

Annual Review of Linguistics

The Origin and Dispersal of Uralic: Distributional Typological View

Johanna Nichols^{1,2,3}

¹Department of Slavic Languages and Literatures, University of California, Berkeley, California 94720-2979, USA; email: johanna@berkeley.edu

²Linguistic Convergence Laboratory, National Research University Higher School of Economics, Moscow 101000, Russia

³Department of General Linguistics, University of Helsinki, FI-00014 Helsinki, Finland

Annu. Rev. Linguist. 2021. 7:351–69

The *Annual Review of Linguistics* is online at
linguistics.annualreviews.org

<https://doi.org/10.1146/annurev-linguistics-011619-030405>

Copyright © 2021 by Annual Reviews.
All rights reserved

Keywords

Uralic, Indo-European, 4.2 ka event, Seima-Turbino, protolanguage, Proto-Uralic homeland

Abstract

Recent progress in comparative linguistics, distributional typology, and linguistic geography allows a unified model of Uralic prehistory to take shape. Proto-Uralic first introduced an eastern grammatical profile to central and western Eurasia, where it has remained quite stable. Proto-Indo-European and Proto-Uralic had no connection, either genealogical or areal, until the spreading Indo-Iranian branch of Indo-European came into contact with the already-diverged branches of Uralic about 4,000 years ago. A severe and widespread drought beginning about 4,200 years ago cleared the way for a rapid spread of Uralic-speaking people along the Volga and across south-western Siberia. It also contributed to the sudden rise of the Seima-Turbino bronze-trading complex, one component of the Uralic spread mechanism. After the initial spread, the Uralic daughter languages retained their Volga homelands remarkably stably while also extending far to the north in a recurrent Eurasian pattern.

ANNUAL REVIEWS CONNECT

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

1. INTRODUCTION

This review outlines a distributional-typological and linguistic-geographic model of the origin and spread of the Uralic language family within its larger northern Eurasian linguistic picture. It can supplement but not replace an expert Uralicist assessment; those by Aikio (2020) (chiefly on reconstruction) and Saarikivi (2020) (chiefly on early branch histories) are excellent, but questions are still open on early Uralic sociolinguistics, demography, language typology, ultimate origins, and wider Eurasian connections. Family-wide coverage of current typological issues is rare (for welcome exceptions, see Miestamo et al. 2015, Klumpp et al. 2018, Pajusalu et al. 2018, Janhunen 2020), and comprehensive typologically informed grammars are few (but increasing; see Nikolaeva 2014, Feist 2015, Wilbur 2015, Toldova et al. 2018). Uralic linguists need to be able to present linguistically well-informed hypotheses to geneticists, archaeologists, and ethnohistorians. Another critical need is for Uralic evidence to better inform current discussions of where the Uralic homeland was located in relation to the Indo-European (IE) homeland; this issue is central to the model sketched out here, which is based on that of Grünthal et al. (manuscript in preparation).

The Uralic family is often estimated to be about 4,500 years old (Kallio 2006; 2015, pp. 80–81) but sometimes 6,000 years or more. A firm terminus ante quem comes from a contact episode that brought early Indo-Iranian loanwords into the individual Uralic branches early enough in their histories that the loans reconstruct to the branch protolanguages, while the inventory and phonology of loans in each branch show that at that time, the branches had formed and had just begun their separate phonological evolutions (Holopainen 2019). That episode is dated at about 4,000 years ago on the basis of archaeological evidence and the attestation dates of earliest Indic and Iranian. The estimate of 4,500 years for Proto-Uralic (henceforth PU) is based on the estimate (commonly voiced by linguists, though never systematically tested) that 500 years is enough for sister dialects to evolve into separate, mutually unintelligible languages. As the inflectional morphemes even of the geographically distant Proto-Finnic and Proto-Samoyedic branches are well preserved, the branch protolanguages must have been more dialect-like than language-like 4,000 years ago.

Setting a terminus post quem depends on whether the Samoyedic branch, which retains a low number of PU words, is deemed to have gone through an unusual sociolinguistic event that caused massive lexical loss and borrowing from an unknown language, or to have separated from the others so early that normal rates of lexical loss have winnowed the vocabulary down.

The family is divided into nine uncontroversial elementary branches, which are shown in **Table 1**. What, if any, higher-level groupings exist is debated (see Section 3.2). Mari and Hungarian are one-language branches, Permic and Mordvin are minimally subdivided into two very close daughters each, and the others are internally of about Romance-like or Germanic-like age—that is, about 2,000–2,500 years of internal divergence into daughters numbering from 3 to 10. Finnish, Estonian, and Hungarian are national languages; the others are minority languages, and most are endangered. A current survey of the languages and family can be found in the handbook by Bákro-Nagy et al. (2020); older but still useful is Abondolo's (1998) *Uralic Languages*. A map of the modern ranges is available at <https://bedlan.net/uralic/> (for a map of historical ranges, see Grünthal & Salminen 1993).

Typologically, all the Uralic languages belong, to a greater or lesser extent, to the Inner Asian structural type characterized by head-final ordering at all levels; agglutinative morphology, which comprises monoexponential derivational and inflectional morphemes and transparent morpheme boundaries; fairly simple syllable structure; and absence of tones or contrastive stress placement (see Section 4.3).

Table 1 Uralic daughter branches and languages (surviving or recently extinct only), listed east to west and (within branches) south to north

Branch	Languages	Possible groups ^b	Binary division ^b
Samoyedic	Kamas, Selkup, Nenets, ^a Enets, ^a Nganasan		Samoyedic
Hungarian	Hungarian	} Ugric	} Finno-Ugric
Mansi	South, East, West, North Mansi		
Khanty	South, East, North Khanty		
Permic	Udmurt, Komi ^a		
Mari	Meadow Mari, Hill Mari		
Mordvin	Moksha, Erzya	} Western Uralic	
Finnic	South Estonian, Livonian, Estonian, Ingrian, Karelian, Votic, Veps, Finnish		
Saami	South, Ume, Pite, Lule, North, Inari, Skolt, Akkala, Kildin, Ter		

^aThese languages have dialects that are often classified as separate languages.

^bFor these possible higher-level groupings, see Section 3.2.

2. CHRONOLOGICAL FRAMEWORK

The time frame relevant to Uralic origins extends from about 6,000 years ago (henceforth, 6000 BP¹), when inner Eurasia north of the steppe was pre-Neolithic or barely Neolithic and PU had not yet split, to early historical times, when states and colonial interests began to determine language history. By about 2,000 years ago, the major branches of the Uralic language family were formed and in place, their own internal subdivisions were mostly formed, and branch-specific spreads, mostly northward, were underway.

For pre-Neolithic times, we know little about the linguistic population of northern inner Eurasia. Relevant genetic factors include a counterclockwise pattern of gene spread from Southeast Asia, which brought genes and possibly also languages across the north from east to west beginning in Paleolithic times (Rootsi et al. 2007). Other genetic work also points to westward movements of Siberian genes across the north affecting even the westernmost Uralic languages (e.g., Ilumäe et al. 2016, Tambets et al. 2018, Balanovsky 2019, Saag et al. 2019). The major north Eurasian language spreads also have gone from east to west, at least from Neolithic times on.

In Neolithic to early Bronze Age times, a standing pattern of east-to-west and south-to-north language spreads began to take shape in northern Asia, driven by the dynamic between steppe nomads and China (Barfield 1989): Protection money from China enabled warlords to attract followers and field armies, and this dynamic eventually made westward spreads possible. Northward expansions, however, occurred all across northern Eurasia, including the spread of Uralic. They were probably driven by the higher rates of population growth and technological advances to the south of the forest zone and, later, by the spread of large-scale reindeer herding.

The Eneolithic (or Copper Age) and the Bronze Age saw spreads of people and languages across the western steppe as cattle herding spread to the steppe periphery from eastern Europe and ancient metallurgical centers created trade networks and drove expansions. The early spread of IE (Section 3.1) emanated in part from areas near the Carpathian and Caucasian metallurgical centers. When this system, and with it Old European civilization, collapsed, ore deposits in the

¹Throughout this review, prehistoric times are expressed in approximate calendar years before present (BP).

southern Urals and the Altai area became known and fueled economic growth in Southwest Asia as well as the spread of the Yamnaya cultural system (late Proto-Indo-European-speaking) from the lower Volga area across the steppe (Anthony 2007, pp. 300–39; Anthony & Ringe 2015) (Section 3.1). Bronze Age spreads of an IE branch from the Ural and Altai centers were an important element in the spread of early Uralic.

