

Reconstructing the Proto-Uralic Case System With Regard to Proto-Indo-European

Silobrčić, Ivan

Master's thesis / Diplomski rad

2023

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Humanities and Social Sciences / Sveučilište u Zagrebu, Filozofski fakultet**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:131:668918>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-07-18**



Sveučilište u Zagrebu
Filozofski fakultet
University of Zagreb
Faculty of Humanities
and Social Sciences

Repository / Repozitorij:

[ODRAZ - open repository of the University of Zagreb
Faculty of Humanities and Social Sciences](#)



University of Zagreb
Faculty of Social Sciences and Humanities
Department of Linguistics
Chair of Historical and Comparative Linguistics

Reconstructing the Proto-Uralic Case System With Regard to Proto-Indo-European

Master's thesis

Student: Ivan Silobrčić

Mentor: Ranko Matasović, Ph.D., full professor

Zagreb, 23rd of June 2023

Sveučilište u Zagrebu
Filozofski fakultet
Odsjek za lingvistiku
Katedra za poredbenu lingvistiku

Rekonstrukcija prauralskog padežnog sustava u odnosu na praindoeuropski

Diplomski rad

Student: Ivan Silobrčić

Mentor: dr. sc. Ranko Matasović, redoviti profesor

U Zagrebu, 23. lipnja 2023.

Table of contents:

The theoretical problems of reconstructing the Proto-Uralic case system	3
A brief etymological glossary of the Proto-Uralic cases	6
The grammatical cases	10
The nominative *- \emptyset	10
The accusative *-m	16
The genitive *-n	20
The directional cases	21
The locative *-nA	22
The separative *-tA	23
The ‘lative paradigm’	23
On Uralic cases not reconstructable into Proto-Uralic	25
The problems of the morphological reconstruction	30
Assessing how realistic reconstructing Proto-Uralic is, with regard to Proto-Indo-Uralic	33
Conclusion	38
References	40
Key-words / Ključne riječi	48
Summary / Sažetak	48
Kratak životopis	49
Appendix: Handout for the thesis defence	50

The theoretical problems of reconstructing the Proto-Uralic case system

Nominal morphology in the Uralic languages is word-final phonology and word-final syntax, and historical morphology is the archeology of word-final phonology and word-final syntax, puzzle-like as this field of study is. This thesis will attempt to give an up-to-date look at how the case system of the Uralic protolanguage is reconstructed by Uralicists, being a description of the problems involved in reconstructing the cases in a language family that is proverbially, if somewhat misleadingly thought of as being richly generative of cases. It will focus on the historical formation of the Proto-Uralic case system, both its phonological and its syntactic formations, viz. the phonology of the desinences and the question of clause alignment in Proto-Uralic, all with particular regard to Proto-Indo-European for two particular reasons. Proto-Indo-European is the most thoroughly reconstructed protolanguage and can thus offer its methodology into reconstructing the nominal morphology of other protolanguages; and, perhaps more pertinently, because an increasingly divisive debate is taking place around three or so schools of thought (or *tendencies* of thought) as to whether or not Proto-Indo-European and Proto-Uralic derive from an identifiable common protolanguage. As will be shown below, Uralicists who have researched this tend to be on the whole more than sceptical toward the Indo-Uralic hypothesis (e.g. Aikio, Holopainen, Parpola, Kallio, formerly also Janhunen), whereas Indo-Europeanists who have delved into this issue are generally in favour of it (e.g. Kortlandt, Kloekhorst, Pronk, Lubotsky). The Proto-Uralic cases will thus be reconstructed from both ends, from its descendants – the attested Uralic languages (using only very representative samples), and, more critically, from the presumed ancestor that is Proto-Indo-Uralic as offered by Kortlandt, Kümmel and Kroonen. Accordingly, we will be exploring both the morphological reconstruction of the Proto-Uralic morphemes, and the morphosyntactic reconstruction; the first being the cause of the second, as a new language forms as a result of a confluence of phonological innovations.

Uralic, despite being the second-most studied language family of all (Aikio 2022: 3), probably lacks this tipping point in terms of quantity (to say nothing of quality) of its researchers; this is best seen by the influence that Luobbal Sámmol Sámmol Ánte (Ante Aikio) of the Sámi University of Applied Sciences in Kautokeino, Finnmark, Norway has had on Uralic studies. Hence, we will be comparing sources prior to his work, but his chapter (chapter 1: Proto-Uralic) in the *Oxford Guide to the Uralic Languages* (2022, ed. Marianne Bakró-Nagy, Johanna Laakso, Elena Skribnik), alongside his published and unpublished work will be our principal sources. He has been praised for breathing new air into the study

of Proto-Uralic, as someone who has been able to make very critical use of the seminal works of Proto-Uralic studies which, together, form the general doctrine upon which Proto-Uralic studies rest¹. A rather comprehensive reworking of Uralic studies had been ongoing for the last 40 or so years, or, even less than that (Aikio 2022: 3).

There are differing views regarding how real the reconstructed forms are – we cannot know empirically something unless we check it empirically – which is most often impossible in linguistic reconstruction, almost inevitably, in spite of historical linguists’ (and archeologists’) best efforts. Whether they are mere abstractions which account for the attested forms and nothing more, or real forms, is a matter of approach to the problem (cf. Nehring 1961). All unattested forms are abstractions if we trust the empirical sciences (although the debates surrounding how to reconstruct those forms are proof enough that they are questionable, especially in Uralic, such as in the gap between Rédei’s Proto-Uralic and Aikio’s Proto-Uralic, see below). We will be assuming that they are abstractions; when referring to the *lexical* reflexes of Proto-Uralic, we will usually restrain ourselves to citing Hungarian and Finnish (Holopainen 2020: 296 rightly finds this to be precisely the problem in new attempts are reconstructing Proto-Uralic, viz. the unavailability, or poor availability of most other Uralic languages) for the sake of simplicity. They are the best studied Uralic languages, and few linguists, save for dedicated Uralicists, show much knowledge of most other Uralic languages, as can even be seen by the more than vague nomenclature given to them: Sámi isn’t even remotely *one* language, and neither is Khanty. Which is why, when citing examples from other sources, we will regularly be checking all of them up at the Uralic Etymological Database, or Uralonet, which is run by the Institute of Linguistics (*Nyelvtudományi Kutatóközpont*) of the Hungarian Academy of Sciences, which is based off of the material gathered in Rédei’s *Uralisches Etymologisches Wörterbuch* (1986–1988). A much improved *Uralic Etymological Dictionary* (forthcoming) is being prepared by Ante Aikio, who has published online a draft version of the dictionary in 2020, covering the

¹ These principal works are Rédei’s *Uralisches Etymologisches Wörterbuch* (1986–1988, in seven slim volumes), Sammallahti’s chapter *Historical Phonology of the Uralic Languages* and Raun’s chapter *Proto-Uralic Comparative Historical Morphosyntax* (alongside other chapters in the collection *The Uralic Languages: Description, History and Foreign Influences* 1988, ed. Denis Sinor), as well as the severely outdated but useful *Comparative Grammar of the Uralic Languages* (1960) by Collinder (who freely admits in his introduction that his interpretations as to how he reconstructs Proto-Uralic are *more than* open to questioning, being the first such systematic attempt), and Hajdú’s short but useful synoptic paper *The Main Characteristic Features of the Uralic Languages* (1983).

Proto-Uralic lemmas beginning with vowels and with **č*²; it is only the second such attempt, following and heavily correcting Rédei's *Uralisches Etymologisches Wörterbuch*.

In this thesis, we will be simplifying as well, like everyone else, all in an attempt to form a synoptic view on the state of how the case system of the Uralic protolanguage stands, in its two-sided reconstruction: its morphonological and in its morphosyntactic sides to the figurative coin that is its case system. The stages of reconstruction are, simply: setting up correspondences, establishing proto-phonemes and assigning them phonetic values, and then proto-morphemes, and assigning them grammatical values, a set of tasks performed by multiple researchers with sometimes identical results, and sometimes completely different ones (cf. the section on the 'lative paradigm' below). This is why we will, after a brief etymological glossary of the reconstructed case morphemes, be discussing the state of research of Proto-Uralic phonology, as well as the relevance of the Indo-Uralic hypothesis (mentioning, too, the homeland problem as well as potential lexical correspondences which might serve as a confirmation of the hypothesis). Morphology comes in three parts: sound (form, realisation), grammar (syntactic function) and semantics (what it conveys, its meaning), all of which are more than open to shifts and readjustments in a reconstructed (unattested) system which we necessarily cannot know entirely.

We will not adopt one particular perspective, but rather compare those which have been explored and developed, against linguistic trends which come and go; it is an increasingly popular academic activity to criticise the Leiden school of thought, but this criticism ought to be applied more generally, especially to any such schools of thought, in particular to a debate as lively as the one around the Indo-Uralic hypothesis, and within Uralic studies themselves. The comparative method is the primary method of reconstruction, aided by internal reconstruction, the application of linguistic universals, language typology, all of which we will be mentioning. Simply put, there is no doxa in Indo-European linguistics, much less so in Uralic linguistics; theories, not doctrines abound in historical linguistics³.

² Available on the following link:

https://www.academia.edu/41659514/URALIC_ETYMOLOGICAL_DICTIONARY_draft_version_of_entries_A_Č.

³ For the sake of additional clarity, we will also be colour-coding the reconstructed morphemes and lemmas thusly: Proto-Uralic **k*, Proto-Indo-European **s*, Proto-Indo-Uralic **m*. Note, too, that the smaller Uralic languages, as well as all of its branches, are known for having a plethora of names – this stems from conflicting traditions in nomenclature, as well as in classifying these branches. Thus the older, especially German-language literature talks of a Cheremis language, today usually referred to as Mari. Similarly (in the German, Russian and English literature) Ostyak > Khanty, Vogul > Mansi, Lapp > Sámi, Zyryan > Komi.

A brief etymological glossary of the Proto-Uralic cases

In this section, we propose a brief etymological glossary of the reconstructable grammatical cases in Proto-Uralic. Each case will be analysed from a representative sample of its reflexes in the attested Uralic languages, as well looking into the matter of to what extent they can be related to the Proto-Indo-European case system, and, consequently, to Proto-Indo-Uralic. Proto-Uralic as currently reconstructed has the same problems as reconstructing Proto-Indo-European, viz. that multiple strata of the language are reconstructed at once which have, so far, only partially being combed out of each other (Hyllested 2009: 112), presupposing that the reconstructed language truly is the form which split off into the daughter languages, and not a scientific abstraction. Pre-Proto-Uralic (or Early and Late Proto-Uralic), unlike Pre-Proto-Indo-European, does not seem to have been noticed much yet, seldom mentioned as it is in the relevant literature.

All the attested Uralic languages have case, number as well as possessors on their nominals, and all of them could be zero-suffixed in Proto-Uralic (Abondolo 2006a: 18, Aikio 2022: 14). Most Uralic languages have a three-way opposition to spatial case systems, viz. stasis (or true *locative*) as opposed to motion, whose cases denoted motion away from (various *separatives*) and motion towards (various *latives*) (Abondolo 2006a: 18-19, Aikio 2022: 14), with Aikio reconstructing two possible additional “adverbial cases” into Proto-Uralic (Aikio 2022: 14). The attested (which in practice mostly means *contemporary*, see below) Uralic case systems have much in common, especially as regards the common to most of them cases. This makes reconstruction feasible, after having established the sound correspondences between the Uralic languages and branches, as well as their relative chronology – in spite of the lack of older attestations, which would shed much needed light on how Proto-Uralic made use of adverbial cases, irrelevant as that seems to be to the Indo-Uralic hypothesis. Prepositions are rare in Uralic, appearing only in Finnish, Sámi and Estonian, and even there they are rare (Hajdú 1983: 105); cases are agglutinated alongside other postpositions in Proto-Uralic, viz. they did not exclude other postpositions (Abondolo 2006a: 21). Agreement between head nouns and its adjuncts is rare; the Uralic languages that have agreement are on the western edges of the language family (such as Finnish), viz. in contact with Indo-European languages which do have agreement (Matasović 2008: 38). Thus Finnish repeats and rhymes nominal groups, as in *tuo-hon sama-an iso-on rakennukse-en that-ILL same-ILL big-ILL building-ILL* ‘to that same old building’, with all elements in this

nominal group in the illative case (Abondolo 2006a: 32). The basic synchronic Uralic case inventory *minimum* is three, in Hungarian⁴, Mari and Ob-Ugrian, more concretely in North Khanty (Abondolo 2006a: 23). As discussed below, many grammatical and adverbial cases can be attributed to the simplest form of casogenesis (or *Kasusursprung*) – that is to say, in Uralic, postpositions becoming enclitical in an agglutinative language. One area where the otherwise seldom used oldest Hungarian corpus has been of use has been in showing that many non-grammatical cases, written separately in the Old Hungarian corpus, appeared as unfused postpositions (Kittilä, Laakso, Ylikoski 2022: 881), thus showing that many Uralic cases are of relatively recent origin (cf. below on language contact between Uralic and Tocharian and Old Lithuanian). Cases, when they are affixes (and not tones as in some Nilotic languages) are adpositions (usually postpositions, probably for phonemic reasons) which have successfully removed word boundaries, thus becoming morphemes (Kulikov 2009: 440), viz. lexemes which reduce their semantic field by becoming bound morphemes, losing their phonetic stress even, thus becoming moveable parts of words (Heine 2009: 459 et passim). For example, the Proto-Balto-Finnic lexeme *kansa ‘people’ becoming the Estonian comitative-instrumental case suffix *-ga/-ka* ‘with’, as in *sõbra-ga* ‘with (a) friend’ (idem.). One of the most contentious parts, as discussed below, in reconstructing Proto-Uralic is the vocalic inventory of Proto-Uralic which, consequently, makes unclear how exactly the noun stems alternate in line with the vocalic alterations, as required by vowel harmony, of the case suffixes (Hajdú 1983: 105); thus more or less only the consonants of these monosyllabic (thereby monophonemic) case endings can be reconstructed.

The traditional conceptualisation of the Uralic branches has been that of dividing a Finno-Ugric branch from a Samoyed one, this split having been the dominant model for branching out the Uralic languages from Proto-Uralic, and, consequently, for reconstructing Proto-Uralic; fundamental disagreements regarding Uralic taxonomy exist, and there seems to be no consensus on this whatsoever (only the traditional doxa of Finno-Ugric vs. Samoyed) (Aikio 2022: 3). Since Häkkinen 2007, he and many others have leaned toward the split being around a Finno-Permic branch and a Ugro-Samoyedic one⁵. The issue here being

⁴ Cf. Anhava’s 2010 paper on the traditional and modern criteria used by Finnish and Hungarian grammarians in determining the number of cases in Finnish and Hungarian: the former is traditionally thought to have a paradigm of 15 cases, with a handful of non-productive cases, viz. cases which are not applied to adjectives as well in agreement (Anhava 2010: 239), whereas the latter has 17 or 18 cases if the strictest criteria are applied, and up to 28 if one removes more or less any restrictions (ibid. 243), or even fewer, as will be shown below, if one considers only so-called grammatical (non-adverbial) cases to be true cases.

⁵ For instance, PU ?**čiḡä-mə* (Holopainen 2020: 297) (cf. Aikio: **čV(w)ḡä(-mi)*) ‘heart’ > Finnish *sydän* ‘ibid.’, Hungarian *szív* ‘ibid.’, shows that Finnish and Hungarian developed an initial s- parallel to each other, but this

that the taxonomy of the primary branches of Uralic, much like the taxonomy of the proto-daughter-languages of PIE, has yet to be ordered (Hajdú 1983: 101; Aikio 2022: 3), which makes all the difference in reconstructing Proto-Uralic (and Proto-Indo-European) (for a summary, see Saarikivi 2021, and the often cited Aikio 2012 paper) as well as carefully unweaving areal contact from genetic relations (Pereltsvaig warns of this in Pereltsvaig 2012: 210, as Russian may have influenced Finnish through Merya, Muroma and Meshchera (all extinct Uralic languages), rather than assuming any Proto-Indo-Uralic characteristics too hastily). Thus Aikio refutes the traditional Finno-Ugric vs. Samoyed split which, as recently as 2006, Abondolo claimed there was universal agreement on, that the first node after Proto-Uralic was Samoyedic vs. Finno-Ugric (Abondolo 2006a: 1). This issue seems to have come from exaggerating the phonological deviance of Samoyed (Aikio 2022: 4), underestimating both the substrate influence in Hungarian, and the discovery that Ugric (Hungarian, Khanty, Mansi) is core Uralic – Abondolo is of the opinion that Proto-Ugric seems to be the core branch, which ‘split off’ last (or, rather, didn’t split off, having no ‘offshoots’) (Abondolo 2006a: 3, citing Viitso 1996), as Balto-Slavic is for Proto-Indo-European. Researchers have, trailing on that idea that Ugric is the core of Proto-Uralic, since proposed that the first node was the Finno-Permic vs. Ugro-Samoyedic split. This is almost impossible to ascertain; the corpus material, imprecise as it is, incomplete (as is the case in describing the Uralic languages beside the three largest ones) as it is, is a hurdle difficult to deal with; reconstruction is a very conjectural, formal science as the number of facts at our disposal is limited, and the most we can do to augment the corpus of facts is applying the historical method as best we can (or wait for archeologists to find ancient inscriptions). All languages (not just Uralic) are insufficiently studied and, in the case of the Uralic languages, show an astonishing degree of synchronic variation; the proto-languages we reconstruct must be rigorous logical deductions, but, given how hotly debated they are, evidently do not live up to their desired standards. Indo-European studies have the distinct advantage of being studied far and wide, and a sufficiently large number of researchers are able to correct one another and constructively engage in academic dialogue, alongside having a plethora of ancient attestations⁶. But they show just how widely different results one can

and similar phenomena do not, despite appearances, imply a Finno-Ugric protolanguage (Holopainen 2020: 297; cf. Katičić 1970: 138-139 on how to distinguish between common innovation and parallel innovation).

