

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/309763268>

Robbeets, Martine 2017. Japanese, Korean and the Transeurasian languages. In: Hickey, Raymond (ed.) The Cambridge handbook...

Chapter · November 2016

CITATIONS

0

READS

26

1 author:



[Martine Robbeets](#)

Max Planck Institute for the Science of Human History

37 PUBLICATIONS 21 CITATIONS

[SEE PROFILE](#)

22

The Transeurasian Languages

Martine Robbeets

22.1 Introduction

The present contribution is concerned with the areal concentration of a number of linguistic features in the Transeurasian languages and its historical motivation. The label ‘Transeurasian’ was coined by Johanson and Robbeets (2010: 1–2) with reference to a large group of geographically adjacent languages, traditionally known as ‘Altaic’, that share a significant number of linguistic properties and include up to five different linguistic families: Japonic, Koreanic, Tungusic, Mongolic and Turkic. The question whether all similarities between the Transeurasian languages should be accounted for by language contact or whether some are the residue of a common ancestor is one of the most debated issues of historical comparative linguistics (see Robbeets 2005 for an overview of the debate). Since the term ‘linguistic area’ implies that the shared properties are the result of borrowing, I will refrain from *a priori* attaching it to the Transeurasian region and rely on the concept of ‘areality’ instead, that is, the geographical concentration of linguistic features, independent of how these features developed historically. Only after evaluating 27 structural features shared across the Transeurasian languages will I consider how the insights from the data are relevant to historical statements about the way the languages may have come to share these features, considering diffusion, genealogical relationship or an interaction of both factors as possible explanations.

In spite of the strong polarization in the Transeurasian field between so-called ‘retentionists’, who view the similarities as arising from common descent, and ‘diffusionists’, who view them as arising from areal interaction, detailed characterizations of Transeurasian as a linguistic area are surprisingly rare in the linguistic literature. Poppe (1964) analysed Altaic as a ‘language type’ on the basis of a list of structural parallels shared between Korean, Tungusic, Mongolic and Turkic languages, and Rickmeyer (1989) elaborated on this research, adding data from Japanese. Even if these

contributions provide an impressive list of shared features, they do not strictly identify Transeurasian as a language area because they do not (i) delimit the language type in relation to its neighbours, (ii) list deviations from the prototypical type in the peripheries, (iii) consider the extent to which the features in question are common or rare across the world as a whole, or (iv) attempt to distinguish contact-induced from genealogically motivated features.

In this chapter I attempt a partial answer to these concerns by providing a typological profile of selected Transeurasian languages, along with their oldest linguistically reliable historical varieties, and by comparing this profile with the behaviour of languages immediately outside the Transeurasian region. In order to examine external boundaries, I have included adjacent languages to the east (Ainu and Nivkh in the northeast and Rukai in the southeast), to the south (Mandarin Chinese) and to the north (Kolyma Yukaghir, Ket and Eastern Khanty).¹ These languages are taken as horizontal comparative points representative of surrounding areas such as the Siberian area (Nivkh, Kolyma Yukaghir, Ket, Eastern Khanty) or the Mainland Southeast Asia area (Mandarin) and neighbouring families such as Austronesian (Mantauran Rukai), Sino-Tibetan (Mandarin), Yukaghiric (Kolyma Yukaghir), Yeniseic (Ket), Uralic/Ob-Ugric (Eastern Khanty) or Ainuic (isolate Ainu). Although Eastern Khanty can be taken as a representative of the Uralic languages, the main boundary to the west, I have paid less attention to additional western boundaries, excluding sample languages from the Caucasus region or from the Indo-European languages because of the limited space available here.

The vertical comparison points in my analysis consist of a list of 27 features, denoted F1–F27, chosen to maximize positive (+) values for Transeurasian as opposed to neighbouring languages. Although all features reflect a certain internal coherence, about half of them (i.e. 13) display deviations from the prototypical type in the peripheries. Where possible, I add an estimation of the degree to which the feature under discussion is common or rare across the world's languages, relying on the counts in *The World Atlas of Language Structures* (Haspelmath et al. 2005) or on other typological research to be specified below.

Given the controversy between diffusionists and retentionists, we cannot simply amass a number of shared features among the Transeurasian languages and allow geographical adjacency to imply the probability of diffusion, without requiring any linguistic support for this. Therefore, historical evidence suggesting the diffusion or the retention of traits may be particularly telling in this particular case. For representatives of the contemporary varieties of the five families belonging to the Transeurasian continuum,

¹ The following sources were consulted for retrieving linguistic data underlying the feature values in neighbouring languages: Gruzdeva (1998) for Nivkh; Maslova (2003a) for Kolyma Yukaghir; Werner (1997), Vajda (2004) and Georg (2007) for Ket; Filchenko (2007) for Eastern Khanty; Li and Thompson (1989) for Mandarin; Zeitoun (2007) for Mantauran Rukai; Shibatani (1990) and Tamura (2000) for Ainu.

I have chosen Turkish (Turkic), Khalkha Mongolian (Mongolic), Evenki (Tungusic), Korean (Koreanic) and Japanese (Japonic) as horizontal comparison points.² However, in order to allow a diachronic perspective, their profile will be supplemented by values from the oldest linguistically reliable historical varieties of the individual families, namely Old Turkic (eighth to fourteenth centuries), Middle Mongolian (thirteenth to seventeenth centuries) and/or Written Mongolian, Manchu (seventeenth to nineteenth centuries), Middle Korean (fifteenth to sixteenth centuries) and Old Japanese (eighth century).³ If a diachronic variety does not openly or productively reflect a certain feature but nevertheless preserves a trace of it, indicating that the value was positive in an earlier stage of the language, the historical variety will be marked with a plus. In this way, we can obtain a glimpse of the unrecorded typological past of the language in question.

The organization of this chapter is as follows. In Section 22.2 I will set up a typological profile of the Transeurasian languages in relation to that of the selected languages immediately outside the continuum. The linguistic levels discussed will include phonology, lexicon and semantics, morphology and syntax. I intend to treat grammaticalization patterns as a distinct level of analysis because, rather than representing a static feature value, they are concerned with a dynamic force, leading languages to change from a less to a more grammatical status. I will close Section 22.2 with a tabular overview, summarizing the presence of the 27 examined features in the selected languages by way of plus (+) and minus (–) values. In Section 22.3, I will consider how the insights from these data are relevant for general statements about areality, paying attention to the delimitation of areality, peripheral deviations from the prototype, changes in areality and the distinction between diffused and inherited features. In Section 22.4, I will conclude the chapter.

22.2 Typological Profile of the Transeurasian Languages

22.2.1 Phonology

F1 Predominantly polysyllabic root structure The Transeurasian languages, together with their historical varieties, display a preponderance of polysyllabic roots, as do most languages in North Asia. Contemporary and Old Japanese possess a relatively great number of monosyllabic roots, many of which are attributed to root-internal consonant loss and subsequent

² The following sources were consulted for retrieving linguistic data underlying the feature values in contemporary Transeurasian languages: Göksel and Kerslake (2005) for Turkish; Janhunen (2012) for Khalkha Mongolian; Bulatova and Grenoble (1999) and Nedjalkov (1997) for Evenki; Martin (1992) and Sohn (1994) for Korean; Martin (1988) and Kaiser et al. (2001) for Japanese.

³ The following sources were consulted for retrieving linguistic data underlying the feature values in historical Transeurasian languages: Erdal (2004) for Old Turkic; Street (1957), Weiers (1966) and Rybatzki (2003) for Middle Mongolian; Poppe (1954) for Written Mongolian; Gorelova (2002) for Manchu; Martin (1992) and Lee and Ramsey (2011) for Middle Korean; Vovin (2005, 2009) and Frellesvig (2010) for Old Japanese.

vowel contraction ([Whitman 1990](#)). These phonological reductions argue against Janhunen's (1997) suggestion that Japanese derives from an originally monosyllabic language. Austronesian languages such as Rukai are typically polysyllabic as well. Mandarin is the only language in the tables that is marked with a negative value. Similar to the languages of Mainland Southeast Asia it is predominantly monosyllabic, but, in comparison to Classical Chinese, it has developed a greater number of polysyllabic roots through compounding (Norman 1988: 86). As such, Japanese and Chinese occupy an intermediate position between the languages of North and Southeast Asia.

F2 Absence of complex tonal distinctions None of the Transeurasian languages is tonal in the sense that each syllable is characterized by a distinctive pitch pattern. This is also true for Austronesian languages, such as Rukai. With the exception of Ket, which has been attributed a tone system with five oppositions in recent descriptions by Vajda (2004) and Georg (2007), the neighbouring languages of North Asia are typically non-tonal as well. However, Nivkh, Japanese and some varieties of Korean have suprasegmental systems which can be seen as transitional between tonal and non-tonal languages. Nivkh makes distinctive use of two types of tones, whereas Middle Japanese and Middle Korean use a system of pitch-accent that differentiates words according to the position of one prominent syllable after which the pitch drops. This system survives in Contemporary Japanese, but it has been lost in Contemporary Standard Korean, where it developed into a vowel length distinction. The two-way tone distinction and the pitch-accent system are highly restricted in comparison to complex tonal systems such as in Ket and Mandarin, where each syllable is marked with one out of five distinctive tones. Tonal languages are not only extremely widespread throughout Southeast Asia, but also 42 per cent of languages in Maddieson's (2005: 58–61) sample of 526 languages across the world are tonal.

F3 Presence of vowel harmony Vowel harmony can be defined as a phenomenon whereby vowels within a domain agree with each other in terms of one or more features (Ko 2012: 7). It is a characteristic feature of the Transeurasian languages, except Japanese, but it is also present in most Uralic languages, including Khanty, and in many other languages in North Asia such as in Yukaghir, Nivkh and Ainu. Ket lacks vowel harmony and so do Rukai and Mandarin, as such reflecting prototypical Austronesian and Mainland Southeast Asian behaviour, respectively. In Old Japanese, however, there is a restriction on the shape of root morphemes, whereby the vowel o_2 cannot occur in a root together with the vowels u , o_1 or a . This phenomenon, known as Arisaka's law, has been taken as a kind of vowel harmony, but it has been rejected from comparisons with other Transeurasian languages because it applies to roots rather than to suffixes and because it does not reflect palatal harmony, the type of harmony which was attributed to the Transeurasian languages until recently (e.g. Frellesvig 2010: 44). However, in lexicalized verb stems incorporating derivational suffixes, as

well as in the noun inflectional suffixes such as the plural suffix and the genitive suffix, there are traces of $a \sim o_2$ vowel alternation according to the quality of the vowels in the preceding root, e.g. OJ *no_2* genitive versus OJ *-na-* petrified in compounds such as OJ *mi_1-na-moto* [< water-GEN-base] ‘source, the headwaters’ (Rickmeyer 1989: 316; Robbeets 2015).

F4 Presence of tongue root vowel harmony Among the various types of vowel harmony, the most frequently attested ones across the languages of the world are palatal harmony, labial harmony, height harmony and tongue-root harmony. Palatal harmony requires all vowels within a domain to be exclusively front or back. It can be found in most Uralic languages such as in Khanty as well as in the Turkic languages (e.g. Tk. *ip-ler* [rope-PL] ‘ropes’ versus *pul-lar* [stamp-PL] ‘stamps’). Since the western Mongolic languages Oirat and Kalmuck display palatal harmony as well, it has been proposed that the original system of Mongolic harmony was palatal (Poppe 1955; Svantesson 1985). However, Ko (2012) demonstrated that the original vowel harmony in Mongolic was in fact based on opposition between the advanced versus retracted position of the tongue root, rather than on a palatal contrast. He argued that the tongue root retraction system in Khalkha (e.g. *od-o:s* [feather-ABL] versus *ɔd-ɔ:s* [star-ABL]) represents retention rather than innovation. Furthermore, he supported the view that Tungusic vowel harmony is retracted tongue root (RTR) based, as it is in Manchu and Evenki, and that the reduced vowel harmony in Contemporary Korean derives from a tongue root based system in Middle Korean. As far as the harmony-like opposition between o_2 and u , o_1 or a in Old Japanese is concerned, the recent reconstruction of a seven-vowel system in Proto-Japonic by Frellesvig and Whitman (2008) implies an underlying opposition between pJ $*i$, $*ə$ and $*u$, $*o$, $*a$, which does not exclude an original RTR-based contrast. Whereas Vovin (1993: 50–51) and Bugaeva (2015: 26–28) reconstruct palatal harmony in Ainu, Shibatani (1990: 15) speculates that the Ainu opposition between o and u , a might have its origin in tongue root harmony, but here the indications are even weaker than in the Japanese case. According to Maslova (2003a: 35), Yukaghir might be more appropriately described as having tongue root harmony than palatal harmony. Chukchi also displays tongue root harmony. Although Gruzdeva (1998: 10) suggests that Nivkh leaves traces of height harmony, Janhunen (1981) and Ko, Whitman and Joseph (2014) interpret the system in terms of tongue root harmony. Cross-linguistically, tongue root harmony seems to be concentrated in Niger-Congo and Nilo-Saharan languages. Outside Africa and the Northeast Asian region the phenomenon seems to be rather rare: only Native American languages such as Nez Perce and Coeur d’Alene Salishan are known to have the feature (Ko 2012: 11–12). A rough estimate would be that less than 10 per cent of the world’s languages have a tongue root vowel harmony system.

F5 Absence of initial velar nasal In most Turkic languages as well as in Mongolic languages and Korean, the velar nasal η - cannot appear in word-initial position. Japanese lacks a velar nasal phoneme. In the Tungusic languages, with the exception of Manchu, however, η - can appear word-initially, but generally restricted to a specific phonological environment, notably when it is followed by the sonorants n, r, l, m, y , e.g. Evk. *ηene*- ‘to go’, Ma. *genu*- ‘to go together’, Evk. *ηe:le*-, Ma. *gele*- ‘to fear’, etc. According to Poppe (1964: 4) the initial velar nasal in Tungusic is the result of secondary assimilation of pTg. *g-, which implies that originally * η - was absent in Tungusic as well. The assimilation was probably triggered by influence from languages in the Siberian area, such as Nivkh, which allow initial velar nasals (Anderson 2006). It is under the same influence that initial η became allowed in Dolgan (Turkic), e.g. *ηassa* ‘pipe’. In Khanty, Ket, Kolyma Yukaghir, Ainu and Mandarin, η - does not occur in word-initial position. Rukai allows an initial velar nasal, e.g. *ηa/ai* ‘saliva’. In Anderson’s (2005: 42) sample of 468 languages, 69 per cent lack an initial velar nasal. Among the languages of the world that have a velar nasal phoneme, as is the case for most Transeurasian languages, only 35 per cent do not use it in word-initial position.