Later, beginning in the early centuries of this era, the Migration Period involved mass migration from the western steppe (moving away from Hun and then Avar oppression) to northern and western Europe, leading to Russian northward expansion into Finnic-speaking lands. The early Viking Age brought about contacts of early North Germanic with ancestral Saami languages and created or strengthened Germanic-Finnic contacts, with assimilation in both directions, around the Baltic coast. From the Middle Ages on, Russian expansion to the north and east in pursuit of control and expansion of the fur trade (not covered in this review) brought ecological and cultural destruction across the northern forest and the first stages in the decline of the indigenous Uralic languages.

3. INDO-EUROPEAN AND URALIC

3.1. Indo-European Background

A discussion of the Uralic spread should begin by examining the relationship of early Uralic to early IE. The IE family is about 6,000 years old (Anthony 2007, Parpola 2012, Anthony & Ringe 2015, Chang et al. 2015).² Proto-Indo-European (PIE) had two economic and cultural centers: one in the west (near the Dnieper, close to the Caucasus-Carpathian metallurgical center and Old European culture) and one in the east, along the lower Volga (in the Khvalynsk culture and its successors). The eastern center is the origin of important and archaeologically visible elements of PIE religion and culture and of most of the genes. The western center is likely to have been the proximate source of wool technology and wheeled transport technology—the two markers of the IE spread that are critical to determining the age and homeland of PIE. All IE branches have cognate terminology for wool and wool technology; all but Anatolian, the first to branch off, have terms for wheeled transport (e.g., ‘wheel’, ‘axle’, ‘convey’). PIE necessarily originated in a place where wool technology appeared not long before wheeled transport, and only the western steppe and the time frame of around 6000 to 5500 BP match that description (Darden 2001).

The PIE language is generally assumed to have come from the eastern center, but the point has not been argued. The map of the PIE homeland from Anthony & Ringe (2015, p. 209) shows the entire area from west of the Dnieper to the southwestern Urals and from the southern dry steppe to the forest-steppe zone in the north as PIE homeland, but this map covers a time period of 6200 to 5000 BP. Before the use of wheeled transport and the development of nomadic steppe pastoralism, this area would probably have been too large for the range of a single language. Just where in this large range the PIE language spread from has not been explicitly argued by linguists. The pre-transport spread of the Anatolian branch (the first split in the PIE tree, c. 6000 BP) proceeded south from near the Dnieper to the Danube and through the Balkan peninsula to Anatolia, and this trajectory appears implausible if its PIE source was spoken east of the Volga but plausible and in fact likely if it was spoken around the lower Dnieper.

Beginning around 5300 BP, some 700 years after the Anatolian separation, late PIE language and culture spread rapidly and dramatically from the lower Volga across the entire western steppe,

²The cited works by Anthony (2007) and Anthony & Ringe (2015) are the sources of statements here about IE that are not otherwise attributed.

carried by the Yamnaya culture, which brought wheeled transport, pastoral nomadism, new genes, late PIE speech, and reorganized and expanded political, economic, and kinship networks across the steppe. At the time, this was the biggest economic, cultural, and social leap ever seen on and around the western steppe and in southeastern Europe. Neighboring societies tended to shift to PIE speech (and economy, culture, and networks), and all languages nearby must have borrowed PIE vocabulary, saliently including terms for wheels, wheeled transport, and domesticated horses. No Uralic language has native terms for these things; the earliest stratum of IE loans in Uralic languages is from early Indo-Iranian c. 4,000 years ago (as discussed in Section 3.2).

At the western steppe periphery, a fusion of IE language, economy, culture elements, and genes together with east European agriculture created the archaeological Corded Ware complex, which subsequently spread back east through the forest-steppe zone at the northern steppe periphery. The easternmost Corded Ware culture, the Fatyanovo culture (c. 4800–4200 BP), spoke a Pre-Proto-Balto-Slavic language whose only modern descendants are the Slavic and Baltic languages. (For the Fatyanovo culture, see Nordqvist & Heyd 2020; all statements about its language are mine.) Distinctively Baltic names of rivers are found throughout the forest-steppe and forest zones of the Dnieper basin—today’s Belarus and nearby, as far east as the Oka basin [see Toporov & Trubačev (1962), who reconstruct at a time depth of about 4000+ to 3000 BP]. This area was the westernmost Fatyanovo range and was (Pre-)Baltic-speaking until protohistorical times when the fifth-century Slavic expansion absorbed and displaced the Baltic population, whose languages survive in today’s Latvian and Lithuanian to the northwest.

The Fatyanovo culture originated in the Dnieper area and spread eastward along the middle Volga to the Urals, where its easternmost portion is known as the Balanovo culture (for these and other archaeological terms, see **Figure 1**). The Balanovo and eastern Fatyanovo economies centered on copper mining and then on bronze production. Balanovo appears to have assimilated some of the indigenous hunter-fisher groups, and they too were involved in mining and metallurgy. To the west, the Fatyanovo culture was based on small-scale farming and livestock husbandry. Had an IE language from the Volga-Ural portion of the range survived to modern times, it would probably be considered a separate branch of IE, though its relationship to Balto-Slavic would also be clear. Early in its formation, the language of the Fatyanovo culture would have been a dialect chain; by the time it came to an end some 600 years later, separate languages would have developed, though probably still with local mutual intelligibility. For convenience, I refer to its western portion, in the Dnieper basin, as Baltic or Pre-Baltic, and to the eastern portion as Para-Baltic.³ Loans from Para-Baltic into adjacent languages, had any survived, would easily be recognized as IE and Balto-Slavic-like. Such borrowing must have happened as adjacent hunter-fisher societies interacted with the Fatyanovo culture and sometimes kept a few livestock, and terminology for livestock and other IE cultural innovations would surely have been borrowed. The fact that no such loans survive in the Uralic languages indicates that today’s Uralic languages do not descend from the northern neighbors of Fatyanovo.

Drastic changes occurred along the middle Volga and southern Urals beginning c. 4200 BP when two important near-simultaneous events occurred: the 4.2 ka event and the Seima-Turbino transcultural complex. The 4.2 ka event, a climate episode that lasted from about 4200 BP to about 3900 BP, brought global drought to continental interiors, accompanied by some cooling;

³Following common historical-linguistic usage, the prefix *Para-* labels languages or branches that are sisters of another language or branch; Para-Baltic is a sister to all of Baltic (or, more properly, all of Balto-Slavic) and is not just a Baltic daughter language. The ancestor of Balto-Slavic and Para-Baltic would be called Macro-Balto-Slavic. Pre-Baltic is any prehistoric stage of some or any Baltic languages (distinct from Proto-Baltic, which is a linguistic reconstruction that is unspecified as to time or space).

