

Robbeets, M. (2017). Shared verb morphology in the Transeurasian languages: Copy or cognate? In M. Robbeets (Ed.), *Stability and borrowability* (pp. 136-153). London [u.a.]: Routledge.

21 SHARED VERB MORPHOLOGY IN THE TRANSEURASIAN LANGUAGES:

COPY OR COGNATE?

Martine Robbeets

The genealogical relationship of the Japonic, Koreanic, Tungusic, Mongolic and Turkic languages, here referred to as “Transeurasian”, is among the most disputed issues of historical comparative linguistics. The major objections raised against the relatedness of these languages are, first, that they do not have enough bound morphology in common, and, second, that all similarities can be attributed to code-copying. Using the traditional comparative method as a basic tool, both objections are examined in this paper. Comparing copying patterns with genealogical patterns in a cross-linguistic sample of languages, twelve guidelines for the distinction between the effects of contact and inheritance in shared morphology are developed. Applying these criteria to the verb morphology shared by the Transeurasian languages, it is argued that the common morphology can best be accounted for by inheritance from a common ancestor.

1 Introduction

The genealogical relationship of the Transeurasian languages is among the most disputed classifications of historical comparative linguistics. Rather than using the traditional term “Altaic”, the term “Transeurasian” is proposed in reference to a large group of geographically adjacent languages that share a significant number of linguistic properties and include at most five linguistic families: Japonic, Koreanic, Tungusic, Mongolic and Turkic. The major objections raised against the overall genealogical relatedness of these languages are first, that they have some but not enough bound morphology in common, and second, that all similarities can be attributed to code-copying. In this chapter both objections will be examined.

On the basis of a comparison between copying patterns and genealogical patterns in a cross-linguistic sample of languages, the first part of this chapter develops twelve guidelines that help to distinguish between the effects of contact and those of inheritance in shared morphology in general. In the second part these guidelines will be applied to the

verb morphology shared by the Transeurasian languages. By way of conclusion, an answer will be provided to the question whether the shared verb morphology in the Transeurasian languages can best be accounted for by contact or by inheritance.

2 Guidelines for the distinction between copies and cognates

Contrasting case studies of contact-induced morphology—including mixed languages—with cases of inherited morphology leads to the following guidelines. The guidelines one to six are indications against inheritance, while the guidelines seven to twelve increase the probability of a genealogical explanation.

2.1 Attachment of shared morphemes: to shared roots only vs. also to unrelatable roots

An indication of a copy is a restriction of shared morphemes to shared roots only. Agia Varvara Romani, a Romani dialect spoken in a suburb of Athens, for instance, copied the Turkish nonfocal present paradigm, but all copied morphemes are hosted by verbs copied from Turkish (Iglá 1996, 214-216; Friedman 2009, 112). Bakker and Hekking (this volume, xx) find that contact with Spanish has substantially affected the morphology of three Amerindian languages, Quechua, Guaraní and Otomi, but there are hardly any cases where the copied derivational or inflectional markers are found on native lexical entities; they are mainly restricted to nouns, verbs and adjectives copied from Spanish. As far as contact between Hittite and Luvian, the Indo-European languages of Anatolia, is concerned, Folke Josephson (this volume, xx) remarks that the evidence for copied bound morphology is restricted to some case endings and nominal derivational suffixes that were copied into Hittite only when attached to Luvian host lexemes. Kossmann (2010) argues that the borrowing of morphological paradigms together with foreign lexicon is a well-attested phenomenon in the languages of the world. Therefore, only shared affixes that can attach to native, unrelatable roots will be taken into account as genealogical evidence.

Table 1: Agia Varvara Romani copy of the present paradigm of Turkish *çalış-* ‘work’

	Turkish model (work-PRS-PERS)	Romani copy
1 SG	<i>çalış-ir-im</i>	<i>calusurum</i> ‘I work’
2 SG	<i>çalış-ir-sin</i>	<i>calusursun</i> ‘you work’
3 SG	<i>çalış-ir</i>	<i>calusur</i> ‘he works’

1 PL	<i>čališ-ir-iz</i>	<i>calusurus</i> 'we work'
2 PL	<i>čališ-ir-siniz</i>	<i>calusursunus</i> 'you work'
3 PL	<i>čališ-ir(-lar)</i>	<i>calusur(lar)</i> 'they work'

2.2 Shared suffix complexity: suffix strings vs. simplex morphemes

Shared suffix strings that are morphologically segmentable in one language, but not in the others, are the result of copying. This is for instance the case in the example in Table 1, where Agia Varvara Romani copies the Turkish present tense along with the entire paradigm of Turkish person marking. A similar example illustrated in Table 2 is the copying of the Yakut presumptive-assertive paradigm as a presumptive in Uchur Evenki and as an assertive in Sebjan-Küöl Even. The copied strings consist of the Yakut presumptive-assertive mood in *-TAx-* along with the entire paradigm of Yakut person marking (Malchukov 2003, 244; 2006, 126-27; Pakendorf 2009, 98-105 and 109-110; Comrie 2010, 26). These person markers do not occur as simplex morphemes in Even, Evenki or elsewhere in Tungusic. The correlations can easily be unmasked as copies because the suffix strings are morphologically complex in one language, but not in the other.

