Chapter 45. Transeurasian as a Continuum of Diffusion   /Draft of January 13, 2018/
Edward Vajda

Abstract. Intermingling of Turkic, Mongolic, and Tungusic speakers over many centuries left multiple overlapping layers of contact-induced language change in their wake. While the dynamics of pastoralist mobility spread linguistic traits far and wide, it remains unresolved whether contact alone (together with coincidental resemblance) can account for all of the shared features in the families traditionally grouped as "Altaic", or whether some homologies represent evidence of deeper common ancestry. Without arguing strongly for or against either possibility, this article considers how typological parallels may have diffused among pastoral Inner Eurasia's four autochthonous families – Uralic, Turkic, Mongolic, and Tungusic – and also into Yeniseian, Yukaghir, Chukchi-Kamchatkan, Nivkh, Ainu, Koreanic, and Japonic – families and isolates that interacted less pervasively with steppe and forest pastoralists.

1. Introduction

While debate continues unabated over whether Turkic, Mongolic, and Tungusic constitute a genealogical unity either together or as separate branches of a larger family, the study of language contact among Inner Eurasia's indigenous languages has proven a highly fruitful avenue for historical-linguistic research. A key question in supporting or rejecting the existence of an Altaic, Macro-Altaic, or Transeurasian language family is how to separate contact effects from plausible evidence of deeper genealogical relatedness. Given the mobility of pastoral populations and their tendency to absorb or otherwise interact with neighboring groups, identifying new facts about language contact, borrowing, and diffusion can make a particularly relevant contribution toward resolving issues of genealogical classification in this area of the world. The present article focuses on the issue of diffusion among the indigenous languages of Inner and North Asia. Section Two surveys a number of typological traits shared not only between Turkic, Mongolic and Tungusic, but also with Uralic, the area's other widespread pastoral language family. To what degree are Transeurasian linguistic traits also found in languages of the Uralic family? The hypothesis of Ural-Altaic as a genetic unity has been universally abandoned, mainly because Uralic itself has been clearly demonstrated as a family, so that similarities between Uralic and Turkic, Mongolic and Tungusic are now viewed as resulting from language contact, even where the timing and mechanism of that contact has not been clearly established. If Transeurasian languages share multiple diffused traits with the unrelated Uralic, however, what evidence is there that the traits they share among themselves are not likewise due to contact? Section Three broadens the typological comparison to assess the degree to which the so-called Paleosiberian or Paleo-Asiatic families and isolates have absorbed Ural-Altaic linguist features through borrowing and diffusion. Linguistic features borrowed by these unrelated languages from Uralic, Turkic, Mongolic or Tungusic pastoral groups are usually easily identified and provide a new perspective on the historical reasons and structural outcomes of linguistic diffusion in Inner Eurasia. Section Four extends the same analysis to Koreanic and Japonic, both of which share more extensive similarities with the three traditional Altaic families than is true of other Asian languages, despite being spoken outside the main Inner forest-steppe

diffusion zone. The task of demonstrating whether these similarities are plausibly due to genetic inheritance or result instead from early mutual contact in Manchuria or the Korean peninsula is identified as a key research objective. Section Five summarizes how the known evidence of contact between and beyond Transeurasian languages contributes to the debate about their genetic relatedness, and also suggests additional directions for future investigation.

2. Inner and Northern Eurasia as a language area

Though still often useful as shorthand designations for disparate groups of peoples or languages, the terms "Paleosiberian", "Paleo-Asiatic", "Ural-Altaic" and even the word "Altaic" itself are the residue of an earlier phase of historical linguistics that did not fully differentiate between linguistic relatedness and shared cultural traits, nor between similarities arising from contact and those due to genetic inheritance. The first two terms have never been more than a handy way to group together several unrelated isolates or microfamilies in order to set them apart from the major pastoral groups of northern and interior Asia and their widely spoken languages. As far as concerns the latter two terms, so far only Uralic has proven to be a demonstrated family, while the issue of whether Altaic (in either its micro or macro version) is also a valid genealogical unit remains a matter of serious disagreement among historical linguists. Janhunen (2014) has argued convincingly that many if not most "Altaic" typological traits are also shared with Uralic and appear to be the result of a series of overlapping zones of diffusion rather than genuine evidence of genealogical relationship. Given this historiographic background, it is worthwhile assessing the extent to which Turkic, Mongolic and Tungusic actually share typological traits that clearly set them apart within the broader Ural-Altaic zone. If Uralic shares many striking typological features with Altaic languages because of early contact, the same features could plausibly be shared between the three Altaic families likewise because of contact. This section examines the degree to which this is arguably true and also highlights homologies between Turkic, Mongolic and Tungusic that are absent from Uralic and cannot be so easily identified as resulting from language contact.

Languages from Finland to Mongolia and historical Manchuria, and sometimes further east to Korea and Japan, share a broad typological similarity involving SOV word order, postpositions, and the exclusive use of suffixes in both nominal and verbal inflectional morphology. This includes the use possessive suffixes, in sharp contrast to the use of possessive prefixes in Yeniseian languages, located north of the forest-steppe zone. There is also the prevalence of root-to-suffix vowel harmony. Because Mongolic, Tungusic, Koreanic and Japonic were likely spoken in a contiguous or near contiguous zone in Manchuria and the Korean Peninsula (Janhunen 2014), it is plausible that the similarities could be due either to contact or shared genetic inheritance, or to a combination of both. There is no firm evidence, however, that Proto-Turkic was spoken farther to the east than western or central Mongolia or adjacent parts of southern Siberia. The Uralic homeland is likely to have been still farther west, somewhere in the forest-steppe zone near the Ural Mountains. It remains unclear how Uralic and Turkic share a contiguous geographic point of origin with eastern Transeurasian languages. Similarly unclear is the geography of language contact that could be responsible for similarities shared between Uralic and Turkic, on the one hand, and the eastern Transeurasian languages, on the other.