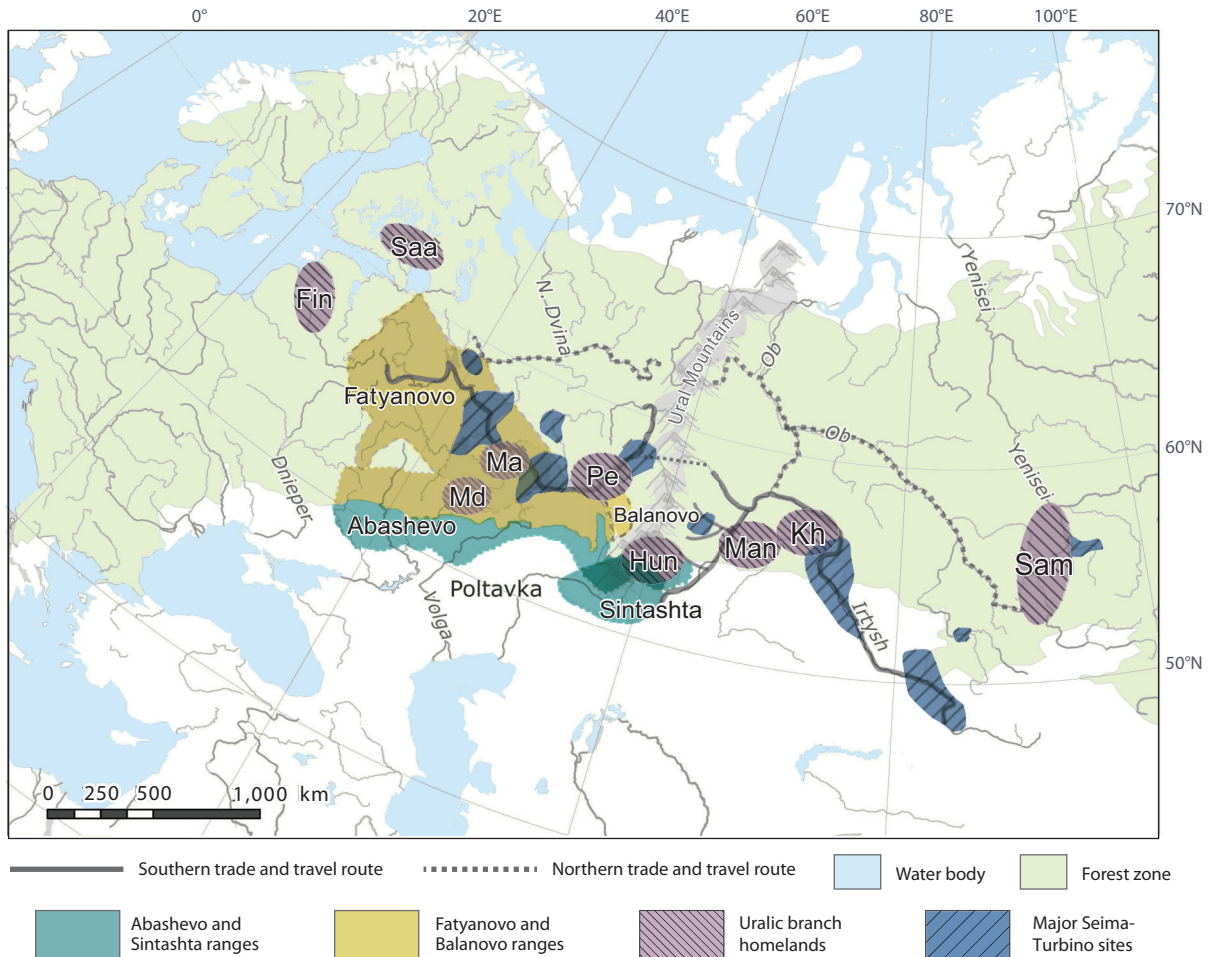


Figure 1

Linguistic and archaeological distributions c. 4200–4000 BP. The light green band across the entire area indicates the recent and modern forest zone (steppe to south, tundra to north). The languages and cultures arose in the following chronological order: Fatyanovo, Balanovo, and then Abashevo for the Baltic- and Para-Baltic-speaking post-Corded Ware Indo-European cultures; Poltavka and then Sintashta for the Indo-Iranian-speaking cultures. Dark blue indicates Seima-Turbino major sites. Purple ovals indicate core locations of Uralic branch ancestors: Saami (Saa), Finnic (Fin), Mordvin (Md), Mari (Ma), Permian (Pe), Hungarian (Hun), Mansi (Man), Khanty (Kh), and Samoyedic (Sam). The core locations are known secondary staging grounds for Saami and Finnic. Thick lines along rivers indicate the southern trade and travel route. Dotted lines indicate the northern route. Figure based on research by Grünthal et al. (manuscript in preparation) (CC BY), Saarikivi (2020), Lang (2018), and Anthony (2007). Map provided by Nora Fabritius.

this affected lands from the arctic to Southwest Asia, including most of the modern Uralic range (Helama & Oinonen 2019, Perşoiu et al. 2019). West of approximately the Oka, Europe may have been cooler and wetter than usual. The Seima-Turbino complex was a trade network extending from the Altai region to northern Europe, marked most clearly by bronze artifacts—chiefly ones that were not only tools but also symbols of power, such as swords, axe heads, and spearheads (Anthony 2007, pp. 443–48; Marchenko et al. 2017). Grünthal et al. (manuscript in preparation) argue that the two developments were connected: During the drought, the nomadic pastoral economy of the steppe collapsed, and herders crowded into the southern Ural foothills, where the

drought was less severely felt (Anthony 2007, pp. 389–93). Competition and warfare were intense, and longer and closer proximity to livestock probably triggered bouts of disease. Formerly nomadic steppe pastoralists (Indo-Iranian-speaking) formed permanent settlements and turned to metallurgy. Production of bronze and bronze artifacts intensified, and trade networks stretched to the Altai, Southwest Asia, and Europe.

Climate amelioration began c. 3900 BP, and the Seima-Turbino bronze trade came to an end. East of the Oka, the previous Fatyanovo-Balanovo culture and its offshoots were replaced by evidently Uralic-speaking cultures, which continued the economic emphases of the former Fatyanovo and Balanovo cultures. To the west, in the Dnieper drainage, Pre-Baltic languages remained. The pastoral nomadic economy of the steppe recovered, and Indo-Iranian languages spread rapidly across the entire western steppe and most of the Kazakh steppe and began to make incursions in Southwest Asia. This spread began before the climate amelioration in the southeastern Urals where the Indo-Iranian metallurgical centers were located; that is, the southeastern Urals were the dispersal center and proto-homeland of Proto-Indo-Iranian and then Proto-Iranian. Before that, the ancestor of Proto-Indo-Iranian had almost certainly been the language of the nomadic pastoral Poltavka culture of the Caspian steppe (Mallory 1997; Anthony 2007, p. 306). Poltavka was a descendant of the Yamnaya culture, which had remained a nomadic pastoral culture all along.

Around this same time, the Indo-Iranian contact episode began for the Uralic languages. Indo-Iranian input ranges from Pre-Indo-Iranian to Proto-Indo-Iranian to early Iranian and varies from branch to branch of Uralic; the contact extended over time and space and affected an already incipiently diversified early Uralic (Holopainen 2019). Importantly, it affected not PU but, rather, its separate branch ancestors.

3.2. The Proto-Uralic Homeland

At some point in this general time frame, PU dispersed and the nine elementary branches spread to their reconstructed branch homelands (Saarikivi 2020; Grünthal et al., manuscript in preparation). The spread appears to have been rapid, large, and sweeping: It brought branch protolanguages to their core locations, which stretched from the southern Urals to the Volga-Oka confluence, with no evidence of substratal effects on grammar, lexicon, or toponymy (except in the easternmost branch, Samoyedic) and no evidence of isolation by distance effects on lexicon or grammar, as would have been expected in a more gradual expansion. These facts plus the distribution of the Indo-Iranian loans point to a rapid spread not long before the Indo-Iranian contact episode. Samoyedic is exceptional in showing evidence of substratal lexical influence while preserving a conservative inflectional morphology and phoneme inventory, in preserving notably fewer PU lexical items than the other branches, and in having very few Indo-Iranian loans [Holopainen (2019) finds only 4–8 in Samoyedic languages compared with 28–76 in other branches]. In the other branches, the number of Indo-Iranian words varies but always in proportion to the total number of PU words retained; thus, apart from Samoyedic, the variation is due to branch-internal and language-specific vocabulary loss rates and not to different degrees of contact intensity. The traditional explanation of the vocabulary discrepancies posited an initial bifurcation of PU into Samoyedic versus Finno-Ugric, as early as 6000 BP or before, allowing enough time for vocabulary attrition in Samoyedic; however, this analysis is at odds with the conservative morphology of Proto-Samoyedic and a general lack of sound changes that would distinguish Proto-Samoyedic from the putative Proto-Finno-Ugric. The limited but independent Indo-Iranian borrowings in Samoyedic indicate that it was geographically distant from the rest of the family and from the Indo-Iranian sphere of influence. The Ugric languages are intermediate in their lexical retention rate and number of Indo-Iranian borrowings, and overall the evidence from sound changes and lexical sharing indicates less cohesion and contact, from the start, among the languages now

east of the Urals—Ugric and Samoyedic—than among those along the Volga west of the Urals (Grünthal et al., manuscript in preparation). The higher-level structure of the Uralic family tree and the status of Samoyedic in it are still unsettled (Salminen 2001, Häkkinen 2009, Lehtinen et al. 2014, Kallio 2015, Aikio 2020); the chief contenders are the traditional model with the initial bifurcation of Samoyedic versus Finno-Ugric and subsequent west-branching bifurcations, and a rake or star phylogeny with an initial split into the nine elementary branches. It does, however, seem clear that on any view, diversification and geographical dispersal began earlier among the eastern branches.

What we need to know is whether the lexical loss rate of Samoyedic is significantly greater than for the other branches. If it is, then something out of the ordinary must have happened with Samoyedic. If not, then the lexical losses are within the expected range of variation, and nothing special needs to be posited—neither a sociolinguistic discontinuity nor a much earlier separation. The lexical loss rate for Samoyedic cannot be assessed using available PU word lists because, for all of those, a word is defined as PU if it is attested in Samoyedic and one other branch, so the Samoyedic retention rate is 100% by definition. The working word list of 500 PU items used by Aikio (2020) is based on other criteria; when that word list is complete and published, it will be possible to judge whether Samoyedic has or has not gone through a diachronic singularity.