F6 Absence of initial r- Throughout the Transeurasian languages, the consonant r - is not allowed to occur word-initially, except in borrowings (e.g. J *rajio*, K *latiwo*, Even *radio*, Khal. *radio*, Tk. *radio* ‘radio’). This is also true for Kolyma Yukaghir. Ket lacks a phoneme $/r/$ altogether. Although initial * r - is not reconstructed for Proto-Uralic, Khanty is atypical in this sense, e.g. *rayta* ‘to drop, slide’ and *räy* ‘garbage’. Nivkh, Ainu, Mandarin and Rukai also have native words in initial r -.

F7 Absence of initial consonant clusters None of the Transeurasian languages tolerate initial consonant clusters, although medial clusters are tolerated in Turkic, Mongolic, Tungusic and Korean but, on the face of it, not in Japanese. On the basis of morphological, etymological, dialectal and textual evidence, however, it is safe to assume that the Old Japanese obstruents OJ b, d, g, z resulted from the rephonologization of nasal obstruent clusters pJ * np , pJ * nt , pJ * nk , pJ * ns (Robbeets 2008). Reminiscent of how the Transeurasian languages do not allow for consonant clusters in initial position, Old Japanese did not permit word-initial voiced obstruents except in mimetic adverbs. From the ninth century onwards, as loans from Chinese began to have a major impact, the restriction was relaxed and initial voiced obstruents began to appear in borrowings and in contracted native forms. The avoidance of consonant clusters is further characteristic of Uralic languages, such as Khanty. Similarly, Yukaghir, Ket, Ainu and Rukai tolerate only single consonants in word-initial position. Word-initial clusters may comprise at most two consonants in Nivkh, e.g. *mra* ‘fault’ and *ksynz* ‘witch’. Although Mandarin lacks consonant clusters, there is strong evidence that in Old Chinese (first millennium BCE) a variety of consonant clusters could occur at the beginning of the syllable as well (Norman 1988:

9–10). The simplification and eventual loss of consonant clusters appears to be a tendency affecting most of the Mainland Southeast Asia area. It is possible that early contacts between Chinese and Transeurasian, that has never tolerated initial clusters, have triggered the development along these lines.

F8 Presence of voice distinction for stops Turkic, Mongolic and Tungusic languages share a voiced–voiceless opposition for stops. In Contemporary and Middle Korean, stops display an opposition between lax (p), aspirated (ph) and tensed (p'). Even if the lax stops become lightly voiced between voiced sounds, there is no phonemic voicing distinction. The Japanese voicing distinction for stops is a secondary development. As mentioned in the previous paragraph, voiced stops derive from prenasalized voiceless stops, so originally Japanese lacked a voicing distinction. Khanty lacks a voicing distinction for stops, a feature characteristic of Proto-Uralic, although many contemporary Uralic languages have developed an original singleton–geminate contrast into a voicing distinction. For example, although the contrast between /p/ and /pp/ in Proto-Uralic **lapa* 'flat surface, leaf' and **tappa*- 'to stamp with feet, to hit, knock' is maintained in the Finnish reflexes *lapa* 'shoulderblade, leaf surface' and *tappa*- 'to beat to death, kill', it has usually developed into a distinction between /b/ and /p/ such as in Estonian *laba* 'surface' and *tapa*- 'kill, slaughter'. Ket and Yukaghir display a voicing distinction, but languages to the extreme northeast such as Ainu, Nivkh and Chukchi do not. Mandarin, like Nivkh, has a distinction between aspirated and unaspirated stops, but lacks a voiced–voiceless opposition. Characteristic of most Austronesian languages, Rukai also displays voice distinction for stops.

22.2.2 Lexicon and Semantics

F9 Preference for a non-verbal strategy with (extra-family) verbal borrowing As far as the mechanisms of loan verb accommodation are concerned, most recipient languages can be categorized into two distinct groups: borrowed verbs either arrive as verbs, needing no formal accommodation, or they arrive as non-verbs and need formal accommodation. In Wohlgemuth's (2009) terminology, the first group represents *direct insertion*, while the second group represents either *indirect insertion*, when the formal accommodation involves a verbalizer, or *light verb strategy*, when the borrowed verb is integrated into a complex predicate. Turkic, Mongolic, Korean and Japanese can be assigned to the second group because they display a clear preference for the non-verbal strategy (Wohlgemuth 2009: 159, 161); for instance, Tk. *klik-le*- and *klik et*- << English *click*; Khal. *zee-l*- << Mandarin *zhài* 'borrow, lend'; K *coking ha*-, J *zyogingu suru* 'to jog' << English *jog*; J *demo-r*- << English *demonstrate*. Whereas the Northern Tungusic languages prefer to borrow verbs through direct insertion, e.g. Evk. *vypolñaj*- << Russian *vypolnja-t'* 'to fulfil, carry out', the Southern Tungusic languages use

verbalizers, e.g. Ud. *tancewa-la-* << Russian *tancewa-t'* 'to dance' and Na. *voprosa-la-* << Russian *voproša-t'* 'to enquire, question'. In contrast to the Transeurasian languages, Uralic languages such as Khanty and Austronesian languages such as Rukai, Ainu and Mandarin show a strong preference for direct insertion (Tamura 2000: 267; Wohlgemuth 2009: 158, 161). Yukaghir and Nivkh did not integrate any recognizable verbal borrowings from Russian or from other foreign languages into their lexicons (Fubito Endo, personal communication; Ekaterina Gruzdeva, personal communication). In Wohlgemuth's (2009: 157) sample, 55 per cent of languages worldwide are found to use direct insertion, while the remainder prefer non-verbal strategies such as indirect insertion and the light verb strategy.

F10 Presence of a two-way proximal/distal distinction in demonstrative pronouns Although Old Turkic displays a two-way distinction in its demonstratives, i.e. OT *bo* / *bun-* 'this' versus *ol* / *an-* 'that', many contemporary Turkic languages such as Turkish make a three-way distinction, e.g. Tk. *bu* 'this', *şu* 'that', *o* 'that (over there)'. Demonstrative pronouns in earlier and contemporary varieties of Mongolic and Tungusic exhibit a proximal/distal distinction: MMo. *ene* 'this' versus *tere* 'that', Khal. *e-* 'this' versus *te-* 'that', Ma. *ere* 'this' versus *tere* 'that' and Evk. *er(i)* 'this' versus *tar(i)* 'that'. Demonstrative pronouns in Contemporary and Middle Korean, however, show a proximal/mesial/distal opposition: K *i* 'this', *ku* 'that', *ce* 'that over there' and MK *i* 'this', *ku* 'that', *tye* 'that over there'. This is also true for Contemporary Japanese: J *ko-* 'this', *so-* 'that', *a-* 'that over there'. In contrast to most accounts of Old Japanese demonstratives, which posit a three-way contrast between OJ *ko*₂ 'this', *so*₂ 'that' and *ka* 'that over there', Frellesvig (2010: 139–142) argued that OJ *ka* was not a productive member of the demonstrative system and that pre-Old Japanese had a simple proximal/distal distinction. While Khanty distinguishes between proximal *timi* 'this (here)' and distal *tomi* 'that (there)', Yukaghir, Ket and Ainu have a three-way opposition, each demonstrative pronoun denoting a different degree of proximity: Yukaghir *tiŋ* 'this' (proximal), *adiŋ* ~ *ediŋ* 'that' (mesial), *taŋ* 'that' (distal); Ket *tu-* 'this, that' (neutral), *ki-* 'this, that' (proximal); *qa-* 'this, that' (distal) and Ainu *ta an* 'this' (distal), *ne an* 'that' (mesial), *to an okai* 'that over there' (distal). Nivkh makes as many as five distinctions: *tyd'* 'this' (near and visible), *hyd'* 'this, that' (distant), *ad'* 'that' (more distant and visible), *aixnt* 'that' (most distant), *kud'* 'that' (absent).⁴ Rukai distinguishes four demonstrative pronouns in terms of visibility and distance: *'ina* 'this' (proximal), *ana* 'that' (mesial), *ona* 'that over there' (distal but visible), *dhona* 'that over there' (distal and invisible). Mandarin has a two-way distinction between proximal *zhè(ge)* 'this' and distal *nà(ge)* 'that', which developed from a three-way distinction in Classical Chinese between neutral, proximal and distal. In Diessel's (2005: 170–173) sample

⁴ Note that this analysis deviates from the feature values given for distance contrasts in demonstratives by Diessel (2005: 170–173), since he marks Ainu, Nivkh, Yukaghir and Turkish as having a two-way contrast.

of 234 languages, 54 per cent exhibit a two-way distance contrast in demonstratives, while 38 per cent exhibit a three-way contrast.

F11 Inclusive/exclusive distinction in first person plural pronouns Among the Turkic languages, there are no unique pronominal forms that distinguish inclusive from exclusive person forms. Although Old Turkic and most presently spoken varieties of Turkic distinguish between a first person plural (Tk./OT *biz* ‘we’) and an augmented plural form (Tk. / OT *biz-ler* ‘we (as a group)’), Nevskaya (2010: 124) argues for a collective interpretation of the augmented plural, denoting ‘an isolated group of people who want to oppose themselves to the others’, rather than an inclusive interpretation as suggested by Grönbech (1936: 81). The Middle Mongolian distinction between exclusive *ba* and inclusive *bida* is formally preserved in the Khalkha oblique paradigm in the variation between formally exclusive *man-* and formally inclusive *bidn-*, but the functional distinction has been lost. In the Tungusic languages, however, the inclusive/exclusive opposition is generally well preserved, e.g. exclusive Ma. *be*, Evk. *bu* versus inclusive Ma. *muse*, Evk. *mut* ~ *mit*. Similar to the Turkic languages, Middle and Contemporary Korean distinguish between a first person plural (K/MK *wuli* ‘we’) and an augmented plural form (K *wulitul*, MK *wuli-tolh* ‘we (as a group)’ in which K *tul*, MK *-tolh* is a collective marker. Contemporary Japanese lacks an inclusive/exclusive distinction and cannot derive an augmented plural from the first person plural *watasi-tati* [I-PL] ‘we’. Old Japanese also lacks the distinction, but the stem OJ *wa-* ‘I, we’ can be used as a first person plural in the possessive case form, but it can also be augmented with a collective marker *-ra* ~ *-re* to OJ *ware* ‘we’, a form which in its turn has been augmented into *ware-ra* ‘we’ later in Japanese. As is the case for many Uralic languages, Khanty marks a dual distinction but not an inclusive/exclusive distinction on its personal pronouns. While Ket and Yukaghir lack the distinction, Nivkh distinguishes between exclusive *n’yŋ* and inclusive *mer* ~ *mir*. Although Ainu personal affixes on the verb have an inclusive/exclusive distinction, the first personal pronoun *aoka(i)* only has a single form.⁵ The distinction found in the first person plural pronouns between exclusive *wōmen* and inclusive *zānmen* ‘we’ of Beijing and certain other northern Chinese dialects may be due to Transeurasian influence. Such a distinction was not found in Old Chinese, and it began to appear in North China during the period of Altaic rule. It is significant in this regard that both Middle Mongolian, spoken under the Yuan dynasty, and Manchu distinguish exclusive and inclusive forms. Rukai distinguishes exclusive *-nai* ~ *nai*- [NOM] from inclusive *-mita* ~ *ta*- [NOM], a feature characteristic of Austronesian languages. In Cysouw’s (2005: 166–167) sample of 200 languages, 31 per cent distinguish between inclusive and exclusive with independent pronouns.

⁵ Note that my evaluation differs from Cysouw’s (2005: 166–167) analysis, which marks Ainu as having an inclusive/exclusive distinction with independent pronouns.

F12 Property words may be verbally or nominally encoded Cross-linguistically adjectives have no prototypical encoding strategy of their own: they will align themselves either with verbs or with nominals (Stassen 1997: 30). Across the Transeurasian languages, the encoding of property words appears to be mixed because, at least in the earlier stages, both the nominal and the verbal strategy is used (Robbeets 2015). Generally, this mixed encoding is split in the sense that most property words have only a single encoding option, with the exception of some instances of switched encoding mentioned under feature value 13. In Old Turkic, most property words are nominally encoded, but there seems to be a tendency to apply the verbal strategy to words expressing time-unstable properties such as OTk. *bädü-* ‘to be(come) big, great’, OTk. *isi-* ‘to be hot’, OTk. *kat-* ‘to be hard, firm, tough’, OTk. *kiz-* ‘to be red’, OTk. *tumli-* ‘to be cold’, OTk. *us-* ‘to be thirsty’, OTk. *tigra-* ‘to be tough’, etc. Contemporary Turkic languages maintain only few reflexes of these verbal property words, e.g. Tk. *büyü-* ‘to be(come) large’, but in the majority of cases, the earlier verbal property word has been derived through a deverbal noun suffix into a nominal adjective (e.g. Tk. *büyük* ‘big’). Similarly, most property words are nominally encoded in Mongolic, but there is a tendency to apply the verbal strategy to less permanent properties in Middle Mongolian such as MMo. *ayu-* ‘to be(come) afraid’, MMo. *čat-* ‘to be ripe, be(come) saturated’, MMo. *hiče-* ‘to be ashamed’, WMo. *qala-* ‘to be(come) warm’, MMo. *sohta-* ‘to be drunk’, etc. Contemporary Mongolic languages such as Khalkha maintain only a few reflexes of these verbal property words, e.g. Khal. *ayu-* ‘be afraid’. The same is true for Tungusic, where contemporary languages such as Manchu (e.g. Ma. *aka-* ‘to be sad’, Ma. *bere-* ‘to be lame’, Ma. *ebi-* ‘to be satiated’) and Evenki (e.g. Evk. *ukti-* ‘to be hungry’, Evk. *uwi-* ‘to be satiated’, Evk. *buli-* ‘to be sad’) may occasionally exhibit verbal encoding.⁶ In Korean, there are property words such as K *kanan ha-* ‘to be poor’ and *phikon ha-* ‘to be tired’ that consist of a nominal root and the auxiliary *ha-* ‘to be in the state of’ and whose bases are called ‘adjectival nouns’ (Martin 1992: 189, 190; Sohn 1994: 219–220). However, the majority of property words are inflected in essentially the same way as verbs, e.g. K *kwut-*, MK *kwut-* ‘to be(come) hard’, K *noph-*, MK *nwoph-* ‘to be high’, etc. Some Japanese property words, such as J *sizuka*, OJ *siduka* ‘quiet’, J/OJ *tasika* ‘trustworthy’ are encoded exclusively nominally, while others such as J/OJ *taka-* ‘to be high’, J/OJ *kata-* ‘to be hard, tough’ are essentially inflected in a similar way to verbs. In line with most Uralic languages, property words in Khanty are exclusively nominally encoded. This is also true for Ket. In Yukaghir, Ainu and Nivkh, however, property words are exclusively verbally encoded. As in the case with most Transeurasian languages, Ainu property verbs express both the property and the process leading to the property, e.g. *pirka* ‘to be(come) good’. In line

⁶ Note that my analysis deviates from the feature values inserted for predicative adjectives by Stassen (2005b: 480–481), in which Evenki and Manchu are marked as exclusively verbal encoding, in line with the traditional view.

with Mainland Southeast Asian and Austronesian languages, Mandarin and Rukai use verbal encodings for property words. In Stassen's (2005b: 478–481) sample of 386 languages, 27 per cent have mixed encoding in predicative adjectives.

