Jaakko Häkkinen
Sipoo

On locating Proto-Uralic

In recent years the debate regarding the Proto-Uralic homeland has again intensified. However, not all the relevant arguments have been considered thoroughly. Therefore, in the present article their validity and weight are evaluated. The article also develops further concepts and methodology for reconstruction of stages of Uralic, making it possible to compare Uralic stages to the Indo-Iranian loanword layers with higher resolution than before. As a result, the paper locates Late Proto-Uralic and successive stages in the Central Ural Region, matching the Koptyaki Culture (dated to the early 2nd millennium BCE) and its local predecessor.

https://doi.org/10.33339/fuf.120910 *FUF* 68: 43–100 (2023)

1. Introduction

The debate on the Proto-Uralic homeland goes back a long way (for earlier research history, see K. Häkkinen 1996: 65–76) and it is still ongoing, the strongest candidates being located on either side of the Ural Mountains, either a European homeland in the Volga-Ural Region or a Siberian homeland in West Siberia (up to the Yenisei). Recent articles supporting the Volga-Ural homeland are Kallio (2006; 2015b), J. Häkkinen (2009), and Parpola (2012b; 2017; 2022). Recent articles supporting the West Siberian homeland are Janhunen (2009), Nichols & Rhodes (2018), Nichols (2021), Grünthal et al. (2022), and Saarikivi (2022). In practice, the actual distance between the candidates for the homeland is sometimes minimal; compare the homeland in the Kama-Ural Region in Parpola (2022: 270, 264) to the homeland reaching to the east from the Central Trans-Urals in Saarikivi (2022: 56).

In this article I intend to go through all the relevant arguments for locating the Late Proto-Uralic homeland. Due to ongoing advances in the fields of Uralic etymology and historical phonology, the criteria are stricter in this critical examination than in many earlier articles – including my own previous attempt on the topic (J. Häkkinen 2009).

It is crucial to define what is relevant evidence and what is not. First, many earlier pieces of evidence have been discarded due to flaws which weaken their evidentiary value. If a word has too narrow a distribution, too irregular sound correspondences between cognates, or the original meaning cannot be reconstructed reliably, it has no value for locating the homeland.

This leads to the second point: the only relevant chronological stage here is Late Proto-Uralic – the moment in time right before the disintegration. Not only are pieces of evidence that are too late discarded, but so are pieces of evidence that are too early. The suggested Ural-Altaic features of an areal-typological nature and possible distant contacts or even relatedness with Indo-European, Yukaghir, or Eskimo-Aleut would in any case precede Late Proto-Uralic by several millennia. These phenomena are just as irrelevant for locating Late Proto-Uralic as the location of Late Proto-Uralic is for locating Late Proto-Finnic (in Estonia) or Late Proto-Samoyedic (in South Siberia).

Arguments for locating the homeland can be divided into two categories: compelling arguments and suggesting arguments. Compelling arguments are undeniable, and only if they cannot help to locate the homeland must we turn to suggesting arguments. Among the compelling arguments are the loanword layers from identifiable and locatable donor languages (Section 2) and the paleolinguistic evidence (Section 3). There are several suggesting arguments, but those deserve a shorter treatment (Section 4).

As byproducts of this critical examination, a new model for the disintegration of Late Proto-Uralic will be proposed (Subsection 5.1), as well as the dating of Late Proto-Uralic and the successive stages of reconstruction (Subsection 5.2). After that, a conclusion is drawn from the locating arguments (Subsection 5.3), followed by a brief review of possible archaeological counterparts (Subsections 5.4 and 5.5).

The new model for disintegration has consequences for the question of which words should be counted as Proto-Uralic. When is the distribution of a word wide enough to be counted as Proto-Uralic? In the present critical examination I apply a criterion that in the west, a cognate must be found in at least Mordvin, Finnic, or Saami, and in the east, a cognate must be found at least in Mansi or Khanty. I will acknowledge the special status of Samoyedic, yet I will argue in Subsection 5.3 for why the presented disintegration model allows a word to not be present in Samoyedic, although in that case a regular cognate in Mansi or Khanty is required. For example, the famous words for 'bee' and 'honey' do not have cognates beyond Hungarian in the east, making them too suspect as Late Proto-Uralic words.

As will be demonstrated later, there is probably more than half a millennium (but less than a full millennium) between the first regional division and the first (macro-)branch-specific sound changes. Such circumstances naturally lead to the existence of a multitude of words which appeared after Late Proto-Uralic but still cannot be distinguished from the Late Proto-Uralic words by the phonological criteria. Such circumstances require applying quite a strict distributional criterion.[1]

1. I would like to thank Luobbal Sámmol Sámmol Ánte, anonymous referees, and the editors for their valuable comments.

1.1. On notations and labels

In the first syllable, there are eight generally accepted vowel qualities in Proto-Uralic: the front vowels **ä*, **e*, **i*, **ü* and the back vowels **a*, **o*, **u*, **ę* (**į*) (Aikio 2022: 5). There were only two certain vowel qualities beyond the first syllable, namely **a* and **ə*, the primary distinction probably being a full vowel vs. a reduced vowel (Kallio 2012: 163–165). Possibly a phonetically front allophone appeared after a front vowel in the first syllable, and a phonetically back allophone after a back vowel. There is also some evidence supporting the possible existence of **o* in non-initial syllables (Aikio 2015: 37–38; 2022: 9).

In the Uralic Phonetic Alphabet (UPA), the symbol /ə/ denotes a phonetic value different from the International Phonetic Alphabet (IPA), that is, it is usually a front counterpart of the back /ə̑/, and both of those UPA symbols are semilabial and mid-high (the three vowel heights in the UPA do not neatly correspond to the four vowel heights in the IPA; see Iivonen 2012: 18).

The reconstruction stages following Late Proto-Uralic, as well as other related concepts, will be labeled and explained when they are encountered in the present critical examination. Labels for units at different stages of disintegration following the uniform proto-language are: (1) center = regionally separated unit > (2) pre-dialect = substitutionally differing unit > (3) proto-dialect = phonologically differing unit (see Subsection 5.1).

Datings are specified with the conventional marking "BCE", denoting calendar years before the common calendrical starting point. Datings are rough approximations, which is shown by the use of even centuries. Abbreviations are explained when they first appear, and a list of abbreviations can be found at the end of the article.

The Ural Mountains are generally divided into five regions in Russia and one in Kazakhstan. From the north, the regions are the Polar Urals, the Sub-Polar Urals (the Nether Urals), the Northern Urals, the Central Urals (the Middle Urals), and the Southern Urals. The southernmost region in Kazakhstan is called the Mughalzhar Hills. The most relevant regions for this article are the Central Urals (parallel to the Lower and Middle Kama up to Perm) and the Southern Urals (parallel to the Middle Volga from the Kama fork down to the Kazakhstan border). The label "Trans-Urals" denotes the eastern slopes of the Urals and a narrow strand of lowland adjacent to them.

1.2. Different types of borrowing

As loanwords occupy a great share of the following critical scrutiny, an important topic must be clarified first. The distinction between borrowing into the uniform proto-language and into later stages has been generally acknowledged, but it is necessary to increase the resolution here. There are several types of borrowing concerning the situation after the disintegration of the proto-language, and the evidentiary value varies between them. I illustrate the different types with examples from the Indo-Iranian loanwords:

1. **Parallel borrowing**: Difference in sound substitutions between recipient languages, while the original word in the donor language is the same. U **počaw* 'reindeer' (> Saami, Finnic?, Mari, Permic) represents substitution with *o* and U **päčVw/γ* 'reindeer calf' (> Mansi, Khanty) substitution with *ä* of the very same Proto-Iranian **patsu-* 'cattle' (Holopainen 2019: 184–185, 196).
2. **Variant borrowing**: Difference between variants of the original word in the donor language. U **s/šǯrńa* 'gold' (> Hungarian, Mansi, Khanty) vs. U **serńa* 'gold' (> Mordvin) belong to this type, as the latter was borrowed from a different grade of the same Proto-Iranian word, where there was a syllabic resonant instead of a vowel: PIr **dzəranya-* vs. PIr **dzr̥Hnya-* 'gold' (Holopainen 2019: 232–234).
3. **Separate borrowing**: Different donor languages or different chronological stages of the same donor lineage. U **sa/ora* 'lake' (> Hungarian, Mansi, Khanty) vs. Proto-Permic **sariď z* 'sea' (Holopainen 2019: 217–219) must represent two donor languages/stages separated by a great gap in time, because the original U **a* has changed into the PPe **u* (Metsäranta 2020: 94).
4. **Irregular developments** in a recipient language, which cannot be entirely explained by differences between sound substitutions, donor variants, or donor languages, for example U **ćarwi* 'horn' vs. PKh **ćerpa* 'horn' (Holopainen 2019: 220–222).

Considering the evidentiary value of these types, (1) *parallel borrowing* is solid evidence for the disintegration of the proto-language: if the very same word has been borrowed in parallel, showing different sound substitutions between branches, then we can be certain that the proto-language indeed

disintegrated during the borrowing. However, there is a distributional restriction for this rule: if one of these substitutions has a wide enough distribution in Uralic, it still can be reconstructed for Late Proto-Uralic, as the other substitution with narrow distribution could be later. In such case we would actually be dealing with separate borrowings.

(2) *Variant borrowing* is almost as strong evidence for the disintegration of Late Proto-Uralic, because it requires at least two centers within the proto-language speech community. This type of borrowing can be considered the second strongest evidence pointing to the already disintegrated proto-language. The same distributional restriction applies here as in the previous type.

In the case of (3) *separate borrowing*, when the words are from different points in time, we cannot exclude the possibility that the older or more widespread loanword was earlier known also in the other language, until it became replaced by the younger rather similar-looking loanword. Therefore, this type is somewhat weaker evidence for the disintegration of Late Proto-Uralic.

(4) *Irregular developments* occurring in an individual language or branch after the borrowing event cannot testify against the status of the word in the uniform proto-language, although it certainly can distort our attempts to reconstruct the word in the proto-language. Irregular cognates in some branches cannot diminish the value of regular cognates in other branches: if those branches are distant enough in the taxonomic model of the language family, the word can be reconstructed for the proto-language.

2. Early Indo-European loanword layers

Within the Early Indo-European loanword layers, I include early stages of two separate donor lineages, which descend from Late Proto-Indo-European (LPIE): the Indo-Iranian lineage consisting of Early (EPIIr) > Middle (MPIIr) > Late Proto-Indo-Iranian (LPIIr) > Proto-Iranian (PIr), and the Northwest Indo-European lineage consisting of Archaic Indo-European (AIE) > Northwest Indo-European (NwIE). At the moment the number of convincing loanwords in the Early Proto-Indo-Iranian layer is about a dozen, and greater still in the Late Proto-Indo-Iranian and Proto-Iranian layers (Grünthal et al. 2022: Appendix 2 "Indo-Iranian loans in Uralic"). There are several dozen proposed Archaic and Northwest Indo-European loanwords, but their reassessment is still an ongoing process.

2.1. Indo-European and Uralic

Since the generally accepted rejuvenation of Late Proto-Uralic (Kallio 2006; J. Häkkinen 2009), this stage is no longer contemporaneous with Late Proto-Indo-European. That does not disqualify the earliest proposed Indo-European loanwords in Uralic, it only requires a new label for them. I use the label Archaic Indo-European as an umbrella term for the loanwords resembling LPIE with wide distribution in Uralic, and the label Northwest Indo-European for the (usually likewise LPIE resembling) loanwords, which have a western (Finno-Permic) distribution in Uralic and cognates mainly in the Northwest Indo-European branches. Most if not all of the Archaic Indo-European loanwords could probably be explained as borrowed from the ancestor of Northwest Indo-European.

Northwest Indo-European is not a proto-dialect in the same sense as Late Proto-Indo-Iranian, but rather a continuum of phonologically conservative Indo-European varieties roughly corresponding to the wide area of the Corded Ware Cultures. Northwest Indo-European branches (at least Balto-Slavic, Germanic, Celtic, and Italic) share quite a lot of words that lack cognates in more distant branches. Many such words also show phonological or phonotactic features foreign to Late Proto-Indo-European and tend to denote local flora, fauna, and livelihoods, which points towards a substrate origin from unknown ancient languages (Mallory & Adams 2006: 78–79).

It has also been noted that while Anatolian, Indo-Iranian, and Greek must be assumed to have gone through some branch-specific sound changes already during the 3rd millennium BCE, in the northwest, the phonological distinctions occur only during the 2nd millennium BCE (Mallory & Adams 2006: 103–104). Even the loanwords borrowed into Uralic branches seem to testify that distinguishably Balto-Slavic and Germanic phonological features appear later than the recognizable Indo-Iranian features (Kallio 2009: 38–40; forthcoming).

In principle, both of these Early Indo-European donor lineages weigh heavily on locating Proto-Uralic. In practice, however, we suffer from limitations in the resolution between the reconstruction stages and in the quality of the loanwords. These topics are considered in Subsections 2.3 and 2.4.

2.2. Changing views in the Uralic studies

During the late 20th century, in the Uralic studies there prevailed the so-called Moderate Continuity Theory, in which the arrival of Finnic and Saami in the Baltic Sea Region was connected to the spread of the Typical Combed Ware ca. 4000 BCE, and (Late) Proto-Indo-European was seen as roughly contemporaneous with (Late) Proto-Uralic. In that framework, the Indo-Iranian loanword layers were seen as clearly later than Proto-Uralic (on the research history, see Aikio & Aikio 2001).

At the beginning of the current century, the accumulating evidence especially from Germanic loanwords led to a later dating for the phonological divergence between Finnic and Saami (Koivulehto 2002; Aikio 2006; Kallio 2009; 2015a). At the same time, discontent towards the traditional Uralic taxonomic model also grew, i.e. the family tree in which Samoyedic was the first branch to split away (K. Häkkinen 1984; Salminen 1989; 2002; J. Häkkinen 2007).

It was also more pronouncedly emphasized that linguistic continuity could not be reliably tracked from archaeological continuity (Aikio & Aikio 2001; Mallory 2001; J. Häkkinen 2010). The utter unreliability of the continuity argument was finally revealed by gravely contradicting datings achieved by the very same method from the very same data, when Proto-Uralic was claimed to have been spoken in Finland already right after the Ice Age (on the history of this scientific debate in Finland, see Tirkkonen 2012).

Together all these factors led to a paradigm shift. When Samoyedic no longer had veto power to dismiss words from being Late Proto-Uralic, it was possible to consider even the Late Proto-Indo-Iranian loanwords as being borrowed into Late Proto-Uralic. Consequently, earlier datings were rejected and later datings around 2000 BCE were accepted for Late Proto-Uralic (Kallio 2006; J. Häkkinen 2009; Parpola 2012b). The level of phonology (enough regular cognates in the Uralic branches) subdued the level of distribution as the paramount dating criterion.

However, it has been recently demonstrated that distribution cannot be so easily overruled when the number of convincing loanwords in a layer is high enough. That is, when there appear to be more than a dozen Early Proto-Indo-Iranian loanwords in Uralic and none of these has a cognate in Samoyedic, it becomes highly improbable that all the words could have simply disappeared since Late Proto-Uralic (Grünthal et al. 2022: 10). At the same time, the stage where a shift occurs from regular cognates to

parallel borrowings appears to be much later, occurring only around the Proto-Iranian stage (Holopainen 2019: 343).

Neither of these arguments can be explained away or ignored. The reason for the apparently contradictory results must lie in the fundamental difference between the lexical and phonological levels. Therefore, based on the evidence from the Indo-Iranian loanword layers, I will construct a new model for the disintegration of Late Proto-Uralic, presented in Subsection 5.1.

2.3. Resolution of the donor lineages

By resolution I mean the density of successive distinguishable reconstruction stages in a donor lineage, seen in the loanwords through the Uralic filter. In a donor lineage, there might have occurred changes concerning word-initial consonant clusters or voiced obstruents, but many such changes would remain invisible due to the restrictions caused by Uralic phonology and phonotactics: an initial cluster would have been substituted by a single consonant, and a voiced obstruent would have been substituted by a voiceless obstruent. Similarly, the presence of the Indo-European palatovelars and labiovelars in a donor lineage would be difficult to recognize reliably (Holopainen 2021).

For the taxonomy of Proto-Indo-European, I follow the well-argued consensus view, according to which Anatolian was the first branch to split away, followed by Tocharian (Jasanoff 2017; Ringe 2017). The remaining core is here called Late Proto-Indo-European (LPIE), following Anthony & Ringe (2015: 201).

Concerning the disintegration of Late Proto-Indo-European, centumization and satemization are no longer considered clade-defining changes: the first occurred independently across different branches, and the second has spread secondarily (Ringe 2017). There are more exceptions to satemization in Balto-Slavic than in Indo-Iranian, which points to its secondary spread (Kim 2018: 1975). In the position after **s* LPIE **ḱ* was depalatalized to **k* before satemization in Balto-Slavic (Matasović 2005: 148), showing that satemization was not the earliest Balto-Slavic sound change.