This article will consider a new idea to help explain the diffusion of Ural-Altaic features from west to east. These features could have spread from an early branch of Uralic, now extinct, that once extended far to the east, into central Mongolia, from where it influenced both Turkic

and Mongolic through contact. Possible evidence of this comes from several sources. First, there appear to be Uralic loans into Early Mongolic that cannot otherwise be easily explained. These include Proto-Uralic *kele 'tongue', 'language' (Rédei 1988: 144), which yields Khalkha Mongolian *xel* 'tongue', 'language', and Finno-Ugric *kućз 'birch tree' (Rédei 1988: 211), which is a plausible source for loanwords into Mongolic (cf. modern Buryat *xuh* 'birch tree'). The same Uralic word also appears to have been borrowed into Proto-Yeniseian to give Arin *kus* and Northern Ket *ù:se* 'birch tree'. This word for 'birch tree' is absent from the Khanty and Samoyedic (Selkup, Enets) neighbors of the Yeniseians, so the source could only have been a different, unattested eastern branch of Uralic. The Yeniseian homeland was located in the area of Siberia between north-central Mongolia, the Yenisei headwaters, and the southern tip of Lake Baikal. This area was not contiguous with the Manchurian homeland of Mongolic, so that shared diffusion must have come from a third language occupying the intervening territory of northern and eastern Mongolia itself.

Though extant documentation of Xiongnu words from Chinese records is too sparse to relate them convincingly to any known language family, despite much debate on the subject, it is worth considering how indirect evidence from language contact might shed light on the mystery of Xiongnu linguistic identity. While there is some evident that the Jie (Kjet) people in the later Xiongnu Confederation spoke a Pumpokolic variety of Yeniseian (Vovin, de la Vaissière & Vajda 2016), it is unlikely that the ethnic core of this influential and widespread political entity spoke a Yeniseian language, given the complete lack of Yeniseian influence on any Transeurasian languages outside the immediate area of south Siberia (Yeniseian influence on South Siberian Turkic is relatively late). The possibility that an extinct branch of Uralic was spoken within the Xiongnu Confederation as the politically dominant language could however explain the diffusion of Uralic typological traits and lexical items into early Turkic and Mongolic (as well as into Yeniseian). A Uralic hypothesis of Xiongnu origins would go far in explaining many shared aspects of these languages that cannot otherwise be accounted for as resulting from contact yet do not seem plausibly due to genetic inheritance either. The presence in Mongolia by about 2400 years ago of Y-dna haplotype N1c (Keyser-Tracqui et al. 2004), which is found in all Uralic-speaking populations, also supports the presence of Uralic speakers in this area. Uralic involvement in the Xiongnu Confederacy was already suggested by Di Cosmo (2002: 166), though the possibility has never been seriously investigated from a linguistic perspective. The idea that an early Uralic linguistic presence (Xiongnu Era or earlier) in the area of present-day Mongolia influenced Turkic, Mongolic and Tungusic (as well as Yeniseian, to the north) could still be correct even if the ethnic core of the Xiongnu polity spoke a completely different language and the Chinese attestation of Xiongnu words are not themselves of Uralic origin.

In his survey of possible homologies in Uralic-Altaic nominal inflectional morphology, Janhunen (2014: 327) discusses instances where Uralic and Turkic pattern together, in contrast to other languages of the Transeurasian zone. These include the tendency for case systems to contain three basic local cases (dative, locative, ablative) in Uralic and Turkic, but only two (dative-locative and ablative) in the eastern Transeurasian languages. Janhunen (2014: 331) further notes that Uralic shares a greater number of typological parallels with the three core Altaic families of Turkic, Mongolic and Tungusic than with Koreanic and Japonic. This pattern could also be explained, at least in part, by Xiongnu Era influence from a Uralic language, which did not extend far enough eastward to strongly affect the latter two families.

Another factor to consider is vowel harmony. While the typological presence of vowel harmony is prevalent throughout Inner Eurasia, the type found in Uralic and Turkic differs from

that of Mongolic and Tungusic. The latter two families have tongue-root harmony, with tongue height (actually, tongue root advancement or retraction) being the underlying mechanism. This type has traditionally been called dominant-recessive, which is an infelicitous term in that the "recessive" set of vowels /i, u, e/ are made by advancing the tongue root in relation to the "dominant set" /e, o, a/, for which the tongue root is retracted. Uralic, on the other hand, has palatal (front-back) harmony, as does Turkic. Early Turkic could have acquired palatal harmony from Uralic during Xiongnu times (2300 – 1900 years before present), while the tongue root height systems of Mongolic and Tungusic could be evidence of an earlier genealogical unity or could have spread somehow through early contact between these families. As with certain homologies in nominal inflection, the feature of vowel harmony across Transeurasian (or Ural-Altaic) languages appears to involve two different, but overlapping features and not a single feature that unites the entire group. The two different types of vowel harmony could later have undergone what Janhunen (2014) calls "shared drift", in other words, becoming typologically more similar due to contact.