Table 2: Uchur Evenki and Sebjan-Küöl Even copy of the Yakut assertive-presumptive paradigm

	Yakut model (go-PRES-PERS)	Uchur Evenki copy (kill-CONN-PRES)	Sebjan-Küöl Even copy (spend.the.night-CONN-ASS)
1 SG	<i>bar-day-im</i>	<i>wa:-r-dayim</i>	<i>a:ηηa-j-dagim</i>
2 SG	<i>bar-day-iŋ</i>	<i>wa:-r-dayiŋ</i>	<i>a:ηηa-j-dagiŋ</i>
3 SG	<i>bar-day-a</i>	<i>wa:-r-daya</i>	<i>a:ηηa-j-daga</i>
1 PL	<i>bar-dax-pit</i>	<i>wa:-r-dakput</i>	<i>a:ηηa-j-dakpit</i>
2 PL	<i>bar-dax-xit</i>	<i>wa:-r-dakkit</i>	<i>a:ηηa-j-dakkit</i>
3 PL	<i>bar-dax-tara</i>	<i>wa:-r-daktara</i>	<i>a:ηηa-j-daktara</i>

2.3 Shared function: restricted to secondary semantics vs. also primary semantics

When the semantic correspondence concerns a meaning that is demonstrably secondary to one of the participating morphemes, we are dealing with a copy. In the case of Yakut influence in Northern Tungusic in Table 2, the Yakut suffix *-TAx* functions as a non-finite conditional-temporal marker, apart from marking the presumptive-assertive mood in the

finite clause. It is safe to assume that the meaning of Yakut *-Tax* developed from temporal to conditional to presumptive to assertive. The Evenki and Even forms being restricted to finite use and the Even form in Pakendorf's corpus being restricted to finite assertive, the semantics shared with Yakut are secondary. This observation indicates that the similarities have a non-genealogical source.

2.4 Shared form: contradiction vs. confirmation of established sound correspondences

If previous research has provided a system of sound correspondences on the basis of lexical comparison, as it has in the case of the Transeurasian languages (Starostin et al. 2003, 24-25 and 92-93; Robbeets 2005, 373-376), we should examine whether the compared morphemes obey these rules. In Acadian French spoken on Prince Edward Island in Canada, for instance, the English loan *back* replaces the French prefix *re-* with verbs such as *revenir* 'come back' as in *venir back*, *arriver back*, *mettre back* (King 1999, 116-125). Here Grimm's law in Indo-European can prevent us from misinterpreting Acadian French *back* and English *back* as cognates. Lexical comparison shows that English word-initial *b-* corresponds regularly with French *f-*, as in *barley* and *farine* 'flour', *brother* and *frère* 'brother', *bottom* and *fond* 'bottom', *brass* and *ferre* 'iron' etc. Thus, contradiction of established sound correspondences is an indication against genealogical retention.

2.5 Distribution: limited to contact zones vs. spread among low contact languages

The limited distribution of morphemes within a particular contact zone is indicative of copying. The Yakut influence on Northern Tungusic verb morphology is restricted to the Yakut-Tungusic contact zone and does not occur in the Evenki and Even dialects to the East. The Albanian admirative present *-ka*, discussed by Friedman (this volume, xx), is copied into Romanian, but restricted to the Frasheriole Aromanian dialect of Gorna Belica, without spreading to other dialects spoken in that same village. The limitation of shared verbal morphology to contact zones is an indication against inheritance.

2.6 Shared paradigms: specific morphosyntactic subsystems vs. subsystems in general

The examples of copied verbal morphology discussed so far are in line with Seifart's (2010) hypothesis that if morphemes are copied at all, it is often the case that more than one form is copied. In the workshop that inspired the publication of this volume he found, for instance, that in Resígaro certain morphological subsystems such as nominal classification

and number marking have been copied entirely, while others have hardly been influenced at all. His observation that paradigmatic copying is restricted to specific morphosyntactic subsystems seems to apply more generally. The split can be between nominal and verbal morphology as in the languages of Arnhem Land (Heath 1978, 105) and in Michif (Thomason & Kaufman 1988, 228-233; Bakker 1997, 97-102; Comrie 2008, 21-22) or between inflectional and derivational morphology as in Northern Tungusic (Pakendorf 2009) and in Ma'a (Thomason 1983; Thomason & Kaufman 1988, 223-228), but it may also affect different component parts of these subsystems such as finite vs. non-finite as in Mednyi Aleut (Thomason & Kaufman 1988, 233-238; Sekerina 1994; Thomason 1997; Comrie 2008, 24-31; 2010, 28-30). In other words, shared paradigmatic morphology that is restricted to specific morphosyntactic subsystems is an indication against inheritance.

2.7 Shared infrequent grammaticalization: selective vs. global

If shared morphology answers to the guidelines seven to twelve, the probability of common ancestorship increases. An indication of inheritance is globally shared infrequent grammaticalization. Most cases of so-called contact grammaticalization referred to in the literature involve selective semantic copying, or, in Heine and Kuteva's (2005, 7) terms "replication". A classic example is the copying of progressive aspect on the verb for 'to carry' in southern Basque under Spanish influence (Jendraschek 2007, 157). In Spanish, the verb *llevar* 'carry' can express progressive aspect when attached to a gerund of a lexical verb as in *llevar estudiando* 'be studying'. In the southern varieties of Basque, where all speakers are bilingual in Spanish, the grammaticalized meaning of the Spanish verb is copied on the Basque verb *eraman* 'to carry' resulting in similar progressive constructions. The Basque verb has thus maintained its native form *eraman* and has only copied the progressive meaning from Spanish. Therefore we can say that shared grammaticalization due to contact is selective: it shares function, not form.