The *ruki*-rule in Indo-Iranian was triggered also by the secondary **i* developing from an earlier syllabic laryngeal and the secondary **r* from **l* (Lubotsky 2018: 1881), so it was not among the earliest Indo-Iranian sound changes. Also, the results in Nuristani differ from those in Iranian and Indic (Hegedűs 2012). In Balto-Slavic the *ruki*-rule is regular only in Slavic, while

in Baltic it is more restricted (Kim 2018: 1976), and therefore it seems to have spread there only after the disintegration of Proto-Balto-Slavic. In Balto-Slavic the *ruki*-rule is also later than satemization (Matasović 2005: 148).

Interestingly, some Baltic loanwords in Finnic show the *ruki* reflex even though the modern East Baltic languages do not, e.g. LPFi **laiha* 'thin, slender' < MPFi **lajša* ← dialectal Balto-Slavic **laišas* ~ Lithuanian *líesas* 'thin' < **laisas* (Kallio 2008: 267). Due to the secondary nature of this change, it is natural to assume that **laišas* would represent some eastern Balto-Slavic dialect spoken closest to Indo-Iranian, while **laisas* could represent a more central Balto-Slavic dialect (Proto-Latvo-Lithuanian?). There are possible traces of an even more diverse continuum of Balto-Slavic varieties than has been recognized thus far, based on recurring irregularities in the loanwords borrowed into the West Uralic branches (J. Häkkinen 2022).

Archaic Indo-European and Northwest Indo-European were still phonologically very similar to Late Proto-Indo-European, at least as far as we can see through the Uralic filter. However, there seem to be no certain examples of preserved palatovelars in the loanwords borrowed into Uralic (Holopainen 2021: 199). This could point to post-Proto-Indo-European donor languages, as the palatovelars either merged with the plain velars (in *centum*-dialects) or changed to palatalized affricates or sibilants (in *satem*-dialects). As Germanic has gone through centumization and Balto-Slavic through satemization, there could be ancient loanwords from both types of dialects in Uralic. However, it is questionable whether we could distinguish even the centumized or satemized consonants from the Late Proto-Indo-European consonants through the Uralic filter:

1. LPIE **ḱ*, **ǵ*, and **ǵʰ* could have been substituted by either U **k* or U **ć*, and LPIE **kʷ*, **gʷ*, and **gʷʰ* could have been substituted by U **k* or **ku*.
2. The centumized **k*, **g*, and **gʰ* (< **ḱ*, **ǵ*, and **ǵʰ*) would have been substituted by U **k*, and the preserved **kʷ*, **gʷ*, and **gʷʰ* could have been substituted by U **k* or **ku*.
3. The satemized **ć*, **ʒ́*, and **ʒ́ʰ* would have been substituted by U **ć*, and the **k*, **g*, and **gʰ* (< **kʷ*, **gʷ*, and **gʷʰ*) would have been substituted by U **k*.

Consequently, even if there were centumization- or satemization-related sound changes in an Indo-European donor language, we could not, in the absence of other branch-specific sound changes to guide our interpretation, reliably distinguish them from the Late Proto-Indo-European stage.

Moreover, even if there were loanwords from some early *centum-* or *satem-* dialects adopted into Uralic, their accurate dating would still be practically impossible to determine: the centumization- and satemization-related sound changes could have occurred in different Indo-European branches at different times, from right after the Late Proto-Indo-European stage (around or before 3000 BCE) to much later dates (around or after 2000 BCE).

Because this time span also includes the disintegration of Late Proto-Uralic, such a lack of resolution prevents us from estimating whether those kinds of loanwords were borrowed into Late Proto-Uralic or already into separate Uralic pre-dialects, and therefore their value for locating Proto-Uralic would be gravely diminished. Temporally relevant would be only those loanwords which have been borrowed from a datable reconstruction stage, and among those, spatially relevant would be only those loanwords which would immediately precede the dispersal of Late Proto-Uralic.

The chronological resolution is high only in the Indo-Iranian lineage (see Subsection 2.5), while in the Northwest Indo-European lineage the lack of distinguishable sound changes (visible through the Uralic filter) continues up to the 2nd millennium BCE, making that lineage practically worthless for locating Late Proto-Uralic (Figure 1; for the chronology, see Subsection 5.2).

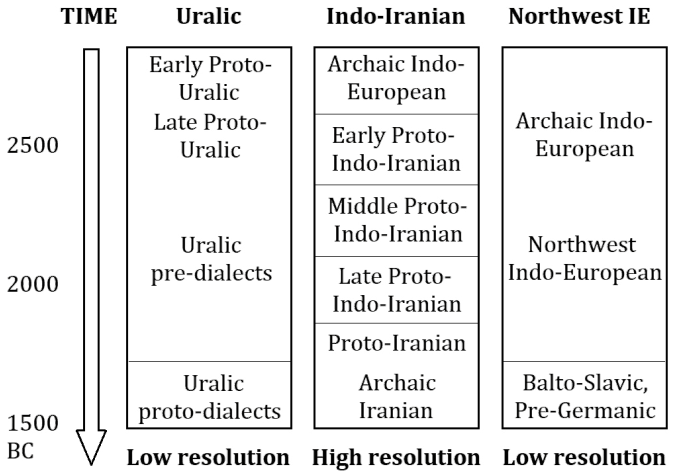


Figure 1: Successive reconstruction stages are not always phonologically distinguishable; only those which are separated by a thin horizontal line are. In the Indo-Iranian lineage the resolution is highest, but between the Uralic and the Northwest Indo-European lineages it is difficult to date the loanwords precisely.

At the moment, there is no reliable method to distinguish loanwords adopted into Late Proto-Uralic from loanwords adopted earlier or later, and consequently I am forced to dismiss the Northwest Indo-European lineage from this scrutiny.

2.4. Quality of loanwords

There are several dozen proposed loanwords which could be borrowed from Late Proto-Indo-European, Archaic Indo-European, or Northwest Indo-European (e.g., Koivulehto 1991; 2001), so the existence of such loanword layers does not suffer from the lack of quantity but from the possible lack of quality. There are recent critical assessments about the words with Uralic **š* as the assumed substitute for the Indo-European laryngeals (Hyllested 2014), about the most widespread Indo-European loanwords in Uralic (Simon 2020), and about the words containing alleged Late Proto-Indo-European palatovelars (Holopainen 2021). Many of the assumed early Indo-European loanwords have already been proven to be improbable, but there are still plenty of proposed loanwords waiting to be assessed more thoroughly.

We must bear in mind that even if some loanwords showed unexpected sound correspondences, it would not automatically make them false. Some phonological mismatches could be caused by phonological developments occurring already before Late Proto-Indo-European or Late Proto-Uralic, so they could reflect very ancient contacts. Other mismatches could reflect a lost dialect of some Indo-European branch or just an arbitrary and unexpected sporadic sound substitution. If there are multiple occurrences of such an unexpected sound correspondence (becoming a recurring irregularity), it becomes more probable that there is some real phenomenon behind them. As an example, there are at least two occasions showing unexpected U **ć* reflecting LPIE/LPIIr **s*, which I here suggest possibly belong to the Late Proto-Indo-Iranian loanword layer:

U **poća* 'penis' >
Saami **puoćę* ~ Hungarian *fasz* (Sammallahti 1988: 548)
← LPIIr **pásas* (> Sanskrit *pásas-*; Holopainen 2019: 185) < LPIE **péses* 'penis' (Mallory & Adams 2006: 183–184)

U **moćkə-* 'wash' > Finnic, Mordvin, Mari, Permic, Hungarian, Samoyedic (Sammallahti 1988: 538)

← LPIIr **mazg-* (> Sanskrit *májjati* 'sinks') < LPIE **mesg-* 'dip underwater, dive' (Mallory & Adams 2006: 403; Pokorny 2007: 2107); cf. Balto-Slavic **mazgo-* 'wash' < **mosg-* (Derksen 2015: 308)

Both of these words could represent the well-known **o*-substitution for the Late Proto-Indo-Iranian **a* (see Holopainen 2019: 49–50). The second word has earlier been considered a Proto-Indo-European loanword, but the Uralic **ć* was problematic without any known further examples (Simon 2020: 248). This word could also come from Balto-Slavic, but it is more probable that both of these words come from the same donor language, namely Late Proto-Indo-Iranian. There are also other loanwords in this layer for which a cognate is present in Samoyedic (see Subsection 2.6.2).

One possible explanation for the unexpected substitution with **ć* is connected to the fact that Uralic **ć* appears frequently in the Late Proto-Indo-Iranian loanwords as the substitute for **ć*, **ʒ́*, and **ʒ́^{h}*. Perhaps that sound was therefore associated with that particular donor language and hypercorrectly appeared even in some words where the donor language had plain **s*?

My original intention was to include in this scrutiny also loanwords from the Northwest Indo-European donor lineage, but after consulting with Luobbal Sámmol Sámmol Ánte (Aikio, personal communication), I concluded that a critical reassessment of the quality of these loanwords has only begun. Another reason for omitting these loanwords was the low resolution between the Uralic and the Northwest Indo-European donor lineages (see Subsection 2.3).

Moreover, the evidentiary value of Archaic Indo-European loanwords would in any case be weak. Even though Late Proto-Indo-European and Northwest Indo-European were spoken in Europe, the picture is complicated by the location of Pre-Proto-Tocharian. This eastward expansion is generally connected to the movement from the European steppe to South Siberia, where the Afanasyevo Culture was formed (Anthony & Ringe 2015: 209). As this movement is dated already to the 4th millennium BCE, the Archaic Indo-European spoken there could easily explain loanwords into Early or Late Proto-Uralic, if these stages were spoken in South Siberia.

Therefore, no decisive evidence could be gained from the Archaic Indo-European loanwords as long as we cannot reliably demonstrate that they belong to the Northwest Indo-European lineage instead of the Tocharian lineage, or at least demonstrate a temporal gap between the AIE and the NwIE loanword layers being so short that it would require also a regional vicinity and make it probable that they represent successive stages of the very same NwIE lineage.

2.5. Reconstruction stages of the Indo-Iranian lineage

The early part of the Indo-Iranian lineage is here divided into four reconstruction stages: Early (EPIIr), Middle (MPIIr), and Late Proto-Indo-Iranian (LPIIr), as well as Proto-Iranian (PIr). The following list contains mainly sound changes distinguishable through the Uralic filter, as seen in the Indo-Iranian loanwords borrowed into Uralic, so the list is not comprehensive. The sound changes are taken from Ollett (2014), Cantera (2017), and Lubotsky (2018):

1. Early Proto-Indo-Iranian:
   1.1. Interconsonantal $*h > *i$
   1.2. Brugmann's Law: $*o > *\bar{o}$ in open syllables
   1.3. Laryngeal coloring: $*e+h_2 > *a+h_2$, $*e+h_3 > *o+h_3$
   1.4. $*l > *r$
2. Middle Proto-Indo-Iranian:
   2.1. Satemization: $*\acute{k}$, $*\acute{g}$, $*\acute{g}^h > *\acute{c}$, $*\acute{ʒ}$, $*\acute{ʒ}^h$, while $*k^w$, $*g^w$, $*g^{wh}$ merge into $*k$, $*g$, $*g^h$
   2.2. The *ruki*-rule: $*s > *š$ next to *r*, *u*, *K*, *i* (also the secondary $*i$ and $*r$)
   2.3. Palatalization of the velar stops before the remaining $*e$: $*k$, $*g$, $*g^h > *\check{c}$, $*\check{ʒ}$, $*\check{ʒ}^h$
3. Late Proto-Indo-Iranian:
   3.1. The merger of non-high vowels and syllabic nasals: $*a$, $*e$, $*o$, $*m̥$, $*n̥ > *a$
   3.2. The merger of remaining laryngeals into $*H$
4. PIr: The depalatalization of affricates $*\acute{c}$, $*\acute{ʒ} > *ts$, $*dz$

Satemization is rather difficult to date. It must precede the stage 2.3, but it could also be somewhat earlier (but not the earliest; see Subsection 2.3). Even though the laryngeal coloring on an adjacent short $*e$ ($*e+h_1 > *e+h_1$; $*e+h_2 > *a+h_2$; $*e+h_3 > *o+h_3$) occurred in every Indo-European branch and is therefore often considered already a Proto-Indo-European development

(Byrd 2018: 2063–2064), other instances leading to similar results are clearly later. Greek shares with Armenian and probably Phrygian similar coloring of a word-initial laryngeal before a consonant (*h_1C > *eC; *h_2C > *aC; *h_3C > *oC), but only in Greek do we see a similar change concerning the syllabic laryngeal and the word-final laryngeal (*h_1 > *e; *h_2 > *a; *h_3 > *o; Beekes 2011: 146–153). These cases show that such colorings have occurred independently more than once, when the conditions (the coloring quality of the laryngeals) still remained.

Even the oldest type, the laryngeal coloring of short **e*, could be later than Late Proto-Indo-European. It has been argued that in Indo-Iranian it seems to be younger than Brugmann's Law (**o* > **ō* in open syllables; Lubotsky 1990; 2018: 1877). Ollett (2014) has shown that in Indo-Iranian the laryngeal coloring preceded the palatalization of the velar stops, but he could not decisively refute that Brugmann's Law came first. According to Lubotsky (2018: 1877), the change of interconsonantal **H* > **i* must be even older than Brugmann's Law in Indo-Iranian, at least in the final syllable.

Therefore, I place the laryngeal coloring within the Early Proto-Indo-Iranian stage. This leaves room for the option that also in the Northwest Indo-European lineage the laryngeal coloring could be partially a later phenomenon, which should be taken into consideration when assessing possible Archaic or Northwest Indo-European loanwords in Uralic. If we could stratify the convincing loanwords into pre- and post-coloring stages, the resolution in that donor lineage would increase.

Interestingly, among the loanwords into Uralic, there are no certain examples of the remaining **e* with reflexes of the new palatalized affricates (**če-*, **ǯe-*, **ǯʰe*), so there might be a gap in the contacts during the Middle Proto-Indo-Iranian stage. However, there are also only three more or less convincing Early Proto-Indo-Iranian loanwords with **ke-* in Uralic, so this gap could be only illusory, caused by the low number of loanwords beginning with these secondary MPIIr affricates in the first place. Another possible explanation is that the Late Proto-Indo-Iranian vowel merger occurred very soon after the palatalization of velars.

Yet there is at least one possible loanword before the vowel merger: Late Proto-Finnic **herä-* < PrePFi **šera-* 'wake up' ← MPIIr **Hǯer-* < LPIE **h₁ger-* 'awake' (Holopainen 2019: 258). There are competing etymologies, and Uralic **š* is not considered the most expected substitute (because Uralic also had a consonant **č*), but there are parallel examples for the **š*-substitution of an initial Indo-Iranian affricate:

West Uralic **šukta* 'swidden field'
← PIr **tsuxta* 'burned' (Holopainen 2019: 264–265)
LPFi **hadas* < PrePFi **šatas* 'sprout, germ (of seed)'
← PIr **dzaHta-*, verbal adjective from the root **dzanH-* 'be born, grow' (could also be Germanic; Holopainen 2019: 257–258)
Possibly U *šę̄rńa* 'gold'
← PIr **dzəranya-* 'gold' (Holopainen 2019: 232–234; other initial sibilants are possible for this Uralic word, see Subsection 2.5.2).

2.6. Connecting the Uralic and the Indo-Iranian reconstruction stages

2.6.1. Evidence from the lexical level

It has long been known that Samoyedic seems to be the lexical outlier within the Uralic language family. Samoyedic seemed to share the smallest number of words with all the other branches, and this situation was modeled as a division between Finno-Ugric and Samoyedic, leading to the stratification into the earlier Proto-Uralic (cognate present in Samoyedic) vs. the later Proto-Finno-Ugric layer (no cognate present in Samoyedic) in vocabulary (Sammallahti 1988).

However, the lexical level alone cannot reliably testify to the disintegration of Late Proto-Uralic, because there are other possible explanations as to why some branches appear to share less or more inherited words with other branches than expected. Historical phonology is a more reliable level and leaves less room for different interpretations (J. Häkkinen 2012).

Nevertheless, concerning the Indo-Iranian loanwords, it seems remarkable that in the Early Proto-Indo-Iranian loanword layer (14 words), there are no words with a cognate in Samoyedic (Grünthal et al. 2022: Appendix 2 "Indo-Iranian loans in Uralic"). It is statistically highly improbable that Samoyedic would have first borrowed and later lost all the Early Proto-Indo-Iranian loanwords. It seems necessary to assume that Samoyedic was already outside the contact zone during the earliest distinguishable Indo-Iranian contacts. However, Samoyedic surprisingly participated in contacts during the Late Proto-Indo-Iranian loanword layer (see the following subsection).

