Peter S. PIISPANEN (Stockholm)

The Uralic-Yukaghiric connection revisited: Sound Correspondences of Geminate Clusters

This paper presents and discusses regular correspondences between Uralic geminate items and Yukaghiric with proposed sound change laws and new and some modified older cognate suggestions (twenty-four nouns and eight verbs). Geminate items were found to contain surprisingly stable, relatively unchanging vowels in Yukaghiric in regard to the Proto-Uralic form. The results suggest that degemination – taking place in all cases except in a few forms that can otherwise be explained – was an early process in Yukaghiric and occurred after or while many vowel changes had already taken place in the Yukaghiric vocabulary. The data shows that the relationship between Uralic and Yukaghiric is more extensive than previously believed. Some very early possible sound changes are discussed. Furthermore, a correspondence to Proto-Uralic *-ü- has been found in Late Proto-Yukaghiric *-ö-. Also, it is shown that the early suffixation in Yukaghir to Uralic-like stems has produced several modern words through grammaticalization.

1. Introduction

1.1. Yukaghir languages

The Yukaghir languages are spoken in the extreme north-east of Siberia in Russia. While the Yukaghir languages were earlier spoken in a very large area (Collinder 1965; see the striped area of Figure 1[1]), there currently – since the 19th century – remain only two spoken forms; these are Kolyma Yukaghir (KY), found around the Kolyma River in the south (southern solid area) and Tundra Yukaghir (TY), found in the far north in the coastal area (northern solid area).

Figure 1. Current and historical geographic areas of the Yukaghir languages.

The two languages, which have traditionally been seen as representatives of two dialects at the extreme ends of a dialect continuum (Nikolaeva 2008), are lexically very different and mutually unintelligible, which is why it may be more proper to regard these as separate languages, the common denotation of Yukaghir notwithstanding.[2]

---

1. This historical area constituted most of the modern Chukotka Autonomous Okrug, the northeastern parts of the Sakha Republic and the northern Magadan Oblast. This puts the area approximately between mainly the river Lena in the west and the river Anadyr in the east (Dolgikh 1960).
2. For example, Kolyma Yukaghir and Tundra Yukaghir had a cognancy rate of only 44% on the Swadesh 100 word list in one study (Nikolaeva 2008: 327).

1.2. The Historical Yukaghir

Historical records reveal that earlier there were 13 shamanistic tribes of nomadic hunters and gatherers in this geographic area within clan systems which constituted the ethnic Yukaghir group (Dolgikh 1960: 379–442).[3] The Omok and Chuvantsy varieties, for example, survived up to perhaps the 18th century. In 2002 the total number of remaining Yukaghirs was only 1,509 people (Russian 2002 Census), of which only 50 could speak Kolyma Yukaghir and 150 Tundra Yukaghir (Nikolaeva 2006: vii). As of the late 1980s, the majority of ethnic Yukaghirs have been using Russian or Yakut, a Turkic language, as well as other indigenous languages of the areas: Chukchi, Evenki and Even (Odé 2009). According to later Russian observations in 2009, there were only five to ten Kolyma Yukaghir and 60 to 70 Tundra Yukaghir speakers left. As such, the two remaining Yukaghir languages are moribund (Wurm 2001: 27, 51).[4]

2.   Yukaghir etymology

2.1. Data collection

The Yukaghir lexicon was first written down in the late 17th century by military personnel in the area, and then these writings were handed over to Russian scholars.[5] Over the years several other collections were made with each source describing only one local idiolect. Thus, several smaller dictionaries of the Yukaghir languages have been presented, as well as two larger ones, all used as sources in Nikolaeva's recent dictionary. The grammars of Tundra Yukaghir and Kolyma Yukaghir were presented in 2003 (Maslova 2003a; 2003b, respectively). While important and voluminous lexicographic work on Tundra Yukaghir was published in Russian in 1990 and 2001 (Kurilov 1990; 2001), the most recent and possibly authoritative lexicographic work on Yukaghir consists of Nikolaeva's *A Historical Dictionary of Yukaghir* (Nikolaeva 2006). It contains 2,623 entries (totaling of about six or seven thousand words), and summarizes data from all of the previously reported idiolects into a comparative format. Additionally, and equally importantly, reconstructed (later) proto-Yukaghir (PY) root stem forms are presented therein for the first time. While etymological notes for some loanwords from either Russian or from (and to) some other local language are given, standard etymologies for the Yukaghir vocabulary is not usually given because the connections of Yukaghir and its earlier history of language contact remains unknown.

---

3.   These were: Vadul-Anais, Odul, Chuvan, Anaoul, Lavren, Olyuben, Omok, Penjin, Khodynt, Khoromoy, Shoromboy, Yandin and Yandyr.
4.   Merlijn De Smit, Jarmo Lainio and Jenny Larsson are gratefully acknowledged for their valuable input on the manuscript during preparation. I also give thanks to the editor and anonymous reviewers for their valuable input for improvisations.
5.   Some of these manuscripts described now extinct Yukaghir languages.

2.2. The *Uralic-Yukaghiric* genetic relationship hypothesis

Yukaghir was originally thought to be a language isolate. In 1907, Paasonen (1907) noted some similarities between Yukaghir and the Uralic languages, as did Pedersen (1931). Contributions by many researchers followed: Winkler (1912: 115–127) argued for a genetic relationship based on similarities in personal pronouns. Lewy (1928: 274–289) found 27 lexical correspondences between the language groups. Bouda (1940: 71–93) presented 100 Yukaghir to Uralic/Finno-Ugric correspondences, while in 1940, Collinder noted several parallels between the two groups and argued for a distant genetic relationship, summarizing a total of 90 possible cognates in the Yukaghir languages in his later dictionary (Collinder 1955). Some of these were derived from the same root morpheme. Up to perhaps five hundred rather poorly founded correspondences were reported by Sauvegeot (1969: 354). Wickman (1969) argued for a genetic relationship between Indo-European, Ural-Altaic and Uralo-Yukaghiric. The possible relationship between the two, as well as the nature of such a relationship, has been extensively discussed, for example, in Collinder (1940; 1957; 1965a; 1965b), Bouda (1940), Angere (1956), Tailleur (1959), Nikolaeva (1988a) and Fortescue (1998). In summary, the main counter-arguments against a genetic relationship are the relatively low number of proposed cognates, the lack of many solid sound laws, different morphological forms between Yukaghir and most Uralic languages and the remote geographical distances between the two language groups of Uralic (Proto-Uralic (PU) is usually considered to have originated somewhere quite near the Ural Mountains at any one of six different locations; Campbell 2004: 405–408) and Yukaghir (presently in North-east Siberia). However, regarding the number of proposed cognates, it must be pointed out that the situation is quite similar even within the Uralic language group with Finno-Ugric and Samoyed sharing only a bit more than 200 cognates (Janhunen 2009: 58); due to the higher time depth far fewer cognates could be expected for genetically related Uralic-Yukaghiric than between the individual Uralic languages. As to the lack of solid sound laws, this is a difficult, although not impossible task to formulate (i.e. sound laws) given the low number of items on the incompletely known list of cognates. Some sound correspondences have quite convincingly been presented in detail even if the nature of the relationship between Uralic and Yukaghiric is unclear. More phonological work is reportedly being carried out at the moment (Häkkinen 2012: 91), and will likely result in very interesting reading in the coming years. Then again, regarding the remote geographic distances between the languages, even the Uralic languages are spread out over very large geographical distances, where related, intervening earlier languages have died out, which is why this argument may hold little merit.

2.3. The *Uralic-Yukaghiric* borrowing hypothesis

In contrast, by another slightly more recent school of thought (see Krejnovič 1958, Rédei 1999 – also includes an overview of Uralic-Yukaghir research – and, recently Häkkinen 2012), the lexical correspondences between various branches of Uralic and Yukaghir are derived from the phenomena of contact and through loanwords only, apparently assuming no genetic relationship between the two groups. Rédei's (1999: 10) conclusions – following the stated edict that a genetic relationship must be supported by sound lexical, phonological and morphological correspondences – were that only typological correspondences can be demonstrated between Uralic and Yukaghir, but not decisive grammatical/morphological correspondences, which is why lexical correspondences must entirely be due to borrowing from four different Uralic sources to Yukaghir. While being a sound principle it does require vigorous research before a final verdict can be made. Presented (in Häkkinen, building on Rédei) were suggested loanwords from (Pre-)Proto-Uralic into Early Proto-Yukaghir (32 words) and from East-Uralic into Middle Proto-Yukaghir (24 words), as well as a list of previously suggested cognates (30 words), which were rejected by distribution or semantic or phonologic considerations. The analysis of the Samoyed cognates (38 words) was postponed for future study.

2.4. A matter of controversy

Today, there is no universal consensus on the matter, possibly because so little of the history of Yukaghir is known. Irina Nikolaeva's (2006) dictionary notes possible Uralic correspondences and gives appropriate suggestions for sound correspondences for several Yukaghir items (see also Nikolaeva 1988b), basing such arguments on her own analysis as well as on those previously presented in scientific literature. Given this state of affairs, some researchers now do accept that the two groups are indeed genetically related (Uralo-Siberian is listed as a language family, for example, in Fortescue 1998). However, this school of thought does not believe Yukaghir to be Uralic, but rather a para-Uralic language group. As such, the Yukaghir languages and the Uralic languages are both believed to originate from an even earlier proto-language, forming a larger Uralo-Yukaghir language family.

    As genetically related languages, the original language has been called *Proto-Uralo-Yukaghiric*, which is a geographically neutral term, or *Proto-Sibero-Uralic*, which is the geographic term. *Sibero-Uralic* was proposed, for example, by Fortescue (1998), where he also included Chukotko-Kamchatkan and Eskimo-Aleut. In the case of a common origin, the Uralic and Yukaghir languages can be traced directly back to their common proto-language.

    As to borrowed lexicon, recently, Häkkinen (2012) spoke of Pre-Proto-Uralic as the donor language to the Yukaghir lexicon. Given that (Late) Proto-Yukaghir is reconstructed from a relatively small number of quite recently attested languages,

Proto-Uralic is, chronologically speaking, a far older language, which is why many of the suggested Yukaghir cognates can be traced directly back to a Uralic form. This would be the case both with the borrowed lexicon and with genetically related material.

3.   Possible problems with the lexical borrowing scenario

3.1.   Regarding direct borrowing problems

As noted above, there are two opposing schools of thought regarding the Uralic–Yukaghir relationship. In the first school, the two language groups are believed to be genetically related from an even earlier common language (referred to as *Pre-Proto-Uralic* or *Proto-Uralo-Siberian*) (for example Collinder 1940; 1955; 1957; 1965a; 1965b; Fortescue 1998). In the second school, all cognates of Uralic vocabulary in Yukaghir, as well as morphological features, are believed to originate from extensive language contact situations (for example Rédei 1999 and Häkkinen 2012).

It is certainly true that "one of the most important confounding factors for any type of approach to genealogical classification is lexical borrowing." (Haspelmath 2008), and "unfortunately, the further back in time, the more difficult it becomes to distinguish inheritance from transfer." (Nelson-Sathi et al., 2010). The situation is no doubt difficult to analyze, and more research on the subject is required. While I can personally see good arguments and merits for both schools, I will, however – if I am allowed to play the devil's advocate for a moment – point out certain problems below with the "borrowing only" theory that need to be taken into future account.