Globally shared grammaticalization, however, is more probably than not genealogically motivated. The term "global" is reminiscent of Johanson's (2002) code-copying terminology, referring to a full correspondence, including form and function. A classical example of globally shared grammaticalization comes from the Romance future markers. Romance languages share a root for the verb 'to have' in form and function, such as French *avoir*, Spanish *haber*, Portuguese *haver* and Italian *avere* as well as the grammaticalized future marker as in French *chante-rons*, Spanish *canta-remos*, Portuguese *canta-remos* and Italian *cante-remo* 'we will sing' (Pinkster 1987: 203-214, Klausenburger

2000). This indicates that the process of grammaticalization was already well on its way in common Romance because it is the best way to explain why the same path in the formation of a new future was followed in form and function by so many Romance languages. Corroboration comes from attestations in Vulgar Latin of the 6th and 7th centuries of forms such as *daras* 'you will give' and *pussediravit* 'shall possess'. It is very unlikely that this example should be attributed to language contact because the shared grammaticalization involves globally corresponding morphemes.

Moreover, the example is unlikely to be the result of universals of grammaticalization because the have-to-future development is common in Romance languages, but it is not a frequent pathway cross-linguistically. Since code-copying and universals are unlikely, inheritance is the best explanation for globally shared infrequent grammaticalization.

2.8 Shared function: categorial clarity vs. categorial opacity

Categorial clarity refers to morphosemantic transparency whereby the grammatical function of a morpheme can be understood without considering the context of its broader morphosyntactic environment. Gardani (2008, 88) concludes that none of the copied morphemes involved in his research on inflectional borrowing can be viewed as categorially opaque. Heath (1978: 111-112) finds that in the Australian languages in the Arnhem Land area the verbal inflectional suffixes marking tense, aspect, mood and negativity are totally immune for copying. He explains this observation by an interaction of impeding factors, one of them being opacity. In Nu, for instance, inflectional suffixes make sense only in the light of other morphemes such as prefixes and particles. The so-called 'past-2' in Nu can be past continuous, past negative, past potential, or past negative potential, dependent on which prefixes and negative particles are present in the verb complex. It appears that, if corresponding morphemes are categorially opaque, they are likely to be inherited.

2.9 Shared function: noncumulative vs. cumulative

Cumulative morphemes, i.e. morphemes with an unanalyzeable form that simultaneously blend several distinct morpho-syntactic features together, are relatively impervious to copying. Gardani (2008, 89) finds that 70% of all cases of copied inflectional morphemes involved in his study are monofunctional. Moreover, in cases where forms with cumulative exponence are copied, the cumulation tends to be reduced in the copying language. When the Frasheriote Aromanian dialect copied the Albanian portmanteau morpheme *-ka*

expressing third singular admirative present, for instance, it reinterpreted the morpheme regardless of person and number. Therefore, shared cumulation is regarded as an indication of inheritance.

2.10 Shared variant allomorphy: reduced vs. complete

The copying language tends to replace phonologically conditioned alternants by fewer allomorphs. Russian and Mednyi Aleut share a great amount of finite verb morphology, as illustrated in Table 3. Russian has two conjugational classes, differentiated by the inflectional person-number suffixes in the present tense. The most explicit difference is that the third person plural in the first conjugation is *-(j)ut*, whereas in the second conjugation it is *-(j)at*. Mednyi Aleut has reduced the allomorphy: it has only *-jut* in the third plural (Sekerina 1994, 25; Comrie 2008, 27; 2010, 29). Corresponding morphemes that share complete variant allomorphy are likely to be inherited.

Table 3: Shared finite verb morphology in Mednyi Aleut and Russian

	Russian model		Mednyi Aleut copy 'work'
	1 'work'	2 'speak'	
1SG	<i>rabota-ju</i>	<i>govar-ju</i>	<i>aba-ju</i>
2 SG	<i>rabota-eš'</i>	<i>govar-iš'</i>	<i>aba-iš</i>
3 SG	<i>rabota-et</i>	<i>govar-it</i>	<i>aba-it</i>
1 PL	<i>rabota-em</i>	<i>govar-im</i>	<i>aba-im</i>
2 PL	<i>rabota-ete</i>	<i>govar-ite</i>	<i>aba-iti</i>
3 PL	<i>rabota-jut</i>	<i>govar-jat</i>	<i>aba-jut</i>

2.11 Comparative setting: binary vs. multiple

The sets of copied verb morphology discussed so far all have a binary setting in common. Morphological copying typically goes from a model language into a basic language. Occasionally it may progress into a third language as in the cases discussed by Josephson (this volume, xx) and Gutiérrez-Morales (this volume, xx). One example involves the Classical Persian volitional prefix *be-* copied into Kurdish and from there into Neo-Aramaic dialects, the other example involves the Spanish nominalizer *-ero* copied into Nahuatl and from there into Sierra Popoluca. Copying processes in three stages, like these, are relatively

rare. The more language families involved in the comparison of a bound morpheme, the more likely inheritance becomes.

2.12 Attestation of etymology members: full vs. gaps

Code-copying is typically unidirectional and linear, progressing from one contact language into the other and then, perhaps, into the next. Genealogical divergence, by contrast, can be pictured as the rings formed when a stone is thrown into the water. Innovations start in the centre and push the older forms towards the periphery. This observation explains why some very conservative inherited items leave traces in remote areas, but are barely attested elsewhere in the linguistic continuum. In this way gaps in the attestation of members of an etymology may be relevant. The absence of a morpheme in an intermediate contact language can be taken as an indication of genealogical relatedness.

3 Overview of shared verb morphology in the Transeurasian languages

On the basis of previous research, etymologies can be advanced for twenty-one different verb suffixes relating Japanese to the Transeurasian languages. An overview is given in the Tables 4 to 11. Within the limited space available it is impossible to justify every individual proto-form involved in the comparisons. For this purpose reference is made to Robbeets (2007a/b; 2009; 2010) and to the manuscript in preparation on Japanese and the Transeurasian verb system. Section 4 will provide an evaluation of the etymological proposals from the viewpoint of the twelve guidelines to distinguish copies and cognates proposed above.