2.6.2. Evidence from the sound substitutions of loanwords

As was presented in Subsection 1.2, parallel borrowings are the strongest evidence against the uniform proto-language. This works also the other way around: the lack of parallel borrowings bears evidentiary value in the cases where we have several possible sound substitutions. It is well known that there were four possible substitutes for Late Proto-Indo-Iranian **a* on the Uralic side: **a*, **ä*, **o*, and **ę* (Holopainen 2019: 49–50). No phonological conditions have been found to explain the choice of the vowel – it appears to have been purely arbitrary (Holopainen 2019: 327–328). As LPIIr/PIr **a* is also very frequent in loanwords, it works perfectly as a diagnostic vowel. In the LPIIr layer we still see many loanwords which have a wide distribution and in which all the Uralic branches agree with the same vowel substitution – even Samoyedic, when there is a cognate:

U **pę̄ŋka* 'psychedelic mushroom' >
Mordvin **paŋgə̑* ~ Mari **poŋgə̑* ~ Mansi **pīŋk* ~ Khanty **pāŋk* ~ Samoyedic **pę̄ŋkå* (Aikio 2015: 59)
← LPIIr **b(ʰ)anga-* 'narcotic plant' > Middle Persian *bang*, *mang* 'henbane (*Hyoscyamus niger*)' (Holopainen 2019: 186–188)

U **ćęra*, **ćęr|kə* >
Saami **ćuorē* 'light gray (of reindeer hair)', **ćuorkę-dē* 'gray (of human hair)' ~ Samoyedic **sį̄rå* 'snow', **sęr* 'white, ice' (Aikio 2020: 125–126)
← LPIIr **ćar-* (> Sanskrit *śārá-* 'colored') < LPIE **ḱer-* 'grayish blue/green' (Mallory & Adams 2006: 333; Pokorny 2007: 1547–1548 includes words like Armenian 'snow' and Slavic 'hoarfrost'). New proposition; earlier loan etymologies have been discarded by Holopainen (2018: 157–158; 2019: 231–232).

U **ćęta* '100' >
Saami **ćuotē* ~ Finnic **sata* ~ Mordvin **śadə̑* ~ Mari **šüδə* ~ Permic **śo* ~ Hungarian *száz* ~ Mansi **sį̄t* ~ Khanty **sāt* (Aikio 2015: 60)
← LPIIr **ćatá-m* '100' (Holopainen 2019: 242–244)

U **kęnta(w)* 'log, fallen tree' >
Saami **kuontə̄* ~ Finnic **kanto* ~ Mordvin **kandə̑* ~ Mansi **kį̄ntā* ~ Khanty **kānt* (Aikio 2015: 59)
← LPIIr **skandʰá-* > Old Indic *skándhas-* 'twig, branch' (Holopainen 2019: 120)

U **tora-* 'fight' >
Saami **toarɔ̄-* ~ Finnic **tora-* ~ ?Mordvin **ťuŕə-* ~ Samoyedic **tåro-* (Aikio 2015: 62)
← LPIIr **tara-* 'overcome' (Holopainen 2019: 282–285)

U **ćoma-* 'be starving, get tired' >
Mordvin **śumə̂-* ~ Mari **šŭma-* ~ Permic **śuma-* ~ Hungarian *szom-* (Aikio 2015: 61)
← LPIIr **ćamH-* 'be extinguished, die out' (Holopainen 2019: 213)

U **onća* 'part, share' >
Finnic **osa* ~ PSa **oańćē* ~ Mari **užaš* ~ Hungarian *ágyék* ~ Mansi **ūńć* (Aikio 2015: 61)
← LPIIr **Hanća-* 'share of a fortune, loot' (Holopainen 2019: 170–171)

U **ora* 'awl' >
Saami **oarē* ~ Finnic **ora* ~ PMd **urə̂* ~ Hu *ár* (Aikio 2015: 61)
← LPIIr **HāŕāH-* 'goad' (Holopainen 2019: 163–164)

U **ćaδa-* 'rain' >
Finnic **sata-* ~ Samoyedic **sårå-* (Aikio 2015: 56)
← LPIIr **ćad-* 'fall' (Holopainen 2019: 224)

It is highly improbable that of the four available substitutions for the Indo-Iranian **a*, all separate and already distinct Uralic branches would have independently chosen the same one in every single loanword. Admittedly, only four of these nine words have cognates in Samoyedic, in which the original **a* and **o* cannot always be distinguished, but we can still distinguish from these sounds the original **ę* and **ä*. All four of these Samoyedic cognates agree with the rest of Uralic. Consequently, the evidence points to a borrowing situation preceding a wider regional dispersal of Uralic. There must have still existed a narrow Uralic speech community during the borrowing of these Late Proto-Indo-Iranian loanwords.

However, even though parallel borrowing from Indo-Iranian seems improbable, a possibility of borrowing from one Uralic pre-dialect to another should also be considered. Nevertheless, the more words there are with a wide distribution and regular cognates, the more improbable it is that all of them could be due to intra-Uralic borrowing. These nine words cannot be dismissed by this explanation.

During the slightly later Proto-Iranian loanword layer, separate regional pre-dialects already existed. For Proto-Iranian **a*, we find different substitutions in different branches, each of them usually having quite a narrow distribution:

U **kɇrtańa* 'iron' > Mari, Mordvin?
Parallel borrowing U **kärta* > Permic, Khanty?
← PIr **kártana-* 'a cut' or PIr **kr̥tí-* 'dagger, knife' (Holopainen 2019: 121–125)

U **počaw* 'reindeer' > Saami, Finnic?, Mari, Permic
Parallel borrowing U **päčVw/γ* > Mansi, Khanty
← PIr **patsuka* 'livestock' (Holopainen 2019: 184–185, 196)

U **saŋka/*soŋka* 'old' > Mari, Hungarian (ambiguous borrowing: both **a* and **o* are possible for Mari and Hungarian)
← PIr **sanaka* 'old' (Holopainen 2019: 235)

U **s/šɇrńa* 'gold' > Hungarian, Mansi, Khanty
Possible parallel borrowing U **ć/šɇrńa* > Mari
Variant borrowing U **serńa* > Mordvin
← PIr **dzŕ̥Hnya- ~ *dzəranya-* 'gold' (or even AIr/OIr **zaranya*?) (Holopainen 2019: 232–234)

The last example does not represent PIr **a*, but it also shows variation between different pre-dialects. In Mari, the Malmyž dialect (MariM) has best preserved the original **s*, while in the other dialects the **s* has coalesced into **š* (Metsäranta 2020: 36). However, in *Tscheremissisches Wörterbuch* (Moisio & Saarinen 2008) the closest dialect to Malmyž is marked as Ok (Bolshoy Kiľmez), but it only has *s* in some of those words (and *š* in others), in which we find Malmyž *s* in the *Mari Nyelvjárási Szótár* (Beke 1997–2001). It seems that the Malmyž dialect in the latter dictionary has best preserved the original **s*, and the Bolshoy Kiľmez dialect in the first dictionary second best.

In Beke's dictionary, we find Malmyž *soŋgo* 'old' (page 2449) and Malmyž *šörtńö* 'gold' (page 2492), the latter pointing towards either the original **ć* or **š*, being in any case an unexpected reflex. As **s* and **š* have coalesced in all the East Uralic branches, theoretically we could derive also them from U **šɇrńa*. In Subsection 2.5, other examples showing U **š* as the substitute for the Indo-Iranian initial affricates were considered.

2.6.3. Evidence from sound changes in different branches

The first sound changes appeared at the macro-branch level, although it must be noted that it was no longer a question of a uniform language but instead shared innovations between already regionally separate yet adjacent pre-dialects of the Uralic speech community. Especially diagnostic are the East Uralic sibilant changes shared by Hungarian, Mansi, Khanty, and Samoyedic (drafted in J. Häkkinen 2007: 71–73; reply to recent critique in J. Häkkinen 2023):

1. U **s* and **š* merge into **š*
2. U **š* > **ɬ* (voiceless lateral fricative) or **ϑ* (voiceless dental spirant)
3. U **ć* > **s*

This chain of three subsequent changes in a certain order makes it practically impossible to assume independent development in different branches or later spreading from one branch to another. These changes must have occurred in a narrow area within a short time span (for wider argumentation, see J. Häkkinen 2023). Based on the previous subsection, the participants were already neighboring pre-dialects when these sound changes occurred, because the loanwords showing different substitutions between branches have also gone through the East Uralic sound changes. Two examples seen already in the previous subsection are the following:

U **saŋka*/**soŋka* ‘old’ > EU **ɬaŋka*/**ɬoŋka* >
Hungarian *agg* ‘old’

U **s*/*šǝrńa* ‘gold’ > EU **ɬǝrńa* >
Hungarian *arany* ‘gold’ ~ Mansi **tarəń* ‘copper’ ~ Khanty **ɬarńə* ‘copper’.

Similarly, the westernmost branches Saami, Finnic, and Mordvin seem to share some sound changes. However, the current evidence points also here to shared innovations between already regionally separate but still adjacent centers or pre-dialects. First, the merger of **δ̆* into **δ* intervocalically was possibly shared also by Mari, although it is also possible that the development of the two spirants in Mari differed from each other (Metsäranta 2020: 39). Uralic **ǝ* usually yields Mari **ü*, but sometimes it behaves like **a*,

which changed into Late Proto-Mari **o* regularly before velar consonants and occasionally in some other environments (Metsäranta 2020: 81, 314). However, as also Uralic **o* has been preserved in Mari before the velar nasal, there the change could have been directly **ę* > **o*.

Even the three westernmost branches do not fully agree on the development of Uralic **ę*: even though they share the changes **aj–ə* > **i̯j–ə* and **ę-a* > **a–a* (in Finnic **ę* and the **a* have totally merged), only Saami and Mordvin seem to share a later merger of **a–ə* into **o–a* (Aikio 2015: 39). However, it is to be expected that also pairwise changes appeared. Similarly, showing different pairwise distributions, Finnic shares with Mordvin the merger of the word-final **m* into **n*, and Finnic shares with Saami certain developments concerning round vowels in the second syllable (Kuokkala 2018).

Here is one example from the previous subsection showing that the West Uralic changes were later than the Late Proto-Indo-Iranian loanwords:

U **ćęta* '100' > WU **ćata* >
Saami **ćuotē* ~ Finnic **sata* ~ Mordvin **śadə̂*

It is also possible that some loanwords with Proto-Iranian **a* were originally substituted by Uralic **ę*, which then changed into West Uralic **a*, but we cannot know for certain, because at that stage there probably were no longer loanwords having regular cognates both in the western and other Uralic branches:

(U **węsa* 'calf' >) WU **vasa* >
Finnic **vasa* ~ Mordvin **vaz*
← Iranian **wasá* 'calf' (Holopainen 2019: 300–301)

There are also later Iranian loanwords borrowed even after some of the branch-specific sound changes, and often also the Iranian original is clearly younger and easy to distinguish from the Proto-Iranian word. Here I give only few examples:

Samoyedic **pu/ilə* 'bridge'
← Middle Iranian **puhl* 'bridge' < LPIIr/PIr **pr̥tu-* (Holopainen 2019: 195). If this word was borrowed before the Samoyedic sound changes, it would have been ***puj* in Late Proto-Samoyedic.

Mansi **širγV* 'sword'

← Iranian (Alanic) **cirγ* 'sword' < PIr **tiγra-* (Holopainen 2019: 259). Late Proto-Uralic **i* > **ä* in Mansi, so this borrowing is clearly later than the Mansi vowel changes. The initial consonant cannot reflect the Proto-Iranian stage.

Permic **das* '10'

← Iranian **das* '10' < PIr **daca-* (Holopainen 2019: 379). Late Proto-Uralic **a* > **u* in Permic, so this borrowing is clearly later than the Permic vowel changes. Also, Permic **s* cannot reflect Proto-Iranian **c*.

All the evidence presented in the preceding subsections will be considered when I set forth a new model for the disintegration of Late Proto-Uralic in Subsection 5.1. Consequences for the location of Late Proto-Uralic will be considered in Subsection 5.3.

3. Paleolinguistic arguments

Due to the tightened criteria, this section contains only few items. For example, the words for 'bee', 'honey', and 'oak' do not have cognates in the easternmost branches, while the words for 'fir' and 'larch' cannot be reconstructed for Late Proto-Uralic due to sound correspondences that are too irregular.

3.1. Uralic **sęksa* 'Siberian pine'

For a long time, tree names have been used to locate the Proto-Uralic homeland. Many trees are too widespread to be diagnostic (birch, pine, spruce, willow, alder, rowan, bird cherry), but there are two groups of trees spreading in opposite directions and meeting in the Volga-Ural Region: the western deciduous trees (oak, elm, maple, linden, hazel, ash) and the eastern coniferous trees (Siberian pine, fir, larch). Consequently, neither of these tree groups could disqualify the Volga-Ural homeland, but the western trees could disqualify the more distant Siberian homeland, and the eastern trees could disqualify the Upper Volga and the more western homelands, if their names could be reconstructed for Late Proto-Uralic. At the present, there is only one name for an eastern tree that fulfills the required criteria:

U **sę̮ksa* 'Siberian pine (*Pinus sibirica*, earlier *Pinus cembra sibirica*)' >
Permic **sus* 'Siberian pine; juniper (in Udmurt)' ~ Mansi **tī̮t* 'Siberian pine' ~ Khanty **ɪ̮̄γəɪ* 'Siberian pine' ~ Samoyedic **tī̮tåjŋ* 'Siberian pine' (UEW: 445; Aikio 2015: 60).

Even though there are no cognates in the westernmost branches, the regular sound correspondences point to a very old word. The distribution of the Siberian pine (AgroAtlas: *Pinus sibirica*) excludes Europe except for the north-easternmost part north from the Upper Kama Region, but it also excludes the southern half of West Siberia. While the pollen samples were earlier often dated by stratigraphy, now there are an increasing number of radiocarbon datings available. According to the new results, the Siberian pine appeared in the Upper Kama Region only ca. 1300 BCE (Lapteva et al. 2017: 330), so this piece of evidence narrows the Late Proto-Uralic homeland down to the Central Ural Region or the northern half of West Siberia.

3.2. Uralic **ćę̮lə* 'elm'

The following is the name for a western tree with the widest Uralic distribution:

U **ćę̮lə* 'elm (*Ulmus*)' >
Mordvin **śäľəŋ́* 'elm' ~ Mari **šolə* 'elm' ~ Hungarian *szil* 'elm' (UEW: 458–459; Aikio 2014: 67).
New possible cognate: Mansi **sī̮lt* 'linden bast' > TY TCh *sāľť*, KL *sāľt*, KM KU Pe VN VS VNZ LM *sę̮̄ľt*, LL *sę̮ľt* 'linden bast' (Kannisto 2013: 748a), LU So *sālt* 'willow bast?' (in compound words only; Kannisto 2013: 741b).

Aikio omits the otherwise regular Finnic cognate **salaga* 'crack willow (*Salix fragilis*) etc.' as a Germanic loanword. The Finnic **jalaga* 'elm' would be an otherwise suitable cognate, but here the problem is the lack of further examples of the required irregular change **ć* > ***j*. On the other hand, I have recently proposed a possible Para-Slavic etymology for the Mordvin word and the Finnic word **halaga* 'bay willow (*Salix pentandra*)', speculating on the possibility of including also the Finnic **jalaga* and **salaga* there (J. Häkkinen 2022: 132–133).

The Mordvin **ä* is unexpected but explainable, as Aikio (2020: 48–49) presents a parallel case for the change U **ę* > Mordvin **a* > **ä* next to **ć*: U **ęćə-* 'set' > Mordvin **aśəm* > **äśəm*. The Hungarian and Mansi words can be regularly derived from U **ćęlə*, but the final **-t* in Mansi should be explained as a secondary element. However, such unexplained extra elements are occasionally seen at the end of Uralic nouns and verbs in Mansi (cf. Aikio 2020: 6–8: **āľāt*, **-ūjt-*; 11: **ūləć*). In the Mansi varieties the palatalized **ľ* is more common than the expected **l*, but this does not appear to be a big problem.

In Mari, **ę* > **ü* is the most frequent outcome, although **ę* > **o* is common in front of velar consonants (Metsäranta 2020: 80–81), and it occurs occasionally also in other environments: U **δ̌ęmə* 'bird cherry' > Mari **ľom-bə* | U **ęppə* 'father-in-law' > Mari **owə* | U **lęmpə* 'pond, bog' > Mari **lo/åp* 'hollow, lowland' (Metsäranta 2020: 314–315). Therefore, the cognates of Mordvin, Mari, Hungarian, and Mansi can all be derived from the common proto-form.

