NIKLAS METSÄRANTA   (Helsinki)

'SISTER', 'DAUGHTER' AND 'BROTHER' — ETYMOLOGICAL DISCUSSION OF A FEW BORROWED URALIC KINSHIP TERMS

Abstract. In this article, three Indo-European loan etymologies for Uralic kinship terms meaning 'sister', 'daughter' and 'brother' are discussed and a thorough etymological analysis of the words is given based on the latest research. An attempt is made to try and untangle the multitude of ways in which the relationship of these words can be interpreted and to provide the most probable scenario for whence the words entered the lexicons of Uralic languages. Both the phonological and the semantic side — which has often been greatly neglected in Uralic etymology — of the loan etymologies are explored.

Keywords: Uralic languages, etymology, loanword research, kinship terms.

1. Introduction

In this article, three Indo-European loan etymologies for Uralic kinship terms meaning 'sister' (MdE *sazor* 'younger sister', MdM *sazər* ~ MariM *šüžar*, MariH *šǯžar* ~ Udm *suzer,* Komi *sozor* 'Fadenbruch (im Gewebe)'), 'daughter' (Fi *tytär*, Veps *tütär*, Est *tütar*, Liv *tidār* ~ SaaS *dektier*, *daktere* 'married daughter' ~ MdE *ťejťeŕ*, MdM *śťiŕ*) and 'brother' (Fi *veli* 'brother', Veps *veľľ*, Est dial. *veli* 'brother; bride's brother in a wedding', Liv *veʼḷ* ~ SaaN *viellja*, SaaS *vielle*, SaaSk *villj*) are discussed in detail.

There are only a few borrowed kinship terms in Uralic languages that, based on their distribution in more than one branch of Uralic, could be considered shared loanwords into the common ancestor of these languages. These three examples are all confined to the more western half of the Uralic languages, which already seems to indicate a post-Proto-Uralic origin. In this article, a detailed etymological analysis of the words utilising the most up-to-date research is given. From what follows, it should become obvious that the etymological relationship of these words can be interpreted in a multitude of ways. An attempt is made to untangle this multitude and to provide the most probable scenario for whence the words entered the lexicons of the Uralic languages in question. The discussion focuses on the relative

chronology of the loanwords; the absolute chronology falls outside the scope of this article.

Historical phonology has often dominated the discussion in loanword research. In this article, too, phonology is given a lot of attention, but also the semantic side — often neglected in etymological research in Uralic studies — is considered carefully. In particular, the semantic parallels provided to establish the loan etymology for 'brother' (Early Proto-Finnic *weljə* 'brother' ← Proto-Germanic *sweljan-* 'brother-in-law' (LägLoS III 386—387)) are reflected upon meticulously by making use of the kinship-term material collected for Kinura.[1]

2. Etymologies

2.1. 'Sister'

One of the few potential shared loanwords is the word for 'younger sister', MdE *sazor*, MdM *sazər*, MariM *šüžar*, MariH *šə̂žar* and Udm *suzer*, Komi *sozor* 'Fadenbruch (im Gewebe)' < FP *sasare*, that according to the traditional view was borrowed into Proto-Finno-Permic from Proto-Indo-Iranian or Early Proto-Iranian *swasar-* 'sister' (UEW 752—753). It should be noted that the Finno-Permic languages seem to not have shared any actual innovations, and as such a Proto-Finno-Permic proto-language might not be an actual node (Salminen 2002) and thus should be approached with scepticism. In terms of lexicon (Proto-)Finno-Permic, in the way it is used in the UEW, is largely a designation of distribution. As Proto-Finno-Permic, as well as most other intermediary proto-languages between Proto-Uralic and separate branch-level proto-languages such as Proto-Finno-Ugric, Finno-Volgaic etc., have lost favour, the chronology of loanwords like PII *swasar-* 'sister' has to be reconsidered. In recent research, the solution seems to have been to categorise them as loans into Proto-Uralic, with the caveat that either the reflex of PU *sasarV* was later lost in all other branches or that the word was borrowed into neighbouring Proto-Uralic dialects that later developed into Mordvin, Mari and Permic (Holopainen 2019 : 222—224).

The above-mentioned Mordvin, Mari and Permic words have sometimes been considered in tandem with Finnic words for 'sister', Fi *sisar*, Veps *sizar*, Votic *sẹsar*, Est *sõsar*, EstS *sysaŕ*, Liv *sõzār* etc. The Finnic words are usually thought to reflect a loan from Baltic, cf. Lith *sesuõ*: *seseȓs* 'sister' (SSA 3: 187—188), that according to the authors of SSA could also underlie MariM *šüžar*, MariH *šə̂žar*. This possibility is not further discussed by Holopainen (2019),