Table 4: Shared actionality suffixes in the Transeurasian languages

	(a) pJ	(b) pK	(c) pTg	(d) pMo	(e) pTk	pTEA
(1)	*-ra- denominal verb N/O		*-lA:- denominal verb N	*-lA- denominal verb N/O	*-lA- denominal verb N/O	*-lA- denominal verb N/O
(2)	*-na- process N/A	*-nu/o- process A/V	*-nA- natural proces N/A/V	*-n(A/i)- spontaneity A/V	*-n- spontaneity A/V	*-nA- process A/V
(3)	*-ka- iconic O	*-ki- iconic O	*-ki- iconic O	*-ki- iconic O	-kl- iconic O	*-ki- iconic O
(4)	*-ma- inclination N/A/V	*-mu/o- inclination N/A/(V?)	*-mA- intention N/A	*-mA- intention A/V		*-mA- intention N/A/V
(5)	*-ka-	*-k(u/o)-	*-gA-		*-k-	*-gA-

	inchoative N/A/V	inchoative N/A/(V?)	inchoative N-V		inchoative N/A/V	inchoative N/A/V
--	---------------------	------------------------	-------------------	--	---------------------	---------------------

Table 5: Shared diathetical suffixes in the Transeurasian languages

	(a) pJ	(b) pK	(c) pTg	(d) pMo	(e) pTk	pTEA
(6)	pJ *-ya- auxiliary fientive passive		pTg *-dA- auxiliary fientive	pMo *-dA- auxiliary fientive passive	pTk *-(A)d- auxiliary fientive passive	pTEA *-dA- auxiliary fientive passive
(7)	pJ *-ta- caus.-pass.	pK *-ti- caus.-pass.	pTg *-ti- caus.-pass.	pMo *-ti- causative	pTk *-tl- caus.-pass.	pTEA *-ti- caus.-pass.
(8)	pJ *-pa- anticaus. fientive	pK *-pu/o- anticaus.	pTg *-p- refl.-anticaus.	pMo *-pu- refl.-anticaus.	pTk *-(p)U- anticaus. fientive	pTEA *-pə- reflexive
(9)	pJ *-ra- anticaus. fientive	pK *-(u)l- anticaus. fientive	pTg *-rA- anticaus.	pMo *-rA- anticaus.	pTk *-(l)r- anticaus. fientive	pTEA *-rA- anticaus. fientive
(10)	pJ *-(C)i- caus.-pass.	pK *-ki- caus.-pass	pTg *-ki- causative	pMo *ki- 'do, make'	pTk *ki(l)- 'do, make'	pTEA *ki- 'do, make'

Table 6: Shared (imperfective) nominals developing to (non-past) finite suffixes in the Transeurasian languages

	(a) pJ	(b) pK	(c) pTg	(d) pMo	(e) pTk	pTEA
(11)	pJ *-n (ad)nominal finite	pK *-n (ad)nominal finite	pTg *-n (ad)nominal finite	pMo *-n (ad)nominal finite	pTk *-n (ad)nominal finite	pA *-n (ad)nominal finite
(12)	pJ *-m (ad)nominal finite	pK *-m nominal finite	pTg *-mA (ad)nominal finite	pMo *-m(A) (ad)nominal finite	pTk *-m(A) (ad)nominal -	pA *-mA (ad)nominal -
(13)	pJ *-rə (ad)nominal finite	pK *-l (ad)nominal finite	pTg *-rA (ad)nominal finite	pMo *-r (ad)nominal -	pTk *-(A)r (ad)nominal finite	pA *-rA (ad)nominal -

Table 7: Shared (perfective) nominals developing to (past) finite suffixes in the Transeurasian languages

	(a) pJ	(b) pK	(c) pTg	(d) pMo	(e) pTk	pTEA
(14)	*-ka RES NML (ad)nominal finite	*-kA- RES	*-gA ~ *-kA RES NML (ad)nominal finite	*-gA ~ *-kA RES NML (ad)nominal finite	*-gA ~ *-kA RES NML (ad)nominal PF non-PST finite	*-gA ~ *-kA RES NML
(15)	*-sa RES NML nominal finite	*-s RES NML nominal	*-sA RES NML nominal finite	*-sA RES NML nominal	*-sA RES NML nominal finite	*-sa RES NML

	A / (COP)					
--	-----------	--	--	--	--	--

Table 8: Shared deverbal nouns developing to converbs in the Transeurasian languages

	(a) pJ	(b) pK	(c) pTg	(d) pMo	(e) pTk	pTEA
(16)	*-i NML/ CONV	*-i NML / adverb.	*-i NML	*-i NML / adverb.	*-i NML/CONV	*(-)i 'fact, thing'
(17)	*-ku NML/ CONV	*-k(o) NML / adverb.	*-ku / -kü *-gu / -gü NML / CONV	*-ku / -kü *-gu / -gü NML / adverb.	*-ku / -kü *-gu / -gü NML	*-kU *-gU NML / CONV

Table 9: Shared negation markers in the Transeurasian languages

	(a) pJ	(b) pK	(c) pTg	(d) pMo	(e) pTk	pTEA
(18)	*ana- 'not exist' PREP	*an- 'not be' PREP	*ana- 'not exist' PREP			*ana- 'not exist' PREP
(19)		*e-? 'not be' PREP	*e- 'not be' PREP	*e- 'not be' PREP	*ä- 'not be'	*e- 'not be' PREP

Table 10: Shared markers relating to person agreement in the Transeurasian languages