The different meaning of the proposed Mansi cognate requires some attention. First, names for trees do not necessarily follow our modern logic of biological taxonomy, instead they can be motivated by the function of a tree. Bast was taken from both linden and elm, and the Finnish words *pärnä*/*pernä* and *kynneppää* can refer to both trees; the former word can also mean 'bast' (Vilppula 1984: 196–198; SSA 2: *pärnä*). Furthermore, *lehmus* 'linden' can mean 'soft wood' and *niini* 'bast' can mean 'linden' in some Finnic varieties (SSA 2: *lehmus*, *niini*). These words seem adequate parallel cases to justify the semantic shift 'elm' > 'bast' for the Mansi word.

Second, a semantic shift is to be expected when a language has spread outside the natural habitat of the tree. Mansi is presently spoken in Northwest Siberia, an area to which the elm subspecies have never spread; the easternmost extension ends in the Central Ural Region (AgroAtlas: *Ulmus glabra*, *Ulmus laevis*), and even to the south from the current Mansi region around Chelyabinsk only sporadic pollen finds appear (Lapteva & Korona 2012: 329).

Linden, on the other hand, has spread to Siberia beyond the Central Ural Region, advancing by the present time well past the Irtysh–Tobol confluence (AgroAtlas: *Tilia cordata*), although its presence has remained marginal there (less than 1% of the pollen sum; Volkova et al. 2016: 309). Linden is the primary bast tree in the Southern Mansi region. In the region of the Northern Mansi varieties along the rivers Sosva and Lozva even

linden is not present, so the meaning there has again shifted to 'willow bast'. On the European side of the Urals, the elm arrived in the Middle Kama Region from the southwest already ca. 5000 BCE (Shumilovskikh et al. 2020: 533).

Concerning the names for 'Siberian pine' and 'elm', they both cover only part of the Uralic branches. Naturally, languages tend to lose words when they are no longer needed (semantic shift is, of course, another option). Therefore, we would not even expect to find names for eastern trees in the westernmost languages or names for western trees in the easternmost languages. Consequently, the loss of the word for (or at least the meaning) 'elm' in Mansi, Khanty, and Samoyedic is quite as expected as the loss of the word for 'Siberian pine' in Finnic, Saami, and Mordvin. Such tree names have the best possible excuse for their disappearance from languages, compared to words from any other semantic category.

Against this background, a regular cognate for U **ćęlə* 'elm' in Hungarian and Mansi – the latter showing a credible semantic shift – justifies the Finno-Ugric status for the word. Admittedly, Samoyedic seems to be the first Uralic branch to form a regionally separate center, but the associative arguments (see Subsection 5.3) still require the presence of Samoyedic in the immediate vicinity of Finno-Ugric for a long time. Consequently, the name for 'elm' would exclude Siberia as a whole from the possible regions for the Late Proto-Uralic homeland. On the other hand, if we omitted both words due to a lack of cognates in the other end of the language family, we would no longer have any diagnostic tree names left in Late Proto-Uralic.

Interestingly, the eastern trees have different evidentiary value in the cases of the European vs. the Siberian homeland, because these trees originate in Siberia. Considering the Siberian homeland, names for the eastern trees could have appeared in the language already much earlier than at the actual Late Proto-Uralic stage. This in turn would lead to a paradoxical conclusion that it would no longer be necessary to locate Late Proto-Uralic in Siberia – it would be enough to locate the very distant Pre-Proto-Uralic there. Considering the European homeland in the Volga-Ural Region, the names for both the western and the eastern trees could not have been acquired before the trees themselves spread to easternmost Europe from the opposite directions. Therefore, in the Volga-Ural homeland the appearance of the tree names would probably be closer to the actual Late Proto-Uralic stage.

3.3. Uralic **wVć(k)V* 'metal'

This word is highly diverse in Uralic, and no single reconstruction has convincing cognates with wide enough distribution. Therefore, the dispersal of the word has recently been considered later than Late Proto-Uralic (Aikio 2015: 42–43). Grünthal et al. (2022: supplementary file: 12) have tried to defend a more suitable reconstruction for the Samoyedic cognate, **wäsä*, by assuming that the Nganasan cognate is irregular, but even if we accepted LPSy **wäsä* instead of **wäsa*, there would be no certain cognates with **ä* elsewhere: Saami **veaškē* would regularly come from Uralic **wećka*, and even though it might also come from **wäćka*, also **waćka* could be possible due to sporadic palatalization **a* > **e* caused by the adjacent **ć*. The Finnic **vaski* could come from both Uralic **waćkə* or **wäćka*.

Mordvin and Hungarian rather point to **a* in the first syllable, but Mari and Khanty even point to a secondary **a* not corresponding to the Uralic **a*. In Permic and Mansi the word only appears as the latter part of a compound word and has therefore been badly eroded, although **e* seems a possible original vowel there (J. Häkkinen 2023). In conclusion, the word cannot be reliably reconstructed for Late Proto-Uralic. Even if it were a Late Proto-Uralic word, its locational value would be weak, because the reconstructed meaning appears to be 'metal' in general, and copper was used widely in Northern Eurasia already several millennia before Late Proto-Uralic.

4. Other arguments

4.1. Lack of loanwords from certain languages

Nichols (2021: 355) and Grünthal et al. (2022: 8) write that the lack of Para-Baltic loanwords from the Fatyanovo Culture would testify against a homeland in the Volga-Ural Region. However, there are several reasons why such an argument is not valid.

First, we do not know for certain which language lineages were spoken within the Fatyanovo Culture. Even the Indo-Iranian lineage might have come from the Fatyanovo sphere: at least the Sintashta population was very similar to the Corded Ware populations both at the autosomal level (Saag et al. 2021: 5) and at the Y-chromosomal level (Underhill et al. 2014: 3, 5; Saag et al. 2021: 3).

Second, even though the Northwest Indo-European lineage was probably spoken (also) within the Fatyanovo Culture, it remained phonologically archaic up to the 2nd millennium BCE (see Subsection 2.3), so it is anachronistic to require Para-Baltic loanwords at the 3rd millennium BCE. Instead, there are plenty of proposed Archaic and Northwest Indo-European loanwords which can be connected to that cultural context, although at the moment we cannot stratify them reliably.

Third, the lack of loanwords is not a valid argument, because there is no universal law which would compel adjacent languages to borrow certain words (or any words at all) from each other. Borrowing a word is a complex sociolinguistic situation – it is not an automatic consequence of two languages spoken close to each other. Sometimes words get borrowed very easily, sometimes not.

Fourth, even if there originally were loanwords, there is a possibility that the speakers in the contact zone shifted their language to another, thus losing such loanwords along with their whole language. The Middle Volga was an especially complex area, where at many times several overlapping cultures coexisted (Carpelan & Parpola 2001: 86–89). The situation was not necessarily as simple as the Uralic speakers borrowing loanwords from all of their neighbors and then that language expanding westwards, but there could have occurred language shifts back and forth, following several different cultural influences in different directions. To complicate the situation even more, there are also traces of Paleo-North European languages in the region (J. Häkkinen 2009: 47–49; Aikio 2015: 43–47; Zhivlov 2015), so the closest neighbor of the Uralic speech community in the west was not necessarily an Indo-European language.

Consequently, even though the Fatyanovo Culture belonged to the Corded Ware Cultures, we should not assume that there was only one language present within its whole wide region. Balto-Slavic was only one of possibly several languages spoken in the region, and it becomes recognizable only during the 2nd millennium BCE (Häkkinen 2022: 138–141; Kallio, forthcoming).

Another related argument can be seen in Grünthal et al. (2022: supplementary file, page 13), where the authors write that because the only metal name in Proto-Uralic meant only 'metal' in general, this could be seen as an argument against the homeland close to the rich metallurgical center in the Southern Urals. However, the authors ignore the fact that there were metallurgical centers also near the assumed Siberian homeland candidates

and that there were Indo-Europeans also in South Siberia already since 3300 BCE within the Afanasyevo Culture (which knew bronze metallurgy and even meteoritic iron; see Koryakova & Epimakhov 2006: 188–189). If the Late Proto-Uralic homeland was anywhere near there, should we not expect (Indo-European) metal names borrowed into Uralic also there? Of course, such expectations are in any case unfounded, as argued above.

Moreover, metal names are mainly temporal arguments, and they would have value for locating the Proto-Uralic homeland only if a word could be shown to have been borrowed from a certain locatable language already into Late Proto-Uralic. However, there are no metal names fulfilling these criteria (see Subsection 3.3). I have recently proposed an Indo-Iranian origin for the first part of a Mansi–Permic compound metal name, but it is in any case post-Proto-Uralic (J. Häkkinen 2023).

4.2. The taxonomic argument

The taxonomic argument is based on the deepest division within a language family. However, it is easy to find counterexamples like the Turkic, Indo-Iranian, or Celtic homelands, which are not located in the region where the deepest division is nowadays observed – not even within the present distribution of these branches (Kallio 2015b: 84). This argument could only work in cases in which the homeland falls within the present region of the language family – and not necessarily even always when that condition is fulfilled. Consequently, as this argument requires that we already know where the homeland is before we can apply it reliably, it is practically redundant.

Furthermore, the views on the taxonomic structure of the Uralic language family have been notoriously diverse: there are many different results based on many different pieces of evidence. However, sharing of the inherited lexicon, not to speak of only a short list of selected items (like numerals) thereof, cannot be considered a reliable datum, because there are possible distorting processes leading to either increased or decreased lexical sharing between branches (J. Häkkinen 2012). Therefore, the phonological level should always be taken as the starting point, although the lexical level cannot be totally dismissed, as will be shown with the new model for the disintegration of Proto-Uralic (Subsection 5.1).

Related to the taxonomic argument, Saarikivi (2022: 57) writes that the Ugric group is more diverse than other Uralic groups, and therefore the

Uralic homeland could be located close to the homelands of Hungarian, Mansi, and Khanty. However, Ugric is not necessarily a taxonomic branch but instead a unit of areal convergence between the three branches. Even if it were a true branch, its greater diversity could not be interpreted as evidence about its greater temporal depth, because the structure and depth of a branch are purely stochastic. The variables are: (1) How long after the proto-language did the branch-specific sound changes occur? (2) How many consecutive divisions occurred within the branch? (3) How many of these sub-branches survived, without being leveled by closely related dialects or replaced by other languages?

There are several possible results from these variables. It is very well possible that most of the branches within a language family descended from a single recent macro-branch; and it is also possible that a single branch would have greater temporal depth within a language family than a macro-branch with several sub-branches. How it really was, cannot be deduced straightforwardly from the taxonomic structure of the language family – width cannot testify for depth.

Saarikivi (2022: 57) is probably correct when he writes, "If there is such a thing as Proto-Ugric, it is, without doubt, even older than Proto-Samoyedic." This means that the disintegration of Proto-Ugric (if it was a branch) would be older than the disintegration of Late Proto-Samoyedic. Yet this branch-internal disintegration does not necessarily correlate with the external (family-wide) disintegration, for the reasons stated above.

4.3. Distances and tendencies

It has been argued that the Volga homeland is improbable, because it would require a movement over a huge distance to the region of Late Proto-Samoyedic (Grünthal et al. 2022: 8). Nevertheless, the known regional distribution of the Uralic branches is what it is, and no matter where the homeland was, some branch has had a greater distance to traverse than some other branch. We could equally well use this argument against the homeland in South Siberia, claiming it implausible that Samoyedic has not moved at all, while the Saami branch has moved over a huge distance. If anything, the distance argument could only be seen to support a somewhat central homeland, where the total distance for any single branch is not extremely long – that would be somewhere around the Volga-Ural Region.

It has also been argued that the general direction of movements in Eurasia has been from the east to the west (Grünthal et al. 2022: 8). However, it is easy to find counterexamples, like the spread of Pre-Proto-Tocharian and Indo-Iranian. A tendency is not a law, and therefore its evidentiary value is weak.

Quite similar is an argument based on language sinks: such regions would pull languages rather than push them, and as a language sink, the Middle Volga Region would be an improbable homeland (Grünthal et al. 2022: 8). Still, a proposed Uralic homeland in the Minusinsk Basin in South Siberia is also a language sink (Nichols & Rhodes 2018: 8). Again, this is merely a tendency, not a law.

Moreover, at least in the archaeological data it is well known that the region of the Upper–Middle Volga has for a long time been an expansion center (Carpelan & Parpola 2001: 79–83). Certainly there were also languages connected to these consecutive cultural expansions, but all those earlier languages later disappeared under the Uralic expansion. Even though the earlier languages have not survived, we should not ignore their earlier existence: by constructing tendencies based only on the very few surviving language families, we cannot reach the complex reality of the past.

4.4. Lack of a non-Uralic substrate

This could be a potentially illuminating argument, but at the present state of the art, we know too little about the lost languages, their distribution, and how to even trace them properly. This criterion also works in one direction only: the presence of a non-Uralic substrate in a language can testify that there is no continuity from Late Proto-Uralic in that particular region, but the absence of a visible non-Uralic substrate cannot testify reliably that Late Proto-Uralic was spoken in that region.

Moreover, this criterion could only work for language families within which some extant branch has remained in the original location of the homeland, but in order to be able to fulfill that demand, we should already know where the original homeland was. Consequently, this argument is redundant.

5. Late Proto-Uralic – when and where?

After all the relevant arguments and pieces of evidence have been considered, it is a time for a conclusion. As byproducts, a new model of disintegration and taxonomy will be proposed for the Uralic language family (Subsection 5.1), followed by the most resolute possible dating for the Uralic reconstruction stages (Subsection 5.2). Next, the Late Proto-Uralic homeland will be located (Subsection 5.3). Finally, the rules for connecting the linguistic results to the archaeological results will be briefly discussed (Subsection 5.4) and possible counterparts for the Uralic reconstruction stages will be proposed (Subsection 5.5).

5.1. The disintegration of Proto-Uralic

The disintegration of Proto-Uralic is connected to both the dating and the locating of the proto-language. Linguists who have touched on the Uralic homeland problem (myself included), have not always been able to properly distinguish between different levels of testimony: evidence from the lexical level, from the level of sound substitution of loanwords, and from the level of sound changes. This has sometimes led to favoring one level of evidence over another in order to solve an apparently contradictory picture. In the present critical examination I aim to remedy this problem, and to develop a model which not only allows us to be aware of and distinguish between all the levels of linguistic evidence, but also to use them together to "triangulate" for the most accurate chronological reconstruction possible.

In Section 2 some seemingly contradicting results from the Indo-Iranian loanword layers were presented:

1. There are no certain Indo-Iranian loanwords in Late Proto-Uralic.
2. Based on the lexical level, Samoyedic was already outside the Indo-Iranian contact zone during the Early Proto-Indo-Iranian loanword layer.
3. Based on the sound substitutions, Samoyedic borrowed loanwords together with the other Uralic branches and even agreed with their sound substitutions during the Late Proto-Indo-Iranian loanword layer.
4. Based on the sound substitutions, different Uralic branches show parallel borrowings from the Proto-Iranian loanword layer onward.
5. Based on the (macro-)branch-specific sound changes, the phonological developments began only after all the previous stages.

A new model of disintegration is needed to take all this evidence into consideration. I propose the following alternative labels for these five reconstruction stages:

U1 = LPU (Late Proto-Uralic) = uniform proto-language
U2 = AU (Ancient Uralic) = reorientating proto-language
U3 = BU (Bicentric Uralic) = bicentric proto-language
U4 = CU (Common Uralic) = substitutionally separated pre-dialects
U5 = DU (Diverging Uralic) = phonologically separated proto-dialects

It is a matter of personal taste whether one favors numerical or alphabetical labels. It is also debatable whether the first three stages could perhaps be included within the concept of Proto-Uralic, but I find it most clarifying to give every stage its own label. The label Common Uralic was used by Grünthal et al. (2022), but the other labels after the stage U1 are new. Figure 2 illustrates the disintegration of Proto-Uralic.

Centers are part of the same speech community, although regionally separated. Pre-dialects are regionally separated and show independent sound substitutions. Proto-dialects are regionally separated, show independent sound substitutions, and show (macro-)branch-specific sound changes. The next stage would be the branch-specific protolanguages, divided into early, middle, and late proto-stages, if necessary.

Familiarity with family-tree modeling might prevent us from pursuing or comprehending such a model, because in a line-drawn family tree, one trunk abruptly divides into two branches. However, linguistic reality is rarely so simple. A more adaptable illustration for the more complex process of disintegration is a "family funnel", which allows us to stratify features alternating between wide and narrow distribution (Figure 3).

Naturally, later contact phenomena and convergence by chance (like erosion or loss of vowels in unstressed syllables) can occur between branches. Here the focus is only on features so ancient that they have had an impact on views about the taxonomic model of the Uralic language family. An abrupt disintegration means that one branch has immediately moved further from others, and a rigid disintegration means that no shared isoglosses appear between branches after the initial division. Based on the evidence from the Indo-Iranian loanword layers, the disintegration of Late Proto-Uralic was neither abrupt nor rigid.