⁶ Which the Uralic languages distinctly lack (a single attempt has been made to make use of the oldest Uralic material, in Lühr 2019, where she compares headedness in Hittite, Vedic and Old Hungarian corpuses, in spite of the powerful substrate influence in all three languages). Briefly, the oldest Uralic corpus material consists of a single incomplete sentence (and longest fragment) in Old Hungarian from the establishing charter of the abbey at Tihany on Lake Balaton in Hungary (1055, held at Pannonhalma), reading: “*feheruuaru rea meneh hodu utu rea*”, meaning “up to the military road going to Fehérvár”, and the birch bark letter n. 292 in some archaic form

get, depending on the approach one takes in organising the daughter languages⁷. Aikio, who is currently writing the *Uralic Etymological Dictionary*, for lack of any convincing evidence in organising these branches, takes the academically most honest position of considering the nine branches all to be co-equal branches splitting from Proto-Uralic (Aikio 2022: 3, 4), namely: Finnic (Baltic Finnic)⁸, Sámi (North Finnic)⁹, Mordvin (West Finnic)¹⁰, Mari (separate Finnic)¹¹, Permic¹², Mansi/Vogul (Ob-Ugric)¹³, Khanty/Ostyak (Ob-Ugric), Hungarian (Ugric)¹⁴ and Samoyed¹⁵ (Glottolog, Ethnologue).

of Karelian (beginning of the 13th century), reading: “юмолануолиінимижѝ / ноулиствханолиомобоу / юмоласоудънихоховѝ”, with various readings possible. The longest non-fragmentary text in a Uralic language is in Old Hungarian, the *Halotti beszéd és könyörgés*, or funeral sermon and prayer translated from a Latin original, from around 1200 AD (Raun 1988: 555). Other, non-contemporary Uralic texts are a Hungarian poem from around 1300, *Ómagyar Mária-siralom*, or lamentations of Mary, followed by a few texts in Old Permic (an old form of Komi) as used by Saint Stephen of Perm in the 14th century. Finnish and Estonian Bible translations were produced in the 16th century, with Mikael Agricola’s Reformation-era Bible translation into Finnish from the 1540’s being of note. Sámi texts first appear in the 17th century, and most other Uralic languages become to a greater or lesser extent codified, and appear in written form in the 19th, mostly in the 20th century (). The main databases (in terms of corpus material and in terms of lexicographical material) for the few old Uralic texts are, inevitably, those that regard the Hungarian and Finnish languages. The former is covered by the *Nyelvtudományi Kutatóközpont* (Hungarian Research Centre for Linguistics) (available online on the following link: <http://oldhungariancorpus.nytud.hu/en-fragments.html>), the two sections *Old Hungarian Codices* and *Miscellaneous Minor Texts*, and the latter is covered by the *Kotimaisten kielten keskus* (Institute for the Languages of Finland) (available online on the following link: https://www.kotus.fi/aineistot/tietoa_aineistoista/sahkoiset_aineistot_kootusti), with such dictionaries as the Dictionary of Old Literary Finnish (*Vanhan kirjasuomen sanakirja*) (<https://kaino.kotus.fi/vks/>), to be consulted with the Juntilla, Holopainen, Pystynen 2020.

⁷ For instance, Schleicher’s *Stammbaumtheorie* organises by descendants, and Schmidt’s *Wellentheorie* by isoglosses (viz. linguistic geography), which in many instances more precise than the former method, as Lehmann (1996: 24-25) explains the pitfalls of conceptualising comparative linguistics using comparative anatomy, and language families using biological metaphors (Katičić 1970: 138 *et passim* mentions how the Iranian aspirated stops point to different directions using those two methods: the tree model leads one to Indo-Iranian, but the wave model leads one to Irano-Slavo-Balto-Celtic, as they all treat their aspirated stops the same way). For instance, Karlsson, in his brief descriptive grammar of the Finnish language, shows the relative synchronic similarity to Finnish from the perspective of the other Uralic languages in the following concentric manner (Karlsson 1999: 1): Finnish < Estonian, Karelian, Vepsian, Ludian, Votian, Livonian < Sámi < Mordvin, Mari < Komi, Udmurt < Hanti, Mansi, Hungarian.

⁸ Consisting of Estonian ((*Northern*) Estonian) and Southern Estonian (*Seto*, *Võro*(, *Ugala*)) as well as Finnish (and its dialects and the Old Finnish corpus), *Veps*, *Ingrian*, *Votic*, *Livonian*, *Eastern Kven*, *Western Kven*, and *Karelian* (and the Old Karelian birch bark letter n. 292), *Ludic* and *Olonets Karelian*.

⁹ Divided into West Sámi, which includes Southwest Sámi (*Southern Sámi*, *Ume Sami*) and Northwest Sámi (*Lule Sámi*, *Northern Sámi*, *Pite Sámi*), and East Sámi, which includes Peninsular Sámi (*Kildin Sámi*, *Ter Sámi*) and Mainland Sámi (*Akkala Sámi*[†], *Inari Sámi*, *Kainu Sámi*[†], *Kemi Sámi*[†], *Skolt Sámi*).

¹⁰ That is to say *Mordvin* (*Erzya* and *Moksha*) and, possibly, *Marya/Merya*[†], *Meshchera*[†] and *Muroma*[†] languages.

¹¹ Of which there are two variants, *Meadow Mari* and *Hill Mari*.

¹² Consisting of *Udmurt*, *Komi-Permyak*, *Komi-Yodzyak* and *Komi-Zyrian* as well as *Old Permic*, an old form of Komi.

¹³ That is to say *Eastern Mansi* (*Vogul*), *Western Mansi* (*Vogul*), *Southern Mansi*[†] and *Northern Mansi*[†].

¹⁴ As well as its dialects and the Old Hungarian corpus.

¹⁵ Consisting of North Samoyedic (*Forest Nenets*, *Tundra Nenets*, *Enets* and *Nganasan*) and South Samoyedic (*Ket Selkup*, *Taz Selkup* and *Tym Selkup*) as well as Sayan Samoyedic (*Kamas*[†], *Karagas*[†], *Koibal*[†], *Motor/Mator*[†], *Soyot*[†] and *Taigi*[†]).

The grammatical cases

The nominative *-∅

It is agreed that the Proto-Uralic nominative case was zero-marked, noted as *-∅ (Collinder 1960: 282, Aikio 2022: 14), as is typologically usually the case for nominative cases (Proto-Indo-European being an exception, cf. the section on ergativity below) and therefore called more precisely an *absolutive* case (Abondolo 2006a: 18). Raun (1988: 558) proposes that an Ur-case, an *absolutus* as he terms it, may have preceded the Proto-Uralic case system, and that this *absolutus* could express “any case-like function”, as well as designating both the subject and the indefinite object, later branching out its functions into “up to six” morphemic cases, of which three grammatical and three local, that is to say nominative, accusative, and genitive, and locative, ablative (here, below, called the separative) and lative which will be the subjects of the following pages. The Proto-Uralic plural marker *-t is agglutinated to this zero-morpheme nominative, and a dual marker can tentatively also be reconstructed for this zero-marked nominative, namely *-k(V) (Aikio 2022: 14).

The problem of whether Proto-Uralic was a nominative-accusative language or an ergative-absolutive one is key to linking it to Proto-Indo-European. Aikio (2022: 21), and many before him regard Proto-Uralic as having been almost certainly nominative-accusative, given that all of the attested descendant languages are – with the possible sole exception of East Khanty, which is capable of a quasi-ergative construction, but this is in all likelihood an innovation (Aikio 2022: 21; Matasović 2008: 22). Many have also noted that there is exceptionally little understanding of the Uralic, and of the Proto-Uralic sentence (Abondolo 2006a: 33, mentioning Koizumi 1994; Aikio 2022: 21). Aikio assesses de Smit’s 2015 attempt to find traces of an earlier ergative system in Proto-Uralic to be “far from conclusive”, but allows the *possibility* of Pre-Proto-Uralic as having been an ergative language, especially in light of other factors which link Proto-Uralic to Proto-Indo-European. The problem here being that the oldest well-attested Indo-Uralic languages, that is to say Vedic, Hittite and Old Hungarian all have powerful substrate influences (and all three are used as examples for research in headedness in Lühr 2019); it is to be noted that, in assessing how Proto-Uralic and Proto-Indo-European might align in terms of clause-alignment, one must be weary of pitfalls in the historical method, such as a phenomenon noted as Sanskritophilia, and an overemphasis on data from Hittite, seeing as substrate influence may

be decisive in this matter.

Morphology is what links Proto-Indo-European and Proto-Uralic (Pereltsvaig, Lewis 2015: 200)¹⁶, not just the vocabulary, whose status is highly contested, and much less the phonology (see below). Proto-Indo-European, like Proto-Uralic, was a consequently suffixing language, and so are many other Eurasian languages, such as those belonging to the Micro-Altai languages, whose common relationship is highly contested (Matasović 2008: 6) but, the PIE, unlike PU, was fusional (Matasović 2008: 7), which is suggestive of a case system of a longer duration, which has had the time to fuse the morphemes for cases, gender and number. Both Proto-Indo-European and Proto-Uralic are reconstructed as being SOV languages¹⁷ (Aikio 2022: 21; Raun 1988: 569), and SOV languages are generally postpositional, as, again, both Proto-Indo-European and Proto-Uralic are reconstructed as being (Lühr 2019: 163); bafflingly little else can be said of Proto-Uralic syntax, as “much of what has been climbed regarding Proto-Uralic syntax is not based on any systematic research at all, but instead on educated guesswork or even speculation” (Aikio 2022: 21).

Proto-Uralic, having no gender, also had a separate morpheme for the plural, **-t* (and **-j* for the accusative and genitive cases (Aikio 2022: 14), and seemingly none for the directional cases) which was agglutinated to the case morpheme. The origin of this plural suffix in Proto-Uralic may be parallel to the origin of the PIE plural suffix, **-es*, namely that the Proto-Uralic plural marker **-t* < PU second person marker **-t*, and the PIE plural marker **-es* < PIE second person singular marker **-s*, both stemming, presumably, from a common morpheme semantically marking a non-first-person, and giving the PIE form by way of assibilation (Pereltsvaig, Lewis 2015: 200). The PIE neuter nominative plural **-ŋ* (**-h₂*) would then be a later, separate development (Kortlandt 2020b: 3) as, per Kortlandt in his very panoptic view of the matter, the case endings in PIE largely formed after the PIE and Anatolian branches split, as their common ancestor (Proto-Indo-Anatolian)¹⁸ lacked reconstructable genitive, dative or oblique plural endings (thus having unmarked plurals). Hittite partly preserved this state of affairs as there is a singular-plural genitive ending, and is

¹⁶ Listing the following morphemes (with a more thorough list below by Kortlandt): the pronominal roots: **m-* = the first person, **t-* = the second person, **i-* = the third person; the case markings: **-m* = accusative and **-ta* = ablative/partitive; the negative particle **ne*; the interrogative pronoun **k^w-* ‘who?, which?’

¹⁷ For instance, Hungarian is not a SOV language, but a Topic-Focus-Verb language (with the topic being optional), stemming from a Proto-Hungarian SOV (which is maintained in the other Ugric languages, viz. Khanty and Mansi) (Lühr 2019: 163), viz. Proto-Hungarian subject > Old Hungarian topic, Proto-Hungarian object > Old Hungarian focus (Lühr 2019: 179).

¹⁸ It has been noted early on (1921) that Anatolian was not an Indo-European daughter language, but a sister language, thus positing Indo-Anatolian as being the proto-language because of a number of features which Hittite did not have, and the IE languages do: the feminine, the aorist and the perfect, for instance (Kloekhorst, Pronk 2019: 2).

thus reconstructable as such (Kortlandt 2010: 40). Late Proto-Indo-European (the hypothetical/abstract final day before no longer being Proto-Indo-European) cases must have been the nominative, vocative, accusative, genitive, ablative, dative, locative and instrumental; presumably a *simpler* case system existed in Early PIE, which would then be able to fan out into *more* cases (mostly of prepositional origin, particularly as regards the plural forms¹⁹), in Late Proto-Indo-European (Matasović 2008: 8; for a highly detailed venture into this, see Kloekhorst 2023)²⁰.

Grammatical gender is semantically meaningless – it is grammar in pure form, a marker to be followed around when reconstructing; agreement is a marker which interconnects clauses and/or nominal or other groups in a sentence, and across sentences. The Uralic languages lack gender, with Selkup being the sole exception (Collinder 1960: 242); thus Proto-Uralic is taken to have no gender which, moreover, is a typologically rare phenomenon²¹. Studying gender is a complex matter, as inflexional languages tend to fuse gender with other categories, as in Latin: *lūp-ī* ‘wolves’, where the *-ī* morpheme simultaneously denotes gender, number and case, that is to say the masculine, plural and nominative (Matasović 2004: 21). The feminine derives from “very old abstractions and inanimates of the neuter gender”, as in the nominative-accusative dual neuter consonant stems, e.g. PIE **h₃ek^w-ih₁* ‘two eyes’ (Ancient Greek ὄσσε, Old Church Slavonic *очу*), the same ending being the nominative-accusative dual feminine **-eh₂*, as in PIE **k^woyneh₂-ih₁* (Old Church Slavonic *цѣнь* ‘two prices’), which is not the same as e.g. PIE **pod-ih₁* ‘two feet’ (Ancient Greek πόδε), as the PIE **-ih₁* = common dual ending **-h₁*, *agglutinated* to the zero-grade collective suffix of PIE **-ēy-*, which is PIE **-i-* (Matasović 2004: 166). Gender in Proto-Indo-European gave rise to noun classes, in effect declension classes (morphologically different from one another), which are apparently absent in PU, which had nothing even resembling gender, nor noun classes. Number is a category in Proto-Uralic which, it would seem, sheds little light on reconstructing the cases. Number is generally not expressed on the

¹⁹ Some case endings in PIE are reconstructible as being of prepositional origin, such as (for instance, in the *o*-stem paradigm) dat. pl. **-b^hos/*-mos*, loc. pl. **-su* and instr. pl. **-b^his* as well as dat. du. **-b^h(y)oh₁/?**-moh₁*, abl. du. **-b^h(y)oh₁/?**-moh₁* and instr. du. **-b^h(y)oh₁/?**-moh₁*.***

²⁰ Indo-Anatolian is deemed to be all but proven, cf. the list of features which, in fact, only take time to prove, semantic, morphological, sound and syntactic changes in Anatolian or Hittite (Kloekhorst, Pronk 2019: 4-5, 6), with Tocharian splitting off first from this Anatolian-less Proto-Indo-European (Peyrot 2019b: 188), Peyrot calling this non-Anatolian PIE Indo-Tocharian; the nod after the Tocharian split he calls Indo-Italo-Celtic.

²¹ Many North-East Caucasian languages, much like the Bantu languages, have vast noun classes or genders (and nominal classifiers in many East Asian languages constitute a para-gender class), as do the vast majority of the Indo-European languages (save for Armenian, Persian and some Western Iranian languages); in fact, many Northern Eurasian languages lack gender (Basque, Uralic, Tungusic, Turkic, Mongolic, Chukotko-Kamchatkan, Korean-Japanese-Okinawan, Ainu, Sino-Tibetan), save for Ket, a Yenisean language, which does have it (Matasović 2004: 20, 25, 26).

noun in the Uralic languages if there is a numeral, except in Finnish where it is with the partitive case (Collinder 1960: 235-236); in Hungarian linguistics, this is known as *numerus absolutus* (Raun 1988: 556). Other than this instance, two nominal plural morphemes can be reconstructed in PU, but their relationship is unclear (the above-mentioned *-t and *-j morphemes) (Aikio 2022: 15).

Clause alignment in Proto-Indo-European is suggestive of ergativity, and gender in Proto-Indo-European was related to marking ergativity (Matasović 2004: 18). If a language morphologically treats both its subjects of transitive and intransitive verbs alike, but distinctly from its direct objects of transitive verbs, it is a language with accusative clause agreement – for instance, Latin, where the subject is marked as a nominative, and the object is marked as an accusative; most Indo-European languages belong to this category, as do the Uralic languages, be it with an accusative case, or a different oblique case, such as a partitive or genitive. The ergative clause alignment, by contrast, is when a language morphologically treats the subjects of its transitive verbs one way, and the subjects of its intransitive verbs identically to its direct objects – such as the isolate Basque, as well as many other Eurasian isolates, or many African languages; the former category is marked as ergative, the latter as absolutive^{22,23} (Matasović 2004: 181; Matasović 2000: 329). Moreover, clause alignment does not necessitate case-marking; a language can be ergative and caseless (Matasović 2000: 329-330). Early Proto-Indo-European is reconstructed with the nominative marker *-s, the sigmatic nominative, denoting agents; neuter nouns cannot be agents in Hittite, so, in Early Proto-Indo-European too, they could not be marked with this sigmatic nominative. Therefore, this PIE *-s denoted transitive verb subjects, and *-∅ (zero-marking) denoted intransitive verb subjects and objects – which is a description of an ergative language. Furthermore, marking nominatives with anything but a zero-morpheme is itself suggestive of a non-nominative-accusative clause-aligning language (Matasović 2004: 182). As Kortlandt

²² Put schematically:

nom.-acc. c. a.: subj. tr. verb = subj. intr. verb ≠ dir. obj. tr. verb

erg.-abs. c. a.: subj. tr. verb ≠ subj. intr. verb = dir. obj. tr. verb

sta.-act. c. a.: subj. tr. verb (if agent) = subj. intr. verb (if agent) ≠ subj. (which is not an agent) & obj. (which is not an agent)

²³ There also exists the stative-active clause alignment, wherein *some* intransitive verbs appear in an ergative construction; many languages in the Americas and on New Guinea are of this sort. It is deemed unlikely that PIE was actively-aligned (Matasović 2000: 332-333); pairs of semantically identical but gender-different words need not be evidence of an active clause system (agent and patient pairs), as has been claimed for PIE, as one lemma in this pair may simply be an archaism with respect to the other, or of a difference in register (as, for instance, in Croatian *oganj* and *vatra*) (Matasović 2004: 186). All three clause-alignment possibilities have been argued for Proto-Indo-European – even the stative-active clause alignment, by Gamkrelidze and Ivanov, which is supported by no evidence (Matasović *Clause Alignment in Proto-Indo-European*: 2-7), and more than one of these models can coexist in a single language.

has brought attention to, already Christianus Cornelius Uhlenbeck (*Agens und Patiens im Kasussystem den indogermanischen Sprachen*, 1901) noted that the nominative neuter and the accusative neuter were treated identically in PIE, both in the singular and in the plural, which resembles an absolutive case (or *Passivus* in Uhlenbeck's paper), which corresponded to the PIE bare-stem nouns (save for the athematic PIE *o*-stem nouns), and, on the other hand, the subject of transitive verbs (or *Aktivus* in the same paper) was marked by PIE **-s*, which is reminiscent of an ergative case (Kortlandt 2010: 33). Thus ergative PIE **-os* > the nominal *o*-stem nouns, as well as the accusative singular PIE **-om* > nominative singular in neuter *o*-stem nouns, which also gave the genitive plural PIE **-om* (Kortlandt 2020b: 3).