While a suggested borrowed lexicon in Yukaghir has been divided into cultural spheres (Rédei 1999: 30–31), the lexicon from three of the claimed sources ((earlier) North Samoyed, Nenets and Selkup) – supposedly constituting different chronological waves of language contacts between Yukaghirs and speakers of different Samoyedic languages – were all grouped into one table with many semantic categories. The categories of proposed borrowings from different cultural spheres of the fourth source, Ugric, in a second table, are the same as in the previous table (all tables are found in Rédei 1999). As such, the lexicon in Yukaghir with Uralic correspondences does not appear to clearly constitute a particular cultural subgroup of borrowed vocabulary of any given chronological period or culture.[6] Rather, Uralic correspondences are found quite extensively in function words and the core vocabulary – which, it is well-known, is more resistant towards borrowing – such as, tentatively, in personal pronouns, numbers, demonstrative pronouns and terms of kinship (many of which

---

6.   In stark contrast to what has been seen in cultural lexicon borrowing with systematic sound correspondences, for example, in some Indo-European studies, see for example Suhonen (1988) and Larsson (2001). In Finnish (Fin.), for example, there are clear borrowing strata from the Germanic and Baltic languages in cultural areas such as ships and seafaring).

bear unclassified similarities), as well as natural phenomena and dozens of common verbs. And yet, it is known, that "everyday low-prestige words, function words and grammatical formatives are usually not borrowed" (Harbert 2007: 23).[7] To expand a bit on this: For example, the Yukaghir 1st pers. sing. and 2nd pers. sing. *mət* 'I' and *tət* 'you, sing.', respectively, certainly appear similar to the corresponding PU **mE/*mon* 'I' and **tE/*ton* 'you, sing.' (the literature presents two main proto-forms; UEW 294–295, 539–540, respectively, and Janhunen 1982: 273–274. More on the front vowels of pronoun stems in Honti 1995). Likewise, the Uralic -*i*- infix, signifying plurality, may be found in the Yukaghir 1st pers. pl. *mit* 'we' and 2nd pers. pl. *tit* 'you, pl.'. While third person pronouns are occasionally borrowed (for example, English *they* from Old Norse *þeir* 'they'), it is believed to be much rarer with first and second person pronouns. In general terms, there appears to be a difference between closed and paradigmatically tightly linked pronoun systems (such as in the Uralic languages) where borrowing is unlikely, and more open systems (various honorifics, etc.) where borrowing may occur much more easily. While plural pronouns were not marked in the Proto-Uralic form, there are quite similar plural markings to the Yukaghir forms through secondary development in Saamic (N. Saami *mii* 'we' and *dii* 'you, pl.') and North Finnish (*met* 'we' and *tet* 'you, pl.').

3.2. Regarding numerals and pronouns

In regard to numerals, for example, Yukaghir *irke* 'one' bears a structure showing a possible similar development from PU **ükte/*ikte* 'one' (UEW 81) as what is seen in Moksha Mordvin (MM) *ifkä* 'one' (suggestion: **ükte > *ütke > *itke > irke*; perhaps similar to the mutation **t > r/_V* displayed in, for example, Nivkh (Gruzdeva 1988), as well as similar word-initial transformations seen in Celtic), while TY *kiji* 'two' resembles dialectal Mansi *kitiɣ* 'two' and PS **kite* 'two' from PU **käktä* 'two' (Janhunen 1981: 272; UEW 118). Likewise, PFU **kolme* 'three' (UEW 174), PFP **kolmanti* 'third' (Sammallahti 1988: 552) and KY *jalme* 'three' and *jalməštə(gi)* 'third',[8] as well as PFU **ńeljä* 'four' (UEW 315–316) and KY *jelek* 'four' (this pair also mentioned in Nikolaeva 2006: 188), bear noteworthy similarities. As to groups of demonstrative pronouns, the PU roots **ta-* (ex: dialectal Finnish *taa* 'this one'; UEW 505) and **to-* (ex: Finnish *tuolla* 'there'; UEW 526–528) bear a striking similarity to the PY demonstrative pronoun root **ta-* (ex: Yukaghir *tada:* 'there'). Furthermore, interrogative pronouns bear similarities: PU **ke-/*ki-* 'who' (UEW 140–141), Fin. *ken* 'who', Yukaghir *kin* 'who' as well as PU **ku-/*ko-* 'which, what' (UEW 191), Fin. *kun* 'when', *kuka* 'who', *koska* 'when', KY *qadi* 'which', *qajn* 'when' etc. Some of these also bear similarities to the forms of some Altaic languages as well as Indo-European.

---

7. Although the borrowing of pronouns does occur at times. See Lainio (1995) for a few examples in Swedish.
8. Also cf. Finnish *kolmesti* 'three times'.

3.3. Regarding geography

The Yukaghir vocabulary with Uralic correspondences consists of Paleolithic vocabulary not bound to any specific cultural or semantic field. For the most part such found correspondences are phonologically systematically regular. Tentative exceptions to regularity (personal observations) could be explained as further consecutive changes, changing the vowel qualities among other things, which are explainable by common phonological processes, such as labialization, palatalization etc., found throughout the Uralic languages, for example in going from Proto-Uralic to Komi-Zyrian (Rédei 1988). Could such extensive borrowing have taken place over such long (at least 2,000 km) geographic distances through several chronological waves from several geographically separate, but quite closely related languages, into approximately the same area of speakers in Northeast Siberia where Yukaghir was spoken, and thus infuse local Yukaghir forms with borrowed vocabulary which then spread out to many of the other Yukaghir dialects? Or instead, did a few larger migrating groups of speakers of a genetically related language migrate westwards to produce all of the Uralic languages? The distance does not lend to the rejection either of these possibilities.

3.4. Regarding morphology

Regarding the morphological similarities of Yukaghir and Samoyed (Rédei 1999: 12–13), which are quite confounding, these could result from either extensive language contact situations or a common linguistic genetic relationship. Most noteworthy are the following case-marking similarities (Mikola 1988: 236–237) nom. -Ø (PS & KY), gen. *-n (PS & PU) & -n/-d (KY), loc. *-kana (PS) & -gə (KY) and abl. *kata (PS) & -gət (KY). Further notes of interest on the subject of possessive markers have been published (Seefloth 2000). Why do these two language groups have morphological similarities if it is a question of borrowed vocabulary? In fact, one could argue for, based on the morphological structures alone, for example, the following possible chain of events:

> Pre-Proto-Uralic → Early East (Pre-)Proto-Uralic (+ West (Pre-)Proto-Uralic → Proto-Finno-Permic → Proto-Finno-Volgaic[9] → Proto-Finnic) → Late East Proto-Uralic (+ Early Proto-Yukaghir → Middle Proto-Yukaghir → Late Proto-Yukaghir) → Proto-Samoyed + Proto-(Finno)-Ugric.

Such a tentative model shows that it is fairly simple to construct a reasonable genetic tree, which explains quite much, while also offering some other advantages. As such,

---

9. Proto-Finno-Volgaic is included here as a purely lexical layer. I believe that this layer of vocabulary may consist of quite a great deal of borrowings and innovations resulting from extensive language contact situations.

Yukaghir and Samoyed could contain morphological and phonological features, as well as some lexical features, from the oldest, common (North-)East Uralic forms. This way there would also be some lexical traces left identifiable as West Uralic, while most cognates would be found with East Uralic.

It has been suggested, in defense of a borrowing scenario, that languages do not develop from a proto-language into very different morphological directions while retaining most inherited words unchanged (Häkkinen 2012: 94). However, innovation in one linguistic sub-system may co-occur with archaism in another: Saamic lacks the outer local cases and thus has a more "archaic" case system but has undergone very significant changes in its vowel system, which Finnic has not. Thus, in contrast, Finnish has developed several new cases from the far fewer original Uralic ones, while retaining an archaic vocabulary (for example briefly in J. Häkkinen 2011).[10] Likewise, the morphology of the similar negative verbs of Uralic and Yukaghir need to be addressed. For these reasons it would not be too odd if a genetically related Yukaghir language had kept morphology in common only with its possibly closest branch (also in geographic terms), Samoyedic, while retaining quite a few items of the unchanged lexicon.

3.5. Regarding typology

According to Nichols: "Features having fair or better grammatical stability, moderate or better genetic stability, fair to moderate areal consistency, and a scale of patterning continental or larger in size can give a good picture of the long-standing affinities and disparities among large areas." Structural markers of particular interest are: "head/dependent marking, alignment, inclusive/exclusive pronouns, plurality neutralization, and prepositional phrases". (Nichols 1992: 185.) Perhaps not surprisingly, a comparison between some Uralic languages (Komi, Hungarian and Finnish) to Yukaghir reveals a rather strong, although not complete, correlation between such morphological properties, where Finnish is the outlier. It must be pointed out that some of these features seem quite common to the Altaic languages in general. According to Nichols's typological approach this may be taken as a hint of a valid genetic relationship between Uralic and Yukaghir.

---

10.   Here it may be useful to also compare the number and type of cases of Yukaghir and Proto-Uralic. Yukaghir has nominative, rhemative (also called the predicative case), accusative, genitive, dative, instrumental, comitative, locative, ablative and prolative (Nikolaeva 1988a: 140), with some having differences between determinative and indeterminative use, while Proto-Uralic had at least six known cases: nominative, accusative, genitive, locative, ablative and lative (Itkonen 1966: 69). What is noteworthy is that practically all the Proto-Uralic cases can be found in Yukaghir which also has a few more cases.

3.6. Regarding phonology

Furthermore, as regards to phonology, it has been suggested that borrowing may be hampered by the structural incompatibility between the loaning and borrowing languages. For example, French regularly does not borrow verbs due to its own peculiar verbal morphology which is incompatible with that of many other languages (Thomason & Kaufman 1988: 348). For Yukaghir to so extensively borrow lexical items of all word classes from (Pre-)Proto-Uralic, and for such words to stay relatively unchanged into modern times, the two languages must have indeed born many similarities. If they are structurally so similar in phonology and morphology, as well as share a common lexicon, it may instead be suggestive of a valid genetic relationship between the two.

3.7. Regarding borrowing facts

Additionally, scholars usually agree that nouns are more easily borrowed than verbs (Hock and Joseph 2009). For example, a study of a Quechua corpus revealed 42–49% of nouns vs. only 11–27% of verbs loaned (van Hout & Muysken 1994). If all Yukaghiric cognates were borrowed, with (Pre-)Proto-Uralic as the loaning language, why are there so many verbs compared to nouns (37.5% verbs and 62.5% nouns in Häkkinen 2012) among the borrowed corpora? It has indeed been pointed out that the cognates found are amazingly similar. However, naturally, the most similar cognates are also the ones easiest to detect why the cognates known may only be part of the full cognate set. Quite possibly, then, there are more cognates that are less similar to be found, and then in particular within the noun class. This would actually be true for both the borrowing theory, if types of common borrowed loanword strata in other languages bear any merit, and if a genetic relationship is valid, since nouns are a much larger word class than verbs. It should also be pointed out that the Quechua study further concludes that frequently used vocabulary, i.e. a core vocabulary, is more resistant than general lexicon towards being exchanged. Why then would the Yukaghir core vocabulary be so extensively exchanged with loanwords in all word classes and types?

3.8. Summary of problems with the borrowing hypothesis

All of these factors (the lexicon in question, phonology, morphology and structural markers), then, could suggest, instead of a mere lexical borrowing situation, either a direct genetic relationship with Uralic or these factors are the remnants of a language shift situation.[11] However, in a language shift scenario the problem is that typically

11.   In such a case, given the prevalence of a basic corresponding vocabulary, the original language would have been (Pre-)Proto-Uralic, which switched over to Yukaghiric.

very little is left of the original vocabulary – while phonologic and syntactic traces can more commonly be found – and when there is vocabulary remaining it often relates to specialized semantic fields (e.g. words for flora and fauna, which are not found in the language shifted to); as mentioned, the Uralic-Yukaghiric cognates are not limited to any specific semantic field, and Yukaghir does not clearly show the aforementioned traces.

In this article I intend to present new cognates (borrowed or inherited) between the language groups, which show apparent Yukaghir cognancy with geminate Uralic lexicon.

4.   Lexical comparisons and suggested cognates

Given that the genetic relationship between Uralic and Yukaghir is possibly real, it should be possible to find systematic sound correspondences for most phonemes between the languages and to adequately explain the exceptions. Suggestions regarding some consonants have been presented (Nikolaeva 2006 and others), while the vowels have only partially been researched (Häkkinen 2012). Such general correspondences, of course, need to be separate from mere happenstances of similarity between the two language groups, and the nature and synchronic nature of *wander words* or other borrowed items need to be specified. No doubt, both in finding new cognates and in describing the sound correspondences between two tentatively genetically related languages, a very rigorous and systematic approach is necessary, and permissive semantic and phonological correspondences are to be avoided as far as possible.[12]

In the extended literature, approximately 80 to 90 reasonable cognate suggestions between Uralic and Yukaghir can be found, as well as a bit over 60 suggestive correspondences between later Uralic languages and Yukaghir.[13] While for example Collinder and Nikolaeva tentatively seem to assume, although it is rarely directly stated, that these are cognates, all of these were considered either loanwords or unrelated items in Häkkinen's work. Häkkinen does indeed justly disqualify or cast doubt on several earlier cognate suggestions on phonological or semantic grounds. The division of lexicon into Early (EY) or Middle Proto-Yukaghir (MY) is ambitious, reasonable and an excellent start, although it may suffer from the fact that the phonological form of several of the vowels can be explained from common phonological effects, such as the palatalization and labialization effects found in several Uralic and

---

12.   And arguably, proposed cognates with permissive semantic shifts do not really help prove the genetic relationship. This is, for example, a problem with many of the proposed cognates, one of the arguments of Vajda's interesting Yenisey–Dene connection hypothesis, throughout numerous publications, as has been pointed out by Lyle Campbell in his criticism (Campbell 2011). The Yenisey–Dene connection is a proposed genetic relationship of much higher time depth than the Uralic–Yukaghiric case.
13.   This indicates that such correspondences are not to be found in the Samoyed branch. At most 60 such suggestions are for Proto-Finno-Ugric items, as well as one Proto-Finno-Permic and two Proto-Finno-Volgaic items. The suggested correspondences to Samoyed vocabulary (38 words) will likely be treated in the future.

the Yukaghiric languages, as well as nasalization and uvularization effects found in other languages, which is why some of the presented sound correspondences seem less certain at least to this author. Given the small number of items in every vowel correspondence group presented this unfortunately makes the possible margin of error even larger.