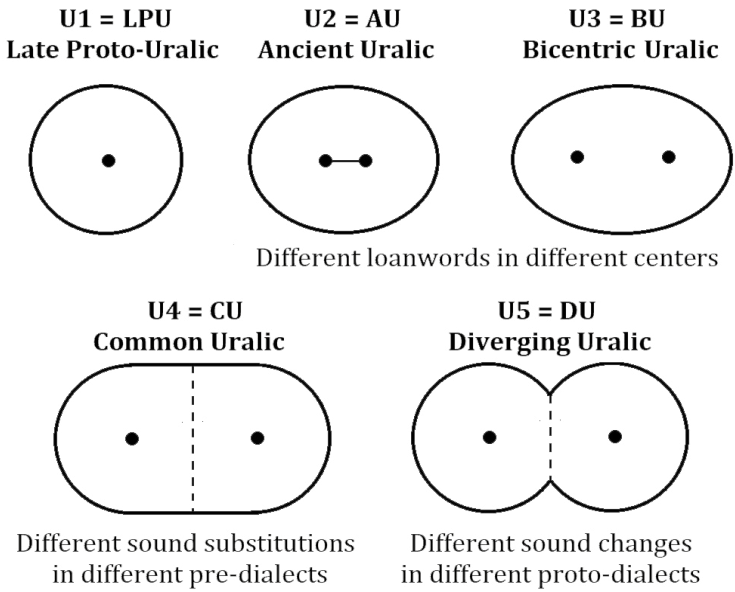


Figure 2: Disintegration of Proto-Uralic: five successive reconstruction stages based on the evidence from the Indo-Iranian loanword layers

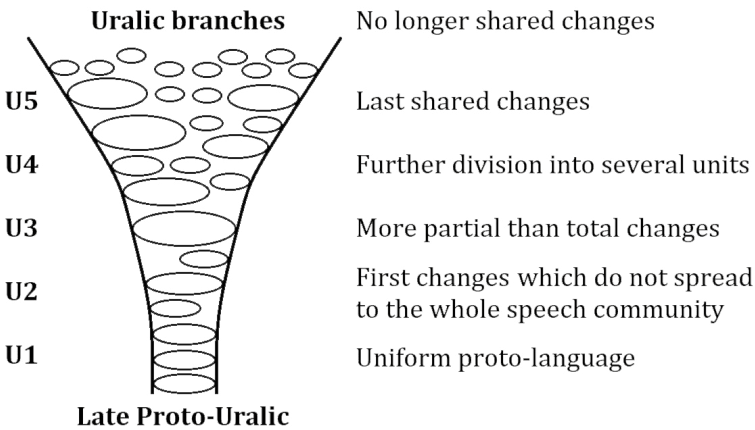


Figure 3: The family funnel illustrates non-abrupt, non-rigid disintegration of a proto-language. Isoglosses (oval discs) may contain information from different levels of language: lexicon, sound substitutions, and sound changes.

5.2. Indo-Iranian evidence for dating Late Proto-Uralic

The absolute chronology of Indo-Iranian is quite firm in the later end, anchored by the following pieces of evidence, from the latest to the earliest:

1. The first Indic and Iranian writings appear at the mid-first millennium BCE, and they already show many branch-specific sound changes.
2. The Indic language of the *Ṛgveda* is considerably more archaic than Classical Sanskrit, even though the written attestations are not earlier, and thus the oral formation of the *Ṛgveda* has been dated already to the end of the 2nd millennium BCE (Cardona 2017). The same goes with the Iranian Old Avestan language as compared to the Younger Avestan language (Skjærvø 2017). Both of these more archaic varieties were transmitted in liturgical contexts for a long time before they were written down, while at the same time colloquial Indic and Iranian varieties went through more phonological changes.
3. Indic words in the Mitanni and Hittite writings from ca. 1400 BCE are even more archaic than Vedic Sanskrit, close to Proto-Indic (Witzel 2001: 49).
4. The chariot vocabulary shared by Indic and Iranian and certain ritualistic features described in the *Ṛgveda* and *Avesta* are best matched by the archaeological remains of the Sintashta Culture in the Southern Urals ca. 2100–1800 BCE (Anthony 2007: 408–411). Late Proto-Indo-Iranian can therefore be dated and located there.
5. Beyond that, the dating becomes more imprecise. Disintegration of Late Proto-Indo-European is dated to the late 4th millennium BCE (Anthony & Ringe 2015), but the Indo-Iranian sound changes probably began to occur only during the latter half of the 3rd millennium BCE.

The datings of these Indo-Iranian stages can be transferred to the Uralic side through the disintegration model based on Indo-Iranian loanword layers (Table 1).

Table 1: Uralic reconstruction stages and their approximate dating transferred from the Indo-Iranian chronology through the loanword layers

Indo-Iranian stage	Dating and Uralic stage	Disintegration of Proto-Uralic			
Archaic Indo-European	~2800 BCE U0 = EPU	Early Proto-Uralic			
Archaic Indo-European	~2500 BCE U1 = LPU	Late Proto-Uralic			
Early Proto-Indo-Iranian	~2300 BCE U2 = AU	Finno-Ugric			Samoyedic
Late Proto-Indo-Iranian	~2000 BCE U3 = BU	Finno-Ugric			Samoyedic
Proto-Iranian/ Indic	~1800 BCE U4 = CU	Pre-dialects			
Archaic Iranian/Indic	~1500 BCE U5 = DU	West Uralic		East Uralic	
Archaic Iranian/Indic	~1200 BCE	Saami, Finnic, Mordvin	Mari, Permic	Hungarian, Mansi, Khanty	Samoyedic

5.3. Evidence for locating Late Proto-Uralic

For locating purposes, the earlier stages of the Indo-Iranian lineage carry the most weight, being closer in time to Late Proto-Uralic. Based on the connection between the chariot-related vocabulary and the ceremonial practices described in the *Ṛgveda* and *Avesta* on the one hand, and the archaeological remains of chariots and graves on the other hand, Late Proto-Indo-Iranian is connected to the Sintashta Culture (ca. 2100–1800 BCE), and the language only spread beyond the river Tobol around 2000 BCE, when the Sintashta-rooted Andronovo Complex spread to Southwest Siberia and Northern Central Asia (Anthony 2007: 389–390, 397; E. Kuz'mina 2007: 451).

All the stages leading to Late Proto-Indo-Iranian developed on the European side of the Urals (E. Kuz'mina 2007: 305). There are no serious challenging views for the original European homeland of the Indo-Iranians, which is significant considering the location of the Uralic stages. Traditionally the Indo-Iranian lineage has been connected to archaeological

cultures of the European steppe, like the Poltavka Culture (Mallory & Adams 1997: 439–440), but also influence from the Abashevo Culture into the Sintashta Culture has been acknowledged (Anthony 2007: 382–387).

Even though the Poltavka Culture began already ca. 2600 BCE (Morgunova & Khokhlova 2013), there is no reason to believe that the Indo-Iranian phonological developments began so early. Most of the Indo-Iranian loanwords were borrowed into Uralic only after the vowel merger in Late Proto-Indo-Iranian, and the rest of them need not be much earlier, reflecting already most of the Middle Proto-Indo-Iranian sound changes (see Subsection 2.5). Therefore, even though already the Poltavka Culture spread beyond the Southern Urals, it is highly uncertain to try to explain the Early Proto-Indo-Iranian loanwords in Uralic by its extension beyond the Urals, as implied by Grünthal et al. (2022: 10). More likely Early Proto-Indo-Iranian developed only slightly before the appearance of the Sintashta Culture in the Southern Ural Region ca. 2100 BCE.

Moreover, it is not certain that the Indo-Iranian lineage can be connected to the Poltavka Culture at all. The Sintashta Culture has roots both in the steppe cultures and in the Abashevo Culture, and the latter has roots also in the Fatyanovo Culture (belonging to the Corded Ware Cultures; Anthony 2007: 383). Based on the recent genetic results, the Sintashta population was most similar to the populations of the Fatyanovo Culture and the other Corded Ware Cultures, both at the autosomal level (Saag et al. 2021: 5) and at the Y-chromosomal level (Underhill et al. 2014: 3, 5; Saag et al. 2021: 3). Consequently, we cannot exclude the Fatyanovo Culture as the possible origin of the Indo-Iranian lineage.

For Uralic, the exact cultural counterpart for Early Proto-Indo-Iranian is irrelevant, because all the candidates (the Poltavka, the Abashevo, and the Fatyanovo-Balanovo Culture) coexisted in just about the same Volga-Ural Region at the late 3rd millennium BCE. The partial overlapping of the Fatyanovo-Balanovo and the Abashevo Cultures could explain the regionally spread features shared by Balto-Slavic and Indo-Iranian, like satemization and the *ruki*-rule (see Section 2.3), as suggested by Parpola (2022: 264).

Although the Early Proto-Indo-Iranian loanword layer directly requires locating only the Finno-Ugric center in the vicinity of the Southern Urals (likely to the north from the Indo-Iranians, in the Central Ural Region), there are associated arguments requiring also the presence of the Samoyedic center right next to it for a long time:

1) Samoyedic still shares some Late Proto-Indo-Iranian loanwords with other Uralic branches during the stage U3, showing in different words the same arbitrary substitute for the LPIIr **a* as the Finno-Ugric branches. These shared loanwords would be impossible to explain if Samoyedic were already located in South Siberia (see also Kallio 2015b: 82, footnote 5).

2) Samoyedic appears to have participated in the chain of three subsequent sibilant changes shared with Hungarian, Mansi, and Khanty, which requires the presence of Samoyedic next to the three other branches (which descend from the Finno-Ugric center) still during the stage U5.

3) Hungarian, Mansi, and Khanty share Early and Late Proto-Indo-Iranian loanwords (with the same sound substitutions) with the western branches, so they must still be located around the Central Ural Region through the stages U2 and U3. After that, they could have moved to Siberia together with Samoyedic, in which case the East Uralic sibilant changes at the stage U5 could have occurred in Siberia. On the other hand, it is also possible that all these branches were still in the Central Ural Region during that stage.

We may note a strikingly compatible pattern between the distribution of the Indo-Iranian loanwords in Uralic and the reconstructed dispersal of Indo-Iranian: first, during the stage U2, when Early Proto-Indo-Iranian was spoken in the Volga-Ural Region, the Samoyedic center was outside the contact zone. Later, during the stage U3, when Late Proto-Indo-Iranian spread to the east and was spoken in the Southern Trans-Urals between the headwaters of the rivers Ural and Tobol within the Sintashta Culture, the Samoyedic center got involved in the Indo-Iranian contacts. Therefore, we can locate the Finno-Ugric center in the western part of the Central Ural Region and the Samoyedic center in the eastern part of the Central Ural Region (see Figure 4 below).

The crucial question is whether the uniform Late Proto-Uralic (the stage U1) was spoken in that very same area or somewhere else. It seems impossible that the disintegration of Late Proto-Uralic could have occurred in some distant location, like around the Middle Irtysh Region or even beyond. If that was the case, Samoyedic would have remained there, while only Finno-Ugric would have moved to the Central Ural Region, and we could not explain how Samoyedic could share with Finno-Ugric some Late Proto-Indo-Iranian loanwords with the very same arbitrary sound substitutions at the stage U3. Neither could we explain how Samoyedic could have participated in the chain of three successive sibilant changes

together with the ancestral stages of Hungarian, Mansi, and Khanty at the stage U5.

Consequently, there is no avoiding the inevitable conclusion: the disintegration of Late Proto-Uralic must have begun in the Central Ural Region. Of course, it is possible that its immediate ancestor arrived from South Siberia only slightly before the beginning of the disintegration, but that possible stage in South Siberia could not be labeled Late Proto-Uralic. Earlier stages of the Uralic lineage fall beyond the scope of this scrutiny, but hopefully future research produces more results about that topic. At the moment scholars might locate the distant Pre-Proto-Uralic both in Siberia (Aikio 2022: 26–27) and in the Volga-Ural Region (Parpola 2022).

This conclusion has an important consequence: through the above-presented associative arguments, every piece of evidence which is absent in Samoyedic yet shows a Finno-Ugric distribution, phonologically regular enough, and semantically credible cognates, now has an impact on the location of Late Proto-Uralic itself. By anchoring the Finno-Ugric center, any such piece of evidence anchors also the Samoyedic center in the immediate vicinity of the Finno-Ugric center until after the stage U3, and in the vicinity of the other East Uralic branches until after the stage U5. I shall label this factor the "Uralic bundle effect".

While the Indo-Iranian loanword layers pull Late Proto-Uralic to the west, the area of the Siberian pine holds the reins for that pull. Even today, the natural habitat of the Siberian pine on the European side of the Urals does not reach south from the Upper Kama Region, although random occurrences might appear in a wider area in the north-eastern part of European Russia (AgroAtlas: *Pinus sibirica*). However, random occurrences can hardly explain the preservation of the tree name in the daughter languages for over four millennia – clearly the languages must have been spoken very close to the natural habitat of the Siberian pine.

Pollen of this tree appears in the Upper Kama Region only ca. 1300 BCE (Lapteva et al. 2017), which requires the presence of the speakers of Late Proto-Uralic firmly in the Central Urals, excluding the Middle Volga homeland and every homeland candidate further to the west. However, the South Siberian homeland encounters problems, too. Even though the Siberian pine has for a long time been present in the Sayan Region (Blyakharchuk & Chernova 2013), it is not present in the southern part of West Siberia (AgroAtlas: *Pinus sibirica*), through which the Uralic language could probably be supposed to have extended were the homeland

in South Siberia. To secure the presence of the Siberian pine along the expansion route, the language should have moved first to the north along the Yenisei, and only then to the west.

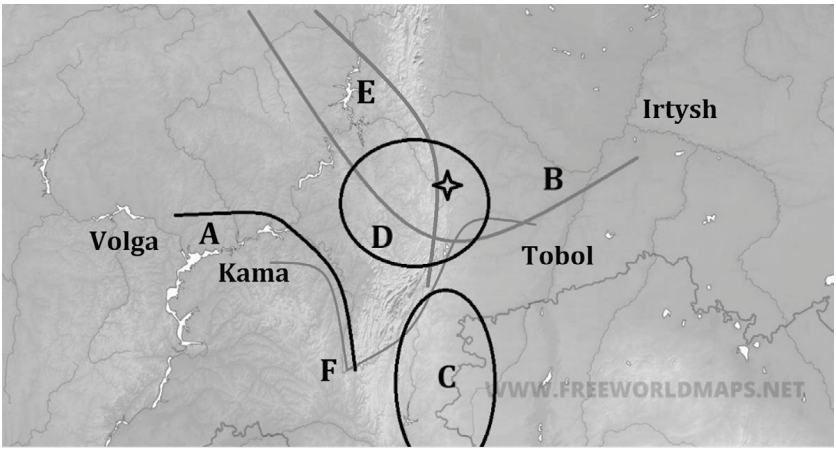
The Finno-Ugric name for 'elm' excludes the West Siberian homeland, but the Central Ural Region and the European homelands are acceptable. Perhaps the word was never even adopted into Samoyedic, but in any case, the "Uralic bundle effect" requires Samoyedic in the immediate vicinity, i.e. in the eastern slopes of the Central Urals. During the 3rd millennium BCE, the Central Ural Region was the only region where the Siberian pine and elm met each other (Figure 4).

However, names for these trees did not necessarily appear in the language at the same moment. It is possible that one of them appeared earlier in Pre-Proto-Uralic either in the more western or in the more eastern region, and the other one was adopted later. Nevertheless, no matter which scenario we favor, the evidence always pulls Late Proto-Uralic back to the Central Ural Region.

Even if one rejects the name for the elm as uncertain due to the semantic shift in Mansi, the final result would not change. In that case, the name for the Siberian pine could have been borrowed already earlier in Siberia, but the Indo-Iranian loanword layers still require Late Proto-Uralic and the subsequent stages in the Central Ural Region (Figure 4). The Koptyaki Culture appears to have been in the right place at the right time (see Subsection 5.5).

The boundaries of the Siberian pine, elm, and forest-steppe on the map in Figure 4 are based on the present distribution, but the latter two have not changed for many millennia, although the boundary of Siberian pine was somewhat further to the east during the Late Proto-Uralic stage, as described above.

In the Central Trans-Urals, the forest-steppe reaches up to the river Iset (the boundary as drawn in O. Kuz'mina 2021: 1209) and has remained rather stable for several millennia, since long before the Late Proto-Uralic stage (Lapteva & Korona 2012: 329–330). Distinguishably Mansi placenames reach from the present Mansi region to the south, between the headwaters of the Neiva and Iset (Matveev 2011: 445) – this region is exactly on the Central Ural Passage. Consequently, there is no need to locate the ancestral stages of Hungarian, Mansi, and Khanty in any more southern environment in order to explain the horse-related vocabulary.



