

A comparative approach of verbal morphology in Transeurasian

Martine Robbeets

Abstract

This chapter starts from a point of consensus in the Transeurasian debate, namely that common paradigmatic morphology could substantially help unravel the question. It aims at giving an overview of the verbal morphology shared between the Transeurasian languages. In addition to regular correspondences in form and function, it argues that the Transeurasian verb morphology displays a certain degree of paradigmaticity, based on relationships of grammatical patterning among different morphosyntactic subsystems. Moreover, shared irregularities such as peculiar allomorphies and functional idiosyncrasies are taken as a strong indication of affiliation of the languages concerned. The chapter concludes with an assessment of chance similarity, leading to the viewpoint that it is more sensible to attribute the correlations in the verbal morphology to inheritance than to take refuge in non-genealogical explanations.

Keywords: Transeurasian, verbal morphology, paradigm, paradigmaticity, allomorphy, shared idiosyncrasy, inheritance

30.1 Introduction

Supporters and critics of Transeurasian relatedness seem to agree on at least this one point, namely that patterned morphology could substantially help settle the controversy. Vovin (2005c: 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”, whereas Dybo and Starostin (2008: 125) agree that “regular paradigmatic correspondences in morphology are necessarily indicative of genetic relationship”. Since consensus between supporters and critics is a rare commodity in the Transeurasian debate, I will devote this chapter to an overview of the verb morphology shared between the Transeurasian languages and indicate to what extent the correlations may be regarded as “paradigmatic”.

Since the beginnings of the historical comparative study of the Transeurasian languages, the emphasis has always been on lexical research. Although the field of contemporary Transeurasian linguistics was practically founded with the posthumous publication of Ramstedt’s “*Formenlehre*” in 1952, few linguists followed in his footsteps, preferring to turn their attention rather to the comparison of sounds, words or typological structure. The publication of the first part “*Vergleichende Lautlehre*” of Poppe’s (1960b) “*Vergleichende Grammatik der altaischen Sprachen*” was intended to precede a second part on comparative morphology, but unfortunately, the comparative phonology was the only volume to appear. Some exceptions to the underrepresentation of comparative verbal morphology include Poppe (1972), Nasilov (1978) and Kormušin (1984), but similar to Ramstedt’s “*Formenlehre*”, none of these contributions took Japanese data into account.

Only few comparative studies of Transeurasian verbal morphology include Japanese as well as Korean comparanda. Miller compared deverbal verb suffixes (1981), denominal verb suffixes (1982), negation (1971: 245–284, 1985) and gerunds (1971: 285–291). Menges treats transitivity pairs (1975: 32–35), negation (1975: 96–109, 1984: 262–263) and gerunds (1975: 110–111). Street dealt with denominal verbs (1978: 44, 204, 219), actionality (1978: 74–75, 113, 115, 188, 178–179, 219–220, 247–249, 230–231), diatheses (1978: 53, 55, 71, 75, 179, 181–182, 185–188, 199–

200, 208, 230, 239–242, 257–259) and converbs (1978: 200, 251–252). Finch (1987) reconstructed Altaic verb classes. Vovin (1998a) provided a sketch of comparative Altaic verb morphology including Japanese, but in (2001) he restricted the same evidence to a comparison of Japanese, Korean and the Tungusic languages and later in (2005) he distanced himself from his former positions by completely rejecting all evidence in support of Transeurasian affiliation. “*The Etymological Dictionary of the Altaic Languages*”—including Japanese and Korean—published by Starostin et al. (2003: 173–229) provided an overview of morphological elements that look similar on the surface, but being a dictionary rather than a comparative grammar, it did not investigate these look-alikes in more detail.

In this chapter, I will present correlations in the verbal morphology between the Transeurasian languages, summarizing my findings on this topic as discussed in more detail in Robbeets (2015). In Section 30.2, I introduce the dataset, providing a tabular overview of common verbal morphology. I intend to reserve Section 30.3 for a discussion of the formal correlations while I will focus on functional correlations and indications of paradigmaticity in Sections 30.4 and 30.5, respectively. In Section 30.6, I will assess the likelihood that the correlations in the verb morphology are purely coincidental. Finally, in Section 30.7, I will conclude this chapter.

30.2 Overview of the shared verbal morphology of Transeurasian

In Robbeets (2015), I undertook a detailed comparative study of elements of verb morphology shared between the Transeurasian languages. In Table 30.1, I present an overview of the different etymologies for auxiliaries and suffixes. The final column of this table provides numbered sound correspondences in line with the consonant and vowel correspondences in Table 40.2 and 40.3. Due to the limited space available, I

restrict myself to a small number of examples in this chapter, but I refer to Robbeets (2015) for a detailed description of the underlying forms.

Table 34.1 Overview of verb morphology shared between the Transeurasian languages

	PTEA	PJ	PK	PTg	PMo	PTk	Sound Corr. No
01	*ana- negation	*ana- negation	*an- negation	*ana- negation		[*an-] negation	32, 28, 32
02	*ə- negation			*e- negation	*e-se- negation	*e- negation	34
03	*-lA- manipulative	*-ra- manipulative		*-lA:- manipulative	*-lA- manipulative	*-lA- manipulative	31, 32/34
04	*-nA- processive	*-na- processive	*-nO- processive	*-nA- processive	*-nA- processive	*-(X)n- processive	28, 32
05	*(-)ki- 'do, make' iconic	*-ka- iconic	*-ki- iconic	*-ki- iconic	*(-)ki- 'do, make' iconic	*ki(-)l- /-kI- 'do make' iconic	14, 40
06	*-mA- inclination	*-ma- inclination	*-mO- inclination	*-mA- inclination	*-mA- inclination		26, 32/34
07	*-gA- inchoative	*-ka- inchoative	*-k(O)- inchoative	*-gA- inchoative	*-gA- inchoative	*-(X)k-~- (X)g- inchoative	15 32/34
08	*-ti- causative	*-ta- causative passive	*-ti- causative passive	*-ti- causative passive	*-ti- causative	*-tI- causative passive	8, 40
09	*-pU- reflexive anticausative	*-pa- reflexive anticausative	*-pO- anticausative	*-p- reflexive anticausative	*-βU- reflexive anticausative	*-U- reflexive anticausative	2, 38/39
10	*-dA- fientive	*-ya- fientive passive		*dA:- fientive	*-dA- fientive passive	*-(A)d- fientive anticausative	10, 32/34
11	*-rA- anticausative	*-ra- anticausative	*-(u)l- anticausative	*-rA- anticausative	*-rA- anticausative	*-rA- anticausative	29, 32/34
12	*-gi- creative causative	*-(k)i- creative causative anticausative	*-ki- creative causative passive	-gi:- creative causative			15, 40