For example PU *wanča* 'root' (UEW 548) is given as a loanword into Early Proto-Yukaghir since it is found in Late Proto-Yukaghir as *wonč-* 'root'. However, *(k)wa- > *(k)wo- > *(k)wu-* is a common labialization effect found in, for example, Komi-Zyrian (Rédei 1988), meaning that the regular correspondence in Yukaghir could well be *wanč-* (which according to Häkkinen's scheme would instead make it Middle Proto-Yukaghir) and then later be labialized into PY *wonč-*. Consequently, the developmental stages seem uncertain. It seems more cognates need to be found to safely assess the phonological development. It is perhaps noteworthy that all the corresponding lexicons involve Paleolithic vocabulary. Relatively few detailed sound correspondence or sound change laws have yet been presented. Most proposed cognates are either semantically identical, synonymous, or very closely related in meaning, while others have seemingly permissive semantic shifts, which is why the research presented here was done.

5.   Research methodology

In this study, focus has been – while making no assumptions about either borrowing or a genetic relationship – on establishing tentative sound correspondences between previously proposed cognates, to say something about the chronology of such sound changes, and to use these to more firmly establish both the existence and possibly, in the best case scenario, details regarding the nature of the relationship of the languages. A sub-goal has been comparing the Kolyma and Tundra Yukaghir lexicon to modern Finnish – a Uralic language – because the Finnish lexicon is well studied among scholars of the Uralic languages and constitutes a good etymological basis of comparison. After methodologically summarizing all of the apparent sound correspondences for the suggested, related items in tabular form, both vowel and consonant correspondences could be found. A critical examination has been necessary during every step of the analysis. It seems to this author that both vowels and consonants of such apparent sound correspondences have changed according to common phonological principles, including those that have been conditioned by their immediate phonological environment. As such, the changes are explainable by common phonological principles found in various languages around the world, but most specifically within the Uralic languages. Finally, a few earlier cognate propositions could be rejected,[14]

---

14.   Such as PU *lēme*, cf. Fin. *liemi* 'broth' being a cognate of Kolyma-Yukaghir *leppul* 'blood'. Instead, the PU *leppä* 'alder' (UEW 689), which is a red tree, seems a more likely and phonologically acceptable cognate, as was also suggested by Nikolaeva. Such a semantic development appears similar to the development of Fin. *punainen* 'red' from PU *puna* 'hair' (UEW 402).

while, interestingly, quite a few new ones, including cognate terms of kinship and other nouns, could be found by back-tracking the lexicon from the sound change correspondences found using dictionaries.[15]

Two examples of new cognate suggestions are: 1. PFP *äje* 'father, big' (UEW 609), cf. Fin. *äijä* 'guy, old man', cf. Est. *äi* 'grandfather', cf. Udm. *aj* 'father, grandfather', KZ *aj* 'father, grandfather', N. Saami *ag'gja* 'grandfather, old man, fellow' and cf. Kar. *äijä* 'many' has a likely correspondence in PY *o:jə*, cf. KY *oye* 'father'. Another likely correspondence is between 2. PFP *jure* 'root (end)' (UEW 639), cf. Fin. *juuri* 'root, foot, origin', cf. Est. *juur* 'root', Erzya Mordvin (EM) & MM *jur* 'end of stem', Udmurt (Udm.) *jir* 'head', KZ *jur* 'head', to PY *jo:-* 'head' and KY *jo:* 'head'; this has undergone the exact same semantic shift as in Komi-Zyrian and Udmurt and was discussed in UEW 639. While there are more items, due to the large resulting dataset and the space limit of this paper, the Yukaghir vocabulary corresponding to the Uralic vocabulary limited to consonant geminates is only presented and discussed here. Up to sixteen new nouns (including the two above) and seven new verbs as cognates between Uralic and Yukaghir were uncovered and are presented below.

It will be shown that a few modern Yukaghir items have arisen through the grammaticalization of suffixes attached to earlier proto-forms apparently similar to Proto-Uralic word stems. It has been shown that Late Proto-Yukaghir had a very rich system of up to 183 inflectional and derivational affixes (Nikoleava 2006: 79–83), differing in productivity, and traceable in the modern KY and TY lexicon. As a comparison, there are 139 reconstructable Indo-European and/or Proto-Germanic noun- and adjective-forming suffixes (attached to root stems). However, only 42 reasonably productive forms were left, for example, in Old English (Lass 1994: 199), the rest having disappeared, been exchanged by a far fewer number of newer suffixes, or simply been grammaticalized into the lexicon into productively unrecognizable forms. In Yukaghir, as is not uncommon for older languages, many of the older suffixes also seem to have become unproductive and can only be traced in the phonological forms of the lexicon. This trend, the loss of suffixes through lexical grammaticalization, may be representative in the development into a newer language.

All suggested cognates naturally must be adequately described in lexical, phonological and semantic terms. In certain cases, throughout the analyses, there are etymological or phonological uncertainties and difficulties that are discussed further. New cognate suggestions, to the best of my knowledge, are marked as such. All cognate suggestions are numbered and underlined for easy reference. Some older suggestions are expanded upon and given references. A number given after a suggested correspondence refers to the entry number in Nikolaeva's (2006) *A historical dictionary of Yukaghir* (format: noted in entry number, example: "noted in 2150" means that the correspondence was noted in Nikolaeva's Dictionary entry number 2150), where this correspondence was also discussed, while other references, such as the monumental

---

15.   Also, see the section on correspondences to PU *ü* for more examples.

Uralische Etymologische Wörterbuch (UEW 1988–1991) and, on rare occasions, the *Uralic Etymological Database* (UED online) are normally given together with the page(s) (format: UEW page number(s)) and entry number (format: UED entry number), respectively, for the proto-items. The following basic etymological dictionaries have also been consulted for a further tracing of Finnish and Yukaghir: Collinder (1955), K. Häkkinen (2011), SSA 1992–2000 and SKES 1955–1981. Many other earlier dictionaries and etymological works have been consulted, to produce the above work, some of which are found in the Etymological References section of the bibliography.

6.   Correspondences of geminate clusters

6.1. Basic correspondences

Proto-Uralic geminate items have a correspondence in degeminated or rhotacized forms in Yukaghir. All the other sound changes involved are non-controversial and can be found in literature dealing with Yukaghir phonology (for example Rédei 1999: 25–28; Nikolaeva 2006: 29–78). There are factors explaining the sometimes varied correspondences. Curiously, it was found that the vowels of geminate Proto-Uralic or Proto-Finno-Ugric (PFU) items are often unchanged in the assumed Late Proto-Yukaghiric correspondence or they have been subject only to minor, chronologically late, and phonologically quite expected vowel changes. Thus, as a general matter, the vowels of Proto-Uralic items containing the geminate consonant clusters *-pp-*, *-kk-* and *-tt-*[16] were found to be relatively stable in their Yukaghir correspondence, although the analysis is complicated and made more difficult by the presence of incompletely understood proto-items. The end result was always a degeminated item in Yukaghir. Some correspondences in the modern languages of Finnish, Moksha Mordvin, Udmurt, Kolyma Yukaghir, Tundra Yukaghir and others, as given in the abbreviations, are also presented. In some cases, all the vowels of the Proto-Uralic item are not known, but given that Proto-Uralic does show signs of vowel harmony, and that the forms in the daughter languages are known, educated assumptions on what such vowels can or must have been in Proto-Uralic can be made.

6.2. Non-trivial correspondences

As it has been stated, cases of regular correspondences between geminate Uralic items and degeminated Yukaghiric items can be found. However, there is a phonological, non-trivial curiosity in the dataset where a few apparent Yukaghir cognates have, instead of having degeminated forms, heterorganic consonant clusters at the position

---

16.   The study is limited to these three plosive types since these are actually the only type of geminate consonant clusters possible in Uralic.

where the Uralic item has the geminate. These anomalous items can be explained by one of two possible hypotheses. In the first hypothesis – if one assumes that the PU forms represent genetically related precursors to the chronologically later PY forms – a few apparent transformations of a plosive into a rhotic or other consonant in a condesedly constrained phonological environment have seemingly occurred. The lexical form would suggest that it is indeed a transformation and not a deletion followed by an insertion. The transformations *-kk- > *-rk- and *-pp- > *-rp-, respectively, apparently occur when the geminate, to undergo degemination, is surrounded by high back vowel *-u- on both sides; such PU proto-forms are perhaps more common than believed, as they follow vowel harmony which seems to have been a Uralic feature. In all other cases, including where only one surrounding *-u- is found, a simple degemination occurs. This take would suggest the following limited, phonological transformations:

$k > r$ / [u]_k[u] and $p > r$ / [u]_p[u]

As such, PU cluster *[u]kk[u] will find a correspondence in OY or MY *[u]rk[u] (sometimes followed by > *urqu or > *uɣu), while the PU cluster *[u]pp[u] will find a correspondence in OY or MY *[u]rp[u]. If the transformation constitutes a solid, regular sound law, it may suggest clearer proto-forms for a few PU-items where the r-transformation has occurred in the Yukaghir correspondence with regard to its vowels.

However, this hypothesis (the so-called r transformation) would not explain the presence of PY items containing *-mp-, *-pt- and *-pk- whose regular PU correspondences are *-pp-, *-pp- and *-pp-, respectively. In fact, formulating a full phonological rule to describe all these correspondences as transformations of geminates becoming the heterorganic clusters found in PY might be a very difficult or even impossible task, and would be difficult to motivate.

In defense of the r transformation hypothesis it must be mentioned that transformations of the aforementioned type are not entirely without parallel in other languages. In Liverpool English, for example, a singular intervocalic plosive /t/ is rhotacized into an alveolar tap or an alveolar lateral flap in certain phonological environments. As such, the transformation /t/ > [ɹ or ɾ] often occurs with a phonologic environment of higher back vowels, as seen in the examples: 'get off' [gɛɹɒf] and 'lot of' [lɒɹəv] (Watson 2007: 353).

There are certain loanwords in Finnic where the plosives /p/, /k/ and /t/, often before /r/ or /l/ were transformed into /u/ in western Finnish dialects (and now part of standard Finnish[17]) but were retained in certain eastern and southeastern Finnish dialects (Kettunen 1940). Examples include: Proto-Finnic *kopra > western Finnish *koura* 'hollow of the hand', Proto-Finnic *nakra- > western Finnish *nauraa* 'to

---

17.   The standard Finnish of today is a complex mixture of vocabulary taken from both Western and Eastern dialect sources.

laugh', Proto-Scandinavic *arþra*/*arðra* → Proto-Finnic *atra* > western Finnish *aura* 'plow', Proto-Finnic *hapras* > western Finnish *hauras* 'frail', Proto-Germanic *næþlō-* → Proto-Finnic *neekla* > western Finnish *neula* 'needle', Proto-Finnic *kakra* > western Finnish *kaura* 'oat', Proto-Finnic *putro* > western Finnish *puuro* 'porridge', Proto-Germanic *naglan* → Proto-Finnic *nakla-* > western Finnish *naulata* 'to nail' (Itkonen 1992; Kulonen 1995; 2000).

In phonological terms this is:

p, t, k > [*u*] / V_r or *l*.

The Finnish and English examples do indicate some kind of transformative phonological process involving /u/ and /r/ and the plosives /p/, /t/ and /k/ in various languages. The great difficulty in formulating a full phonological description for such a transformative process in Yukaghiric – in order to explain early Uralic–Yukaghiric correspondences – may mean that it is not a valid hypothesis at all. However, it will still be employed as a working hypothesis throughout the analyses.