All in all therefore, nom. and gen. sg. PIE **-s* denoted the ergative in Pre-Proto-Indo-European (Matasović 2012: 15; also in Matasović 2000: 330). As for masc. and fem. inanimate nouns in Proto-Indo-European (e.g. **pōds* 'foot'), the old ergative became the new nominative, and as for neuter nouns (e.g. **h₃nom_ŋ* 'name'), the old absolutive became the new nominative-accusative, viz. all of these nouns had at one point ergative case-inflection (Matasović 2012: 15-16). The original ergative would then be preserved in the gen. sg. PIE **-s* (Matasović 2012: 16, admitting that all of this is conjectural), as, for instance, in Modern Greek gen. sg. *ονόματο-ς* 'of (a) name'. Nonetheless, the ergative hypothesis for Proto-Indo-European has the advantage of not being based on putative typological universals (Matasović *Clause Alignment in Proto-Indo-European* 7), instead connecting independently reconstructed facts about PIE, namely: that Proto-Indo-European has a nominative case marker (unlike almost all other Eurasian languages; few languages have ever been found to do this); that there is a vocative form (PIE **-e*) which is non-identical to the nominative; that some inanimate nouns are gendered instead of being neuter; that the nom. sg. and the gen. sg. are similar, that gen. sg. and abl. sg. are syncretic except in thematic nouns; that *only o*-stem and static-stem nouns are nouns where, when gendered, are animate, and when neuter, are inanimate; that the case markers for personal pronouns are noticeably different from the case markers for nouns, therefore not directly derived from them; that Proto-Indo-European verbs have two separate sets of personal markers (per categories); that the 3rd person singular personal marker of the verbal categories is similar to the nom.-acc. neuter demonstrative pronouns; and that the perfect and thematic present endings evidently hark back to an Early Proto-Indo-European prototype which, taken together, point to ergativity in Proto-Indo-European being typologically possible (Matasović *Clause Alignment in Proto-Indo-European*: 7-16, 25). Proto-Uralic was in all likelihood not ergative (Aikio 2022: 21), and with Proto-Indo-Uralic being even less likely an

ergative language, Proto-Indo-European could have something to do with intense language contact with NW Caucasian languages (Matasović *Clause Alignment in Proto-Indo-European*: 26-29), which archaeological evidence does seem to support, despite heavy criticism of this methodology (Campbell 2015; see below).

Matasović thus reconstructs the following order in which an ergative Early Proto-Indo-European may have become a nominative Late PIE. First (stage I), the plurals were unmarked (as in much of northern Eurasia), animate nouns were inflected, that is to say inanimates had a collective plural (PIE $*-h_2$, as in Sanskrit युगानि (*yugā-ni*), Ancient Greek ζυγά). There were two grammatical cases, viz. the sigmatic ergative $*-s$ (which served also as a genitive and an ablative) and the unmarked absolutive $*-\emptyset$, with two declension types (one with no accent shift, as in the future PIE $*ph_2tér-\emptyset$ ‘father’, $*ph_2tér-s$, and the other with an accent shift, as in PIE $*h_3nóm̥n̥-\emptyset$ ‘name’, $*h_3n̥mén-s$, this latter variant being derived from the ergative/absolutive form). Also, no gender and probably no case and number agreement can be reconstructed at this earliest stage, just like in Uralic, Altaic and Yukaghir, viz. Micro-Altaic – thus two cases, the ergative/genitive-ablative $*-s$ (as in what would become PIE $*h_3n̥men-s$, $*h_2ner-s$ ‘man’) and the absolutive $*-\emptyset$ ($*h_3n̥men-\emptyset$, $*h_2ner-\emptyset$), and plural nominative-accusative $*-es$ ($*h_3n̥men-h_2$, $*h_2ner-es$), much like in Proto-Uralic (Matasović *Clause Alignment in Proto-Indo-European* 8-10).

The following stage (stage II) would consist of an absolutive $*-\emptyset$ agglutinated to this ergative/genitive-ablative $*-s$ morpheme, which would give rise to the possessive genitive $*-s$ ($*-ós$, as in what would become PIE $*ph_2trós$), and then the ergative and absolutive would no longer be functional (Matasović *Clause Alignment in Proto-Indo-European*: 11-12). This (stage III) would be followed by the nominative and accusative being introduced, possibly by adstrate or substrate, or even some former personal pronouns; thus the absolutive $*-\emptyset$ > the *patiens* of transitive verbs which would equal the new accusative, and the common gender accusative $*-\emptyset$ > the new vocative (when it wasn’t what would become PIE $*-e$). The genitive $*-s$ and the ablative $*-s$ are confounded, as they are also similar functionally (Matasović *Clause Alignment in Proto-Indo-European* 13). This would be followed (stage IV) by Szemerényi’s Law, viz. $*-Rs\#$ > PIE $*-\bar{V}R\#$, thus $*ph_2ter-s$ > $*ph_2tēr$, and the accusative $*-m$ either denoted definiteness, or was a directive (direction) case. Therefore, the accusative $*-m$ was formed, the absolutive was replaced by the accusative, and therefore the old neuter accusative became the neuter vocative; PIE $*-e$ was added to the vocative, agglutinated in fact (Matasović *Clause Alignment in Proto-Indo-European* 15-16).

Classical Proto-Indo-European was already influenced by this North Caucasian

substrate (Lühr 2019: 181), which, in Kortlandt’s judgement, *made* Proto-Indo-European into an ergative language, with grammatical gender and adjectival agreement (Kortlandt 2010: 36; cf. Bjørn 2019: 40 for a visualisation of this). It has been noted that various crises are what made the Illyrians speak in Vulgar Latin suddenly, or the Brittons with regard to Vulgar Latin, and then Saxon; thusly one can surmise how Pre-Proto-Indo-European speakers, speakers of a nominative-accusative languages, became the speakers of PIE, an ergative language (Matasović 2012: 3). Bomhard (2019) presents the evidence in favour of the thesis that the difficulties presented by the various attempts at reconstructing Proto-Indo-Uralic are caused by Proto-Indo-European having a North Caucasian substrate shift its morphological profile. Kortlandt envisages PIE as being influenced by a North-West Caucasian substrate language even without the extralinguistic evidence.

The accusative *-m

The Proto-Uralic accusative case-marker PU *-m is of supreme importance to linking Proto-Uralic to Proto-Indo-European, which had a PIE *-m morpheme for accusative-marking. Collinder (1960: 282; Aikio 2022: 4, 14 concurs) offers an accusative *-m, as does Abondolo (2006a: 18) who restricts its use to the definite direct object of finite verbs (inflected verbs) for person. This case is reconstructed on the basis of the Mari accusative -m, Samoyed (Tundra Nenets accusative -m, Taz Selkup accusative -əm₂) accusatives, as well as “eventual traces” or “correspondences” in Finnic, Permic and, “perhaps” in Mordvin (Abondolo (2006a: 19), such as the Mordva accusative -n^j, with the genitive -n^j developed from it independently due to Mordva prosody, adding that no visible traces of such an accusative were to be found in Hungarian nor in Khanty (Raun 1988: 558, Abondolo 2006a: 19). Thus the core (Hungarian and Permic, that is to say Komi and Udmurt) does not preserve the accusative *-m, except for South Mansi. The periphery and thereabouts have reflexes of the PUE *-m, which in Finnic *-m# > Proto-Finnic *-n#; thus there is no accusative Finnic (Kittilä, Laakso, Ylikoski 2022: 881), which is ultraconservative in terms of vowels (and, to some extent, of the consonants) because it (PU *-m) fused with the genitive PU *-n into a Proto-Finnic *-n (Kittilä, Laakso, Ylikoski 2022: 881, 883; (Abondolo 2006a: 19).

Designated as an *accusative* (marking the definite object of the sentence), Raun quotes Vértés (1960: 192) who suggested that it may rather have been a “deictic-determining element, related to the first person marker,” (Raun 1988: 558) citing the examples of the word ‘fish’ in the accusative in North Sámi *guolle-m*, Mari *kolə-m*, Mansi *kul-ma* and Tundra Nenets *xāl'em*. Being an ultrastable phoneme, this resonant is also the accusative case marker in PIE, and is thus reconstructed into Proto-Indo-Uralic as well. Kortlandt (2010: 36) reconstructed this morpheme as a Proto-Indo-Uralic lative-accusative **-m*, later revising it, in effect simplifying it and bringing into more into line with PIE and PU as an accusative **-m* (Kortlandt 2020a: 2-3), already noted as such in Pereltsvaig, Lewis (2015: 200; Kortlandt 2020b: 3). This and the genitive (discussed below) are the obvious candidates for the Indo-Uralic hypothesis, alongside two dozen or so other morphemes whose grammatical values can most often only be approximated²⁴.

Such examples as Russian *ветром снесло крышу* ‘the roof was blown away by the wind’, where the inanimate agent is in the instrumental (*ветром* ‘wind (instr. sg.)’) and the object is in the accusative (*крышу* ‘roof (acc. sg.)’) is taken as evidence that the “the accusative is of Indo-Uralic origin and therefore older than the ergative.” (Kortlandt 2010: 36) In fact, Finnish and Russian share a feature (thus Finnic or Uralic and Balto-Slavic or Proto-Indo-European, as Finnish and Russian are, within their language branches and families, taken to be conservative), namely that they both mark the direct object of an affirmative sentence with the accusative case (as in *Hän luki kirja-n*. ‘He read a book’ = *Он читал книгу-у*. ‘ibid.’), and they both mark the direct object of a negative sentence with the partitive case and the genitive case (as in *Hän ei lukenut kirja-a*. ‘He didn’t read a book’ = *Он не читал книгу-и*. ‘ibid.’) (Pereltsvaig 2012: 209).

The Proto-Indo-European accusative singular marker PIE **-m* referred to common nouns in Late PIE; presumably, this PIE **-m* was a directional particle or postposition. If it’s true that Proto-Indo-European was ergative, then subjects of transitive verbs have PIE **-s* and subjects of intransitive verbs and objects PIE **-ø*, and therefore neuter PIE **-ø* in Late

²⁴ In 2010 Kortlandt listed the following Proto-Indo-Uralic morphemes (Kortlandt 2010: 36): genitive **-n*, lative-accusative **-m*, dative-locative **-i*, ablative-instrumental **-t*, dual **-ki*, personal pronouns **mi*, **ti*, **me*, **te*, reflexive **u*, negative **n-*, interrogative **k-*. In the same year he gave the following list for the PIU morphemes (Kortlandt 2010: 405-406): first person **m*, second person **t*, demonstrative **i/*e*, demonstrative **t*, demonstrative **s*, dual **ki*, plural **t*, participle **nt*, particle **l*, verbal noun **s*, reflexive **u/*w*. In 2020 he slightly revised this to the following list (Kortlandt 2020a 2-3): 1st person **m*, 2nd person **t*, demonstrative **i/*e*, demonstrative **t*, demonstrative **s*, dual **ki*, plural **t*, plural **i*, accusative **m*, genitive **n*, dative **ka*, locative **ru*, locative **n*, locative **i*, ablative **t*, diminutive **k*, nominaliser **i*, nominaliser **m*, participle **n*, participle **t*, participle **nt*, participle **l*, verbal noun **s*, conative **sk*, reflexive **u/*w*, negative **n*, interrogative **k*.

Proto-Indo-European signified that they never were subjects of transitive verbs and, if Hittite can be taken to be a reliable source on these matters (being influenced by substrate languages), no neuter nouns can be subjects of transitive clauses in Hittite (Matasović 2004: 182)²⁵. Early IE languages don't have markings on inanimate actors (Matasović 2004: 183). Kortlandt offers the idea that the sigmatic nominative singular marker PIE *-s < animate ergative < ablative *-s < *-t-i (!) (and, similarly, the nominal nominative plural PIE *-es < *-et-i) (Kortlandt 2020b: 3). It is conceivable the accusative singular marker PIE *-m was applicable to all nouns at some early stage, but that's unprovable; it is, however, more likely that PIE *-m marked definite objects only, and some neuter nouns would be marked as indefinite objects (abstractions and similar semantically indefinite categories among nouns). Thus neuter nouns in Proto-Indo-European morphologically do not distinguish between the nominative and accusative cases, forming a common nominative-accusative, and unmarked, at that (Matasović 2004: 183, 186). And since the nominative case is, looked at typologically, usually unmarked, it would suggest that the bare-stem nominative-accusative for the athematic (lacking a Proto-Indo-European thematic/ablauting vowel) neuter nouns is the earliest reconstructable nominative, viz. the asigmatic Proto-Indo-European nominative *-∅

²⁵ For instance, in Hittite transitive sentences with neuter nouns as subjects are in the ergative, with the derivational suffix *-anza-*, but this isn't ergative-case marking as *-anza-* is a derivational morpheme, not an inflectional one (Matasović 2000: 330); nonetheless, this suffix was an adjective formant, viz. either agentive or patientive and “[i]t seems then that the meaning depends on the transitivity of the verb. If the verb is transitive, then the *-ens-* derivative is patientive; otherwise, it is agentive. This distribution is also found with the Hittite participles in *-ant-* and is usually assumed to be original for the PIE **-nt-* participles. (Lubotsky 2019: 160), and, quoting Kortlandt (2010: 397, working on Bojan Čop's Indo-Uralic) PIU **-ti* assibilated into **-si*, implying that the suffix **-ens* started as **-ensi* < **-ent-i*, thus reflecting the original locative case form of the PIU participle (Lubotsky 2019: 160). This agentive also existed in Hieroglyphic Luwian sg. *-antis*, pl. ?; Cuneiform Luwian sg. *-antiš*, pl. *-antinsi*; Lycian sg. ?, pl. *-ēti* (Lopuhää-Zwakenberg 2019: 133); as to the Anatolian languages, this agentive is present only in Lycian and in Neo-Hittite, and *-anza* is an interpretation of the *-ant-* suffix. The Lycian plural agentive *-ēti* < *-et-i* is a hapax, and is structurally identical to the Hittite form, thus deriving it from Proto-Anatolian **-ont-* (Lopuhää-Zwakenberg 2019: 145). Hoffner and Melchert (2008: 72-73) consider it unclear whether or not it was fully grammaticalised in Proto-Anatolian; neuter nouns which are the subjects of transitive sentences in Hittite have an ergative ending: erg. sg. *-anza*, erg.pl. *-antēš*. This ending derives from *-ant-*, which “animatises” neuter nouns referring to inanimate things and forms the participle, from PIE **-ent-*. So, erg. sg. *-anza*, erg.pl. *-antēš* really is ergative, as it is grammatically motivated, e.g. neuter *MUNUS.MEŠ hazkara(i)-* ‘female dancers’ > ergative *MUNUS.MEŠ hazkaranza*. Some have rejected that Hittite was an ergative language, because the erg. sg. *-anza*, erg. pl. *-antēš* morpheme ‘animatises’ the noun to *make* it ergative, this morpheme originally (and synchronically in Hittite) being a derivational suffix of verbal origin, forming, among other things, participles, as in *haššant-* ‘born’, *huwant-* ‘running’, but it can be used separately, too, as in *anda hūppanduš NA₄.HIA* ‘stones gathered together’. As to split ergativity, the subject of transitive and intransitive verbs is equal to the nominative in Hittite, and the direct object of transitive and intransitive verbs is equal to the accusative. But, neuter subjects of transitive verbs (eg. *to give*) were marked as ergative. Thus neuter nouns cannot function as the subject of an intransitive verb in Hittite. Common nouns and pronouns were marked as nominative-accusative, while neuter nouns as absolutive/ergative, as in *laman* ‘name’ (neuter n-stem), *haštae* ‘bone’ (neuter ai-stem) can be ergative, but not absolutive (Hoffner, Melchert 2008: 66-68). Thus this agentive morpheme PIE **-ens-* is reconstructed for several PIE lemmas: **meh₁-ns-* ‘moon’, **g^{wh}r-ens-(o)-* ‘heat’, **dh₁-ens-* ‘dense’, **trh₂-ns* ‘through?’, **d^hwens-* ‘to scatter, sprinkle’, **k_{ens-}* ‘to declare’, and **g^hh₂-ens-* ‘goose’ (literally meaning ‘gaping [animal]’, as **g^heh₂-* ‘to gape’, viz. **g^hh₂-(e)w-* > Ancient Greek *χάος*) (Lubotsky 2019: 151).

(Matasović 2004: 186).

Lopuhää-Zwakenberg (2019: 147, 148) summarises the case endings for clause alignment in Proto-Classical Indo-European (that is post-Anatolian IE) thusly: agent masculine *-os, agent neuter *-om, subject masculine *-os, subject neuter *-om, patiens masculine *-om, neuter *-om, and in Proto-Anatolian as: agent common gender *-os, agent neuter n/a, subject common gender *-os, subject neuter *-on, patiens common gender *-on, neuter *-on, thereby positing the reconstructed forms for the common ancestor of Proto-Classical Indo-European and Proto-Anatolian, Proto-Indo-European, as: agent masculine *-os, agent neuter n/a, subject masculine *-os, subject neuter *-om, patiens masculine *-om and neuter *-om. Note the inventory of the cases in Proto-Indo-European, as defined by the Leiden school of historical linguistics, as excluding Anatolian (Kortlandt 2020b: 4-5, this table being a tabular form of Kortlandt's text, only showing the singular forms in PIE, as the plural forms seem to have no relation to PU and are derivable differently)²⁶:

	athematic	animate thematic	neuter thematic
<i>nominative sg.</i>	*-(s)	*-os	*-om
<i>accusative sg.</i> ²⁷	*-(m)	*-om	*-om
<i>genitive sg.</i> ²⁸	*-iŋ, *-(e/o)s	*-os	*-om
<i>ablative sg.</i> ²⁹	*-(e/o)s	*-os	*-os
<i>instrumental sg.</i> ³⁰	*-(e/o)t	*-ot	*-ot

²⁶ Note Kortlandt's notation of the so-called laryngeals: *h₁ he writes as *ʔ (a glottal stop), *h₂ as *ʕ (a pharyngeal fricative) and *h₃ a *ʕ^w (a pharyngeal fricative with lip rounding), taking into consideration that in Proto-Anatolian and Hittite *h₁ (*ʔ) gave ø, *h₂ (*ʕ) gave ħ and *h₃ (*ʕ^w) also gave ħ.