The initial divergence of PU, then, had already occurred by 4000 BP, the approximate time frame of the Indo-Iranian contact episode. As of about 3900 BP or earlier, the Fatyanovo-Balanovo culture was replaced by the Textile Ware culture from around the Volga-Oka confluence in the west almost to the Kama in the east, and east of that by the Akozino-Mälar and Anan'ino cultures extending to the Urals. These cultures were evidently Uralic-speaking (Parpola 2012); all three were continuous to historically attested Uralic-speaking societies. Akozino-Mälar and Anan'ino produced bronze items that were traded as far as the Baltic coast and Scandinavia; the Textile Ware culture was based on small-scale farming and livestock husbandry. Subsequently, ancestral Finnic and Saami began to spread westward from the Textile Ware area to the southeastern Baltic coastal area (today's Estonia) and today's southeastern Finland, respectively; still later, ancestral Saami spread northwest, reaching western Norway and all of northern Fenno-Scandia, and ancestral Finnic later spread northward, displacing or absorbing most of the prior Saami languages (Aikio 2012; Lang 2018; Grünthal et al., manuscript in preparation). Also after the initial Uralic divergence, most of the daughter languages engendered northward spreads across the entire northern forest and, with the later arrival of reindeer herding, into and across the tundra. These northward spreads proceeded more gradually except for spurts triggered by large-scale reindeer herding and climate changes (Aikio 2012; Khanina et al. 2018; Khanina & Gusev, manuscript in preparation).

The mainstream view among archaeologists since the mid-twentieth century places the PU homeland somewhere in the middle Volga region. This view is based primarily on two arguments: internal linguistic diversity and connections to PIE.⁴ The criterion of greatest diversity as evidence of homeland location (known as the principle of least moves or center of gravity) seeks

⁴Tree names are sometimes invoked as evidence for the PU homeland, but in fact the reconstructible tree names have little diagnostic value for PU. The same three coniferous genera (pine, fir, spruce) occur in both the north European and west Siberian boreal forests, and there are PU terms for all three; any plausible homeland between the upper Volga and the Yenisei is in this range. There are no PU terms for broadleaf trees; the middle Volga is in the transition zone from central European broadleaf to north European coniferous forests, so if the PU homeland had been there, the central Uralic branches might well have had cognate terms for broadleaf trees, but all other branch homelands are in coniferous taiga zones, so any such PU words would probably have been lost from them (including from Samoyedic, which defines Proto-Uralic). A term reconstructed as meaning 'Siberian stone pine, *Pinus sibirica*' refers to a species found only east of the Urals. It is likely a species-specific name as this pine yields edible nuts that are a valued trade item today and might conceivably

the homeland near the projection of the root of the family tree on the map [Sapir 1949 (1916); Dyen 1956; Diebold 1960; Nichols 1997, pp. 371–72]. In practice, this means that the homeland is near where the earliest-diverging branches meet. The principle requires knowledge of the family's higher-level branching structure, which has been debated for Uralic. On the traditional view, in which the initial split was between Samoyedic and Finno-Ugric, the projection of the root is where any Samoyedic language meets any Finno-Ugric language. That configuration occurs at the western edge of the Samoyedic range, not far west of the Yenisei, where eastern Khanty meets southern Selkup and Forest Nenets; and also farther north where Komi meets Forest Nenets. On the view of the Uralic family tree as a simultaneous split of the nine elementary branches, there is no projection of the root, and the principle is inapplicable. In practice, advocates of a middle Volga homeland often mention the area of greatest diversity in the sense of the greatest number of different branches: Mordvin, Mari, and Permic are all found along the middle Volga. However, this view does not reflect the technical sense of phylogenetic diversity underlying the original criterion, and in any case three distinct branches also approximate each other in the eastern Baltic area (where Saami, Finnic, Mordvin, and a Mari-like or Permic-like language were present until the late Middle Ages; see Rahkonen 2013, Saarikivi 2020) and in the lower Yenisei region where Samoyedic, Komi, and Khanty meet. Stronger support for the middle Volga homeland is usually voiced in terms of early connections between PU and PIE. Lexical resemblances between PIE and PU have long been noted, and recent comparative work has pared the list down to about a dozen carefully reconstructed words, which are viewed variously as early loans (Koivulehto 2001) or cognates (Helimski 2001). The comparative work has been done with sophistication and rigor, but the number of resemblant words is not enough to exceed chance, given the size of the total wordstock surveyed and the semantic and formal latitude allowed in the search (Nichols & Rhodes 2018; see also Nichols 2010; regarding these and other statistical problems, see Simon 2020).⁵

There is thus no demonstrable relatedness, whether by borrowing or by descent, between PIE and PU; the resemblances are indistinguishable from the chance look-alikes that can be found between any two languages. In reality, of course, sets like PU **weti*: PIE **watar-/weten-* 'water' and PU **nimi*: PIE **h₁nomn-* 'name', and the resemblances in personal pronouns and person suffixes discussed below, are striking and raise questions. It is conceivable that they are genuine sharings and very ancient—so ancient that too few survive to exceed chance—and that they antedate PIE stem formation with extensions and/or the development of the PU CVCV stem canon. But such developments, if real, preceded PIE and PU and also the Pre-PIE and Pre-PU that can be reached by internal reconstruction, so it is not clear how or whether they bear on either homeland.

Stronger evidence against a middle Volga homeland for PU is the absence of any early IE loans into Uralic before the Indo-Iranian contact episode, which was demonstrably post-PU. The Yamnaya spread and the later Fatyanovo-Balanovo spread along the middle Volga brought major cultural and technological advances, and it is hard to believe that words for domestic livestock, wheeled transport, wool, and early IE cultural institutions would not have been borrowed by neighboring languages including PU.

Furthermore, it is unlikely that Proto-Samoyedic could have spread well to the east of the Iranian cultural sphere by the time of the Indo-Iranian contact episode. In the northern forest zone, where the early Uralic branch homelands are located, travel was usually on rivers, which

have been traded beyond the tree's natural range in ancient times, so the reconstruction does not entirely rule out a homeland west of the Urals. However, the term is found only in Samoyedic, Ugric, and Permic, which are in the tree's range (Permic is only barely so).

⁵ Comparisons of Uralic or Proto-Uralic with other languages and families of northern Eurasia (e.g., Yukagir, Turkic-Tungusic-Mongolic) have the same statistical problem but often less rigor.

were direct, level, and navigable by canoe in summer and by sleigh or skis in winter. The Volga offered ideal rapid connections west of the Urals. East of the Urals, the major watercourses run south to north or southeast to northwest, so east–west travel to the Proto-Samoyedic homeland in or near the Minusinsk basin area on the upper Yenisei would have been indirect or overland, and positing a western homeland offers no rationale for the migration there in the first place. The very lack of cohesion shown in the lexical distributions among Ugric and Samoyedic languages (Grünthal et al., manuscript in preparation) is evidence against a timely eastward migration of Samoyedic from the Volga or even a beeline migration impelled by knowledge of resources available in the Altai-Sayan area. A slower and shorter spread from somewhere east of the Urals would allow more easily for Proto-Samoyedic to distance itself from the others. Alternatively, it would make possible a beeline migration from near the Urals to the upper Yenisei, impelled by locally available knowledge about the Altai. Either way, distancing ancestral Samoyedic from all but the most marginal contacts with Iranian-speaking populations (which, at the time of the Indo-Iranian contact episode, were found in and near the southern Urals on both sides and were beginning to spread into the northern Kazakh steppe) is implausible with a middle Volga starting point.

It has long been noted that PU preserves vocabulary for the natural world of the taiga environment but not for tundra, steppe, or coastal ecologies. Uralic spreads to the tundra and arctic coasts are demonstrably secondary and branch-specific or language-specific. Absence of tundra or maritime vocabulary is confirmation of an origin farther south but is not required as evidence. Knowledge of steppe ecologies is another matter: It is possible that PU had some familiarity with steppe ecology and some relevant vocabulary but that environmental terms have survived only for the environment familiar to all the daughter branches, which is the taiga. An origin at the southern edge of the attested Uralic range fits the standing northward spread pattern better. There is toponymic evidence that the Ob-Ugric languages are secondary arrivals in their current northern ranges (Matveev 1962, pp. 292–97), and the Ugric languages have terms for ‘horse’ and aspects of horse culture and nomadism, some from Iranian and some from an unknown source. Thus, the Ugric languages, at least, may have originated in the south of the composite attested Ugric range.