	(a) pJ	(b) pK	(c) pTg	(d) pMo	(e) pTk	pTEA
(20)	*-n(u)- stem-final obliquus	-	*-n- stem-final obliquus	*-n(a)- stem-final obliquus	*-n- stem-final obliquus	*-n- stem-final obliquus

Table 11: Shared interrogative markers in the Transeurasian languages

	(a) pJ	(b) pK	(c) pTg	(d) pMo	(e) pTk	pTEA
(21)	pJ *ka interrog. particle	pK *ka interrog. particle	*ka- interrog. pronoun root	*ke- interrog. pronoun root	*ka- interrog. pronoun root	*ka- interrogative pronoun root

4 Evaluation: copy or cognate?

4.1 The unlikelihood of code-copying

For the etymological evidence summarized in the Tables 4 to 11, none of the indications of codecopying (2.1 to 2.6) are fulfilled. First, it is noticeable that the shared morphemes are not restricted to etymologically equivalent roots. They derive comparable meaning even from bases that are semantically very similar but etymologically clearly distinct. This is

remarkable, for instance, in the case of etymology (1) for the common denominal verb suffix pTEA *-*la*- that adds the meaning 'carry out a difficult action in order to achieve the base noun' in derivations such as OJ *ipo* 'hut' -> OJ *ipor*- 'lodge in a hut', Ud. *aŋa* 'night shelter' -> *aŋala*- 'make a night shelter', WMo. *ger* 'house' -> *gerle*- 'found a house of his own, marry', OTk. *äv* 'house' -> *ävlä*- 'furnish (sb.) with dwellings, marry (sb.) off', or in the case of etymology (5) for an inchoative denominal verb suffix pTEA *-*ga*- in derivations such as OJ *mi₁du* 'water' -> *mi₁duk*- 'get soaked (in water)', MK *mul* 'water' -> MK *mulk*- 'be watery, be thin', and OTk. *šī* ~ *čī* 'moist' > *čī-k*- 'to get moist'.

Second, the majority of the morphemes compared in the Tables 4 to 11 are either monophonemic or monosyllabic and as such there is no internal evidence that suggests that they are further segmentable into compound morphemes.

Third, the functional correspondences of the etymologies are not restricted to a meaning that is demonstrably secondary to one of the participating morphemes. In agreement with Nichols (this volume, xx) who shows that our faith in the stability of pronouns in Eurasia is misplaced, I am reluctant to explain most correspondences in the personal pronoun paradigm as inherited. It seems, for instance, that some personal pronouns are the result of grammaticalization from original demonstratives, while others are not. There is some internal evidence that suggests that the first person pronoun in Old Turkic developed from a proximal demonstrative. Erdal (2004, 195) observes that in Old Turkic "*bän*, *bini* 'I; me' and *biz* 'we' share a morpheme *b^o* with *bo*, the demonstrative of close deixis (presumably pointing at the domain of the speaker) and *bärü* 'hither; since' which signals movement towards the here and now of the speaker." The supposed development contradicts a genealogical connection with formally similar first person pronouns in Mongolic, Tungusic and Japanese, which do not bear evidence to such a grammaticalization process.

Fourth, the phonological correspondences generally confirm the sound correspondences established on the basis of lexical comparison.

Fifth, the reconstruction of the morphemes in the individual branches is based upon well-distributed morphemes, that are not restricted to a certain contact zone. For Proto-Turkic, I take Chuvash and Old Turkic into account, for Tungusic, Northern Tungusic, Southeastern Tungusic and Southwestern Tungusic, for Japanese, Old Japanese and Ryukyuan languages and so on.

Finally, the etymological data in the list are not restricted to specific morphosyntactic subsystems. The comparisons (1) to (10) of derivational morphemes do not outnumber the

inflectional evidence in the etymologies (11) to (21). Finite as well as non-finite morphology is compared. Various categories such as actionality, diathesis, negation, tense and agreement are represented. Therefore, it appears that there are no observable imbalances across morphosyntactic subsystems.

4.2 The likelihood of inheritance

4.2.1 Globally shared infrequent grammaticalization

Besides the indications against code-copying presented in 4.1, the evidence also displays a number of characteristics suggesting that the languages under inspection are genealogically related. For seven etymologies input and outcome of grammaticalization is shared in form and meaning, with the grammaticalization pathway being crosslinguistically rather uncommon. It concerns the etymologies (7) and (10), reflecting grammaticalization from causative into passive and the etymologies (11) to (15), reflecting “finitization”, the grammaticalization from non-finite into finite suffixes. Etymology (7), for instance, leads to the reconstruction of a common causative-passive suffix pTEA **-ti*. Although the development of passives from causatives occurs cross-linguistically, Haspelmath (1990, 49) finds that “passives from causatives do not seem to be as frequent as the passives from active auxiliaries and passives from NP-reflexives.” Therefore, we can say that the following form-function matches reflect an infrequent grammaticalization process, reminiscent of the development of *habere* into future in the Romance languages

(7a) OJ *-t-* < pJ **-ta-*

causative

OJ *ke₂-* ‘get extinguished (intr.)’ -> OJ *ke₂t-* ‘make vanish (tr.)’

EOJ *panar-* ‘get distant, be expelled (intr.)’ -> OJ *panat-* ‘separate, alienate (tr.)’

Passive

pJ **ayama-* ‘mistake (tr.)’ in OJ *ayamar-* ‘err, make a mistake, apologize (intr.)’ => OJ *ayamat-* ‘err, make a mistake (intr.)’

OJ *kudar-* ‘go down (intr.)’ ~ OJ *kudas-* ‘take down(tr.)’ -> OJ *kutat-* ‘come down, end, deteriorate (intr.)’