F13 Some property words exhibit switched encoding Some property words in the Transeurasian languages, especially in the earlier varieties, further exhibit traces of switching, whereby the same property word can have both nominal and verbal encoding: e.g. OT *ač* 'hungry' / *ač*- 'to be hungry', OT *keč* 'late, slow' / *keč*- 'to be late, slow'; MMo. *bulqa* 'hostile; hostility' / *bulqa*- 'to be hostile'; Ma. *jalu* 'full' / *jalu*- 'to be full', Ma. *sula* 'loose, free' / *sula*- 'to be loose, be free'; MK *toso*- versus MK *toso ho*- 'to be warm'; OJ *taka* 'high' / *taka*- 'to be high', OJ *opo* 'big' / OJ *opo*- 'to be big', etc. None of the neighbouring languages, except Tundra Yukaghir, exhibits such behaviour. There, two property words, i.e. *juku* 'small' and *t'ama* 'big', occur as noun modifiers without overt adnominalizers, e.g. *t'ama-d'ohoje* (big-sword) 'sabre', in addition to having a verbal encoding, for example, in the deverbal inchoative *t'ama-mu*- (be.big-INCH) 'to grow, become big' (Maslova 2003b: 14). Logically, the proportion of languages exhibiting mixed and switched encoding will be lower than 27 per cent, i.e. the proportion of languages with mixed encoding in general.

F14 Partial emphatic reduplication of nominal property words Partial emphatic reduplication is a phenomenon whereby the first consonant (if present) and vowel of a nominal property word are repeated with the addition of another consonant to indicate the presence of the property to the utmost degree. Whaley and Li (2000) found that it is widespread in Turkic, Mongolic and Tungusic, e.g. Tk. *bem-beyaz* 'snow white', OT *kap-kara* 'quite black', Kal. *xob-xoldu*: 'frozen through', WMo. *ub-ulayan* 'completely red', Evk. *ab-aya* 'very good'. I have not been able to find examples in Manchu, but the phenomenon is present in Sibe, a presently spoken variety of Manchu, e.g. *fak-farxun* 'extremely dark'. In Tungusic, emphatic reduplication is restricted to Sibe, Kile-Nanai, Solon Evenki and Oroqen, i.e. the languages spoken on Chinese soil, which have been under strong influence from the Mongolic languages Khalkha and Dagur. On the basis of this distribution and because the greatest flexibility, in terms of both the number of reduplicated words and the type of concepts they denote, is found in Turkic, Whaley and Li (2000: 358) argued for a diffusion of the feature from Turkic to Mongolic to Tungusic. Japanese, Korean and the neighbouring languages under examination do not display partial emphatic reduplication. In Rukai, however, descriptive verbs are partially reduplicated in comparative constructions (see feature 23).

22.2.3 Morphology

F15 Morphology is agglutinative Agglutinative languages connect morphemes linearly such that there is a one-to-one relationship between

a morpheme and its meaning. The Transeurasian languages belong to a North Asian and European belt of agglutinative languages together with the Uralic languages, including Khanty and other languages of the Siberian area such as Ket, Yukaghir, Chukchi and Nivkh. Ainu has agglutinative morphology and so do the Austronesian languages, including Rukai. Chinese, the only analytic language under examination, has triggered a decrease of agglutinating features in Tungusic as one moves from north to east and further south. Manchu is the most analytic among the Transeurasian languages; it treats case forms, for instance, as particles rather than suffixes.

F16 Inflectional morphology is predominantly suffixing Across the strongly suffixing Transeurasian languages, prefixation is rare and restricted to derivational morphology, such as the partial emphatic reduplication (see feature 14 above) and some derivational prefixes in Korean (e.g. *K yel-* ‘young, new’ in *yel-cwungi* ‘a chick out of its shell’) and in Japanese (e.g. *J ma-* intensive in *ma-siro* ‘snow white’). As is the case for most Uralic languages, Khanty is strongly suffixing and so is Yukaghir. Nivkh is considered to be weakly suffixing. In Ket, nominal inflectional morphology is strongly suffixing, whereas verb inflection is predominantly prefixing. In Ainu and Rukai, inflection makes use of both prefixes and suffixes. Probably due to Transeurasian influence, Mandarin is hard to assign unequivocally to either the isolating or weakly suffixing type, but Sinitic varieties in general tend towards the isolating pole. In Dryer’s (2005a: 110–113) sample of 894 languages, 43 per cent are strongly suffixing.

F17 Absence of obligatory numeral classifiers Although in Turkic and Mongolic some nouns of low countability may be accompanied by a unit of measure by means of which they can be counted, e.g. Tk. *sekiz bardak su* [eight glass water] ‘eight glasses of water’, OT *yeti tutum talkan* [seven handful parched.grain] ‘seven handfuls of parched grain’, Khal. *gourben debter nom* [three volume book] ‘three volumes of books’, etc., these languages do not make obligatory use of sortal numeral classifiers. Similar to the use of collective suffixes for counting people in Old Turkic (OT -(A)gU in e.g. *üčägü* ‘three together’) and Middle Mongolian (MMo. -’ULA ~ AIA in e.g. *qoya’ula* ‘two together’), the Tungusic languages use a variety of collective suffixes following numerals from ‘two’ to ‘ten’ such as Evk. -*kt(e)* and -*ni* for counting people (e.g. *d’u-kte* ‘two (people together), (we, you, they) two’), Evk. -*gdA/ -ngnA* for counting objects, Evk. -*lla* for counting the number of days (e.g. *nada-lla* ‘seven days, a week’), Evk. -*nu / -pu* for counting the number of tents (e.g. *ilan-nu* ‘three tents’) and Evk. -*musa* denoting the number of places or directions. However, only Manchu has developed about 70 sortal numeral classifiers, which divide the inventory of count nouns into semantic classes, each of which is associated with a different classifier, such as *fesin* which is used for objects equipped with a handle, e.g. *ilan fesin loho* [three CLAS sword] ‘three swords’. These words have

original lexical meanings, e.g. *fesin* 'haft, shaft, handle', but under Chinese influence they have grammaticalized into classifiers, which are not obligatory in Manchu. *Loho ilan* [sword three] 'three swords', for instance, is equally possible. Whereas the standard pattern in Middle Korean was to modify a noun with a preposed numeral, e.g. *twu kalh* [two knife] 'two knives', the most common pattern in Contemporary Korean makes use of a classifier, e.g. *pus se:k calwu* [writing.brush three CLAS] in which *calwu* denotes long objects with handles. However, the original pattern surfaces in expressions such as *K twu nala* 'two countries' and the use of classifiers remains optional in Korean, e.g. *kalh hana-ka issta* [knife one-NOM be.pre-sent] 'there is one knife'.⁷ With Chinese influence inundating the language from Middle Korean times onwards, the classifiers developed from native words under Chinese influence or were borrowed as such from Chinese. Note that Middle Korean leaves traces of specialized suffixes to count days, e.g. **(o/u)l* in *saol* 'three days', *naol* 'four days' etc. and that some Korean dialects use a suffix *-i* to count persons, e.g. *se:-i* 'three people', *ne:-i* 'four people', which recalls the use of collective suffixes in the other Transeurasian languages. While there is an extensive list of obligatory classifiers in Contemporary Japanese, e.g. *enpitu san-bon* [pencil three-CLASS] 'three pencils', the system of classifiers is much less developed in Old Japanese, where Chinese influence is restricted to a minimum. Numerals could be used with nouns, without intervening classifiers, e.g. *OJ nana se* [seven rapid] 'seven rapids' and the so-called 'classifiers' are restricted to roughly six suffixes, i.e. *-ka* for counting days starting from the numeral 'two', *-tu* / *-ti* for counting objects, *-ri* for persons, *-mo₂to₂* for grassy plants, *-pe₁* for layers and *-ka* for plants. It is not unlikely that these suffixes originate in collective suffixes. Numeral classifiers are absent in Uralic languages such as Khanty, as well as in Yukaghir and Ket. Ainu makes use of a small set of obligatory classifiers such as *-n* / *-iw* for persons, *-pe* / *-p* for animals and things, and *rerko* for counting days starting from the numeral 'two' (with irregular forms *tutko* 'two days' and *rerko* 'three days'). Nivkh distinguishes between 26 semantic classes with different numeral forms for each class. The obligatory use of classifiers is a widespread feature shared by Mandarin and the languages of Southeast Asia, but the use of classifiers in Classical Chinese was the exception rather than the rule. In Rukai the use of classifiers is optional in the sense that it uses a set of unaffixed numerals without classifiers, as well as a set of bound numerals which combine with five different sortal classifiers to form verbs. In Gil's (2005: 226–229) sample of 400 languages, 80 per cent lack obligatory numeral classifiers.

F18 Presence of *mi-Ti* opposition in first versus second person singular pronouns Nichols (2012) observes that *mi-Ti* pronominal

⁷ Note that my evaluation differs from Gil's (2005: 228–229) interpretation that Korean has obligatory numeral classifiers.

paradigms with first person labial nasal *m* and second person apical or palatal obstruent *t*, *c*, *s*, etc. are much more common in northern Eurasia than elsewhere in the world. Janhunen (2013: 213) adds that there is a smaller group of *mi-Ti* languages extending from Uralic in the west, to Turkic, Mongolic and Tungusic in the east to Yukaghir in the north, in which not only the initial consonant but also the root vowel of the singular stems shows a basic similarity, in that it contains a non-low unrounded front vowel *i* or *e*. Although *m* is absent in the nominative first person singular in the Turkic, Mongolic and Tungusic languages, e.g. Tk. *ben*, OT *ben*, Khal. *bii*, MMo. *bi*, Ma. *bi*, Evk. *bi:*, it has developed in oblique forms through assimilation to the nasal oblique suffix *-n*, e.g. OT *min-*, Khal. *min-ii* (GEN), MMo. *mi-nu* (GEN), Ma. *min-*, Evk. *min-*. The second person singular forms all reflect a voiceless dental *T*, i.e. Tk. *sen*, OT *sen*, Khal. *cii*, MMo. *ci*, Ma. *si*, Evk. *si:*. The Korean pronouns are, among others, first singular K/MK *na* and second singular K/MK *ne*. In Japanese, *J watasi* and *OJ wa* are among others used in the first singular, while a variety of contemporary pronouns and *OJ na* are used in the second singular. Although the Proto-Uralic first and second singular pronouns **mun* and **tun* reflect a *mi-Ti* distinction (Janhunen 1982: 35), Khanty is deviant in having first singular *mă* and second singular *nõŋ*. In Yukaghir, however, the *mi-Ti* opposition is present in first singular *met* versus second singular *tet*. In Nivkh, the distinction is absent in the singular pronouns, first person *n'i* versus second person *či*, but it is present in the opposition between the first plural inclusive *mir/mer* and the second plural pronoun *čiŋ*. The opposition is not found in Ket, Ainu, Chinese and Rukai. In Nichols and Peterson's (2005: 546–551) sample of 230 languages, 13 per cent display a *mi-Ti* opposition in first versus second person pronouns. Logically, languages reflecting a *mi-Ti* opposition will represent an even smaller proportion.

F19 Formation of a secondary oblique stem of personal pronouns With the exception of Korean, the Transeurasian languages share a tendency to form a secondary oblique stem of the personal pronouns by means of a suffix which phonologically may be identified as the dental nasal *-n*. In most contemporary Turkic languages, the nominative and oblique forms have merged, e.g. Tk. *ben* for the first singular nominative and oblique, but in Old Turkic the first singular nominative *bān* is distinguished from the oblique stem *min-*, which can be derived from an original pTk. **bi-n-* [1SG-OBL]. Similarly, the Mongolic and Tungusic languages derive oblique pronominal stems from the nominative roots through a nasal suffix, for instance, in the first person plural pronouns MMo. *ba* [NOM] versus *man-* [OBL] and Khal. *bid* [NOM] versus *bidn-* [OBL] and in the first person singular pronouns Ma. *bi* [NOM] versus *min-* [OBL], Evk. *bi:* [NOM] versus *min-* [OBL]. There are no oblique pronominal stems in Contemporary Japanese, but Old Japanese leaves traces of an oblique nasal suffix in some case forms, for example in the eastern OJ first person

singular dative *wa-nu-ni* in alternation with western OJ *wa-ni*. Vovin (2005: 229–230) further found that an original Japonic pronominal oblique **-n-* is well supported by Northern Ryukyuan dialects, where the first person pronoun uses *waa-* as the nominative and genitive base and extended *waN-* in the oblique cases. The erosion of the pronominal paradigm in Korean and Japanese may be due to the gradual de-pronominalization in the recorded history of these languages, whereby the system of personal pronouns became replaced by various terms of address and self-reference, probably under Chinese influence. The oblique nasal suffix is an important element in the Uralic pronominal paradigm as well, for example the Khanty first person pronoun *mā* [NOM] versus *mān-* [OBL]. Ket, Yukaghir, Ainu and Mandarin, however, do not derive secondary oblique stems. The third person singular pronoun in Nivkh has both regular and suppletive case forms, e.g. *if-ən* [3SG-NOM] versus *if-toX ~ e-rx* [3SG-DAT/ADD], but here the oblique form is not derived from the nominative base. Rukai personal pronouns have different shapes for nominative, topic, genitive and oblique cases, e.g. the first person singular *-lrao* [NOM], *ilrae* [TOP], *-li* [GEN] versus *-iae* [OBL], in which the oblique seems to be formally derived from the nominative base by means of the same *i-...-e* marking as in the topic form.