---

[1] Kinura is an interdisciplinary research project (funded by Kone Foundation) studying the evolution of kinship relations and contacts of Uralic speaking populations. The linguistic side of the project has initially involved the collection of kinship terms from all branches of Uralic languages (at present around 20 languages ranging from Saami to Samoyedic). In addition to collecting the kinship terms themselves, other information has been collected, including the etymologies featured in research literature. In order to study the borrowability of different kinship terms, the etymology of kinship terms are given a numerical value between 0 and 4, based on the scale outlined by Haspelmath and Tadmor (2009 : 12—13) by which the reliability of the loan etymology is assessed. This article has been written in part to provide a reasoning for why our assessment of the probability of borrowing of certain terms, especially 'brother' in Finnic and Saami, might differ from what has been considered probable and why instead of counting the words for 'sister' in Mordvin, Mari and Permic as one loanword, we rather count them as three separate loanwords.

although it is acknowledged that a Baltic origin alongside Indo-Iranian has been considered possible for both Mari and Permic in earlier literature, cf. Joki 1973. The vowel correspondences between the Finnic languages are irregular, as Finnish, Veps and South Estonian reflect PF *sisar*, while Estonian, Votic and Livonian reflect PF *sẹsar*, which has been taken to indicate that they actually reflect two separate loanwords from Baltic (Kallio 2018 : 255). However, it is unclear how separate borrowing would actually explain the forms with *i*. Whatever the explanation behind the *i*, it can hardly be a secondary post-Proto-Finnic development considering its distribution in South Estonian.

It has been argued that deriving the Mordvin, Mari and Permic words from Baltic is complicated, and as at least Mordvin and Permic clearly reflect an earlier PU *a-a* stem, an Indo-Iranian origin is probable and the Baltic origin of the words even impossible (Holopainen 2019 : 222—223). The Mari word is interpreted to reflect an independent borrowing from ultimately the same Indo-Iranian source as the Mordvin and Permic words (Holopainen 2019 : 222—223). The situation is admittedly complicated and the phonology of the Mordvin, Mari and Permic words could be interpreted in many ways. Let us take a closer look at some of the phonological details presented.

Erzya *sazor* and Moksha *sazər* can reflect PU *a-a* as mentioned, although it is not the only possibility and PMd *a* is equally likely to reflect PU *ẹ* (Aikio 2015 : 39). This observation does not seem to provide any real new insights into the etymology of the Mordvin word and, on phonological grounds, there is no reason to dispute the Indo-Iranian or Iranian origin of the Mordvin word.

The Permic words can be interpreted in a number of different ways. The cognate of Udmurt *suzer*, Komi *sozor* 'Fadenbruch (im Gewebe)', J *sózər* is not a kinship term at present, but given that the phonological match between Komi and Udmurt is unproblematic and that the semantic difference can be explained with parallels, cf. Russian *сестра* 'sister; Fadenbruch' (UEW 753), there seems to be no reason to doubt the cognate relationship. Furthermore, Komi-Zyrian *soć* 'sister', Upper Sysola *soć*, J *sočo* 'elder sister' might reflect some obscured hypocorism on *sozor*, indicating that at some point in time the Komi word probably also meant 'sister' (КЭСК 262).[2] Unlike the Mordvin words, the Permic words do not regularly reflect PU *a-a,* however. In inherited vocabulary, the vowel correspondence Udm *u* ~ Komi *o*, J *ó* in fact regularly reflects PU *e-ä*, cf. PU *elä-* 'to live' > Udm *ulị-*, Komi J *ol-*, PU *pesä* 'nest' > Udm *puz* 'egg; testicle', Komi J *poz* 'nest' (Metsäranta 2020 : 327—328 for further examples). One possible explanation could be to assume an earlier chronology, meaning that the Permic word would have been borrowed already from Pre-Proto-Indo-Iranian before PIE *e* changed into PII *a* (КЭСК 260), reflecting a form more closely resembling PIE *swésōr-* 'sister'. Based on phonology alone, even borrowing from Baltic, cf. Proto-Baltic *ses(-e)r* (Derksen 2015 : 395), is not outside the realm of possibility, albeit this is unlikely, because there is no clear indication that Pre-Proto-Permic was ever in direct contact with Baltic (which is not to say one could not find singular Baltic loan etymologies in the research literature). Furthermore, the few Baltic and Germanic loanwords that are thought to have reached the Permic languages via Finnic do not hold up to closer scrutiny (Metsäranta 2020 : 226—246).

---

[2] It is certainly possible that a hypocorism could in time oust the kinterm it is based on, as the hypocoristic forms are often much more frequent in common parlance. For example in Finnish, the hypocoristic *sisko* 'sister' is more commonly used in speech, while *sisar* has a more literary air.

An Indo-Iranian or Iranian source for Udm *suzer* and Komi *sozor* still seems like a possibility, if we take a look at the Indo-Iranian loanwords in the Permic languages as a whole. There are dozens of Indo-Iranian or Iranian loanwords of varying ages in the Permic languages. They seem to be divisible into three distinct layers based on the substitution of Indo-Iranian/Iranian *a*. There is a layer of loans where *a* is reflected as PP *u*, cf. PI *acwa* 'horse' → PP *už* 'stallion' > Udm Komi *už*, PI *gada-* 'thief' → PP *gu-* 'to steal' > Komi *gu-*, PII *Hrawpāćá-* / PI *rawpāca-* 'fox' → PP *ruć* 'fox' > Udm *žiči*, Komi *ruć*. These words have been, generally speaking, subject to all the same sound changes affecting inherited PU vocabulary in Permic, most notably the regular change PU *a* > PP *u*, so it seems reasonable to assume they have been borrowed into Pre-Proto-Permic, a stage that was still phonologically, as far as we can tell, mostly identical with Proto-Uralic: PI *acwa* → Pre-Proto-Permic *ača* > PP *už*, PI *gada-* → Pre-Proto-Permic *gaδa-* > PP *gu-*, PII *Hrawpāćá-* / PI *rawpāca-* → Pre-Proto-Permic *rapać* > PP *ruć*.