6.3. Further hypotheses

There is another completely different but obvious second hypothesis on the aforementioned matter. In many cases, the assumption that Proto-Uralic is older language than Proto-Yukaghir allows one to posit a chain of phonological developments for the Yukaghiric items as if they had originated directly from Proto-Uralic. However, as this does not always work, one also inevitably finds that the need to explain that the complex phonological processes of the previous section would be completely eliminated if one assumes that heterorganic clusters such as *-rk-*, *-rp-*, *-mp-*, *-pt-* and *-pk-* (that is, plosive and likely any sonorant or two different plosives) were already part of proto-items before becoming the Proto-Uralic and Proto-Yukaghir forms. In other words, they would have remained unchanged into Proto-Yukaghir but become changed into geminates in Proto-Uralic. The simplification of a heterorganic cluster to a homorganic geminate following some simple rules could easily have occurred in Proto-Uralic with, for example, *-pk-* > PU *-pp-* and *-rp-*/*-mp-* > PU *-pp-*, while leaving the forms intact only into Proto-Yukaghir, which would then represent a more archaic form in regard to such clusters. Naturally, such an assumption would fall in line with the school of thought that Yukaghir is a para-Uralic language entity (with both originating from common Pre-Proto-Uralic). This second hypothesis of a genetic relationship will be discussed in footnotes throughout the lexical analyses. It is noteworthy that neither of the two hypotheses (borrowing versus genetic relationship) invalidates actual cognancy.

Also, correspondences to PU *-ü-* in PY have been found in this study, and are also presented in the next section.

7.   The PU *-pp- cluster correspondence

7.1.   Known PU *-pp- to PY correspondences

Six PU items containing the *-pp- cluster with previously suggested PY correspond-
ences (Nikolaeva 2006) were found. The phonological form of these correspondences
further suggests the presence of other similar cognates, as presented directly after
these items.

3. N. Saami *čuop'pâ- -pp-* 'to chop, to cut, to amputate', MM *ćapa-* 'to hit', EM
ćapa- 'to clap the hands', KZ *ćapkį-* 'to throw with not outstretched hand, Udm.
ćapkį- 'to beat, to strike, to clap the hands, Hung. *csap-* 'to hit, to slap, to throw' –
PFU ***ćappa*** 'to chop, to beat with popping sound' (UEW 29) – PY ***sapa*** 'to strike,
to hit' – KY *šapaɣədaj-* 'to strike, to hit' (noted in 2150), New: TY *sawɣə-* 'to crack,
to knock'.

These verbs have the same semantic meanings in both branches. No palataliza-
tion from the preceding affricate *ć* or labialization from the preceding labial *p* on
the vowels is noted. The following is suggested: *ćappa > *čappa > *sappa > *sapa*.
While the normal correspondence is between PU *ć- and PY *č- the slightly aberrant
development in this case parallels others found within Yukaghiric itself; for example
there are irregular correspondences of KY *š-* and TY *č-*, as in KY *šömör* 'top' and
TY *čumur* 'back, hill' (noted in 314) , which both likely go back to an earlier *ć-.[18]
These old suggestions seem valid. The TY item is a new addition to this cognancy
set by me as it seems to belong here from both a phonological (through the common
root *sapaɣə- > *sawaɣə- > given proto-item in 2176: *sawɣə-;[19] with the common
KY *š- <–> TY *s-* correspondence intact) and a semantic (quite closely paralleling the
meanings in, for example, PFU, Udmurt and Erzya) viewpoint.

4. Fin. *suippu* 'wedge, tip', N. Saami *čuppa* 'cacumen pilei v. caliptrae', KZ *čup*
'woman's breast', Udm. *čup* 'nipple', Hung. *csúp* 'tip, (mountain)top, hill' – New:
PFU ***ćuppV*** 'wedge, tip, point' (UEW 44) – PY ***čupo*** 'sharp' – TY *čupone* 'sharp'
(noted in 362).

These items are only slightly semantically shifted in PY from 'tip/point' to
'sharp'. As in the previous item, no palatalization or labialization effects are noted.
Previously, PFU ***ćuppa*** 'narrow' (UEW 44) has been suggested as the cognate in this
pair, but this is neither phonologically nor semantically exact. Another old suggestion
is a correspondence with the closely related PFU ***ćuppV*** 'wedge, tip, point', which
can be assumed to be ***ćuppo*** or ***ćuppu***, as indicated by the Finnish item, which is

---

18.   There are also the previously suggested correspondences between PU *šaŋkV* '(to) taste, smell',
Hung. *szag-* 'taste, smell', *szagol-* 'to smell', Selkup *sanga-*, *haaku-* 'to taste' and PY *čoŋ-*, KY *čoŋu:-*
'tasty, sweet', *čoŋčə* 'fat, lard' (noted in UED 946) and PU *ćäćä* 'trap for birds, hare or fox' (UEW
30–31), Ter Saami *šiešše* 'frame made from tree branches for bird snare', Mari *ćüćaš* 'bird noose at the
end of a bent tree branch', Khanty *sesəɣ* 'trap for grouse or fox' and PY *sas-*, KY *šašil* 'triangular trap
for hares and willow grouse' (noted in 2169) to take into account.
19.   Compare to the phonological development in Tungusic *papa-/*paba-* 'to work' > Northern Tun-
gusic *hawa-* (noted in 1993).

why I have given a corrected suggestion above. Due to the lack of an *r* transformation realized in the phonological environment of the second proposition, **ćuppo* is the most likely origin of PY **ćupo* from both phonological and semantic standpoints. The suggestion is: **ćuppo* > **čuppo* > **čupo*.

5. N. Saami *lappad* 'long tree trunks used on the sides to stop a camp fire from spreading', Mari *lə̂pə̂*, *ə̂lə̂p* 'driftwood, hay, straw', KZ *lɛp* 'driftwood pile', Mansi *lōpē* 'fallen tree, flooded tree trunk', Hung. *láp* 'moor, swamp', PFU **lappa* 'floating log' (UEW 257) – PY **law-* – KY *labut* 'floating log or snag', TY *lawur* 'heaps of tree trunks on the bank of a river' (noted in 1015).

The KY meaning parallels particularly well with those found in Mari and KZ, while the TY meaning parallels well with that found in Mansi. The PFU item, referenced as **lVppV* (UEW 257), may have been **lappa* 'floating log', as here suggested by me, which could coincide with PU **lappa* 'buckle' (UEW 236–237), cf. Fin. *lappa* 'buckle' (no Yukaghir correspondence found), another almost flat object.[20] Perhaps PU **lappa* had the meaning of 'uneven flat object'. In this case the following is suggested: **lappa* > **lapa* > **lawa* > **law-*, i.e. only a late *vowelization* of the intervocalic plosive.[21]

6. Fin. *repi-*, *reväise-* 'to tear', *reve-* 'to be ripped apart', Est. *räbi-/rebi-* 'to tug, to tear, to crack', N. Saami *râppâ-* 'to open, to take a cover off', Mansi *ript-* 'to decrease, to be destroyed, to crush, to wound', Hung. *reped-* 'to burst, to crack, to gape' – New: PFU **reppe* 'to burst, to split' – PY **lepe* 'to break off, to chip off' (noted in 1038).

These items are semantically shifted in PY to achieve an end result of loosened materials by applying force as in PU. Since PY has no initial **r-*, like some other East Asiatic languages, the correspondence to PU **r-* seems quite natural. Word-initial PU **r-* is quite rare and marginal as well. No labialization effect by the preceding labial *p* is noted. The reference is usually **rVppV-/*reppV-* (UEW 427), and I have given the suggested full proto-form above. The following is suggested: **reppe* > **leppe* > **lepe*.

7. Fin. *leppä* 'alder', Est. *lepp* 'alder', N. Saami *læi'be* 'alder', EM *ľepe* 'alder', MM *ľepä* 'alder' – PU **leppä* 'alder' (UEW 689) – PY **lep(k)-* 'blood' – KY *leppul* 'blood' (noted in 1040).

This is an interesting and possible item. Alternatively, it has been proposed that Yukaghir *leppul* 'blood' is a cognate of PU **leme/*lēme* 'broth' (UEW 245), cf. Finnish *liemi* 'broth', although this is phonologically problematic. The comparison to **leppä* 'alder', which is a reddish tree, seems more fitting (as suggested by Nikolaeva), particularly since the PY **-k-* appears to have been an affix.[22] The suggestion is thus:

---

20. This is contrasted with PFU **koppa* 'something concave or convex' (UEW 181).
21. This intervocalic process can be compared, for example with Japanese which underwent such a transformation in the 10th century. The following transformation occurred there: $p > w$ / V_[i, e, a, o], but $p > \emptyset$ / V_[u] (Frellesvig 2010: 202).
22. If it were not an infix, as suggested, an earlier proto-item would be **lepkä* from which PU **leppä* (assimilation) and PY **lepk-* would be derived, although given the rich Yukaghir suffix system this seems unnecessary and unlikely. KY *leppul* is, no doubt, also suffixed.

leppä > *leppa* > *lepa* > *lepaka* > *lepkV* > *leppV* > *leppul* 'blood'. The last step could result from labialization of the form *-*pa* > *-*po* > -*pu*, except that in this case, the *u* is epenthetic with the added suffix. The semantic development would well parallel Fin. *punainen* 'red' from PFU ***puna** 'hair' (UEW 402) in stating that something is red by talking about a reddish item.

   8. Fin. *lippi* 'trowel, small flat plate', Est. *lipp* 'thin board, washer', *lipits* 'spatula, stir stick, trowel', EM *l'ipiš*, *l'ipuža* 'one of the boards between the charge threads of a loom' – PFV ***lippV** 'shovel, board used as a spade' (UEW 690–691) – PY ***li:pə** – KY *li:pə* 'spade made of poplar for shoveling snow while putting up traps' (noted in 1059).

   While a bit controversial because the compared item is Proto-Finno-Volgaic the semantic and phonological connection between PFV and KY holds well, and cognancy seems valid. The suggestion is: *lippa > (*lipa >) **li:pa* > **li:pə* > *li:pə*.

7.2. New PU *-pp-* to PY correspondences

The above correspondences suggest that further similar cognates may be present. Indeed, the following five new cognate forms suggested by me can be found:
   9. New: Fin. *lappea* 'flat', *lappio* 'flat surface', Est. *lapp* 'flat', N. Saami *lap'pâd* 'past, without hitting, a miss', EM *lapuža* 'flat, area', MM *lapš* 'flat, area', Mari *lap* 'low', *lapka* 'flat, low', KZ *peliş-lop* 'rudder blade', *lap* 'flat, area', Hung. lap 'flatland, lowland, valley, level', Nenets *lapcā-* 'simply, to deforest' – PU ***lappe** 'flat' (UEW 237) – PY ***lewe:** 'land, earth' – KY *lebe:* 'land, earth', TY *lewejn-burebe* 'nature, homeland, lit. cover of the earth'.

   The KY and TY meanings parallel well those found in Hungarian, KZ, Moksha and other languages in relating flat surfaces to areas. As such, PU *lappa appears to be associated with being somewhat flat, while PU *lappe is completely flat. In this case the following is suggested: *lappe > *lape > *lawe > *lewe > **lewe:*, i.e. indicating only a late regressive long-distance vowel assimilation and vowel lengthening due to prosodic reasons.[23] Previously, the PY item has been suggested to be a loanword from Tungusic **lebe:(n)* 'swamp, marsh' (noted in 1047), which may be invalid given this suggestion.

   Here, an analysis of terms of kinship offers its own insights and problems. **10.** New: PU ***appe** 'father-in-law' (UEW 14), Fin. *appi(-ukko)* 'father-in-law', N. Saami *vuop'pâ* -*pp*- 'father-in-law', Mari *owấ* 'father-in-law', Hung. *ipa* 'father-in-law' and PU ***apV** 'elder female relative, aunt, older sister' (UEW 139), KZ *ob* 'father's sister', Nenets *ńāba* 'foster mother, second wife of the father, wife of older brother' should be mentioned – in analog with the above, a cognate in PY would be found as either

---

23.   Interesting to note is also that there is a Proto-Uralic *lapta 'flat' (UED 466), e.g. Nenets *lapta* 'flatland'. Could this *lapta have been progressively assimilated into the form *lappa before becoming the form found in most other Uralic languages?