22.2.4 Syntax

F20 SOV (Subject-Object-Verb) sentence order Syntactically, the Transeurasian languages pattern as typical SOV languages but the sentence order is not rigid. SOV is also among the characteristic features of the Uralic languages, here represented by Khanty. Languages to the north such as Yukaghir and Ket or to the northeast such as Ainu and Nivkh are almost all SOV languages, while those to the southeast are virtually all SVO languages. Mandarin, and in fact all major varieties of Chinese, corresponds with Southeast Asia with respect to verb-object order. Like most Austronesian languages, Rukai tends to be verb-initial, but the word order is non-rigid, switching freely between VSO and VOS. In Dryer's (2005b: 330–333) sample of 1228 languages, 40 per cent have SOV sentence order.

F21 GAN (genitive-noun / adjective-noun) phrase order A modifier-before-headword word order in the sentence (SOV) is expected to correlate with a modifier-head order within the noun phrase (GAN), whereby adjectives, genitives and modifiers in general occur before the nouns to which they refer. This is the case for the Transeurasian languages, the Uralic languages including Khanty, and other languages of North Asia such as Yukaghir, Ket, Ainu and Nivkh. Mandarin, however, runs against the implicational expectation, since genitives and adjectives occur before the nouns to which they refer in spite of SVO sentence order. This combination of feature values is absent from almost all the other languages of Southeast Asia and has probably arisen under the influence of the Transeurasian languages. Rukai combines an

adjective-noun order with a noun-genitive order (ANG). In Dryer's (2005c: 350–357) sample of 1105 languages, 55 per cent have genitive-noun order, while out of 1213 languages, 28 per cent have adjective-noun order.

F22 Extensive use of converbs Converbs, also known as gerunds or adverbial participles, can be defined as non-finite verb forms whose main function is to mark adverbial subordination (Haspelmath 1995: 3). Originally coined by the Altaic scholar Ramstedt, the term converb was adopted from Transeurasian linguistics to denote a cross-linguistic category. The Transeurasian languages are converb-prominent languages in the sense that they use converbs rather than adverbial subordinators as found in many European languages (Alpatov and Podlesskaya 1995; Bisang 1995, 1998; Johanson 1995; Malchukov 2012; Nedjalkov 1995; Sohn 2009):

- (1) Turkish
Ali gel-ince *şaşırd-ı*
 Ali come-CONV be.surprised-PST-3SG
 'When Ali came, he was surprised' (Johanson 1995: 314)
- (2) Khalkha
Ger-ees-ee *gar-aad*
 house-ABL-REFL exit-PFV.CONV
deuc-en *jil-iin* *daraa* *ol-d-lao*
 forty-AND year-GEN after find-PASS-FIN
 'She went away from home and was found forty years later'
 (Janhunen 2012: 280)
- (3) Even
Dagam-mi, *kunte-le* *d'u-v* *it-ti-n*
 approach-CONV clearing-LOC house-ACC see-PST-3SG
 'When he came nearer, he saw a house on a clearing' (Malchukov 2012: 213)
- (4) Korean
Kiho-nun *nol-ko* *ca-ss-eyo*
 Kiho-TOP play-CONV sleep-PST-POL
 'Kiho played and then slept' (Sohn 2009: 300)
- (5) Japanese
Taroo-ga *bangohan-o* *tabe-te* *furo-ni* *hait-ta*
 Taroo-NOM dinner-ACC eat-CONV bath-DAT enter-PST
 'Taroo took a bath after he ate dinner' (Alpatov and Podlesskaya 1995: 473)

Although the Uralic languages are characterized by an extensive use of converbs, Khanty is rather atypical in this sense because it has only a single converb in *-min*, which is the least frequent non-finite verb form. Yukaghir and Nivkh also use a variety of converbs to link clauses. Ainu, however, employs subordinating conjunctions. Ket has no converbs or serial verb

constructions of any kind. In Mandarin, verbs or verbal phrases are merely juxtaposed, the relation between the items being largely unmarked. Rukai marks adverbial subordination through a variety of means such as subordinating conjunctions, changes in word order and nominalized verb forms.

F23 Use of a locative existential construction to encode predicative possession The Transeurasian languages show a clear preference to express the concept ‘X has Y’ on the basis of an existential sentence, whereby the possessed noun phrase functions as the grammatical subject of the ‘exist’-predicate, while the possessor noun phrase is in a dative/locative case form. Although locative possessive constructions were standard in Old Turkic, Turkish uses genitive existential sentences as well as locative existential sentences. ‘I have a book’, for instance, can be expressed by *Ben-de bir kitab var* [I-LOC a book exist] or by *Ben-im bir kitab-ım var* [I-GEN a book-1SG.POSS exist]. Middle Mongolian and Khalkha make use of either a conjunctive possessive which construes the possessor noun phrase as the grammatical subject of the copula and marks the possessed with the comitative *-tai*, e.g. Khalkha *Bi nom-tai bai-n* [I book-COM be-DUR], or a locative possessive, e.g. *Nad-ed nom bai-n* [I-DAT book be-DUR]. As is the case for most Tungusic languages, Manchu and Evenki employ locative existential constructions, e.g. Ekv. *Min-du: kniga bisi-n* [I-DAT book be-3SG], Ma. *Min-de bithe bi* [I-DAT book be]. Korean uses a locative existential construction, e.g. *K Na-hanthey chayk-i issta* [I-LOC book-NOM exist], but the possessor can also be construed as the topic of the noun phrase, e.g. *Na-nun chayk-i issta* [I-TOP book-NOM exist]. This is also true for Japanese, e.g. *Watashi-ni hon-ga aru* [I-DAT book-NOM exist] and *Watashi-wa hon-ga aru* [I-TOP book-NOM exist]. Topic possessives may have developed under the influence of Chinese, since they represent the standard strategy in Mandarin. Among the strategies used to encode predicative possession in the Uralic languages, we find locative possession such as in Finnish and Hungarian, genitive possession such as in Nenets, and possession encoded by a transitive verb ‘to have’ such as in Khanty. Whereas Yukaghir employs a conjunctive possessive and Ainu a *have*-possessive, Ket and Nivkh use locational possessives. Although many Austronesian languages employ topic possessives, Rukai makes use of locative and genitive possessive constructions. In Stassen’s (2005a: 474–477) sample of 240 languages, 20 per cent use a locative existential construction to encode predicative possession.

F24 Use of the ablative case form to encode predicative comparison The Transeurasian languages all form comparative constructions in which the standard noun phrase is constructed in the ablative case form, e.g. Tk. *bu araba-dan daha büyük* [this car-ABL more big] ‘bigger than this car’, OT *barča-da üzä-räk* [everything-ABL high-COMP] ‘higher than anything else’, Khal. *ene xun-ees iluu* [this person-ABL good] ‘better than this

person', MMo. *qola-sa qola* [far-ABL far] 'farther than far', Evk. *oron-duk gugdamar* [deer-ABL tall-COMP] 'taller than a deer', Ma. *ere niyalma ci sain* [this person ABL good] 'better than this person', OJ *ware-yo₁ri mo₂ mantusi-ki, pi₁to₂* [I-ABL PT be.poor-ADN person] 'people poorer than me' and J *chikyu:-yori omoi* [globe-ABL be.heavy] 'heavier than the globe'. In literary Korean the ablative marker *eyse* 'from' can be used in comparative constructions, e.g. *K i eyse te khu-n salang* [this ABL more be.big-ADN love] 'a greater love than this', but it is more common to use a comparative particle *pota* 'than', e.g. *K kicha pota ppaluta* [train PT be.fast] 'faster than a train', MK *nyey pwota thak.wel hota* [past PT superior be] 'superior to the past'. This particle has grammaticalized from the verb MK *pwo-* 'to see' and the transferentive *-ta-ka*, which signals the interruption of an event before its completion, i.e. 'when one looks at'. It replaced earlier particles for comparison in Middle Korean, such as MK *tukwo* 'than' and *lawwa* 'than'. The Uralic languages differ from one another with regard to comparative constructions: languages to the west, such as Finnish and Hungarian, use particle comparatives as in European languages, and languages to the east, such as Nenets and Udmurt, mark the comparative standard with the ablative case ending, as in the Transeurasian languages. In Khanty, the marker of comparison is a postposition *niŋə* 'since, from', which has ablative-like semantics but differs from the standard ablative case ending *-oy* or the ablative-relative ending *-i*. Yukaghir and Ket mark the comparative standard with the ablative case ending. In Nivkh, the comparative suffix *-yk* is traditionally considered as a separate case form as there is no evidence to relate it to the formally similar locative/ablative suffix *-(u)ye*; *-(u)x* (Gruzdeva, personal communication). Ainu forms comparative constructions by means of the particle *kasuno* 'than'. In comparative constructions in Mandarin the standard noun phrase is constructed as the direct object of a verb 'to exceed'. In Rukai, a comparative construction is formed through partial reduplication (CVV) of the descriptive verb stem. In Stassen's (2005c: 490–493) sample of 167 languages, 47 per cent use locational comparatives, but the proportion of languages that specifically use the ablative case form to encode predicative comparison is logically expected to be lower.

22.2.5 Grammaticalization

F25 Direct insubordination One of the driving forces of morphosyntactic change in the Transeurasian languages is a recurrent tendency to grammaticalize non-finite suffixes to finite suffixes (Robbeets 2009, 2016). In line with Evans (2008: 367), I call this development 'insubordination', i.e. the conventionalized main clause use of what appear to be formally subordinate clauses, but it can be further specified as 'direct' insubordination because non-finite suffixes are directly reanalysed as finite ones, without the omission of a specific matrix predicate (Robbeets 2016). Deverbal noun suffixes such as OTK *-(A)r* in OTk. *tug-* 'to be born, to rise

(of sun) (intr.)' > *tugar* 'sunrise, east'; MMo. *-m* in MMo. *quri-* 'to come together (intr.)' > *qurim* 'feast'; Ma. *-rA* in *mute-* 'to be able' > *mutere* 'ability'; MK *-(u/o)m* in *yel-* 'to bear' > *yelum* and OJ *-sa* in *naga-* 'to be long' > *nagasa* 'length' develop over intermediate stages of clausal nominalizers and relativizers into finite suffixes, as illustrated in the following examples:

- (6) Old Turkic
Ölüm-tä oz-upan ögir-ä savin-ü yori-r.
 death- escape- rejoice- be.happy- go.on-FIN
 ABL CONV CONV CONV
 'Having been saved from death it happily goes on with its life.'
 (Erdal 2004: 325)
- (7) Middle Mongolian
udurit-basu ber ulu busire-m.
 guide-COND PT NEG believe-FIN
 'Even if you guide them, they don't believe.' (Weiers 1966: 144)
- (8) 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)
- (9) Korean
onul-un swuep-i eps-um.
 today-TOP class-NOM not.exist-FIN
 'No class today.'
- (10) Old Japanese
punapi₁to₂-wo mi₁-ru-ga to₂mo₂si-sa
 boat.people-ACC see-NML-GEN be.enviable-FIN
 'How enviable it is to see the boat-people!' (Wrona 2008: 206)

The Uralic languages also display a recurrent tendency towards direct insubordination. Deverbal noun suffixes such as Proto-Uralic **-k*, **-pÄ*, **-mä* and **-šÄ* are thought to have developed into finite markers for present (**-k*, **-pÄ*) and past (**-mä*, **-šÄ*) tense, either in Proto-Uralic or after the separation of the daughter languages (Collinder 1965: 110–115; Janhunen 1982: 36–37). Eastern Khanty preserves only a faint trace of this development since the finite form of the negative verb can be marked with the perfective participle *-əm*, as illustrated in example (11). However, the phenomenon is well preserved in the Mansi cognate deverbal noun suffix *-əm* in *uul-* 'to sleep' > *uuləm* 'sleep', which has developed into the finite past, illustrated in example (12). Nikolaeva (1999) also observes the development in the Northern Khanty dialects.

- (11) Eastern Khanty
mata waja *lök ant-im*
 some animal track NEG-FIN
 'There is not a single animal track' (Filchenko 2007: 429)

- (12) Mansi
am joht-um-m
 I come-FIN-1SG
 'I have come' (Collinder 1965: 113)

In Nivkh, there is a single instance of direct insubordination, but the phenomenon does not seem to be recurrent. It concerns the deverbal action noun and infinitive suffix *-d'* which has developed over participial use into a finite form *-d'*:⁸

- (13) Nivkh

<i>if</i>	<i>hum-d'</i>	<i>hyjm-d'</i>
he	live-NML	know-FIN

 'He knows the living one/ (his) life.' (Malchukov 2013: 200)

The remaining neighbouring languages under discussion display strategies other than direct insubordination in grammaticalizing non-finite suffixes to finite suffixes. In Yukaghir and Mandarin, for instance, clausal nominalization in a construction with a copula is the main source for developing new finite constructions. Many Sinitic languages use focus constructions consisting of a nominalizer plus a copula verb; dropping the copula then paves the way for developing finite constructions. The Mandarin *shi* ... *de* focus constructions, for instance, consist of a copula *shi* and a nominalizer *de*, whereas the finite stance construction appears without the copula (Yap and Matthews 2008: 20). Similar processes are found in the Siberian area, for instance in Yukaghir (Malchukov 2013: 192–195). In Kolyma Yukaghir, the deverbal action noun suffix *-l* in *pala:-* 'to escape' > *pala:l* '(a situation of) escaping' has developed into a finite form in subject focus constructions, as illustrated in (14). The intransitive subject 'I' takes a focus marker *-ek*, which is also used to mark nominal predicates, thus pointing to its origin as a copula-like form. As such, the example in (14) can be derived from a cleft-like construction 'It is me sitting'.