Alongside this substitution pattern, we find another one where Indo-Iranian/Iranian *a* (or *ā*) is reflected as PP *o* or PP *ɔ* (this distinction probably has its roots in the vowel quality of the second syllable before second-syllable vowels were eventually lost in Permic, but this can be ignored for now). In inherited PU vocabulary, these are the regular reflexes of PU *e-ə* and PU *e-ä* respectively. Traditionally a handful of loans have been thought to reflect this pattern, cf. Udm *dumi-* 'to bind', Komi *dom*, J *dom* 'leash, tether', Komi dial. (Ud.) *dom-* 'to bind' < PP *dɔmi-* ← PI *dáHman-* 'Band, Seil, Fessel' (Holopainen 2019 : 80), Komi *zon(m-)* 'boy' < PP *zɔn(m-)* ← Iranian *zana-* > Oss *zænæg* 'child' (Holopainen 2019 : 384), although the stratification has not been all that consistent, the former being classified as an Indo-Iranian loan and the latter as a later Iranian loan despite reflecting the same Proto-Permic vowel.

First of all, it seems that the number of Iranian loanwords that fit this pattern is notably higher than previously thought. Recently, a new Iranian loan etymology has been proposed for around ten Permic words, cf. Udm *ud* 'sprout, shoot (of cereal)', Komi *od* 'spring verdancy; shoot, sprout' < PP *ɔ/od* ← Iran *ādu-* 'corn, cereal, grain' (Metsäranta 2020 : 175—177), Udm *gur* 'oven', Komi *gor(j-)* 'sauna stove', J *gor* 'oven' < PP *gɔr* ← Iran *gārįa* 'stone heap' > Waxi *γ̃or ~ γor* (Metsäranta 2020 : 181—183), Udm *gureź* 'mountain, hill', Komi *goruv* 'base of a mountain' (-*uv* 'lower part') < PP *gɔr-* 'mountain' ← Iran *gari-* 'mountain' (op. cit. 183—184) where the same substitution PP *ɔ/*o* ← Iranian *a/*ā* occurs. Phonologically speaking, the solution for the substitution could be that the vowel developments eventually leading up to Proto-Permic had already started at this point. PU *a* had perhaps been labialised into *o*, while PU *e*, which would eventually develop into a back vowel in Permic — the vowel reflected as Udm *u*, Komi *o* — had started backing and was perhaps an "*a*-like" vowel or at least similar enough to fit as the substitution for Iranian *a/*ā* (as Iranian only had only three vowel qualities (Cantera 2017 : 482—483), it should be obvious that "*a*-like" has to be understood rather loosely here). Chronologically speaking, this layer has been labelled "Middle-Proto-Permic" and is thought to represent a loanword layer younger than where PII/PI *a* is substituted with Pre-Proto-Permic *a* > PP *u* (Metsäranta 2020 : 160—202).

The third pattern is one where Iranian *a* is reflected as PP *a*, cf. Udm Komi *das* '10' < PP *das* ← ?Alanic *das* '10' > Oss *dæs*, Udm Komi *zarńi* 'gold'

< PP *zarńi ← Iran *zaranįa 'gold' > Oss *zærin*. This is quite clearly the most recent of the Iranian loanword layers, as there are no further vowel changes affecting the words belonging to this layer. To summarise the stratification of (Indo-)Iranian loans based on the substitution of Indo-Iranian/Iranian *a:

❖ the oldest layer: PU *a > Pre-Proto-Permic *a (← PII/PI *a) > MPP *o > PP *u

❖ the middle layer: PU *e > Pre-Proto-Permic *e > MPP *a (← PII/PI *a) > PP *o/ɔ

❖ the youngest layer: PP *a ← Iranian *a

One could also perhaps toy with the idea that the middle layer is actually the oldest and was borrowed already into Proto-Uralic or Pre-Proto-Permic from Pre-Proto-Indo-Iranian before the change *e > *a. This conclusion seems less likely on distributional factors alone, i.e. the words are confined to Permic. Pre-Proto-Permic being a geographically and linguistically distinct entity at the time of Pre-Proto-Indo-Iranian is not a likely possibility given the antiquity of the latter. Following the stratification and reasoning presented above, Udm *suzer*, Komi *sozor* could be analysed as a Middle-Proto-Permic loan: PP *sɔzVr < MPP *sasar. It has been said that PII *s- was still retained in Proto-Iranian (Mayrhofer 1989 : 7). Rather than Proto-Indo-Iranian *swasar-, the MPP *sasar could then represent a later Iranian borrowing like the other Iranian borrowings in Permic with the same vowel substitution Iranian *a → MPP *a > PP *ɔ/o. The chronology fits Mordvin phonologically as well. At the very least, as the vowel correspondence between Mordvin and Permic is not regular, they should be treated as separate loanwords rather than shared Proto-Uralic loanwords.