²⁷ Kortlandt (2010: 36) reconstructed a lative-accusative PIU *-m, later revising it to an accusative PIU *-m (Kortlandt 2020a: 2). This is a simplification, as there is, in fact, little suggesting that PIE would have had a lative or para-lative case; this is a particularity of PU.

²⁸ Reconstructing, as is expected, a PIU genitive *-n (Kortlandt 2010: 36, Kortlandt 2020a 2-3)

²⁹ Cf. the following footnote.

³⁰ No instrumental case can be reconstructed for PU (the instrumental in the daughter languages are formed lexically, as in Estonian, cf. above); in PIE, the Proto-Indo-European instrumental construction was not equal to the Proto-Indo-European comitative one (as in Latin *pisces hamo capio* (instrumental), vs. *eo tecum* (comitative) (Matasović 2008: 10), but they shared a common case morpheme, *-(e/o)t. Kortlandt at first reconstructed this as an ablative-instrumental PIU *-t (Kortlandt 2010: 36), then restricting the formerly ablative-instrumental *-t as an ablative *t (Kortlandt 2020a 2-3).

<i>dative-locative sg.</i> ³¹	*-(i) ³²	*-o ³³	*-o ³⁴
--	---------------------	-------------------	-------------------

The genitive *-n

Reconstructed by Collinder (1960: 282; Aikio 2022: 14 concurs) as the genitive *-n, it is a problematic reconstruction, with some researchers not reconstructing it in Proto-Uralic at all (Raun 1988: 558). Aikio (2022: 14) notes that the plural that goes with this morpheme is *-j, not *-t, which is agglutinated only to the (zero-marked) nominative. The Proto-Uralic genitive *-n is far less attested than the Proto-Uralic accusative *-m, viz. Finnish *-n*, Erzya Mordva *-enʹ*, Meadow Mari *-ən* and Taz Selkup *-ən₂* (Abondolo 2006a: 19), and there is no genitive of any sort in either Permian (Udmurt, Komi) nor in Hungarian (Abondolo 2006a: 23, Raun 1988: 558, Kittilä, Laakso, Ylikoski 2022: 881), where the genitive is formed from a different root or in a different manner altogether. Nonetheless, a proto-genitive *may* have left traces in these two branches, and this is why it can be reconstructed all the same (Raun 1988: 558). This *-n is, at once, a subordinate suffix and a genitive marker, a pronominaliser with nouns, and an adverb formant with verbs (Abondolo 2006a: 18), which seems to indicate that the case-ending is derived from one of its other reconstructable functions, which semantically makes sense. This genitive and its reflexes in the Uralic languages and the synchronic attributive adjective of the same languages often coincide. Thus, ‘of the fish’: Finnish *kala-n*, Sámi *guolle-n*, Mari *kolə-n*, Selkup *qəli-n* (Raun 1988: 559). Kortlandt, much like for the nominative, sees a PIU genitive *-n as directly giving form to the Proto-Indo-European (which was supposed to be lost in Proto-Indo-European) and the Proto-Uralic genitives (Kortlandt 2010: 36; Kortlandt 2020a 2-3). It is to be noted that this *-n is an ultrastable phoneme, meaning that these resonant phonemes historically seldom become a different phoneme, compared to how often other phonemes are able to do this in various phonotactic combinations; moreover, it is not uncommon for word-final nasals to change from one into another.

³¹ At first offering a PIU dative-locative *-i (Kortlandt 2010: 36), later revising it into a locative *i, alongside two more locative morphemes, locative *ru and locative *n, and reconstructing a dative *ka (Kortlandt 2020a 2-3).

³² Which, in Classical (viz. post-Tocharian) IE became dative *(e)i and locative *(i).

³³ Which, in Classical IE became dative *-o?ei, locative *-o?i.

³⁴ Which, in Classical IE became dative *-o?ei, locative *-o?i.

The directional cases

What can be said generally about the Proto-Uralic adverbial case inventory is that it was a directional case system, affixing this grammatical information with three types of directionality: where?, where from? and where to? (Hajdú 1983: 105). Similarly to how Proto-Uralic casogenesis is related to some seemingly completely unreconstructable postpositions (due to the shortness of the particles in question, in fact their monosyllabicity), this same manner of generating case suffixes can be neatly studied in Old Lithuanian, which created four case suffixes (for the inessive, illative, adessive and allative cases) out of various Indo-European postpositions (Kulikov 2009: 445-447; a similarly interesting paper on *ad nauseam* casogenesis in Ossetic and Uralic is Holopainen, Ylikoski 2022). Thus, depending on the analysis, one can say that Proto-Uralic had six cases (the three grammatical cases mentioned above, and the three directional cases below), eight (if we add at least two adverbial cases, as does Aikio, cf. below), some indefinite number beyond eight (almost a dozen lative cases have been offered, cf. below on the ‘lative paradigm’), or just two, that is to say the only non-nominative grammatical cases or, as some have done for Hungarian, consider this language to have had no cases whatsoever, only somewhat synchronically visible fossilised postpositions. The non-grammatical cases can be considered to be adverbial, directional or semantic cases (Kittilä, Laakso, Ylikoski 2022: 880), as opposed to supposedly *purely* grammatical cases. The fact that the Uralic languages have a wide variety of number of cases and of degrees to which they spread out their directional cases systems goes to show that for a long time after the disintegration of Proto-Uralic the cases were continued to be felt as postpositions (all the more so given the word-initial stress across almost all of Uralic), and as is shown, for instance, in Finnish and Hungarian, where a number of non-productive (viz. synchronically semantically-motivated) cases are still in use (cf. Anhava 2010), as well as evidence from Old Hungarian indicating the same (see above). In a sense, the accusative **-m* and the genitive **-n* were *the only* cases in Proto-Uralic; the nominative is a non-case (an ‘absolute’, and unmarked), and the directional cases are adverbial, unless we accept a localist hypothesis of how the cases formed (wherein all the cases are, in as abstract a sense as possible, *local*, in an approach that is almost cognitive) – this may well be the case for Proto-Uralic, but it is completely unverifiable, as the grammatical cases, too, are not without their share of doubt cast upon their reconstruction (see above).

The locative *-nA

The locative is reconstructed as marking the *stasis*, or lack of movement in a three-way directional case system designating movement, the other two markers being of two sorts, or directions of movement: moving away from (the *separative*, see below) and moving toward (the *lative*, see below). Note that Proto-Indo-European adverbs were most often formed from the accusative singular and locative singular forms of nouns (accusative and locative, meaning movement and location) much of the same can be said of any instance of casogenesis, so too in Proto-Uralic. Reconstructing this case is relatively straightforward, as its reflexes are to be found in each of the nine Uralic branches, “although its original local meaning prevailingly appears in adverbs and postpositions,” and when it is used as a case morpheme, it has specialised onto the Finnish and Sámi essive case (Finnish *alla*, from a previous **alna* ‘below’, Sámi *vuollen*), the inessive case in Komi (Komi *ulin*), the superessive case in Hungarian, and it has branched out into locative, instrumental, essive, causal-final and more cases in Northern Khanty (also Mari *ülnə*, Mansi *joln*, meaning ‘under, below’) (Raun 1988: 559). Collinder (1960: 282) reconstructs this locative as *-*na* ~ *-*nä*, Abondolo (2006a: 18) concurring as well as Aikio (2022: 14), noting it as *-*nA*. This Proto-Uralic case morpheme is well attested across the Uralic languages in various functions derived from its original semantic function of denoting *stasis*, viz. Hungarian *tav-on* ‘on a lake’, Finnish *talo-ssa* (< **talo-s-na*) ‘in a house’ (Abondolo 2006a: 20), also deriving the Finnish essive (meaning ‘functioning as a’) from this PU locative *-*nA*, as in *koto-na* ‘at home’ (Abondolo 2006a: 20)³⁵. Tocharian B had a phonologically similar locative case ending *-ne*, which, much like in Old Lithuanian, may well be due to substantial Uralic substrate influence or language contact (cf. Peyrot 2019a). Kortlandt, for his part, offers a similarly bewildering number of locative or locative/lative-like cases in PIU, viz. dative **ka*, locative **ru*, locative **n*, dative-locative **-i*, locative **i* (Kortlandt 2020a 2-3), with Proto-Indo-European having only had one, **-i*, which he claims is a direct descendant of the dative-locative **-i* of the PIU locatives (Kortlandt 2020b: 3).

³⁵ Raun also mentions a “second locative”, reconstructable as *-*t(t)*, attested only in Mansi as a locative case ending, as in Mansi *timät* ‘at this place’, as an adverb in Khanty, as in *tot* ‘there’, and in Hungarian as an aberrant toponymic locative, as in *Kolozsvárt* ‘in Kolozsvár’ (Raun 1988: 560). All in all, the locative cases are multiform and variegated in their uses in the attested Uralic languages, thus offering many different possibilities as to their reconstruction, as “[t]wo locative suffixes and even more lative suffixes have been under discussion” in Proto-Uralic studies (Raun 1988: 559).

The separative *-tA

A case designating *moving away from*, forming part of a three-way locative, and a two-way directional case system in Proto-Uralic from is reconstructed by Collinder (1960: 282) as the separative *-ta ~ *-tä, concurred by Aikio (2022: 14), and by Abondolo (2006a: 18) as the separative *-tA ~ *-tI (see below on problems concerning Proto-Uralic vocalism). This case is morphologically clearly similar to the PIE ablative singular of the *o*-stem nouns³⁶, *-os, as much assibilation is reconstructed into Proto-Indo-Uralic lemmas and morphemes when becoming Proto-Indo-European ones (see above). But no other cases, save for the phonetically ultrastable accusative and genitive case endings in PU and PIE, are obviously comparable in the two protolanguages (and the Tocharian B reflex (al)lative -ś(c), Tocharian A -ac are unclear, whether or not it has any relations to a Uralic substrate, cf. Peyrot 2019a, and above). Raun calls this *-TA or *-tə as either a separative or an ablative, as the case that designated, to put it simply, a genitive with an ablative meaning. Adding that the simpler solution to reconstructing this morpheme as “*-δa (with a non-alternating *δ)” from the Finnic (where it can have a partitive function, with *-tA ~ *-tI > the Estonian partitive -d, as in *võid* ‘butter-partitive’), Sámi, Mordvin (as the ablative Moksha Mordva -də), Mari, Mansi, Khanty, Hungarian and Samoyed reflexes, viz. ‘from beneath’: Finnish *alta*, Sámi *vuolde*, Mordvin *aldo*, Mari *üləts* (with -ts < PU *-t), Hungarian *alól*, Samoyed *nyld* (Raun 1988: 559, Abondolo 2006a: 20).

The ‘lative paradigm’

The morphemes reconstructed as meaning *motion toward*, “or ‘moving along’ but also ‘becoming somebody or something’” (Raun 1988: 560) are called the lative case, being in most respects the complementary opposite of the separative case. Abondolo offers two forms, each with two possible reconstructions in PU as *-k (and/or *-ŋ) and *-cʲ (and/or *-ŋʲ) (Abondolo 2006a: 18), noting that it was the most used case (Abondolo 2006a: 20). Raun calls them suffixes, due to them having been more than one, noting that the “most popular

³⁶ One of only two thematic, viz. non-consonantal PIE nominal stems. The *o*-stem and the *eh₂*-stem are called thematic stems due to them incorporating a thematic vowel, but with no ablauting across the cases in the root nor in the ending (e.g. the zero-grade *wǵkʷos ‘wolf’, *wǵkʷosyo, *wǵkʷōy &c.); all the other stems are *athematic* and, therefore, consonantal (since *i/*y and *u/*w are glides) and both roots and endings (and even the suffixes) here were open to ablaut, and not necessarily in line with each other (e.g. *pont-eH-s ‘path’, *pnt-H-e/os with a zero-grade in the root but a full-grade in the suffix).

among them has been **-k(V)*” (Raun 1988: 560), morphologically similar, if not identical to the dual marker. Aikio (2022: 12) mentions Ylikoski (2011: 256) explaining this lative paradigm as a “persistent assumption” that Proto-Uralic synchronically had numerous latives, viz. directional case morphemes all of which are **C* (**-k*, **-n*, **-ń*, **-ŋ*, **-s*, **-j* among others), which brings up the question of the imprecision of case names, and whether or not to reconstruct such a richly diverse (and unclear) directional case system, these being adpositions with adverbial meanings which are diversely attested across the Uralic language. Their number may well have been even greater than what is confidently reconstructable, but their mutual relations are unclarified (Raun 1988: 560); this is certainly the reason why Uralic is categorised as a “case-increasing language” (Kulikov 2009: 455), being a postpositioning language which is capable of creating new cases (and shedding them, as, for instance the lative is obsolete in Finnish, fossilised in some adverbs, e.g. *kauas* ‘far away’). Abondolo (2006a: 19) notes that the richer the case system, the more the norm it is that the lative category hypertrophies. Aikio (2022: 14) offers **-ŋ* as the lative case (with an example: **jǎŋi-n ũli-ŋ* ‘ice-genitive on-lative ‘on the ice’ (Aikio 2022: 22)), alongside only two adverbial cases as offered by him: the translative **-ksi* (in Finnic, Mordvin and Mari) and the caritive (or abessive or privative) **-ktAk* (in Sámi, Finnic, Mari, Permic and Ob-Ugric) (Aikio 2022: 14-15), dismissing other potential cases as unverifiable as such, only formally reconstructable into a set of contextless suffixes, or postpositions, or cases. Examples abound: the lative Ingrian *-k*, in *alak* ‘under’, the lative Sámi *-kV*, in *deikě* ‘hither’, the lative Mari *ũlkə* ‘under’, the Mordvin prolativ *-ka*, *-ga/-ge*, *-va* which “may be of the same origin”, Komi *ulę* ‘down’, Udmurt *uk* ‘down’; in Mansi the PU **-k* is to be found in “both in adverbs and in the *-jy* of the translative case, e.g. *pijanēr-jy* ‘to a pioneer’”, Northern Khanty lative suffix *-a*, *-ä*, as in *xōta* ‘into a house’, Eastern Khanty translative adverbs in *-γ*, as in *katâγ* ‘to a house’, Hungarian translatives *-vé*, *-vá*, in *emberré* ‘to a man’ (possibly also (or from a **-j*), the Hungarian latives and adverbials in *-é*, *-á*, as in *belé* ‘into’, perhaps even Selkup *-qo* as agglutinated into a genitive, as in *qorqytqo* (from *qorqy* ‘bear’) (Raun 1988: 560).

As mentioned above, most Uralic languages are concatenative with regard to case, number and possessor, the three categories that can mark a noun³⁷; because Uralic nominal morphology is not fusional, that is to say it is segmentable, it is very reconstructable – too reconstructable, in fact, hence the cascade of lative suffixes/postpositions, and the

³⁷ As in Finnish *talo-i-sta-mme* house-PL-ELA-POSS.1PL, which is parallel to Hungarian *háza-i-nk-ból* house-PL-POSS.1PL-ELA and to Komi *керка-яс-сьы-ным* house-PL-ELA-POSS-1PL (Kittilä, Laakso, Ylikoski 2022: 882).

impossibility of determining which among them were cases in Proto-Uralic, and which were not.

On Uralic cases not reconstructable into Proto-Uralic

It would seem that the proverbially large case inventory of some Uralic languages is not representative of Proto-Uralic (Kittilä, Laakso, Ylikoski 2022: 881), in spite the endless possibilities as regards reconstructing them, especially the latives. The attested Uralic languages have a plethora of cases which are not represented across the language family, but which do point to the fact of a very productive casogenesis in Uralic and, by implication, in Proto-Uralic; for instance, the Permic languages tend to have many cases, ie. standard Udmurt has 15, standard Komi has 16-23 (Kittilä, Laakso, Ylikoski 2022: 881). Nonetheless, most Uralic languages (and there are some 40 or so, depending on the criteria for distinguishing them) have few cases, and this is ultimately why few are reconstructed in Proto-Uralic (coupled with the uncertainty as regards the number of lative cases): Erzya Mordvin has “about” ten cases, but North Sámi has only six, and North Khanty has the fewest of any Uralic language, namely three (Kittilä, Laakso, Ylikoski 2022: 881). Finnish, proverbial for its spatial and directional cases, has six directional cases, which are neatly complementary, viz. inessive *-ssa/-ssä*, elative *-sta/-stä*, illative *-Vn, -hVn*, ablative *-seen/-siin*, adesive *-lla/-llä* and allative *-lle* (Karlsson 1999: 107; Hummo, Ojutkangas 2006), summarised in a tabular form in the following way, which may hint at how, at least, the postpositions (if not the lative paradigm) in Proto-Uralic may have functioned (in unison forming a clear spatial picture):

		<i>location</i>	
		<i>inside</i>	<i>outside</i>
<i>direction</i>	<i>static</i> (locative)	in-essive	ad-essive
	<i>away from</i> (separative)	e-lative	ab-lative
	<i>towards</i> (lative)	il-lative	al-lative

A number of other grammatical cases are found in Finnish (and in Estonian), such as the essive (denoting a *similaris*) in Finnish, “a (temporary) state or function, sometimes circumstances, conditions or causes” (Karlsson 1999: 123), and the translative denotes “a state, property, function or position into which something or someone enters, or the end point of a movement or change” (Karlsson 1999: 125), beside the abessive (denoting ‘without’), comitative (denoting ‘with, accompanied by’) and instructive (in fixed plural expressions which, in Finnish, are rare, appearing mainly in fixed expressions like idioms) (Karlsson 1999: 127), as well as partitive cases (denoting incomplete nouns), instructive cases (denoting means by which an action is performed) and prolative cases (denoting thanks to what an action is performed, by way of), cf. Anhava 2010: 244. Likewise, not all non-Proto-Uralic cases in the attested Uralic languages are of a directional semantic sort (for a summary of the adverbial cases in Uralic, see the table at Kittilä, Laakso, Ylikoski 2022: 890); these non-grammatical cases are local cases, possessive cases or comitative/instrumental (Kittilä, Laakso, Ylikoski 2022: 880). For instance, Estonian has a comitative-instrumental *-ga/-ka*, Hungarian has a sociative *-stul/-stül*, which are of a purely individual derivation, derivable synchronically at times, among many others throughout the attested Uralic languages.