- (14) Kolyma Yukaghir
Met-ek moda-l
 I-FOC sit-FIN
 'I sit' (Malchukov 2013: 194)

Ket displays yet another strategy to develop finite markers, namely to reduce the matrix predicate to an affix on the former dependent verb. In example (15), for instance, the matrix verb *bimbata* 'it is audible' is reduced to a present suffix *-beta* ~ *-bata* on verbs expressing sound production (Malchukov 2013: 196–197):

- (15) Ket
 (15a) *tam bis'ej in'nej bi-mbata*
 PT what sound **be.audible-FIN**
 'a certain sound is audible' (Werner 1997: 170)

⁸ Note that Kortlandt (2004: 4) identified the Nivkh suffix *-d'* with the Indo-Uralic participial suffix **-nt*, considering it as evidence of a common origin.

(15b) *p-kutəl'ej-beta*

1SG.POSS-whistle-FIN

'I whistle'

(Werner 1997: 187)

In Ainu, deverbal noun suffixes appear to be functioning as both derivational suffixes and syntactic clausal nominalizers, but there is no indication that they have developed into finite endings. Ainu lacks other non-finite markers such as participial or converb affixes that could be open to developing into finite markers. Similarly, Rukai does not exhibit traces of direct insubordination.

The languages of the world use a variety of mechanisms for developing finite function on formerly non-finite forms such as (i) verbalization of nominal predicates plus finite copula with subsequent copula erosion; (ii) reduction of a finite verb to affix; (iii) insubordination through ellipsis of a matrix clause and (iv) direct insubordination.

F26 Grammaticalization from negative verb to verbal negator The historical development of negation in the Transeurasian languages involves a recurrent development of an independent negative verb into a negative auxiliary verb, which may move from preposed to postposed position and eventually assume suffix status (Robbeets 2014). All Tungusic languages except Manchu have preserved evidence supporting the reconstruction of a negative verb pTg. **e*- 'not to be, not to exist'. There are some instances of independent use of the negative verb, i.e. without a lexical verb, where it means 'not to exist, not to live' as in the Evenki example in (16a). In example (16b), the negative verb acts as a finite auxiliary to the lexical verb, which assumes an invariant adnominal form. In spite of SOV word order, the finite negative verb is preposed to the lexical verb. In emotive sentences, such as in example (16c), the negative auxiliary may move to a postposed position. The Nanai example in (17) represents the final stage of the negative cycle, i.e. fusion, whereby the auxiliary negative verb has assumed the status of derivational suffix on the lexical verb and its phonological form is reduced to lengthening of the stem-final vowel. Although its predecessor Jurchen preserves traces of pTg. **e*- 'not to be, not to exist', Manchu does not, but a similar negative cycle can be reconstructed for the verbal negator Ma. *aku*..

(16) Evenki

a. *esile e-dyeli-m tadu-gla*

now NEG-FUT-1SG there-ENCL

'Now I will not be (live) there.' (Nedjalkov 1994: 27)

b. *nungan nekun-mi e-ce-n suru-v-re.*

he younger.brother-POSS.REFL NEG-PST-3SG go.away-CAUS-ADN

'He did not lead his younger brother away.' (Nedjalkov 1994: 11)

brother away.'

c. *nungan songo-ro e-ce-n*

he cry-ADN NEG-PST-3SG

'He did not cry [– what's the use of crying?].' (Nedjalkov 1994: 8)

(17) Nanai

xola:-ci-si

read.NEG-PST-2SG

'You didn't read.'

Like Old Turkic, Turkish has a verbal negative suffix *-mA-* that can be derived from an original negative verb pTk. **ma-* 'not to exist'. The verbal origin is supported by the occurrence of a negative postposition *mar* in Chuvash, which contains a deverbal noun suffix **-r* and can take a nominal argument such as the directive case in debitive constructions. The Middle Mongolian negative verb stem *ese-* 'not to be, not to exist' survived in a number of conjugated forms, such as with the past marker *-be-* in example (18a), but gradually the negative auxiliary became used as an invariant form, transferring its entire inflection to the lexical verb, i.e. the past marker *-be* is attached to *ire-* 'to come' in example (18b).

(18) Written Mongolian

a. *ükü-be-üü* *ese-be-üü*

die-PST-INTER NEG-PST-INTER

'Did [he] die or did [he] not?' (Poppe 1954: 175)

b. *manu bayši ese ire-be*

our teacher NEG come-PST

'Our teacher did not come.' (Poppe 1954: 175)

The Middle and Contemporary Korean verbal negator MK *a-ni*, K *an(i)* can be derived from an original negative verb **an-* and the suffix MK *-i* that derives both nouns and adverbs from verbs. Gradually, the negator *ani* is being replaced by an analytic construction consisting of *ani* augmented by the finite auxiliary MK *ho-*, K *ha-* 'to do, be', which usually contracts to *anh-* 'not to do' in Contemporary Korean. This seems to reflect the start of the next negative cycle, whereby the grammaticalized verbal negator is replaced by a new negative construction in which a negative verb is restored in its function as finite auxiliary to the lexical verb, which is nominalized with the suffix *-ci*.

(19) Middle Korean

ˊes-tyey a-ni wo-no-n-ywo

why NEG come-PROC-ADN-INTER

'Why is [he] not coming?' (Martin 1992: 420)

(20) Korean

apenim un ka-ci anh-usy-e

father TOP go-NML NEG-HON-FIN

'Father is not going.' (Robbeets 2015)

Old and Contemporary Japanese use an independent negative existential adjective *na-* 'to be non-existent, not to exist', illustrated in (21a), which is thought to derive from the same origin as the Old Japanese negative suffix *-(a)n-*, illustrated in (21b) (Martin 1988: 821). As such, an original negative

verb pJ **ana-* ‘not to exist’ seems to have developed into a negative suffix. Negative imperative constructions with *na-* preserve a trace of the originally preposed position of the negative auxiliary.

(21) Old Japanese

- a. *s-uru* *sube₁-no₂* *na-sa*
do-ADN way-GEN NEG-NML
‘Nothing can be done.’

(Vovin 2009: 483)

- b. *ki₁mi₁-ga* *k-i₁-mas-an-u*
lord-GEN come-CONV-deign-NEG-ADN
‘You did not come, [my] lord.’

(Vovin 2009: 788)

Similar to the Transeurasian languages, one of the characteristics of the Uralic languages is the expression of negation by means of a construction, comprising a fully inflected negative auxiliary and a largely invariant lexical verb (Comrie 1981; Honti 1997; Janhunen 1982: 37; Payne 1985: 215–221; Suihkonen 2002: 173). The construction may develop in ways which result in a redistribution of inflectional categories between the negative verb and the lexical verb. Eventually, as is the case for the negative particles Khanty *əntə* or Estonian *ei*, the negative auxiliary may become totally free of inflections and turn into an invariant verbal negator, which recalls the situation in Mongolic in (18b). However, there are no examples in Uralic in which the negative auxiliary ultimately becomes a suffix, as it does in Turkic, Tungusic and Japanese. In Yukaghir, clausal negation is expressed by a proclitic *el-*, which usually precedes the verb, but in spite of the formal similarity with the Proto-Uralic negative auxiliary **e-*, there are no language-internal indications that it originated in a negative verb or auxiliary (Fubito Endo, personal communication). Clausal negation in Nivkh is expressed in three ways: (i) a construction with the negative verb *k’au-* ‘not to be, not to have’ preceded by the lexical verb in an invariant dative case form; (ii) an incorporation of the negative verb *ḡavr-/gavr-* ‘not to be, not to have’ incorporated into the verbal form; (iii) a negative suffix *-rla / -tla*. Given the formal similarity between the negative verb *k’au-* and the negative affix *-ḡavr-*, it is not unlikely that they go back to the same source (Ekaterina Gruzdeva, personal communication). Among the many negatives used in Mandarin, the most general and neutral negation is expressed by *bu*, whereas the existential negative *mei* ‘there is not, has not’ is used to negate the completion of an event ‘not yet’. Both *mei* and *bu* originate as verbs (van Gelderen 2008: 225). Ainu uses a negative particle that precedes the verb (e.g. *somo ku-oman* [NEG 1SG-go] ‘I do not go’). The non-verbal status of *somo* is indicated by the fact that it does not take any personal affix. There are also no indications that the Ket negative particle *bə:n* or the Rukai negative suffix *-ka* originate in a negative verb.

The expression of negation via negative auxiliaries is worldwide a minor type to begin with, found in only 40 (17 per cent) out of 240 languages in

Dahl's (1979) sample, which is areally biased towards Uralic and Altaic languages, in 45 (4 per cent) out of 1,011 languages in Dryer's (2005d) sample, and in 16 (5 per cent) out of the 297 languages in Miestamo's (2005) sample. As a consequence, the particular development of negative verbs to auxiliaries to particles or suffixes is hence even rarer.

F27 A morphologically simplex first person plural pronoun is complemented by the grammaticalization of the first person pronoun augmented with a collective-plural marker When dealing with the inclusive/exclusive distinction in first person plural pronouns in Section 22.2.2 (F11), it was mentioned that most Turkic languages and Korean complement their first person plural pronoun (Tk./OT *biz* 'we'; K/ MK *wuli* 'we') with an augmented collective-plural form (Tk./OT *biz-ler* 'we (as a group)'; K *wuli-tul*, MK *wuli-tolh* 'we (as a group)'). A similar tendency has been found in the history of Japanese, where the first person singular / plural OJ *wa*- 'I, we' coexists with the same form augmented by a collective marker OJ *wa-re* 'we', a form which in its turn was later augmented into *ware-ra* 'we' in the history of Japanese. Etymologically, the Middle Mongolian inclusive *bida*, reflected in the Khalkha formally inclusive oblique *bidn-*, derives from the first person singular MMo. *bi* 'I' and a plural suffix *-da*, which also occurs in the plural demonstratives pronouns MMo. *e-de* 'these' versus *te-de* 'those' (Doerfer 1985: 2; Domii 2006; Nevskaya 2010: 119).⁹ Domii argues that originally, **ba* and **bi-da* complemented each other as plural pronouns and that the distinction between exclusive and inclusive meaning was a secondary development. The Tungusic exclusives Evk. *bu* and Ma. *be* can be derived from the first person plural pTg. **bö* and an augmented plural **bö-(x)e*, respectively (Doerfer 1978: 81–83, 95–96; Janhunen 2013: 217), whereas the inclusive Evk. *mut* ~ *mit* may go back to pTg. **bö* plus the collective suffix pTg. **-ti* (Benzing 1955: 1020) and the inclusive Ma. *muse* may be an extension of this root with the collective suffix *-sa* (Benzing 1955: 1017–1018). This analysis suggests that successive cycles of plural augmentation on morphologically simplex (or simplified) plural pronouns have triggered the secondary development of an inclusive/exclusive distinction in Tungusic. As far as the Uralic languages are concerned, Khanty makes a commonly found distinction between pronouns in the first person singular (*mä* 'I'), dual (*min* 'both you and me') and plural (*məŋ* ~ *miŋ* 'we'), but it does not reflect any trace of plural augmentation on the first person plural pronoun. Similarly, no traces of plural augmentation on first person plural pronouns are found in Ket or Yukaghir. In Rukai, the first person plural inclusive *-mita* [NOM] is formally underivable from the exclusive *-nai* ~ *nai-* [NOM]. The personal pronouns in Ainu have all grammaticalized from person affixes followed by any one of

⁹ An alternative analysis, deriving the inclusive MMo. *bida* from the first person singular pronoun **bi* 'I' plus the second plural pronoun **ta* 'you (many)' is proposed by Janhunen (2013: 215), but the voicing of the medial dental stop would represent an irregular development.

several existential verbs meaning ‘exist’. The first person plural pronoun *aoka(i)*, for instance, consists of the first person plural inclusive transitive subject affix *a-* and the verb *oka* ‘to exist’. In Nivkh, however, all plural personal pronouns can optionally be augmented with a plural suffix; the first person plural exclusive pronoun, for instance, appears either as *n'yŋ-ø* or as *n'yŋ-gu* [1PL-PL] ‘we’. In Mandarin, two separate roots for the first person singular *wǒ* and *zán* pluralized, along with the suffix *-men*, into the derived exclusive *wǒmen* and inclusive *zánmen* ‘we’. However, since in Classical Chinese *wǒ* ‘I’ could be used as a first person plural ‘we’ as well, the exclusive *wǒmen* can be regarded as an instance of plural augmentation.

22.2.6 Overview

In the body of this chapter, I have set up a list of 27 feature labels, chosen to maximize positive values for the Transeurasian languages. These features, inserted as vertical comparison points in the tables below, have been examined for selected representatives among the Transeurasian languages and their linguistic neighbours, which are inserted as horizontal comparative points. In the tables, I summarize the observations made above by introducing plus (+) and minus (–) values in the corresponding cells. This then leads to a quantification of the number of plus values in the last row.

As far as the feature values for the Transeurasian languages are concerned, Table 22.1 shows the following tendencies. First, the typological coherence seems to be greater for historical than for the contemporary stages of the languages investigated. This suggests that Transeurasian areality has decreased over the last millennium. Second, maximal coherence is found in the Mongolic and Tungusic languages, with minor deviations from the prototype in the Turkic languages in the west and somewhat more in the Japonic and Koreanic languages in the eastern periphery. Third, the deviation from the prototype in the east does not reflect a gradual loss as we proceed from Korean to Japanese, but rather an *en bloc* reduction of features or even a slight increase for Old Japanese.