It has already been noted that the vowels in MariM *šüžar*, MariH *šə̂žar* do not regularly reflect PU *a and it looks like the Mari word might be an independent borrowing from an (Indo-)Iranian source (Holopainen 2019 : 223). The possibilities have not been explored beyond this. The labial front vowel in Meadow Mari *šüžar* is quite clearly a secondary development. Here one should consider the Mari dialectal forms in their totality: East (Ob Oka Ok Okr Mm) *šužar*, East (Malmyž) *suzar*, Central (Ms, Mmu) *šüžar*, Central (Mm₄) *šüžar*, *šužar*, Volga *šŭžar*, Upša *šǚžar*, Northwest *šŏžar*, West *šəžar* (TschWb 755; MNySz 2579). The majority of the Mari dialectal forms regularly point to PM *sŭzar*, as first-syllable East *u*, Volga *ŭ*, Northwest *ŏ* and West *ə̂* are all regular reflexes of PM *ŭ* (Luobbal Sámmol Sámmol Ánte (Aikio) 2014a : 126). There are two regular sources of PM *ŭ* that are applicable here (the third one involves an adjacent labial element). In inherited vocabulary the labial reduced back vowel in Proto-Mari reflects either Pre-Mari *u*, cf. PU *sula- 'to thaw' > PM *sŭle- or Pre-Mari disharmonic *i-a, cf. PU *wišara 'green' > PM *ŭžar (Luobbal Sámmol Sámmol Ánte (Aikio) 2014a).

As the Mari word cannot reflect an earlier *a at any reconstruction level, considering the Mari word an independent (Indo-)Iranian loan does not seem to provide a convincing explanation for its phonology. One could perhaps argue that Indo-Iranian or Iranian *sw- was rendered as Pre-Mari *su-, which would seemingly explain the unexpected vowel correspondence. This substitution is without parallels, however, and as such remains an *ad hoc* solution. Unlike in Permic, there also are very few if any independent Indo-Iranian/Iranian loans in Mari, although such loans are not completely absent, cf. Iran *karta- 'sword,

sabre' → Pre-Mari *kärtä > PM *kerδə > M kerδe 'sword, sabre', Iran *pitu- 'meat' → Pre-Mari *pitV > PM *pĭćə > M pə̂ze (Metsäranta 2020 : 203—205).

Given the phonology of the Mari word, one theoretical possibility that should probably be mentioned is that the word could be a loanword from Permic, more precisely from Udmurt. The reasoning for such a late timing is that presumably borrowing from Proto-Udmurt *sozer would have yielded PM **sozar, cf. PM *oδə 'Udmurt' ← PUdm *od-mort > Udm ud-murt (Bereczki 1992 : 122). The raising of PUdm *o to u is recent enough that it has even affected some Russian loans in Udmurt kuso 'scythe' ← Ru коса (Лыткин 1964 : 19). It remains speculation whether or not Udm u could have been substituted with PM *ŭ, as there are no parallels among the loanwords that have been traditionally thought to be (Proto-)Udmurt loans in Mari. In this scenario, one could not speak about the loan as Proto-Mari *per se*, but it is interesting how uniform even some Russian loans in Mari are in terms of vowel correspondences, cf. East unə̂ka, Upša ŭnŭka, Northwest onoka (!), West ə̂nə̂ka 'grandchild' (TschWb 873) (← Ru внук). This is to illustrate that even recent loans can exhibit vowel patterns reminiscent of Proto-Mari despite post-dating them. This could, among other things, be due to interdialectal borrowing.

Interestingly the Finnic words mentioned earlier, namely Fi sisar, Veps sizar, EstS sysaŕ, could in theory reflect the same proto-form *sisar as Proto-Mari *sŭzar. Shared Baltic loans dating back to the common ancestor of Finnic and Mari is not an appealing possibility. Parallel borrowing from a Baltic source can probably also be ruled out since Mari has very few independent Baltic loanwords and most of the fewer than 10 proposed Baltic loans have a superior explanation (Grünthal 2012 : 310).