(7b) MK *-t-* < pK **-ti-*; MK *-chi-* < pK **ti-ki*; MK *-chu-* < pK **ti-kwu*

causative

K *kulu-*, MK *kulu-* ‘be wrong (intr.)’-> K *kuluchi-*, MK *kulu·ch(u)-* ‘ruin (tr.)’

K *sos-*, MK *swos-* ‘rise (intr.)’ -> K *soschi-* ‘raise (tr.)’

pK **a-* ‘exist’ in MK *-e/ -a-* infinitive -> MK *et-* ‘get, receive’

passive

K *kunh-*, MK *kunh-* ‘stop (tr.)’ -> K *kuchi-*, MK *kunchi-* ‘put an end to (tr.); come to an end (intr.)’

K *coch-*, MK *cwos-* ‘pursue (tr.)’ -> K *cchocki-*, MK *cwoschi-* ‘be pursued (intr.)’

pK **mo-* ‘bring together’ in MK *mwoy-* ‘accompany, escort (someone respected)(tr.)’ (incorporates *-i-* causative), *mwosi-* ‘accompany (tr.)’, *mwoy·ho-* ‘gather, bring together (tr.)’ -> MK *mwot-* ‘come together (intr.)’

(7c) Even *-c- / -t- ~ -ci-*, Ma. *-tA- ~ -cA-*, Evk. *-t- ~ -ci-*, Neg. *-c- / -t- ~ -ci-*, Ud. *-si-*, Na. *-ci- ~ -si-* < pTg **-t- ~ -ti-*

causative

Even *oli-* ‘boil (tr.)’ -> *oli:t-* ‘bring to boil (tr.)’

Even *huk-* ‘be hot (intr.)’ -> *hukci-* ‘warm, heat up (tr.)’

passive

Even *el-* ‘stand up’ -> *elat-* ~ *elac-* ‘stand, be standing’

Even *hor-* ‘catch, capture (tr.)’ -> *horci-* ‘be caught, be captured’

(7d) WMo. *-ci-* < pMo **-ti-*

equipollent causative

WMo. *ebde-re-* ‘break down, fall to pieces (intr.)’ -> WMo. *ebdeci-* ‘break, destroy, ruin (tr.)’

WMo. *jada-ra-* ‘unfold, unwrap, loosen (intr.)’ -> WMo. *jadaci-* ‘untie, unroll, undo (tr.)’

WMo. *nuyu-ra-* ‘be folded, stoop (intr.)’ -> WMo. *nuyuci-* ‘fold (tr.)’

(7e) OTk. *-(X)t-* < pTk **-ti-*

causative

OTk. *ari-* ‘be(come) clean, pure (intr.)’ -> OTk. *arit-* ‘clean, purify (tr.)’

OTk. *bak-* ‘look at (intr.)’ -> OTk. *bakit-* ‘to make someone look at something (tr.)’

passive

OTk. *kov-* ‘follow, chase’ -> OTk. *kovit-* ‘get chased’

OTk. *te-* ‘say (tr.)’ -> OTk. *tet-* ‘be said to be, be called, be considered (intr.)’

4.2.2 Categorical opacity

Nine etymologies have members that are categorially opaque. In the examples (7) and (10) the morpheme can be interpreted as a causative or a passive marker, depending on the morphosyntactic environment in which it occurs. In examples (11) to (15) the function of the morpheme as a participle, a deverbal noun suffix or as a finite verb form can be understood only by its broader morphosyntactic context. In examples (16) and (17), the interpretation of the suffix as a deverbal noun marker or as a converb is also dependent on the sentence in which it occurs.

4.2.3 Cumulative exponence

Two etymologies reflect cumulative exponence because they share simultaneously expressed distinct functions. The etymologies (14) and (15) combine the derivation of a deverbal adjective or noun with resultative aspect. Etymology (14), for instance, leads to the reconstruction of a common resultative deverbal noun **-gA*, lexicalized in Japanese, Mongolic and Turkic but still productive in Korean and Tungusic. The fact that all languages, except Korean, share the same form in addition to sharing two cumulative functions, deverbal noun and resultative, is an indication of inheritance.

(14a) OJ *-ka* < pJ **-ka* resultative nominalizer

OJ *me₂* ‘eye’ and OJ *tura-* ‘be though’ -> OJ *me₂turaka* ‘strange’ in e.g. OJ *me₂turaka pi₁to₂* ‘strange person’

pJ **sintu-* ‘go down’ in OJ *sidum-* ‘submerge, sink (intr.), get quiet’ -> OJ *siduka* ‘quiet’

pJ **kasu-* ‘become faint’ in OJ *kasum-* ‘be hazy, get misty, get dim’ -> OJ *kasuka* ‘faint (of sound or color)’

(14b) MK *-•ke-/-•ka-/-•Ge-/-•Ga- / -•e- / -•a-* resultative < pK **-ka-* resultative

MK *ˈti•na-ke-n ˈnyey ˈnwuy* (Martin 1992, 603)

pass.by-RES-PCP old world

‘a long past ancient world’

(14c) Ma. *-hA* ~ *-kA*, Evk. *-kA*, Even *-kA*, Ud. *-gA* ~ *-kA*, Na. *-kA* < pTg **-gA-* / **-kA-* resultative nominalizer

Ma. *ere abala-me gehe-he gucu-sa* (Gorelova 2002, 257)

this hunt-CONV go-PERF.PCP companion-PL

‘companions, who have gone hunting’

(14d) WMo. *-GA / -kA(i)* persistent perfect < pMo **-gA / *-kA*- resultative nominalizer
 WMo. *kebte*- ‘lie down, recline (intr.)’ -> WMo. *kebtege* ‘horizontal (adj. and adv.)’
 WMo. *edü*- ‘begin, start, commence (tr.)’ -> WMo. *edüge* ‘now, at present, contemporary (adj. and adv.)’

