets 2005: 158-159; Frellesvig 2010: 112) would mostly agree with Ōno (1953) that the *a*- stem of consonant verbs is nothing but a surface stem that diachronically reflects resegmentation of suffixes with initial **a*-. However, if -*(a)zi*, -*(a)na*, -*(a)ba*, -*(a)s*-, -*(a)sime*$_2$-, -*(a)ye*-, -*(a)re*-, -*(a)n*-, -*(a)zu* -*(a)m*- and -*(a)masi* all go back to suffixes with an initial pJ **a*, one could wonder why proto-Japanese had so many **a* initial suffixes in comparison to suffixes beginning with phonemes other than /a/. The answer to this question lies in the assumption that a certain number of these suffixes go back to reinforced periphrastic constructions with

the original copula pJ *_a_- ‘to be’. The tentative OJ -(_a_)_m_-, for instance, incorporates a simplex tentative marker pJ *-_ma_-, which has lexicalized in verb pairs such as OJ _nade_- B ‘to pat, stroke’ -> OJ _nadame$_2$_- B ‘to soothe, placate, pacify (tr.)’, OJ _kakus_- B’ ‘to hide, conceal (tr.)’, OJ _kakur_- B’ ‘hide (intr.)’ -> OJ _kakumap_- B ‘to shelter, give refuge to (tr.)’; OJ _mo$_2$t_- B ‘to hold, have (tr.)’ -> OJ _moto$_2$me$_2$_- B ‘to seek, pursue, desire, request’; OJ _nozok_- A ‘to peek, peer’ -> OJ _nozom_- A ‘hope for, wish for, look for, look over, view (tr.)’; etc. The tentative construction in OJ -_ke$_1$m-u_ in example (1) derives from a word-final past suffix OJ -_ki$_1$_ followed by a nominalized form of the tentative OJ _am_-, which suggests an independent copular origin for OJ _am_-, i.e. < pJ *_a-ma_- be-TENT-.

(1) OJ _simo$_1$-no$_2$_ _pur-i-k-e$_1$m-u_
frost-GEN fall-CONV-PST.FIN-TENT-NML
‘the fact that the frost would have fallen’
(MYS V: 804; Vovin 2009: 805)

Similar derivations can be made for the passive markers OJ -_(a)ye_- and -_(a)re_-, since these forms incorporate simplex suffixes pJ *-_ya_- passive and anticausative pJ *-_ra_-. The passive pJ *-_ya_- has lexicalized in verb pairs such as OJ _i_- ?A ‘to shoot’ -> OJ _iy_-, _iye_- ?A ‘to get shot’; OJ _ki$_1$k_- A ‘to hear’ -> OJ _ki$_1$ko$_2$ye_- A ‘to be heard, be audible’ and OJ _mi$_1$_- B ‘to see’ -> OJ _mi$_1$yar_- ‘to view the distance, overlook, survey’, OJ _mi$_1$ye_- B ‘to be seen, seem, be visible’ (see also Unger this volume: xx), while the anticausative pJ *-_ra_- is reflected in pairs such as OJ _aka_- A ‘to be clear, bright, red’ -> OJ _akar_- A ‘to brighten, redden (intr.)’ ; OJ _kap_- A ‘to transfer, exchange, buy (tr.)’ -> OJ _kapar_- A ‘to change, be substituted for (intr.)’

and OJ *oti*- B ‘to fall’ (<**ətə-Ci*-) **->** OJ *otor*- ?B ‘to be inferior, fall behind (intr.)’. Therefore, OJ *-(a)ye*- and *-(a)re*- can be derived from the copula pJ **a*- ‘to be’ plus the original passive **-ya*- or anticausative **-ra*- and the intransitivity polarizer **-Ci*-.

Further internal support for the reconstruction of the copula pJ **a*- ‘to be’ is found in J *aru* B, OJ *ar*- ‘to exist’ (< **a*- + **-ra*- anticausative); OJ *-aku* bound noun (< **a*- + ***-ku** nominalizer); J *eru*, OJ *e*- ‘to get, obtain’ (< **a*- + **-Ci*- causative-anticausative) and in the Ryukyuan perfective participles, e.g. Shodon *-an, -ar, -am* (< **a*- + pJ **-n*, **-ra*, **-m* nominalizer). Given the traces of switched adjective encoding in Japanese (Robbeets: forthcoming a), we can probably reconstruct an alternation between verbally encoded pJ **a*- ‘to be’ and nominally encoded pJ **a* ‘being’, the latter of which grammaticalized into the nominalizing suffix pJ **-a* hypothesized by Sakakura (1966: 286-303), e.g. in OJ *par*- (B) ‘open ground, clear land (for cultivation)’ -> OJ *para* (2.3) ‘field, plain, prairie’, OJ *nap*- (B) ‘twist, make rope’ -> OJ *napa* (2.3) ‘rope, cord’, etc.

Martin (1992: 70, 1996: 13, 2006: 222) has compared the Japanese copula to the Korean converb suffix K *-e/a*, MK *-·e/a*. Many of the Korean monosyllabic high-accent stems that end in a vowel lose the accent in common paradigmatic forms but retain it before the converb *·e/a*. This seems to indicate that the converb was originally a bound stem. Indications that the vowel harmonic alternant *-·e* developed after the grammaticalization of the copula pK **a*- into the converb suffix have been preserved in some dialects and early texts, which use only *-a*, regardless of the preceding vowel. Given that switched adjective encoding is also preserved in Middle Korean (Robbeets: forthcoming a), we can reconstruct an alternation between verbally encoded pK **a*- ‘to be’ and

nominally encoded pK *_a_ 'being', the latter grammaticalizing into the converb marker. Moreover, it can be noted that the rising tone in the verb MK ¨_et_- 'to get, receive' indicates the contraction of two syllables. The front vowel of the verb may have resulted from the contraction of the copula pK *_a_- 'to be' with a causative-passive suffix *-_ti_- (Robbeets 2007).

A straightforward cognate for pJ *_a_- 'to be' is pMo *_a_- 'to be', which is reflected as (SH) MMo. _a_- 'to stay, live, be', WMo. _a_- 'to be', Khal. _a_- 'to be', Dag. _a:_- 'to be' and Mog. ʌ- 'to be'. Kane's (2009: 158) reconstruction of Khitan *_a_- 'to be' suggests that the root was already present in proto-Khitan-Mongolic.

A Turkic cognate may be the suffix OTk. -_A_-, which derives verbs from nouns (Erdal 1991: 418-429), e.g. OTk. _kor_ 'loss, damage' -> _kora_- 'to suffer loss, to get destroyed'. Quite commonly, the bases are deverbal nouns, derived, for instance, with the deverbal noun suffix -_(X)n_, e.g. OTk. _es_- 'to blow (gently) (intr. /tr.)' -> _esin_ 'a breeze' -> _esn-e_- 'to blow (of a breeze) (intr.)'. Erdal (1991: 434) finds that derivations of -_Xn_ deverbal nouns with the default denominal verbalizer OTk. -_lA_- are remarkably rare, while derivations with -_A_- are surprisingly numerous. The observation that the verbalizer -_A_- is always associated with -_Xn_ deverbal nouns suggests that the origin of -_A_- may be a copula 'to be'. This hypothesis is supported by the proposal in Section 3.2.5 that the so-called "aorist" OTk. -_Ar_ may derive from a copula pTk *_a_- 'to be' and the deverbal noun suffix *-_r(V)._ Note that the replacement of verb suffixes by a reinforced periphrastic construction consisting of a copula and the same suffix, similar to the development reconstructed for OJ -_(a)m_-, -_(a)ye_- and -_(a)re_-, seems to be a recurrent tendency

across the Transeurasian languages.[2]

*3.2. The nominalizer pTEA *-rA as the origin of the rentaikei (adnominal form)*

*3.2.1. pJ *(wo)-ra*

The deverbal noun suffix pJ *-*ra* can be reconstructed as a suffix that derived nominal and adnominal forms from verbal adjectives such as in the examples in (2a). Clausal (ad)nominalization made use of a suffix pJ *-*oro* reflected as the so-called "*rentaikei*" -*uru*/ -*ru* / -*u* in Old Japanese and as *-*uru* / -*ru* in the Ryukyuan languages, which may go back to a complex form pJ **wo-ra* consisting of a copula **wo*- and the deverbal noun suffix *-*ra*.[3] This is illustrated by the Old Japanese complement clause in (2b) and the relative clause in (2c). This analysis suggests that originally, lexical nouns were derived by adding the suffix pJ *-*ra* directly to verbal adjective stems, whereas verb stems involved in clausal nominalizations incorporated the copula **wo*- 'to sit, be'. The relative clauses with OJ -*uru* developed one step further to mark syntactically independent

[2] Joseph (this volume) mentions a similar development in Polish, whereby the past tense endings are renewed by past tense forms of the copula *być* 'to be'.