As far as the feature values for representative neighbouring languages are concerned, Table 22.2 shows the following tendencies. First, the neighbouring languages show significantly stronger deviations from the prototype than do any of the investigated Transeurasian varieties. This suggests that it is meaningful to apply the concept of ‘areality’ to the Transeurasian languages in the sense that they reflect a geographical concentration of linguistic features that sets them apart from the selected neighbouring languages. Second, Khanty and Yukaghir show more typological similarity with the Transeurasian prototype than do other neighbouring languages. Note that for at least three of the examined features (i.e. F6, F18, F22), Khanty yields a minus value where the Uralic prototype would yield a plus value. This suggests that ‘areality’ may also apply in a wider, but less

Table 22.1 Feature values for selected Transeurasian languages along with their historical stages

	Tk.	(pre-)OT	Khal.	(pre-)MMo.	Evk.	(pre-)Ma.	K	(pre-)MK	J	OJ	(pre-)
F1	+	+	+	+	+	+	+	+	+	+	+
F2	+	+	+	+	+	+	+	+	+	+	+
F3	+	+	+	+	+	+	+	+	+	?	+
F4	-	-	+	+	+	+	-	+	-	?	+
F5	+	+	+	+	-	+	+	+	+	+	+
F6	+	+	+	+	+	+	+	+	+	+	+
F7	+	+	+	+	+	+	+	+	+	+	+
F8	+	+	+	+	+	+	-	+	+	+	+
F9	+	?	+	?	-	?	+	?	+	?	+
F10	-	+	+	+	+	+	-	-	-	+	+
F11	-	-	-	+	+	+	+	+	-	+	+
F12	-	+	-	+	+	+	-	+	-	+	+
F13	-	+	-	+	-	+	-	-	-	-	+
F14	+	+	+	+	+	-	+	+	+	+	+
F15	+	+	+	+	+	+	+	+	+	+	+
F16	+	+	+	+	+	+	+	+	+	+	+
F17	+	+	+	+	+	+	+	+	-	+	+
F18	-	+	+	+	+	+	-	-	-	-	+
F19	-	+	+	+	+	+	-	-	-	+	+
F20	+	+	+	+	+	+	+	+	+	+	+
F21	+	+	+	+	+	+	+	+	+	+	+
F22	+	+	+	+	+	+	+	+	+	+	+
F23	+	+	+	+	+	+	+	+	+	+	+
F24	+	+	+	+	+	+	+	+	+	+	+
F25	+	+	+	+	+	+	+	-	+	+	+
F26	+	+	+	+	+	+	+	+	+	+	+
F27	+	+	+	+	+	+	+	+	+	+	+
	20	24	24	26	24	25	19	19	18	20	

Table 22.2 *Feature values for representative neighbouring languages*

	Khan.	Ket	Yuk.	Niv.	Ain.	Ch.	Ruk.
F1	+	+	+	+	+	–	+
F2	+	–	+	+	+	–	+
F3	+	–	+	+	+	–	–
F4	–	–	+	?	?	–	–
F5	+	+	+	–	+	+	–
F6	–	+	+	–	–	–	–
F7	+	+	+	–	+	+	+
F8	–	+	+	–	–	–	+
F9	–	–	?	–	–	–	–
F10	+	–	–	–	–	+	–
F11	–	–	–	+	–	+	+
F12	–	–	–	–	–	–	–
F13	–	–	+	–	–	–	–
F14	–	–	–	–	–	–	–
F15	+	+	+	+	+	–	+
F16	+	–	+	–	–	–	–
F17	+	+	+	–	–	–	+
F18	–	–	+	–	–	–	–
F19	+	–	–	–	–	–	+
F20	+	+	+	+	+	–	–
F21	+	+	+	+	+	+	–
F22	–	–	+	+	–	–	–
F23	–	+	–	+	–	–	+
F24	–	+	+	–	–	–	+
F25	+	–	–	+	–	–	–
F26	+	–	–	+	–	+	–
F27	–	–	–	+	–	+	–
	14	11	17	12	8	7	8

coherent sense to the belt of Transeurasian-Yukaghiric-Uralic languages. Third, the investigated languages of North Asia have more typological features in common than those in Southeast Asia, i.e. Mandarin and Rukai. This suggests a third ring of areality that is the least uniform, involving the languages of North Asia.

22.3 Interpretation of the Observations

22.3.1 Delimitation of Areality

The Transeurasian continuum has clear boundaries which delimit the language type in relation to its neighbours both to the north (Yeniseic, Yukaghiric) and east (Nivkh, Ainu) as well as to the south (Sinitic, Austronesian). Although the observations above are in line with Janhunen's (2009: 61–62) findings about a certain internal uniformity in the larger Ural-Altaic belt, they also suggest including Yukaghir in this larger belt, and they indicate additional boundaries in areality between the Uralic and the Transeurasian languages as such. Among the features that enable us to delimit the Transeurasian languages in relation to their Uralic

neighbours are: F4 tongue root harmony in Transeurasian (and Yukaghir) versus palatal harmony in Uralic; F8 voicing distinction for stops in Transeurasian (and Yukaghir) versus original singleton–geminate distinction in Uralic; F9 non-verbal strategy of verbal borrowing in Transeurasian versus direct insertion in Uralic; F11 inclusive/exclusive distinction in Transeurasian versus none in Uralic (and Yukaghir); F12/F13 mixed and switched encoding of property words in Transeurasian (and perhaps originally in Yukaghir) versus nominal encoding in Uralic; F14 partial emphatic reduplication in Transeurasian versus none in Uralic (and Yukaghir); F18 absence of initial *m* in the nominative first person singular versus presence in Uralic (and Yukaghir); F25 development of a negative verb into a suffix in Transeurasian versus none in Uralic (and Yukaghir); and F27 augmented first person plural pronoun in Transeurasian versus none in Uralic (and Yukaghir). For some features such as F23 and F24, Uralic makes use of a larger variety of strategies than the Transeurasian languages, where all languages uniformly use locative possession or ablative comparatives. It is remarkable that Yukaghir aligns with Uralic rather than with Transeurasian in more than half of the delimiting features (i.e. F9, F11, F14, F18, F25 and F27), although it is geographically adjacent to Transeurasian languages such as Yakut (Turkic) and some Northern Tungusic languages, but not to the Uralic languages. In my opinion, this observation is probably not coincidental, but it might reflect the alleged genetic relatedness between Uralic and Yukaghir proposed by, among others, Collinder (1965).

22.3.2 Deviations from the Prototype

Along the margins of the Transeurasian continuum, we can observe examples of gradual loss of Transeurasian features in the western and eastern peripheries, as well as gradual adoption of Transeurasian features, as in the case of Mandarin.

Examples of original Transeurasian features changing in the western periphery under Uralic influence are: F4 Transeurasian tongue root harmony, which aligns with the Uralic languages as palatal harmony in Turkic; F12 gradual loss of verbal encoding of property words – mirroring Uralic nominal encoding – as one proceeds from older to contemporary varieties and from Tungusic in the east to Turkic in the west; and F18 the secondary development of *m*-initials yielding a *mi*–*Ti* opposition in first versus second person singular pronouns in Turkic, Mongolic and Tungusic.

Changes in areality in the eastern peripheries may take place under the influence of the languages to the extreme northeast of the Siberian area or under Chinese influence. Examples of original Transeurasian features in Tungusic and Mongolic changing under Siberian influence are: F5 the secondary assimilation of pTg. *g- into an initial velar nasal in Tungusic, in line with Nivkh; and F11/F27 the secondary development of an

inclusive/exclusive distinction on augmented plural pronouns, mirroring the situation in Ainu and Nivkh. Examples of Korean and Japanese features aligning with the extreme northeast Siberian area are the lack of voicing distinction in Korean and Old Japanese, in line with Ainu, Nivkh and Chukchi, and F10 the development of a mesial demonstrative distinction in Japanese and its presence in Korean, similar to the situation in Yukaghir, Ainu and Nivkh.

Chinese features seem to have diffused into Manchu, Korean and Japanese, for instance: F1 the gradual increase of monosyllabic roots in Japanese; F2 the development of simple tone systems in Japanese and Korean; F3/F4 the alleged erosion of tongue root vowel harmony in Old Japanese; F12 the relatively strong proportion of verbally encoded property words in Japanese and Korean in comparison to the other Transeurasian languages; F15 the increase of analytic features in Manchu in comparison to the other Tungusic languages; F17 the increase of sortal numeral classifiers in Manchu vis-à-vis the other Tungusic languages and in Japanese and Korean vis-à-vis older varieties of the languages; F19 the gradual de-pronominalization, which has taken place in the recorded history of Japanese and Korean; and F23 the development of topic possessives in Korean and Japanese. Note that some Transeurasian languages to the centre of the continuum, such as several Turkic and Mongolic languages of the Amdo Qinghai region, have also lost prototypical Transeurasian features under the influence of Chinese and other languages of the area (Janhunen 2007).

However, the above observations support previous studies by Hashimoto (1986), Norman (1988: 10–12, 20) and Comrie (2008), arguing that the Transeurasian languages have also left a serious mark on the linguistic structure of Chinese. The following developments illustrate how Chinese may have changed some of its original Mainland Southeast Asian features under Transeurasian influence: F1 the development of a greater number of polysyllabic roots compared to Classical Chinese; F7 the simplification and loss of consonant clusters compared to Old Chinese; F10 the development of a two-way distinction in demonstratives compared to the three-way distinction in Classical Chinese; F11 the development of an inclusive/exclusive distinction in first person plural pronouns in Beijing and certain other northern Chinese dialects, which was not found in Old Chinese; F16 the weak suffixing tendency of Mandarin as opposed to other Sinitic languages; and F21 the rare combination of SVO sentence order and GAN noun phrase order in Mandarin, absent in almost all the other languages of Southeast Asia. Geographically, Chinese is located between the Transeurasian languages and the languages of Mainland Southeast Asia, an intermediate position, which it also occupies from the point of view of typology.

Finally, some features in the Siberian languages to the extreme north-east seem to have diffused directly from Southeast Asia, without

a Transeurasian intermediary: F2 the occurrence of two distinctive tones in Nivkh in comparison to the relatively simple pitch-accent systems of Japanese and Korean; F12 the exclusively verbal encoding of property words in Yukaghir, Nivkh and Ainu, similar to Mandarin, but different from mixed encoding in Japanese and Korean; and F17 the obligatory use of an extensive list of classifiers in the Nivkh lexicon, and a smaller one in Ainu, recalling the widespread and archaic use of classifiers in Southeast Asia, as opposed to their relatively late development in Japanese and Korean. This observation may gain relevance in the light of theories that derive Ainu from the south (e.g. Bengston and Blažek 2009; Murayama 1992; Vovin 1993).

22.3.3 Diffused versus Inherited Features

A simplistic interpretation of the observations would be to assert that the properties of the Transeurasian language type are universally so common that their parallel occurrence in several adjacent language families is coincidental. This is certainly not the case, however, because the Transeurasian continuum has clear boundaries which delimit the language type in relation to its neighbours both to the west (Uralic), north (Yeniseic, Yukaghiric), east (Nivkh, Ainu) and to the south (Sinitic, Austronesian). Moreover, the relatively low frequency of some features indicates that the shared properties are not due to mere universal principles in linguistic structuring. Above I have provided an estimation of the frequency of 19 out of 27 features. Seven features are not very common (i.e. F5, F9, F10, F16, F20, F21, F23) in the sense that they occur in less than half (50 per cent) but more than a third (33 per cent) of the languages worldwide. Nine features are relatively uncommon in the sense that they occur in less than a third (33 per cent) of the languages worldwide (e.g. F4, F11, F12, F13, F18, F19, F23, F25). Phenomena that are relatively infrequent and randomly spread across the world's languages but frequent and geographically concentrated in a specific group of languages provide evidence of a historical connection – be it areal or genealogical – between the languages concerned (Croft 1990: 206–207). The strength of the argument increases when a number of features correlate in a particular part of the world, but not in the world as a whole.

It is important to note that the typological similarities among the Transeurasian languages are accompanied by a significant number of correspondences in the lexicon (see Robbeets 2005) as well as in verb morphology (see Robbeets 2007a, 2007b, 2010, 2012) in such a way that – in my own judgement – these languages are likely to be genealogically related. The most plausible family tree, representing the overall relationships is given in Figure 22.1. The affiliation of the Transeurasian languages remains debated, but even critics such as Janhunen (1996: 220) would

agree that before the first millennium BCE the homelands of the individual language families concerned were all located in a compact area in southern Manchuria, along with the homelands of Ainuic and Nivkh speakers.

Although some of the shared features discussed above, such as F11 inclusive/exclusive distinction in first person plural pronouns, F14 partial emphatic reduplication of nominal property words, or F18 *mi-Ti* opposition in first versus second person singular pronouns are almost certainly contact-induced, others appear to be the residue of common ancestral features, as suggested by the following six observations.

(i) Geography: isolated position of Japanese Although the Sea of Japan and the Tsushima Strait form a strong geographical boundary separating Japanese from the other Transeurasian languages, Japanese is typologically closer to the Transeurasian languages than geographically less isolated languages such as Ket, Yukaghir, Ainu and Nivkh. Even within a prehistoric contact scenario, this suggests that the Transeurasian characteristics in Japanese did not exclusively arise through diffusion because Ainuic and Nivkh were also present in southern Manchuria.

(ii) History: older varieties are more prototypically Transeurasian A comparison of typological uniformity between historical and contemporary stages of the languages investigated suggests that Transeurasian areality has decreased over the last millennium. While influences diffusing from adjacent areas such as Mainland Southeast Asia, Siberian and Uralic have demonstrably displaced earlier Transeurasian features in certain contact zones, I find no evidence of Transeurasian features having displaced earlier Chinese, Siberian or Uralic features inherent to the continuum from Japanese to Turkic. Among the examples of displacement of features in contact zones, for instance, we find that initial velar nasals have developed in Tungusic under Siberian influence (F5), simple tone systems and classifiers have developed in Japanese and Korean under Chinese influence (F2 and F17), and palatal harmony has developed in Turkic under Uralic influence (F4). We furthermore note that nominal encoding of property words has increased in Turkic, Mongolic and Tungusic under Uralic influence, while verbal encoding has increased in Japanese and Korean under Siberian and Southeast Asian influence (F12). However, we find no evidence of Transeurasian features entering, for instance, from the Turkic languages and diffusing all over the Transeurasian area, while displacing original and prototypical Sinitic features. This suggests that Transeurasian features are inherent to these languages.