MdE sazor, MdM sazər, MariM šŭžar, MariH šə̂žar and Udm suzer all refer to 'younger sister'. There is no age distinction on the Indo-European side and the word in Baltic and (Indo-)Iranian refers to 'sister' in general. It has been suggested that initially the word on the Uralic side also meant 'sister' more broadly, and later the meaning 'younger sister' developed due to Turkic influence (Holopainen 2019 : 224). One open question is how the same semantic shift took place in all three branches presumably independent of each other. The point of contention is not so much the semantic shift itself that mirrors the system present in Turkic, but whether or not the change happening in Mordvin, Mari and Udmurt separately is credible and how exactly this would have come about. A possible explanation is that in all three branches, the word for 'elder sister' was at some point borrowed from (Old) Chuvash (later also from Tatar in case of Udmurt, apaj 'elder sister' ← Tatar apaj 'honorific vocative for elder sister' (Csúcs 1990 : 105—106)), cf. MdM aka 'elder sister' ← Chuvash (Mészáros 2001 : 172), MariM aka 'elder sister', MariH äkä ← Chuvash aɢaj 'elder sister' (Räsänen 1920 : 112), Udm aka 'elder sister; father's sister, uncle's wife' ← Chuvash aɢi̯, aɢi, akka 'elder sister' (Wichmann 1903 : 38). Perhaps this in turn triggered the semantic shift from 'sister' to 'younger sister' of the existing word in all three languages. This convergence can hardly be coincidental although it could have its own circumstancial background independent of the (Indo-)Iranian loan etymology.

In sum, the Mordvin, Mari and Permic words for 'sister' should be seen as three separate instances of borrowing; they are undoubtedly ultimately of Indo-European origin, but in terms of phonology the words cannot be regarded as shared loanwords into Proto-Uralic but rather separate loan-

words into already phonologically dispersed dialects. This is especially true for Permic.

2.2. 'Daughter'

The words Fi *tytär* 'daughter', Est *tütar*, Veps *tütär*, Liv *tidār* ~ SaaS *dektier*, *daktere* 'married daughter' ~ MdE *ťejťeŕ* 'daughter', M *śťiŕ* have often been treated as cognates (Kalima 1936 : 173—174; SKES 1463; SSA 3 : 349). Such a cognate relationship at least implies that the borrowing from Baltic, cf. Lith *duktẽ* (gen. *dukteȓs*), took place when the ancestor of Finnic, Saami and Mordvin still formed a single proto-language. This underlying assumption of a cognate relationship also seems to guide the path which sound development is thought to have taken (see YSuS (https://sanat.csc.fi/wiki/Etymologiadata:YSuS)), which derives the modern Finnic words from Late Proto-Finnic **tüťtär* and ultimately from Early Proto-Finnic **tüktärə*). Recent research also describes the Finnic words as Baltic loans (Junttila 2015 : 96), but does not comment further on a cognate relationship between the Finnic, Saami and Mordvin words.

Some differing opinions on the relationships between these words have been voiced especially regarding the Saami member, for which at least four competing views have been put forth. According to one view, SaaS *dektier*, *daktere* reached Saami via Finnic, i.e. the Saami word was borrowed from Finnish (Wiklund 1896 : 42). According to another view, the Saami word was borrowed directly from Scandinavian languages, cf. Swedish *dotter* (Qvigstad 1893 : 125). According to the most recent proposal, the word is indeed of Scandinavian origin but borrowed from their predecessor, i.e. PScand **duhter-* (Luobbal Sámmol Sámmol Ánte (Aikio) 2020 : 17). Besides borrowing from Finnic or some form of Scandinavian, it has been proposed that the Saami word was borrowed from Baltic into Pre-Proto-Saami independently (Sammallahti 1984 : 139).

Let us start the unpacking with Finnic. The phonological problem with deriving the Finnic words from an earlier EPF **kt* cluster is that its expected outcome in (most of) Finnic would be **ht* rather than the geminate **ťt* we actually find in LPF. This unexpectedness naturally did not escape the attention of earlier researchers, and different solutions have been sought to explain it. One solution has been in essence to ignore the unexpected outcome by referring to a few Finnic words that, based on their proposed cognates elsewhere in Uralic, also seem to reflect an irregular EPF **kt* > LPF **ťt* change, e.g. Fi *mätäs* 'tussock',[3] *pettää* 'to churn butter',[4] and assuming that *tytär* belongs

---

[3] The cognate set, Fi *mätäs* 'tussock', Karelian *mätäs*, Veps *mättaz*, *mätäz*, Votic *mätäz* 'hill, hillock; tussock, peat', Est *mätas* 'tussock, peat', Liv *mätāl* 'tussock' < LFP **mäťtäs* ~ SaaN *miekta* 'tussock' < PS **miektẹ* ~ NenT *ńet*, Selkup *mäkte*, *mekte*, *mäktə*, Kamassian *mekte*, *bäkte* < PSam **mektə* is thought to reflect PU **mäktə* 'Rasenhügel, Hügel' (UEW 266). Although PSam **mektə* is remarkably similar to both the Saami words and the reconstructed PU protoform, it appears, however, to be irregular as well because one would expect the **kt* cluster to regularly simplify into PSam **t*, cf. PU **ẹkta-* 'to hang' > PSam **ịta-* > NenT *ŋida-*, Ngan *ŋiti-*, SlkTaz *įtị-*. The apparent similarity could thus be coincidental. Also the derivational relationship between LFP **mäťtäs* and PU **mäktə* would be peculiar and atypical of inherited vocabulary.