The next section examines diffusion of these and other Ural-Altaic features into languages beyond the Transeurasian zone to argue that they too appear to have originated from two different regions of Inner Eurasia. The first is a Uralic point of origin in South Siberia, later overlapping with Turkic, and involving both steppe and forest pastoralism. The second is a Manchurian point of origin involving the spread of Tungusic-speaking reindeer breeders into northeastern Asia.

3. Pastoral Eurasia and the "Paleo-Asiatic" languages

Studying diffusion of linguistic features beyond Uralic, Turkic, Mongolic or Tungusic into the languages spoken by North Asia's far-flung remnant hunter-gatherers – Yeniseian, Yukaghir, Yupik, Chukchi-Kamchatkan, Nivkh, and Ainu – can give a fresh perspective on the issue how the same traits may have spread through contact among the pastoral peoples themselves. While the possibility that Uralic (in contrast to the five families subsumed under Transeurasian) might be related to either Yukaghir or even to Eskimo-Aleut has never been fully demonstrated nor conclusively rejected, the various "Paleosiberian" or "Paleo-Asiatic" languages are not seriously thought to be genealogically relatable to Transeurasian languages, except conceivably at some level too deep to be demonstrated using the Comparative Method. For this reason, considering the presence of shared traits between these six isolates or microfamilies and the four widespread families of Eurasia spoken by pastoral nomads has implications for understanding the Transeurasian diffusion zone.

3.1. Yeniseian

Yeniseian (Yeniseic, Yenisseian) is a microfamily containing Ket and the extinct Yugh, Kott, Assan, Arin and Pumpokol languages spoken by Inner Eurasia's last hunter-gatherers across a vast area stretching from northern Mongolia and the southwestern corner of Lake Baikal across much of western Siberia. The family is named after the Yenisei River, near which all known speakers of this family nomadized at the time when Russians entered central Siberia in the early 1600s. Though genealogically completely unrelatable to any Transeurasian language, modern Ket shows evidence of significant structural accommodation to the neighboring Uralic, Turkic and Tungusic languages (Vajda 2009, 2018a). The social mechanism behind this diffusion was

the induction of young brides from the surrounding pastoral peoples into the Yeniseian-speaking hunting bands, a process that continued for many centuries and probably began thousands of years ago. On the basis of the original Yeniseian prototype of a strongly prefixing templatic polysynthetic verb and possessive prefixes on nouns, modern Ket has restructured its morphology to resemble the suffixal agglutination found in all of the surrounding languages. All productive patterns of Ket verb form creation are now predominantly suffixing, with the original verb-final root having undergone grammaticalization as a marker of transitivity or aspect. A suffixing case system with strong semantic resemblances to that found in South Siberian Turkic has developed out of nested possessive constructions (Vajda 2013). Shared case features include a prolative (or prosecutive) to denote motion along or through, as well as an ablative, locative, dative, adessive, comitative-instrumental, and caritive (denoting lack or absence). Unlike either Turkic or Uralic, however, which have an overtly marked accusative case, Yeniseian nominal morphology does not in any way explicitly mark its nominative/accusative alignment, the difference between grammatical subject and object being instead expressed verb internally. Comparison of Ket-Yugh and Kott-Assan, the family two best documented primary branches, show that the suffixing case system, as well as the shift from prefixing to suffixing in the finite verb, was already under way in Proto Yeniseian and therefore goes back at least 2,500 years if not deeper into history, to the earliest interactions of the family with pastoral peoples of the Transeurasian zone. In phonology, Yeniseian shares with Turkic an aversion to word-initial sonorants. Phonetically, the lateral anlauts in Yeniseian words like Ket *laŋat* 'hand' or Yugh *lə'χ* 'dirt' were actually pronounced as fricative [ɬ] or affricate [tɬ] by native speakers before the spread of Russian as a universal first language. There is no evidence that any form of vowel harmony has spread to Yeniseian, though the phonemic tones found in Ket monosyllables have reduced to a word-initial pitch accent in polysyllables, possibly under the influence of the agglutinative character of the surrounding non-tonal languages.

There are relatively few loanwords into Yeniseian from Transeurasian languages. Also, no grammatical affixes in either the verbal or nominal morphology were borrowed. Instead, the restructuring of the morphology occurred entirely through grammaticalization or re-grammaticalization of native elements, as well as by surface pronunciation adjustments such as word-initial prefixes becoming enclitics on the preceding word. The important point here is that despite significant areal influence, the original typological profile and lexical stock of Yeniseian is still easily identifiable from internal reconstruction, as well as through external comparison with Na-Dene languages, a family with which Yeniseian appears to share a common genetic origin (Vajda 2018b). Yeniseian therefore provides an example of a language family where diffused Transeurasian features can clearly be distinguished from features inherited genetically. The study of modern Ket has also shown that attempts to understand the reasons for unusual combinations of typological features (such as the mix of prefixing and suffixing patterns in the finite verb) can lead to breakthroughs in understanding a language's contact history as well as its external genealogical relations.