*ewe(:) or *epe(:). There is PY **awa** 'elder', from which we may have KY *aboo* 'elder' and KY *abuďa* 'older sister, female cousin', but probably also TY *abučie* 'mother-in-law' and TY *abuo* 'older sister'. Some of these forms could have early gender suffixes or be semantically shifted cognates. Furthermore, there is PY **epe:**, producing KY *epe:* 'grandmother, father's older sister', which is a possible cognate with PU **apV**, unless PU **appe** and PY **epe:** are the cognates which would indeed stem from phonological development analogous to the above. The presence of PU **anV(-ppV)** 'mother-in-law' (UEW 9–10, and discussed as a compound in Janhunen 1981: 267) complicates the analysis further.

Furthermore, it would seem as if TY *apanalaa* 'wife, older woman' (from PY **apanəla:** as per 109) also belongs in this cognancy set; this compound may consist of *apa + *-n- 'genitive marker' + either *a/əla: 'a theoretically possible female counterpart to TY *aka:* 'older brother' (in 29)[24] and root for TD *alayo* 'mother' (in 38), or, *nal 'i.e. KY *ńe:l* 'daughter-in-law, sister-in-law', which is palatalized, likely from PU-like **naje** 'woman' (noted in 1337) through epenthesis, vowel lengthening and palatalization' + *a: 'a suffix'. In any case, most of these terms of kinship are likely related through one or two original roots, but the cognancies are unclear.[25]

11. New: Khanty *kopək* 'lung', Nenets *kopui, kapui, kapy* 'lung', PU **koppV**/**kuppV** 'lung' (UEW 182–183) – PY **körp-** – KY *kurpul* 'lung'.

These items are semantically identical in all aspects. The suggestion is: *kuppV > *kurpV > (*kurp-) > kurpul* (suffixation with a second syllable consonant of epenthetic form).[26] The transformation of *r* before a *p* in this phonological environment mirrors the case of the transformation of *r* before the other plosive *k* in the same phonological environment, as seen realized in the other lexicon. No labialization effects are noted.

12. New: L. Saami *tab'tē* 'spleen', Mari *lepə* 'spleen', KZ *lop* 'spleen', Udm. *lup, lub* 'spleen', Khanty *lepətne* 'spleen', Hung. *lép* 'spleen', Nenets *rapśā* 'spleen', PU **läppV**/**leppV** (~ **δäppV, -e-**) 'spleen' (UEW 242) – PY **ľa:jə** – KY *ja:jə, ľa:jə, ńa:jə* 'spleen'.[27]

---

24.   This is suggested to be a possible borrowing from Even *aka:* 'older brother' (noted in 29), although the synharmonism is noted to be irregular.

25.   If this is not the case, another possibility exists. The presence of an earlier proto-item *apke could produce PU *appe through simple progressive assimilation, as well as KY *aboo* and *abuďa* through: *apke > (*appe >) *ape (a reduction often seen in Yukaghir) > *abe (voicing) > aboo and *apke > *apoke (vowel insertion) > *apuke (labialization) > *abode > abuďa.

26.   This must assume that the PY *körp- is incorrect in its vowel. I would expect such a form from PU *kü(r)p-, which would correspond to an observed correspondence of internal PU *-ü- to PY *-ö- in other lexical data, and which is not the case here. However, as the Yukaghir form has *k, which, like *g, is through Yukaghir synharmonism associated with front vowels (as *q and *γ show up with back vowels), the PY form, not the PU form, may actually have been *kürp-. An alternative take to explain this cognancy would be to assume an earlier proto-item *kurpV from which PU *kuppV would have arisen by regressive assimilation but which was left unchanged with regard to the consonant cluster into Proto-Yukaghir. Alternatively, while requiring more parallel changes, the earlier proto-item could have been *kuprV which would have been progressively assimilated into PU *kuppV and been metathesized into PY *kurp-. The structure of *kurpul* suggests that it was a monosyllabic root which was suffixed and received an epenthetic *u*.

27.   And not the elsewhere suggested Saami-linked item *δäδwä 'lung' (Nikolaeva 1988a: 223).

These items are semantically identical in all aspects but require more phonological work than the above examples. The suggestion is: *läppe > *läpe > *lä:wə > *la:jə > *ľa:jə > ja:jə, ľa:jə, ńa:jə.[28]

13. New: Fin. leppeä 'mild, smooth, friendly', Est. lebe 'mild, silent, quiet', L. Saami lāhpōkis 'behaved, polite, friendly, good-natured', MM ľäpä 'soft', KZ liwə 'lukewarm, soft', Khanty lewət 'weak, soft', PFU *leppV 'soft' (UEW 242) – PY *limpə 'soft' – TY libene 'tender, soft'.

These PY and PFU sub-items are semantically practically identical in both the physical aspect and in mannerism. However, the suggested PY form is problematic with regard to *-mp-. While this is indeed regularly found in TY as -b- it is possible that, in this case, the PY form should be just *-p- as there are no traces of the m in any of the other words originating from the PY item. This would suggest: *leppe > *lepe > *ľepe > *lipə.[29]

14. New: Fin. tappaa 'to kill', Est. tapa- 'to kill', EM & MM tapa- 'to beat, to smash, to trample', Udm. tap- 'to be firmly trodden', tapị̄rtị̄- 'to stomp', KZ tap-tap kar- 'to knock, to beat', Hung. topp 'step', top- 'dial. to trample', tapod- 'to step, to crush', Nenets tapar- 'to step with the foot, to push, to provide a leg' – PU *tappa 'to stamp with one's feet, to hit, knock' (UEW 509) – PY *tibe – KY tibege 'to stomp one's feet, to produce hollow sounds'.

This semantic meaning seems to be almost identical between KY and those found in the older Uralic languages, including stomping like in Nenets, Hungarian, Udmurt and Mordvinic and producing sounds like in KZ while in Balto-Finnic the meaning has changed quite radically. The suggestion is: *tappa > *tapa > *taba > *ťaba > *tibe > tibege-.[30]

Two more new suggestions could be 15. New: ?Fin. suppilo 'funnel', Mari šuwə̑š 'bellows, leather bag, hose', Udm. śepị̄s 'leather bag', KZ śepị̄s 'bag', ?Khanty sous 'basket' – PFP *śoppV-sV 'sack' (UEW 776) – PY *söγ- – KY šögi: 'sack'. Semantically a sack and a bag are both movable containers. I suggest that the Finnish item belongs here both from phonological (with -lo being a suffix) and semantic perspectives (a funnel is at least structurally similar to a bellows and a hose). However, as to Yukaghiric, this is phonologically very difficult to support and may be only a chance similarity between PFP and PY.[31]

Another similar curiosity could be 16. New: KZ ćipị̄š 'salt box made of birch, basket carried on the back', Hung. csupor 'pot, pitcher', PU *ćuppV 'vessel, pot made of birch' (UEW 45) – PY *čoqo- 'pot made of birch'. Semantically, all these

28.   The last of these likely results from the hypercorrection of ja:jə by the speaker. The transformation PY *j- > KY ń- is often observed.
29.   However, another suggestion is the possibility of an earlier proto-item *lempV which could have been regressively assimilated into becoming PU *leppV but labialized in PY into *limpə. This would instead suggest the path: *lempV > PU *leppV and PY *limpə.
30.   However, it must also be noted that there is an Even verb tibelde-, tibele:n- 'to stomp one's feet, to trample down'. The directions of borrowing or the nature of cognancy is therefore unclear.
31.   Also speaking towards a chance similarity is the fact the PY root *seγ-/*söγ- is connected to concepts such as 'to enter, to get inside, to be inside'. As such, the KY noun šögi: 'sack' may have just arisen from the verbal forms and be non-connected to the Uralic items.

items concern vessels or pots made of birch (although the meaning has been lost in Hungarian where other materials are also possible), but this too is difficult to support phonologically.[32]

Fittingly, the phonological changes are supported by the fact that the PY lexicon containing *q* often has undergone the following transformations: PY **qi* & **qo* > KY **qa* and PY **qü* > KY **qe*.[33] That is, unless PY **čoqo* is formed specifically to differentiate it from the homonym **čupo* 'sharp' (as mentioned above), in which case the common proto-item **ćuppV* could still be viable. An alternative explanation can be given, similar to other *r* transformations in the study, if the proto-form is assumed to have been like PU **ćuppu*, giving the chain: **ćuppu* > **ćurpu* > **ćurku* > **ćurqu* > **ćuqu* > **čuqu* > PY **čoqo*-.[34]

In summary, while the most common correspondence of PU **-pp-* in PY is **-p-* (degemination), **-b-* (degem. → voicing) and in isolated cases **-w-* and **-j-*. Forms in Yukaghir containing rhotic sounds in the geminate position of a corresponding Uralic item suggest either a) an unusual *r* transformation under specific phonological conditions in Yukaghir or b) that a rhotic, heterorganic consonant cluster in an earlier, common proto-item was simplified in becoming the Uralic item, but retained in Yukaghir.

8.   The PU **-kk-* cluster correspondence

8.1. Known PU **-kk-* to PY correspondences

In the data, only one older cognate suggestion including the PU **-kk-* cluster could be found:

17. Fin. *kokka* 'protruding tip, upstanding, front ship', Kar. *kokka* 'stempost, hook', Est. *kokka* 'dial. (fishing) hook', N. Saami *goakke* 'pickaxe', Khanty *kaɣəw* 'wooden hook' – PU **kokka* 'hook, protruding edge' (UEW 171–172) – PY **kö:kə* – KY *kö:kə* 'fish or animal head' (noted in 860) or, alternatively, new: PY **kökö*-, KY *koknə* 'hook used to hang the kettle on a tripod over the fire'.

The semantic narrowing both in Finnish and both suggested KY items seems clear enough. Degemination works out phonologically, but the vowel differences are difficult to explain since no change is required as the forms CVCV and CV:Cə are both regularly synharmonic, and thus prosodically allowed, forms. The suggestion is

---

32.   These last two items could, however, be explained if the proto-item had been: 1) **śopkV* or even **śüpkV* (even if this second suggestion seems problematic in becoming the PU first vowel), giving PU **śoppV(-sV)* through assimilation and **śokV* before PY **söγ-*, and 2) **ćupkV*, directly giving PU **ćuppV* and in Yukaghir **ćupkV* > **ćukV* > **ćuqV* > PY **čoqo*- through a final regressive long-distance assimilation or by the uvular, lowering influence of the *q*, respectively.
33.   And uvular *q* is, in fact, known to lower following vowels, for example, in Kabardian, a Caucasian language (Choi 1991).
34.   If the rhotic sound was already present in an earlier common proto-item like **ćurpV*, we would arrive, with fewer complex phonological explanations, to homorganicity through regressive assimilation at PU **ćuppV* and with the Yukaghiric item left unchanged until it became **ćurkV* (>> PY **čoqo*-).

thus: *kokka > (*koka >) *ko:ka > *kö:kə.[35] KY *kö:kə* 'fish or animal head' would have been semantically narrowed down similarly to how the roots for 'end', 'root' or 'base' have often come to be regarded as a *head* in both the Uralic and Yukaghiric languages (as exemplified earlier in this paper). Interestingly, there is also PY ***kökö**-, KY *köknə*, which semantically fits well with the Karelian and Khanty items, which, suggested by me, might thus well instead or also constitute the cognate of the PU root *kokka unless the Yukaghir form is actually borrowed from Tungusic *goko (i.e. Yakut *köxö* 'hook') (as noted in 861). In such a case, the suggestion would be: *kokka > *koka > *koko > PY *kökö.

8.2. New PU *-kk- to PY correspondences

This section suggests up to eight other less obvious new cognates:

18. New: Mari *ĉŷĝ̂rγe-* 'to curve, to throw oneself', Udm. *ćukges* 'bent', Khanty *tŏkanə* 'bent (of a tree)', Hung. *csukorod-* 'to contract, to crouch, to squat' – PFU ***ćukkV**/***ćokkV** 'to curve, bend' (UEW 42) – PY ***čarqə-** – TY *čarqaľuu* 'bent'.