(iii) Distribution: maximal coherence in Mongolic and Tungusic Maximal structural uniformity is found in the Mongolic and Tungusic languages. This distributional pattern conforms to the expectations for the Mongolic languages within a diffusional scenario, since they constitute the centre of the linguistic continuum, but it is not what one

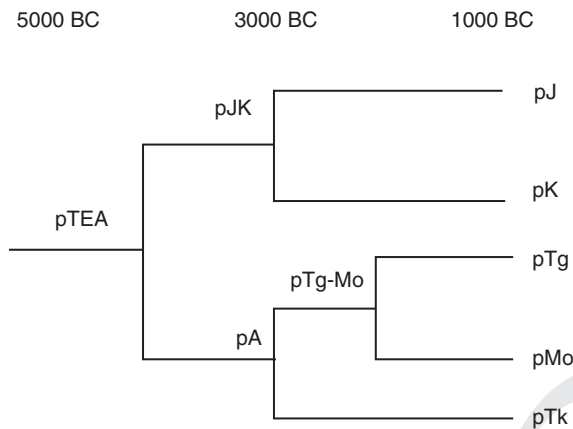


Figure 22.1 Family tree of the Transeurasian languages

would expect for the Tungusic languages, extending towards the north-eastern periphery. The structural coherence in Mongolic and Tungusic recalls the separation of Mongolo-Tungusic in Figure 22.1 as a distinct genealogical unit.

(iv) Distribution: en bloc reduction of features in Korean and Japanese Similarly, the collective rather than gradual reduction – if not slight increase – of features as we proceed from Korean to (Old) Japanese is not what we would expect within a scenario of gradual diffusion. It is furthermore difficult to explain how some Transeurasian features, such as F10 and F19, that show a gap in Korean, have diffused into Japanese without a Korean intermediary.

(v) Cyclicity: recurrent grammaticalization The features involving shared patterns of grammaticalization in Section 22.2.5 are particularly good candidates for genealogical motivation because they are recurrent in different forms and at various chronological stages of the same language. Aikhenvald (2013) characterized contact-induced grammaticalization as ‘change against the grain’ or atypical grammaticalization, while she regarded genealogically motivated grammaticalization as ‘change that reinforces similarities’ because it tends to maintain uniformity between related languages. Given that languages tend to renew their formal encodings in cyclic processes of grammaticalization while maintaining their inherited grammatical categories, new forms are thus expected to grammaticalize along shared conceptual pathways to restore old categories (Heath 1998: 729). Consequently, genealogically motivated grammaticalization is expected to recur on different formal encodings at various points in time, while contact-induced grammaticalization is expected to be restricted to a single formal encoding (or to a very limited number of encodings) during a certain period of contact. The repeated waves of grammaticalization and replacement involved in features F25–F27 imply that the parallel patterns are genealogically motivated.

(vi) Isomorphism: shared features combine with formal correspondences The observation that some structural features shared among the Transeurasian languages combine with a formal correspondence of the marker reflecting the particular feature is also indicative of genealogical retention. This is, for instance, the case for: F9 the non-verbal strategy of verbal borrowing employing a deverbal noun suffix of the common shape **-lA-* (Tk. *-lA-*, Khal. *-l-*, Ud. *-lA-*, J *-r(a)-*) to accommodate for verbal borrowings; F19 the formation of a secondary oblique stem of personal pronouns through a common suffix **-n-* in all Transeurasian languages, except Korean; F25 direct insubordination involving deverbal noun and finite suffixes of the common shape **-rA-*, **-mA-*, **-n-*, **-xA* ~ **-kA* and **-sA* (Robbeets 2009, 2015, 2016) across all Transeurasian languages; and F26 the grammaticalization from negative verb to verbal negator, involving common negative verbs of the common shape **ana-*, **e-* and **ma-* across the Transeurasian languages (Robbeets 2014). In instances like these in which isostructuralism coincides with form-function isomorphism, the structural correspondence is likely to be genealogically motivated, especially when it concerns an instance of shared grammaticalization (Robbeets 2013). Note that the Uralic languages also display oblique personal pronouns in *-n-*, direct insubordination in **-k-*, **-mə* and **-šĀ* and grammaticalization of negative verbs in **e-*, an observation which seems to point to remote genealogical ties between the Uralic and the Transeurasian languages.

22.4 Conclusion

In this chapter I have tried to show that the Transeurasian languages form an internally homogeneous linguistic continuum. For this purpose, I have examined the areal concentration of 27 features in the Transeurasian languages, providing a typological profile of some contemporary varieties in relation to historical stages of the languages involved and to selected languages immediately outside the continuum. Comparison with neighbours to the north (Yeniseic, Yukaghiric), south (Sinitic, Austronesian), east (Nivkh, Ainu) and west (Uralic) makes it possible to set up boundaries which delimit the Transeurasian prototype. Along the margins of the Transeurasian continuum, I have found examples of gradual loss of Transeurasian features in the western and eastern peripheries, as well as gradual adoption of Transeurasian features, as in the case of Mandarin. The data further suggest that the Transeurasian continuum in its turn is part of a larger Uralic-Yukaghiric-Transeurasian belt of languages, which again is part of a larger area of North Asian languages. Although it is meaningful to apply the concept of 'areality' to the Transeurasian languages in the sense of a historically motivated geographical concentration of linguistic features, I prefer avoiding the label 'area' with reference to these languages because this would imply that all shared properties are the

result of diffusion. Observations relating to geography, history, distribution, cyclicity of grammaticalization and combined isomorphism indicate that this is not the case.

A fuller study would need to take more feature values into account and to insert a larger variety of Transeurasian languages as comparative points. Neighbouring languages should also be more diversified, and adjacent languages in the west such as Indo-European languages or languages of the Caucasus region should be included. One should also pay attention to structural dependencies between the features and to considering whether particular features can be easier accounted for by diffusion or by genealogical retention. For the latter purpose, it would be particularly interesting to take common diachronic mechanisms, such as shared patterns of grammaticalization into fuller account. Although this chapter perhaps raises as many new questions as it answers, I hope to have contributed here to the understanding of areality among the Transeurasian languages.

Abbreviations

Linguistic forms

ABL	ablative
ACC	accusative
ADD	additive
ADN	adnominalizer
CAUS	causative
CLASS	classifier
COM	comitative
COMP	comparative
CONV	converb
COND	conditional
NEG	negative
NML	nominalizer
NOM	nominative
OBL	oblique
PFV	perfective
PL	plural
POL	polite
POSS	possessive
PROC	processive
PST	past
PT	particle
REFL	reflexive
SG	singular
TOP	topic

Languages

Ain.	Ainu
Ch.	Mandarin Chinese
Evk.	Evenki
J	Japanese
K	Korean
Khal.	Khalkha
Khan.	Khanty
Ket	Ket
Ma.	Manchu
MK	Middle Korean
MMo.	Middle Mongolian
Niv.	Nivkh
WMo.	Written Mongolian
OJ	Old Japanese
OT	Old Turkic
pJ	Proto-Japonic
pK	Proto-Koreanic
pMo.	Proto-Mongolic
pTg.	Proto-Tungusic
pTk.	Proto-Turkic
Ruk.	Mantauran Rukai
Yuk.	Yukaghir

Acknowledgement

The research presented here was realized with financial assistance of the DFG (Deutsche Forschungsgemeinschaft), supporting a collaborative research project 'Die transeurasiatischen Sprachen: Kontakt in der Familie' at the University of Mainz from January 2010 to December 2013. I thank Anna Bugaeva, Fubito Endo, Andrey Filchenko, Ekaterina Gruzdeva, Seongyeon Ko, Fukui Rei and Elisabeth Zeitoun as well as the editor, Raymond Hickey, for their helpful feedback.

References

- Aikhenvald, Alexandra, 2013. Areal diffusion and parallelism in drift: Shared grammaticalization patterns. In Robbeets and Cuyckens (eds), pp. 23–42.
- Alpatov, Vladimir and Vera Podlesskaya, 1995. Converbs in Japanese. In Haspelmath and König (eds), pp. 465–486.
- Anderson, Gregory, 2005. The velar nasal (ŋ). In Haspelmath et al. (eds), pp. 42–45.
- Anderson, Gregory, 2006. Towards a typology of the Siberian linguistic area. In Yaron Matras, April McMahon and Nigel Vincent (eds), *Linguistic Areas: Convergence in Historical and Typological Perspective*, pp. 266–300. Basingstoke, UK: Palgrave Macmillan.
- Bengston, John D. and Václav Blažek, 2009. Ainu and Austric: Evidence of genetic relationship. *Journal of Language Relationship* 2: 1–24.
- Benzing, Johannes, 1955. Die tungusischen Sprachen: Versuch einer vergleichenden Grammatik. *Abhandlungen der geistes- und sozialwissenschaftlichen Klasse* 11: 949–1099.
- Bisang, Walter, 1995. Verb serialization and converbs: Differences and similarities. In Haspelmath and König (eds), pp. 137–188.
- Bisang, Walter, 1998. Structural similarities of clause combining in Turkic, Mongolian, Manchu-Tungusic and Japanese: A typological alternative to the hypothesis of a genetic relationship. In Lars Johanson (ed.), *The Mainz Meeting: Proceedings of the Seventh International Conference on Turkish Linguistics*, pp. 199–223. Wiesbaden: Harrassowitz.
- Bugaeva, Anna, 2015. Causative constructions in Ainu: A typological perspective with remarks on the diachrony. *STUF – Language Typology and Universals (Sprachtypologie und Universalienforschung)* 68 (4): 439–484.
- Bulatova, Nadežda Ja. and Leonore A. Grenoble 1999. *Evenki*. Languages of the World/Materials, vol. 141. Munich: LINCOM.
- Collinder, Björn, 1965. *An Introduction to the Uralic Languages*. Berkeley, CA: University of California Press.
- Comrie, Bernard, 1981. Negation and other verb categories in the Uralic languages. In Osmo Ikola (ed.), *Congressus Quintis Internationalis Fenno-Ugristarum*, vol. VI, pp. 350–355. Turku: Suomen Kielen Seura.

- Comrie, Bernard, 2008. The areal typology of Chinese: Between North and Southeast Asia. In Redouane Djamouri, Barbara Meisterernst and Rint Sybesma (eds), *Chinese Linguistics in Leipzig*, pp. 1–21. Collection des Cahiers de Linguistique Asie Orientale, vol. 12. Paris: École des Hautes Études en Sciences Sociales, Centre de Recherches Linguistiques sur l'Asie Orientale.
- Croft, William, 1990. *Typology and Universals*. Cambridge: Cambridge University Press.
- Cysouw, Michael, 2005. Inclusive/exclusive forms for 'we'. In Haspelmath et al. (eds), pp. 162–169.
- Dahl, Östen, 1979. Typology of sentence negation. *Linguistics* 17: 79–106.
- Diessel, Holger, 2005. Distance contrasts in demonstratives. In Haspelmath et al. (eds), pp. 170–173.
- Doerfer, Gerhard, 1978. Urtungusisch ö. In Gerhard Doerfer and Michael Weiers (eds), *Beiträge zur nordasiatischen Kulturgeschichte*, pp. 66–116. Tungusica, vol. 1. Wiesbaden: Harrassowitz.
- Doerfer, Gerhard, 1985. *Mongolo-Tungusica*. Wiesbaden: Steiner.
- Domii, Tumurtoogoo, 2006. The inclusive and the exclusive in Mongolian. In T. Shagdarsuren (ed.), *Mongol ulsin ix sürgүүлйн: Erdem šinžilgeenij bičig*, pp. 77–78. Acta Mongolica, vol. 6.267. Ulanbaatar: National University of Mongolia.
- Dryer, Matthew S., 2005a. Prefixing versus suffixing in inflectional morphology. In Haspelmath et al. (eds), pp. 110–113.
- Dryer, Matthew S., 2005b. Order of subject, object, and verb. In Haspelmath et al. (eds), pp. 330–333.
- Dryer, Matthew S., 2005c. Order of genitive and noun; Order of adjective and noun. In Haspelmath et al. (eds), pp. 350–357.
- Dryer, Matthew, 2005d. Negative morphemes. In Haspelmath et al. (eds), pp. 454–457.
- Erdal, Marcel, 2004. *A Grammar of Old Turkic*. Leiden: Brill.
- Evans, Nicholas, 2008. Insubordination and its uses. In Irina Nikolaeva (ed.), *Finiteness: Theoretical and Empirical Foundations*, pp. 366–431. Oxford: Oxford University Press.
- Filchenko, Andrey Yury, 2007. *A Grammar of Eastern Khanty*. PhD dissertation, Rice University, Houston.
- Frellesvig, Bjarke, 2010. *A History of the Japanese Language*. Cambridge: Cambridge University Press.
- Frellesvig, Bjarke and John Whitman, 2008. Evidence for seven vowels in Proto-Japanese. In Bjarke Frellesvig and John Whitman (eds), *Proto-Japanese: Issues and Prospects*, pp. 15–41. Current Issues in Linguistic Theory, vol. 294. Amsterdam: John Benjamins.
- van Gelderen, Elly, 2008. Negative cycles. *Linguistic Typology* 12: 195–243.
- Georg, Stefan, 2007. *A Descriptive Grammar of Ket (Yenisei-Ostyak)*, part 1: *Introduction, Phonology, Morphology*. Folkestone: Global Oriental.