[4] Fi dial. *pettää* 'to churn butter', Karelian *pettäjäine* 'freshly churnt, unsalted butter', Votic *pettäjäiset* 'churning of butter', Est *pett*, *petipiim* 'churned milk', Liv *pietk*, *pietkəm-semďā* 'to mix butter with a stirring utensil' ~ MdE *pivťems*, *pifťems*, MdM *pišťəms*, *pifťəms* 'to churn', MdE *pivťi*, MdM *pifťä* 'piston of churn'.

to this group of irregular words (Uotila 1985 : 314). It should be noted that a supposedly irregular change having parallels is an oxymoron or at least a confused methodology. It might very well be that the "parallels" turn out to have a different explanation altogether.

Another solution, proposed by Lauri Posti (1953—1954 : 44—46), was that PF *-tt-* was substituted for Baltic *-kt-* in words where the main stress came after the cluster. Part of the explanation is that the regular Finnic change *kt* > *tt* after unstressed syllables had already taken place and as a result *kt* no longer occurred after an unstressed syllable. This argument is not contradicted by known Baltic loans, cf. Fi *juhta* 'beast of burden' ← Baltic, cf. Lith *jùngtas* 'connected, united', Fi *suhta* 'ratio, proportion' ← Baltic, cf. Lith *sùktas* 'twisted', which reflect substitution of Baltic *kt* as PF *ht* (< *kt*) and have initial stress (Junttila 2015 : 176). No clear parallels for the substitution of Baltic *kt* as PF *tt* are presented by Posti or later research, however, and thus it remains speculation even when there is no evidence to directly contradict it.

A third solution, namely borrowing via South Estonian, has not been explicitly proposed (Junttila 2015 : 176). The reasoning behind the idea that the word spread originally from South Estonian to other Finnic varieties seems to lie in the fact that the first dialectal split within Finnic is thought to have taken place between Inland Finnic (predecessor of the South Estonian varieties Võro, Seto, Leivu, etc.) and Coastal Finnic (all other Finnic varieties), with PF *kt* developing into *tt* in Inland Finnic and to *ht* in Coastal Finnic (Kallio 2014 : 156). A scenario where Baltic *dukter-* was borrowed into Inland Finnic prior to the *kt* > *tt* change and later from Inland Finnic to Coastal Finnic, Baltic *dukter-* → PF *tüktär* > Inland *tüttär* → Coastal Finnic *tüttär*, would seem to explain the unexpected geminate in the descendants of Coastal Finnic. This idea would still need to be explored further, as a singular example is not sufficiently convincing in itself. All of the proposed solutions for LPF *tüťtär*, i.e. 1) irregular EPF *kt* > LPF *ťt*, 2) substitution of Baltic *kt* with *tt* after unstressed syllables and 3) borrowing via Inland Finnic, rely on scanty evidence. It is not readily obvious which of the proposed solutions is preferable, but solutions 2 and 3 have basically the same implications for a cognate relationship between the Finnic, Saami and Mordvin words. As Saami and Mordvin cannot reflect an earlier geminate *tt*, the Finnic words should probably be regarded as separate loanwords at the very least.

In the following, the different explanations given for the origin of the Saami word are discussed in more detail. Along with South Saami *dektier*, *daktere* 'married daughter', the word is attested in Ume Saami and Pite Saami and reflects PS *tektēr* (Lehtiranta 1989 : 130—131). Sammallahti (1984 : 139—140) proposes that the Saami word was borrowed directly from Baltic without Finnic mediation. This is in theory possible, but then again there are only a handful of words of Baltic origin in Saami that do not have a cognate in Finnic (e.g. SaaS *saertie* 'heart (as food)' < Pre-Proto-Saami *šärtä* < Baltic *šerdā*). According to Aikio (Luobbal Sámmol Sámmol Ánte (Aikio) 2012 : 74) out of the 32 Baltic loans in Saami, only 8 lack a cognate in Finnish. Aikio also points out that such a low fraction does not serve as proof of independent contacts between Pre-Proto-Saami and Baltic, because it might easily be the case that the mediating word in Finnish was simply lost at a later stage, as it certainly cannot be assumed that Finnish has retained 100% of its Pre-Proto-Finnic vocabulary. Given that the majority of Baltic loans in Saami are shared with Finnish, the most

probable explanation would seem to be that Pre-Proto-Saami was never in direct contact with Baltic and the Baltic loans in Saami were secondarily diffused through Pre-Proto-Finnic, which in turn was in direct contact with Baltic (Luobbal Sámmol Sámmol Ánte (Aikio) 2012 : 73—75).

It has also been argued that South Saami *dektier*, *daktere* was borrowed via Finnish from Baltic (Wiklund 1896 : 42), which given what was said about Baltic loans in general in the previous paragraph, is a reasonable assumption. General observations of what is typical of a certain loanword layer are not sufficient evidence alone, however, especially since borrowing directly from Finnish or even LPF *tüťtär* provides no explanation for the *kt* cluster in Saami (Sammallahti 1984 : 139) or the vowel correspondence. Even if the process of etymological nativisation could somehow be involved, the vowel *y* (< *ü*) has almost never been etymologically nativised (Aikio 2007 : 30—31), adding to the unlikelihood. Assuming an even more ancient borrowing from EPF *tüktärə* into Pre-Proto-Saami ignores the fact that there is no way of of ascertaining, independently of the loan etymology itself, that the Finnic actually reflects an earlier *kt* cluster. Claiming that the Saami word proves that Early Proto-Finnic had a *kt* cluster and that the Saami word is therefore a loanword from EPF, is a circular argument.