These semantic meanings are very similar throughout the languages. For the *u* to be lowered to such a degree, either the following, uvular *q* must have lowered the vowel severely and/or then it is another case of regressive long-distance vowel assimilation. The suggestion is the case of *r* transformation: *ćukku > *čurku > *čurqu > *čurqə > PY *čarqə.[36]

19. New: Fin. *johta-* 'to lead', ?EM *jaka-* 'to go', Hung. *iktat-* 'to turn on, to use, to install' or Fin. *rakas* 'darling, beloved', Khanty *rǎχ-* 'to approach', *rǎχi* 'beloved', Mansi *raw-*, *row-* 'to let sneak up closely (from the wild)', Hung. *rokon* 'relative, close' – PFU ***jakka** 'to reach, to go' (UEW 88) or PFU ***rakka** 'to be near, to approach' (UEW 418–419) – PY ***ľaqa-** – KY *jaqa-* 'to reach, come, arrive'.

The semantics are clear between KY and the older Uralic languages, while Finnish has experienced the largest shift. These work slightly better phonologically than the previously proposed correspondence with PFU ***läkte** 'to leave, to go out' (UEW 239–240) (noted in 1004). From a purely semantic viewpoint ***jakka** seems to be a somewhat closer fit of the two PFU alternatives. From a phonological viewpoint, however, ***rakka** is a better fit than ***jakka** since the transformation PY *j- > KY *ľ-* required for ***jakka** is only somewhat prevalent.[37] ***rakka**, on the other hand,

---

35.   There is also Yakut *ko:ko* 'fish head', which may be a Yukaghir borrowing.

36.   An alternative take is to assume an earlier proto-item *ćurkV which was regressively assimilated into PU *ćukkV, but which was left unchanged with the regard to the heterorganic cluster into Proto-Yukaghir (i.e *ćurkV > *čurkV > *čurqV > *čarqə > PY *čarqə).

37.   The most common transformation is PY *ľ- > KY *j-*, while the transformation *j- > KY *ń-* is also commonly observed. Isolated examples of the rarer PY *j- > KY *ľ-* are: PY *jolo- > KY *ľolože* 'to leave' & PY *ju:lə > KY *ľuuliďe* 'cunning' (both from Krejnovič's materials), PY *jompə > KY *ľo:ledelie* 'wound' & KY *ľoler-* 'to wound oneself" (both from Jochelsen's materials), PY *jowlə > KY *loullu-* 'to ask (TR)' & KY *lolo-* 'to pray, to beseech' (the second from Jochelsen's materials) and PY *jöwlə > KY *loulaγi:-* 'to feel sorry for, to treat smb well, to feel affection for (TR)'. Nikolaeva notes for the PY *j-initial lexicon that the initial KY *l-* in some forms may be due to the assimilative influence of the consonant in the second syllable (Nikolaeva 2006: 194).

corresponds well with PY *ľaqa since PY, completely lacking word-initial *r-, has a natural correspondence of *l- for Uralic items with *r-. There are thus two suggestions to entertain: *jakka > *jaka > *jaqa > *ľaqa- or *rakka > *lakka > *laka > *ľaka > *ľaqa-. In both suggestions the uvular q would possibly further lower the vowels, but these are all already low.

Some less obvious cognate possibilities of similar PU *-kk- types, since they involve more complex phonological changes, could be found. Interestingly the vowels seem relatively stable even in these correspondences.

20. New: ?Fin. *suukko* 'kiss',[38] Kar. *čokkoa-* 'to sting', Mansi *śoχəl-* 'to kiss', Hung. *csókol-* 'to kiss', *csók* 'kiss' – ?PFU ***ćukkV(-lV)*** 'to kiss' (UED 1757), POU ***ćukkV(-lV)*** (UEW 838–839) – PY ***juɣ-*** – KY *jugi:-* 'to kiss'.

The suggestion would be: *ćukku > *ćurku > *ćuɣu > *juɣ(u) > jugi:- . The corresponding PY *-ɣ- form suggests that the PFU vowel must have been *-u-, unless unspecific vowels also triggers the r transformation, and hence the above suggestion. The different initial consonants in the PFU and PY forms are difficult to explain if the Uralic form is used as the default.[39] Semantically the meanings are all the same between the languages except for in Karelian where it has undergone a radical shift. The suggestion may be valid unless the PY item is a borrowing from Tungusic *ńuka:n- 'to kiss' (as noted in 731).

21. New: Fin. *pelkkä* 'only, naked, pure', Mari *pelkə* 'coaxingly, benign, conservative', ?KZ *peľk* 'clean, pure, pretty', PFV ***peľkkä*** 'clean' (UEW 728) – PY ***pel-*** – KY *pele:-* 'to wipe dry', and TY *pilie-* 'to clean'.

The suggestion would be: (*peľkkä >) *peľkä > *peľkə > *peľə > *pel- > pele:-, pilie-. Semantically the meanings in PFV and KY are identical, while it has been verbalized in Yukaghiric.

22. New: Khanty *pĕɣtə* 'black', Mansi *piti* 'black', Hung. *fekete* 'black', POU ***pEkkV-ttV*** 'black' (UEW 882) – PY ***puɣuče:*** – KY *pugučie* 'black fly'.

The PY item is probably suffixed. The suggestion, with my suggested vowel forms, would be: *pukku(-ttV) > *purku- > *puɣu(-čV) > *puɣuče: > pugučie. That is unless one or both of the surrounding vowels were another vowel which also triggered the r transformation as seen with POU *ćukkV(-lV) above;[40] in such a case the PY form, with *-u-, would have originated through labialization by *p-. The PY *-če in this case may originate from an affricated *-tV. While it seems quite implausible from a semantic viewpoint that the literal word for black could be found in the same insect in English and in Yukaghir the phonology does hold up reasonably well. The final syllable in both the PFU and PY items are probably suffixes, possibly even the same suffix at different chronological stages.

---

38.   Although this more likely originates in the base lexical item *suu* 'mouth' with the added diminutive suffix *-kko*, i.e. 'small mouth'.

39.   A simpler phonological explanation, however, would be provided by the assumption of an earlier, common, onomatopoetic, proto-item *ťurku(-lV) which would have become PU *ćukku(-lV) through affricativization and regressive assimilation as well as *ťuɣu(-lV) > PY *juɣ(u) > KY *jugi:-.

40.   Again, an alternative explanation would be to assume an earlier, common proto-item *pErkV(-ttV), which would have become PU *pEkkV-(ttV) through regressive assimilation as well as *pEɣV(-čV) > PY *puɣuče: > KY *pugučie through labialization by *p-.

23. New: ?Fin. *sukku* 'crushed condition', Khanty *jöŋ-săk* 'frazil ice', *săk* 'finely crumbled', *săkaɣəl-* 'to suffer damage, to become broken', Hung. *szak* 'little piece, part, rubble', *szakad-* 'to tear, to break', POU **sakkV/*sukkV* 'piece, part' (UEW 891), PFU **śakkV/*śukkV* 'piece, part' (UEW 457–458) – PY **suk(sə)-* > KY *šukšə* 'piece of painted willow bark used for cleaning a gun'.

The later Uralic items are related to 'tearing crumbles', perhaps through a verbalization process, with the original meaning having been related to the nouns referring to 'part' or 'smaller piece'. I suspect that the **-šə* in KY is a causative suffix, i.e. 'thing' + suffix → 'thing used to get clean' ← 'rubble-remover' ← 'the thing that cleans by tearing away dirt'.[41] This assumes that the KY word has been semantically specified and limited in modern times to guns. The suggestion is: **śukkV > *śukV → *suk(sə)- > šukšə*.

Two less convincing cases of possible cognates are presented below without much discussion:

24. New: Fin. *sukki* 'clever, captious', N. Saami *čǫk'kâ -kk-* 'summit, mountain top', Udm. *ćuk* 'hill, top', KZ *ćuk* 'small hill, mound, hillock', Khanty *t'äk* 'top, projecting end', Mansi *ćakə.l* 'heap, hill', Nenets *soχo* 'high, round mound', PU **ćukkV* 'hill, top' (UEW 42) – PY **čo:jə* – TY *čuoje* 'hill';

A possibility: **ćukkV > *ćukV > ću:jV* > PY **čo:jə*. Semantically, the PU and TY meanings are identical.

25. New: Fin. *kalkku* 'testicle', Kar. *kalkku* 'testicle', KZ *koľk* 'testicle', PFP **kaľkkV* 'testicle' (UEW 644) – PY **kuɣe* – KY *kuge* 'testicle';

A possibility: (**kaľkku* >) **kaľku > *karku > *kaɣu > *kuɣu > *kuɣe > kuge*. Semantically, the meaning is identical throughout all the languages.

In summary, the most common correspondence of PU **-kk-* in PY are **-k-* (degemination), **-q-* (degem. → backing of plosive), **-g-* (degem. → voicing) and possibly **-j-* in isolated cases. Forms in Yukaghir containing rhotic sounds in the geminate position of a corresponding Uralic item suggest either a) an unusual *r* transformation under specific phonological conditions in Yukaghir or b) that a rhotic, heterorganic consonant cluster in an earlier, common proto-item was simplified in becoming the Uralic item, but was retained in Yukaghir (for example: the earliest proto-item **-rk-* > PU **-kk-*, but was retained as PY **-rk-* (and occasionally followed by > **-ɣ-*).

9.   The PU **-tt-* cluster correspondence

9.1.   Known and new PU **-tt-* to PY correspondences

Only one new cognate suggestion containing PU **-tt-* could be found in the data set:

26. New: Udm. *jiti̯* 'to bind, to connect', KZ *jit-* 'to connect, to enforce, to sew', Nenets *jūt'e-* 'to attach with stitches' – PU **jutta* 'to bind, to attach' (UEW 106) – PY **joðo-* – KY *jodo-* 'to bind'.

---

41.   And, indeed, **-šə* has been presented as a transitive causative suffix in Late Proto-Yukaghir (Nikolaeva 2006: 83).

Semantically and phonologically these are more fitting between all the languages than the previously suggested (noted in 697) cognancy to PFU *jorkV- 'to turn, to wind' or PFU *jorV 'to roll' (UEW 102). Here, a lowering effect on the vowel(s) can be noted. Beyond this, no new cognate suggestions related to any of thirty found P(F)U items containing the *-tt- cluster could be found.

The correspondence of PU *-tt- would seem to be *-t- (degemination) or *-δ- (degem. → aspiration) in PY.

10.   The Late Proto-Yukaghiric *-ö- correspondence to PU *-ü-

10.1. New PU *-ü- to PY correspondences

So far, the literature has not reported any correspondences in Yukaghir to the Proto-Uralic vowel *ü (examples are missing in the summary of Häkkinen 2012). In this study, some adequate correspondences have been found and are therefore presented here in greater detail. The following new, suggested cognates treating this correspondence, namely PU *-ü- <–> PY *-ö-, are given here:

27. New: Fin. *sylki* 'spit', *sylkeä* 'to spit', Est. *sülg* 'spit', N. Saami *čoḷgâ -lg-* 'saliva', EM *śeḷge* 'spit', MM *śeḷgä* 'spit', Mari *šüwala* 'to spit' → PFV *śulke* 'to spit' (UEW 778), Udm. *śal* 'spit', KZ *śęlal-* 'spit', Khanty *sŏjəɣ-* 'to spit', Mansi *süĺk-* 'to spit' < PFU *śüĺke*/*śiĺke* '(to) spit' (UEW 479) <–> PY *söɣe:/*söŋke:* > KY *sögie* 'saliva'. The suggestion is: *śüĺke > *śöĺke > *söɣe(:)*. Semantically, the meanings are identically related to the noun referring to *saliva/spit* and the verbal action of the *spitting of saliva* in all the languages.

28. New: Fin. *kynsi* 'nail, claw', Est. *küüs* 'nail', N. Saami *gâʒʒa -ʒʒ-* 'nail, claw' EM *kenže* 'claw, nail, hoof', Mari *kəč* 'nail', Udm. *giʒi* 'nail, claw', KZ *giž* 'nail, claw', Khanty *kŏ nÚč* 'nail, claw', Mansi *künš* 'nail, claw', Nenets *χada* 'nail, claw' < PU *künče* 'nail' (UEW 157) <–> PY *önć-* > KY *uďiĺ* 'nail' & PY *kö:nćikiĺ* > KY *kuoďikiĺ* 'two small nails on the rear of the front legs of a reindeer, i.e. the dewclaws of a reindeer'. Semantically, the items throughout all the PU and PY languages are closely related to 'claws' and 'nails'. Phonologically the loss of the initial consonant into PY may seem odd, but this can also be seen in other old suggested and structurally very similar cognancy correspondences, such as for example: Fin. *kusi* 'urine', N. Saami *goʒʒa-* 'urine', KZ *kuʒ́* 'urine', Hung. *húgy* 'urine', Kamas. *kənze* 'urine', PU *kuńće* 'urine' – PY *ončə-*, KY *o:ži:* 'water'.[42]

29. New: Fin. *kynsiä* 'to claw', and multiple verbal forms of the previous item < PU *künče-* 'to claw' (UEW 157) <–> PY *könčə-* > KY *köže-* 'to scrape, to scratch, by a bear'. The suggestion is: *künče > *könče > *könčə*. Semantically, the meanings are all identical and have become somewhat semantically limited in KY.