13a	*-rA lexical NMLZ	*-ra lexical NMLZ	*-l lexical NMLZ	*-rA lexical NMLZ	*-r lexical NMLZ	*-rV lexical NMLZ	29, 32/34
13b		*-wo-ra clausal NMLZ	*-wo-l clausal NMLZ	*-rA clausal NMLZ	*-r clausal NMLZ		29, 32/34
13c		*-wo-ra relativizer	*-wo-l relativizer	*-rA relativizer		*-rV relativizer	29, 32/34
13e		*-wo-ra finite	*-wo-l finite	*-rA finite	*-r finite	*-rV finite	29, 32/34
14a	*-mA lexical NMLZ	*-m lexical NMLZ	*-m lexical NMLZ	*-mA lexical NMLZ	*-mA ~ *-m lexical NMLZ	*-mA ~ *-m lexical NMLZ	26, 32/34
14b		*-wo-m clausal NMLZ	*-wo-m clausal NMLZ	*-mA clausal NMLZ	*-mA ~ *-m clausal NMLZ		26, 32/34
14c		*-wo-m finite	*-wo-m finite	*-mA finite	*-mA ~ *-m finite		26, 32/34
15a	*-n lexical NMLZ	*-n lexical NMLZ	*-n lexical NMLZ	*-nA ~ *-n lexical NMLZ	*-n lexical NMLZ	*-n lexical NMLZ	28
15b		*-wo-n clausal NMLZ	*-wo-n clausal NMLZ	*-nA ~ *-n clausal NMLZ	*-n clausal NMLZ	*-n clausal NMLZ	28
15c		*-wo-n relativizer	*-wo-n relativizer			*-n relativizer	28
15d		*-wo-n finite	*-wo-n finite	*-nA ~ *-n finite	*-n finite	*-n finite	28
16a	*-xA ~ *-kA resultative lexical NMLZ	*-ka resultative lexical NMLZ	*-ka(-)i resultative lexical NMLZ	*-xA: ~ *-kA: resultative lexical NMLZ	*-xA ~ *-kA resultative lexical NMLZ	*-xA ~ *-kA resultative lexical NMLZ	22 32/34
16b				*-xA: ~ *-kA: clausal NMLZ	*-xA ~ *-kA clausal NMLZ		22 32/34
16c		*-ka relativizer		*-xA: ~ *-kA: relativizer	*-xA ~ *-kA relativizer	*-xA ~ *-kA PFV.FUT relativizer	22 32/34
16d		*-ka finite		*-xA: ~ *-kA: past finite	*-xA ~ *-kA past finite	*-xA ~ *-kA future finite	22 32/34
17	*-sA resultative lexical NMLZ	*-sa resultative lexical NMLZ		*-sA ~ *-si: < *sA-i: resultative lexical NMLZ	*-sA ~ *-si: < *sA-i: resultative lexical NMLZ	*-sA	24 32/34

				*-sA ~ *-si: clausal NMLZ	*-sA ~ *-si: clausal NMLZ	*-sA perfective clausal NMLZ	24 32/34
				*-sA ~ *-si: relativizer			24 32/34
		*-sa finite		*-sA ~ *-si: finite	*-sA ~ *-si: finite	*-sA past finite	24 32/34
18	*-i ~ ∅ nominalizer	*-i ~ ∅ nominalizer infinitive converb	*-i ~ ∅ nominalizer converb adverb	*-i: ~ ∅ nominalizer	*-i ~ ∅ nominalizer converb adverb	*-I ~ ∅ nominalizer infinitive converb adverb	40
19	*-xU ~ *-kU nominalizer infinitive	*-ku nominalizer infinitive converb adverb	*-k(Λ) / *-ku nominalizer infinitive converb adverb	*-xu: ~ -ku: nominalizer converb adverb	*-xU ~ -kU nominalizer infinitive converb	*-xU ~ -kU nominalizer infinitive	22 38/39
20	*∅ imperative	*∅ imperative	*∅ imperative	*∅ imperative	*∅ imperative	*∅ imperative	

The reconstruction of the individual morphemes for Proto-Turkic, Proto-Mongolic, Proto-Tungusic, Proto-Koreanic and Proto-Japanic relies on material from their earliest unambiguously written stages—Old Japanese, Late Middle Korean, Middle Mongolian, Written Mongolian and Old Turkic—and is supplemented by the most relevant contemporary varieties. A contemporary variety is considered relatively relevant for reconstruction purposes if it derives from a major node in the family tree for which historical records are either lacking or do not provide satisfactory morphological information. The reconstruction of Proto-Turkic morphemes is based on the main sources of evidence of the earliest split into the western Oghuric branch and the eastern Common Turkic branch, namely Chuvash and Old Turkic.

The reconstruction of Proto-Mongolic morphemes is mainly based on Middle Mongolian and Written Mongolian. When supplementary information is needed, however, it is retrieved from contemporary varieties. Moreover, I added evidence of verbal morphology in Khitan texts by Kane (2009: 144–158), where available.