[3] The copula **wo*- 'to sit, be' can be reconstructed on the basis of J *iru* A, OJ *wi*- 'to sit, be' (< **wo*- + *-*Ci*-), J *oru* A, OJ *wor*- 'to be, exist' (< **wo*- + *-*ra*- anticausative) and OJ *wos*- ?A 'deign to control/rule/eat/drink/ wear' (< **wo*- + *-*sa*- causative). Moreover, Yanagida and Whitman (2009: 127-129, 134) suggest that the object marker OJ *wo* in nominalized clauses of the type O-*wo* S-*ga* V is really a focus marker and that it has grammaticalized from an original copular verb. Even if Old Japanese makes no distinction between o_1 (< **o*) and o_2 (< **ə*) after *w,* attestations such as Shuri *jijuN* reflect an original pR **wir*- 'to sit', which derives from pJ **wo-(C)i*- 'to sit, be' (Thorpe 1983: 328-29; Pellard 2011: 10).

sentences as in (2d), signaling the evaluative nature of the proposition in discourse. In such cases, the finite form could be accompanied by focus particles specifying the exact nature of the speakers' reaction, such as question, exclamation, confirmation, explanation etc., a phenomenon known as *kakari-musubi* in Japanese.

(2) Reflexes of the deverbal noun suffix pJ *-*ra* in Western Old Japanese

a. Lexical (ad)nominalizer

OJ *aka*- 'to be red' -> *akara* 'red'

OJ *sakasi*- 'to be wise' -> *sakasira* 'wisdom'

OJ *kanasi*- 'to be sad' -> *kanasira* 'sadness'

b. Clausal nominalizer

punapi$_1$to$_2$-wo *mi$_1$-**ru**-ga* *to$_2$mo$_2$si-sa*

boat.people-ACC see-**NML**-GEN enviable-NML

'it is enviable to see the boat-people'

(MYS 15: 3658; Wrona 2008: 206)

c. Relativizer

*op-i$_1$-k-**uru*** *mo$_2$no$_2$*

pursue-CONV-come-**ADN** thing

'the things that pursue [us]'

(MYS 5: 804; Vovin 2009: 613)

d. Finite marker

ide ika-ni *kokodaku* *ko$_1$p-**uru***

Oh why-DAT so.much love-**FIN**

'Oh, why do I love her this much?'

(MYS 12: 2889; Wrona 2008: 206)

*3.2.2. pK *-(o-)l*

The contemporary Korean adnominalizer K *-(u)l* is usually called "prospective" but in Middle Korean, *-(·u/o)l* generally functions as a time neutral marker (Martin 2002: 376). The examples in (3a) suggest that pK *-*l* began as a lexical nominalizer applied to verb stems to create nouns. Gradually, it became used for clausal nominalization as in (3b) and relativization in (3c). The so-called "modulator" MK -·w^u/*o*-, which has been derived from an original copula pK **o*- 'to be' by Martin (1996: 13, 83; 2006: 222), appears before MK *-(·u/o)l* in case the modified noun is semantically the object of the adnominalized verb, as in example (3c). This suggests that, originally, lexical nouns and clausal nominalizations of intransitive verbs were derived by adding the suffix **-l* directly to verbal stems, whereas clausal nominalizations of transitive verbs could incorporate the copula **o*- 'to be'. The relative clauses marked with pK *-*(o)-l* developed one step further to mark syntactically independent sentences which added supplementary information in discourse such as question (e.g. MK -*(·u/o)l-·kwo* finite interrogative), confirmation / explanation (e.g. MK -*(·u/o)·l i* finite explicit) and exclamation MK -*(·u/o)·l-a* (finite subjunctive attentive*).* As illustrated in (3d), the subjunctive attentive is morphologically segmentable into the imperfective adnominalizer and the vocative particle *a,* which usually follows nouns.

(3) Reflexes of the deverbal nominalizer pK **-l* in Middle Korean

a. Lexical nominalizer

MK *kuch-* ‘to stop’ -> *ku·chul* ‘cessation’

MK *hhoyng ho-* ‘to travel’ -> *hhoyng hol* ‘traveler’

MK *ciž-* ‘to make, build, compose, manufacture’ -> MK *·cil* ‘procedure’ (< **cižul*)

b. Clausal nominalizer

¨se *·twoy* *tu· li-**l**-s* *HHWA-PPYENG-··ul*

three measure contain-**NML**-GEN vase-ACC

nwo-··khwo

place-CONV

‘Placing a vase with a capacity of three cupfuls’

(1459 Wel 10: 119 a; Martin 1992: 873)

c. Relativizer

ccywung-soyng-oy *nip-wu-**l*** *wos*

common.people-NOM wear-MOD-**ADN** clothing

‘clothes that the common people wear’

(1459 Wel 8: 65; Lee & Ramsey 2011: 206)

d. Finite marker

·QILQ-SIM-··u·lwo *kwoyGwoy* *ho-··**l**-a*

wholehearted-ADV silence do-**FIN**-VOC

‘Be utterly quiet!’

(1464 Kumkang 12a; Martin 1992: 851)

Although MK *-(·u/o)l* generally functions as a time neutral adnominalizer, it is interesting to note some quirky behavior in some time expressions such as K *onul*, MK *wo·nol* ‘today’ and K *wolhay*, MK *wol ·hoy* ‘this year’. These

lexicalizations contain an adnominal form of the verb K *o-*, MK *wo-* 'to come', deriving from **o-[l] ·nal* [come-ADN day] and **o-l ·hoy* [come-ADN year], respectively.[4] Since 'today' and 'this year' are not equivalent to 'the coming day' and 'the coming year', but rather should be interpreted as 'the day that has (just) come' and 'the year that has (just) come', these expressions suggest an original (recent) perfective interpretation of pK **-l* following telic verbs. Compare the use of MK *·wo-no-n ·hoy* [come-PROC-ADN year] for 'next year', i.e. 'the year that is coming' and MK *·ni-·ke-n ·hoy* [depart-RES-ADN year] for 'last year', i.e. 'the year that has departed'.

*3.2.3. pTg *-rA*

The deverbal noun suffix pTg **-rA* can be reconstructed as a suffix that derived nominal and adnominal forms from verb stems such as in the Even and Evenki examples in (4a). The (ad)nominalizers were then extended to the clause level to mark complement and relative clauses, such as in the Manchu examples in (4b/c). The relative clauses developed one step further to mark syntactically independent sentences as in the example in (4d).

(4) Reflexes of the deverbal noun suffix pTg **-rA* in Tungusic

a. Lexical (ad)nominalizer in Even and Evenki

Even *da:l-* 'to be sweet, pleasant, light' -> *dalra* 'sweet, tasty'

Evk. *langa-* 'to break a tooth' -> *langara* 'toothless'

b. Clausal nominalizer in Manchu

[4] MK /l/ drops before /n/ and the other apicals /s/, /c/ and /t/.

mama-de *ala-**ra**-de,*

old.woman-DAT tell-**NML**-DAT

mama *hendu-me...*

old.woman say-CONV

'When [he] tells [it] to the old woman, the old woman says: …'

(Gorelova 2002: 257)

c. Relativizer in Manchu

*bargiyata-**ra*** *niyalma*

protect-**ADN** people

'people who protect [him]' (Gorelova 2002: 485)

d. Finite marker in Manchu

si *nene-me* *isinji-ci* *uthai* *sin-de* *bu-**re***

you be.first-CONV come-CONV at.once you-DAT give-**FIN**

'If you come first, I shall give [it] to you straight away'

(Gorelova 2002: 256)

It is relevant to note that in some Northern Tungusic languages, such as Even and Evenki, the finite temporal interpretation depends on actional semantics of the verb: derived from telic verbs, *-rA* refers to the recent past, whereas derived from atelic verbs, it refers to the present; e.g. Even *em-re-n* [come-FIN-3SG] '(he) has just come' vs. *hong-ra-n* [weep-FIN-3SG] 'he weeps' (Malchukov 2000: 443). This suggests that the proto-Tungusic clausal adnominalizer pTg *-*rA* was interpreted as imperfective or perfective according to the telicity of the verb base and recalls the etymology of Korean 'today' and 'this year'.

*3.2.4. pMo *-r*

The deverbal noun suffix pMo *-*r* reflected in MMo. / WMo. *-(U)r* began as a suffix that derived nominal and adnominal forms from verb stems, such as in the examples in (5a). There are relics of clausal nominalization in the final converb in *-rA,* illustrated in (5b)*,* which can be derived from *-*r* marking a complement clause plus the dative suffix in *-*A*. Similarly, the preparative converb on *-run* is a compound of *-*r* and the genitive suffix in *-*un* (Poppe 1954: 59, 98, 180). There is no evidence that pMo *-*r* developed further to a finite marker in Mongolic proper, but Khitan uses a past tense suffix *-r*, illustrated in (5c), which is probably related. Recalling the situation in Korean and Tungusic, the past tense use in Khitan is only preserved in telic expressions, such as 'become', 'become appointed', 'become awarded', 'compose an edict', 'write this text' (Kane 2009: 145-146).