In summary, the PU homeland needs to be placed east of the Urals, in the south of the attested or reconstructed Uralic range (specifically, east of the southern Urals), in or near a taiga ecology, possibly within striking range of the steppe, with easy access to the Volga and the southern Ural mining and metallurgical centers. The western dialects need to have been positioned to spread readily along the Volga when the Fatyanovo and Abashevo cultures faltered; a gradual diffusion of Finno-Ugric speakers into Para-Baltic-speaking societies until they gained numerical superiority seems unlikely to have produced Finno-Ugric dialects with no evidence of substratal vocabulary, toponymy, grammar, or phonology. The eastern dialects were positioned to spread into environments with less direct connections and less effective network reaches. Ancestral Samoyedic may have been in contact with a non-Uralic society that was familiar with Altai metallurgy or had access to it, which would have drawn Samoyedic eastward and out of the Indo-Iranian orbit; such a scenario might also account for the lexical losses and replacements in Samoyedic. The spread positioned the rest of the family to take in loan vocabulary from Indo-Iranian in good numbers, over some time, and independently in each branch. The entire family appears to have been well equipped, culturally and technologically, to step into trade networks, metallurgy, and farming and herding when the occasion presented itself.

One possibility would be a homeland along the middle Ishym and/or Irtysh and/or upper Ob’, and perhaps the Tobol. Some such range is comparable in size to a single-language range in historical western Siberia. Shortly after the initial spread, the dialects along the Volga and the Ugric dialects were within range of Iranian societies. Settlements along the Volga were in direct contact with Indo-Iranian. Ugric communities were far enough west for some contact with

Indo-Iranian-speaking communities near the Urals and/or far enough south for contact with Indo-Iranian-speaking nomadic herding societies on the Kazakh steppe.

On this scenario, early Ugric stayed close to the homeland, Samoyedic moved to the east, and the rest of the family moved somewhat west and was close to the southeastern Ural foothills when bronze work and bronze trading began to expand under Indo-Iranian control. As the herding and farming economies west and southwest of the Urals faltered in the 4.2 ka event, the western Uralic dialects spread rapidly westward along the Volga. In approximately the same time frame as the Indo-Iranian episode, the Seima-Turbino waterborne bronze trade network extended from its earliest sites on the upper Irtysh, near the Altai copper and tin mines, to the Baltic, leaving hoards and burial sites along the middle and upper Volga, the middle Kama, and the middle and upper Irtysh, scattered sites from the Irtysh to the upper Yenisei, and a few sites along the northern river route (Northern Sosva, Pechora, Izhma, Vychegda, Sukhona). In the west, both the Volga and northern routes connected via lakes and rivers to the Gulf of Finland. At least by this time, early Uralic speakers were likely involved in the Seima-Turbino network, probably as indigenous boatmen who provided the logistics and possibly also as miners and metalworkers.

In summary, a PIE homeland on or near the lower Volga is possible for earliest PIE (though it makes the Anatolian migration through the Balkan peninsula less straightforward) and is demonstrable for the Yamnaya spread, but a homeland along the Volga, or more generally anywhere west of the Urals, is not possible for PU, as shown by valid counterevidence and the absence of any supporting evidence. Research on the PU homeland needs to start afresh with no assumptions about connections to PIE. One possibility has been sketched out here.

4. MORE ON URALIC ORIGINS

4.1. Lactose Tolerance

Earlier publications report high rates of lactase persistence/lactose tolerance (ability to digest milk in adults) among today's Mordvin and Udmurt population [see Kaiser (2004), reporting work by Leena Peltonen that I have not located] and in prehistoric Corded Ware populations (Allentoft et al. 2015, Mathieson et al. 2015). According to Kaiser (2004), the mutation along the Volga was dated by Peltonen at 6600–4800 BP—early for the Uralic spread and even for the Corded Ware eastward spread. Appreciable rates in a population indicate regular consumption of dairy products and are inexplicable in early Uralic populations unless those populations absorbed sizable numbers of dairying people. For this to have arisen in the Uralic spread along the Volga would require Uralic-speaking societies to have absorbed significant numbers of Corded Ware people and probably to have been outnumbered by the latter. Acquisition of the gene earlier in Siberia is implausible. Even if, speculatively, the gene appeared in the Botai culture of the late Neolithic northern Kazakh steppe, where the horse was domesticated early and dairy products were used (Outram et al. 2009, de Barros Damgaard et al. 2018), it is very unlikely to have spread to Uralic-speaking foragers in numbers sufficient to account for the high rates that occurred later on along the middle Volga. The history of lactose tolerance needs more study, but it does appear that early Uralic, or at least Finno-Ugric, is part of that history.

4.2. Eastern Typological Traits

PU had (and the modern Uralic languages mostly still have) a number of typological traits that are typical of the greater North Pacific Rim area and not of western Eurasia.⁶ These traits include

⁶The Pacific Rim, as defined by Bickel et al. (2017), comprises the Pacific-facing coastal and near-coastal areas up to the far side of the major coast range; the North Pacific Rim extends from Japan to southern California.

high frequency of nonfinite verb forms in nonmain clauses, extensive use of inflectional person marking, personal pronoun roots that contain no semantic feature of person (which is marked only inflectionally), salient head marking, relatively high frequency of flexible noun-verb roots in the lexicon, and traces of nonaccusative alignment; the base of verb derivational paradigms is usually the intransitive form (Grünthal et al., manuscript in preparation).

These traits are shared with many Siberian languages but are (otherwise) rare in Europe. More precisely, the patterns, as types, are found throughout Siberia and nearby; in Uralic, the morphemes marking them are generally native to Uralic. Several of them are attenuated in the western languages, chiefly Saami and Finnic, but on the whole they are fairly well preserved. Thus, PU and the modern languages have their closest typological affinities with Siberian and North Pacific Rim languages, and it is at least plausible that Pre-PU migrated to the PU homeland from farther east. It is difficult to be more precise because much of the evidence for ancient Siberian typology was obliterated in the later westward and northward spreads of Turkic, Tungusic, and Mongolic.

It is also difficult to be precise about the areal affinity of PIE, the other old inner Eurasian language family, because its spread obliterated all of its sisters and near neighbors. In the broader Eurasian perspective it appears to be more Eurasian than Near Eastern or Anatolian and to be more western than eastern (Nichols 2007). It may be relevant that although the Saami branch had no direct contact with Baltic and had close contact with Germanic only after its spread across Scandinavia and after at least some differentiation of its major branches, Saami appears to show much the same degree of westernization as Finnic, which was in close contact with Baltic and Germanic from the early stages of its westward spread (Grünthal 2012, Junttila 2012). Western features include loss of object indexation in the verb, attrition of possessive person inflection, and development of a personal pronoun paradigm whereby the root carries person-number meaning and the inflection marks only case and not person. The westernization in Saami may suggest that the extinct pre-Saami substratal languages were also of a western type and that, therefore, their grammatical influence on Saami was similar to that of Baltic and Germanic on Finnic. This would mean that at least some non-IE languages of Fenno-Scandia belonged to an areal type similar to that of IE.

4.3. The Inner Asian Type

Uralic languages also have a number of traits that define what I call the Inner Asian type.⁷ These traits are not all typical of the North Pacific Rim or Siberian linguistic populations but rather are hallmarks of a smaller set of families: prototypically, Tungusic, Turkic, Mongolic, and Uralic or at least eastern Uralic; to a lesser extent, Korean, Japonic, and Eskimo-Aleut. Chief among these traits are the following:

- Head-final word and morpheme order at all levels: verb-final clauses, noun-final NPs, postpositions, possessor preceding possessed noun; synthetic compounds with the second element as head; primarily or exclusively suffixing morphology (inflectional and derivational)
- Strongly configurational syntax: strict word order, strict phrase boundaries, recursive embedding, phrase structure within phrases

The greater Pacific Rim, also termed Trans-Pacific and Circum-Pacific, extends inward to include all of North America and Siberia at least to the Lena and perhaps farther. It is difficult to set a western limit in northern Eurasia because there is no single coastal range and there is considerable ecological and human-geographical continuity all across northern Eurasia.

⁷This type is sometimes called the Ural-Altaic type or the Altaic type, but these terms have often been used as names of unproven macrofamilies, and I avoid that ambiguity.

(These first two define a morphosyntax that is self-similar at all levels and structured by few and simple principles.)

- Monoexponential (separative) encoding of inflectional and derivational categories (one category per morpheme; few or no portmanteau morphs)
- Transparent morpheme boundaries; minimal sandhi

(These two define what is often called the agglutinative type.)

- Simple syllable structure; in particular, simple onsets
- Minimal or no prosodic contrasts (such as tone or contrastive stress positions)
- Context-free parsing of certain kinds; in particular, minimal or no effects of person hierarchies on the coding or interpretation of combinations of arguments, and minimal lexical classes determining inflectional paradigms
- A tendency to develop vowel harmony, either palatal-velar or ATR (advanced tongue root); this trait is well developed in several modern Uralic languages but does not reconstruct to PU

Cross-linguistically, strict word order and strict phrase boundaries correlate negatively with rich morphology, but the Inner Asian languages are distinctive in combining these traits with rich inflectional and derivational morphology.