For the historical study of Tungusic languages, records are unfortunately rather scarce for most of the languages. The oldest historical records are restricted to the Manchu branch of Tungusic, written in Jurchen, the official language of the Jin Dynasty (1115–1234), and in its immediate descendant, Manchu, the official language of the Qing Dynasty (1644–1911). Due to the fragmentary attestation of Jurchen, I retrieve morphological information about Tungusic mainly from Manchu as representative of the Manchuric branch, from Even and Evenki, as representatives of the eastern and western subbranches of Northern Tungusic, and from Udehe and Nanai as representatives of the eastern and western subbranches of Southern Tungusic, respectively.

The reconstruction of Proto-Koreanic morphemes is based on Middle Korean and contemporary Korean. As happened with Mongolic under Chingis Khan, the Silla unification erased all coexisting varieties of Koreanic. In this study, the label “Middle Korean” mainly refers to Late Middle Korean, the language written down after the invention of the Korean script in 1446.

The reconstruction of Proto-Japonic morphemes is based on representatives of its two branches, i.e. Mainland Japanese and Ryukyuan. The better attested earliest source of Mainland Japanese is (Western) Old Japanese, representing the language spoken in Central Japan in the Nara period (710–794). As far as the Ryukyuan languages are concerned, evidence comes from Northern Ryukyuan languages such as Shodon or Shuri or from Southern Ryukyuan languages such as Hirara or Hateruma.

30.3 Correlations in form

30.3.1 Regular sound correspondences

The consonants and vowels involved in the comparisons correspond regularly according to the correspondences given in Table 40.2 and 40.3, respectively. The vowel harmony archiphonemes in Table 30.1 represent the following alternations: PK $*O = *i \sim \text{ɶ}$; PTg $*A = *a \sim e$ [RTR $*a \sim \text{ə}$]; PMo $*A = *a \sim e$ [RTR $*a \sim \text{ə}$]; PMo $*U = *ü \sim u$ [RTR $*u \sim \text{ʊ}$]; PTK $*A = *a \sim e$, PTK $*U = ü \sim u$, PTK $*I = i \sim \text{ɪ}$ and PTK $*X = i \sim \text{ɪ} \sim u \sim \text{ü}$. The reinterpretation of the quality of the Koreanic, Tungusic and Mongolic vowels in the light of the Retracted Tongue Root (RTR) interpretation by Joseph et al. (this volume: Chapter 29) leads to the reconstruction of RTR harmony in the original Transeurasian suffixes with archiphonemes PTEA $*A = *a \sim \text{ə}$ and PTEA $*U = *u \sim \text{ʊ}$. It is not unlikely that Turkic shifted to a front-back sound harmony system and that the RTR system became distorted in Japonic, due to areal influences at the periphery of the Transeurasian family. However, a number of Japonic verb suffixes in Table 30.1 have left traces of a vowel harmonic-like alternation, notably PJ $*-ra-$ \sim $-r\text{ə}-$ denominal verbalizer, PJ $*-na-$ \sim $-n\text{ə}-$ processive, PJ $*-ma-$ \sim $-m\text{ə}-$ inclination, PJ $*-pa-$ \sim $-p\text{ə}-$ reflexive-anticausative and PJ $*-ra$ \sim $-r\text{ə}$ nominalizer. This opposition may imply an original RTR based contrast.¹

According to the vowel correspondences, the vowel harmony between PTEA $*-a-$ and $*-\text{ə}-$ regularly merges into PJ $*-a-$, while the harmony between PTEA $*-u-$ and $-\text{ʊ}-$ regularly merges into PJ $*-u-$. This merger and the widespread $*A$ -vocalism in the Proto-Transeurasian suffixes in general may explain why the large majority of Japanese suffixes display a -vocalism. The irregular low central vowel in the Japonic iconic suffix $*-ka-$ and causative-anticausative suffix $*-ta-$ may be the result of analogy, seeking resonance with the other suffixes in the verbal paradigm.

The consistently reduced vowel reflexes ($*-i-$ / $-\text{ɶ}-$) in the Korean suffixes are probably due to vowel reduction in unstressed position, reminiscent of the frequent

weakening of the final vowel involved in lexical items. Morphological items, especially suffixes, tend to be unstressed and are therefore liable to undergo irregular reductive developments such as the loss of word-final vowels in some reflexes of the lexical nominalizers PTEA **-rA* and **-mA*. Vowel erosion in word-final suffixes is to be expected especially following sonorants such as /m/ and /r/, because their high sonority allows them to be articulated without final vowel. This may explain why the vowels are retained in the reflexes of other word-final suffixes such as PTEA **-sA* and **-xA ~ *-kA*.

34.3.2 Shared allomorphy

Two velar suffixes, PTEA **-xA ~ *-kA* for the resultative nominalizer in (16) and, PTEA **-xU ~ *-kU* for the nominalizer/infinite in (19) reveal the original conditioning factor behind the velar fricative correspondence 22 in Table 40.2. They display a peculiar allomorphy in a phonological environment that often involves continuants. Since the velar involved becomes a voiceless obstruent in this environment, I suggested that it arose through de-fricativation (Robbeets 2015: 402, 406, 416, 483). The allomorphy is illustrated in example (1) for the resultative nominalizer, whereby Japonic and Koreanic merged the original allomorphy in a single voiceless velar suffix **-kA*, while the Altaic languages preserved traces of the original allomorphy in continuant environment.