- Gil, David, 2005. Numeral classifiers. In Haspelmath et al. (eds), pp. 226–229.
- Göksel, Aslı and Celia Kerslake, 2005. *Turkish: A Comprehensive Grammar*. London: Routledge.
- Gorelova, Liliya, 2002. *Manchu Grammar*. Leiden: Brill.
- Grönbech, Karl, 1936. *Der Türkische Sprachbau*. Copenhagen: Levin and Munksgaard.
- Gruzdeva, Ekaterina, 1998. *Nivkh*. Languages of the World/Materials, vol. 111. Munich: LINCOM.
- Hashimoto, Mantaro, 1986. The Altaicization of Northern Chinese. In John McCoy and Timothy Light (eds), *Contributions to Sino-Tibetan Studies*, pp. 76–97. Leiden: Brill.
- Haspelmath, Martin, 1995. The converb as a cross-linguistically valid category. In Haspelmath and König (eds), pp. 1–56.
- Haspelmath, Martin et al. (eds), 2005. *The World Atlas of Language Structures*. Oxford: Oxford University Press.
- Haspelmath, Martin and Ekkehard König (eds), 1995. *Converbs in Cross-linguistic Perspective: Structure and Meaning of Adverbial Verb Forms – Adverbial Particles, Gerunds*. Berlin: Mouton de Gruyter.
- Heath, Jeffrey, 1998. Hermit crabs: Formal renewal of morphology by phonologically mediated affix substitution. *Language* 74: 728–759.
- Honti, László, 1997. Die Negation im Uralischen, parts I–III. *Linguistica Uralica* 2: 81–96, 161–176, 241–252.
- Janhunen, Juha, 1981. Korean vowel system in North Asian perspective. *Hangeul* 172: 129–146.
- Janhunen, Juha, 1982. On the structure of Proto-Uralic. *Finnisch-ugrische Forschungen* 44: 23–42.
- Janhunen, Juha, 1996. *Manchuria: An Ethnic History*. Mémoires de la Société Finno-Ougrienne, vol. 222. Helsinki: Suomalais-Ugrilainen Seura.
- Janhunen, Juha, 1997. Problems of primary root structure in pre-proto-Japanic. *International Journal of Central Asian Studies* 2: 14–30.
- Janhunen, Juha, 2007. Typological interaction in the Qinghai Linguistic Complex. *Studia Orientalia* 101: 85–103.
- Janhunen, Juha, 2009. Proto-Uralic: What, where, and when? *Mémoires de la Société Finno-Ougrienne* 258: 57–78.
- Janhunen, Juha, 2012. *Mongolian*. Amsterdam: John Benjamins.
- Janhunen, Juha, 2013. Personal pronouns in Core Altaic. In Robbeets and Cuyckens (eds), pp. 211–226.
- Johanson, Lars, 1995. On Turkic converb clauses. In Haspelmath and König (eds), pp. 313–348.
- Johanson, Lars and Martine Robbeets, 2010. Introduction. In Lars Johanson and Martine Robbeets (eds), *Transeurasian Verbal Morphology in a Comparative Perspective: Genealogy, Contact, Chance*, pp. 1–5. *Turcologica*, vol. 78. Wiesbaden: Harrassowitz.

- Ko, Seongyeon, John Whitman and Andrew Joseph, 2014. Comparative consequences of the tongue root harmony analysis for proto-Tungusic, proto-Mongolic, and proto-Korean. In Martine Robbeets and Walter Bisang (eds), *Paradigm Change in the Transeurasian Languages and Beyond*, pp. 141–176. Amsterdam: John Benjamins.
- Lee, Ki-Mun and Robert Ramsey, 2011. *A History of the Korean Language*. Cambridge: Cambridge University Press.
- Li, Charles N. and Sandra A. Thompson 1989. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley, CA: University of California Press.
- Frellesvig, Bjarke, 2010. *A History of the Japanese Language*. Cambridge: Cambridge University Press.
- Kaiser, Stefan, Yasuko Ichikawa, Noriko Kobayashi and Hirofumi Yamamoto, 2001. *Japanese: A Comprehensive Grammar*. London: Routledge.
- Ko, Seongyeon, 2012. *Tongue Root Harmony and Vowel Contrast in Northeast Asian Languages*. PhD dissertation, Cornell University, New York.
- Kortlandt, Frederik, 2004. Nivkh as a Uralo-Siberian language. In Adam Hyllested, Anders Richardt Jørgensen, Jenny Helena Larsson and Thomas Olander (eds), *Per aspera ad asteriscos: Festschrift in Honour of Jens E. Rasmussen*, pp. 285–289. Innsbruck: IBS.
- Maddieson, Ian, 2005. Tone. In Haspelmath et al. (eds), pp. 58–61.
- Malchukov, Andrej, 2012. Tungusic converbs and a typology of taxis. In Andrej Malchukov and Lindsay J. Whaley (eds), *Recent Advances in Tungusic Linguistics*, pp. 213–228. Turcologica, vol. 89. Wiesbaden: Harrassowitz.
- Malchukov, Andrej, 2013. Verbalization and insubordination in Siberian languages. In Robbeets and Cuyckens (eds), pp. 177–208.
- Martin, Samuel Elmo, 1988. *A Reference Grammar of Japanese*. Tokyo: Tuttle.
- Martin, Samuel Elmo, 1992. *A Reference Grammar of Korean*. Tokyo: Tuttle.
- Maslova, Elena, 2003a. *A Grammar of Kolyma Yukaghir*. Mouton Grammar Library, vol. 27. Berlin: Mouton de Gruyter.
- Maslova, Elena, 2003b. *Tundra Yukaghir*. Languages of the World/Materials, vol. 372. Munich: LINCOM.
- Miestamo, Matti, 2005. *Standard Negation: The Negation of Declarative Verbal Main Clauses in a Typological Perspective*. Empirical Approaches to Language Typology, vol. 31. Berlin: Mouton de Gruyter.
- Murayama, Shichirō, 1992. *Ainugo no kigen [Origins of the Ainu Language]*. Tokyo: San'ichi Shobo.
- Nedjalkov, Igor V., 1994. Negation in Evenki. In Peter Kahrel and René van den Berg (eds), *Typological Studies in Negation*, pp. 1–34. Amsterdam: John Benjamins.
- Nedjalkov, Igor V., 1995. Converbs in Evenki. In Haspelmath and König (eds), pp. 441–464.
- Nedjalkov, Igor V., 1997. *Evenki: Descriptive Grammar*. London: Routledge.

- Neuskaya, Irina, 2010. Inclusive and Exclusive in Altaic Languages.
In Lars Johanson and Martine Robbeets (eds), *Transeurasian Verbal Morphology in a Comparative Perspective: Genealogy, Contact, Chance*, pp. 115–128. *Turcologica*, vol. 78. Wiesbaden: Harrassowitz.
- Nichols, Johanna, 2012. Selection for M : T pronominals in Eurasia.
In Lars Johanson and Martine Robbeets (eds), *Copies versus Cognates in Bound Morphology*, pp. 47–70. *Brill's Studies in Language, Cognition and Culture*, vol. 2. Leiden: Brill.
- Nichols, Johanna and David Peterson, 2005. Personal pronouns.
In Haspelmath et al. (eds), pp. 546–553.
- Nikolaeva, Irina, 1999. *Ostyak*. *Languages of the World/Materials*, vol. 305. Munich: LINCOM.
- Norman, Jerry, 1988. *Chinese*. *Cambridge Language Surveys*. Cambridge: Cambridge University Press.
- Payne, John R., 1985. Negation. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*, vol. 1: *Clause Structure*, pp. 197–242. Cambridge: Cambridge University Press.
- Poppe, Nicholas, 1954. *Grammar of Written Mongolian*. Wiesbaden: Harrassowitz.
- Poppe, Nicholas, 1955. *Introduction to Mongolian Comparative Studies*. *Mémoires de la Société Finno-Ougrienne*, vol. 110. Helsinki: Suomalais-Ugrilainen Seura.
- Poppe, Nicholas, 1964. Der altaische Sprachtyp. In B. Spuler et al. (eds), *Mongolistik*, pp. 1–16. *Handbuch der Orientalistik*, vol. 5.2. Leiden: Brill.
- Rickmeyer, Jens, 1989. Japanisch und der altaische Sprachtyp. Eine Synopsis struktureller Entsprechungen. *Bochumer Jahrbuch zur Ostasienforschung* 12: 313–323.
- Robbeets, Martine, 2005. *Is Japanese Related to Korean, Tungusic, Mongolic and Turkic?* *Turcologica*, vol. 64. Wiesbaden: Harrassowitz.
- Robbeets, Martine, 2007a. How the actional suffix chain connects Japanese to Altaic. *Turkic Languages* 11 (1): 3–58.
- Robbeets, Martine, 2007b. The causative-passive in the Trans-Eurasian languages. *Turkic Languages* 11 (2): 235–278.
- Robbeets, Martine, 2008. If Japanese is Altaic, why is it so simple?
In Alexander Lubotsky, Jos Schaeken and Jeroen Wiedenhof (eds), *Evidence and Counter-evidence: Essays in Honour of Frederik Kortlandt*, vol. 2: *General Linguistics*. *Studies in Slavic and General Linguistics*, vol. 33. Amsterdam: Rodopi.
- Robbeets, Martine, 2009. Insubordination in Altaic. *Journal of Philology*, vol. 31. *Ural-Altaic Studies* 1: 61–79.
- Robbeets, Martine, 2010. Transeurasian: Can verbal morphology end the controversy? In Lars Johanson and Martine Robbeets (eds), *Transeurasian Verbal Morphology in a Comparative Perspective: Genealogy, Contact, Chance*, pp. 81–114. *Turcologica*, vol. 78. Wiesbaden: Harrassowitz.

- Robbeets, Martine, 2012. Shared verb morphology in the Transeurasian languages: Copy or cognate? In Lars Johanson and Martine Robbeets (eds), *Copies vs. Cognates in Bound Morphology*, pp. 427–446. Brill's Studies in Language, Cognition and Culture, vol. 3. Leiden: Brill.
- Robbeets, Martine, 2013. Genealogically motivated grammaticalization. In Robbeets and Cuyckens (eds), pp. 147–175.
- Robbeets, Martine, 2014. The development of negation in the Transeurasian languages. In Pirkko Suihkonen and Lindsay J. Whaley (eds), *On Diversity and Complexity of Languages Spoken in Europe and North and Central Asia*, pp. 401–420. Studies in Language Companion Series, vol. 164. Amsterdam: John Benjamins.
- Robbeets, Martine, 2015. Diachrony of Verb Morphology: Japanese and the Transeurasian Languages. Trends in Linguistics Studies and Monographs, vol. 291. Berlin: Mouton de Gruyter.
- Robbeets, Martine, 2016. Insubordination and the establishment of genealogical relationship. In Nicholas Evans and Honore Watanabe (eds), *Insubordination*, pp. 209–246. Typological Studies in Language, vol. 115. Amsterdam: John Benjamins.
- Robbeets, Martine and Hubert Cuyckens (eds), 2013. *Shared Grammaticalization: with Special Focus on the Transeurasian Languages*. Studies in Language Companion Series, vol. 132. Amsterdam: John Benjamins.
- Rybatzki, Volker, 2003. Middle Mongol. In Juha Janhunen (ed.), *The Mongolic Languages*, pp. 57–82. London: Routledge.
- Shibatani, Masayoshi, 1990. The languages of Japan. Cambridge Language Surveys. Cambridge: Cambridge University Press.
- Sohn, Ho-min, 1994. *Korean*. London: Routledge.
- Sohn, Ho-min, 2009. The semantics of clause linking in Korean. In R. M. W. Dixon and Alexandra Aikhenvald (eds), *The Semantics of Clause Linking: A Cross-linguistic Typology*, pp. 285–317. Oxford: Oxford University Press.
- Stassen, Leon, 1997. *Intransitive Predication*. Oxford: Clarendon Press.
- Stassen, Leon, 2005a. Predicative possession. In Haspelmath et al. (eds), pp. 474–477.
- Stassen, Leon, 2005b. Predicative adjectives. In Haspelmath et al. (eds), pp. 478–481.
- Stassen, Leon, 2005c. Comparative constructions. In Haspelmath et al. (eds), pp. 490–493.
- Street, John, 1957. The Language of the Secret History of the Mongols. New Haven, CT: American Oriental Society.
- Suihkonen, Pirkko, 2002. The Uralic languages. *Fennia* 180 (1/2): 165–176.
- Svantesson, Jan-Olof, 1985. Vowel harmony shift in Mongolian. *Lingua* 67 (4): 283–327.
- Tamura, Suzuko, 2000. *The Ainu Language*. ICHEL Linguistic Studies, vol. 2. Tokyo: Sanseidō.

- Vajda, Edward J., 2004. *Ket. Languages of the World/Materials*, vol. 204. Munich: LINCOM.
- Vovin, Alexander, 1993. *A Reconstruction of Proto-Ainu*. Brill's Japanese Studies Library, vol. 4. Leiden: Brill.
- Vovin, Alexander, 2005. *A Descriptive and Comparative Grammar of Western Old Japanese*, part 1: *Sources, Script and Phonology, Lexicon, Nominals*. Languages of Asia, vol. 3. Folkestone: Global Oriental.
- Vovin, Alexander, 2009. *A Descriptive and Comparative Grammar of Western Old Japanese*, part 2: *Adjectives, Verbs, Adverbs, Conjunctions, Particles, Postpositions*. Languages of Asia, vol. 8. Folkestone: Global Oriental.
- Weiers, Michael, 1966. *Untersuchungen zu einer Historischen Grammatik des Prällassischen Schriftmongolisch*. PhD dissertation, Rheinischen Friedrich-Wilhelms-Universität, Bonn.
- Werner, Heinrich, 1997. *Die ketische Sprache*. Wiesbaden: Harrassowitz.
- Whaley, Lindsay J. and Fengxiang Li, 2000. *Emphatic reduplication in Oroqen and its Altaic context*. *Linguistics* 38 (2): 355–372.
- Whitman, John Bradford, 1990. *A rule of medial *-r- loss in pre-Old Japanese*. In Philip Baldi (ed.), *Linguistic Change and Reconstruction Methodology*, pp. 511–545. Trends in Linguistics: Studies and Monographs, vol. 45. Berlin: Mouton de Gruyter.
- Wohlgemuth, Jan, 2009. *A Typology of Verbal Borrowings*. Berlin: Mouton de Gruyter.
- Wrona, Janick, 2008. *The nominal and adnominal forms in Old Japanese: Consequences for a reconstruction of pre-Old Japanese syntax*. In Bjarke Frellesvig and John Whitman (eds), *Proto-Japanese: Issues and Prospects*, pp. 193–215. Current Issues in Linguistic Theory, vol. 294. Amsterdam: John Benjamins.
- Yap, Foong Ha and Stephen Matthews, 2008. *The development of nominalizers in East Asian and Tibeto-Burman languages*. In María José López-Couso and Elena Seoane (eds), *Rethinking Grammaticalization: New Perspectives*, pp. 309–341. Amsterdam: John Benjamins.
- Zeitoun, Elizabeth, 2007. *A Grammar of Mantauryan (Rukai)*. Taipei: Institute of Linguistics, Academia Sinica.

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.