These further suggest two new, less clear possibilities:

---

42.   Semantic parallels can be made with, for example, Swedish *att kasta vatten* 'to urinate, lit. to throw water'.

30. New: Mari *nörgä* 'sprout, young shoot', Udm. *ńẹr* 'rod, branch, whip', KZ *ńẹr* 'sprout, young shoot', Khanty *ńĕr* 'a grove in a burned place', *ńăr* 'hardwood grove', Mansi *ńär* 'pipe, rod, branch', Hung. *nyír* 'birch', Nenets *ńerū* 'willow', PU ***ńürkä*** (assumed from **ńVrkV* (UEW 331)) 'rod; young shoot' > PY ****ńörkə*** > KY *nörgəmtu*: 'small dry twigs of larch used for kindling fire; brushwood'. The suggestion is: **ńürkä > *ńörke > *ńörkə*. Semantically, this item has become specialized in most Uralic branches, as it has in KY. This may clarify this previously given cognancy (noted in 1520).[43]

31. New: Fin. *kyynär(-pää)* 'elbow', Est. *küünar* 'forearm, ulna', *küünarpää* 'elbow', N. Saami *gâr'dnjel*, *gâi'gŋer* 'elbow', EM & MM *keńeŕ* 'forearm, elbow', Mari *kəńer* 'ulna', *kə̂ńer-wuj* 'elbow', Udm. *gịr-pum* 'elbow, ulna', KZ *gịr* 'half a yard', *gịrźa* 'elbow', Khanty *kŏnǓŋi* 'elbow', Mansi *koänɣel'* 'elbow', Hung. *könyök* 'elbow', PFU ***küńä*** 'elbow' (UEW 158–159) > PY ****qač-*** > TY *qačil* 'elbow, riverbend'

This case requires further step-wise clarification. The following is suggested: PU **küńä > *köńə > *köńə-če* (suffixation[44]) *> *köńče > *köćə > *qöćə > *qaćə* (uvularization effect on the first vowel: **qö- > *qa-* (front)) *> *qaći* (patalazation effect on the second vowel) > PY ****qač-*** > TY *qačil* (suffixation). Semantically the TY item is identical in meaning to that found in the various Uralic languages with further extended meanings as a descriptive term.

As for word-initial PU **ü-* we may have a case of cognancy with Fin. *yksi* 'one', N. Saami *okta* 'one', MM *ifkä* 'one', Udm. *odịg* 'one', KZ *ẹťik* 'one', Mansi *ük* 'one', PFU ***ükte*/*ikte*** 'one' (UEW 81). A regular correspondence in PY may be **i-* as found in KY *irke* 'one', where the form would quite closely mirror that found in Moksha, if this is even a cognate and not highly anomalous, although it may alternatively be a question of a direct correspondence between two **i-* items.

11. Regarding the chronology of sound changes

11.1. Basic chronology of change

Recently, The Yukaghiric proto-language has been divided into Early Proto-Yukaghir (EY), Middle Proto-Yukaghir (MY) and Late Proto-Yukaghir (PY). Clearly, (Late) Proto-Yukaghir had no geminate plosive clusters. It may also have had the vowel *ü*, although this is lacking in both Kolyma and Tundra Yukaghir.

There are no traces of the geminate cluster even in Late Proto-Yukaghir (except for a few cases where the PY item has a heterorganic cluster consisting of a sonorant and a plosive, and where the PU item has a homorganic, geminate structure) which suggests that degemination was quite early, most likely in either Early or Middle

---

43.   However, there is also Even *ńirguqi* 'brushwood'. As such, this may be an Even loanword.
44.   Cf. PY **molqə* 'joint' > KY *molɣiče* 'lit. end of joint, i.e. joint'. Cf. PU **küńä/*kińä* (UEW 158–159) > Finnish *kyynärpää* 'lit. head of the elbow, i.e. elbow'. The suffix *-če* here would turn the original **küńä* 'elbow' into 'end of the elbow', a semantic transformation very similar to what is seen in Finnish.

Proto-Yukaghir, and thus occurred simultaneously as some vowel changes were going on. A possibility is that the geminated, stressed word was quite resistant to vowel changes, which explains why a later changed degeminated form would have relatively unchanged vowels in comparison to the Uralic form. As is seen in the presented dataset, there have sometimes been quite varied vowel changes from Late Proto-Yukaghir, long after the plosive degemination had occurred.

11.2. Further hypotheses of change

The hypothesis of Yukaghir as a para-Uralic entity would considerably simplify the phonology involved in explaining apparent cognancy. In this hypothesis, all proto-items studied would originally have had either geminate plosive clusters or heterorganic consonant clusters consisting of a sonorant and a plosive or two different plosives. While the geminate clusters would have been retained in Uralic, they would have been degeminated in Yukaghir. The heterorganic clusters, however, were retained in Yukaghir – thus exemplifying archaic lexical forms – but were sometimes assimilated into homorganic, geminate consonant clusters in Uralic.[45] It is thus possible that certain heterorganic clusters were regularly simplified into geminate items in Proto-Uralic as influenced by surrounding vowels, which would constitute a regular sound law for certain combinations. This simplification into PU by geminization, however, would have been a limited event since there were items containing either two different plosives or a sonorant and a plosive in Proto-Uralic.[46] The phonological simplification of such clusters in Uralic would have occurred at a very early stage going from Pre-Proto-Uralic to Proto-Uralic.

Schematically in going from a tentative pre-Proto-Uralic into Proto-Uralic and Early Proto-Yukaghir (C = consonant, V = vowel, S = sonorant, P_1 = plosive, P_2 = another plosive) in this hypothesis is the following:

$*(C)VP_1P_1V$ > PU unchanged and EY $*(C)VP_1V$
$*(C)VP_1SV$ / $*(C)VSP_1V$ > PU unchanged or PU $*(C)VP_1P_1V$ and EY unchanged
$*(C)VP_1P_2V$ > PU unchanged or PU $*(C)VP_1P_1V$ and EY unchanged

---

45.   Assimilation would have occurred in either a progressive or regressive fashion. The governing factor in this would seem to have been the retaining of the dominant plosive of the cluster (with *p* having priority over *t* and *k*, and with the sonorant always being assimilated into becoming a plosive) in the cases when it happened.

46.   Examples: PU **kupsa* 'to extinguish' (UEW 214–215) cf. Estonian *kustu-* 'to extinguish', KZ *kus* 'to extinguish' & Nganasan *kabtu-* 'to extinguish'. PU **kakta/*käktä* 'two' (UEW 118–119), cf. Finnish *kaksi* 'two', KZ *ki̯k* 'two' & Nganasan *siti* 'two'.

11.3. Some notes regarding vowel changes

A noteworthy observation is PU *-*ü*-, which has a chronologically early correspond-ence, possibly in MY *-*ö*- (examples in the previous section) and PU *-*ä*-, which has the correspondence -*e*- in EY and -*a*- (front) in MY, as well as PU *-*ä*, which has the correspondence of (MY?) Proto-Yukaghir *-*e* or *-*ə* depending on prosody (personal observations). All of these seem to either predate or be contemporaneous with the plosive degemination in EY or MY. Degeminated items did, of course, some-times have further quite expected sound changes in becoming the KY and TY forms. Furthermore, in several cases, it seems as if the loss of one of the original consonants is compensated by lengthening one of the vowels in Late Proto-Yukaghir, although this is likely controlled by prosody. Correspondences do suggest that Yukaghir is a language group with relatively archaic features.

12.   Summary and conclusions

12.1. Regarding the development from early geminate plosive clusters

As is clear, most data is available regarding the *-*pp*- clusters, with only a few exam-ples of the *-*kk*- and *-*tt*- clusters. In almost all cases, the Yukaghir items were found to be degeminated. The following correspondences apply:

PU *-*pp*- – PY *-*p*- (degemination), *-*b*- (degem. → voicing) and isolated cases of *-*w*- and *-*j*- (but also spurious *-*rp*-, *-*pt*-, *-*pk*- and *-*mp*- in a few cases, unless these are part of earlier, common proto-items that got assimilated into *-*pp*- in Proto-Uralic only).

PU *-*tt*- – PY *-*t*- or *-*δ*- (degem. → aspiration).

PU *-*kk*- – PY *-*k*- (degemination), *-*q*- (degem. → backing of plosive), *-*γ*- (degem. → aspiration) and possibly *-*j*-, depending on several factors (but also *-*rk*- , unless this is part of earlier proto-items that got assimilated into *-*kk*- only in Proto-Uralic).

The factors governing the exact phonological outcome and features in PY – for items with PU cognancy – requires further investigation in order to provide a full and sat-isfactory description.

A phonological curiosity is found in what appears to be unusual *r* transforma-tions: $p > r$ / $[u]\_p[u]$ and $k > r$ / $[u]\_k[u]$. While no such examples are found in the data, it seems conceivable that this *r*-transformation, if correct, would also have occurred with *-*tt*- clusters surrounded by back high vowels. This transformation,

however, does not adequately explain the apparent cognancy between homorganic Uralic items and Yukaghir items with non-rhotic, heterorganic consonant clusters, which is why an alternative hypothesis must be considered.

12.2. Regarding found cognates

Using the principles found in the comparison, several new likely cognates – all within semantically acceptable limits – some more obvious, some less obvious, were found. If the tentative second hypothesis of earlier, common proto-items presented throughout the text is correct, then it appears as if at least some of the Proto-Uralic geminate clusters could be the result of early assimilations which did not occur in the Yukaghir branch. The advantage of this hypothesis is that the aforementioned phonological curiosity, the so-called *r* transformation given above, could be completely eliminated in favor of a simplified phonological explanation.

In this hypothesis, the assumption of earlier common proto-items existing with heterorganic consonant clusters – thus assuming a genetic relationship between Uralic and Yukaghir at high time depth – directly gives the homorganic geminate clusters in Proto-Uralic through simple assimilation governed by very basic rules, while such complex clusters would have been archaically retained partly or completely in the modern Yukaghir languages.

Furthermore, the regular sound correspondence PU *-*ü*- to PY *-*ö*- is found and suggested with some quite convincing cases. A few items bear traces of early suffixation which have been grammaticalized into the modern lexical items.

Despite a general lack of historical, archeological, genetic data etc. using the outlined methodology it has been possible to find new cognates between the language groups and, furthermore, very tentatively a small number of weakly established even earlier proto-items, which, if valid, would suggest a genetic relationship. The results do indicate that correspondences between Uralic and Yukaghir are more widespread than previously thought, and that further cognates can be systematically found. Triple the amount (24) of possible noun cognates in relation to verb cognates (8) are presented and discussed in this paper, totaling many new cognate suggestions of which seven are non-geminate. Naturally, any such new cognates will give valuable insights in future research on the sound changes and nature of the Uralic–Yukaghiric relationship.

In summary, the lexical forms of consonants of some apparent cognate correspondences seem to imply – unless complex and previously non-described transformations or epenthetic modifications can be subscribed to – that certain phonological changes have occurred in the Uralic branch while not in the Yukaghir branch and vice versa. This would eventually lead to the conclusion that the two groups stand in the relationship of genetic para-languages.

Abbreviations

EM	= Erzya Mordvin		MY	= Middle Proto-Yukaghir	
Est.	= Estonian		N. Saami	= Northern Saami	
EY	= Early Proto-Yukaghir		PFP	= Proto-Finno-Permic	
Fin.	= Finnish		PFU	= Proto-Finno-Ugric	
Hung.	= Hungarian		PFV	= Proto-Finno-Volgaic	
Kamas.	= Kamassian		PS	= Proto-Samoyed	
Kar.	= Karelian		PU	= Proto-Uralic	
KY	= Kolyma Yukaghir		PY	= Late Proto-Yukaghir	
KZ	= Komi-Zyrian		TY	= Tundra Yukaghir	
MM	= Moksha Mordvin		Udm.	= Udmurt	

References

Angere, Johannes 1956: *Die uralo-jukagirische Frage. Ein Beitrag zum Problem der sprachlichen Urverwandschaft*. Stockholm: Almqvist & Viksell.