(1) PTEA **-xA* resultative nominalizer → **-kA*/r_ or n_ or β_

a. Proto-Japonic **-kA* resultative nominalizer

PJ **ata-* ~ *atu-* ‘to be warm’ in place name Atami (< **ata-umi* ‘warm sea’), OJ *atu-* B ‘to be warm’ → OJ *atataka* (4.11) ‘warm’

- b. Proto-Koreanic *-*kA-i* resultative nominalizer²

K *cci-* ‘to steam (tr.)’ → *ccikey* ‘thin stew’

- c. Proto-Tungusic *-*xA*: resultative nominalizer → *-*kA*:/ n_ or β_

Ma. *ji-* ‘to come’ (<**di-*) → *ji-he mangi* (come-PFV after) ‘after [he] had come’

Ma. *dosi-* ‘to enter’ (<**dosin-*) → *dosi-ka mangi* (come-PFV after) ‘after [he] had entered’

- d. Proto-Mongolic *-*xA* resultative nominalizer → *-*kA*/ r_ or β_

WMo. *ide-* ‘to eat, consume (tr.)’ → *idege* ‘food’

WMo. *toyuri-* ‘to go about, circle, surround (tr.)’ → *toyurqa* ‘the encirclement of the tent’

- e. Proto-Turkic *-*xA* resultative nominalizer → *-*kA*/ r_ or β_

OTk. *kīs-* ‘to pinch, squeeze, reduce (tr.)’ → *kīsga* ‘short’

OTk. *öp-* ‘to kiss, sip or suck in the air or a liquid’ (< **öβ-*) → *öpke* ‘generated in the lung; lung, anger’

The standard Manchu perfective adnominalizer *-ha* ~ *-he* ~ *-ho* has an allomorph *-ka* ~ *-ke* ~ *-ko* for 185 verbs, which can be traced back to earlier verbs in stem-final nasal or *-b-* (<intervocalic *-β). The verb Ma. *dosi-* ‘to enter’ in example (1c) can be

derived from an earlier **dosin-*, considering the preservation of the original stem-final nasal in the causative Ma. *dosim-bu-*. The Written Mongolian resultative nominalizer *-ya ~ -ge* has a voiceless allomorph WMo. *-qa ~ -ke* following *-r-* and *-b-* (<intervocalic **-β-*), which is reflected in a few derivational pairs.³ Similarly, there are some instances in which the suffix OTk. *-gA* seems to devoice to *-kA* following *-r-* and *-p-* (< **β*). The continuants in Tk. *öfke* ‘anger’, *öyken* ‘lung’, Az. *öxbä* ‘lung’, Gag. *üfke*, Tkm. *öyken*, Sal. *öhhen*, cognate to OTk. *öpke* ‘generated in the lung; lung, anger’ in example (1d) may point to a bilabial fricative in pre-OTk. **öβke*.

From a phonological point of view, it is difficult to motivate the devoicing of a voiced velar stop in a sonorant environment. However, de-fricativization is commonly seen in this environment, for instance in Spanish and Icelandic (Johanson 1979a: 30). The juxtaposition of two continuants, which is articulatorily inconvenient, will be avoided by dropping the incomplete oral closure from the second consonant: the fricative will thus become a stop. Therefore, it is linguistically sensible to derive the voiced velar stop from an earlier fricative in Mongolic and Turkic. Sharing such a peculiar allomorphy in a comparable phonological environment is a strong argument for the genealogical relatedness of the languages concerned.

30.4 Correlations in function

The functional correspondences go beyond general tags such as denominal verbalizer or deverbal noun suffix since they also hold for the specific type of verbalization or nominalization concerned, i.e. manipulative, iconic, processive, inclinational and inchoative verbalization or instrumental, resultative and action/state nominalization. Moreover, there are correlations between the concrete submeanings that make up these functional specializations. The manipulative meaning of the denominal

verbalizer in example (2), for instance, falls apart in several submeanings, such as ‘to achieve, overcome or execute a difficult action with success on the base noun’, ‘to make use of the base noun’ and ‘to sound or feel like the base onomatopoea’, all of which correlate across the individual branches. In addition, there are a number of derivational pairs across the different daughter languages that derive strikingly parallel meaning from—although not formally—semantically equivalent bases. Derivational bases meaning ‘neck, throat’, for instance, are regularly derived as ‘kill by aiming at the neck/throat’, e.g. OJ *kubiir*- A ‘to strangle (tr.)’, OTk. *boguzla*- ‘to cut (somebody’s/ an animal’s) throat (tr.)’, whereas bases meaning ‘hand’ are derived as ‘take by using the hands’, e.g. OJ *to1r*- ~ *to2r*- B ‘to take, hold in hand’, Ud. *cob’olo*- ‘to scoop with one’s hand’, OTk. *adutla*- ‘to scoop up with the palm of one’s hand (tr.)’. Moreover, bases meaning ‘house’ or ‘shelter’ are regularly derived as ‘found/make a house’, e.g. *angala*- ‘to make a night shelter’, WMo. *gerle*- ‘to marry, found a house of his own (intr.)’, OTk. *evle*- ‘to furnish (sb.) with dwellings, marry (sb.) off (tr.)’ and those meaning ‘word’ are derived as ‘speak’, e.g. OJ *katar*- ‘to tell’, WMo. *kelele*- ‘to utter words, say, narrate’, OTk. *sözle*- ‘to speak, say, talk with somebody (tr.)’.

(2) Proto-Transeurasian *-rA manipulative denominal verb suffix

a. Proto-Japonic *-rA manipulative denominal verb suffix

1 ‘to achieve, overcome or execute a difficult action with success on the base noun’

OJ *kiipa* 2.3 ‘limit, brink’ → OJ *kiipar*- B ‘to come to an end, wear out (intr.)’⁴

OJ *kubi* 2.1 ‘neck’ → OJ *kubiir*- A ‘to strangle (tr.)’