A = The north-eastern boundary of the Abashevo Culture (early stage)
B = The southern boundary of natural habitat of Siberian pine
C = The approximate area of the Sintashta Culture
D = The approximate area of the Koptyaki Culture
E = The north-eastern boundary of natural habitat of elm
F = The northern boundary of forest-steppe
✧ = The Central Ural Passage

Figure 4: The most probable homeland for Late Proto-Uralic based on evidence from Indo-Iranian loanwords (A, C) and the names 'Siberian pine' (B) and 'elm' (E). The Koptyaki Culture (D) matches the required spatial and temporal coordinates. Forest-steppe (F) also reached to the area, explaining the Ugric horse-related vocabulary.

5.4. Connecting linguistic and archaeological results

In the following subsection I consider possible extra-linguistic counterparts for the Uralic dispersal. It must be emphasized that what is suggested here, is not a result as much as a starting point for further research. Nevertheless, even the starting point requires certain conditions: (1) that we accept the linguistic results and (2) that we can find an extralinguistic counterpart in the right place at the right time, its later stages spreading in the right direction(s).

Methodological pitfalls of multidisciplinary surveys have been recognized for a long time in Indo-European and Uralic studies (e.g., Mallory 1989; 2001; Aikio & Aikio 2001; Heggarty 2007; J. Häkkinen 2010; Saarikivi & Lavento 2012), and I shall not delve deeper into the methodology here. I only mention a few important points.

First, an archaeological culture could contain several language communities, and a language community could correspond to several cultures. Therefore, if we consider an ancient language as a phenomenon tightly following cultural boundaries or distribution of cultural features, we are not in reality even dealing with language but with some imaginary pseudo-linguistic level: an abstract projection of cultural boundaries, erroneously labeled as "language".

Second, cultures are usually polythetic, which means that a distribution may vary from item to item. How then could we ever guess to which an item a language would best correspond? How could we know whether an ancient speech area matched better with a distribution of ceramic pots, bronze axes, or certain type of graves?

Third, a correlation between a language and an archaeological or a genetic phenomenon is always only momentary. In a different place or at a different time the same phenomenon could be connected to a different language. This is an inevitable conclusion from the fact that language is not inherited dependent on any cultural or genetic phenomena. Assuming otherwise is again dealing with some irreal pseudo-linguistic level. For every step of a language expansion, the counterpart should be looked for independently.

Fourth, there are always several possible counterparts for language. A language always has only one genealogical root (except for real mixed languages), while cultures and populations usually have several roots, and so do their ancestral cultures and populations, etc. Therefore, when trying to follow a language back in time, the probability to choose the right counterpart grows cumulatively lower by each step beyond the starting point (the initial spatial-temporal correlation).

Fifth, archaeological continuity usually corresponds to linguistic discontinuity: archaeological continuity is to some extent evident everywhere, yet the linguistic landscape is mostly a result of quite recent language expansions. The wider the area of the language family, the lower the probability that any random region was the original homeland.

One cannot discern language from culture or DNA, and archaeology or genetics do not have methods for studying language. If language is included in a multidisciplinary survey, then the most reliable linguistic results must be taken as the starting point. If there appears a discrepancy in time or space between the results of different disciplines, one cannot ignore the linguistic results and keep on claiming that a certain language must be dated earlier or located in another place in order to save that particular correlation. The only scientific way is to acknowledge the discrepancy and comprehend that clearly this extra-linguistic phenomenon is a poor match for this particular language. The only way forward is to find a better-matching candidate without discrepancies in time, space, or direction of expansion.

Consequently, when in the following subsections I will present an archaeological phenomenon as a possible counterpart for a language, it only means that there appears an apparent spatial-temporal match. I do not intend that this language is confined within the limits of such a phenomenon, nor that this language is the only possible language within those limits, nor that this language is transmitted to the following generations along with certain archaeological phenomena, nor that this language can be assumed to descend from a certain local or non-local ancestor of that appointed counterpart. Nevertheless, even a connection as thin as assumed here is still a connection to more concrete prehistorical events, which makes it easier to comprehend the context where the speakers of this language lived.

On the Indo-European side, there are some fortunate anchors between the linguistic and archaeological realities, which connect a certain language in a certain place and time: the wagon vocabulary in Late Proto-Indo-European, finding its counterpart in the remains of early wagons within the Yamnaya Culture in the late 4th millennium BCE (Anthony & Ringe 2015), and the chariot vocabulary in Indo-Iranian, finding its counterpart in the remains of the first chariots within the Sintashta Culture ca. 2000 BCE (Anthony 2007: 408–411). On the Uralic side we are not so fortunate, but we can anchor our reconstruction stages through the Indo-European loanword layers, as was demonstrated in Subsection 5.2.

5.5. The Koptyaki Culture and the Seima-Turbino Network

The Koptyaki Culture in the Central Ural Region occupied the natural trade route over the Central Urals: the plateau between the adjacent headwaters of the rivers Iset in Siberia and Chusovaya in Europe. Sites of the culture show bronze items of both the Samus-Kizhirovo type and the Seima-Turbino type, and its chronology and origin remain so far rather unclear. It probably derives partly from the local Ayat Culture, partly from eastern and southern influences (Korochkova et al. 2010; Korochkova 2019; Grigoriev 2021: 22). The Koptyaki Culture had contacts with the cultures to the south, unlike the contemporaneous cultures in the West Siberian southern taiga zone (Korochkova 2012: 146). This agrees with the Indo-Iranian loanword layers in Uralic.

Within the Koptyaki Culture, considerable variation is visible from the Middle Kama Region in the west to the Tobol Region in the east, but the ceramics are considered the common denominator (Korochkova 2019: 734). Chronologically probably only the later stages U3–U5 could be connected to the Koptyaki Culture itself, and Late Proto-Uralic might be connected to its poorly known local predecessor (local for the reasons explained in Subsection 5.3). The Koptyaki Culture was succeeded by the Cherkaskul and the Mezhovskaya Cultures (see below; Grigoriev 2021: 22).

Interestingly, the Central Ural Plateau was exactly the passage through which the main trade route of the Seima-Turbino Network ran. According to the distribution maps of the Seima-Turbino items, the main river route from the Sayan Region to Europe was the Irtysh–Tobol–Iset–Chusovaya–Kama–Volga (Carpelan & Parpola 2001: 99–111). Another route from Tobol was slightly more northern but ended very close in the watershed area, where the recently found site of Shaytanskoye Ozero is located: along the Tobol–Tura–Neiva–Revda–Chusovaya (Chernykh et al. 2017: 48). One might suspect that the people of the Koptyaki Culture profited greatly from the use of these routes by the traders of the Seima-Turbino Network.

It was Carpelan who first proposed that the Seima-Turbino Network was connected to the spread of Samoyedic from the Volga Region to South Siberia (Carpelan & Parpola 2001: 109). Kallio (2006) considered its connection to the spread of Proto-Uralic, followed by J. Häkkinen (2009) and Parpola (2012b; 2017). Recently Grünthal et al. (2022) connected it to the spread of Finno-Ugric westwards from West Siberia.

While the Abashevo metallurgy derived from the Circum-Pontic metallurgical tradition and was based on arsenical bronze, the Seima-Turbino metallurgy was based on tin bronze. Tin came from and was first utilized in the Altay-Sayan Region, but copper deposits were found in a wider area in the Ural Region and Kazakhstan (Koryakova & Epimakhov 2006: 28–29).

The initial stage of the network in Southwest Siberia ca. 2200–2000 BCE cannot be associated with the Uralic speakers, for the reasons presented in Subsection 5.3, but perhaps the later stages in Europe could. After securing the passage over the Central Urals (datings from the Shaytanskoye Ozero site right on the passage are ca. 2000–1650 BCE; Korochkova 2019: 733), the Seima-Turbino Network established new centers in the Kama Region and the Mid-Upper Volga Region ca. 1900–1600 BCE (Chernykh et al. 2017: 51–52; Marchenko et al. 2017).

Interestingly, the Pepkino Kurgan on the southern side of the Middle Volga in Mari El contained the remains of 28 Abashevo warriors who were killed probably around 2000 BCE (Chernykh et al. 2017: 53). As the Seima-Turbino sites appear further to the west soon after that, it seems possible that the Seima-Turbino Network managed to take control over the Volga route. However, there are traces of hostilities also between the people of the Abashevo and the Balanovo Cultures, so it is only speculation that the Abashevo mass grave and the expansion of the Seima-Turbino Network would be causally connected.

At the same time with the Seima-Turbino expansion, widespread east–west contacts can be seen also in the shared features of ceramics between the Krotovo Culture (in the Middle Irtysh Region), the Garino Culture (in the Middle Kama Region), and the Chirkovo Culture (in the Mid-Upper Volga Region; Vybornov et al. 2019: 19). These locations match with the Seima-Turbino centers in Rostovka, Turbino, and Yurino, respectively.

As ceramics in prehistoric Eurasia are widely considered to have been the realm of women, perhaps this spread of certain features in ceramics reflects the exchange of brides between the groups participating in the Seima-Turbino Network, or perhaps the bronze traders brought their families with them. Together the shared extension of both bronze items and ceramic wares seems to testify to the movement of both men and women within the wide network, which offers an adequate background also for a possibility of a language expansion.

Could the Uralic speakers, as the gatekeepers of the Central Ural Passage, have demanded their share of the network on the European side of the Urals? They already had established contacts with the neighbors to the south and southwest, as testified by the Indo-Iranian loanwords. Joining in the network and expanding it could be the ultimate reason behind the Uralic expansion to the west. The Middle–Upper Volga centers could correspond to the West Uralic branches (Saami, Finnic, Mordvin, West Chudic, and Meryanic, and perhaps Mari), and the Middle Kama center could correspond to the Central Uralic branches (Permic and perhaps Mari). The original region of Mari is uncertain: this branch seems to share a surprisingly low number of pairwise words and common innovations with both Permic (Metsäranta 2020: 285–286, 290–291) and Mordvin (Itkonen 1997: 259), so perhaps Mari was for a long time separated from those branches by unknown Uralic or non-Uralic neighboring languages.

The East Uralic branches (Hungarian, Mansi, Khanty, and Samoyedic) could be connected to the Cherkaskul Culture, which existed in the very same Central Ural Region between Middle Kama and Tobol after the Koptyaki Culture, ca. 1800–1500 BCE, as suggested by Parpola (2012a). The influence of this culture is visible also in the Upper Ob and Irtysh Region, which is considered as the homeland of Late Proto-Samoyedic. The Cherkaskul Culture is included among the Andronoid Cultures, the label reflecting a strong influence from the steppe Andronovo Complex; especially in ceramics the connection to the Fyodorovka Culture is clear (Grigoriev 2021: 24). There also appeared an expansion from the Cherkaskul Culture to the southern directions, to both sides of the Urals (Korochkova 2011: 28–29).

The Ugric branches possibly continued together in the Central Ural Region within the following Mezhovskaya Culture ca. 1500–1000 BCE. This archaeological framework presented by Parpola (2012a) matches nicely with the datings achieved from the Indo-Iranian loanword layers: the East Uralic sibilant changes were probably fully developed by 1500 BCE at the latest. The leap of Samoyedic to South Siberia separated it from the other Uralic branches for millennia to come, until much later the northward-advancing Samoyeds met the eastward-advancing Khanty between the Ob and Yenisei.

For Hungarian, Mansi, and Khanty the eastern side of the Urals is far enough: from there Hungarian moved first to the south and later to the west, while Mansi and Khanty moved to the north and northeast. The

forest-steppe zone extends to the river Iset, so the Ugric horse-related vocabulary from some unknown language could have been adopted right there in the Central Ural Region.

There are already recent models of dispersal agreeing with the rejuvenated Proto-Uralic (Parpola 2012a; 2012b; 2017; 2022; Lang 2020), and it is not possible to go through the whole dispersal process here. Suffice it to say that connecting the spread of Saami and Finnic westwards from the Upper Volga Region to the later stages of the Netted Ware tradition at the end of the 2nd millennium BCE seems possible, although the linguistic results could also agree with a somewhat later dispersal.

To conclude, the Seima-Turbino Network is only a partial match for the early Uralic expansion: its later western extension might be connected to the early dispersal of Uralic westwards from the Central Ural Region, but its earlier eastern core region cannot be related to the expansion of Late Proto-Uralic or the subsequent stages. Even during the stage U5 closing to the mid-second millennium BCE, the Uralic proto-dialects appear to have spread only within a narrow strand, reaching from the Volga–Oka Region through the Lower/Middle Kama Region to the Central Ural Region. At the same time, the Seima-Turbino Network had already reached its ultimate width from Mongolia to Finland.

Nevertheless, as people and items spread quicker than languages, it is possible that the Uralic speakers were somehow involved in the whole width of the Seima-Turbino Network, but their number and proportional density was sufficient to expand their language only within a few centers close to their core area. Parpola (2012b: 159–160) has earlier proposed that only the European side of the network was Uralic-speaking, while the Siberian side spoke Indo-Iranian. However, the new datings for the Seima-Turbino Network in Siberia (Chernykh et al. 2017; Marchenko et al. 2017) are too early for Indo-Iranian, so the language in the Siberian part of the network must have been originally something else.

Mallory (2001) has proposed that the Indo-Iranian influence and a new societal structure could be behind the Uralic expansion. There are indeed several Late Proto-Indo-Iranian loanwords which could reflect a bronze-trade context: metal tools **ora* 'awl' and **waćara* 'hammer, ax'; **ćišta* 'wax', which could be connected to bronze casting; numerals **ćęta* '100' and **ćasra* '1000', as well as **a/ęrwa* 'value', possibly connected to high-volume bronze trade; and **asora* 'lord, prince' connected to the new social organization (all these loanwords are from Holopainen 2019).

Additionally, I have recently suggested an Indo-Iranian etymology for the first component of a compound metal name U **ä(j)sVn-weć(k)V* 'tin and/or lead' found only in Permic and Mansi but still preceding the East Uralic sound changes. U **ša/okara* 'armor', found only in Khanty, could be already a Proto-Iranian loanword, yet still earlier than the East Uralic sound changes (J. Häkkinen 2023). The oldest body armor in the relevant region is a lamellar armor made of horn blades, found in the Seima-Turbino burial ground in Rostovka, in the Middle Irtysh Region (Koryakova & Epimakhov 2006: 107). Recent datings from Rostovka range mostly between 2200–2000 BCE (Marchenko et al. 2017; Chernykh et al. 2017).

However, there is nothing in these words pointing specifically towards the Seima-Turbino Network: the words could be related to other metallurgical traditions or cultures as well, for example to the Abashevo and the Sintashta Cultures. It seems probable that there were several factors behind the Uralic expansion: (1) contacts with the Indo-Iranians to the south and southwest, (2) contacts with the Seima-Turbino Network to the east, and (3) a critical location controlling the Central Ural Passage, which was of paramount importance for the trade routes running in the Eurasian forest zone.

Based on the known later development (prevailing of the Uralic languages both to the west and to the east from the Urals instead of Indo-Iranian, Paleo-West Siberian, or Paleo-North European languages), it seems that the Uralic speakers managed to capitalize on their strategic position to the maximum. The first wave of expansion was directed to the west from the Central Ural Region around the second quarter of the 2nd millennium BCE, and the second wave was advancing westwards from the Upper Volga Region (Finnic, Saami, Mordvin, Meryanic, and West Chudic; see Rahkonen 2013) from the late 2nd millennium BCE onward. In the east, only Samoyedic leaped far away to the Sayan Region probably at some point during the latter half of the 2nd millennium BCE, while the Ugric branches remained in the vicinity of the Central Ural Region for a long time. We cannot exclude the possibility of Para-Samoyedic languages existing earlier in Southwest Siberia, but the traces probably would have been wiped away by later successive expansions of Iranian, Yeniseian, and Turkic languages.

6. Conclusion

In this critical examination I have considered every relevant piece of evidence fulfilling the criteria for Late Proto-Uralic word and carrying locational evidentiary value, and I have mapped out the most accurate possible regions concerning individual pieces of evidence. The evidence leaves room for only one possible homeland for Late Proto-Uralic: the Central Ural Region.

However, the sphere of this homeland reaches towards the Middle Kama in the west and Middle Tobol in the east, partially overlapping with some recent homeland propositions (Parpola 2022: 270, 264; Saarikivi 2022: 56). The evidence dismisses homeland candidates further to the west (e.g. the Middle Volga Region and the Upper Volga Region) or to the east (e.g. the Middle Irtysh Region and the Sayan Region). Nevertheless, some of these rejected regions could be the homeland of some earlier stage preceding Late Proto-Uralic.

Another result of this scrutiny is a more resolute stratification of the Indo-Iranian loanword layers. Based on these loanword layers, a new model for the disintegration of Late Proto-Uralic is presented. This kind of flexible model is necessary to account for the non-abrupt, non-rigid disintegration process of Late Proto-Uralic. Moreover, through these loanword layers the Uralic reconstruction stages are anchored to the Indo-Iranian chronology. It is argued that even though Late Proto-Uralic was divided into two centers (Samoyedic and Finno-Ugric) already soon after ca. 2500 BCE, both centers must have remained close to each other until ca. 2000 BCE.