(5) Reflexes of the deverbal noun suffix pMo *-*r* in Mongolic

a. Lexical (ad)nominalizer in Written Mongolian

WMo. *amu-* 'to rest, relax; be relieved (intr.)' -> *amur* 'peace, rest; easy'

WMo. *belčige-* 'to pasture, graze (tr.)' -> *belčiger* 'pasture, grazing grounds'

WMo. *irüge-* 'bless, pray, wish well (tr. /intr.)' -> *irüger* 'prayer, blessing'

b. Clausal nominalizer in Written Mongolian

eke-yügen *eri-**re*** *od-bai*

mother-ACC search-**CONV** go-PST.FIN

'He went to find her mother'

(Sárközi 2004: 47)

c. Finite marker in Khitan

*puu giuuŋ shï po-**or***

fu gong shi become-**PST.FIN**

'He was appointed a fu gong shi'

(Kane 2009: 146)

*3.2.5. pTk *-rV*

The deverbal noun suffix pTk *-*rV* is reflected as a suffix that derived nominal and adnominal forms in Old Turkic, such as in the examples in (6a). It is formed with -*Ar* after most simple consonant stems, with -*Ur* or -*Ir* after diathetic consonants stems and with -*yUr* or -*r* after vowel stems.[5] The (ad)nominalizers were then extended to the clause level to mark clausal (ad)nominalization as in (6b). Relative clauses developed one step further to mark syntactically independent sentences with present continuous meaning as in (6c), also known under the label "aorist".

(6) Reflexes of the deverbal noun suffix pTk *-*rV* in Old Turkic

a. Lexical nominalizer

[5] The allomorphs -*yUr* /-*Ur* and -*Ar* may derive from suffix strings in which pTk *-*rV* follows the copular verbs **u*- 'to become' and **a*- 'to be', respectively, while the allomorph -*Ir* would reflect a stem-final -*i*. The final vowel in the reconstruction of pTk *-*rV* is supported by the negative adnominalizer OTk. -*mA-z*, in which -*z* follows the negative suffix, where the positive adnominalizer has -*r* after vowels. The assumption that in coda position pTk *-*r* became -*z* in Eastern Old Turkic, but was preserved as -*r* in Western Old Turkic and its modern representative Chuvash could very well apply to the development of OTk. -*mA-z*, if we assume that the suffix originally had an additional vowel (Erdal 2004: 84-85).

OTk. *kïs-* ‘to compress, squeeze, pinch’ -> *kïsïr* ‘having the sexual organs constricted, sterile, barren (of woman, animal)’

OTk. *tug-* ‘to be born, to rise (of sun) (intr.)’ -> *tugar* ‘sunrise, east’

OTk. *yat-* ‘to lie down (intr.)’ -> *yatar* / *yatur* ‘(something) lying down, invalid’

b. Relativizer

ak-ïp *kel-**ir*** *sogïk* *suv*

flow-CONV come-**ADN** cold water

‘cold water flowing forth (or coming up)’

(Erdal 2004: 284-285)

c. Finite marker

ölüm-tä *oz-upan* *ögir-ä* *savin-ü* *yorï-**r.***

death-ABL escape-CONV rejoice-CONV be.happy-CONV go.on-**FIN**

‘Having been saved from death it happily goes on with its life.’

(Erdal 2004: 325)

A possible cognate for this suffix is Chu *-r*, which occurs in similar lexical nominalizations, as a perfective relativizer as well as a finite past form; see example (7). In the common ancestor of Chuvash and Old Turkic, the finite temporal interpretation of the relativizer may have depended on the actional semantics of the verb base: derived from telic verbs it may have derived perfective and past forms, whereas derived from atelic verbs it may have derived imperfective and continuous present forms. In Chuvash, the original functional distinction disappeared in favor of perfective and past meaning, while Old Turkic favored imperfective and continuous present meaning.

(7) Reflexes of the deverbal noun suffix pTk *-*rV* in Chuvash

a. Lexical nominalizer

Chu. *xĕs-* 'to compress, squeeze, pinch (tr.)' -> *xĕsĕr* 'sterile, barren'

b. Relativizer

xura *vărman* *vitĕr* *tux-**r**-ăm* *čux-ne*

black forest through go.out-**PFV.NML**-POSS.1SG time-DAT

'When I went out through the black forest' (Benzing 1959: 742)

c. Finite marker

văl *sirĕ* *palla-**r**-ĕ*

he you.OBL recognize-**PST.FIN**-POSS.3SG

'He recognized you' (Krüger 1961: 146)

Note that the cognate Yakut suffix *-Ar-* expresses the so-called "synthetic imperfect". Johanson (this volume: xx) finds that the preterite use of this marker is highly remarkable for a Turkic language and considers the possibility that it has been copied from Even. Alternatively, under the present analysis, the preterite use may represent an inherently Turkic feature.

*3.3. The substantivized adnominal pTEA *-rA-i as the origin of the izenkei (subjunctive form)*

*3.3.1. pJ *(wo)-ra-i*

In Old Japanese, the so-called "*izenkei*" or subjunctive suffix *-ure* was used as a finite form signaling the strong evaluative nature of the proposition as illustrated

in (8). In line with Unger's (2000: 664) analysis, this form can be derived from the adnominal form discussed in Section 3.2.1, reflecting a suffix string pJ **wo-ra-i*, in which *-i* is a substantivizer following adjectival nouns. This substantivizer was added to nominal adjectives such as OJ *taka* 'high' -> *take*$_2$ 'peak'; OJ *awo* 'blue' -> *awi* 'indigo plant' and OJ *aka* 'red' -> *ake*$_2$ 'red object, red cloth' and is assumed to be cognate with the bound noun OJ *i* 'fact (that); that (which)' (Whitman 1985: 44, 246; Martin 1987: 64-65, 420). Thus, the subjunctive form is derived from a substantivizer 'fact (that)' attached to an adnominal form of the verb, which grammaticalized into a finite verb marker to signal the speaker's evaluation of the proposition.

(8) The use of the Old Japanese subjunctive *-ure*

ip-u	*ko$_2$to$_2$*	*yam-i$_1$*	*…*	*ino$_2$ti*	*taye-n-**ure**.*
say-ADN		thing stop-CONV	…	life	cease-PERF-SUBJ

'[he] stopped speaking and [his] … life ended'

(MYS V: 904; Vovin 2009: 704)

*3.3.2. pK *(o)-l-i*

The Korean ending of explicit statement K *-uli*, MK *-(·u/o)·l i* is illustrated in (9). Parallel to the derivation of the Old Japanese subjunctive *-ure*, it derives from the adnominalizer MK *-(·u/o)l* and a bound noun MK *i* 'fact (that); that (which)'.

(9) The use of the Middle Korean explicit *-(·u/o)·li*

·NGWOY-¨TTWOW	*'y-n*	*·tol*	*a·ni*	*cwocco·W-oli*
heretical	be-ADN	things	NEG	follow-EXP

‘They will not follow any false doctrines’

(1449 Kok 99; Martin 1992: 856-857)

*3.3.3. pTg *-rA-i*

In the Southern Tungusic languages, the suffix *-*ri:* is gradually replacing the reflex of the adnominalizer pTg *-*rA* and even spreading to some finite forms. In Udehe, this has happened in the third person finite form, as illustrated in example (10). Parallel to the derivation of the Japanese and Korean forms, the Tungusic suffix pTg *-*ri:* can be derived from the (ad)nominalizer *-*ra* and a substantivizer pTg *-*i:* (Menges 1968: 205). The substantivizer derives ordinal from cardinal numbers, e.g. Evk. *ïlan* ‘three’ -> *ïli:* ‘the third one, third’ (< pTg **ila-i*: three-thing) and may be related to the third person pronoun Manchu *i* (Benzing 1955: 1051).

(10) The use of the Udehe finite -*i*

*wasia mäusa-wa zawa-**i**-ni*

Vasya gun-ACC take-**FIN**-3SG

‘Vasya is taking the gun’

(Nikolaeva 1999: 146)

*3.4. The nominalizer pTEA *-mA as the origin of the shūshikei (finite indicative form)*

*3.4.1. pJ *(-wo)-m*

The deverbal noun suffix pJ *-*m* can be reconstructed as a suffix that derived nominal and adnominal forms from verbal adjectives such as in the examples in (11a). The evidence comes from the accent class 2.5, the origin of which was linked with the loss of nominalizer pJ *-*m* — frequently denoting colors — in verbal adjectives (Polivanov 1924: 126; Vovin 2008: 142-150). The vowel alternation in some adjectives, such as OJ *kura*- 'to be dark' ~ OJ *kuro*$_1$ 'black' and OJ *sira*-~ OJ *siro*$_1$ 2.5. 'white' suggests that the copula **wo*- 'to be' may be involved in the derivation, yielding **kura-wo-m* [thick-COP-NML] and **sira-wo-m* [thick-COP-NML] respectively. Clausal nominalization of verbs makes use of a suffix pJ *-*om* reflected as -*u* in Old Japanese and as *-*um* in the Ryukyuan languages, which may go back to a complex form pJ **wo-m* consisting of a copula **wo*- and the deverbal noun suffix *-*m.* Although the standard use of the so-called "*shūshikei*" OJ -*u* and Ryukyuan *-*um* is finite indicative as in (11c), we find a few relic examples of clausal nominalization following the negative suffix OJ -*(a)z*- in constructions where the negative nominalizer -*(a)zu* occurs before the converb *ni* of the defective copula *n*- as in (11b).