It has been noted that there are formal similarities between many Uralic comitative-instrumental cases and abessives, but this is inconclusive, and therefore not reconstructed into Proto-Uralic (Kittilä, Laakso, Ylikoski 2022: 892); nonetheless, Aikio, does offer an abessive in Proto-Uralic (see above). This shows that what is a case is quite arbitrary; adverb derivation and case morphemes are a vague zone in the Uralic languages (Kittilä, Laakso, Ylikoski 2022: 879). Aikio says as much, since the postpositions are seemingly equal to the cases in Proto-Uralic³⁸, save for the accusative **-m* and the genitive **-n*; some legitimately claim that Hungarian has no cases at all (or, that Hungarian has anywhere between 17 and 27 cases) (Kittilä, Laakso, Ylikoski 2022: 879), since they are but fused postpositions, or just enclitical, as the language is invariably stressed word-initially (Kittilä, Laakso, Ylikoski 2022: 880).

One of the principal problems caused by the small, and temporally quite shallow Uralic corpus (Aikio 2022: 3, 11) is that the phonological reconstruction of Proto-Uralic is a subject yet to form a consensus upon. It is true that in most, if not almost all reconstructed

³⁸ Moreover, a small number of relational nouns (functioning as adverbs or postpositions) can be reconstructed into PU, thus creating a set of nouns in PU which “carried” the grammatical information of the case: for instance, Aikio (2022: 13), who reconstructs five unambiguous relational nouns (**üli-* ‘space on or above’, **jla-* ‘space under or below’, **edi-* ‘space in front of’, **läši-* ‘space near or opposite to’ and **müñä-* ‘space behind (one’s back)’), notes that such forms as the locative of **üli-* ‘space on or above’, **ül-nä* ‘on’, its ablative **ül-tä* ‘off (from)’ and its lative **üli-ŋ* ‘onto’ forms were all fossilised postpositions in themselves.

languages there will always be some degree of uncertainty as regards their phonemic inventories by the very nature of studying languages with no native speakers; Proto-Uralic is somewhat particular in both the relatively large number of daughter languages which can be made use of for reconstruction, as contrasted with the time-depth of Proto-Uralic and lack of ancient texts, or even epigraphy; hence its particular status among reconstructed languages. Here we will touch upon this issue, as it is central to how the researchers of Proto-Uralic have gone about using the extant Uralic languages to give shape to the Proto-Uralic case morphemes.

Sampsa Holopainen (2020) thoroughly assessed the state of the phonological reconstruction of Proto-Uralic, which has as of recently been principally aimed at confirming or denying the Indo-Uralic hypothesis. Most Proto-Uralic phonemes which are problematic do not concern the reconstructed case morphemes; the problematic reconstruction being those of PU *ś and whether or not it is equal to *ć, as well as *š and its identity or less to *č (Holopainen 2020: 295) and, of course, of the Proto-Uralic vowels – with the initial vowels, meaning those invariably under stress, being a category apart from all the other ones; thus, again, *not* concerning the case morphemes except via vowel harmony. The reconstructed vocalic inventory of Proto-Uralic is in stark distinction with PIE’s reconstructed ablaut and minuscule vocalism, the phonology of the central and core branches (Mari and Ugric) and the so-called Proto-Uralic “proto-phonetics”, such as the status of *δ and δ’ (which Aikio has offered as being reconstructable as dental fricatives) (Holopainen 2020: 295).

For a synoptic reconstruction of the Proto-Uralic consonants, see the summary by Aikio (2022: 6); briefly, the consonant inventory is reconstructed, by Aikio (with seemingly minor, but crucial differences in Abondolo 2006a or Sammallahti 1988), thusly (Aikio 2022: 5-8, basing his, as most other sources generally do, his notation on the Uralic Phonetic Alphabet, itself generally based on Finnish orthography)³⁹:

	labial	dental/alveolar	postalveolar/retroflex	palatal	alveopalatal	velar
			x		al	

³⁹ Simply put, this is based on the fact that most Uralic languages have labi(odent)als [p] & (f), dentals [t] & [s], postalveolar/retroflex [tʃ] & [ʃ], palatals [tʲ] & [sʲ], velars [k] & (x), hence the general idea of the consonant inventory (Abondolo 2006a: 10). Likewise, pulmonic egressive airstream is common in Uralic, save for Finnish, where some pulmonic ingressive airstream is heard; quantity, voice and palatalisation are present across Uralic, but are represented in a skewed way across it; there are quantitative consonant distinctions in Finnish, Estonian, Selkup, and voice distinctions in Hungarian and Permian; and palatalisation in some form or other in all Uralic languages, save for Finnish (Abondolo 2006a: 8-9).

glides	*w			*j		
nasals	*m	*n			*ń	*ŋ
stops	*p	*t				*k
affricates			*č			
fricatives/sibilants		*s	*š		*ś/?*ć	
lateral		*l				
trill		*r				
<i>unclear</i>		*d/δ			*dʹ/δʹ/?*lʹ	*x

The vowels are a completely different matter, and of relatively lesser importance to the cases than the consonants as Proto-Uralic was stressed word-initially⁴⁰. The cases were postponed (as they are agglutinated postpositions), and because of vowel harmony, no problems like an Auslautgesetz are reconstructible into Proto-Uralic. Of the Proto-Uralic vowels Daniel Abondolo judges that “[p]robably no academic field is more contentious than the prehistory of the Uralic vowels...” (Abondolo 2006a: 16). Aikio, as is visible from his *Uralic Etymological Dictionary*, admits much more than Collinder did that the vowels are too problematic to reconstruct with much certainty, thus writing a placeholder *V more often than most in his reconstructions – this is, academically speaking, the most honest position. His reconstruction of the inventory consists of the following phonemes, in essence like Collinder’s reconstruction⁴¹ but with a *i̯ (see footnote 42) instead of Collinder’s *õ and *y

⁴⁰ Word-stress (or ictus) is word-initial in all of Uralic, with some, explainable, exceptions: Estonian tolerates non-initial icti on loanwords, a belt of west-central Uralic languages, namely Moksha Mordva, Mari and all of Permian save for Komi, and Samoyedic (Abondolo 2006a: 8). All Uralic languages have their stress accent on the first syllable, and Proto-Uralic does too, and the postpositional casual morphemes are invariably pronounced with less stress; hence the vowel of the case morphemes being interchangeable. Vowel harmony is reconstructed in Proto-Uralic despite not being synchronically present in very many Uralic languages (Hajdú 1983: 102), as a system of distinguishing word-boundaries when the stress accent is inevitably word-initial, which may seem non-transparent due to consonant gradation – a phenomenon which, for instance in Finnish, means *p*, *t*, *k* in syllable-initial position become, in final syllable position in oblique cases: *kV#* > *øV#*, e.g. *sika* > *siassa*, *pV#* > *vV#*, e.g. *tupa* > *tuvasa* (the difference being an open syllable and a closed syllable), *tV#* > *dV#*, e.g. *katu* > *kadulla*, *kkV#* > *kV#*, e.g. *kukka* > *kukassa*, *ppV#* > *pV#*, e.g. *kuppi* > *kupista*, *ttV#* > *tV#*, e.g. *hattu* > *hatutta*, similarly for *lt* > *ll*, *nt* > *nn*, *rt* > *rr*, *mp* > *mm*, *ht* > *hd*, *rk* > *r*, *nk* > *ng* (Hajdú 1983: 103).

⁴¹ Slightly different opinions are found in Collinder. Briefly, Collinder (1960: 184) reconstructed nine Proto-Uralic vowels, namely: the low unrounded back vowel *a, the low unrounded front vowel *ä, the middle-high rounded back vowel *o, the middle-high unrounded back vowel *õ, the middle-high unrounded front vowel *e, the high unrounded back vowel *y, the high rounded back vowel *u, the high rounded front vowel *ü, the high unrounded front vowel *i, noting that Proto-Finnic has preserved the entire Proto-Uralic vocalic system (Aikio 2022: 5 notes that the modern Finnic languages, in particular Finnish itself, are “almost bizarrely archaic [in that they] often preserve the quality of Proto-Uralic vowels as such”), save for *y > *i and *õ > *o except for Lappish (Sámi), where it became *u (Collinder 1960: 186, 185), and, notable in explaining the non-archaic vowel system of Hungarian, despite it being the core of Uralic, “[i]n Hungarian there has been a

(Aikio 2022: 5); moreover, although in many Uralic languages the vowels can be long or short, this is unclear in Proto-Uralic (Abondolo 2006a: 12):

	front		back	
closed	*i	*ü	*j ⁴²	*u
mid	*e		*o	
open	*ä		*a	

It is probably impossible to make many meaningful comparisons with the Proto-Indo-European vowels, which too suffer from unclear statuses; for instance, the traditionally reconstructed PIE *i, *u, *e, *o may well have been *i, *u, *æ, *â, and the Leiden school claims that there was no *a phoneme in Proto-Indo-European (Matasović

general tendency to lower the tongue: *o > a; *u > o; *y > o, a; *e (sometimes) > ä; *i > e; *ü > ö.” (Collinder 1960: 186).

⁴² Being defined somewhere along the lines of [u]~[i] according to Sammallahti 1988 or [ɤ]~[ɘ]~[ʌ]~[ɜ] according to Aikio (Aikio 202: 5). For Collinder’s *õ, he said: “In the second syllable of non-derived stems, no other vowels occurred in PU or PFU than *a ~ *ä and *õ ~ *e,” and it may be that *õ occurred only in the first syllable (Collinder 1960: 187).

The problems of the morphonological reconstruction

Reconstructing phonemes is a technical matter, akin to reverse-engineering a language using as much evidence as possible from its sister languages, and from typological data. Reconstructing a morpheme is a more nebulous process; its semantics are volatile, its use can

⁴³ In short (in an update on the historical phonology of the Uralic languages as described by Sammallahti 1988), the generally accepted velars and the sibilants in PIE are **k, *k', *k^w, *g, *g', *g^w, *g^h, *g^h, *g^{wh}* and **h₁, *h₂, *h₃*, while in Proto-Uralic they are reconstructed as **k* (voiceless velar stop), **ɣ* (voiced velar fricative) (overlooking the matter of the voiced consonants, which may be reflected in the geminate vs. non-geminate consonants in Anatolian, which may, through Indo-Anatolian, actually point to a state of affairs more in line with Proto-Uralic). Proto-Indo-European had **s*, while Proto-Uralic had **s, *ś, *š* (and the formerly reconstructed **ć* and **č*) (Holopainen 2020: 293). It is accepted by Kortlandt that there were no **k, *g, *g^h* phonemes in PIE, as the opposition between **k, *g, *g^h* and **k^w, *g^w, *g^{wh}* was annulled when **u / *s + *k / *k^w*, and when **k / *k^w + *r / *H*, then **k / *k^w > k* (Kortlandt 2020b: 2); in this context, Holopainen cites several attempts at connecting these two sets of velars and sibilants: Kümmel's 2009 attempts by suggesting the following phonological correspondences: PIE & PU **k^w* and **ko-, *ku-, *k* and **k; *g* and **ŋ; *g* and **k, *ć; *H* and **x(=*ɣ)*, Kümmel 2019 revising this into PIE & PU: **k* and **ć, *k'/*k* and **k, *g* and **k'/*ŋ, *k^w* and **ku, uk, *g^w* and **ŋu, uŋ*. Hyllested's 2009 attempt, which is not Indo-Uralic, but Nostratic was: PIE & PU: **k* and **k, *k'* and **k, *g* and **k, *g^h* and **k, *g^h* and **j* (offering a PIU **g_j [ǰ]*), **k^w* and **ku, *h_{1-3}}* and **k, *g^w, *g^w* and **ć, *č* (PIU **g_w/N*). Zhivlov 2017, also Nostratically, offered this: Nostratic > PIE, PU: KUCV- > KweC-, ku; Cu/üCCV- > Ce(C)C- or K^we(C)C-, ku; Cu/üTV- > CewT- or K^wewT-, Cu; Cu/üRV- > CRew- or K^wRew-, ku and kü; CujV- > Cey- or K^wey-, kuj; CoCCV- > Ce(C)C- or K^we(C)C-, Co (Holopainen 2020: 296). Abondolo warns against any such comparison, and Holopainen demonstrated the wide variety of attempted connections that researchers have been able to come up with in order to square this circle by noting that comparing Proto-Uralic and Proto-Indo-European phonology is a trap: both of their vocalic systems are questionably reconstructed, and the obstruents (stops) in PU, few in number, can be formally compared to Proto-Indo-European roots far too easily and in far too many formally feasible ways. For example a theoretical PU root **kVt(V)-* can easily be compared to some PIE root like **kVt-, *kVd-, *gVd^h-, *g^hVd-, *g^hVd^h-, *g^wVt-, *h₁Vd-* and with other laryngeals, **skV(i?)d-* and with other *s mobilae*, and so on (Abondolo 2006a: 7), which, therefore, is proof of nothing. As to the sibilants, Hyllested 2009 offered PIE **s ~ PU *ć, *ś, *s, *r* (< **VsV*), and Uralic **r* may have arisen from intervocalic rhotacism; Kümmel 2009, 2015 offered, simply, PIE **s ~ PU *s*, PIE **k ~ PU *ć*; and Zhivlov 2017 PIE **s ~ PU *s, *ś, *š, *č* (Holopainen 2020: 296). Bojan Čop demonstrated in 1972, that PIE **g* (**ǵ, *g^w*) correspond to PU **ŋ*, which is in line with the Indo-Uralic hypothesis wherein it is shown that PIE **d* < PIU **[?]n*, PIE **ǵ* < PIU **[?]ŋ*, viz. preglottalised nasals giving glottalised stops (the mediae), viz. PIE **yeg/ǵ-* 'ice' and PU **jāŋe* 'ice' (Kroonen 2019: 111). Kümmel (2019: 124) summarises this state of affairs as Proto-Indo-Uralic **t, *d?*, **d* > Proto-Uralic **t, *t, *n* (*1) or **t*; Proto-Indo-Uralic **t, *d?*, **d* > Proto-Indo-Anatolian **t, *d, *d* > Proto-Anatolian **t: *t, *t*, Proto-Indo-Anatolian **t, *d, *d* > PIE **t, *d, *d*. The laryngeals are a trickier matter, possibly having to do with language contact between Proto-Uralic and Pre-Proto-Tocharian; Kortlandt (2010: 39) concludes that PU **q* (or **x* in Sammallahti's notation) historically behaves just like the PIE laryngeals, as in the case of the much discussed verb PU **toxi / *toye-* 'to give etc.' (e.g. Kortlandt 2010: 389, cf. footnote 17). The almost invariably ultra-stable liquid and nasal consonants (viz. the resonants) provide, as is expected, common ground between PIE and PU; it is, then, perhaps no coincidence, that the PU accusative **-m* and the PIE accusative **-m* case-markers, belong to this category of phonemes (Holopainen 2020: 294), which may undermine its reliability as a correct, or even relevant case morpheme, given that a very large number of Proto-Indo-European proto-phonemes have been associated with a very small number of Proto-Uralic ones, especially since the morpheme structures are quite similar, that is to say monosyllabic (Aikio 2022: 11, 26), all the more true for the phonetically tiny case proto-morphemes, the formally reconstructable homophonous suffixes, also due to the small phoneme inventory of Proto-Uralic (Raun 1988: 556).

be difficult to pinpoint, even in a living language replete with examples. The obvious issue, aside from the problem of time-depth (which is a partially solvable problem, nonetheless), is a matter of methodology. It has been noted that in reconstructing Proto-Uralic (or any other proto-language), one must be wary of emulating Indo-European studies; much as Sanskrit was used by the *Juggrammatiker* to reconstruct a fake Proto-Indo-Indo-European, thusly, it has been noted, Finnish was relied upon in excess to create an unscientific Proto-Uralic⁴⁴ – cherry-picking is a bane on science, and this is in full view in historical linguistics as seen in these three examples. The other issue here is that we can only make use of what we know, as in the case of Old Church Slav(on)ic in relation to Proto-Slavic; completely different conclusions in reconstructing Proto-Slavic and Common Slavic would have been made had we not had any texts in Old Church Slav(on)ic (Katičić 1970: 139-140). Thus it is good to keep in mind that the conclusions we come to depend on the material we have, not knowing how much we are missing. Naturally, the ideal of linguistics emulating the exact sciences is attained either rarely or, according to one strict perspective, never, by the very nature not just of the methodology. Methodology is highly open to modes and tendencies in academia, ideological (few academic subjects are more ideologically fraught than Indo-European studies, which has since calmed down in Europe for the most part, but is still raging in India (Pereltsvaig, Lewis 2015: 19)) or personal (the transnational study history may be a positive a step in that direction, as well as the consciousness of the structuralism of one's theories, as in Bomhard's work), as well as to some sharp differences in analysis, which can be open to quite subjective assessments of a given problem. Weaknesses in reconstructing due to faulty methodology and unexamined assumptions about how conservative certain languages or branches may be, and about language as such, can be a hindrance which takes generations to unravel (such as the above-mentioned case of overstating the archaism of Finnish with regard to Proto-Uralic).

This is to say nothing of the various schools of thought among linguists and universities; the Indo-Uralic hypothesis has a long history, with the Indo-Europeanists generally being in favour of the hypothesis (Kümmel 2009, 2019, Kortlandt 1989, 2010, Kloekhorst 2008 among others), with Clackson 2007 being an exception. In other words, this is a split in reflecting schools of thought regarding the reliability of the historical method in reconstructing protophonemes; as usual, the Moscow School is known for generating (or resuscitating) Nostraticism, the Leiden School for (re)constructing Indo-Uralicism, and the

⁴⁴ Quoting Austerlitz 1993: 291, “the choice of Finnish as the Finno-Ugric prototype may have impeded progress in the field (Austerlitz 1993: 291)...” (Abondolo 2006b: 149)

Finnish School (at the Universities of Helsinki, Turku and Oulu), if it may be called that, is on the whole much too rigorous to accept the abundance of possibilities in proving genetic relationships between the major languages families of (northern) Eurasia. While Abondolo or Aikio underline the differences between Proto-Uralic and Proto-Indo-European (such as the fact that Proto-Uralic and Proto-Indo-European morphological typologies are *vastly* different (Abondolo 2006a: 7)), Kortlandt or Lubotsky tend to underline the similarities. This divide stems, at least in part, from how these schools of thought conceptualise the protolanguages: as abstractions, or as real languages which are being uncovered. This chasm seems to derive from the extent to which various researchers judge historical phonology or historical morphosyntax to be reliable; for instance, Sammallahti takes the view that historical phonology is the ideal tool for establishing genetic relationships, in determining linguistic families, considering historical morphology as untrustworthy, as borrowing and analogising can take place, as well as typological similarities which can be misleadingly similar (Sammallahti 1988: 478); others, on the other hand, take historical grammar, not phonology, for measuring language families or macro-families, to be the most reliable method, such as the oft-mentioned above Pereltsvaig. It is not uncommon for researchers in historical linguistics to admit, as does Abondolo (2006a: 8), that “our present primitive understanding of the mechanisms of language changes” are insufficient to resolve many problems in historical linguistics, as can be amply seen when dealing with macro-families, where *widely* differing views are on offer.