3.2. Yukaghir

What is often called "the Yukaghir language" (or Yukagir, though more commonly spelled with 'gh' to indicate that the letter 'g' is pronounced as velar [g] rather than affricate [ʤ]) is actually a family (Yukaghiric) represented by two surviving members: Tundra (or Northern) Yukaghir and Kolyma (or Forest) Yukaghir. These two distinct but obviously genetically related languages are

spoken by a dwindling number of elders in remote parts of the Sakha Republic. The autonyms of the two groups are Wadul for Tundra Yukaghir and Odul for Forest Yukaghir, words that represent a native plural noun meaning 'mighty ones'. The origin of the exonym "Yukaghir" is unknown. When Russian adventurers and fur-tax collectors first entered the area in the 17th century, there were probably several other related languages spoken in northeastern Siberia. Unfortunately, diseases brought by Europeans, coupled with economic dislocation from tsarist fur tax extraction, decimated most Yukaghiric groups before their languages could be documented even rudimentarily. Though so little survives of this originally widespread family, much progress has recently been made in reconstructing Proto-Yukaghir (Nikolaeva 2006) and in documenting the considerable lexical differences that exist between the two surviving varieties. The idea of Uralo-Yukaghir as a genetic family has neither been demonstrated nor conclusively refuted and remains an active area of investigation. Aikio (2014) argues that the parallels reflect prehistoric Samoyedic-Yukaghir language contact, suggesting that there are genuine historical reasons for at least some of the similarities between Uralic and Yukaghir, if only of an areal nature. The study of Yukaghir, therefore, continues to be germane to Ural-Altaic (and Transeurasian) linguistics.

Compared to the other "Paleo-Asiatic" languages, Yukaghir shares the largest number of basic typological traits with Ural-Altaic, which also suggests that the family's history of contact and possibly deeper genealogical relationship with the larger families of Inner Eurasia should be further investigated. Borrowing, possibly from now extinct languages, is likely to be implicated in the strong lexical differences observed between the family's two surviving members. Widespread Inner or North Eurasian typological traits include a predominantly suffixing-agglutinative morphology, SOV word order, the use of postpositions rather than prepositions, and a relatively simple phoneme system. Yukaghir phonology has both palatal and labial harmony, which spreads from the first vowel of the stem. This type of vowel harmony patterns with Uralic (and Turkic) rather than Tungusic, where harmony is instead based on tongue root height. The prevalence of disyllabic stems that permit a variety of internal consonant clusters is also reminiscent of Uralic. The lack of a straightforward nominative/accusative distinction in the nominal morphology, however, puts Yukaghir in sharp contrast with Uralic and Turkic and more resembles the situation in Mongolic. Another un-Uralic feature of Yukaghir is that adjectives represent a sub-class of verbs rather than a morphologically distinct part of speech, a feature also characteristic of Korean. Yukaghiric therefore is interesting in showing a mix of typological features; some of these (palatal vowel harmony) are shared with Uralic and possibly have a historical explanation (either ancient contact or genetic relationship), while others are found in eastern Transeurasian languages and could be due either to coincidence or, in theory, to some very ancient genealogical connection. Regardless of which explanation is correct, the study of Paleosiberian languages provides a reminder of why it is useful to consider the presence of shared Transeurasian typological traits beyond the primary Inner Asian diffusion zone.

One potentially fruitful prospect for learning more about Yukaghir linguistic prehistory would be to investigate the origin of its unique morphological system of focus marking, which involves a complex interplay between case suffixes on subjects and objects and a focus prefix with agreement suffix on the verb form itself (Comrie 1980: 258-261; Maslova 2003a: 9-10; Maslova 2003b: 6-10). Because this system is highly distinctive both formally and functionally, using internal reconstruction to trace how it arose could shed light on the evolution of Yukaghir verb morphology and case marking. This in turn might reveal more about areal influence from

the family's external contacts and possibly even its deeper genealogical relationships, as the original Yukaghir typological profile becomes clearer.

3.3. Chukchi-Kamchatkan

Chukchi-Kamchatkan (sometimes alternatively called Kamchukotian or Kamchukotic) is the autochthonous family of the Kamchatka Peninsula and the Chukchi District – the extreme northeast portion of the Russian Federation. One branch of the family contains Chukchi and Koryak (also spelled Korak), as well as Kerek and Alyutor – distinct languages that during Soviet times were treated as Koryak dialects. All four of these language forms are closely related and stand in significant contrast to the southern branch of the family, which contains the sole surviving Western Itelmen language. Two other Itelmenic languages (Northeastern and Southern) disappeared with no documentation aside from rudimentary vocabulary lists. Lexical and typological differences between the two branches of the family are such that Itelmen has sometimes been regarded as genealogically unrelated to Chukchi-Koryak (Georg & Volodin 1999). However, as in the case of failed challenges to the status of Uralic as a valid family (Marcantonio 2002), Chukchi-Kamchatkan undoubtedly represents a genetic unity. The family's internal typological dissonance therefore provides a valuable laboratory for studying the effects of differential language contact.

     The stark differences between the two branches appear to be due not to extreme time depth but rather to sharply different contact histories affecting each branch. While the prehistoric language contacts involving Itelmen are probably unrecoverable, the Chukchi-Koryak branch shows evidence of contact from two different sources. The first is from an Eskimoan substrate (about which see section 3.4 below), while the second involve Tungusic or another eastern Transeurasian source, probably connected with the adoption of reindeer breeding. Neither type of contact reached the early Itelmen-speaking populations of Southern Kamchatka. Chukchi-Koryak shows the same tongue-height based vowel harmony found in Tungusic and Mongolic. Because Chukchi-Kamchatkan languages have prefixes as well as suffixes, the harmony triggered by the root vowel spreads leftward as well as rightward in the phonological word. Although Chukchi-Kamchatkan and Yukaghiric both share the broad typological feature of root-to-affix vowel harmony, the system in Chukchi-Kamchatkan appears to reflect diffusion from a Manchurian source, along with Tungusic and Mongolic, while the Yukaghiric system patterns with Uralic and Turkic, probably resulting from early contact somewhere in western or central Siberia.