(11) Reflexes of the deverbal noun suffix pJ *-*m* in Old Japanese

a. Lexical nominalizer

OJ *awo*- B 'to be blue/green' -> *awo* 2.5 'blue/green (n.)' (< **awo-m* blue-NML)

OJ *kura*- B 'to be dark' -> OJ *kuro*$_1$ 2.5 'black' (< **kura-wo-m* black-be-NML)

OJ *sira*- B 'to be white' -> OJ *siro*$_1$ 2.5. 'white' (< **sira-wo-m* white-be-NML)

b. Clausal nominalizer

amata *pa* *ne-z-**u*** *n-i*

many TOP sleep-NEG-**NML** be-CONV

tada *pi₁to₂ yo₁* *no₂* *mi₂*

only one night PT

'not sleeping [with her] many [nights], only one night'

(NK 66; Vovin 2009: 761)

c. Finite marker

aki₁-no₂ *no₁-ni* *sawosika* *nak-i₁-t-**u**.*

autumn-GEN field-LOC male.deer cry-CONV-PERF-**FIN**

'Male deer cried in the autumn field.'

(MYS 25:3678; Vovin 2009: 602)

*3.4.2. pK *-(o)-m*

As in Contemporary Korean, the Middle Korean deverbal noun suffix MK *-(·u/o)m* was used to derive lexical nouns and to nominalize sentences, but lexical nouns were generally derived by adding the suffix directly to verb stems, while clausal nominalizations incorporated the modulator MK -·w^u/*o-*, which has been traced back to the copula pK **o*- 'to be', see (12a). As illustrated in (12b), the suffix has developed to a marker of finiteness, a position in which it is always followed by the vocative particle *a*. In the documentary style of written contemporary Korean, however, K *-(u)m* appears in main clauses without the vocative, often expressing an impersonal proposition, as in *onul-un swuep-i eps-um* [today-TOP class-NOM not.exist-NML] 'No class today.'

(12) Reflexes of the deverbal noun suffix pK *-*m* in Middle Korean

a. Lexical nominalization vs. clausal nominalization

tywoh-on *yel-um* *yel-wu-m-i*

be.good-ADN bear.fruit-NML bear.fruit-MOD-NML-NOM

'the bearing of good fruit'

(1459 Wel 1: 12; Lee & Ramsey 2011: 177)

b. Independent sentence

·na-y *ne* *to·ly-e* *nil·G-wo-·m-a*

I-NOM you accompany-CONV say-MOD-FIN-VOC

'I will tell you.'

(1517 Pak 1: 32b; Martin 1992: 932)

*3.4.3 pTg *-mA*

The deverbal noun suffix pTg *-*mA* can be reconstructed as a suffix that derived nominal and adnominal forms from verb stems, such as the Evenki examples in (13a). It is particularly frequent in the derivation of colour nouns and adjectives, which recalls the color derivations in Japanese. There are relics of clausal nominalization in the converb suffix, which is also used in complement clauses, such as in (13b). Most Tungusic languages display distinct singular and plural forms, e.g. Evk. *-mi* / *-mil*, Nanai *-mi* / *-mari* / *-meri* , Ud. *-mi* / *-mei*, Olč. *-mi* / *-mari* / *-meri*, Oroč *-mi* / *-mai*. This number distinction reflects the contraction of an original clausal nominalizer pTg *-*mA* with the possessive-reflexive suffixes pTg *-*wi* singular and pTg *-*wari* plural, respectively (Benzing 1955: 1090; Menges 1968: 212). In Sibe, a contemporary descendant of Manchu, the corresponding nominalizer *-m* is still productive for marking independent

clauses as in (13c).

(13) Reflexes of the deverbal noun suffix pTg *-*mA* in Tungusic

a. Lexical nominalizer in Evenki

Evk. *bagda-* 'to become white, freeze' -> *bagdama* 'white (adj./n.)'

Evk. *girku-* 'to walk' -> *girkuma* 'pedestrian'

Evk. *koŋno-* 'to be black' -> *koŋnomo* 'black (adj./n.)'

b. Clausal nominalizer in Evenki

bejetken	*alba-ra-n*	*bira-va*	*elbesce-mi*
boy	can.not-FIN-3SG	river-ACC	swim-NML

'The boy could not swim across the river'

(Nedjalkov 1995: 457)

c. Finite marker in Sibe

am	*nane-ni*	*gel*	*xia-ve-mak*	*niumku bahe-**m**.*
big	person-DEF	also	bite-PASS-CONV	disease get-**IPF.FIN**

'Even adults get bitten and get disease.'

(Jang, Jang & Payne (in prep.))

*3.4.4. pMo *-m(A)*

The deverbal noun suffix pMo *-*mA* alternates with *-*m* and can be reconstructed as a suffix that derived nominal and adnominal forms from verb stems, such as in the examples in (14a). There are doublets, such as WMo. *degerem* ~ *degerme* 'robbery, robber' and *toγum* ~ *toγuma* 'sensibleness; good behavior', that indicate that the nominalizers *-*mA* and *-*m* share a common origin. Janhunen (2012: 166-167) notes that the marker of the preconditional converb *-mAA/n* in

the Central Mongolic languages, e.g. *sour-maa/n* [study-CONV] 'only if you study', can be derived from the deverbal noun suffix *-m* and the reflexive possessive marker *-AA/n*. Similar to the Tungusic converb in Section 3.4.3, clausal adverbialization can here be derived from clausal nominalization. In Middle Mongolian texts of the thirteenth and fourteenth century, *-m* is the common ending for the imperfective present indicative (Poppe 1955: 261; Weiers 1966: 143-150); see (14b).

(14) Reflexes of the deverbal noun suffix pMo **-mA ~ *-m* in Mongolic

a. Lexical (ad)nominalizer

WMo. *ulayi-* 'to get red-hot, become red (intr.)' -> *ulayima* 'red, red-hot'

WMo. *toγu-* 'to esteem, value' -> *toγum ~ toγuma* 'sensibleness, good behavior'

pMo **degere-* 'to lift (tr.)' in WMo. *degerede-* 'to be lifted' -> *degerem ~ degerme* 'robbery, robber'

b. Finite marker in Middle Mongolian

udurit-basu ber ulu busire-m.

guide-COND PT NEG believe-IPF.FIN

'Even if you guide them, they don't believe'

(HY; Weiers 1966: 144)

*3.4.5. pTk *-m(A)*

The deverbal noun suffix pTk **-mA* alternates with **-m* and can be reconstructed as a suffix that derived nominal and adnominal forms from verb stems, such as in the Old Turkic examples in (15a). Since OTk. *-mA* and *-(X)m* have similar

functions and produce near doublets such as OTk. *örüm* 'something knitted' ~ *örma* 'plaited', they probably go back to a single origin. There are some rare cases of relative clauses, such as (15b) below, in which the noun *but* 'leg' is governed by the *-mA* form, but there are no indications that this suffix has developed finite function.

(15) Reflexes of the deverbal noun suffix pTk *-*mA* ~ *-*m* in Old Turkic

a. Lexical (ad)nominalizer

OTk. *tut-* 'to hold, grasp, seize (tr.)' -> *tutma* 'chest, coffer' ~ *tutum* 'handful'

OTk. *yar-* 'to split (open) (tr.)' -> *yarma* 'crack' ~ *yarïm* 'half'

OTk. *ör-* 'to plait, knit' -> OTk. *örüm* 'something knitted' ~ *örma* 'plaited'

b. Relativizer

but *kötür-me* *tïnlïg*

leg lift.up-ADN living.being

'a living being lifted up by [its] legs'

(Erdal 1991: 319)

*3.5. The nominalizer pTEA *-i ~ ø as the origin of the renyōkei (conjunctive form)*

*3.5.1. pJ *-i ~ ø*

The deverbal noun suffix pJ *-*i* ~ ø* is reflected in numerous Old Japanese nouns, such as those illustrated in (16a). Derived nouns show clear indications of lexicalization because some verbs lack a nominal form, some meanings have specialized (e.g. *momi* 'cloth rub-dyed solid red'), and the accent has

neutralized.[6] Derived action nouns gradually developed an infinitive-like function in verb compounding, as illustrated in (16b). A further increase of verbal properties on the part of the infinitive ultimately resulted in the development of converbs as in (16c).[7] The forms in (16b/c) are known as the "*renyōkei*" or 'conjunctive form' in traditional Japanese grammar.

(16) Reflexes of the deverbal noun suffix pJ *-*i* ~ *ø* in Old Japanese

a. Lexical nominalizer

OJ *ko$_1$pi$_2$*- 'to love' -> *ko$_1$pi$_2$* 'love'

OJ *mom*- 'to rub (with both hands), massage' -> *momi* 'cloth rub-dyed solid red'

OJ *omo$_{(1)}$p*- 'to think, feel' -> *omo$_{(1)}$pi$_1$* 'thought'

b. Infinitive

mi$_1$-ko$_2$ko$_2$ro$_2$-wo *sizume$_2$-tamap-u*

HON-heart-ACC calm.down.**INF**-grant-FIN

'[she] deigned to calm down [her] august hart'

(MYS V: 813; Vovin 2009: 1005)

---

[6] According to Martin (1987: 211, 1995: 149), the *-i* converb leads to a change of pitch in the verb at the point where the ending is added, e.g. OJ *kum*- 'to assemble' (B = initial L) -> *kumi* 'assemble and' (LH) and OJ *kob*- 'to flatter' (A = initial H) -> *kobi* (HL) 'flatter', while the deverbal noun suffix simply erases that change: the forms are atonicized, e.g. OJ *kumi* 'set' (LL) and *kobi* 'flattery' (HH).