There is serious disagreement as to the number of cases reconstructable in Proto-Uralic, and to how they came to be (Raun 1988: 558); the etymologies of the postpositions in Proto-Uralic, and of the prepositions in Proto-Indo-European, are almost impossible to determine. It is worth noting that Collinder’s reconstruction is dependant upon his Proto-Finno-Ugric set of cases, which he proposes had the same set of cases as the Proto-Uralic ones, adding there was also in Proto-Finno-Ugric also “(probably) [a] lative-prolative. There was perhaps also a dative-lative or dative-terminative” (Collinder 1960: 238). This is, of course, in opposition to how many Uralicists today deny a Finno-Ugric branch (in greater detail: Collinder 1960: 238-239). Abondolo, in the introduction to *The Uralic Languages* (2006), proposes that Proto-Uralic had *at least two* cases in addition to *at least three local* cases (Abondolo 2006a: 18), viz. an accusative **-m* for the definite direct object of finite verbs (inflected verbs) for person, and a **-n* for the genitive, serving also as a subordinate suffix, as a pronominaliser with nouns, and an adverb

formant with verbs, not counting the zero-marked nominative as a case; as well as a locative *-nA, a separative *-tA ~*-tI and perhaps the two latives *-k (and/or *-ŋ) and *-cʲ (and/or *-nʲ). This is generally in accordance with the earliest serious attempt at reconstructing the PU case inventory (Collinder 1960: 282), Proto-Uralic had the following *five* case morphemes: nominative *-∅ as a zero-morpheme, genitive *-n, accusative *-m, locative *-na ~ *-nä, separative *-ta ~ *-tä (in greater detail: Collinder 1960: 282-297). Aikio (2022) generally agrees with these reconstructions, notably enlarging on the problem of the latives (see above).

Assessing how realistic reconstructing Proto-Uralic is, with regard to Proto-Indo-Uralic

No reconstruction of a proto-language can be complete; only a coherent, to a reasonable extent *whole* system can illuminate how a given unattested (or fragmentarily attested) language actually worked grammatically. Reconstructing languages thus has its striking limitations, which also stem from the material and methods used to reconstruct them, that is to say the daughter languages and, when applicable, their ancestor languages. Proto-Uralic is no exception; only elements of this proto-language can be reconstructed, not even close to being the whole system (Raun 1988: 555). Grammatical cases, reconstructed morphosyntactically after a sober assessment of the morphonological reconstruction, show that the syntax of a substrate language is largely *not* preserved when languages mix with adstrates (syntax is how *meaning is organised* in a sentence), and can thus, even in an event of language mixing, be of greater use in reconstructing languages at great time depths. If grammar is the common name of morphology and syntax, and if morphology and syntax are merely different perspectives of the same phenomenon (meaning-making and -arrangement) in language, then the the system of the nominal phrase is quintessential in understanding Proto-Uralic (and, as a consequence, also its relationship to other proto-languages).

One of the purposes of this thesis is to ascertain, after having formed a coherent image of the state of research on the Proto-Uralic case system, to what extent it is plausible to connect Proto-Indo-European and Proto-Uralic using today's methodology in historical linguistics – in other words, how plausible is the Indo-Uralic hypothesis, the macrofamily deemed to be the most likely candidate to be the ancestor to Proto-Uralic in mainstream historical linguistics. When attempting to find a common denominator to Proto-Uralic and Proto-Indo-European, the most obvious issue to solve is whether the speakers of these

protolanguages were in contact with each other, for their speakers to have sprung out from a common archeological culture and protolanguage. Nodal events in prehistory (such as spreading events, bottle-necks in population movements) are difficult to detect, and pinpointing them has been the preoccupation of those researchers trying to match historical linguistics with archeology and archaeogenetics, which has become an increasingly serious interdisciplinary field since about 2015⁴⁵.

Those attempting to prove the Indo-Uralic hypothesis (as in the 2019 collection of papers *The Precursors of Proto-Indo-European: the Indo-Anatolian and Indo-Uralic Hypotheses* [ed. by Alwin Kloekhorst and Tijmen Pronk]) have a strong case to make, especially since Asko Parpola's 2022 paper (an expansion of his and Carpelan's 2001 well-received paper (Parpola 2022: 259)) which concludes that the Volga-Kama culture [5700-5500 BCE] on the Lower Kama river basin (a river in between the Volga and the Ural mountains) appears to be the furthest reachable phase of the Uralic language family, namely Pre-Proto-Uralic (Parpola 2022: 261). To note, despite this, Parpola is not a proponent of the Indo-Uralic hypothesis, and on lexical grounds: he cites Proto-Uralic lemmas such as **meti* 'honey' and **mekši* 'bee' as loanwords from Pre-Proto-Indo-Iranian (himself citing Aikio 2022), as well as the Proto-Uralic lemma for the Siberian pine tree (*Pinus (cembra) sibirica*), namely PU **sīksa*, a tree which grows in Siberia and, in Europe, only up to the above-mentioned Kama Valley (Parpola 2022: 260) which, archaeobotanically, may not have in fact been around the Kama Valley at this time (cf. Grünthal et al 2022: 497). In other

⁴⁵ The Indo-Uralic hypothesis hinges not only on historical grammar, but on how that is relevant to archeological and archaeogenetic data; if the Indo-European and Uralic homelands cannot be pinpointed in such a way that they would both derive from a common archaeological culture (and a common proto-language – Indo-Uralic). Which side of the Ural mountains the Uralic homeland was supposed to be on has been debated almost as much as where the Indo-European homeland was to be located. The argument for the Uralic homeland being east of the Urals, viz. deep in Siberia refers to the supposed, and extremely early in Uralic studies conception of the Finno-Ugric vs. Samoyed split. Carpelan & Parpola 2001 are in favour of a homeland in European Russia, largely based on Proto-Uralic arbonyms, such as *Pinus cembra/Pinus sibirica* **sīksa* (a well-attested phytonym which, nonetheless, is a cause of friction among Uralicists) respectively, identifying the Pit-Comb Ware (or Comb Ceramic) Culture, well west of the Urals stretching onto the northern Baltic Sea coast, as the Uralic homeland. On the other hand, Parpola's 2022 identifying the Kama River basin as the Uralic homeland goes hand in hand with Kortlandt's (Kortlandt's 2010: 37) identification of Proto-Indo-European as an Indo-Uralic branch which has moved from the Caspian Sea coast to the Black Sea coast and was heavily influenced by a North Caucasian substratum language; the linguistic ancestors of Abkhaz, Adyghe and Kabardian, viz. Proto-North-West-Caucasian must have been in contact with Proto-Indo-European (Pereltsvaig, Lewis 2015: 198); it has been deduced, rather than proven that the North-West-Caucasian and the North-East-Caucasian languages pushed uphill due to various invaders; their homelands must've been lower down, in the same steppe where Proto-Indo-European was (Pereltsvaig, Lewis 2015: 198). In any case, Early Proto-Indo-European has been associated with a number of archaeological cultures, but Late Proto-Indo-European must correspond to the vast and expansive Yamnaya culture; this theory has become doctrine. It is, nonetheless, judged that it is possible that Proto-Indo-European was south of the Black Sea and Proto-Uralic was east of the Urals, and they *met* in European Russia (Pereltsvaig, Lewis 2015: 201), although no consensus (save for Late Proto-Indo-European corresponding to the Yamnaya archeological culture) has been established.

words, Parpola (an archeologically-literate linguist) seems to belong to Aikio’s and Holopainen’s school of thought. In any case, precisely given its reconstructable vocabulary, “Proto-Uralic is generally accepted as having been a language of foragers living in the forests to the north of the Pontic Steppes, who, as mentioned above, had no domesticated animals other than dogs” (Pereltsvaig, Lewis: 200). Parpola also cites Martin Kümmel in support of the Kama Valley hypothesis as the Uralic homeland, noting “significant parallels in personal pronouns, three different demonstrative pronouns, two interrogative pronouns, one relative pronoun, suffixes of dual, plural, accusative, ablative and instrumental, and 19 lexical comparisons” (Parpola 2022: 260). He adds that Proto-Indo-Uralic is improbable for *typological* reasons (citing Janhunen 2001: 211), that is to say that at the earliest stages of reconstruction for the Proto-Uralic and Proto-Indo-European, the earliest reconstructable forms of the two protolanguages could not be more different typologically – failing, however, to mention that they are abstract constructs, in all likelihood being multiple, unravelled layers of languages yet to be sieved apart; he thus takes those two protolanguages are being concrete, if fragmentary, languages.

Proto-Indo-Uralic, if not by the cases, may be correct by way of lexical correspondences; if the case-ending cannot prove it, then maybe the nominals themselves, the roots, may. Kortlandt 1989 considers verbs taken (by Holopainen) to be of Proto-Indo-European origin borrowed into Proto-Uralic as actually being Proto-Indo-Uralic lemmas, e.g. such elementary verbs as “to give”, “to wash”, “to bring”, “to drive”, “to do”, “to lead”, “to take” (Pereltsvaig, Lewis 2015: 199), offering several candidates for this. Namely, eight verbs (one presents problems to this point of view, see below) in Proto-Uralic which, since they are very basic verbs (they are from Károly Rédei’s list of PU-PIE cognates, hence the German definitions of the lemmas below) (Kortlandt 2010: 388-389). Thus he offers seven tentatively reconstructed PIU verbs (Kortlandt 2010: 389)⁴⁶. To this Guus

⁴⁶ PU **miye-* ‘geben, verkaufen’ (‘to give, to sell’) (Finnish *myy-* or *myö-*), and PIE **mei/y-* (Latvian *mīt*), thus PIU **miye-* (‘to give, to sell’); PU **muške-* or **moške-* ‘waschen’ (‘to wash’) (Estonian *mõske-*), PIE **mesg-*, (Lithuanian *mazgóti*), thus PIU **muske-* (‘to wash’); PU **toye-* ‘bringen, holen, geben’ (‘bring, to fetch/to pick up, to give’) (Finnish *tuo-*), and PIE **deh₃-*, (Hittite *dā-* ‘take’) (this verb, reconstructed in other sources as **toxi*, in particular has attracted attention for supposedly proving that a PIE laryngeal corresponds to a PU **x*), thus PIU **tagu-* (‘bring, to fetch/to pick up, to give’); PFU (Proto-Finno-Ugric which, as mentioned above, can be equated with Proto-Uralic for taxonomic reasons) **aja-* ‘treiben, jagen’ (‘to drive (away), to chase (away)’) (Finnish *aja-*), and PIE **h₂eǵ-*, (Latin *agere*), thus PIU **gaki-* (‘to drive (away), to chase (away)’); PFU **teke-* ‘tun, machen’ (‘to do, to make’) (Finnish *teke-*, Hungarian *tē(v)-*), and PIE **d^heh₁-*, (Latin *facere*), thus PIU **deka-* (‘to do, to make’); PFU **wetä-* ‘führen, leiten, ziehen’ (‘to lead, to guide, to pull’) (Finnish *vetä-*, Hungarian *vezet-*), and PIE *wed^h-*, (Lithuanian *vèsti*), thus PIU **weda-* (‘to lead, to guide, to pull’); PFU **wiye-* ‘nehmen, tragen’ (‘to take, to carry’) (Finnish *vie-*, Hungarian *vi(v)-*, *visz-*, *vē(v)-*, *vész-*), and PIE *weg^h-*, (Lithuanian *vèžti*), thus PIU **wige-* (‘to take, to carry’); PFU **kan-* ‘streuen, schütten, werfen, graben’ (‘to sprinkle, to pour, to throw, to dig’) (Hungarian *hány-*), and PIE **kh₂en-*, (Sanskrit *khánati* ‘digs’); Kortlandt

Kroonen (2019: 111-112) adds, in line with Kümmel 2012, five PIU etymologies⁴⁷. To this list, Martin Kümmel (2019: 123, 127-128) offers many lexical correspondences between PIA/PIE and PU/PFU (alongside the aforementioned PIE **wed-* ‘to lead’ and PU **wetä-*), refraining from offering PIU forms as such due to the obvious problem of the phonological correspondences, cf. Holopainen 2020⁴⁸. However, according to Sampsa Holopainen’s doctoral dissertation (*Indo-Iranian borrowings in Uralic: Critical overview of the sound substitutions and distribution criterion*, 2019), there are *no* common Proto-Indo-Uralic lexemes, only borrowings from Pre-Proto-Indo-Iranian, which has been linked to the Sintasha culture (2010-1770 calBC) (Parpola 2022: 261), into Proto-Uralic, on phonological grounds. For instance, Holopainen criticises many words which are supposed to be common to PIE and PU, finding it troubling that only *some* correspondences can be established (Holopainen 2020: 295, 297-298; cf. Ruppel 2016 on competing loan-etymologies, the statistical approach by Kassian et al 2015 and the correct historical-comparative method by Hyllested 2019)⁴⁹. Holopainen’s reasoning being that the consonant correspondences do not match up, such as PIE **k*’ and PU **ć* in satəm-language loanwords. Some other possible correspondences would be PIE **r* and PU **δ* (a lateral) and **δ*’, not to mention that **ŋ* itself is extremely problematic (Holopainen 2020: 297), summarising, and concluding that there are *no* sound-correspondences in the velars and the sibilants between PIE and PU (noting that the

excludes this PFU **kan-* from a PIU origin on the grounds that it is not a basic verb and may, in fact, have been borrowed (Kortlandt 2010: 389), thus eliminating it from Rédei’s list.

⁴⁷ PU **äŋ-* ‘burn’ (only Hungarian *ég* and Komi *jŋ*), PIE **ng^w-ni-* ‘fire’, thus PIU **Vn^ŋV-*; PU **näki* ‘see’, PIE **dek-* ‘perceive’, thus PIU **^ŋnVkV-*; PU **ñi(ŋ)kćimi* ‘palate, gills’ (the **(ŋ)* is problematic as it is based only on the Mari reflex; possibly from PU **ñi(ŋ)kV* ‘?’ + PU **ćimi* ‘scales, fish skin’), PIE **dŋǵ^h-* ‘tongue’, thus PIU **^ŋVŋgV-*; PU **jäŋi* ‘ice’, PIE **yeg/ǵ-* ‘ice’, thus PIU **jV^ŋV-*; PU **pane* ‘put, place’, PIE **ped-* ‘step; fall’, thus PIU **pV^ŋnV-* ‘lay down, put down?’.

⁴⁸ PU (or PFU, Proto-Finno-Ugric) **metə* ‘honey’ (Finnish *mesi*, Hungarian *méz*, *méze-*) and PIE (or PIA, Proto-Indo-Anatolian) **medu-* ‘mead, honey’; PFU **im-ta-* ‘to give’ (Finnish *anta-*, Hungarian *ad*) and PIA or PIE **h₁em-* ‘to take’; PFU **juxə-* ‘to drink’ (Finnish *juo-*, Hungarian *iv-*) and PIA or PIE **Hjuh₂-* ‘to eat’; PU or PFU **juta-* ‘to wander, travel’ (Mordva *juta-*, Nenets *jādā*) and PIA or PIE **Hjud-* ‘to move’; PU or PFU **tekə-* ‘to do, make’ (Finnish *teke*, Hungarian *tev-*) and PIA or PIE **deh₁-* ‘to do, put’; PU or PFU **tekV-* ‘to push’ (Hungarian *tűz-*) and PIA or PIE **teq-* ‘to weave’; PU or PFU **ćorwa* ‘horn’ (Finnish *sarvi*, Hungarian *szarv*) and PIA or PIE **kerw-/kru-* ‘horn’; PU or PFU **kätə* ‘hand’ (Finnish *käsi*, Hungarian *kéz*, *keze-*) and PIA or PIE **ges-* ‘hand’; PU or PFU **kälä-* ‘to wade’ (Hungarian *kel-*) and PIA or PIE **qel_h1-* ‘to bring, impel’; PU or PFU **kaďa-* ‘to let, leave’ (Hungarian *hagy*) and PIA or PIE **qel_h1-* ‘impel’; PU or PFU **kokə-* ‘to see’ (Finnish *koke-*) and PIA or PIE **k^wek-* ‘to look’; PU or PFU **kulkə-* ‘to move’ (Hungarian *halad*) and PIA or PIE **k^welkh₁-* ‘to turn’; PU or PFU **lamtə* ‘low’ (Finnish *lansi*) and PIA or PIE **leNd-* ‘low’; PU **niŋä* ‘woman’ (Hungarian *nő*, Khanty *ni(ŋ)*) and PIE **neg(ŋ-o-)* ‘sister’ (only attested in Anatolian).

⁴⁹ PU *päŋi* ‘head’ (Finnish *pää*, Hungarian *fő*) and PIE **b^heg-* ‘to break’ (Vedic *bhánakti*); PUg (Proto-Ugric) **anća* ‘groin’ (Hungarian *ágyék*, Mansi **ünš*) viz. PU **ońća* ‘part’ (Finnish *osa*, Mari *užas*, North Sámi *oažži*), and PIE **ang^w-*, **neg^w-* = **ŋg^w-en* ‘groin’ (Latin *inguen*, Ancient Greek *ἄδην*), saying that it is a Proto-Indo-Iranian loanword (**(H)anća*) into Proto-Uralic; PU **ćidä-mə* (UEW **ćV(w)ďä(-mi)*) ‘heart’ (Finnish *sydän*, Hungarian *szív*) PIE **k^werd-* ‘heart’ (Hittite *ker*, Ancient Greek *κῆρ* &c.); PU **künći* (UEW **künče*) ‘fingernail’ (Finnish *kynsi*, Mari **küč*, Hungarian *köröm*, assuming a PU **nč* > Hungarian *r*) and PIE **h₃eng^w-* ‘claw’ (Latin *unguis*, Old Irish *ingen*, Ancient Greek *ὄνυξ*); PU **suŋi* ‘summer’ and PIE **sm^hH-* ‘summer’.

time-depth may be too significant), thus the idea that PIE *k^w and PU *k/u point to some PIU *ku is, it would seem, invalidated by Holopainen (Holopainen 2020: 314-315). He adds, too, that the PIU reconstructions as offered by those who have attempted to give form to them pay no attention to vowels, which amounts to cherry-picking when reconstructing (Holopainen 2020: 298; for a critical discussion of this, cf. Simon 2020, who claims that few Proto-Indo-European loan etymologies are convincing, as most are lookalikes or borrowed from later Proto-Indo-European into Proto-Finno-Ugric).