A Scandinavian origin for the South Saami word, cf. Swedish, Norwegian *dotter* 'daughter' was first proposed by Qvigstad (1893 : 125). The Scandinavian origin was also already supported by Posti (1953—1954 : 45), who argued that that the Saami word was borrowed before the assimilation of *ht* to *tt* on the Scandinavian side and *kt* was substituted for Scandinavian *ht*. As a parallel for the substitution one can cite SaaS (obs.) *slikt* (< PS *slikte̯*) 'smooth', SaaN *livttis* 'smooth and even' (< PS *liktēs*) ← PScand *slihtaz* > ON *sléttr* 'flat, smooth, even' (Posti 1953—1954 : 45; Luobbal Sámmol Sámmol Ánte (Aikio) 2012 : 110). Based on information communicated by Erkki Itkonen, Posti states that there is nothing in the vocalism of the Saami word to speak against a Scandinavian origin. Recently, another iteration of this Scandinavian loan etymology has been stated as SaaS *dektier, daktere* 'married daughter' < PS *toktēr: toktāre̯* ← PScand *duhter-* (> ON *dóttir* 'daughter') (Luobbal Sámmol Sámmol Ánte (Aikio) 2020 : 17). Aikio also rejects a cognate relationship between the Saami and Finnish words based on the irregular consonant correspondence.

It is curious that the case for a Scandinavian origin seems to have been made without referencing the Ume and Pite Saami forms at all. It is true that South Saami *dektier, daktere* could in theory reflect either PS *o* (< Pre-Proto-Saami *u*) or *e̯* (< Pre-Proto-Saami *i, *e, *ü*), as these Proto-Saami vowels fell together before long consonants, i.e. geminates and consonant clusters, in South Saami (Korhonen 1981 : 81), cf. PS *ke̯lmē-* 'to freeze' > SaaS *gelmedh*, PS *lontē* 'bird' > SaaS *ledtie*. However, Ume Saami *dak̍tèr* and Pite Saami *taktier* confirm that the Proto-Saami first-syllable vowel was *e̯*, as this regularly yields Ume and Pite Saami *a*, cf. PS *ke̯lmē-* 'to freeze' > U *galbmeet*, Pi *kal̍pmiet*, whereas PS *o* regularly yields Ume and Pite Saami *å*, cf. PS *lontē* 'bird' > U *låd̍dee*, Pi *låttie*. Despite the seeming mismatch in vocalism, PS *te̯ktēr* ~ PScand *duhter-*, considering that South, Ume and probably also Pite Saami form their own separate proto-dialectal entity, the Southwest dialect, that exhibits its own particular pattern of Scandinavian loanword nativisation (Luobbal Sámmol Sámmol Ánte (Aikio) 2012 : 77), it is still perhaps fruitful to try and explore the Scandinavian origin, as the alternative explanations are not without their problems, either.

A "labial dissimilation" of PS *o to PS *$ę$ has occurred in a handful of words in Saami (Luobbal Sámmol Sámmol Ánte (Aikio) 2015 : 11). In most cases the original vowel can be determined to have been PS *o through etymology, e.g. PU *$lupsa$ 'dew' > PS *$lopsē$ > PS *$lępsē$ (with dissimilation occurring in all Saami languages, but cognates elsewhere clearly point to PU *u). The common denominator for the dissimilation is an adjacent labial consonant; the only exception seems to have been PS *$koco$- 'to hang', which shows dissimilation in South, Ume and Lule Saami. The dissimilation was not regular, as no conditioning factor can be established and there are a significant number of words with PS *o that lack the dissimilation despite being adjacent to a labial consonant. Also, the distribution of variants with *$ę$ is different for each lexical item, with the dissimilation sometimes affecting all Saami languages, like in PS *$lopsē$ 'dew', sometimes occurring in the Western half, cf. PS *$kocō$- 'to hang' (S U Lu) and sometimes in the Eastern, cf. PS *$monē$ 'egg' (N (both), In Sk K T). Aikio (Luobbal Sámmol Sámmol Ánte (Aikio) 2015) offers as an explanation the near-merger of Proto-Saami *o and *$ę$ when adjacent to labial consonants, meaning that PS *$ę$ could have developed a rounded allophone [$ȯ$] that drifted so close to *o that some of the instances would have been transferred from one phoneme to another. One does wonder whether this near-merger had a hand in the word *$tękt\bar{e}r$ as well, despite the word obviously lacking a labial consonant. The problem is that the whole idea of a near-merger remains speculation and there is no way of verifying whether or not the dissimilation occurred, as there exist only reflexes of dissimilated *$tękt\bar{e}r$.