Only in the first quarter of the 2nd millennium BCE did the Uralic speech area disperse into a narrow strand reaching from the Upper Volga Region (> Saami, Finnic, West Chudic, Meryanic, Mordvin, and possibly Mari) through the Volga-Kama confluence (> Permic, possibly Mari and possible extinct branches between them) to the Central Ural Region (> Hungarian, Mansi, Khanty, and Samoyedic). This dispersal led to several Uralic pre-dialects, probably corresponding to the branch ancestors. During the second quarter of the 2nd millennium BCE, the first macro-branch-specific sound changes occurred, giving rise to the Uralic proto-dialects (probably four regional units: West Uralic, East Uralic, Mari, and Permic). Only ca. 1500 BCE could the individual branches have begun to advance farther from each other, and this concerns also Samoyedic, which must be located in the vicinity of the Central Ural Region until that time.

Abbreviations

AIE	Archaic Indo-European: any dialect following Late Proto-Indo-European but not yet showing branch-specific sound changes
AIr	Archaic Iranian: between Proto-Iranian and Old Iranian
AU = U2	Ancient Uralic: reorientating proto-language
BU = U3	Bicentric Uralic: bicentric proto-language
CU = U4	Common Uralic: several pre-dialects, showing independent sound substitutions
DU = U5	Diverging Uralic: several proto-dialects, showing shared sound changes
EPIE	Early Proto-Indo-European: the common ancestor to all the Indo-European languages
EPIIr	Early Proto-Indo-Iranian
EPU	Early Proto-Uralic: a recent ancestor, preceding Late Proto-Uralic by several centuries; a stage during which the earliest Archaic Indo-European loanwords were possibly borrowed
EU	East Uralic (comprising Hungarian, Mansi, Khanty, and Samoyedic)
IPA	International Phonetic Alphabet
KL	East Mansi dialect of Lower Konda
KM	East Mansi dialect of Middle Konda
KU	East Mansi dialect of Upper Konda
LL	West Mansi dialect of Lower Lozva
LM	West Mansi dialect of Middle Lozva
LPIE	Late Proto-Indo-European: the common ancestor to all the Indo-European branches after Anatolian and Tocharian split off
LPIIr	Late Proto-Indo-Iranian
LPU = U1	Late Proto-Uralic: immediately preceding the disintegration
LU	North Mansi dialect of Upper Lozva
MPIIr	Middle Proto-Indo-Iranian
NwIE	Northwest Indo-European: archaic dialect continuum of the predecessors of at least Balto-Slavic, Germanic, Celtic, and Italic
OIr	Old Iranian, during the 1st millennium BCE
Pe	West Mansi dialect of Pelymka
PIr	Proto-Iranian
PrePU	Pre-Proto-Uralic: a distant ancestor or long continuum preceding Late Proto-Uralic by several millennia
So	North Mansi dialect of Sosva
TCh	South Mansi dialect of Great Chandyri on Tavda
TY	South Mansi dialect of Yanichkova on Tavda
UPA	Uralic Phonetic Alphabet
VN	West Mansi dialect of Sotnikova on North Vagilsk
VNZ	West Mansi dialect of Zaozernaya on North Vagilsk
VS	West Mansi dialect of South Vagilsk
WU	West Uralic (comprising Saami, Finnic, Mordvin, and probably the extinct West Chudic and Meryanic)

References

AgroAtlas = Afonin, A. N. & Greene, S. L. & Dzyubenko, N. I. & Frolov, A. N. (eds.). 2008. *Interactive agricultural ecological atlas of Russia and neighboring countries: Economic plants and their diseases, pests and weeds.* http://www.agroatlas.ru

Aikio, Ánte. 2006. On Germanic-Saami contacts and Saami prehistory. *Journal de la Société Finno-Ougrienne* 91. 9–55. https://doi.org/10.33340/susa.81944

Aikio, Ánte (Luobbal Sámmol Sámmol Ánte). 2014. The Uralic-Yukaghir lexical correspondences: Genetic inheritance, language contact or chance resemblance? *Finnisch-Ugrische Forschungen* 62. 7–76. https://doi.org/10.33339/fuf.86078

Aikio, Ánte (Luobbal Sámmol Sámmol Ánte). 2015. The Finnic 'secondary e-stems' and Proto-Uralic vocalism. *Journal de la Société Finno-Ougrienne* 95. https://doi.org/10.33340/susa.82642

Aikio, Ante (Luobbal Sámmol Sámmol Ánte). 2020. Uralic etymological dictionary: Draft version of entries A–Ć published online on Academia.edu.

Aikio, Ánte (Luobbal Sámmol Sámmol Ánte). 2022. Proto-Uralic. In Bakró-Nagy, Marianne & Laakso, Johanna & Skribnik, Elena (eds.), *The Oxford guide to the Uralic languages*, 3–27. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198767664.003.0001

Aikio, Ante & Aikio, Aslak. 2001. Heimovaelluksista jatkuvuuteen: Suomalaisen väestöhistorian tutkimuksen pirstoutuminen. *Muinaistutkija* 4/2001. 2–21. http://www.sarks.fi/mt/pdf/2001_4.pdf

Anthony, David W. 2007. *The horse, the wheel, and language: How Bronze-Age riders from the Eurasian steppes shaped the modern world*. Princeton: Princeton University Press.

Anthony, David W. & Ringe, Don. 2015. The Indo-European homeland from linguistic and archaeological perspectives. *Annual Review of Linguistics* 2015: 1. 199–219. https://doi.org/10.1146/annurev-linguist-030514-124812

Beekes, Robert S. P. 2011. *Comparative Indo-European linguistics: An introduction.* Revised and corrected by Michiel de Vaan. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Beke, Ödön. 1997–2001. *Mari nyelvjárási szótár* 1–9. Unter Mitarbeit von Zsófia M. Velenyák und József Erdődi, neu redigiert von Gábor Bereczki, bearbeitet von Margarita Kuznecova, herausgegeben von János Pusztay (Bibliotheca Ceremissica 4). Szombathely: Berzsenyi Dániel Tanárképző Főiskola.

Blyakharchuk, Tatiana A. & Chernova, Natalia A. 2013. Vegetation and climate in the Western Sayan Mts according to pollen data from Lugovoe Mire as a background for prehistoric cultural change in southern Middle Siberia. *Quaternary Science Reviews* 75. 22–42. http://dx.doi.org/10.1016/j.quascirev.2013.05.017

Byrd, Andrew Miles. 2018. The phonology of Proto-Indo-European. In Klein, Jared & Joseph, Brian & Fritz, Matthias (eds.), *Handbook of comparative and historical Indo-European linguistics*, Volume 3, 2056–2079. Berlin & Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110542431-042

Cantera, Alberto. 2017. The phonology of Iranian. In Klein, Jared & Joseph, Brian & Fritz, Matthias (eds.), *Handbook of comparative and historical Indo-European linguistics*, Volume 1, 481–503. Berlin &Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110261288-033

Cardona, George. 2017. The documentation of Indic. In Klein, Jared & Joseph, Brian & Fritz, Matthias (eds.), *Handbook of comparative and historical Indo-European Linguistics*, Volume 1. 309–325. Berlin &Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110261288-025

Carpelan, Christian & Parpola, Asko. 2001. Emergence, contacts and dispersal of Proto-Indo-European, Proto-Uralic and Proto-Aryan in archaeological perspective. In Carpelan, Christian & Parpola, Asko & Koskikallio, Petteri (eds.), *Early contacts between Uralic and Indo-European: Linguistic and archaeological considerations* (Mémoires de la Société Finno-Ougrienne 242), 55–150. Helsinki: Société Finno-Ougrienne.

Chernykh, E. N. & Korochkova, O. N. & Orlovskaya, L. B. 2017. Issues in the calendar chronology of the Seima-Turbino Transcultural Phenomenon. *Archaeology, Ethnology & Anthropology of Eurasia*. 45/2. 45–55.

Derksen, Rick. 2015. *Etymological dictionary of the Baltic inherited lexicon* (Leiden Indo-European Etymological Series 13). Leiden & Boston: Brill.

Grigoriev, Stanislav. 2021. Andronovo problem: Studies of cultural genesis in the Eurasian Bronze Age. *Open Archaeology* 2021: 7. 3–36. https://doi.org/10.1515/opar-2020-0123

Grünthal, Riho & Heyd, Volker & Holopainen, Sampsa & Janhunen, Juha A. & Khanina, Olesya & Miestamo, Matti & Nichols, Johanna & Saarikivi, Janne & Sinnemäki, Kaius. 2022. Drastic demographic events triggered the Uralic spread. *Diachronica*, Volume 39, Issue 4. 490–524. https://doi.org/10.1075/dia.20038.gru

Häkkinen, Jaakko. 2007. *Kantauralin murteutuminen vokaalivastaavuuksien valossa*. University of Helsinki. (Master's thesis.) https://helda.helsinki.fi/handle/10138/19684

Häkkinen, Jaakko. 2009. Kantauralin ajoitus ja paikannus: perustelut puntarissa. *Journal de la Société Finno-Ougrienne* 92. 9–56. https://doi.org/10.33340/susa.82020

Häkkinen, Jaakko. 2010. Jatkuvuusperustelut ja saamelaisen kielen leviäminen. *Muinaistutkija* 1/2010. 19–36. http://www.sarks.fi/mt/pdf/2010_1.pdf

Häkkinen, Jaakko. 2012. After the protolanguage: Invisible convergence, false divergence and boundary shift. *Finnisch-Ugrische Forschungen* 61. 7–28. https://doi.org/10.33339/fuf.85674

Häkkinen, Jaakko. 2022. Recurring irregularities in West Uralic 1: Para-Slavic loanwords. In Kaheinen, Kaisla & Leisiö, Larisa & Erkkilä, Riku & Qiu, Toivo E. H. (eds), *Hämeenmaalta Jamalille: Kirja Tapani Salmiselle 07.04.2022*, 127–144. Helsinki: HELDA Open Books. https://doi.org/10.31885/9789515180858.11

Häkkinen, Jaakko. 2023. An old Uralic compound metal name. In Holopainen, Sampsa & Kim, Jeongdo & Metsäranta, Niklas (eds.), *Elämä ja etymologia: Janne Saarikiven 50-vuotisjuhlakirja*. 102–121. Helsinki: HELDA Open Books. https://doi.org/10.31885/9789515182838

Häkkinen, Kaisa. 1984. Wäre es schon an der Zeit, den Stammbaum zu fällen? *Ural-Altaische Jahrbücher, Neue Folge*, Band 4. 1–24.

Häkkinen, Kaisa. 1996. *Suomalaisten esihistoria kielitieteen valossa* (Tietolipas 147). Helsinki: Suomalaisen Kirjallisuuden Seura.

Hegedűs, Irén. 2012. The RUKI-rule in Nuristani. In Whitehead, Benedicte Nielsen & Olander, Thomas & Olsen, Birgit Anette & Rasmussen, Jens Elmegård (eds.), *The sound of Indo-European: Phonetics, phonemics, and morphophonemics* (Copenhagen Studies in Indo-European, Vol. 4). 145–167. Copenhagen: Museum Tusculanum Press.

Heggarty, Paul. 2007. Linguistics for archaeologists: Principles, methods and the case of the Incas. *Cambridge Archaeological Journal* 17:3. 311–340.

Holopainen, Sampsa. 2018. Indo-Iranian loans confined to Saami? In Holopainen, Sampsa & Saarikivi, Janne (eds.), *Perì orthótētos etýmōn* (*Περὶ ὀρθότητος ἐτύμων*): *Uusiutuva uralilainen etymologia* (Uralica Helsingiensia 11), 135–179. Helsinki: Suomalais-Ugrilainen Seura.

Holopainen, Sampsa. 2019. *Indo-Iranian borrowings in Uralic: Critical overview of the sound substitutions and distribution criterion*. University of Helsinki. http://urn.fi/URN:ISBN:978-951-51-5729-4 (Doctoral dissertation.)

Holopainen, Sampsa. 2021. On the question of substitution of palatovelars in Indo-European loanwords into Uralic. *Journal de la Société Finno-Ougrienne* 98. 197–233.

Hyllested, Adam. 2014. *Word exchange at the gates of Europe: Five millennia of language contact*. University of Copenhagen. (Doctoral dissertation.)

Iivonen, Antti. 2012. Kielten vokaalit kuuloanalogisessa vokaalikartassa. *Puhe ja kieli* 32:1. 17–43. https://journal.fi/pk/article/view/6332

Itkonen, Terho. 1997. Reflections on Pre-Uralic and the "Saami-Finnic protolanguage". *Finnisch-Ugrische Forschungen* 54$_{1,3}$. 229–266.

Janhunen, Juha. 2009. Proto-Uralic – what, where, and when? In Ylikoski, Jussi (ed.), *The quasquicentennial of the Finno-Ugrian Society* (Mémoires de la Société Finno-Ougrienne 258). 57–78. Helsinki: Société Finno-Ougrienne. https://www.sgr.fi/sust/sust258/sust258_janhunen.pdf

Jasanoff, Jay H. 2017. The impact of Hittite and Tocharian: Rethinking Indo-European in the 20th century and beyond. In Klein, Jared & Joseph, Brian & Fritz, Matthias (eds.), *Handbook of comparative and historical Indo-European linguistics*, Volume 1, 220–238. Berlin &Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110261288-018

Kallio, Petri. 2006. Suomen kantakielten absoluuttista kronologiaa. *Virittäjä* 1/2006. 2–25. http://www.kotikielenseura.fi/virittaja/hakemistot/jutut/2006_2.pdf

Kallio, Petri. 2008. On the 'Early Baltic' loanwords in Common Finnic. In Lubotsky, Alexander & Schaeken, Jos & Wiedenhof, Jeroen (eds.), *Evidence and counter-evidence: Essays in honour of Frederik Kortlandt 1, Balto-Slavic and Indo-European linguistics* (Studies in Slavic and General Linguistics 32), 265–277. Leiden: Brill. https://doi.org/10.1163/9789401206358_020

Kallio, Petri. 2009. Stratigraphy of Indo-European loanwords in Saami. In Äikäs, Tiina (ed.), *Máttut – máddagat: The roots of Saami ethnicities, societies and spaces/places* (Publications of the Giellagas Institute 12). 30–45. Oulu: Giellagas Institute, University of Oulu.

Kallio, Petri. 2012. The non-initial-syllable vowel reductions from Proto-Uralic to Proto-Finnic. In Hyytiäinen, Tiina & Jalava, Lotta & Saarikivi, Janne & Sandman, Erika (eds.), *Per Urales ad Orientem: Iter polyphonicum multilingue, Festskrift tillägnad Juha Janhunen på hans sextioårsdag den 12 februari 2012* (Mémoires de la Société Finno-Ougrienne 264), 163–175. Helsinki: Société Finno-Ougrienne. https://www.sgr.fi/sust/sust264/sust264_kallio.pdf

Kallio, Petri. 2015a. The stratigraphy of the Germanic loanwords in Finnic. In Askedal, John Ole & Nielsen, Hans Frede (eds.), *Early Germanic languages in contact*, 23–38. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/nss.27.02kal

Kallio, Petri. 2015b. The language contact situation in prehistoric Northeastern Europe. In Mailhammer, Robert & Vennemann, Theo & Olsen, Birgit Anette (eds.), *The linguistic roots of Europe: Origin and development of European languages* (Copenhagen Studies in Indo-European, Vol. 6). 77–102. Copenhagen: Museum Tusculanum Press.

Kallio, Petri. Forthcoming. *The Proto-Balto-Slavic loanword stratum in the Uralic languages*. Manuscript.

Kannisto, Artturi. 2013. *Wogulisches Wörterbuch*. Bearbeitet von Vuokko Eiras, herausgegen von Arto Moisio (Lexica Societatis Fenno-Ugricae XXXV). Helsinki: Société Finno-Ougrienne.

Kim, Ronald I. 2018. The phonology of Balto-Slavic. In Klein, Jared & Joseph, Brian & Fritz, Matthias (eds.), *Handbook of comparative and historical Indo-European linguistics*, Volume 3. 1974–1985. Berlin &Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110542431-036

Koivulehto, Jorma. 1991. *Uralische Evidenz für die Laryngaltheorie* (Österreichische Akademie der Wissenschaften, Philosophisch-Historische Klasse, Sitzungsberichte, 566. Band). Wien: Verlag der österreichischen Akademie der Wissenschaften.

Koivulehto, Jorma. 2001. The Earliest contacts between Indo-European and Uralic speakers in the light of lexical loans. In Carpelan, Christian & Parpola, Asko & Koskikallio, Petteri (eds.), *Early contacts between Uralic and Indo-European: Linguistic and archaeological considerations* (Mémoires de la Société Finno-Ougrienne 242), 235–264. Helsinki: Société Finno-Ougrienne.