Morphology is, probably, more indicative of genetic relatedness than phonology is; although morphology can be deceiving due to typological similarities (perhaps, because, for instance, Tocharian has an Altaic- and Uralic-like case system (which may be due to substrate influence after all, cf. Peyrot 2019a)) phonology may be more so due to its inherently volatile nature; phonology is a branch of linguistics which rests on the assumption of the regularity of sound change, explaining away as best it can its exceptions and approximations. Sound-laws can kick morphological systems out of place, which may or may not be analogically reshaped, or even restored after loss, and only parts of such processes can be deduced with the material available; this is why great time-depth puts into question any and all attempts at linguistic reconstruction. In this sense, historical linguistics seems to be a description of a paramathematical system, para- because it only gives the appearance of being regular; this poses a vast problem to falsifiability, the *sine non qua* of all science, absent from historical linguistics. As Sturtevant explains it, the paradox is that sound change (and variation is simultaneous across multiple phonemes) is regular and causes irregularities in the system, whereas analogy is irregular and gives rise to regularity, viz. simplification in the system (Antilla 1972: 94); thus the essential difference between Nostraticists and non-Nostraticists is how seriously they take sound-correspondences and the regularity of sound-change to be, as discussed above. Irregularities within a given language or protolanguage reflect abandoned systems; aberrant forms are what hints at layers in languages; nevertheless, this is precisely how de Saussure's coefficient sonantique was uncovered, and then confirmed (Hyllested 2009: 113). Languages are indiscrete units, which means that reconstructed languages are even less discrete than living languages, because they are constructs, not slices of time; thus linguistics gets away with vague definitions of language, of dialects etc. (cf. Makoni and Pennycook 2006). Reconstructed protolanguages are fragments of multiple layers of unattested languages, showing all the difficulties of studying living languages, multiplied.

Change is a function of time, and it is unequal everywhere. As has been explained forcefully in Koch 1996, reconstructing historical morphology is, in terms of methodology, a disorienting task, noting, simply, that the best one can do is to “[a]ssess the relative likelihood of each of the compared forms and/or paradigmatic patterns being archaic or innovative” (Koch 1996: 220). Reconstructing historical phonology has temporal limits, after which the historical-comparative method begins to make more *assumptions* than is scientifically permissible (viz., evidence becomes decreasingly relevant to ambitious researchers) – a very sober look at this issue with regard to Uralic is Georg 2022. Explaining away what stands in the way of a hypothesis or theory, in this case the Indo-Uralic hypothesis, increasingly looks like a contest in linguistic creativity, thinking up ways to dispel the obvious obstacles to this hypothesis. The more one does this, the less scientific the whole process becomes, proportionately becoming *undisprovable*, therefore *unscientific*, which is the principle complaint of the Finnish, anti-Indo-Uralic school of thought. Comparative reconstruction is reverse-engineering linguistic change, it is linguistic palaeontology; unravelling allophonic variation, analogy, generalisation, simplification and plain similarity is not a straightforward task, hence the varying results produced by researchers which so readily reveal that, oftentimes, they come to the conclusions they *want* to come to.

Conclusion

In reconstructing languages, it is imperative to keep in mind that all languages are always in flux, and that we only analyse slices. More often than not, slices of standardised forms, which are thus made studyable by being standardised, give a false impression of immutability. Various layers and types of languages (literary languages, standardised languages, colloquial forms, dialects) make reconstruction more complicated, but almost more correct – languages decay, like radioactive atoms, perpetually, and they do not have fixed versions of themselves; moreover, just *how* they decay is only partially predictable, and only approximately reconstructable. This introduces the natural problem of quantity, whose infinity would be ideal for the historical study of languages, but which is near impossible to make use of, even in computer-based attempts made so far. Newer literature in Uralic studies makes a visibly increased use of the smaller Uralic languages, and pays particular attention to truly differentiating them (recognising that there are multi Mansi, Komi, Sámi and Samoyed languages). And, the necessary oversimplification of the corpus, of the subject for the sake of

science; in modern Indo-European studies one must make use of as many of its daughter languages as possible, as useful – thusly with Uralic, too, but this is seldom done (more or less until Aikio, that is; he is incredibly thorough with his Uralic corpus, when compared to Raun or Sammallahti, or even Abondolo or Janhunen). It may be that *oversimplification* produces the very rifts which have split this small academic community into Leidenists, Moscovites and the Finnish Uralicists, as Holopainen has laid it out. Naturally, it would be an easy mistake to overestimate the usefulness of the two Uralic languages which happen to have the greatest numbers of speakers; but this isn't really a problem, as their time-depth makes them, in fact, little more useful than Udmurt, or Khanty, or even Samoyed in reconstructing the case system. In fact, few before Aikio (of course, Collinder, Rédei and Sammallahti excepted, as well as Abondolo) have taken the Uralic languages other than the two major ones – Estonian is taken to be of little use here, as it is, like some Sámi variants, a flecational language, as well as being lexically and in terms of the case inventory sufficiently similar to Finnish to be overshadowed by it (Kittilä, Laakso, Ylikoski 2022: 882) – into serious consideration. Evidently, much higher accuracy can be achieved when reconstructing a language whose predecessor and daughter languages are not only attested, but well attested, even if through reconstruction. Thus we can hint at, with a series of reasonable applications of the historical method, at Illyrian, despite having next to no corpus in it; but, we have Albanian on the one side, with some loanwords in several Balkan languages, and a panoply of ancient Indo-European languages on the other which can be used to bridge the gap; in a similar fashion, Proto-Indo-European is a sort of historical dead-end, since its predecessor is little known – unless we accept, as the Leiden school insists (cf. Kloekhorst, Pronk 2019), Proto-Indo-Anatolian. Similarly, Proto-Balto-Slavic can be quite accurately reconstructed because we can approach it from both sides. If we agree to give credence to Proto-Indo-Anatolian as the precursor to Proto-Indo-European, we can begin to approach Proto-Indo-Uralic – *if* we take one of the reconstructions/abstractions of Proto-Uralic to be scientifically usable⁵⁰.

Thus, in principle, the quantity and quality of the data determines the quality of the reconstruction, if the comparative method is applied soundly. Reconstructing Proto-Indo-Uralic is, for now, deemed next to impossible as it is riddled with uncertainties on

⁵⁰ Kallio 2019 traces the history of trying to find the common ancestor to Proto-Indo-European and Proto-Uralic, starting from the Finnish priest Nils Idman first compared the pronouns in Finnish and Ancient Greek in 1774, noticing the similarities (Kallio 2019: 75). Indo-European was first noticed by the Dutch scholar Marcuz Zuerius van Boxhorn in 1647, and Uralic (in point of fact, only Finnish-Hungarian relatedness) by János Sajnovics (1770) and Sámuel Gyarmati (1799) (Kallio 2019: 75).

all fronts (lexical, phonological, morphological and syntactic) (Kloekhorst, Pronk 2019: 10). Nonetheless, it is being reconstructed by Kortlandt, as had been already started by Illich-Svitych (for Proto-Nostratic) and, building on that material, Bojan Čop (for Proto-Eurasian) (Klemenčič 2019: 88). Proto-Indo-Uralic is deemed highly probable, but not yet scientifically demonstrable (Matasović 2012: 3).

In terms of cases, this is difficult to do. The three grammatical cases are probably insufficient, strictly speaking, to prove Indo-Uralic, as the unmarked nominative is a typologically common phenomenon, and the accusative and genitive case endings are both ultrastable nasal consonants. It is true that the accusative singular case morphemes in Proto-Uralic and Proto-Indo-European seem to be the same, but this is an ultrastable morphemes; the same is said of the genitive in Proto-Uralic. As to the adverbial cases, save for the Proto-Uralic separative which *seems* to hint at a Proto-Indo-European *o*-stem ablative, form a ‘lative paradigm’ which leaves us in the dark as far as the number of them in Proto-Uralic, and the Proto-Indo-European cases other than the nominative, accusative and genitive seem to have little to do with any Proto-Uralic ones, to say nothing of the Proto-Indo-European plural case endings, which are mostly of prepositional origin. None of the reconstructed Proto-Uralic case morphemes are reconstructed without some degree of doubt, even the PU accusative and the genitive cases, due precisely to the relatively small number of reflexes, and the ultrastable nature of those nasal phonemes, let alone the directional cases, which can formally be reconstructed *ad nauseam* into Proto-Uralic. On top of that, the etymology of these monosyllabic morphemes, whose vowels, both in Proto-Uralic and Proto-Indo-European, cannot be reconstructed with much certainty at all, seem to point at the unfortunate conclusion that the data is insufficient. There is not doubt that Proto-Indo-European and Proto-Uralic stem from a common protolanguage, but what this protolanguage was (Proto-Indo-Uralic, Proto-Nostratic, Proto-Eurasian) is open to debate, and it is being debated.

References

- ___ Daniel ABONDOLO 2006a [first edition 1998] *Introduction* in *The Uralic Languages* [ed. Daniel Abondolo]. London, New York: Routledge. 1-42.
- ___ Daniel ABONDOLO 2006b [first edition 1998] *Finnish* in *The Uralic Languages* [ed. Daniel Abondolo]. London, New York: Routledge. 149-183.
- ___ Jaakko ANHAVA 2010 *Criteria for Case Forms in Finnish and Hungarian Grammars* in *Studia Orientalia* 108. 239-244.
- ___ Luobbal Sámmol Sámmol ANTE (Ante AIKIO) 2022 *Proto-Uralic* in *The Oxford Guide to the Uralic Languages* [ed. Marianne Bakró-Nagy, Johanna Laakso, Elena Skribnik]. Oxford: Oxford University Press. 3-27.
- ___ Raimo ANTTILA 1972 *An Introduction to Historical and Comparative Linguistics*. New York: The MacMillan Company.
- ___ Ante AIKIO 2012 *An Essay on Sámi Ethnolinguistic Prehistory* in *Linguistic map of Prehistoric Northern Europe* [ed. Riho Grünthal, Petri Kallio]. Helsinki: Mémoires de la Société Finno-Ougrienne 266. 63-117.
- ___ Ante AIKIO 2014 *The Uralic-Yukaghir lexical correspondences: genetic inheritance, language contact or chance resemblance?* in *Finnisch-Ugrische Forschungen*. 7-76.
- ___ Ante AIKIO 2020 *Uralic Etymological Dictionary (draft version of entries A-Ć)*. Available online at https://www.academia.edu/41659514/URALIC_ETYMOLOGICAL_DICTIONARY_draft_version_of_entries_A_Ć Accessed on 2023-06-22.
- ___ Robert AUSTERLITZ 1993 *Agglutination, demography, Finnish, Turkic* in *Hajdú Péter 70 éves [Festschrift for Péter Hajdú]* [ed. M. Sz. Bakró-Nagy, E. Szij], *Linguistica, series A, Studiae et Dissertationes* 15. Budapest: MTA Nyelvtudomány intézet.
- ___ Allan R. BOMHARD 2019 *The Origins of Proto-Indo-European: The Caucasian Substrate Hypothesis* in *Journal of Indo-European Studies, Volume 47, Number 1 & 2, Spring/Summer 2019*. 1-87.
- ___ Rasmus Gudmundsen BJØRN 2019 *Pronouns and Particles: Indo-Uralic Heritage and Convergence* in *The Precursors of Proto-Indo-European: The Indo-Anatolian and Indo-Uralic Hypotheses* [edited by Alwin Kloekhorst, Tijmen Pronk]. Leiden, Boston: Brill, Rodopi. 30-49
- ___ Lyle CAMPBELL 2015 *Do Languages and Genes Correlate?: Some Methodological*

Issues in Language Dynamics and Change 5. 202-226.

___ James CLACKSON 2007 *Indo-European Linguistics: An Introduction*. Cambridge: Cambridge Textbooks in Linguistics.

___ Björn COLLINDER 1960 *Comparative Grammar of the Uralic Languages*. Stockholm: Almqvist & Wiksell.

___ Merlijn DE SMIT 2015 *Proto-Uralic Ergativity Reconsidered* in *FUM* 38. 1-34.

___ David M. EBERHARD, Gary F. SIMONS, and Charles D. FENNIG (eds.) 2023 *Ethnologue: Languages of the World*. Twenty-sixth edition. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>, accessed on 2023-06-22.

___ Stefan GEORG 2022 *On Perfectly Good-Looking Morphological Comparanda and Their (Sometimes, However, Lacking) Significance for Hypotheses of Language Relationship: Some Marginal Footnotes on the (Still Ongoing?) Altaic Debate in Historical Linguistics and Philology of Central Asia: Essays in Turkic and Mongolic Studies* [Editor Bayarma Khabtagaeva]. Leiden: Brill. 420-428.

___ Riho GRÜNTAL, Volker HEYD, Sampsa HOLOPAINEN, Juha A. JANHUNEN, Olesya KHANINA, Matti MIESTAMO, Johanna NICHOLS, Janne SAARIKIVI, Kaius SINNEMÄKI 2022 *Drastic demographic events triggered the Uralic spread* in *Diachronica*, Volume 39, Issue 4, Aug 2022. 490-524.

___ Péter HAJDÚ 1983 *The Main Characteristic Features of the Uralic Languages* in *Acta Linguistica Academiae Scientiarum Hungaricae* Vol. 33 no. 1 / 4. Budapest: Akadémiai Kiadó. 101-112.

___ Jaakko HÄKKINEN 2007 *Kantauralin murteutuminen vokaalivastaavuuksien valossa*. Pro gradu tutkimus Suomalais-ugrilainen kielentutkimus Suomalais-ugrilainen laitos Helsingin yliopisto.

___ Harald HAMMARSTRÖM, Robert FORKEL, Martin HASPELMATH, Sebastian BANK 2022 *Glottolog* 4.7. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://doi.org/10.5281/zenodo.7398962> (Available online at <http://glottolog.org>, accessed on 2023-06-22)

___ Bernd HEINE 2009 *Grammaticalization of Cases* in *The Oxford Handbook of Cases* [ed. Andrej Malchukov, Andrew Spencer]. Oxford: Oxford University Press. 458-469.

___ Harry A. HOFFNER, Jr., H. Craig MELCHERT 2008 *A Grammar of the Hittite Language: Part 1: Reference Grammar*. Winona Lake, Indiana: Eisenbrauns.

___ Sampsa HOLOPAINEN 2019 *Indo-Iranian borrowings in Uralic: Critical overview of the sound substitutions and distribution criterion*. University of Helsinki, Faculty of Arts.

Available online at <https://core.ac.uk/download/pdf/245131352.pdf> Accessed on 2023-06-22

___ Sampsa HOLOPAINEN 2020 *Criticism of some Indo-Uralic sound-correspondences in 2020 Vácaāṁsi miśrā kṛṇāvmahai: proceedings of the international conference of the Society for Indo-European Studies and IWoBA XII, Ljubljana 4-7 June 2019, celebrating one hundred years of Indo-European comparative linguistics at the University of Ljubljana = Akten der Arbeitstagung der Indogermanischen Gesellschaft und des IWoBA XII vom 4. bis 7. Juni 2019 in Ljubljana aus Anlass der Hundertjahrfeier der dortigen Indogermanistik* [edited by Luka Repanšek, Harald Bichlmeier & Velizar Sadovski]. Hamburg: Baar-Verlag. 293-320.

___ Sampsa HOLOPAINEN, Jussi YLIKOSKI 2022 *On 'on' in Ossetic and Uralic: from adpositions to case suffixes in Siberica et Uralica: In memoriam Eugen Helinski* [ed. Valentin Gusev, Anna Urmanchieva, Aleksandr Anikin]. Szeged: University of Szeged, Department of Altaic Studies, Department of Finno-Ugrian Philology. 15-37.

___ Tuomas HUMMO, Krista OJUTKANGAS 2006 *An Introduction to Finnish Spatial Relations: Local Cases and Adpositions in Grammar from the Human Perspective: Case, Space and Person in Finnish* [ed. Marja-Liisa Helasvuo, Lyle Campbell]. Amsterdam, Philadelphia: John Benjamins Publishing Company.

___ Adam HYLLESTED 2009 *Internal reconstruction vs. external comparison: the case of the Indo-Uralic laryngeals*. in *Internal reconstruction in Indo-European : methods, results, and problems : section papers from the XVI International Conference on Historical Linguistics, University of Copenhagen, 11th-15th August, 2003* [ed. Jens Elmegård Rasmussen, Thomas Olander; with the collaboration of Anders Richardt Jørgensen]. Copenhagen : Museum Tusulanum Press.

___ Juha JANHUNEN 1999 *Critical Remarks on Current Indo-Uralic Comparison in Europa et Sibiria: Beiträge zu Sprache und Kultur der kleineren finnougri-schen, samojedischen und paläosibirischen Völker. Gedenkband für Wolfgang Veenker* [ed. Cornelius Hasselblatt, Paula Jääsalmi-Kruger], Wiesbaden. 211-215.

___ Juha JANHUNEN 2001 *On the paradigms of Uralic comparative studies in Finnisch-Ugrische Forschungen*. 29-41.

___ Santeri JUNTILA, Sampsa HOLOPAINEN, Juho PYSTYNEN 2020 *Digital Etymological Dictionary of the Oldest Vocabulary of Finnish* in *Časopis Instituta za hrvatski jezik i jezikoslovlje*, Vol. 46 No. 2, 2020. 733-747.

___ Petri KALLIO 2001 *Phonetic Uralisms in Indo-European?* in *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations* [ed. Christian Carpelan, Asko Parpola, Petteri Koskikallio]. Mémoires de la Société Finno-Ougrienne 242.

Helsinki. 221-234.