(14e) OTk. *-gA / -kA* < pTk **-gA / *-kA* resultative nominalizer
 OTk. *sal*- ‘move, put in motion, agitate’ -> OTk. *salga* ‘restive’
 OTk. *kis*- ‘pinch, squeeze, reduce (tr.)’ -> OTk. *kisga* ‘short’

4.2.4 Shared variant allomorphy

The seemingly irregular allomorphy in the Turkic, Mongolic and Tungusic members of etymology (14) appears to go back to a regular allomorphy in the ancestral language, whereby a resultative nominalizer pTEA **-gA* became **-kA* when it followed verb stems on *..r-* (but not *..l-*), *..p-*, *..m-* and *..n-*. The Tungusic resultative nominalizer **-gA* has an irregular allomorph **-kA* in Manchu after original *..r-*, *..p-*, *..m-* and *..n-* stems and in Udeghe after consonant stems in general. Manchu *je*- ‘eat’ which is a reflex of an original consonantal verb pTg **jep-* ‘eat’, for instance, has an irregular perfect participle *je-ke* just like Ma. *hafu*- ‘penetrate, go through’, which comes from pTg **xapun-* ‘penetrate’, has an irregular form *hafu-ka*. The corresponding resultative nominalizer Mongolian **-gA* devoices to **-kA* after *..r-* and *..m-*, for instance in WMo. *butarqai* ‘dispersed (adj. and n.)’ from WMo. *butara*- ‘break to pieces (intr.)’ and in WMo. *idemkei* ‘voracious’ from pMo **ide-me-* ‘want to eat’ in WMo. *idemer* ‘edible’, *idemeg* ‘having a good appetite, greedy’. On the basis of Old Turkic sources that permit the distinction between voiced and voiceless velar stops and reflexes in contemporary Turkic languages, a voiceless allomorph pTk **-kA* can be reconstructed in some relic forms in Turkic, such as OTk. *tarka* ‘alone, lonely’ from OTk. *tar*- ‘disperse, send away (tr.)’ and OTk. *öpkä* ‘generated in the lung; lung, anger’ from OTk. *öp*- ‘kiss, sip or suck in the air or a liquid’. The devoicing of the suffix pTk **-gA* seems to be determined by a preceding stem-final *..r-* or *..p-*. Due to lack of original voice distinction in Japanese and Korean, the allomorphs have merged into a single resultative morpheme pJ **-ka* and pK **-kA*-. Given the tendency towards reduction of allomorphy in contact situations, the shared allomorphy across the continental Transeurasian languages is indicative of inheritance.

4.2.5 Multiple setting

Fourteen etymologies, namely (2), (3), (7) to (17) and (21), have members in each of the five individual branches of the Transeurasian family. In a contact scenario, the morphemes would have crossed four linguistic boundaries. The likelihood, that a single verb morpheme progresses from one contact language into the other, repeatedly for four times, is very low to begin with. The probability, however, that fourteen verb morphemes follow the same pathway is close to zero. The observation that the corresponding morphemes are simultaneously attested in five branches strongly favors an internal explanation for the similarities observed.

4.2.6 Gaps

Finally, the distribution of the gaps in certain etymologies may be relevant as well. Three etymologies, namely (1), (6) and (20) lack a Korean member. From the viewpoint of the geographical isolation and the cultural history of Japan, it is difficult to explain the context in which three verb morphemes were borrowed from Tungusic into Japanese without a Korean intermediary. Inherited items, on the other hand, are expected to leave traces in remote areas, even when they do not occur in adjacent regions. The distribution of the gaps in the etymologies can thus be taken as an indication of genealogical relatedness.

5 Conclusion

Returning to the objections against the genealogical relatedness of the Transeurasian languages, this chapter has attempted to show that, first, the languages under consideration have a significant amount of bound verb morphology in common and, second, that it is safer to attribute these similarities to inheritance than to code-copying. The overview of shared verb morphology in the Transeurasian languages in Section 3 proposes etymologies for twenty one different verb suffixes relating Japanese to the Transeurasian languages. When taking into consideration that an inventory of verb morphemes in any individual language will not easily exceed one hundred items, we can say that these languages share a relatively large proportion of their verb morphology.

Second, for the common verb morphemes of the Transeurasian languages, none of the indications of codecopying (2.1 to 2.6) are fulfilled, whereas the evidence displays a number of characteristics (2.7 to 2.12) suggesting that the languages under inspection are genealogically related. This is summarized in Table 12 below. It is therefore safer to identify the shared verb morphemes in the Transeurasian languages as cognates than it is to consider them as copies.

Table 12: Distinction between copied and cognate morphemes applied to the Transeurasian languages

	Copy	in TEA		Cognate	TEA etymologies
2.1	to shared roots only	No	2.7	globally shared infrequent grammaticalization	(7), (10), (11) to (15)
2.2	suffix strings	No	2.8	categorial opacity	(7), (10), (11) to (17)
2.3	secondary semantics only	No	2.9	cumulative function	(14), (15)
2.4	contradiction sound correspondences	No	2.10	complete variant allomorphy	(14), (17?)
2.5	limited to contact zones	No	2.11	multiple comparison	(2), (3), (5), (7) to (17), (21)
2.6	specific morphosyntactic subsystems affected	No	2.12	Gaps	(1), (6), (20)