Koivulehto, Jorma. 2002. Contact with non-Germanic languages II: Relations to the east. In Bandle, Oscar & Braunmüller, Kurt & Jahr, Ernst Hakon & Karker, Allan & Naumann, Hans-Peter & Telemann, Ulf & Elmevik, Lennart & Widmark, Gun (eds.), *The Nordic languages: An international handbook of the history of the North Germanic languages*, 583–594. Berlin & New York: Walter de Gruyter.

Korochkova, Olga Nikolaevna. 2011. *Взаимодействие культур в эпоху бронзы в Среднем Зауралье и подтаежном Тоболо-Иртышье: Факторы, механизмы, динамика* [Interaction of cultures during the Bronze Age in the Central Trans-Urals and the southern taiga Tobol-Irtysh Region: Factors, mechanisms, dynamics]. Russian Academy of Sciences. (Summary of doctoral dissertation.)

Korochkova, Olga Nikolaevna. 2012. Первый металл в культуре таежных аборигенов Западной Сибири и Среднего Урала [The first metal in the culture of the taiga natives of Western Siberia and the Central Urals]. *Вестник НГУ, Серия: История, филология*, Т. 11, Вып. 3. 137–147.

Korochkova, O. N. (Корочкова, О. Н.) 2019. Среднее Зауралье и таежное Приобье в системе связей Западноазиатской металлургической провинции [Middle Trans-Urals and Taiga Ob' in the System of Relations of the West Asian Metallurgical province]. *Вестник Санкт-Петербургского университета: История*, Т. 64, Вып. 2 [*Vestnik of Saint Petersburg University: History*, Vol. 64, iss. 2]. 718–739. https://doi.org/10.21638/11701/spbu02.2019.216

Korochkova, Olga Nikolaevna & Kuzminykh, Sergei Vladimirovich & Serikov, Yuri Borisovich & Stefanov, Vladimir Ivanovich. 2010. Metals from the ritual site of Shaitanskoye Ozero II (Sverdlovsk Oblast, Russia). *Trabajos de prehistoria* 67, N.º 2. 489–499. http://dx.doi.org/10.3989/tp.2010.10052

Koryakova, Ludmila & Epimakhov, Andrej. 2006. *The Urals and Western Siberia in the Bronze and Iron Ages* (Cambridge World Archaeology). Cambridge University Press.

Kuokkala, Juha. 2018. Finnic-Saamic labial vowels in non-initial syllables: An etymological evaluation. In Holopainen, Sampsa & Saarikivi, Janne (eds.), *Perì orthótētos etýmōn* (*Περὶ ὀρθότητος ἐτύμων*): *Uusiutuva uralilainen etymologia* (Uralica Helsingiensia 11), 11–74. Helsinki: Société Finno-Ougrienne.

Kuz'mina, Elena. 2007. *The origin of the Indo-Iranians*. Mallory, J. P. (ed.) (Leiden Indo-European Etymological Dictionaries series 3). Leiden: Brill.

Kuz'mina, O. V. 2021. The Abashevo Culture. *Вестник Санкт-Петербургского университета: История*, Т. 66, Вып. 4 [*Vestnik of Saint Petersburg University: History*, Vol. 66, iss. 4]. 1206–1229. https://doi.org/10.21638/11701/spbu02.2021.411

Lang, Valter. 2020. *Homo Fennicus: Itämerensuomalaisten etnohistoria*. Helsinki: Suomalaisen Kirjallisuuden Seura.

Lapteva, Elena G. & Korona, Olga M. 2012. Holocene vegetation changes and anthropogenic influence in the forest-steppe zone of the Southern Trans-Urals based on pollen and plant macrofossil records from the Sukharysh cave. *Vegetation History and Archaeobotany* (2012) 21. 321–336. https://doi.org/10.1007/s00334-011-0333-z

Lapteva, E. G. & Zaretskaya, N. E. & Kosintsev, P. A. & Lychagina, E. L. & Chernov, A. V. 2017. First data on the Middle to Late Holocene dynamics of vegetation in the Upper Kama Region. *Russian Journal of Ecology* 2017, Vol. 48, No. 4. 326–334.

Lubotsky, Alexander. 1990. La loi de Brugmann et *$*H_3e$-. *La reconstruction des laryngales* (Bibliothèque de la Faculté de Philosophie et Lettres de l'Université de Liège, fascicule CCLIII). 129–136. Liège-Paris. https://hdl.handle.net/1887/2666

Lubotsky, Alexander. 2018. The phonology of Proto-Indo-Iranian. In Klein, Jared & Joseph, Brian & Fritz, Matthias (eds.), *Handbook of comparative and historical Indo-European linguistics*, Volume 3, 1875–1888. Berlin & Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110542431-031

Mallory, J. P. 1989. *In search of the Indo-Europeans: Language, archaeology and myth*. London: Thames and Hudson.

Mallory, J. P. 2001. Uralics and Indo-Europeans: Problems of time and space. In Carpelan, Christian & Parpola, Asko & Koskikallio, Petteri (eds.), *Early contacts between Uralic and Indo-European: Linguistic and archaeological considerations* (Mémoires de la Société Finno-Ougrienne 242). 345–366. Helsinki: Société Finno-Ougrienne.

Mallory, J. P. & Adams, D. Q. (eds.). 1997. *Encyclopedia of Indo-European culture*. London & Chicago: Fitzroy Dearborn Publishers.

Mallory, J. P. & Adams, D. Q. 2006. *The Oxford introduction to Proto-Indo-European and the Proto-Indo-European world*. Oxford & New York: Oxford University Press.

Marchenko, Z. V. & Svyatko S. V. & Molodin V. I. & Grishin A. E. & Rykun M. P. 2017. Radiocarbon chronology of complexes with Seima-Turbino type objects (Bronze Age) in Southwestern Siberia. *Radiocarbon* 2017. 1–17. https://doi.org/10.1017/RDC.2017.24

Matasović, Ranko. 2005. Toward a relative chronology of the earliest Baltic and Slavic sound changes. *Baltistica* XL(2). 147–157.

Matveev, A. V. 2011. *Материалы по мансийской топонимии горной части Северного Урала* [Materials on Mansi toponymy in the mountainous part of the Northern Urals]. Министерство образования и науки Российской Федерации Государственное образовательное учреждение высшего профессионального образования.

Metsäranta, Niklas. 2020. *Periytyminen ja lainautuminen: Marin ja permiläisten kielten sanastontutkimusta*. University of Helsinki. (Doctoral dissertation.) http://urn.fi/URN:ISBN:978-951-51-6870-2

Moisio, Arto & Saarinen, Sirkka. 2008. *Tscheremissisches Wörterbuch*. Aufgezeichnet von Volmari Porkka, Arvid Genetz, Yrjö Wichmann, Martti Räsänen, T. E. Uotila und Erkki Itkonen (Lexica Societatis Fenno-Ugricae XXXII). Helsinki: Suomalais-Ugrilainen Seura.

Morgunova, N. L. & Khokhlova, O. S. 2013. Chronology and periodization of the Pit-Grave Culture in the region between the Volga and Ural rivers based on radiocarbon dating and paleopedological research. In Jull, A. J. T. & Hatté, C. (eds.), *Proceedings of the 21st International Radiocarbon Conference*, Part 2 (*Radiocarbon* Vol 55, Nr 3), 1286–1296. https://doi.org/10.1017/S0033822200048190

Nichols, Johanna. 2021. The origin and dispersal of Uralic: Distributional typological view. *Annual Review of Linguistics* 2021: 7. 351–369.

Nichols, Johanna & Rhodes, Richard A. 2018. Vectors of language spread at the central steppe periphery: Finno-Ugric as catalyst language. In Kroonen, Guus & Iversen, Rune (eds.), *Digging for words*, 58–68. Oxford: BAR Publishing.

Ollett, Andrew. 2014. Evidence of laryngeal coloring in Proto-Indo-Iranian. *Historische Sprachforschung = Historical Linguistics* 127. 150–165.

Parpola, Asko. 2012a. The problem of Samoyed origins in the light of archaeology: On the formation and dispersal of East Uralic (Proto-Ugro-Samoyed). In Hyytiäinen, Tiina & Jalava, Lotta & Saarikivi, Janne & Sandman, Erika (eds.), *Per Urales ad Orientem: Iter polyphonicum multilingue, Festskrift tillägnad Juha Janhunen på hans sextioårsdag den 12 februari 2012* (Mémoires de la Société Finno-Ougrienne 264), 227–250. Helsinki: Société Finno-Ougrienne. https://www.sgr.fi/sust/sust264/sust264_parpola.pdf

Parpola, Asko. 2012b. Formation of the Indo-European and Uralic (Finno-Ugric) language families in the light of archaeology: Revised and integrated "total" correlations. In Grünthal, Riho & Kallio, Petri (eds.), *A linguistic map of prehistoric northern Europe* (Mémoires de la Société Finno-Ougrienne 266), 119–184. Helsinki: Société Finno-Ougrienne. https://www.sgr.fi/sust/sust266/sust266_parpola.pdf

Parpola, Asko. 2017. Finnish *vatsa* ~ Sanskrit *vatsá* and the formation of Indo-Iranian and Uralic languages. *Journal de la Société Finno-Ougrienne* 96. 245–286.

Parpola, Asko. 2022. Location of the Uralic proto-language in the Kama River valley and the Uralic speakers' expansion east and west with the 'Sejma-Turbino Transcultural Phenomenon' 2200–1900 BC. *Археология евразийских степей* 2022(2). https://doi.org/10.24852/2587-6112.2022.2.258.277

Pokorny, Julius. 2007. *Proto-Indo-European etymological dictionary: A revised edition of Julius Pokorny's Indogermanisches Etymologisches Wörterbuch*. Dnghu Association.

Rahkonen, Pauli. 2013. *Southeastern contact area of Finnic languages in the light of onomastics*. University of Helsinki. https://helda.helsinki.fi/handle/10138/38908 (Doctoral dissertation.)

Ringe, Don. 2017. Indo-European dialectology. In Klein, Jared & Joseph, Brian & Fritz, Matthias (eds.), *Handbook of comparative and historical Indo-European linguistics*., Volume 1, 62–75. Berlin & Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110261288-006

Saag, Lehti & Vasilyev, Sergey V. & Varul, Liivi & Kosorukova, Natalia V. & Gerasimov, Dmitri V. & Oshibkina, Svetlana V. & Griffith, Samuel J. & Solnik, Anu & Saag, Lauri & D'Atanasio, Eugenia & Metspalu, Ene & Reidla, Maere & Rootsi, Siiri & Kivisild, Toomas & Scheib, Christiana Lyn & Tambets, Kristiina & Kriiska, Aivar & Metspalu, Mait. 2021. Genetic ancestry changes in Stone to Bronze Age transition in the East European plain. *Science Advances* Vol 7, Issue 4. https://doi.org/10.1126/sciadv.abd6535

Saarikivi, Janne. 2022. The divergence of Proto-Uralic and its offspring: A descendent reconstruction. In Bakró-Nagy, Marianne & Laakso, Johanna & Skribnik, Elena (eds.), *The Oxford guide to the Uralic languages*, 28–58. Oxford University Press.

Saarikivi, Janne & Lavento, Mika. 2012. Linguistics and archaeology: A critical view of an interdisciplinary approach with reference to the prehistory of Northern Scandinavia. In Damm, Charlotte & Saarikivi, Janne (eds.). *Networks, interaction and emerging identities in Fennoscandia and beyond: Papers from the conference held in Tromsø, Norway, October 13–16 2009* (Mémoires de la Société Finno-Ougrienne 265), 177–216. Helsinki: Société Finno-Ougrienne. https://www.sgr.fi/sust/sust265/sust265_saarikivi_and_lavento.pdf

Salminen, Tapani. 1989. Classification of the Uralic languages. In Grünthal, Riho & Penttinen, Sirpa & Salminen, Tapani (eds.), *IFUSCO 1988: Proceedings of the Fifth International Finno-Ugrist Students' Conference, Helsinki, 22–26 May 1988* (Castrenianumin toimitteita 35), 15–24. Helsinki: Suomalais-Ugrilainen Seura & Helsingin yliopisto.

Salminen, Tapani. 2002. Problems in the taxonomy of the Uralic languages in the light of modern comparative studies. *Лингвистический беспредел: Сборник статей к 70-летию А. И. Кузнецовой*. 44–55. Москва: Издательство Московского университета.

Sammallahti, Pekka. 1988. Historical phonology of the Uralic Languages. Sinor, Denis (ed.), *The Uralic languages*, 478–554. Leiden: Brill.

Shumilovskikh, Lyudmila S. & Schmidt, Monika & Pereskokov, Mikhail & Sannikov, Pavel. 2020. Postglacial history of East European boreal forests in the mid-Kama region, pre-Urals, Russia. *Boreas* 49(3). 526–543. https://doi.org/10.1111/bor.12436

Simon, Zsolt. 2020. Urindogermanische Lehnwörter in den uralischen und finno-ugrischen Grundsprachen: Eine Fata Morgana? *Indogermanische Forschungen* 125, issue 1. 239–265.

Skjærvø, Prods Oktor. 2017. The documentation of Iranian. In Klein, Jared & Joseph, Brian & Fritz, Matthias (eds.), *Handbook of comparative and historical Indo-European linguistics*, Volume 1, 471–481. Berlin & Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110261288-032

SSA = Itkonen, Erkki & Kulonen, Ulla-Maija (eds.). 1992–2000. *Suomen sanojen alkuperä: Etymologinen sanakirja* 1–3 (Suomalaisen Kirjallisuuden Seuran Toimituksia 556 & Kotimaisten kielten tutkimuskeskuksen julkaisuja 62). Helsinki: Suomalaisen Kirjallisuuden Seura & Kotimaisten kielten tutkimuskeskus.

Tirkkonen, Jani-Matti. 2012. *"Lopullinen totuus pohjoiseurooppalaisten alkuperästä"? Kalevi Wiik vallankumouksellisen paradigman edustajana juuretkiistassa*. University of Eastern Finland. (Master's thesis.)

UEW = Rédei, Károly. 1986–1991. *Uralisches Etymologisches Wörterbuch*. Band I–II: Unter Mitarbeit von Marianne Bakró-Nagy, Sándor Csúcs, István Erdélyi †, László Honti, Éva Korenchy †, Éva K. Sal und Edit Vértes, Band III: Register Zusammengestellt von Attila Dobó und Éva Fancsaly. Wiesbaden: Harrassowitz.

Underhill, Peter A. & Poznik, G. David & Rootsi, Siiri & Järve, Mari & Lin, Alice A. & Wang, Jianbin & Passarelli, Ben & Kanbar, Jad & Myres, Natalie M. & King, Roy J. & Di Cristofaro, Julie & Sahakyan, Hovhannes & Behar, Doron M. & Kushniarevich, Alena & Šarac, Jelena & Šaric, Tena & Rudan, Pavao & Pathak, Ajai Kumar & Chaubey, Gyaneshwer & Grugni, Viola & Semino, Ornella & Yepiskoposyan, Levon & Bahmanimehr, Ardeshir & Farjadian, Shirin & Balanovsky, Oleg & Khusnutdinova, Elza K. & Herrera, Rene J. & Chiaroni, Jacques & Bustamante, Carlos D. & Quake, Stephen R. & Kivisild, Toomas & Villems, Richard. 2014. The phylogenetic and geographic structure of Y-chromosome haplogroup R1a. *European Journal of Human Genetics* (2014). 1–8. https://doi.org/10.1038/ejhg.2014.50

Vilppula, Matti. 1984. Lisiä lehmuksen, jalavan ja lehmuksen kuoriaineksen nimityksiin. *Virittäjä* 88(2). 193–200.

Volkova, V. S. & Kuz'mina, O. B. & Gnibidenko, Z. N. & Golovina, A. G. 2016. The Paleogene/Neogene boundary in continental deposits of the West Siberian Plain. *Russian Geology and Geophysics* 57. 303–315.

Vybornov, Aleksandr A. & Lychagina, Eugeniya L. & Kulkova, Marianna A. & Melnichuk, Andrey F. 2021. Periodization and Chronology of the Eneolithic in the Upper and Middle Kama Regions. *Materiality and objects: Multi-disciplinary approaches to archaeological material and contexts, Proceedings of the Finnish-Russian Archaeological Symposium, Hämeenlinna, 10-12 October 2019* (*Iskos* 24). 7–20.

Witzel, Michael. 2001. Autochthonous Aryans? The evidence from Old Indian and Iranian texts. *Electric Journal of Vedic Studies* 7:3. 1–93.

Zhivlov, Mikhail. 2015. *Неиндоевропейский субстрат в финно-волжских языках*. (Presentation manuscript.)