3.4. Eskaleut

Though unlikely to have been influenced directly by even the most northern Transeurasian-speaking pastoralists, there are two reasons to include languages belonging to the Eskaleut (Eskimo-Aleut) family in any survey of diffusion and genetic relationship across Inner Eurasia. The first is the still unresolved question of whether this family shares some sort of ancient historical relationship, possibly genetic, with Uralic (Seefloth 2000). Second, languages belonging to the family's Eskimoan branch appear to form a substrate in the northern branch of Chukchi-Kamchatkan due to absorption of Yupik (or other Eskimoan) speakers by Chukchi- or Koryak-speaking coastal groups. Fortescue (2004: 159) provides evidence that Eskimoan speakers once occupied Asia's North Pacific coast from Bering Strait as far south as the isthmus

of the Kamchatkan Peninsula. Understanding the history and diversification of Chukchi-Kamchatkan languages therefore requires consideration of possible effects from this contact, which was still ongoing when Russians penetrated this area in the 18th century.

While most Eskimoan languages are spoken in North America, three varieties were documented in the extreme northeast of Asia, on territory that today is part of the Chukchi District of the Russian Federation. This area still contains pockets of Central Siberian Yupik speakers, who likely migrated back across Bering Strait from Alaska in late prehistory, as well as several communities speaking the more distantly related Naukan Yupik. More intriguing from the perspective of language classification is the recently extinct Sirenik, also spoken on the coast of Chukchi Peninsula. Though sometimes regarded as a highly aberrant variety of Yupik, Sirenik may actually represent a third primary branch of Eskimoan apart from Yupik and languages of the Inuit. Finally, a variety of Iñupiaq was spoken on Big Diomede Island in Bering Strait until the late 1940s when the island's inhabitants were relocated to the Soviet mainland. Even providing for the possibility that Sirenik is an aberrant member of the Yupik sub-branch, the overall dialectal diversity of Eskimoan languages on the Asian side of Bering Strait is clearly deeper than that found in Alaska, Canada and Greenland combined (Fortescue, Jacobson & Kaplan 2010). This fits with the idea that ancestral Eskimoan was formerly spoken more widely in northeast Asia and that the Chukchi (and possibly Koryak) reindeer herders and coastal hunters absorbed earlier speakers of some variety (or varieties) of Eskimoan in late prehistory. Women from Eskimoan groups who were assimilated into Chukchi-speaking populations may conceivably have been the historical source of the distinct female pronunciation of certain Chukchi phonemes, a feature that persisted into the 20th century.

Fortescue (1997, 2011) provides a good account of parallels between the two families that appear to be attributable to Eskimoan contact effects on the Chukchi-Koryak branch of the family. Because this contact did not involve Itelmen, at least some of the divergence between the two branches of Chukchi-Kamchatkan can be attributed to contact with Eskimoan by the latter. For example, ergative alignment seems to have diffused from Eskimoan into Chukchi-Koryak, while being completely absent from Itelmen. If this is true, the Chuckhi-Koryak branch of Chukchi-Kamchatkan provides an example of alignment change via diffusion – something instructive to any approach to historical linguistics seeking to claim that fundamental typological alignment cannot be altered through language contact. An explanation of diffusion (or coincidence) is therefore similarly plausible as regards the nominative/accusative alignment shared widely among Eurasian languages, including Transeurasian. The presence of noun phrase incorporation in the finite verb of Chukchi-Koryak but not of Itelmen, however, remains unexplained, since polysynthesis in Eskaleut languages is of a completely different type.

3.5. Nivkh

The Nivkh (Gilyak) language isolate consists of four closely related dialectal forms spoken on northern Sakhalin Island or adjacent parts of the Asian mainland, on the lower reaches of the Amur River. There are three reasons to consider Nivkh data as part of a broader study of Transeurasian languages. First, Nivkh has sometimes been claimed as related to Transeurasian (or Ural-Altaic) languages in a much larger family, though the evidence is not convincing. The stark typological divide between Nivkh and Tungusic (or even between Nivkh and Yukaghiric) alone would suggest there is no genealogical relationship. Second, Nivkh speakers interacted with Tungusic speakers on the lower Amur, so the possible effects of language contact (Nivkh

substrate affects in Maritime Tungusic or Tungusic diffusion into Nivkh) should be investigated. Finally, there is the possibility that Nivkh is related to Chukchi-Kamchatkan (Fortescue 2011), so that comparisons between the two languages might provide valuable external data to help clarify the origin of submerged morphological features in each of them, thus providing a clearer basis for tracing later episodes of language contact and borrowing.

Nivkh displays certain features not present in any other North Asian language, including a special type of polysynthetic compounding that has been called "dependent-head synthesis". This is a pervasive morphological pattern in the language that governs the structure of both nouns and verbs and also triggers a complex series of alternations between voiced or unaspirated stops, aspirated stops, and voiced or voiceless fricatives (Gruzdeva 1999: 13-15). Another unique Nivkh trait is that basic number words, which appear postposed to the noun they quantify, have multiple allomorphs created by fusing a numeral root with a classifier (Gruzdeva 1999: 23-25). The presence of numeral classifiers in Nivkh patterns typologically with East Asian languages (including Japanese) rather than with continental Transeurasian languages. The rendaku (sequential voicing) pattern observed in Japanese compound words also superficially resembles the consonant mutations found inside of Nivkh dependent-head synthetic compounds, though with nothing approaching the phonological complexity of the Nivkh system. Any parallels between Nivkh and Japanese, however, would likely be due to some ancient areal influence, if not simply to coincidence. Internal reconstruction of unusual morphological features should be attempted for Nivkh, which might help compensate for the shallow depth of the language's dialectal diversity in attempts to link the language to other families.