___ Fred KARLSSON 1999 *Finnish: An Essential Grammar* [tr. Andrew Chesterman]. London, New York: Routledge.

___ Radoslav KATIČIĆ 1970 *A Contribution to the General Theory of Comparative Linguistics*. The Hague / Paris: Mouton.

___ Seppo KITTILÄ, Johanna LAAKSO, Jussi YLIKOSKI 2022 *Case* in *The Oxford Guide to the Uralic Languages* [ed. Marianne Bakró-Nagy, Johanna Laakso, Elena Skribnik]. Oxford: Oxford University Press. 879-893.

___ Petri KALLIO 2019 *Daniel Europaeus and Indo-Uralic* in *The Precursors of Proto-Indo-European: The Indo-Anatolian and Indo-Uralic Hypotheses* [edited by Alwin Kloekhorst, Tijmen Pronk]. Leiden, Boston: Brill, Rodopi. 74-87.

___ Aleksei KASSIAN, Mikhail ZHIVLOV, George STAROSTIN 2015 *Proto-Indo-European-Uralic Comparison from the Probabilistic Point of View* in *Journal of Indo-European Studies* 43. 301-347.

___ Simona KLEMENČIČ 2019 *Bojan Čop's Indo-Uralic Hypothesis and Its Plausibility in The Precursors of Proto-Indo-European: The Indo-Anatolian and Indo-Uralic Hypotheses* [edited by Alwin Kloekhorst, Tijmen Pronk]. Leiden, Boston: Brill, Rodopi. 88-101.

___ Alwin KLOEKHORST 2008 *Some Indo-Uralic aspects of Hittite* in *Journal of Indo-European Studies* 36. 88-95.

___ Alwin KLOEKHORST, Tijmen PRONK 2019 *Introduction: Reconstructing Proto-Indo-Anatolian and Proto-Indo-Uralic* in *The Precursors of Proto-Indo-European: The Indo-Anatolian and Indo-Uralic Hypotheses* [edited by Alwin Kloekhorst, Tijmen Pronk]. Leiden, Boston: Brill, Rodopi. 1-14.

___ Alwin KLOEKHORST 2023 *Proto-Indo-Anatolian, the 'Anatolian split' and the 'Anatolian trek': a comparative linguistic perspective* in *The Indo-European Puzzle Revisited* [edited by Kristian Kristiansen, Guus Kroonen, Eske Willerslev]. Cambridge: Cambridge University Press. 42-60.

___ Harold KOCH 1996 *Reconstruction in Morphology* in *The Comparative Method Reviewed: regularity and irregularity in language change* [ed. Mark Durie, Malcom Ross]. New York, Oxford: Oxford University Press. 218-263.

___ Jorma KOIVULEHTO 2001 *The Earliest Contacts Between Indo-European and Uralic speakers in the light of Lexical Loans in Early Contacts Between Uralic and Indo-European: Linguistic and Archeological Considerations*. Helsinki: Mémoires de la société finno-ougrienne. 235-264.

- ___ Tamotsu KOIZUMI 1994 *Uraru-go toogoron*. Tokyo: Daigaku Shorin [with an English summary. 317-57].
- ___ Frederik Kortlandt 1989 *Eight Indo-Uralic verbs?* in *Münchener Studien zur Sprachwissenschaft* 50. 79-85.
- ___ Frederik Kortlandt 2010 *Studies in Germanic, Indo-European and Indo-Uralic*. Amsterdam, New York: Rodopi.
- ___ Frederik Kortlandt [Manuscript] 2020a *The dissolution of the Eurasiatic macrofamily*. Available online at <https://www.Kortlandt.nl/publications/art346e.pdf>.
- ___ Frederik Kortlandt 2020b *The Disintegration of the Indo-European Language Family in Etymologus: Festschrift for Václav Blažek* [ed. Harald Bichlmeier, Ondřej Šefčík, Roman Sukač] Hamburg: Baar. 267-276.
- ___ Frederik Kortlandt 2021 *How I discovered Proto-Indo-European glottalic stops* in *Journal of Indo-European Studies* 49 (2021). 1-12.
- ___ Guus KROONEN 2019 *The Proto-Indo-European mediae, Proto-Uralic Nasals from a Glottalic Perspective* in *The Precursors of Proto-Indo-European: The Indo-Anatolian and Indo-Uralic Hypotheses* [edited by Alwin Kloekhorst, Tijmen Pronk]. Leiden, Boston: Brill, Rodopi. 111-114.
- ___ Leonid KULIKOV 2009 *Evolution of Case Systems* in *The Oxford Handbook of Cases* [ed. Andrej Malchukov, Andrew Spencer]. Oxford: Oxford University Press. 439-457.
- ___ Martin KÜMMEL 2009 *Verwandte des Indogermanischen? Zur Frage des "Eurasiatischen" und anderer Makrofamilien* in *Die Ausbreitung des Indogermanischen, Arbeitstagung der Indogermanischen Gesellschaft*, Würzburg. 1-10.
- ___ Martin Joachim KÜMMEL 2019 *Thoughts about Pre-Indo-European Stop Systems* in *The Precursors of Proto-Indo-European: The Indo-Anatolian and Indo-Uralic Hypotheses* [edited by Alwin Kloekhorst, Tijmen Pronk]. Leiden, Boston: Brill, Rodopi. 115-130.
- ___ Winfred P. LEHMANN 1996 *Theoretical Bases of Indo-European linguistics*. Winfred P. Lehmann. London, New York : Routledge.
- ___ Milan LOPUHÄÄ-ZWAKENBERG *The Anatolian "Eragative"* in *The Precursors of Proto-Indo-European: The Indo-Anatolian and Indo-Uralic Hypotheses* [edited by Alwin Kloekhorst, Tijmen Pronk]. Leiden, Boston: Brill, Rodopi. 131-150.
- ___ Alexander LUBOTSKY 2019 *The Indo-European Suffix *-ens- and Its Indo-Uralic Origin* in *The Precursors of Proto-Indo-European: The Indo-Anatolian and Indo-Uralic Hypotheses* [edited by Alwin Kloekhorst, Tijmen Pronk]. Leiden, Boston: Brill, Rodopi. 151-162.

- ___ Rosemarie LÜHR *Headedness in Indo-Uralic* in *The Precursors of Proto-Indo-European: The Indo-Anatolian and Indo-Uralic Hypotheses* [edited by Alwin Kloekhorst, Tijmen Pronk]. Leiden, Boston: Brill, Rodopi. 163-185.
- ___ Sinfree MAKONI, Alastair PENNYCOOK 2006 *Disinventing and Reconstituting Languages* in *Disinventing and Reconstituting Languages* [ed. Sinfree Makoni, Alastair Pennycook] Clevedon, Buffalo, Toronto: Multilingual Matter Ltd. 1-41.
- ___ Ranko MATASOVIĆ 2000 *Uses and Misuses of Typology in Indo-European Linguistics* in *Sonderdruck aus 125 Jahre Indogermanistik in Graz: Arbeiten aus der Abteilung "Vergleichende Sprachwissenschaft" Graz*. Graz: Leykam. 328-340.
- ___ Ranko MATASOVIĆ 2004 *Gender in Indo-European*. Heidelberg: Universitätsverlag Winter.
- ___ Ranko MATASOVIĆ 2008 *Areal and Typological Affinities of Proto-Indo-European*. University of Macerata, 1-51.
- ___ Ranko MATASOVIĆ 2012 *Areal Typology of Proto-Indo-European: the Case for Caucasian Connections* in *Transactions of the Philological Society* (Volume 110), 1-28.
- ___ Ranko MATASOVIĆ [manuscript] *Clause Alignment in Proto-Indo-European*. 1-34. Manuscript. Available online at <https://mudrac.ffzg.hr/~rmatasov/ClauseAlignmentInPIE.pdf> Accessed on 2023-06-22.
- ___ Alfons NEHRING 1961 *Zur "Realität" des Urindogermanischen* in *Lingua* 10/1961. 357-368.
- ___ Asko PARPOLA 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* in *Археология евразийских степей* N°2, 2022. 258-277.
- ___ Asya PERELTSVAIG 2012 *Languages of the World: An Introduction*. Cambridge University Press.
- ___ Asya PERELTSVAIG, Martin W. LEWIS 2015 *The Indo-European controversy : facts and fallacies in historical linguistics*. Cambridge : Cambridge University Press, 2015.
- ___ Michaël PEYROT 2019a *The deviant typological profile of the Tocharian branch of Indo-European may be due to Uralic substrate influence* in *Indo-European Linguistics* 7(1). 72-121.
- ___ Michaël PEYROT 2019b *Indo-Uralic, Indo-Anatolian and Indo-Tocharian* in *The Precursors of Proto-Indo-European: The Indo-Anatolian and Indo-Uralic Hypotheses* [edited by Alwin Kloekhorst, Tijmen Pronk]. Leiden, Boston: Brill, Rodopi. 186-202.
- ___ Alo RAUN 1988 *Proto-Uralic Comparative Historical Morphosyntax* in [editor Denis

- Sinor] 1988 *The Uralic Languages: Description, History and Foreign Influences* Leiden, New York, København, Köln: Brill. 555-571.
- ___ Károly RÉDEI 1986–1988 *Uralisches Etymologisches Wörterbuch*. Budapest: Akadémiai Kiadó.
- ___ Klaas RUPPEL 2016 *Eräs etymologisen tutkimuksen logiikan pulma in Finnish Conference of Linguistics (25-27 May 2006)*. Oulu: University of Oulu.
- ___ Janne SAARIKIVI 2021 *The divergence of Proto-Uralic and its offspring: A descendant reconstruction in The Oxford Guide to the Uralic Languages* [ed. Marianne Bakró-Nagy, Johanna Laakso, Elena Skribnik]. Oxford: Oxford University Press. 28-58.
- ___ Pekka SAMMALLAHTI 1988 *Historical Phonology of the Uralic Languages* in [editor Denis Sinor] 1988 *The Uralic Languages: Description, History and Foreign Influences*. Leiden, New York, København, Köln: Brill. 478-554.
- ___ Zsolt SIMON 2020 *Urindogermanische Lehnwörter in den uralischen und finno-ugrischen Grundsprachen: eine Fata Morgana in Indogermanische Forschungen, Vol. 125, No. 1*. 239-265.
- ___ Andrew SPENCER 2008 *Does Hungarian Have a Case System?* in *Case and Grammatical Relations: Studies in Honour of Bernard Comrie* [ed. Greville G. Corbett, Michael Noonan]. Amsterdam: John Benjamins Publishing Company. 35-56.
- ___ Edit VÉRTES 1960 *Beiträge zur Frage des finnisch-ugrischen bezeichneten Akkusativobjekts* in *ALH. X*. 181-194.
- ___ Tiit-Rein VIITSO 1996 *On classifying the Finno-Ugric languages* in *C81FU, Volume IV*. 261-266.
- ___ Jussi YLIKOSKI 2011 *A Survey of the Origins of Directional Case Suffixes in European Uralic* in *Case, Animacy and Semantic Roles* [ed. Seppo Kittilä, Katja Västi, Jussi Ylikoski]. Amsterdam: John Benjamins Publishing Company. 235-280.
- ___ Михаил А. ЖИВЛОВ 2017 *Отражение ностратических огубленных гласных в индоевропейском в Памяти В. М. Иллич-Свитыча: материалы круглого стола* [под редакцией Михаил Н. Саенко]. Москва: Институт славяноведения РАН. 94-98
- ___ Mikhail A. ZHIVLOV 2019 *Indo-Uralic and the Origin of the Indo-European Ablaut* in [edited by Alwin Kloekhorst, Tijmen Pronk] *The Precursors of Proto-Indo-European: The Indo-Anatolian and Indo-Uralic Hypotheses*. Leiden, Boston: Brill, Rodopi. 219-235.

Key-words: Proto-Uralic, nominal morphology, cases, linguistic reconstruction, Uralic languages, Proto-Indo-European

Ključne riječi: prauralski, imenska morfologija, padeži, jezična rekonstrukcija, uralski jezici, praindoeuropski

Summary:

The case system of the Uralic protolanguage and its origins are an unresolved problem in historical linguistics. According to Aikio (2022), Proto-Uralic probably had six reconstructable case morphemes: the grammatical cases, the nominative $*-\emptyset$, the accusative $*-m$ and the genitive $*-n$; the adverbial cases, the locative $*-nA$, the separative $*-tA$ and a complex ‘lative paradigm’. This is generally in accordance with the earliest serious attempt at reconstructing the PU case inventory (Collinder 1960), Proto-Uralic had the following *five* case morphemes: nominative $*-\emptyset$, genitive $*-n$, accusative $*-m$, locative $*-na \sim *-nä$ and separative $*-ta \sim *-tä$. Depending on the analysis, one can say that Proto-Uralic has six cases (the three grammatical cases mentioned above, and the three directional cases below), eight (if we add two additional adverbial cases), some indefinite number beyond eight (almost ten lative cases have been offered, or just two, that is to say the only non-nominative grammatical cases. Their relatedness to Proto-Indo-European cases is only possible as regards the grammatical cases, and those are highly problematic for phonetic reasons (for the accusative and the genitive) and for reasons pertaining to ergativity (for the nominative); the adverbial cases show a minimum correspondence. Thus the Indo-Uralic hypothesis rests on factors significantly different from nominal case morphology.

Sažetak:

Padežni sustav uralskog prajezika kao i izvori tih morfema su nerješeni problem u historijskoj lingvistici. Prema Aikio (2022), prauralski jezik je vjerojatno imao šest padežnih nastavaka: tri gramatička padeža, to jest nominativ $*-\emptyset$, akuzativ $*-m$ te genitiv $*-n$; i priložni padeži, to jest lokativ $*-nA$, separativ $*-tA$ te kompleks padeža koji se naziva ‘lativna paradigma’. Ovaj model je uvelike u suglasnosti s najranijim ozbiljnim pokušajem rekonstrukcije prauralskog padežnog inventara (Collinder 1960), u kojem je prauralski imao pet padežnih nastavaka: nominativ $*-\emptyset$, genitiv $*-n$, akuzativ $*-m$, lokativ $*-na \sim *-nä$ te separativ $*-ta \sim *-tä$. Ovisno o analizi, argumentira se u literaturi da je prauralski imao šest padeža (spomenuta tri gramatička padeža, te tri priložna padeža), osam (ako se broje i još dva priložna padeža, kao

što to nudi Aikio), neki nedifinirani broj više od osam (gotovo deset mogućih lativnih padeža se mogu rekonstruirati), ili samo dva, to jest jedina dva nenominativna gramatička padeža. Genetički odnos s praindoeuropskim padežima je jedino moguć za gramatičke padeže, a i oni predstavljaju probleme u rekonstrukciji, što zbog fonetskih razloga (u slučaju akuzativa i genitiva), što zbog rekonstrukcije ergativnosti (u slučaju nominativa); priložni padeži predstavljaju znatno manje podudarnosti. Zaključuje se da je indouralska hipoteza osnovana više na drugim faktorima, nego li na imensku morfologiju padeža.

Kratak životopis:

Ivan Silobrčić, rođen u Trstu, Talijanska Republika 17. srpnja 1998. Pohađao osnovnu školu na engleskom jeziku u Oslu, Kraljevina Norveška te srednju školu također na engleskom u Zagrebu, Republika Hrvatska. Upisao studij ruskog jezika i književnosti te francuskog jezika i književnosti na Filozofskom fakultetu Sveučilišta u Zagrebu 2016. godine. U srpnju 2019. boravio tjedan dana radi usavršavanja francuskog jezika kao stipendist na Festival des étudiants francophones d'Europe centrale et orientale (Agence universitaire de France) u Plovdivu, Republika Bugarska. U ljetnom semestru 2020./2021. akademske godine boravio radi usavršavanja ruskog jezika kao stipendist Erasmus+ (studentski boravak) programa Europske unije u Tallinnu, Republika Estonija. 2021./2022. akademske godine upisao diplomski studij francuskog jezika i književnosti (smjer: znanstveni) te diplomski studij lingvistike (smjer: poredbeni). 2022./2023. akademske godine stipendist Sveučilišta u Zagrebu u kategoriji stipendija za izvrsnost.

Appendix: Handout for the thesis defence

Reconstructing the Proto-Uralic Case System With Regard to Proto-Indo-European,

Ivan Silobrić – handout (obrana diplomskog rada, 28. VI. 2023.)

gramatički padeži (po Collinder 1960, Raun 1988, Abondolo 2006a, Aikio 2022)

nominativ *absolutus* *-∅ u svim uralskim granama (njih devetero, po Aikio)

akuzativ *-m u finskoj grani, mordvinskoj grani, permskoj grani, marijskoj grani, mansijskoj grani, samojedskoj grani

genitiv *-n u finskoj grani, samskoj grani, mordvinskoj grani, marijskoj grani, samojedskoj grani

→ Problem relativno malog broja refleksa većine padežnih morfema, te fonemski ultrastabilnih prauralskih nastavaka akuzativa i genitiva.

→ Nepoznat broj prauralskih padeža. Gotovo nijedan od rekonstruiranih prauralskih padežnih morfema koji se da rekonstruirati nije van sumnje (najmanje sumnje predstavlja akuzativ). Njihov odnos s praindoeuropskim je dakle spekulativan.

priložni padeži (po Collinder 1960, Raun 1988, Abondolo 2006a, Aikio 2022)

lokativ *-nA u svim uralskim granama osim u mordvinskoj i samojedskoj

separativ^{ablativ} *-tA u svim uralskim granama osim u permskoj i samojedskoj

lativi *-k, *-n, *-ń, *-ŋ, *-s, *-j... po svim uralskim granama

→ *Überkasusursprung*: priložni padeži se mogu formalno rekonstruirati *ad nauseam* u prauralskom, te je zapravo nemoguće odrediti što je prauralskog porijekla.

→ Etimologije prauralskih jednosložnih morfema su nepoznate.

neki od drugih mogućih prauralskih padeža (po Aikio 2022)

translativ *-ksi u finskoj grani, mordvinskoj grani i marijskoj grani

privativ *-ktAk u samskoj grani, u finskoj grani, marijskom, permskoj grani, hantijskoj grani i mansijskoj grani

→ Gramatikalizacija postpozicija, priloga i punoznačnih imenica (aglutinacija/separativnost) je vrlo produktivna kroz uralske jezike, dakle nejasno stanje u prauralskom.

→ Od slabe produktivnosti aglutinacije (sjevernohantijski: tri padeža) do jake produktivnosti aglutinacije (mađarski: maksimum 28 padeža).

→ Primjer jezičnog kontakta kasnog prauralskog s toharskim B, prafinskog sa starolitavskim.