Taken together, these tendencies define a grammatical structure that must be adaptive in situations involving contact, language shift, rapid adult language learning, and communication in a language that one or both parties do not know well. Utterances can be put together and decoded using the minimal resources of words (or stems) and ordering principles with minimal context-specific or lexeme-specific rules. All of these features would be useful in any high-contact situation, so it remains to be explained how and why they took root specifically in northeastern Eurasia. However, once installed, they appear to have been quite stable as a package, changing little even in Saami and Finnic, where there has been extensive Baltic and Germanic contact, or in Hungarian, where there has been extensive Iranian, Slavic, and Germanic contact. The self-similarity probably serves to enhance stability.

The type is very different from the templatic and often polysynthetic type that is common around the North Pacific Rim. It is also quite different from the type of PIE and early IE. As Meillet [2005 (1908), p. 321] writes, “In addition to the meaning expressed in the stem, inflection marks the role played by each word in the sentence; the word is therefore autonomous and is sufficient in itself to indicate its own meaning and role in the discourse. This complete autonomy of each word represents the fundamental property on which the whole structure of the Indo-European sentence rests.” This point is now most often cited in connection with grammaticalization (a word loses its autonomy and becomes a grammatical morpheme) and diachronic loss of inflection, but Meillet seems to have also had in mind the flatter syntax of early IE, where word order was often flexible, phrase boundaries were violable and phrases could be discontinuous, inflection of words was the main indicator of syntactic relations, and there was a good deal of portmanteau, suppletive, and idiosyncratic morphology—a package exploited conspicuously in virtuoso IE poetics.

Several of the Inner Asian properties are attractors in the terms of Nichols 2018: traits that are targets of selection or universal bias and that diffuse under the right sociolinguistic conditions (symbiosis in the terms of Nichols 2018). The attractors include personal pronoun systems with first-person *m* and second-person *t*, *č*, or the like; productive morphological causativization (intransitive verb base, an eastern feature; see Grünthal & Nichols 2016, Nichols 2018); and head-final self-similar morphosyntax. The sociolinguistic situation involves bilingualism with neither

language official or strongly dominant, easy tolerance of code switching, and languages that remain discrete despite code switching. This situation easily allows selection of favored forms and their fixation as loans, even for such things as personal pronouns and phonological canons for pronouns, which are ordinarily not readily borrowed. Fixation of an individual instance of spontaneous selection in code switching must be most likely in small speech communities, where the speech of even one individual can have appreciable impact. The best living example of symbiosis is found in Khamnigan Mongol-Evenki (Janhunen 1990, 1991, 2005).

The clearest cases of the Inner Asian type (including the attractors) are in young families (Turkic and Tungusic are of Romance-like age, and Mongolic, Korean, and Japonic are younger). Uralic is the oldest family in the group, and all properties of the type reconstruct to PU. This means that the type has existed, fully developed, in northern Asia for at least 4,500 years. The PU homeland, wherever it was, was distant from those of the Turkic, Mongolic, and Tungusic languages (which were close neighbors in eastern Manchuria; see Janhunen 1996). Contact and substratal effects at the northern frontiers of the Tungusic, Mongolic, and Turkic spreads do not weaken the type; thus, substratal languages also probably exhibited it. The type, then, may have been widespread in Siberia and nearby; alternatively, the presence of the type is another eastern feature in Uralic. Either way, the sociolinguistics of small speech communities, sparse populations, travel across ethnolinguistic boundaries, and clan lines continuing across ethnolinguistic boundaries may have shaped PU grammar and lexicon and may be reconstructible as PU background. If so, this package may be part of what enabled early Uralic speakers to assume management of trade routes, mining, and manufacture from the IE speakers of the Urals and middle Volga.

4.4. Grammatical Complexity

Language spreads that involve language shift are expected to decomplexify languages, as absorption of an appreciable population of adult second-language learners commonly decomplexifies a language (Trudgill 2011). The mean morphological complexity of Uralic languages is lower than the general north Eurasian mean and comparable to that of languages that have undergone large spreads, such as German, Spanish, Turkish, Yakut, and Mongolian; however, the Samoyedic languages, whose morphology is fairly archaic overall, are among the most complex (Grünthal et al., manuscript in preparation). In those Uralic languages where complexity is highest, much of that complexity is due to post-PU and postbranch developments.

The foregoing concerns enumerative complexity (number of items in the inventory, for some sample of inventories—e.g., number of conjugation classes, number of genders; see Nichols 2019) and canonical complexity (number of nontransparencies in a sample of inflectional paradigms, counted as number of noncanonical patterns—e.g., number of conjugation classes that are unpredictable, number of suppletive patterns, number of portmanteau morphemes, each only for specified inflectional paradigms; for canonical patterns, see Corbett 2007, 2015; for canonical complexity, see Nichols 2020a). A third consideration is relational complexity (Nichols 2020b): the extent to which patterns such as person hierarchies and inversion draw their hierarchical ranking and constraints from general or universal principles rather than specifying them individually for each instance in the grammar. Person hierarchies, as noted, are largely foreign to Inner Asian morphology, in large part because they arise mostly in subject and object indexation, and the Inner Asian languages generally index only subjects. Uralic, unlike the rest, does index objects—but only for number and not for person. Such indexation is known as the objective conjugation in Uralic grammar: a separate paradigm of indexing person-number of subject and number of object, available to transitive verbs with an object present in the clause (Table 2). At first glance, this indexation is unexpected: Cross-linguistically, argument indexation is more common for person alone than

Table 2 Representative Tundra Nenets subjective and objective conjugations of ‘work’ and ‘do, take’

	Singular	Dual	Plural
Subjective conjugation			
First-person subject	mənc°raə-d°m	mənc°ra°-n°ih	mənc°ra°-waq
Second-person subject	mənc°raə-n°	mənc°ra°-d°ih	mənc°ra°-daq
Third-person subject	mənc°ra°	mənc°raə-x°h	mənc°ra°-q
	‘I work’, ‘you (sg) work’ (etc.)	‘we two work’ (etc.)	‘we work’ (etc.)
Singular object conjugation			
First-person subject	meə-w°	me°-m°ih	me°-waq
Second-person subject	meə-r°	me°-r°ih	me°-daq
Third-person subject	me°-da	me°-d°ih	me°-doh
	‘I take him/it’ (etc.)	‘we two take him/it’ (etc.)	‘we take him/it’ (etc.)
Plural object conjugation			
First-person subject	me-yə-n°	me-y°-n°ih	me-y°-naq
Second-person subject	me-yə-d°	me-y°-d°ih	me-y°-daq
Third-person subject	me-y°-da	me-y°-d°ih	me-y°-doh
	(take-PL.OBJ-1sg, etc.)		
	‘I take them’ (etc.)	‘we two take them’ (etc.)	‘we take them’ (etc.)

Conjugations are from Nikolaeva (2014, pp. 78–80). Dual object forms are not shown. Degree symbol indicates zero grade of underlying schwa.

for number alone (and is most common for person and number combined). However, in the Inner Asian type, where hierarchical patterns of any kind are disfavored, number marking makes sense: It allows objects to be indexed while running no risk of person hierarchies arising. Person hierarchies do arise in Uralic languages wherever there is object indexation for person. There is one such case: in the Mordvin languages, which have innovated full person-number indexation for both subject and object and have some neutralization patterns depending on the person of subject and object (see Toldova et al. 2018, pp. 352–56). Though indexing objects at all is generally foreign to Inner Asian languages, it is common around the Pacific Rim. The fact of any object indexation may be an eastern feature in Uralic, and the fact that it involves only number is a way to accommodate eastern object indexation to Inner Asian configurational morphosyntax.

5. CONCLUSIONS

The model of Uralic origin and dispersal sketched out in this review includes the following elements: The Uralic language family dispersed from a homeland in the eastern part of its range and east of the Urals. There is no connection, either areal or genealogical, between Proto-Uralic and Proto-Indo-European. The linguistic configuration around the Urals as of 4500 BP, shortly before the Uralic spread, involved the Fatyanovo-Balanovo-Abashevo complex along the Volga from the upper Dnieper to the Urals, speaking Para-Baltic dialects or languages; the nomadic pastoralist Poltavka culture on the steppe to the south, speaking the ancestor language of Proto-Indo-Iranian; and some interaction with indigenous hunter-fisher populations to the north, whose languages are unknown. The economies of the Balanovo and Abashevo cultures, close to the Urals, involved copper mining and working and then bronze manufacture, and indigenous populations were involved in this work.