2 'to make use of the base noun'

OJ *ipo* 2.3 'hut' → OJ *ipor-* B 'to lodge in a hut'

OJ *ko₂to₂* 2.3 'word, speech, statement' → OJ *katar-* A 'to tell'⁵

OJ *te₂ ~ ta-* 1.3 'hand' → OJ *to₁r- ~ to₂r-* B 'to take, hold in hand'⁶

3 'to sound or feel like the base onomatopoea'

PJ **pika* in J *pika-pika* 'sparkling, glittering, shining (mimetic)' → OJ

pi₁kar- 'to shine'

b. Proto-Tungusic *-*lA-* manipulative denominal verb suffix

1 'to achieve, overcome or execute a difficult action with success on the base noun'

Ud. *anga* 'night shelter' → *angala-* 'to make a night shelter'

Ud. *mamasa* 'wife' → *mamasala-* 'to marry, take a wife'

2 'to make use of the base noun'

Ud. *cob'o* 'handful' → *cob'olo-* 'to scoop with one's hand'

Ud. *kusige* 'knife' → *kusigele-* 'to stab with a knife'

c. Proto-Mongolic *-*lA-* manipulative denominal verb suffix

1 'to achieve, overcome or execute a difficult action with success on the base noun'

WMo. *ger* 'yurt, house' → *gerle-* 'to marry, found a house of his own (intr.)'

WMo. *šibaYun* 'bird' → *šibaYula-* 'to hunt (birds) (tr.)'

2 ‘to make use of the base noun’

MMo. *kelen* ‘tongue, word, speech, statement’ → *kelele-* ‘to utter words, say, narrate (tr. / intr.)’

WMo. *usun* ‘water’ → WMo. *usula-* ‘to water (animals/plants), irrigate (tr.)’

3 ‘to sound or feel like the base onomatopoea’

WMo. *YoYuu* ‘cry of a rooster’ → *YoYuYla-* ‘to cackle’

d. Proto-Turkic *-*la-* manipulative denominal verb suffix

1 ‘to achieve, overcome or execute a difficult action with success on the base noun’

OTk. *boguz* ‘throat’ → *boguzla-* ‘to cut (somebody’s /an animal’s) throat (tr.)’

OTk. *ev* ‘house’ → *evle-* ‘to furnish (sb.) with dwellings, marry (sb.) off (tr.)’

2 ‘to make use of the base noun’

OTk. *adut* ‘palm of one’s hand’ → *adutla-* ‘to scoop up with the palm of one’s hand (tr.)’

OTk. *söz* ‘word, speech, statement’ → *sözle-* ‘to speak, say, talk with somebody (tr.)’

3 ‘to sound or feel like the base onomatopoea’

OTk. *ori:* ‘shout, outcry’ → *orıla-* ‘to shout’

30.5 Paradigmaticity

Whereas lexical comparison has two dimensions, i.e. form and meaning, morphological comparison has a third dimension, notably the patterning of functional oppositions in a paradigm. In a continuum approach of morphology, a paradigm can be defined as “an organized set of derivationally or inflectionally related items that derive a particular semantic or morphosyntactic category from a common base or root” (Robbeets and Bisang 2014). As such, it refers to the full set of forms, inflectional and derivational, that a root enters into. Such a set is not necessarily 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. However, in historical comparison, correlations between small, closed sets of forms are considered to be more diagnostic than those between large, open sets. The Indo-European verb paradigms, for instance, count a strong evidence for relatedness as they form small sets of up to 7 corresponding form slots with positions defined by intersections of the dimensions person and number agreement.

Joseph (2014) introduced the term “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 in Indo-European. It is in this sense that the correlations in the Transeurasian verb morphology can be regarded as paradigmatic: they consist in systematically shared grammaticalization patterns between morphosyntactic subsystems.

Similar to the Indo-European case, the Transeurasian languages to the north and west (Turkic, Mongolic and most Tungusic languages) have grammaticalized person-

number agreement from personal pronouns, but the development being much more recent than in Indo-European, took place in each subfamily independently. The Transeurasian languages to the south and east (Manchu, Korean, Japanese) lack person-number agreement on the verb altogether. As a result of this observation and due to the agglutinative nature of Transeurasian in contrast with the fusional character of Indo-European, the Transeurasian languages did not develop a closed set of correlating person markers on the verb. Nevertheless, they systematically share other grammaticalization patterns, which can be regarded as paradigmatic.

The etymologies 13 to 17 in Table 30.1 suggest systematic correlations in grammatical patterning between non-finite and finite verb paradigms. There is a recurrent tendency to reanalyze non-finite suffixes as finite ones without the omission of a specific matrix predicate, a tendency which I called “direct insubordination” (Robbeets 2015, 2016a, 2017e). The suffixes involved 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—through a pragmatic role in discourse—were extended still further to mark finite forms in independent clauses.

For instance, deverbal noun suffixes such as Hirara *-i* in Hirara *kak-* ‘to write’ → *kak-i* ‘writing’⁷, MK *-(·u/o)l* in MK *hhoyng ho-* ‘to travel’ → *hhoyng hol* ‘traveler’, Ma. *-rA* in *mute-* ‘to be able’ → *mutere* ‘ability’, WMo. *-(u/ü)r* in WMo. *belčige-* ‘to pasture, graze (tr.)’ → *belčiger* ‘pasture, grazing grounds, grass on a pasture’ and OTk. *-(A)r* in OTk. *tug-* ‘to be born, to rise (of sun) (intr.)’ → *tugar* ‘sunrise, east’ developed over intermediate stages of clausal nominalizers and relativizers into finite suffixes, as illustrated in example (3).

(3) a. Hirara

tabaku: fuk-ï padimi-ï

tobacco smoke-NMLZ begin-FIN

‘[I] began smoking’ (Russell 2006: 577)

b. Middle Korean

·qilqsim-·u·lwo kwoyGwoy ho-·la⁸

wholehearted-ADV silence do-SBJV

‘Be utterly quiet!’ (Kumkang 12a; Martin 1992: 851)

c. Manchu

uthai sin-de bu-re.

at.once you-DAT give-FIN

‘I shall give [it] to you straight away.’ (Gorelova 2002: 256)

d. Khitan⁹

puu giuuy shi po-or

fu gong shi become-PST.FIN

‘He was appointed a fu gong shi’ (Kane 2009: 146)

e. Old Turkic

Ölüm-tä oz-upan ögir-ä savin-ü yori-r.

death-ABL escape-CVB rejoice-CVB be.happy-CVB go.on-FIN

‘Having been saved from death it happily goes on with its life.’ (Erdal 2004: 325)