Finally, Nivkh lacks the pervasive root-to-affix vowel harmony found across much of Inner Asia. However, possessor and undergoer pronominal prefixes do have vowel-harmonic allomorphs involving alternations between zero, /i/ and /e/ (Mattissen 2017: 892) that appear to be triggered by the tongue height of the vowel of the root to which they attach. Gruzdeva (1999: 11) claims this is the remnant of a more pervasive system. Another possible explanation is that this marginal feature of Nivkh developed through areal influence from Tungusic languages. However, if Nivkh proves to be related to Chukchi-Kamchatkan, the tongue-height harmony found in Nivkh and Chukchi-Koryak might prove instead to have an ancient common origin. In either case, Nivkh provides another example of tongue-root harmony in eastern Asia, which contrasts to the palatal harmony found in the western portion of the Ural-Altaic zone.

3.6. Ainu

The Ainu language isolate was once spoken in several dialectal forms on northern Honshu, as well as Hokkaido, southern Sakhalin, and the Kurile Islands. The northern Japanese islands would appear to be the original Ainu homeland, so that contact between Nivkh and Ainu in central and southern Sakhalin is probably a relatively late development. There is no evidence to suggest that Nivkh and Ainu are related in a larger family. Past attempts to include Ainu into an expanded Macro-Altaic family or into a still broader Eurasiatic phylum are unconvincing on any level. As with Nivkh, the typological differences between Ainu and the languages of the Transeurasian zone are simply too stark to make such a claim plausible. Nor does Ainu seem to share in any deeper contact-related patterns with Transeurasian languages. There is no evidence that Ainu was ever spoken on the Asian mainland, and the probable location of Proto-Japonic, the easternmost member of the Transeurasian zone, could not have been contiguous. Language contact between Japanese and Ainu is relatively late and has no direct bearing on understanding

early diffusion in the broader Transeurasian zone. Because Ainu was spoken beyond the borders of the Soviet Union, it was not generally included in studies of North Asian languages, despite the logic of including it in the "Paleo-Asiatic" group, given that the Ainu were traditionally hunter-gatherer-fishers with no domesticated animal except the dog. Nevertheless, Ainu should be included in any broad study of North and Inner Asian language families for its potential to contribute to an overall understanding of the area's typological patterns.

Like Nivkh, Ainu is a polysynthetic language (Bugaeva 2017), in sharp typological contrast to all Ural-Altaic/Transeurasian languages. However, Ainu verb structure is completely different from that of Chukchi-Kamchatkan, Ket, or Nivkh, and no plausible genetic or areal parallels can be identified to link Ainu with any of these languages. Ainu does share with most other languages of North Asia the typological features of basic SOV word order, postpositions, and lack of grammatical gender. Ainu also has a relatively simple phonology, but lacks a vowel harmony system of any kind – either functioning or vestigial. There is no reason to pursue the idea of linking Ainu with Transeurasian languages either areally or genetically except as regards relatively late Japanese influence on the Ainu lexicon (Vovin 1993).

4. Korean and Japanese

Research aimed at adding the two easternmost languages in the Transeurasian chain to a hypothetical Altaic (or Macro-Altaic) family came largely after the possibility of Ural-Altaic genetic unity had already been universally abandoned. Despite considerable debate on this topic, no genealogical link between these two languages either together or as part of a larger family is widely accepted. Most linguists recognize the two as language isolates, or refer to them as Koreanic and Japonic to emphasize their known dialectal or historical variation. The time depth of Koreanic is much more shallow than Japonic (or any other Transeurasian language group), dating to no earlier than 500 to 900 years (Janhunen 2014: 312). Because Japanese (or the broader Japonic) has probably been spoken in the Japanese archipelago for about two thousand years, any connection with Koreanic must date prior to the entry of Japanese farming culture into its present homeland. Areal diffusion or shared genetic origin between Koreanic and Japonic would presumably involve a location somewhere in the Korean peninsula or southern Manchuria. Generally, typological parallels between Japonic and Koreanic are found more widely in mainland Asia, though the overt marking of both nominative and accusative case on the noun phrase distinguishes Japonic and Koreanic from other languages in the Transeurasian zone. It cannot be discounted that parallels shared uniquely between Japonic and Koreanic diffused specifically from one to the other, or even from a third linguistic source that no longer exists. It is likely that Manchuria and the Korean peninsula were much more linguistically diverse before the spread of farming and pastoral cultures. There are also, of course, many shared linguistic featured imported into both Korean and Japanese during more than 1,500 years of cultural contact with Chinese civilization that have no bearing in a deeper understanding of genetic or areal connections in the Transeurasian zone.