[7] Infinitives are typically intermediate between deverbal action nouns and converbs in that they occur both in complement clauses and adverbial clauses of purpose (Ylikoski 2003: 200). The verbalization of the form is complete when the complement clause becomes reanalyzed as an adverbial clause (e.g. His crossing the river rapidly frightens you > Crossing the river rapidly, he frightens you), giving rise to a converbial construction.

c. Converb

ip-u *ko₂to₂* *yam-**i₁*** *ino₂ti* *taye-n-ure*

say-ADN thing stop-**CONV** life cease-PERF-SUBJ

'[he] stopped speaking and [his] life ended'

(MYS V: 904; Vovin 2009: 704)

*3.5.2. pK *-i ~ ø*

The deverbal noun suffix MK -·*i* ~ *ø* is incorporated in Middle Korean nouns such as those illustrated in (17a). Historical residue of converbial use of this suffix is left in the ability of some -*i* adverbs to be negated by a sentential negator, such as by the negative adverb *ani* in example (17b).[8] After the converb ceased to be productive, some frequently used converbial forms lexicalized as adverbs, as illustrated in (17c). Note that the verb MK *nwoph*- 'to be high' goes back to an original thematic stem pK **nopkʌ*-, with the vowel reflected in the noun MK *nwo·phoy* 'height', but not in the adverb MK *now·phi* 'highly', which suggests that the converb suffix was productive at a later point in time, when the stem-final vowel had already been dropped.

(17) Reflexes of the deverbal noun suffix pK *-*i* ~ *ø* in Middle Korean

a. Lexical nominalizer

MK *hal*- 'to slander' -> MK *hali* 'slandering'

MK *nwoph*- 'to be high' -> MK *nwo·phoy,* K *nophi* 'height'

[8] Converbial use is also reflected in the ability of some -*i* adverbs to take a subject in the nominative case, e.g. K *eps-i* in *Ku yeca-ka eps-i sa-l su-ka eps-ta* [that woman-NOM not.exist-ADV live-ADN possibility-NOM not.exist.FIN] 'I cannot live without that woman'.

MK *nwu(·)pi-* ‘to quilt’ -> MK *nwu·pi* ‘quilting’

b. Converb

ani *sulph-i* *neki-l* *i* *ep-te-ni*

NEG be.sad-ADV regard-ADN person not.exist-RETR-CONV

‘as there was nobody who did not regard [it] as not sad.’

(Seongha Rhee, p.c.; 1431 Samkang, 26)

c. Adverbializer

MK *nik-* ‘to ripen’ -> MK *ni·ki* ‘thoroughly, ripely’

MK *nwoph-* ‘to be high’ -> MK *now·phi,* K *nophi* ‘highly’

MK *nowoy-* ‘to repeat’ -> MK *nowoy* ‘repeatedly’

*3.5.3. pTg *-i: ~ ø*

The deverbal noun suffix pTg *-*i:* ~ *ø* is mainly reflected in the Northern Tungusic languages, as illustrated by the Even examples in (18). Converbial use of this suffix has not been attested.

(18) Reflexes of the deverbal noun suffix pTg *-*i:* ~ *ø* in Even

Even *das-* ‘to cover, mantle (tr.)’ -> *dasi:* ‘cover, coverage’

Even *dur-* ‘to burn, be on fire, catch fire’ -> *duri*: ‘fire, blaze, forest fire’

Even *ju:pti:-* ‘to double (tr.)’ -> *jupti:* ‘double, twofold’

*3.5.4. pMo *-i ~ ø*

The deverbal noun suffix pMo *-*i* is incorporated in a few Written Mongolian nouns, such as those illustrated in (19a). In Mongolic proper, the suffix is no longer productive as a converb marker but, as illustrated in (19b), Khitan used a

converb in vowel plus -*i*, expressing the meaning 'then, after that' (Kane 2009: 149-150).[9] After the converb ceased to be productive in Mongolic, some frequently used converbial forms probably lexicalized as adverbs, as illustrated in (19c)

(19) Reflexes of the deverbal noun suffix pMo *-*i:* ~ *ø* in Mongolic

a. Lexical nominalizer

WMo. *bü*- 'to be, exist' -> *büi* 'existence, existing'

WMo. *büci*- 'to surround, gather around (tr./ intr.)' -> *büci* 'tie, ribbon, band, lace'

WMo. *muru*- 'to go astray, act contrarily' -> *murui* 'curve, crookedness; bending'

b. Converb in Khitan

*tai zï śiauu sh dem-lege-**ei***

taizi shaoshi grant-PASS-**CONV**

*dieên sieên du dêm gieêm poju-**ii***

palace command chief inspector establish-**CONV**

syiên xuŋ a-ar

zianggün be-PST

'he was given the title of *taizi shaoshi*, was appointed chief inspector of the palace command and had [the position of] court ceremonial.'

(Kane 2009: 152)

[9] The various notations used for the converb marker are <ai>, <ui>, <oi>, <ei>, <ii> and <i>, whereby the final vowel of the preceding stem is often repeated. The repetition of the stem-final vowel is due to the syllabic nature of the Khitan Small Script system.

c. Adverbializer

WMo. *sönü*- ‘to be extinguished, go out (of fire), cease to be’ -> *söni* ‘night, at night’

WMo. *γar*- ‘to go out, pass over, exceed (intr.)’ -> *γarui* ‘more than, beyond, over’

WMo. *daru*- ‘to press, follow, be near’ -> *darui* ‘immediately, at once, thereafter’

*3.5.5. pTk *-I ~ ø*

The deverbal noun suffix pTk *-*I* ~ *ø* is reflected in Old Turkic nouns such as those illustrated in (20a) (Erdal 1991: 340-341). Moreover, Old Turkic makes use of a converb of the shape -*I* after the -*(X)t*- causative suffix and in some exceptional converbial forms such as OTk. *alï*, *barï*, *kalï, keli, siŋi* and *tegi* derived from OTk. *al*- ‘to take’, *bar*- ‘to go to’, *kal*- ‘to remain’, *kel*- ‘to come’, *siŋ*- ‘to penetrate’ and *teg*- ‘to reach’ respectively. This suffix is used as an infinitive-like form in verb compounding as in (20b) as well as in adverbial subordination as in (20c). This seems to reflect a gradual increase in verbal properties, going from action noun to infinitive to converb. Ultimately, petrified converbs, such as those given in (20d) have lexicalized in deverbal adverbs.

(20) Reflexes of the deverbal noun suffix pTk *-*I* ~ *ø* in Old Turkic

a. Lexical nominalizer

OTk *adïr*- ‘to separate (tr.)’ -> *adrï* ‘fork, forked’

OTk. *tög*- ‘to pound, crush (tr.)’ -> *tögi* ‘cleaned and/or crushed cereal’

OTk. *yap*- ‘to cover (tr.)’ -> *yapï* ‘horse-blanket’

b. Infinitive

*aviš tamu-ka bar-**ï** yarlïka-dï*

Avi:ci: hell-DAT go-**INF** deign-PST

'[he] deigned to go in the Avi:ci: Hell'

c. Converb

*toruk at sämrit-**i** [...] yügür-ü bar-mïš*

lean horse make.fatt-**CONV** run-CONV go-INFR

'After a lean horse fattened itself, it went running'

d. Adverbializer

OTk *bar-* 'to go to' -> *barï* 'as much as there is'

OTk. *körüš-* 'to see eachother' -> *körši* 'with a view on, in view of'

OTk. *tak-* 'to attach' -> *takï* 'more, yet, and, too, also'

3.6. The Transeurasian zero imperative as the origin of the meireikei (imperative form)

Whereas consonant final stems add a suffix OJ -e_1 to form the so-called *meireikei* or imperative form (e.g. OJ *kak-*e_1 'write!', *sin-e* 'die!', *ar-e* 'exist!'), the vowel final stems simply use the bare verb stem (e.g. OJ ko_2 'come!', ake_2 'open!', *mi* 'look!'). Most zero imperatives can be reinforced by the exclamatory particle OJ yo_2 (e.g. OJ ake_2 yo_2 'open!', *mi* yo_2 'look!', oki_2 yo_2 'arise!'), which was in the course of changing to an inflectional ending marking of the imperative itself (Vovin 2009: 655; Frellesvig 2010: 101). According to Vovin (2009: 647), the suffix OJ -e_1 can be derived as an instance of monophthongization of the converb suffix pJ *-*i* plus an auxiliary **a* in the imperative. It is, therefore, inviting to identify pJ **a* as the

bare imperative form of the copula pJ *_a_- ‘to be’. As such, bare verb stems seem to represent the original strategy to form imperatives in Japanese.