Bouda, Karl 1940: Die finnisch-ugrisch-samojedische Schicht des Jukagirischen. – *Ungarische Jahrbücher* 20: 80–101.

Campbell, Lyle 2004: *Historical linguistics: An introduction*. 2nd Ed. Edinburgh: Edinburgh University Press.

Campbell, Lyle 2011: Review of the Dene-Yeniseian connection (Kari and Potter). – *International Journal of American Linguistics* 77: 445–451.

Choi, John D. 1991: An acoustic study of Kabardian vowels. – *Journal of International Phonetic Association* 21: 1–12.

Collinder, Björn 1940: *Jukagirisch und Uralisch*. – Uppsala Universitets Årsskrift, 8. Uppsala: Almqvist & Wiksell.

Collinder, Björn 1957: *Uralo-jukagirische Nachlese*. – Uppsala Universitets Årsskrift 12: 105–130.

Collinder, Björn 1965a: *An introduction to the Uralic languages*. Berkeley & Los Angeles: University of California Press.

Collinder, Björn 1965b: *Hat das Uralische Verwandte? Eine Sprachvergleichende Untersuchung*. – Uppsala Universitets Årsskrift 1: 109–180.

Dolgikh, Boris O. 1960: *Rodovoj i plemennoj sostav narodov Sibiri v XVII v.* Moskva: Izdatel'stvo Akademii Nauk SSSR.

Frellesvig, Bjarne 2010: *A History of the Japanese Language*. Cambridge University Press.

Fortescue, Michael 1998: *Language Relations across Bering Strait: Reappraising the Archaeological and Linguistic Evidence*. London & New York: Cassell.

Gruzdeva, Ekaterina 1998: *Nivkh*. München: Lincom Europa.

Harbert, Wayne 2007: *The Germanic Languages*. Cambridge University Press.

Haspelmath, Martin 2008: Loanword typology: Steps toward a systematic cross-linguistic study of lexical borrowability. – Thomas Stolz, Dik Bakker & Rosa Salas Palomo (eds), *Aspects of language contact: New theoretical, methodological and empirical findings with special focus on Romancisation processes*. Berlin: Mouton de Gruyter. 43–62.

Hock, Hans H. & Joseph, Brian D. 2009: *Language history, language change and language relationship. An introduction to historical and comparative linguistics*. 2nd ed. Berlin: Mouton de Gruyter.

Honti, László 1995: Zur Morphotaktik und Morphosyntax der uralischen/finnisch-ugrischen Grundsprache. – *Congressus Internationalis Fenno-Ugristarum* 8(1): 53–82.

Häkkinen, Jaakko 2011: *Is Finnish an archaic language?* <http://www.mv.helsinki.fi/home/jphakkin/Archaic_Finnish.pdf> (14.09.2012).

Häkkinen, Jaakko 2012: Early contacts between Uralic and Yukaghir. – Tiina Hyytiäinen & Lotta Jalava & Janne Saarikivi & Erika Sandman (eds), *Per Urales ad Orientem. Iter polyphonicum multilingue. Festskrift tillägnad Juha Janhunen på hans sextioårsdag den 12 februari 2012.* Mémoires de la Société Finno-Ougrienne 264. Helsinki: Suomalais-Ugrilainen Seura. 91–101.

Itkonen, Erkki 1966: *Suomalais-ugrilaisen kielen- ja historiantutkimuksen alalta*. Tietolipas 20. Helsinki: Suomalaisen Kirjallisuuden Seura.

Janhunen, Juha 1981: Uralilaisen kantakielen sanastosta. – *Journal de la Société Finno-Ougrienne* 77: 219–274.

Janhunen, Juha 2009: Proto-Uralic – what, where, and when? – Jussi Ylikoski (ed.), *The Quasquicentennial of the Finno-Ugrian Society*. Mémoires de la Société Finno-Ougrienne 258. Helsinki: Suomalais-Ugrilainen Seura. 57–78.

Kettunen, Lauri 1940: *Suomen murteet III A: murrekartasto*. Helsinki: Suomalaisen Kirjallisuuden Seura.

Krejnovič, Eruhim Abrahamovič 1958: *Ûkagiriski jâzyk*. Moskva & Leningrad.

Lainio, Jarmo 1995: Sweden Finnish at the cross-roads of code-switches, incorporations and integrations. – Inger Moen & Hanne Simonsen & Helge Lødrup (eds), *Papers from The XVth Scandinavian Conference of Linguistics, Oslo, January 13–15 1995*. Department of Linguistics, University of Oslo. 277–294.

Larsson, Lars-Gunnar 2001: Baltic influence on Finnic languages. – Östen Dahl & Maria Koptjevskaja-Tamm (eds), *The Circum-Baltic Languages: Typology and Contacts. Volume 1: Past and Present*. Amsterdam & Philadelphia. 237–253.

Lass, Roger 1994: *Old English: A historical linguistic companion*. Cambridge: Cambridge University Press.

Lewy, Ernst 1928: Possessivisch und Passivisch. Bemerkungen zum Verbalausdruck in der sprachlichen Typenlehre. – *Ungarische Jahrbücher* 8: 274–289.

Maslova, Elena 2003a: *Tundra Yukaghir*. Languages of the World, Materials 372. München: Lincom Europa.

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

Mikola, Tibor 1988: Geschichte der Samojedischen Sprachen. – Denis Sinor (ed.), *The Uralic Languages. Description, history and foreign influences*. Leiden. 219–263.

Nelson-Sathi, Shijulal & Johann-Mattis List & Hans Geisler & Heiner Fangerau & Russel D. Gray & William Martin & Tal Dagan 2011: Network uncover hidden lexical borrowing in Indo-European language evolution. – *Proceedings of the Royal Society B*. 278: 1794–1803.

Nichols, Johanna 1992: *Linguistic Diversity in Space and Time*. Chicago–London: The University of Chicago Press. (Reprinted in 1999.)

Nikolaeva, Irina 1988a: *The Problem of Uralo-Yukaghir Genetic Relationship (in Russian)*. PhD thesis. Moscow: Institute of Linguistics.

Nikolaeva, Irina 1988b: On the correspondence of Uralic sibilants and affricates in Yukaghir (in Russian). – *Sovetskoe Finnougrovedenie* 2: 81–89.

Nikoleava, Irina 2008: Chuvan and Omok languages? – Alexander Lubotsky, Jos Schaeken & Jeroen Wiederhof (eds), *Evidence and Counter-Evidence. Essays in honour of Frederik Kortlandt. Volume 2: General Linguistics*. Amsterdam: Studies in Slavic and General Linguistics. 313–336.

Odé, Cecilia 2009: *Description of the Tundra Yukaghir project, March 2009–March 2012*. <http://home.medewerker.uva.nl/c.ode/bestanden/Tundra%20yukagir%20project%20 abstract%20sept%202009.pdf> (11.09.2012).

Paasonen, Heikki 1907: Zur Frage von der Urverwandschaft der finnisch-ugrischen und indoeuropäischen Sprachen. – *Finnisch-Ugrische Forschungen* 17: 13–31.

Pedersen, Holger 1931: *Linguistic Science in the Nineteenth Century: Methods and Results*. Translated from the Danish by John Webster Spargo. Cambridge, Massachusetts: Harvard University Press.

Rédei, Károly 1988: Geschichte der Permischen Sprachen. – Denis Sinor (ed.), *The Uralic Languages. Description, history and foreign influences*. Leiden. 351–394.

Rédei, Károly 1999: Zu den uralisch-jukagirischen Sprachkontakten. – *Finnisch-Ugrische Forschungen* 55: 1–58.

Sammallahti, Pekka 1988: Historical Phonology of the Uralic Languages. – Denis Sinor (ed.), *The Uralic Languages. Description, history and foreign influences*. Leiden. 478–554.

Sauvegeot, Aurélien 1969: La position du youkaguir. – *Ural-Altaische Jahrbücher* 41: 344–359.

Seefloth, Uwe 2000: Die Entstehung polypersonaler Paradigmen im Uralo-Siberischen. *Zentralasiatische Studien* 30: 163–191.

Suhonen, Seppo 1988: Die Baltischen Lehnwörter der Finnish-ugrischen Sprachen. – Denis Sinor (ed.), *The Uralic Languages. Description, History and Foreign Influences*. Leiden. 596–615.

Tailleur, Oliver Guy 1959: Plaidoyer pour le youkaghir, branche orientale de la famille ouralienne. – *Lingua* 6: 403–423.

Thomason, Sarah Grey & Kaufman, Terrence 1988: *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.

van Hout, Roeland & Muysken, Pieter 1994: Modeling lexical borrowability. – *Language Variation and Change* 6: 39–62.

Watson, Kevin 2007: Liverpool English. – *Journal of the International Phonetic Association* 37: 3: 351–360.

Wickman, Bo 1969: Die Verwandtschaft des Finno-Ugrischen mit anderen Sprachen. – *Uralaltaische Jahrbücher* 41: 310–316.

Winkler, Heinrich 1912: Samojedisch und Finnisch. – *Finnisch-Ugrische Forschungen* 12: 115–127.

Wurm, Stephen Adolphe 2001: *Atlas of the World's Languages in Danger of Disappearing*. Unesco Publishing.

Etymological References

Finnish and Yukaghiric Etymological dictionaries

Collinder, Björn 1955: *Finno-Ugric Vocabulary. An Etymological Dictionary of the Uralic Languages*. Uppsala. (Includes many Yukaghir correspondences.)

Häkkinen, Kaisa 2011: *Nykysuomen etymologinen sanakirja*. Helsinki: Sanoma Pro.

Itkonen, Erkki 1992: *Suomen sanojen alkuperä. Etymologinen sanakirja. 1. A–K*. Suomalaisen Kirjallisuuden Seuran toimituksia 556 & Kotimaisten kielten tutkimuskeskuksen julkaisuja 62. Helsinki: Suomalaisen Kirjallisuuden Seura & Kotimaisten kielten tutkimuskeskus.

Kulonen, Ulla-Maija 1995: *Suomen sanojen alkuperä. Etymologinen sanakirja. 2. L–P*. Helsinki: Suomalaisen Kirjallisuuden Seura & Kotimaisten kielten tutkimuskeskus.

Kulonen, Ulla-Maija 2000: *Suomen sanojen alkuperä. Etymologinen sanakirja. 3. R–Ö*. Helsinki: Suomalaisen Kirjallisuuden Seura & Kotimaisten kielten tutkimuskeskus.

Kurilov, Gavril Nikolaevich 1990: *Yukaghir-Russian dictionary*. Jakutsk: JaGu.

Kurilov, Gavril Nikolaevich 2001: *Yukaghir-Russian dictionary*. Novosibirsk: Nauka.

Nikolaeva, Irina 2006: *A Historical Dictionary of Yukaghir*. Trends in Linguistics Documentation 25. Berlin & New York: Mouton de Gruyter.

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

SKES = Yrjö Henrik Toivonen, Erkki Itkonen *et al*. (eds) 1955–1981: *Suomen kielen etymologinen sanakirja*. Lexica Societatis Fenno-Ugricae XII. Helsinki: Suomalais-Ugrilainen Seura.

UEW = Károly Rédei (ed.) 1988–1991: *Uralisches Etymologisches Wörterbuch*. Budapest: Akadémiai Kiadó.

UED: The Uralic Etymology Database online – <http://starling.rinet.ru> (includes some Uralic and Yukaghiric correspondences – the database is mostly based on UEW materials).

Yukaghir-related Etymological dictionaries and glossaries

The etymological reference works consulted in Nikolaeva (2006), the source of Yukaghir used for this research, were given with common short-hand: CED (Fortescue, M. et al. 1994), DEWOS (Steinitz, W. 1966–1993), EDAL (Starostin, S. A. et al. 2003), ESRD (Anikin, A. E. 2000), ESRZ (Anikin, A. E. 2003), JRS (Slepcov, P. A. 1972), JU (Collinder, B. 1940 – notes Yukaghir parallels and arguments), HUV (Collinder, B. 1965), LR (Fortescue, M. 1998) and UJN (Collinder, B. 1957). Further, the following glossaries were consulted and collected therein: Angere, J. (1957, based on Kurilov, G. N. 1900), Kurilov, G. N. (2001), Veenker, W. (1989), Endo, F. (1997; 2001), Nyikolajeva, I. (2000), Nikolaeva, I. & Shalugin, V. (2003) and Maslova, E. (2001; 2003).