In the 4.2 ka event, the economies of the steppe and the middle Volga contracted sharply. Steppe herders competed bitterly for the better-watered land around the southern Urals and eventually

turned to bronze manufacture and trade for their livelihood. Bronze trade connections linked the Altai-Sayan metallurgical area, the southern Ural metallurgical center, the cities of Southwest Asia, and societies around the eastern Baltic. The trade in bronze during this period made the long-standing waterborne trade routes of the forest archaeologically visible; the set of sites along the major rivers, known as the Seima-Turbino transcultural complex, is the signature of this trade. Evidently, Uralic speakers were centrally involved in this trade, as after climate amelioration c. 3900 BP the Fatyanovo, Balanovo, and Abashevo cultures were gone and the middle Volga was occupied by Uralic-speaking cultures that continued the economic pursuits of their Fatyanovo-Balanovo-Abashevo predecessors including the bronze trade with northern Europe. The homelands of the more western Uralic branch ancestors are near the Volga, and the later spreads of Saami and Finnic to their staging areas in southeastern Finland and northern Estonia, respectively, came from there, as did large northward spreads of some of the branches.

East of the Urals, Khanty and Mansi had homelands somewhere in the south of the forest zone, and Hungarian was on or near the northern Kazakh or southern Ural steppe; the populations of these homelands have shifted their languages to Siberian Tatar (Turkic), and the Ugric languages survive in the central to northern forest of western Siberia (Mansi, Khanty) or in central Europe (Hungarian). The center of Samoyedic dispersal is the Minusinsk basin (upper Yenisei, Sayan foothills); it is unknown how and when Proto-Samoyedic arrived there and what factors drew it there, but its prehistory may have involved loss of a good deal of Proto-Uralic vocabulary and borrowing of foreign substratal vocabulary. Not long after the Uralic dispersal, the Indo-Iranian contact episode (c. 4000 BP) brought Uralic languages into their first known contact with Indo-European. Over some stretch of time and some expanse of space, Uralic branch ancestors borrowed Indo-Iranian words (28–76 per branch except for Samoyedic, which has 4–8). These words give us a clear absolute date for the Uralic dispersal (not long before about 4000 BP) and show that all the branch ancestors except Samoyedic were within the Indo-Iranian sphere of influence. The further history of the family involves the westward spreads of Saami and Finnic; the northward and westward spread of Saami to cover most of western and northern Scandinavia; the northward spread of Finnish, assimilating most of the Saami speakers in Finland; and northward spreads from other branches (Permic, Ob-Ugric, Samoyedic), absorbing all previous populations so that, as of the late Middle Ages, most of northeastern Europe and western Siberia was Uralic-speaking to the tundra and coast, and Uralic entered history and popular consciousness as a northern language family.

Uralic belonged to a more eastern structural type than that of the Indo-European languages it encountered in its spreads. Proto-Uralic was the westernmost early representative of the Greater Pacific Rim linguistic type. It also belonged to the Inner Asian type, a local northern Asian or eastern Siberian areal type that has spread to cover much of Eurasia as a result of the Uralic spread and those of the Turkic, Tungusic, and Mongolic families. The Inner Asian type is well adapted to the communicative conditions of symbiosis or similar close contact and to the varied and sometimes short-term contacts of languages spoken by small and sparse populations. Uralic is the oldest known representative of the type, which has remained notably stable in the family despite the large spreads and diverse contacts of the family's history. Early Uralic speakers were well prepared by their geography, culture, and technological competence, as well as by their language typology, to step into existing trade networks, the metallurgical industry, and the territories and economies of farming and stockbreeding societies when the occasion presented itself. The 4.2 ka event and its effects on the various Indo-European-speaking societies provided just such an occasion, weakening preexistent economies, expanding trade economies and networks, and triggering the Uralic spread.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might reasonably be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

Supported in part by a Helsinki University Humanities Visiting Professorship (2017–2019) and a grant from the Russian Academic Excellence Project 5-100 to the Higher School of Economics, Moscow. I thank Riho Grünthal, Volker Heyd, Sampsa Holopainen, Juha Janhunen, Olesya Khanina, Matti Miestamo, Janne Saarikivi, and Kaius Sinnemäki for their contributions as my coauthors on the forthcoming work by Grünthal et al. that this review is based on; I also thank the anonymous reviewer for their comments.

LITERATURE CITED

- Abondolo D, ed. 1998. *The Uralic Languages*. London/New York: Routledge
- Aikio A. 2012. An essay on Saami ethnolinguistic prehistory. See Grünthal & Kallio 2012, pp. 63–117
- Aikio A. 2020. Proto-Uralic. See Bákro-Nagy et al. 2020. In press
- Allentoft ME, Sikora M, Sjögren KG, Rasmussen S, Rasmussen M, et al. 2015. Population genomics of Bronze Age Eurasia. *Nature* 522:167–72
- Anthony DW. 2007. *The Horse, the Wheel, and Language: How Bronze Age Riders from the Eurasian Steppes Shaped the Modern World*. Princeton, NJ: Princeton Univ. Press
- Anthony DW, Ringe D. 2015. The Indo-European homeland from linguistic and archaeological perspectives. *Annu. Rev. Linguist.* 1:199–219
- Bákro-Nagy M, Laakso J, Skribnik E, eds. 2020. *The Oxford Handbook of Uralic Languages*. Oxford, UK: Oxford Univ. Press. In press
- Balanovsky O. 2019. *Evolution of the Y-chromosomal pool in East Europe*. Presented at CONTACTS: Archaeology, Genetics & Languages, Suomenlinna, Helsinki, May 9–10
- Barfield T. 1989. *The Perilous Frontier: Nomadic Empires and China, 221 BC to AD 1757*. London: Blackwell
- Bickel B, Nichols J, Zakharko T, Witzlack-Makarevich A, Hildebrandt K, et al. 2017. *The AUTOTYP Database*, release version 0.1.0. <https://github.com/autotyp/autotyp-data>
- Carpelan C, Parpola A, Koskikallio P, eds. 2001. *Early Contacts Between Uralic and Indo-European: Linguistic and Archaeological Considerations*. Helsinki: Suom.-Ugr. Seura
- Chang W, Hall D, Cathcart C, Garrett A. 2015. Ancestry-constrained phylogenetic analysis supports the steppe chronology of Indo-European origins. *Language* 91(1):194–244
- Corbett GG. 2007. Canonical typology, suppletion, and possible words. *Language* 83(1):8–42
- Corbett GG. 2015. Morphosyntactic complexity: a typology of lexical splits. *Language* 91(1):145–93
- Darden BJ. 2001. On the question of the Anatolian origin of Indo-Hittite. In *Greater Anatolia and the Indo-Hittite Language Family*, ed. R Drews, pp. 184–228. Washington, DC: Inst. Study Man
- de Barros Damgaard P, Martiniano R, Kamm J, Moreno-Mayar JV, Kroonen G, et al. 2018. The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science* 360:eaar7711
- Diebold A. 1960. Determining the centers of dispersal of language groups. *Int. J. Am. Linguist.* 26(1):1–10
- Dyen I. 1956. Language distribution and migration theory. *Language* 32(4):611–26
- Feist T. 2015. *A Grammar of Skolt Saami*. Helsinki: Finno-Ugrian Soc.
- Grünthal R. 2012. Baltic loanwords in Mordvin. See Grünthal & Kallio 2012, pp. 297–343
- Grünthal R, Kallio P, eds. 2012. *A Linguistic Map of Prehistoric Northern Europe*. Helsinki: Finno-Ugrian Soc.
- Grünthal R, Nichols J. 2016. Transitivity/detransitivizing typology and language family history. *Ling. Posnan.* 58(2):11–31
- Grünthal R, Salminen T. 1993. *Geographical Distribution of the Uralic Languages*. Helsinki: Finno-Ugrian Soc.
- Häkkinen J. 2009. Kantaauralin ajoitus ja paikannus: perustelut putarissa. *J. Finno-Ugrian Soc.* 92:9–56
- Helama S, Oinonen M. 2019. Exact dating of the Meghalayan lower boundary based on high-latitude tree-ring isotope chronology. *Quat. Sci. Rev.* 214:178–84