As is the case for Japanese, imperatives can be built with special endings, such as desiderative, optative, volitional or exclamatory suffixes in various other Transeurasian languages but most languages seem to share an original imperative that is formed on the basis of the bare verb stem alone. The Korean imperative ending in the intimate style, for instance, is K -*e/a,* MK -*·e/a* (e.g. K *mek-e* ‘eat!’*)*. Similar to the homophonous converb suffix K -*e/a* above, its particular accentuation enables us to derive it from an original auxiliary, i.e. most probably the bare stem of the copula pK *_a_-. This formation parallels that of the imperatives derived from Old Japanese consonant final stems. Among the Tungusic languages only Manchu and Sibe have preserved a zero imperative, e.g. Ma. *gene* ‘go!’, *te* ‘sit down!’, *ala* ‘tell!’. The imperative in the Mongolic languages is also expressed by a bare verb stem, e.g. Khal. *ir*, WMo. *ire* ‘come!’, Khal. *soo*, WMo. *sayu* ‘sit down!’. Likewise, in the Turkic languages, the absence of verb suffixes indicates that the form is to be understood as an imperative, e.g. OTk. *kel* ‘come!’, *két*! ‘go!’. The sole observation that all Transeurasian languages have zero imperatives is weak evidence of relatedness because cross-linguistically imperatives tend to be expressed by the least marked verb forms. However, searching for potential Transeurasian cognates for a closed set of basic Japanese inflections, the paradigmatic opposition of zero-imperatives to the other inflectional suffixes is telling because the imperative fills a cell in the closed set of form slots.

4. Paradigmaticity of the evidence

4.1. Is the evidence paradigmatic?

Table 3 summarizes the formal and functional correlations for the basic inflectional paradigm of Japanese verbs in the Transeurasian languages.

Table 3. The Japanese basic inflectional paradigm in a comparative Transeurasian perspective

	pJ	pK	pTg	pMo/pKMo	pTk	pTEA
Mizenkei pseudostem	**a-* 'to be'	**a-* 'to be'		**a-* 'to be'	**a-* 'to be'	**a-* 'to be'
Rentaikei adnominal	**-ra* lexical NML **-wo-ra* clausal NML clausal ADN FIN	**-l* lexical NML **-o-l* clausal NML clausal ADN FIN	**-rA* lexical NML clausal NML clausal ADN FIN	**-r* lexical NML clausal NML - FIN	**-rV* lexical NML - clausal ADN FIN	**-rA* NML
Izenkei subjunctive	**wo-ra-i* be-ADN-SUB	**o-l-i* be-ADN-SUB	**-rA-i* ADN-SUB			**-rA* NML **i* 'fact'
Shūshikei indicative	**-m* lexical NML **-wo-m* clausal NML FIN	**-m* lexical NML **-wo-m* clausal NML FIN	**-mA* lexical NML clausal NML FIN	**-mA* ~ **-m* lexical NML clausal NML FIN	**-mA* ~ **-m* lexical NML	**-mA* NML
Renyōkei converb	**-i* ~ *ø* lexical NML infinitive converb	**-i* ~ *ø* lexical NML - converb adverb	**-i:* ~ *ø* lexical NML	**-i* ~ *ø* lexical NML - converb adverb	**-i* ~ *ø* lexical NML infinitive converb adverb	**-i* ~ *ø* NML
Meireikei imperative	**ø* imperative **a* be-IMP	**ø* imperative **a* be-IMP	**ø* imperative	**ø* imperative	**ø* imperative	**ø* imperative

4.1.1. Closed set

Among the basic inflected forms of Japanese grammar, we find two nonfinite forms (i.e. *rentaikei* (adnominal) and *renyōkei* (converb)), three finite forms (i.e. *shūshikei* (indicative), *izenkei* (subjunctive) and *meireikei* (imperative)) and one pseudostem (i.e. *mizenkei*). Although the pseudostem cannot be treated on a par with the items belonging to the finite and nonfinite categories, it is still possible to regard the five inflected forms as a paradigm because they form an inherently closed set of inflectionally related items from which the entire Japanese inflectional verb paradigm is deducible. Every single cell in the Japanese paradigm can be matched with a materially corresponding form in the same functional domain, but the cognates do not necessarily make up an inherently closed set in the other Transeurasian languages. Nevertheless, the cognates fall into basic inflectional categories such as finite indicative, subjunctive, imperative and converb.

4.1.2. Ordered set

Transeurasian languages to the north and west (Turkic, Mongolic, Tungusic) have recently grammaticalized person-number agreement from subject pronouns, whereas those to the south and east (Manchu, Korean, Japanese) lack person-number agreement on the verb altogether. As a result, the Transeurasian languages cannot be expected to exhibit paradigmatic correlations defined by intersections of the dimensions person and number agreement, as is the case for the Indo-European languages in Table 1. However, some individual cognates display correlations in grammatical patterning of an ordered set of forms, which

are suggestive of multidimensional paradigmaticity. As illustrated in Table 4, the reflexes of the pTEA nominalizer *-*rA* in Korean, Tungusic, Mongolic and Turkic, for instance, suggest correlations in grammatical patterning defined by the intersections of the dimensions telicity of the verb base and finiteness. Following atelic verb bases reflexes of *-*rA* tend to functions as imperfective relativizers and non-past finite forms, whereas they tend to function as perfective relativizers and past finite forms following telic verb bases.

Table 4. Multidimensional paradigmaticity shared by the reflexes of pTEA *-*rA*

	non-finite	finite
atelic verb base	IPF	non-PST
telic verb base	PF	PST

4.1.3. Quirks

Moreover, the basic inflectional paradigm of Japanese has its little quirks: it displays some peculiar traits or idiosyncrasies that are difficult to explain on the basis of internal linguistic analysis alone. Among these oddities, we first find the "exceptional" use of the *rentaikei* (adnominal) as a finite form, while the *shūshikei* (indicative) can be "exceptionally" used in nonfinite function. Second, it is unusual that so many verb suffixes should begin with the vowel /a/. Third, it is peculiar that the endings used for the derivation and inflection of verbal adjectives differ from those used in verb inflection. Fourth, there are semantic oddities, such as the frequency of color terms in the accent class 2.5.

The first idiosyncrasy can be explained in reference to the tendency to grammaticalize non-finite suffixes to finite suffixes, which is one of the driving

forces of morphosyntactic change in the Transeurasian languages (Robbeets 2009, forthcoming b). The second oddity can be explained in light of the replacement of certain verb suffixes by a reinforced periphrastic construction consisting of a copula *_a_- and the same suffix, a recurrent tendency across the Transeurasian languages. The third peculiarity is, among others, related to the observation that some verb inflections originally required a periphrastic copula construction, while verbal adjectives could attach the same suffix without insertion of the copula pJ *_wo_-. Comparison with Korean suggests that in proto-Japano-Koreanic a copula was required when the modified noun was the semantic object of the modifying verb. Since verbal adjectives cannot govern an object, they never inserted a copula, which resulted in the development of formally different endings in Japanese. The fourth oddity seems to go back to a common Transeurasian semantic peculiarity, whereby the lexical nominalizer pTEA *-_mA_ was frequently used in the derivation of colour nouns from descriptive verbs meaning 'to be(come) the colour'.

A final irregularity shared by all Transeurasian languages except Japanese is the tendency illustrated in Table 4, whereby the aspectual and temporal interpretation of a (non)finite suffix is dependent on the telicity of the base verb. This peculiar alternation can be explained by reference to Bybee's (1985: 147) observation:

> Languages do not show one aspect as clearly unmarked and the other marked because for some verbs (in particular, activity verbs and stative verbs), imperfective is the conceptually unmarked member, while for other verbs (in particular, telic or event verbs), perfective is the conceptually unmarked

member.

It can thus be expected that aspectually neutral nominalizers will be reinterpreted as resultatives following telic verb bases. The grammaticalization from a resultative deverbal noun into a perfective adnominal into a past finite marker is well-attested cross-linguistically (Comrie 1976: 99-101; Bybee 1985: 196; Bybee et al. 1994: 86; Johanson 2000; Malchukov 2000: 447).

4.1.4. Extended paradigmaticity

The paradigmatic coherence goes beyond the internal cohesion between the five cells of the paradigm because it also consists in an external correlation of grammatical patterning between the paradigms of lexical nominalizers and those of clausal nominalizers, relativizers and finite forms, or between the paradigms of lexical nominalizers and those of infinitives and converbs. In other words, the patterned match of the five basic inflected forms is supplemented by external connections between separate paradigms, which are caused by shared tendencies of grammaticalization.

In sum, the basic paradigm of Japanese verb forms is an inherently closed set of inflectionally related items with putative cognates in basic inflectional categories of the other Transeurasian languages. The evidence is paradigmatic in that the cognates reflect an internally ordered organization, share certain idiosyncrasies and display external relations of grammatical patterning.

4.2. Are the paradigms copied?

As argued in Section 2.4, the likelihood that the matches between the basic Japanese verb paradigm and the Transeurasian inflections are purely coincidental is very small. A historical explanation is more likely, but could this be borrowing rather than inheritance? Given the relative resistance to borrowing of verb morphology vis-à-vis nominal morphology and of inflectional morphology vis-à-vis derivational morphology and in view of the diagnostic value of shared quirks and extended paradigmaticity, the likelihood that we are dealing with borrowed morphemes is very low to begin with. However, even if the nature of the paradigmatic correlations in Table 3 seems to be indicative of inheritance, this indication cannot be taken to imply that borrowing is completely excluded as an explanation of the similarities. However, the criteria set out in Section 2.3 further decrease the probability of borrowing to an absolute minimum.

First, it can easily be observed that the attachment of shared suffixes is not restricted to shared verb roots only; the common suffixes apply also to unrelatable verb roots.

Second, the majority of compared suffixes are morphologically simplex. The comparisons involving morphologically complex suffixes, such as in the etymologies for the *rentaikei* (adnominal), the *shūshikei* (indicative) and the *izenkei* (subjunctive), deal with shared constructions that are morphologically analyzable in each and every branch. It is not the case that only one member of the etymology reflects a morphologically segmentable suffix string, while the other members reflect unanalyzable cognates.