This systematic correlation in grammatical patterning between non-finite and finite verb paradigms is intertwined with yet another grammaticalization process, i.e. the development of tense from aspect distinctions. The non-finite forms in the etymologies (13) to (15) generally display aspectually unmarked meaning, while those in the etymologies (1) and (17) have resultative meaning. The

grammaticalization from deverbal noun suffix to adnominalizer to finite suffix involves a change in the part-of-speech status from noun to adjective to verb. In this process of verbalization, an actional interpretation is forced on an originally stative nominal form. This leads to the development of imperfective aspect on adnominalizers and non-past tense distinctions on finite suffixes in the etymologies (13) to (15), while perfective aspect develops on adnominalizers and past tense on finite suffixes in the etymologies (1) and (17). The diachronic evolution of tense from aspect markers is a cross-linguistically well-attested pathway (Comrie 1976: 99–101; Bybee 1985: 196; Bybee et al. 1994: 86; Johanson 2000, 2002a; Malchukov 2000: 447).

In the example (1) above, I illustrated reflexes of the resultative nominalizer PTEA*-xA. Below I show how this suffix developed over intermediate stages of perfective nominalizer and relativizer into a past finite suffix.

- (4) a. Kyushu (Japanese) dialects
- yo-ka*
- good-FIN
- ‘It’s OK.’ (Martin 1987: 803);
- b. Manchu
- ahu:n ji-he turgun-de deo*
- elder.brother come-ADNZ reason-DAT younger.brother
- gene-he*
- go-FIN
- ‘Since the elder brother came, the younger brother went away’
- (Gorelova 2002: 488)

c. Middle Mongolian

Sigi Qutuqu ese abu-’a

Sigi Qutuqu NEG take-FIN

‘Sigi Qutuqu has not accepted anything’ (SH 252; Weiers 1966: 198)

d. Old Turkic

čeviř ay-u bęr-ge men

method explain-CVB give-FIN 1SG

‘I will explain the method for you’ (KP 75, 2)

Comparable to the extended paradigmaticity between personal pronouns and person agreement markers in the Indo-European languages, the Transeurasian languages share the source and outcome of a grammaticalization process for various markers, in form and function. The difference with the Indo-European case lies in the fact that the Indo-European personal paradigm consists of a closed and exhaustive set of seven cells, while the number of markers undergoing the finitization and tense development in Transeurasian is open and non-exhaustive. Indeed, contrary to the number of person pronouns in Indo-European, the number of deverbal noun suffixes is not fixed and not all deverbal noun suffixes have undergone finitization. Therefore, the Transeurasian paradigmatic correlations are less diagnostic of common ancestorship than the Indo-European ones.

In addition to systematic correlations in grammatical patterning, we also find shared idiosyncrasies in the non-finite aspect / finite tense paradigm. Although the etymologies (13) to (15) generally reflect the development of aspectually neutral verbal noun suffixes to imperfective adnominalizers to non-past finite suffixes, there are a number of ambiguous cases in which the cognates reflect perfective and past

meaning instead. Consider, for instance, the idiosyncratic meaning of the Korean imperfective adnominalizer MK *-(u/o)l* in (3) in some petrified lexemes. The time expressions K *onul*, MK *wo·nol* ‘today’ and K *wolhay*, MK *wol·hoy* ‘this year’, contain an adnominal form of the verb K *o-*, MK *wo-* ‘to come’, deriving from **wo-[l]·nal* (come-ADNZ day) and **wo-l·hoy* (come-ADNZ year), respectively.¹⁰ 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 lexicalized expressions suggest a (recent) perfective interpretation. Compare the use of MK *·wo-no-n·hoy* (come-PROC-ADNZ year) for ‘next year’, i.e. ‘the year that is coming’ and MK *·ni·ke-n·hoy* (depart-RES-ADNZ year) for ‘last year’, i.e. ‘the year that has departed’. What looks like a puzzling irregularity when studying Korean in isolation, receives an explanation in a comparative Transeurasian framework. It is striking to observe that for Even *-ra ~ -re* and Evenki *-ra ~ -re*—the cognate of this Korean form in Northern Tungusic languages—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; see example (1e). This suggests that the preceding Proto-Tungusic clausal adnominalizer PTg **-ra* could have imperfective or perfective interpretation according to the telicity of the verb base and explains the etymology of Korean ‘today’ and ‘this year’ as a case of perfective reading following telic verb stems.

The temporal interpretation of Evk. *-ra* depends on the actional semantics of the preceding verb (Nedjalkov 1997: 237). Derived from telic verbs, i.e. verbs of achievement, accomplishment and from activity verbs in which a temporal boundary is arbitrary but implied, *-ra* refers to the recent past, as illustrated in (5) 1 and 2. Derived from atelic verbs, such as verbs derived with the habitual aspect marker -

ngnA and verbs of state, *-rA* refers to the present, as illustrated in (5) 3 and 4.

- (5) Evenki
- 1 *Ami-m* *eme-re-n.*
 father-1SG.POSS come-FIN-3SG
 ‘My father has come’ (Nedjalkov 1997: 237)
- 2 *Nungartin* *amin-du-ver* *bele-re.*
 they father-DAT-REFL.POSS help-FIN
 ‘They helped their father’ (Nedjalkov 1997: 237)
- 3 *Nungartin* *amin-du-ver* *bele-ngne-re*
 they father-DAT-REFL.POSS help-HAB-FIN
 solo-s-toki-n.
 boat.upstream-SMLF-CVB-3SG.POSS
 ‘They always help their father when he goes upstream in a boat’
 (Nedjalkov 1997: 238)
- 4 *Sa:-re-n*
 know-FIN-3SG
 ‘He knows’ (Nedjalkov 1997: 237)

The original correlation, which is summarized in Table 30.2, has attired in the majority of Transeurasian languages, but leaves traces in the ambiguity of imperfective/perfective interpretation of certain markers. Note that the development of past finite meaning on the Khitan cognate in example (3d) probably reflects a similar development.