- Helinski E. 2001. Early Indo-Uralic linguistic relationship: real kinship and imagined contacts. See Carpelan et al. 2001, pp. 235–63
- Holopainen S. 2019. *Indo-Iranian borrowings in Uralic: critical overview of the sound substitutions and distribution criterion*. PhD Diss., Univ. Helsinki, Helsinki
- Ilumäe AM, Reidla M, Chukhryaeva M, Järve M, Post H, et al. 2016. Human Y chromosome haplogroup N: a non-trivial time-resolved phylogeography that cuts across language families. *Am. J. Hum. Genet.* 99:163–73
- Janhunen J. 1990. *Material on Manchurian Khammigan Mongol*. Helsinki: Finno-Ugrian Soc.
- Janhunen J. 1991. *Material on Manchurian Khammigan Evenki*. Helsinki: Finno-Ugrian Soc.
- Janhunen J. 1996. *Manchuria: An Ethnic History*. Helsinki: Suom.-Ugr. Seura
- Janhunen J. 2005. *Khammigan Mongol*. Munich: Lincom
- Janhunen J. 2020. Grammaticalization in Uralic as viewed from a general Eurasian perspective. In *Grammaticalization Scenarios: Cross-Linguistic Variation and Universal Tendencies*, Vol. 1: *Grammaticalization Scenarios from Europe and Asia*, ed. W Bisang, A Malchukov, pp. 361–98. Berlin: De Gruyter Mouton
- Junttila S. 2012. The prehistoric context of the oldest contacts between Baltic and Finnic languages. See Grünthal & Kallio 2012, pp. 261–96
- Kaiser J. 2004. Ural farmers got milk gene first? *Science* 306:1284–85
- Kallio P. 2006. Suomen kantakielen absoluuttista kronologiaa. *Virittäjä* 110:2–25
- Kallio P. 2015. The language contact situation in prehistoric northeastern Europe. In *The Linguistic Roots of Europe*, ed. R Mailhammer, T Vennemann, BA Olsen, pp. 77–102. Copenhagen, Den.: Mus. Tuscul. Press, Univ. Copenhagen
- Khanina O, Koryakov Y, Shluinsky A. 2018. Enets in space and time: a case study in linguistic geography. *Finn.-Ugr. Mitt.* 42:109–35
- Klump G, Mazzitelli LF, Rozhanskiy F, eds. 2018. Typology of Uralic languages: current views and new perspectives. *J. Estonian Finn-Ugr. Linguist.* 9(1). <https://doi.org/10.12697/jeful.2018.9.1.01>
- Koivulehto J. 2001. The earliest contacts between Indo-European and Uralic speakers in the light of lexical loans. See Carpelan et al. 2001, pp. 235–63
- Lang V. 2018. *Läänemeresoome Tulemised*. Tartu, Estonia: Univ. Tartu Press
- Lehtinen J, Honkola T, Korhonen K, Syrjänen K, Wahlberg N, Vesakoski O. 2014. Behind family trees: secondary connections in Uralic language networks. *Lang. Dyn. Change* 4:189–221
- Mallory JP. 1997. Poltavka culture. In *Encyclopedia of Indo-European Culture*, ed. JP Mallory, DQ Adams, pp. 439–40. London: Fitzroy Dearborn
- Marchenko ZV, Syatko SV, Molodin VI, Grishin AE, Rykun MP. 2017. Radiocarbon chronology of complexes with Seima-Turbino type objects (Bronze Age) in southwestern Siberia. *Radiocarbon* 59(5):1381–97
- Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, et al. 2015. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* 528:499–503
- Matveev AK. 1962. Az ugor népek ősi letelepülési helyei. *Nyelvtud. Közlemények* 64:289–97
- Meillet A. 2005 (1908). *The Indo-European Dialects*. Tuscaloosa: Univ. Ala. Press
- Miestamo M, Tamm A, Gáspár K, Wagner-Nagy B, eds. 2015. *Negation in Uralic Languages*. Amsterdam: Benjamins
- Nichols J. 1997. Modeling ancient population structures and movement in linguistics. *Annu. Rev. Anthropol.* 26:359–84
- Nichols J. 2007. A typological geography for Indo-European. In *Proceedings of the 18th Annual UCLA Indo-European Conference*, ed. K Jones-Bley, ME Huld, A Della Volpe, MR Dexter, pp. 191–211. JIES Monogr. 53. Washington, DC: Inst. Study Man
- Nichols J. 2010. Proof of Dene-Yeniseian relatedness. In *The Dene-Yeniseian Connection*, ed. J Kari, BA Potter, pp. 266–78. Fairbanks: Alsk. Nativ. Lang. Cent.
- Nichols J. 2018. Non-linguistic conditions for causativization as a linguistic attractor. *Front. Psychol.* 8:2356
- Nichols J. 2019. Why is gender so complex? New typological considerations. In *Grammatical Gender and Linguistic Complexity*, ed. F Di Garbo, B Wälchli, pp. 63–92. Berlin: Lang. Sci.
- Nichols J. 2020a. Canonical complexity. In *The Complexities of Morphology*, ed. P Arkadiev, F Gardani, pp. 163–92. Oxford, UK: Oxford Univ. Press

- Nichols J. 2020b. For better and/or for worse: complexity and person hierarchies. In *Language Change for the Worse*, ed. D Enke, T Weber, G Seiler, L Hyman, J Nichols. Berlin: Lang. Sci. In press
- Nichols J, Rhodes RA. 2018. Vectors of language spread at the central steppe periphery: Finno-Ugric as catalyst language. In *Digging for Words*, ed. G Kroonen, R Iversen, pp. 58–68. Oxford, UK: BAR
- Nikolaeva I. 2014. *A Grammar of Tundra Nenets*. Berlin: De Gruyter Mouton
- Nordqvist K, Heyd V. 2020. The forgotten child of the wider Corded Ware family: the Fatyanovo culture in context. *Proc. Prehist. Soc.* 86:In press
- Outram AK, Stear NA, Bendrey R, Olsen S, Kasparov A, et al. 2009. The earliest horse harnessing and milking. *Science* 323:1332–35
- Pajusalu K, Uiboed K, Pomozi P, Németh E, Fehér T. 2018. Towards a phonological typology of Uralic languages. *J. Estonian Finno-Ugr. Linguist.* 9(1):187–207
- Parpola A. 2012. Formation of the Indo-European and Uralic (Finno-Ugric) language families in the light of archaeology: revised and integrated “total” correlations. See Grünthal & Kallio 2012, pp. 119–84
- Perşoiu A, Ionita M, Weiss H. 2019. Atmospheric blocking induced by the strengthened Siberian High led to drying in west Asia during the 4.2 ka BP event—a hypothesis. *Clim. Past* 15:781–93
- Rahkonen P. 2013. *Southeastern contact area of Finnic languages in the light of onomastics*. PhD Diss., Univ. Helsinki, Helsinki
- Rootsi S, Zhivotovskiy LA, Baldovič M, Kayser M, Kutuev IA, et al. 2007. A counterclockwise northern route of the Y-chromosome haplogroup N from Southeast Asia toward Europe. *Eur. J. Hum. Genet.* 15(2):204–11
- Saag L, Laneman M, Varul L, Malve M, Valk H, et al. 2019. The arrival of Siberian ancestry connecting the eastern Baltic to Uralic speakers further east. *Curr. Biol.* 29:1701–11.e16
- Saarikivi JS. 2020. Uralic intermediate protolanguages: a descendent reconstruction. See Bákró-Nagy et al. 2020. In press
- Salminen T. 2001. The rise of the Finno-Ugric language family. See Carpelan et al. 2001, pp. 385–96
- Sapir E. 1949 (1916). Time perspective in aboriginal American culture: a study in method. In *Selected Writings of Edward Sapir in Language, Culture, and Personality*, ed. DG Mandelbaum, pp. 389–467. Berkeley/Los Angeles: Univ. Calif. Press
- Simon Z. 2020. Urindogermanische Lehnwörter in den uralischen und finno-ugrischen Grundsprachen: eine Fata Morgana? *Indoger. Forsch.* In press
- Tambets K, Yunusbayev B, Hudjashov G, Ilumäe AM, Rootsi S, Metspalu M. 2018. Genes reveal traces of common recent demographic history for most of the Uralic-speaking populations. *Genome Biol.* 19:139
- Toldova SJ, Xolodilova MA, Tatevosov SG, Kashkin EV, Kozlov AA, et al., eds. 2018. *Èlementy mokšanskogo jazyka v tipologičeskom osveščanii*. Moscow: Buki Vedi
- Toporov VN, Trubačev ON. 1962. *Lingvističeskij analiz gidronimov verxnego podneprov'ja*. Moscow: AN SSSR
- Trudgill P. 2011. *Sociolinguistic Typology: Social Determinants of Linguistic Structure and Complexity*. Oxford, UK: Oxford Univ. Press
- Wilbur J. 2015. *A Grammar of Pite Saami*. Berlin: Lang. Sci.