Abbreviations

Linguistic forms

A	deadjectival
ASS	assertive
CONN	connective vowel
N	denominal
O	de-onomatopoetic
PCP	participle
PERF	perfect
PERS	person agreement
PF	perfective
PREP	preposition
PRS	present
NML	nominalizer
RES	resultative
V	deverbal

Languages

EOJ	Eastern Old Japanese
Evk.	Evenki
K	Korean
Ma.	Manchu
MK	Middle-Korean
Na.	Nanai
Neg.	Negidal
OJ	Old Japanese
OTk.	Old Turkic
pJ	Proto-Japonic
pK	Proto-Koreanic
pMo	Proto-Mongolic
pTEA	Proto-Transeurasian
pTg	Proto-Tungusic
PTk	Proto-Turkic
Ud.	Udeghe
WMo.	Written Mongolian

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.
- Comrie, Bernard. 2008. "Inflectional Morphology and Language Contact, with Special Reference to Mixed Languages." In *Language Contact and Contact Languages*, edited by Peter Siemund and Noemi Kintana, 15-32. Amsterdam: 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*, edited by Lars Johanson and Martine Robbeets, 21-31. Wiesbaden: Harrassowitz.
- Erdal, Marcel. 2004. *A Grammar of Old Turkic*. Leiden: Brill.
- Friedman, Victor. 2009. "Turkish Presents in Romani Dialects." In *Turcological Letters to Bernt Brendemoen*, edited by Éva Csató, Gunvald Ims, Joakim Parslow, Finn Thiesen, and Emel Türker, 109-121. Oslo: Novus.
- Gardani, Francesco. 2008. *Borrowing of Inflectional Morphemes in Language Contact*. Frankfurt: Peter Lang.
- Gorelova, Liliya. 2002. *Manchu Grammar*. Leiden: Brill.
- Haspelmath, Martin. 1990. "The Grammaticization of Passive Morphology." *Studies in Language* 14:25-71.
- Heath, Jeffrey. 1978. *Linguistic Diffusion in Arnhem Land*. Canberra: AIAS.
- Heine, Bernd, and Tania Kuteva. 2005. *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.
- Igla, Birgit. 1996. *Das Romani von Ajia Varvara. Deskriptive und historisch-vergleichende Darstellung eines Zigeunerndialekts*. Wiesbaden: Harrassowitz.
- Jendraschek, Gerd. 2007. "Basque in Contact with Romance Languages." In *Grammars in Contact. A Cross-Linguistic Typology*, edited by Alexandra Aikhenvald and R. M. W. Dixon, 143-162. Oxford: Oxford University Press.
- Johanson, Lars. 2002. *Structural Factors in Turkic Language Contacts*. Richmond: Curzon.
- King, Ruth. 1999. *The Lexical Basis of Grammatical Borrowing: a Prince Edward Island French Case Study*. Amsterdam: John Benjamins.
- Klausenburger, Jürgen. 2000. *Grammaticalization: Studies in Latin and Romance*

- morphosyntax*. (Amsterdam Studies in the Theory and History of Linguistic Science 193.) Amsterdam: John Benjamins.
- Kossmann, Martin. 2010. "Parallel System Borrowing. Parallel Morphological Systems due to the Borrowing of Paradigms." *Diachronica* 27:459-487.
- Malchukov, Andrej. 2003. "Russian Interference in Tungusic languages in an Areal-typological Perspective." In *Studies in Eurolinguistics. Vol 1. Convergence and Divergence of European Languages*, edited by Sture Ureland, 235-251. Berlin: Logos.
- Malchukov, Andrej. 2006. "Yakut Interference in North-Tungusic languages." In *Turkic Languages in Contact*, edited by Hendrik Boeschoten and Lars Johanson, 122-138. Wiesbaden: Harrassowitz.
- Martin, Samuel Elmo. 1992. *A Reference Grammar of Korean*. Tokyo: Tuttle.
- 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.
- Pinkster, Harm 1987. "The strategy and chronology of the development of future and perfect tense auxiliaries in Latin." In *Historical development of auxiliaries*, edited by Harris, Martin and Ramat, Paolo, 193-223. Amsterdam: Mouton de Gruyter.
- Robbeets, Martine. 2005. *Is Japanese Related to Korean, Tungusic, Mongolic and Turkic?* Wiesbaden: Harrassowitz.
- Robbeets, Martine. 2007a. "How the Actional Suffix Chain Connects Japanese to Altaic." *Turkic Languages* 11:3-58.
- Robbeets, Martine. 2007b. "The Causative-passive in the Trans-Eurasian Languages." *Turkic Languages* 11:235-278.
- Robbeets, Martine. 2009. "Insubordination in Altaic." *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*, edited by Lars Johanson and Martine Robbeets, 81-114. Wiesbaden: Harrassowitz.
- Seifart, Frank. 2010. "Interdependences between Copied Derivational and Inflectional Morphemes." Paper presented at the workshop "Bound Morphology in Common: Copy or Cognate?" at the 43th SLE Conference in Vilnius, September 1-5.

- Sekerina, Irina A. 1994. "Copper Island (Mednyj) Aleut (CIA): A Mixed Language." *Languages of the World* 8:14–31.
- Starostin, Sergej, Anna Dybo, and Oleg Mudrak. 2003. *Etymological Dictionary of the Altaic Languages*. Leiden: Brill.
- Thomason, Sarah Grey. 1983. "Genetic Relationship and the Case of Ma'a (Mbugu)." *Studies in African Linguistics* 14:195-231.
- Thomason, Sarah Grey. 1997. "Mednyi Aleut." In *Contact Languages. A Wider Perspective*, edited by Sarah Grey Thomason, 449-468. Amsterdam, Benjamins.
- Thomason, Sarah Grey, and Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.

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.