Third, the polysemy between non-finite and finite function in the etymologies for the *rentaikei* (adnominal), the *shūshikei* (indicative) and the *renyōkei*

(converb) is indicative of a common grammaticalization process. The first two forms originated as deverbal noun suffixes, marking a derivational process at the lexical level, were then extended to function as (ad)nominalizers in dependent clauses at the syntactic level, and eventually grammaticalized still further to mark finite forms in independent clauses. The third form originated as a lexical nominalizer, frequently used for the derivation of action nouns, was then extended to function as an infinitive in verb compounding and grammaticalized still further to mark converbs. Since both the source and target meanings are shared across the different members of the etymology, it is fair to say that shared function is not restricted to secondary, grammaticalized meaning alone. As I previously argued (Robbeets 2013), such instances of "globally" shared grammaticalization, i.e. displaying full correspondence including source function, target function and form, are highly indicative of inheritance.

Fourth, in previous research (e.g. Robbeets 2010, 2012) I have identified up to 21 verb suffixes relating Japanese to the Transeurasian languages, consisting of derivational as well as inflectional markers, finite as well as non-finite forms and various categories such as actionality, diathesis, negation, tense and agreement. Therefore, it appears that there are no observable imbalances across morphosyntactic subsystems.

Finally, most paradigmatic correlations are simultaneously attested in each of the five branches. Within a contact scenario, entire paradigms are supposed to have crossed up to four linguistic boundaries, going, for instance, repeatedly from Turkic into Mongolic, then again from Mongolic into Tungusic until the paradigm has reached Japanese. It is hard to see how these paradigms could have been transferred four times all the way from Turkic to a low-contact

language such as Japanese.

In sum, it is highly unlikely that the paradigmatic evidence summarized in Table 3 can be explained by language contact.

5. Conclusion

In spite of the longstanding debate about the genealogical relationship of Japanese with the Transeurasian languages, linguists seem to agree on at least one point, notably that paradigmatic morphology could substantially help to prove relatedness. Starting from this consensus, I have examined whether Japanese can be said to share paradigmatic morphology with the Transeurasian languages and whether these correlations are significant in such a way that they exclude chance and borrowing as possible explanations.

For this purpose, I have advanced comparative evidence for the five basic inflected forms of Japanese grammar and the so-called *mizenkei* derived stem. It is legitimate to regard the five inflections as a paradigm because they form an inherently closed set of inflectionally related items from which the entire Japanese inflectional verb paradigm is deducible. The comparative evidence is paradigmatic in the sense that the Transeurasian cognates reflect an internally ordered organization, share certain idiosyncrasies and display common extended relations of grammatical patterning.

Weighing non-genealogical explanations, I have argued that the ressemblances are unlikely to be due to chance in view of the number of proposed etymologies, the number of branches in which the inflections have a

match, the relatively small size of the inventory of verb morphemes in an average language, the verification of sound correspondences against regular correspondences in the lexicon, shared polysemy, shared irregularities and internal as well as external paradigmatic cohesion.

Given the relative resistance of inflectional verb morphology to borrowing and in view of the diagnostic value of shared paradigmaticity, the likelihood that the shared paradigms can be accounted for by language contact is relatively low to begin with. Moreover, I have added a number of arguments relating to combinational properties, morphological segmentation, shared grammaticalization, morphosyntactic distribution and linguistic setting that further decrease the probability of borrowing. As such, it is linguistically more reasonable to explain the paradigmatic verb inflections shared between Japanese and the other Transeurasian languages by inheritance than by other factors such as chance or borrowing.

Abbreviations

Linguistic forms

ABL ablative

ACC accusative

ADN adnominalizer

ADV adverbializer

COND conditional

CONV converb

COP	copula
DAT	dative
DEF	definite
EXP	explicit
FIN	finite
GEN	genitive
HON	honorific
INF	infinitive
INFR	inferential
IPF	imperfective
LOC	locative
MOD	modulator
NEG	negation
NML	nominalizer
NOM	nominative
PASS	passive
PERF	perfect
PF	perfective
PL	plural
PST	past
PT	particle
RETR	retrospective
SG	singular
SUB	substantivizer
SUBJ	subjunctive

TOP topic

TENT tentative

VOC vocative

Languages

Dag. Dagur

Evk. Evenki

J Japanese

K Korean

Khal. Khalkha

Ma. Manchu

MK Middle Korean

MMo. Middle Mongolian

Mog. Moghol

OJ Old Japanese

Olč. Olcha

OTk. Old Turkic

pJ proto-Japonic

pK proto-Koreanic

pMo proto-Mongolic

pR proto-Ryukyuan

pTEA proto-Transeurasian

pTg proto-Tungusic

pTk proto-Turkic

Ud. Udehe

WMo. Written Mongolian

Primary sources

HY	1389 Hua-Yi Yiyu
Kok	1449 Welin chenkang ci kok
Kumkang	1464 Kumkang panya phalamil kyeng enhay
MYS	ca. 759 Man'yōshū
NK	720 Nihonshoki kayō
Pak	1517 Pak thongsa enhay
Samkang	1431 Samkang hayngsil-to
Wel	1459 Welin sekpo

References

Bakker, Peter. 1997. *A Language of Our Own: The Genesis of Michif, the Mixed Cree-French Language of the Canadian Métis*. New York: Oxford University Press.

Bauer, Laurie. 1997. Evaluative morphology: A search for universals. *Studies in Language* 21: 533-575.

Beekes, Robert. 1995. *Comparative Indo-European Linguistics. An Introduction*. Amsterdam: John Benjamins.

Benzing, Johannes. 1955. Die tungusischen Sprachen. Versuch einer vergleichenden Grammatik. *Abhandlungen der geistes- und sozialwissenschaftlichen Klasse* 11: 949-1099.

Booij, Geert. 1997. Autonomous morphology and paradigmatic relations. In *Yearbook of Morphology 1996*, Geert Booij & Jaap van Marle (eds), 35-53. Dordrecht: Kluwer.

Bybee, Joan. 1985. *Morphology: A Study of the Relation between Meaning and Form* [Typological Studies in Language 9]. Amsterdam: John Benjamins.

Bybee, Joan, Perkins, Revere & Pagliuca, William. 1994. *The Evolution of Grammar. Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press.

Campbell, Lyle & Poser, William. 2008. *Language Classification. History and Method*. Cambridge: Cambridge University Press.

Comrie, Bernard. 1976. *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press.

Comrie, Bernard. 2008. Inflectional morphology and language contact, with special reference to mixed languages. In *Language Contact and Contact Languages*, Peter Siemund & Noemi Kintana (eds), 15-32. Amsterdam: John Benjamins.

Comrie, Bernard. 2010. The role of verbal morphology in establishing genealogical relations among languages. In *Transeurasian Verbal Morphology in a Comparative Perspective: Genealogy, Contact, Chance* [Turcologica 78], Lars Johanson & Martine Robbeets (eds), 21-31. Wiesbaden: Otto Harrassowitz.

Dressler, Wolfgang. 1989. Prototypical differences between inflection and derivation. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42: 3-10.

Dybo, Anna & Starostin, George. 2008. In defense of the comparative method, or

the end of the Vovin controversy. *Aspects of Comparative Linguistics* 3: 119-258.

Erdal, Marcel. 1991. *Old Turkic Word Formation. A Functional Approach to the Lexicon* [Turcologica 7]. Wiesbaden: Otto Harrassowitz.

Erdal, Marcel. 2004. *A Grammar of Old Turkic*. Leiden: Brill.

Frellesvig, Bjarke. 2010. *A History of the Japanese Language*. Cambridge: Cambridge University Press.

Friedman, Victor. 2009. Turkish presents in Romani dialects. In *Turcological Letters to Bernt Brendemoen*: *Festschrift*, Éva Csató, Gunvald Ims, Joakim Parslow, Finn Thiesen & Emel Türker (eds), 109-121. Oslo: Novus.

Gorelova, Liliya. 2002. *Manchu Grammar*. Leiden: Brill.

Haspelmath, Martin. 2002. *Understanding Morphology*. London: Hodder Education.

Jang, Taeho, Jang, Kyungsook Lim & Payne, Thomas E. (In preparation). *A Grammar of Modern Spoken Xibe*.

Janhunen, Juha. 2012. *Mongolian* [London Oriental and African Language Library 19]. Amsterdam: John Benjamins.

Johanson, Lars. 2000. Viewpoint operators in European languages. In *Tense and Aspect in the Languages of Europe*, Östen Dahl (ed), 27-187. Berlin: Mouton de Gruyter.

Kane, Daniel. 2009. *The Kitan Language and Script*. Leiden: Brill.

Kossmann, Martin. 2010. Parallel system borrowing. Parallel morphological systems due to the borrowing of paradigms. *Diachronica* 27: 459-487.

Lee, Ki-Mun & Ramsey, Robert. 2011. *A History of the Korean Language*. Cambridge: Cambridge University Press.

Malchukov, Andrej. 2000. Perfect, evidentiality and related categories in Tungusic languages. In *Evidentials. Turkic, Iranian and Neighbouring Languages*, Lars Johanson & Bo Utas (eds), 441-469. Berlin: Mouton de Gruyter.