In the realm of phonological typology, Middle Korean appears to have had root-to-affix vowel harmony, a feature widespread across Inner Eurasia. The Korean system has been described as either palatal or tongue-height harmony, with Robbeets (2005: 67) citing Miyake's (1999) more straightforward treatment of the phonology as a two-tiered tongue-height system. Unsurprisingly, this puts Korean in line with the "Manchurian" type of vowel harmony found in Tungusic and elsewhere in the eastern half of northern Asia, which suggests that this feature, not

clearly connectable to Japanese, could conceivably have arisen in Korean as the result of diffusion from Tungusic or other language contact in Manchuria.

Suffixal agglutination of both derivational and inflectional morphemes, which is pervasive in all Transeurasian languages including Koreanic and Japonic, offers the most promising object of study from which to seek paradigmatic morphological evidence in support of Transeurasian as a genetic family. The comparison of verb suffixes taken in Robbeets (2014) represents a fruitful approach. At the very least, her findings could lead to a more refined internal reconstruction of each of the six affixes claimed to be cognate across all of the putative branches of Transeurasian. At the same time, Janhunen (2014) has already convincingly shown that plausible cognate paradigms involving case, number or person agreement suffixes are lacking in Ural-Altaic languages. This result diminishes the likelihood that Transeurasian will ever be demonstrated as a valid genetic family, though certain submerged features of noun classification shared by Mongolic and Tungusic continue to offer a genuine prospect for uncovering a specific genetic link between these two families.

5. Future prospects for Transeurasian studies

This article surveyed evidence of linguistic diffusion among Uralic, Turkic, Mongolic, Tungusic, Koreanic and Japonic and also examined the degree to which some of the same features have spread to the smaller language families and isolates of North Asia. It argued that vowel harmony in Inner Eurasia represents two, partly distinct features, each with a different zone of diffusion, rather than a single unified typological trait. Vowel harmony based on tongue height (or tongue root advancement vs. retraction), which is prevalent in the eastern parts of northern and Inner Asia, appears to have spread from Manchuria. Conversely, the palatal harmony that characterizes Uralic and Turkic may have originally diffused from Uralic. Finally, indirect evidence from language contact was presented to introduce the hypothesis that a hitherto unrecognized Uralic superstrate dating to Xiongnu times or earlier (2500 to 1900 years ago) could have been responsible for the diffusion of Uralic features into Turkic and Mongolic, if not also into Transeurasian languages farther to the east.

Transeurasian undeniably represents a broad zone of diffusion – in fact, a zone characterized by overlapping points of diffusion, each with its own historical origin, some perhaps yet to be demonstrated. This zone also extends to include Uralic, and, to a much lesser extent, to the language families and isolates of Asia's last hunter-gatherers in areas far from the steppe zone. While it can be considered proven beyond doubt that Transeurasian (or Ural-Altaic in a broader sense) represents a language area or interconnected series of overlapping language areas, the question remains whether Transeurasian itself, beneath all of the interacting layers of contact and diffusion, is a language family as well. Scholars such as the present author, who specializes in areas beyond the main languages involved, are unlikely to provide a definitive answer to this question. However, outsiders can sometimes offer a new vantage from which to consider the relevant facts. Because the study of contact phenomena is so obviously crucial to achieving a proper historical understanding of Transeurasian languages, broadening the historical analysis to consider diffusion of linguistic features into additional languages of northern Eurasia can make a useful contribution. While I remain skeptical that Transeurasian will be demonstrated as a genetic family after over a century and a half of inconclusive debate among many of the world's foremost linguists, I am confident that future investigation of this question will continue to bear useful results in the form of a better understanding of language contact and linguistic

typology in this area of the world, regardless of whether this also eventually leads to a resolution of the genetic question.

Recent research on Transeurasian lexical and typological parallels (Robbeets 2005) as well as on shared elements in derivational and inflectional paradigms (Johanson & Robbeets 2010, Robbeets & Bisang 2014) represents a major contribution to the field of Eurasian historical linguistics. These studies stand out for their multidisciplinary breadth, their thorough and balanced historiography, their impeccable application of the Comparative Method, and especially their broad-minded consideration of alternate explanations for parallels shared between the demonstrated language families of Inner Eurasia. But even the best methods and the most dedicated researchers can only succeed in proving a common genetic origin for languages if that origin in fact actually existed. It should be born in mind that the original discipline of Ural-Altaic studies was inherited from an earlier period that did not fully differentiate between evidence of genealogical relationship and the effects left by areal contact. Also inherited from the early days of diachronic linguistics is the still prevalent attitude that proving a genetic relationship between languages represents some sort of pinnacle of achievement in historical linguistics, with evidence of borrowing or other forms of language contact often treated as chaff to be winnowed away in order to get to the desired outcome of establishing a new language family. In reality, genetic relationship is but one of many important potential facts to be investigated in tracing the history of languages. Modern historical linguists should be interested in all aspects of language history. New discoveries about contact, borrowing and diffusion should be recognized as equally important alongside new advances in the genealogical classification of languages. A more refined understanding of grammaticalization and other language-internal developments has no less import for diachronic linguistics than do studies aimed at demonstrating language families. Tracing the origin of morphological oddities or typologically marked combinations of features in languages should therefore be a primary aim of contemporary historical linguistics. Regardless of whether such findings lead to new advances in language classification or to a better understanding of language contact, they are valuable in their own right for the light they shed on how and why languages change through time. Finally, historical studies that have the widest possible aims are precisely those that offer the most realistic chance of uncovering new evidence of a shared common ancestry. For all of these reasons, the definition of success in Transeurasian studies, so eloquently expressed in Robbeets (2005: 26-29) should be expanded to recognize the equal importance to historical linguistics of any new finding involving internal reconstruction in families of the Transeurasian zone, as well as new discoveries regarding ancient language contact among Inner or North Eurasian languages, and not narrowly focused toward proving a genealogical relationship that may or may not have ever existed.