Table 30.2 The impact of verb telicity on aspect and tense interpretation of non-finite and finite suffixes

	non-finite suffix	finite suffix
atelic verb base	imperfective	non-past
telic verb base	perfective	past

In my view, this complex system of correlations in grammatical patterning, including shared idiosyncrasy can be regarded as paradigmatic evidence in the historical comparison of Transeurasian verb morphology.

30.6 How likely is coincidence?

In chapter 40, I relied on a cumulative binomial distribution to estimate the number of comparative sets required in the basic vocabulary to exclude chance as a possible explanation of the similarities. Starting from values for n and p , the cumulative binomial distribution calculates x , the number of binary corresponding items we need to exclude chance with a certain degree of probability. As p , the probability of finding a phonological match by pure chance, and n , the number of trials involved in comparing two datasets increase, x will increase as well.

As the verb suffixes in Table 30.1 consist at best of two phonemes, i.e. CV, the probability of finding a coincidental sound match between two verb suffixes is higher than that of finding one between two basic vocabulary items of CVC structure. We will thus need more morphological comparanda to exclude chance than is the case of the basic vocabulary. Therefore, the typical shortness of morphemes is an important problem that reduces the diagnostic value of morphological comparisons. As a result, similar forms tend to occur by coincidence even in unrelated languages. In addition to the causative-passive reconstruction PTEA **-ti-* in (08) in Table 34.1, for instance, it

is easy to find reconstructions of causatives in **t* in other language families across the world, among others the causatives in Proto-Niger-Congo **-ti*, Proto-Nilo-Saharan **-iṭ* and Proto-Chadic **-d* (Hyman 2014: 111), the transitive-causative in Proto-Dravidian **-tt* (Krishnamurti 2003: 283) and the denominal factitive and deverbal causative in Proto-Uralic **-tā-* / *-tä-* (Janhunen 1982: 33).

Nevertheless, even if the typical shortness of morphemes increases the probability of finding coincidental look-alikes across the languages of the world, this statistical effect is counterbalanced by the factor *n* in the formula, i.e. the number of trials involved in comparing two datasets. The probability that a certain correlation in the verb morphology is due to coincidence decreases with the number of verb markers open to comparison. My birthday being on October 24th, it is easier to find somebody who has their birthday on the same day in a group of 100 people than in a group of 20 people. The probability of finding somebody who is born on the same day therefore decreases with the number of people to select from. 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.

Other chance decreasing factors that need to be considered for the Transeurasian verb morphology include the number of matched segments in the morphological cognates, the number of branches in which a certain morphological item simultaneously has a match, the number of proposed etymologies, the verification of sound correspondences against previously established regularity on the basis of lexical data, shared polysemy, shared idiosyncrasies and extended paradigmaticity.

The greater the number of matched segments in a proposed cognate set, the less likely it is that accident accounts for the similarity. It is exponentially more difficult to find a match between a subsequent consonant and vowel than it is to find a match for a single segment. In the morphological similarity of the **t*-causatives across the languages of the world, only the consonant is matched, while the vowel remains unmatched. In the Transeurasian verb morphology, chance probability decreases because of the consonant and vowel match and because we account for the entire morpheme, not just for some arbitrarily segmented part of it.

Next, the diagnostic power of one etymology that reflects cognates simultaneously in all five branches of the Transeurasian family is much higher than the significance of an etymology that has cognates in only two branches of that family. Hence, the convincing power of “multiple” etymologies such as (04), (05), (07), (08), (09), (11), (13), (14), (15), (16), (18) and (19) that are attested in all five branches of Transeurasian should be taken into account.

Obviously the number of proposed etymologies, the x in the binomial formula, plays a crucial role. The more etymologies, the stronger the evidence. Apart from the matching **t* causatives, there is only one other verb affix with a resemblance between the African macro-groups Niger-Congo, Afro-Asiatic and Nilo-Saharan (Hyman 2014: 111) and some groups such as Indo-European, Kartvelian and Dravidian (Kaiser and Shevoroshkin 1988: 313), notably a causative affix containing a common segment **s*. These two instances may serve as an indication of the number of etymologies for verb affixes that can be expected to match by sheer chance when comparing a handful of random families. Even without explicit chance calculation, one intuitively feels that the twenty etymologies summarized in Table 30.1 are a very striking a number to be attributed to coincidence.

Moreover, for the cross-linguistically corresponding *t-causatives, we are unable to test the assumed consonant correspondences Proto-Niger-Congo *t* :: Proto-Nilo-Saharan *tʼ* :: Proto-Chadic *d* :: Proto-Dravidian *tt* :: Proto-Uralic *t*—let alone the vowel correspondences—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 an independently established set of phonological rules. Obviously, the fact that the formal correspondences in matching morphemes can be verified against pre-existing rules decreases the likelihood of coincidental matches.

In addition, comparing forms with a general functional tag such as “causative” may enhance coincidental matches. However, shared semantic specialization such as the rich common polysemy discussed for the manipulative denominal verbalizer in (03) is more telling.

Finally, shared idiosyncrasies such as the shared allomorphy in example (1) and the imperfective-perfective ambiguity in example (3) further reduce chance similarity. This is also true for the examples of extended paradigmaticity, discussed in Section 30.5.

30.7 Conclusion

Since both supporters and critics of Transeurasian affiliation agree on the diagnostic power of common paradigmatic morphology, I here gave an overview of the verb morphology shared between the Transeurasian languages and indicated to what extent the correlations may be regarded as “paradigmatic”. The correlations in the verb morphology are regular in form and function. From the point of view of form, the comparanda obey to the regular sound correspondences, previously established on the

basis of the lexicon. There are also instances of shared allomorphy, whereby the compared suffixes share a peculiar allomorphy in a phonologically similar environment. From the functional perspective, the correlations go beyond general functional tags as concrete submeanings and detailed functional specializations are shared. Furthermore, the evidence can be viewed as paradigmatic because it displays relationships of grammatical patterning among different morphosyntactic subsystems such as the relationship between non-finite and finite paradigms and that between aspect and tense paradigms. In addition to these paradigmatic correlations, we find shared idiosyncrasies that lack a plausible language-internal motivation, but receive an explanation in the external context of Transeurasian comparison. A seemingly irregular phenomenon in Korean time expressions, for instance, can be explained by the impact of verb telicity on aspect and tense interpretation in the Transeurasian languages.