Martin, Samuel Elmo. 1987. *The Japanese Language through Time*. New Haven: Yale University Press.

Martin, Samuel Elmo. 1992. *A Reference Grammar of Korean*. Tokyo: Tuttle.

Martin, Samuel Elmo. 1995. On the prehistory of the Korean grammar: verb forms. *Korean Studies* 19: 139-150.

Martin, Samuel Elmo. 1996. *Consonant Lenition in Korean and the Macro-Altaic Question*. Honolulu: University of Hawaii Press.

Martin, Samuel Elmo. 2002. Coming and going: deictic verbs in Korean and Japanese. In *Pathways into Korean Language and Culture: Essays in Honor of Young-Key Kim-Renaud*, Sang-Oak Lee & Gregory K. Iverson (eds), 373-381. Seoul: Pagijong Press.

Martin, Samuel Elmo. 2006. What do Japanese and Korean have in common? The history of certain grammaticalizations. *Korean Linguistics* 13: 219-234.

Matras, Yaron. 2009. *Language Contact*. Cambridge: Cambridge University Press.

Menges, Karl Heinrich. 1968. *Die Tungusischen Sprachen* [Handbuch der Orientalistik 1. Der Nahe und der Mittlere Osten 5. Altaistik 3. Tungusologie]. Leiden: Brill.

Moravcsik, Edith A. 1978. Universals of language contact. In *Universals of Human Language*, Joseph Greenberg (ed), 93-122. Stanford: Stanford University Press.

Nedjalkov, Igor. 1995. Converbs in Evenki. In *Converbs in Cross-linguistic Perspective. Structure and Meaning of Adverbial Verb Forms – Adverbial*

Participles, Gerunds [Empirical Approaches to Language Typology 13], Martin Haspelmath & Ekkehard König (eds), 97-136. Berlin: Mouton de Gruyter.

Nichols, Johanna. 1996. The comparative method as heuristic. In *The Comparative Method Reviewed: Regularity and Irregularity in Language Change*, Mark Durie & Malcolm Ross (eds), 39-71. Oxford: Oxford University Press.

Nikolaeva, Irina Alekseevna. 1999. A Grammar of Udehe. PhD dissertation, Leiden University.

Ōno, Susumu. 1953. Nihongo no dōshi no katsuyōkei no kigen ni tsuite. *Kokugo to Kokubungaku* 350: 47-56.

Pakendorf, Brigitte. 2009. Intensive contact and the copying of paradigms: an Even dialect in contact with Sakha (Yakut). *Journal of Language Contact* 2: 85-110.

Pellard, Thomas. 2011. Ryukyuan perspectives on the Proto-Japonic vowel system. *Japanese/Korean Linguistics* 20: 1-15.

Plank, Frans. 1994. Inflection and derivation. In *The Encyclopedia of Language and Linguistics. Vol. 3*, R. E. Asher (ed), 1671-1678. Oxford: Pergamon Press.

Polivanov, Evgenij Dmitrvič. 1924. K rabote o muzykal'noj akcentuacii v japonskom jazyke (v svjazi s malajskim). *Bjulleten' 1-go Sredne-Aziatskogo Gosudarstvennogo Universiteta* 4: 101-108.

Poppe, Nicholas. 1954. *Grammar of Written Mongolian*. Wiesbaden: Otto Harrassowitz.

Poppe, Nicholas. 1955. *Introduction to Mongolian Comparative Studies* [Mémoires de la Société Finno-Ougrienne 110]. Helsinki: Suomalais-Ugrilainen Seura.

Robbeets, Martine. 2005. *Is Japanese related to Korean, Tungusic, Mongolic and*

Turkic? [Turcologica 64]. Wiesbaden: Otto Harrassowitz.

Robbeets, Martine. 2007. The causative-passive in the Trans-Eurasian languages. *Turkic Languages* 11(2): 235-278.

Robbeets, Martine. 2009. Insubordination in Altaic. *Journal of Philology* 31. *Ural-Altaic Studies* 1: 61-79.

Robbeets, Martine. 2010. Transeurasian: Can verbal morphology end the controversy? In *Transeurasian Verbal Morphology in a Comparative Perspective: Genealogy, Contact, Chance* [Turcologica 78], Lars Johanson & Martine Robbeets (eds), 81-114. Wiesbaden: Otto Harrassowitz.

Robbeets, Martine. 2012. Shared verb morphology in the Transeurasian languages: copy or cognate? In *Copies versus Cognates in Bound Morphology* [Brill's Studies in Language, Cognition and Culture 3], Lars Johanson & Martine Robbeets (eds), 427-446. Leiden: Brill.

Robbeets, Martine. 2013. Genealogically motivated grammaticalization. In *Shared Grammaticalization with Special Focus on the Transeurasian Languages* [Studies in Language Companion Series 132], Martine Robbeets & Hubert Cuyckens (eds), 147-175. Amsterdam: John Benjamins.

Robbeets, Martine. forthcoming a. Areality in the Transeurasian languages. In *The Cambridge Handbook of Areal Linguistics* [Cambridge Handbooks in Language and Linguistics], Raymond Hickey (ed). Cambridge: Cambridge University Press.

Robbeets, Martine. forthcoming b. The development of finiteness in the Transeurasian languages. *Linguistics Special Issue*: *Clausal Complementation Strategies*.

Robbeets, Martine. forthcoming c. *Diachrony of Verb Morphology in Japanese and the Transeurasian Languages* [Typological Studies in Language]. Berlin: Mouton-De Gruyter.

Sárközi, Alice. 2004. *Classical Mongolian*. München: Lincom Europa.

Seifart, Frank. 2012. The principle of morphosyntactic subsystem integrity in language contact. Evidence from morphological borrowing in Resígaro (Arawakan). *Diachronica* 29: 471–504.

Sekerina, Irina A. 1994. Copper Island (Mednyj) Aleut (CIA): a mixed language. *Languages of the World* 8: 14–31.

Spencer, Andrew. 1991. *Morphological Theory*: *An Introduction to Word Structure in Generative Grammar*. Oxford: Blackwell.

Starostin, Sergej, Dybo, Anna & Mudrak, Oleg. 2003. *Etymological Dictionary of the Altaic Languages*. Leiden: Brill.

Stump, Gregory T. 2001. *Inflectional Morphology: A Theory of Paradigm Structure* [Cambridge Studies in Linguistics 93]. Cambridge: Cambridge University Press.

Takeuchi, Lone. 1999. *The Structure and History of Japanese*. New York: Longman.

Thomason, Sarah Grey. 1997. Mednyi Aleut. In *Contact Languages. A Wider Perspective*, Sarah Grey Thomason (ed), 449-468. Amsterdam: John Benjamins.

Thomason, Sarah Grey. 2001. *Language Contact: An Introduction*. Washington: Georgetown University.

Thomason, Sarah Grey & Kaufman, Terrence. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.

Thorpe, Maner Lawton. 1983. Ryukyuan Language History. PhD dissertation,

University of Southern California.

Unger, Jim Marshall. 2000. Reconciling comparative and internal reconstruction: the case of Old Japanese /*ti, ri, ni*/. *Language* 76: 655-681.

Vovin, Alexander. 2003. *A Reference Grammar of Classical Japanese Prose*. London: Routledge-Curzon.

Vovin, Alexander. 2005. The end of the Altaic controversy. *Central Asiatic Journal* 49: 71-132.

Vovin, Alexander. 2008. Proto-Japanese beyond the accent system. In *Proto-Japanese. Issues and Prospects* [Current Issues in Linguistic Theory 294], Bjarke Frellesvig & John Whitman (eds), 141-156. Amsterdam: John Benjamins.

Vovin, Alexander. 2009. *A Descriptive and Comparative Grammar of Western Old Japanese. Part 2: Adjectives, Verbs, Adverbs, Conjunctions, Particles, Postpositions* [Languages of Asia 8]. Folkestone: Global Oriental.

Weiers, Michael. 1966. Untersuchungen zu einer historischen Grammatik des präklassischen Schriftmongolisch. PhD dissertation, Rheinischen Friedrich-Wilhelms-Universität Bonn.

Weinrich, Uriel. 1953. *Languages in Contact. Findings and Problems*. New York: Publications of the Linguistic Circle of New York.

Whitman, John Bradford. 1985. The Phonological Basis for the Comparison of Japanese and Korean. PhD dissertation, Harvard University.

Wilkins, David P. 1996. Morphology. In *Contact Linguistics. An International Handbook of Contemporary Research*, Hans Goebl, Peter H. Nelde, Zdeněk Starý & Wolfgang Wölck (eds), 109-117. Berlin: Walter de Gruyter.

Wrona, Janick. 2008. The nominal and adnominal forms in Old Japanese:

Consequences for a reconstruction of pre-Old Japanese syntax. In *Proto-Japanese. Issues and Prospects* [Current Issues in Linguistic Theory 294], Bjarke Frellesvig & John Whitman (eds), 193-215. Amsterdam: John Benjamins

Yanagida, Yuko & Whitman, John. 2009. Alignment and word order in Old Japanese. *Journal of East Asian Linguistics* 18: 101-144.

Ylikoski, Jussi. 2003. Defining non-finites: Action nominals, converbs and infinitives. *SKY Journal of linguistics* 16: 185-237.