References

Aikio, Ante. (2014). The Uralic-Yukaghir lexical correspondences: genetic inheritance, language contact or chance resemblance? *Finnisch-ugrische Forschungen* 62: 7-76.

Bugaeva, Anna. (2017). Ainu. *The Oxford handbook of polysynthesis*, eds. Michael Fortescue, Marianne Mithun, & Nicholas Evans, pp. 882-905. Oxford: Oxford University Press.

Comrie, Bernard. (1980). *The languages of the Soviet Union*. Cambridge: Cambridge University Press.

Di Cosmo, Nicola (2002). *Ancient China and its enemies: the rise of nomadic power in East Asian history*. Cambridge: Cambridge University Press.

Fortescue, Michael. (1997). Eskimo influence in the formation of the Chukotkan ergative clause. *Studies in Language* 21.2: 369-409.

Fortescue, Michael. (2004). How far west into Asia have Eskimo languages been spoken, and which ones? *Études/Inuit/Studies* 28.2: 159-183.

Fortescue, Michael. (2011). The relationship of Nivkh to Chukotko-Kamchatkan revisited. *Lingua* 121: 1359-1376.

Fortescue, Michael, Jacobson, Steven, & Kaplan, Lawrence. (2010). *Comparative Eskimo dictionary with Aleut cognates*, 2nd ed. Fairbanks: ANLC.

Georg, Stefan, & Volodin, Alexander P. (1999). *Die itelmenische Sprache: Grammatik und Texte*. Wiesbaden: Harrassowitz.

Gruzdeva, Ekaterina. (1999). *Nivkh*. Munich: Lincom Europa.

Janhunen, Juha. (2009). Proto-Uralic – what, where, and when? *Mémoires de la Société Finno-Ougrienne* 258: 57-78.

Janhunen, Juha. (2014). Ural-Altaic: the polygenetic origins of nominal morphology in the Transeurasian zone. *Paradigm change in the Transeurasian languages and beyond*, eds. Martine Robbeets & Walter Bisang, pp. 311-335. Amsterdam & Philadelphia: John Benjamins.

Johanson, Lars, & Robbeets, Martine. (2010). *Transeaurasian vverbal morphology in a comparative perspective: genealogy, contact, chance.* Wiesbaden: Harrassowitz.

Keyser-Tracqui, Christine et al. (2004). Does the Tat polymorphism originate in northern Mongolia? *International Congress Series* 1261: 325-327.

Marcantonio, Angela. (2002). *The Uralic language family: facts, myths and statistics*. Oxford: Blackwell.

Maslova, Elena. (2003a). *A grammar of Kolyma Yukaghir*. Berlin & New York: Mouton de Gruyter.

Maslova, Elena. (2003b). *Tundra Yukaghir*. Munich: Lincom Europa.

Mattisson, Johanna. (2017). Nivkh. *The Oxford handbook of polysynthesis*, eds. Michael Fortescue, Marianne Mithun, & Nicholas Evans, pp. 851-881. Oxford: Oxford University Press.

Miyake, Marc Hideo. (1999). *The phonology of Eighth-Century Japanese revisited: another reconstruction based on written records*. Ann Arbor: University of Michigan.

Rédei, Károly. (1988). *Uralisches etymologisches Worterbuch. Band I: Uralishche und finnisch-ugrische Schicht*. Wiesbaden: Harrassowitz.

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

Robbeets, Martine. (2014). The Japanese inflectional paradigm in a Transeurasian perspective. *Paradigm change in the Transeurasian languages and beyond*, eds. Martine Robbeets & Walter Bisang, pp. 197-232. Amsterdam & Philadelphia: John Benjamins.

Robbeets, Martine, & Bisang, Walter (eds.). (2014). *Paradigm change in the Transeurasian languages and beyond*. Amsterdam & Philadelphia: John Benjamins.

Seefloth, Uwe. (2000). Die Entstehung polypersonaler Paradigmen im Uralo-Sibirischen. *Zentralasiatische Studien* 30: 163-191.

Vajda, Edward. (2009). Loanwords in Ket. *Loanwords in the World's languages: a comparative handbook*, eds. Martin Haspelmath & Uri Tadmor, pp. 471–494. Berlin: Mouton de Gruyter.

Vajda, Edward. (2013). Vestigial possessive morphology in Na-Dene and Yeniseian. *Working papers in Athabaskan (Dene) languages 2012*, eds. Sharon Hargus, Edward Vajda, & Daniel Hieber, pp. 79–91. Fairbanks, AK: Alaska Native Language Center.

Vajda, Edward. (2018a). Areal features in Yeniseian grammaticalization. *Unity and diversity in grammaticalization scenarios*, eds. Walter Bisang & Andrej Malchukov, pp. to be provided. Berlin: Language Science Press.

Vajda, Edward. (2018b). Patterns of retention and innovation in Dene-Yeniseian verb morphology. (details to be provided).

Vovin, Alexander. (1993). *A reconstruction of Proto-Ainu*. Leiden: Brill.

Vovin, Alexander, de la Vaissière, Étienne & Vajda, Edward. (2016). "Who were the *Kjet (羯) and what language did they speak?" *Journal Asiatique* 304.1: 123-142.