Even if the typical shortness of morphemes may raise concerns about the chance similarity involved in the comparisons, this statistical effect is counterbalanced by a number of strengths of the Transeurasian evidence. Such chance-decreasing factors include the small size of the dataset of verb morphemes the comparative evidence was selected from, the matching of up to two phonological segments in the morphological cognates, the representation of the majority of etymologies simultaneously in all five branches of the family, the presentation of as much as 20 etymologies, the verification of sound correspondences against previously established regularity on the basis of lexical data, shared polysemy, shared idiosyncrasies and extended paradigmaticity. Shared polysemy, idiosyncrasy and paradigmaticity can also be used as an argument against the borrowing of the proposed cognates in addition to the low borrowability of bound verbal morphology in general. For a more detailed distinction between

borrowing and inheritance as a motivation for the correlations between the Transeurasian languages, I refer to Chapter 47. In sum, I believe that it is linguistically more sensible to explain the above correlations in Transeurasian verb morphology by the inheritance from a common ancestor than to take refuge in explanations such as borrowing and chance similarity.

Acknowledgement

The research leading to these results has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 646612), granted to Martine Robbeets.

¹ Vowel alternations for the denominal verbalizer PJ **-ra-* ~ *-rə-* are reflected in OJ *wa-ra-p-* 'to laugh' and OJ *no₂-ro₂-p-* 'curse', for the processive **-na-* ~ **-nə-* in OJ *ata-na-p-* 'to harm, injure (tr.)' and OJ *udu-no-p-* 'to prize, value (tr.)', for the inclinational **-ma-* ~ **-mə-* in OJ *to₂ga-me₂-* (< **tonka-ma-Ci-*) 'to censure, find fault with (tr.)' and OJ *to₂yo₂-mo-s-* 'to make resound', for the reflexive-anticausative **-pa-* ~ *-pə-* in OJ *kanape₂-* (< **kana-pa-Ci-*) 'to make suitable, possible, make fit, grant (a request)' and OJ *uru-po-p-* 'get damp, get moist, receive profits, get enriched' and for the nominalizer **-ra* ~ **-rə* in OJ *aka-ra* 'red' and OJ *wo₂so₂-ro₂* 'precocious, early ripening'.

² Middle Korean unrounded vowels [ʌ] and [i] are represented by *o* and *u*. In Proto-Korean these vowels are reconstructed as **ʌ* and **i*. The dots in the Middle Korean words represent the distinctive pitch of the following syllable: one dot for high, two dots for rising, and unmarked syllables are treated as low.

³ WMo. [ɣ] is the allophone of /g/ before back vowels *a*, *o*, *u* in the same way as [q] is the allophone of /k/ in that position

⁴ Old Japanese distinguished between two values for later *e*, *i*, *o* in certain syllables, which are indexed with subscripts *i*₁ versus *i*₂, *e*₁ versus *e*₂ and *o*₁ versus *o*₂. Japanese verbs and verbal adjectives can be distinguished according to two prosodic classes, called A and B, corresponding to a high and low initial tone, respectively. Prosodic classes for nouns are more complex and marked with a number notation. Although adequate information about these prosodic patterns is only available from the Middle Japanese stage onwards, I have added the accent classes here because they can be used to support the derivation. Prosodic classes for nouns 1.1, 1.2, 2.1, 2.2 are expected to derive A verbs with an initial high tone, while classes 1.3, 2.3, 2.4, 2.5 are expected to derive B verbs with an initial low tone.

⁵ Vowels and accent class of the derived verb are aberrant.

⁶ Vowels of the derived verb are aberrant.

⁷ Hirara, a southern Ryukyuan language spoken on the Miyako Islands has *-i* (< **iri* < PR **-uru*), where other Ryukyuan languages have maintained the liquid in their reflex of this suffix, e.g. Shodon (Amami) *-ur*, Yamatoma (Amami) *-uru*, and Shuri (Okinawa) *-uru*. The original Ryukyuan suffix **-uru* can be traced back to PJ **wo-ra* (be-NMLZ), which assimilated to PJ *[w]oro and lost its initial labial. As **-oro* started fusing with the converbial **-i* form of the verb, it regularly raised to PR **-uru*. (Robbeets 2015: 339–341).

⁸ The Middle Korean subjunctive attentive ending MK *-(u/o)·la* is morphologically segmentable into the deverbial noun suffix MK *-(u/o)l* and the vocative particle *a*, which follows nouns, e.g. K *palk-un tal-a* (shine-ADNZ moon-VOC) ‘Oh shining moon!’.

⁹ Examples of finite use in Mongolic are lacking, but it is remarkable that among the markers of past tense in Khitan, we find the suffix *-r*, which is only preserved in telic expressions, such as ‘become’, ‘become appointed’, ‘become awarded’, ‘compose an edict’, ‘write this text’ (Kane 2009: 145–146).

¹⁰ MK /l/ drops before /n/ and the other apicals /s/, /c/ and /t/.

This is a draft version of a chapter that appears in Robbeets, M. and A. Savelyev (eds). *The Oxford Guide to the Transeurasian Languages* (OUP, 2020), see <https://global.oup.com/academic/product/the-oxford-guide-to-the-transeurasian-languages-9780198804628>. The research leading to these results has received funding from the European Research Council under the Horizon 2020 Program/ ERC Grant Agreement n. 646612 granted to Martine Robbeets.