We should still perhaps consider the merits of the traditional view according to which the Finnic, Saami and Mordvin words were borrowed into the common proto-language of these languages from an Indo-European source that is most often considered to be Baltic, cf. Lith *duktė̃* (gen. *dukter̃s*), OPruss *duckti* 'daughter' (Kalima 1936 : 173—174; SKES 1463; SSA 3 : 349).[5] The most convincing argument for a shared borrowing from a phonological standpoint is that all three branches can reflect an earlier *$ü$ as a substitution for IE *u. Saami and Mordvin, or at least Erzya *t'ejt'eŕ* (Moksha *śt'iŕ* is rather undiagnostic), can both be regularly derived from Pre-Proto-Saami/Pre-Proto-Mordvin *$tüktärə$. The Finnic words regularly reflect a protoform with a geminate *tt (or *pt), i.e. Pre-Proto-Finnic *$tüttärə$ (or *$tüptärə$) > MPF *$tüttäri$ > LPF *$tüťťär$. The reason behind substituting IE *u with a front-vowel *$ü$ is not altogether clear, but the fact that it has taken place seem-

[5] Kalima (1936 : 173—174) expressed some doubt concerning the Baltic origin by saying that "One can ask whether *tytär* is better explained as a Baltic or an Indo-Iranian loan, cf. Old Indian *duhitā*, Avestan *duɣδar*, New Persian *duχtär*." (translated from Finnish). It is indeed unclear, why the Uralic word could not in theory have been borrowed from a form similar to PII *$d^hugHtar$-. The front-vocalic substitution of IE *u with U *$ü$ is not common among Indo-Iranian loans, there is actually just one half-way decent example of it, i.e. PU *$mükkä$ 'dumb, mute' (Holopainen 2019 : 150), where the front-vowel *$ü$ at least could have some antiquity as also the Saami cognates can be derived from an earlier *$ü$ instead of the front vowel being brought on by secondary fronting in Finnic found in word-pairs such as Fi *tuhma ~ tyhmä*. In Baltic there are around a dozen suggested loan etymologies that require substitution of Baltic *u with Finnic *$ü$ (Junttila 2015). Among the more certain loanwords reconstructable for PF, i.e. Junttila's categories of "Continuously accepted etymologies" & "Via discussion accepted etymologies", there is only one potential parallel for *$tüttär$, namely PF *$tühjä$ 'empty' ← Baltic *tuštjas*, cf. Lith *tuščias* 'empty; poor; unnecessary'). The vowel substitution thus can hardly settle the matter.

ingly uniformly in all three branches is not something that can be easily dismissed.

There are perhaps a few arguments to be made against a shared borrowing from Baltic. First of all, there is at least one group of words of Baltic origin that has a similar distribution, but due to the varying phonological shapes of these words, they were borrowed separately at least into Pre-Proto-Saami and Pre-Proto-Finnic. The words in question are Pre-Proto-Finnic *leppä 'alder' (> Finnish leppä, Est lepp, Liv liepā) and Pre-Proto-Saami *lejpä 'alder' (> SaaN leaibi), which are usually thought to have been borrowed from Baltic, cf. Latvian liepa, Lith líepa, OPruss leipa 'linden' (SSA 2 : 64—65). Mordvin, Erzya l'epe, Moksha l'epä 'alder' could reflect even a third separate Baltic borrowing, Pre-Proto-Mordvin *lippä, although it has been suggested that the Mordvin words were borrowed from Pre-Proto-Finnic (Luobbal Sámmol Sámmol Ánte (Aikio) 2012 : 108). Single words obviously cannot be used to reach any far-reaching conclusions, but perhaps the variance in 'alder' does indicate that by the time of Baltic contacts, Pre-Proto-Saami, Pre-Proto-Finnic and Pre-Proto-Mordvin already formed at least partly separate linguistic entities. Separate Baltic borrowings is a possibility, since although the Mordvinic languages have significantly fewer Baltic loanwords than Finnic, it seems that Mordvinic speakers did have independent prehistoric contacts with Baltic speakers as attested by more than thirty plausible etymologies (Grünthal 2012 : 297) that are also largely absent from Finnic and Saami, indicating that they have been borrowed separately into Mordvinic.

One onomasiological point that could argue against a shared borrowing is that in most Saami languages the word used for 'daughter' is an old inherited word, PU *näjδə 'girl, daughter' > PS *niejtę (Lehtiranta 1989 : 82—83), which also exists in South Saami as nïejte 'girl, unmarried daughter' (ÅDB 204). The two words for 'daughter' in South Saami occupy two different semantic slots, one, the inherited word, referring to 'unmarried daughter' and the other, the loan, to 'married daughter'. As the word for 'daughter' in the Saami languages is an old inherited Uralic word, an argument can be made that onomasiologically speaking, we are dealing with a retention rather than an innovation.[6] A scenario where the Baltic loan first replaced the inherited word for 'daughter' in the common ancestor of Saami, Finnic and Mordvin, only to later be replaced again by a native Uralic word in Proto-Saami, is unlikely.

None of the explanations we have looked at is without its problems. Perhaps the most parsimonious and least problematic solution is to treat the words for 'daughter' as having been borrowed separately into the predecessors of Finnic,

---

[6] This argument is admittedly dependent on the fact that PS *niejtę is actually cognate with LPF *näit-oi 'girl, young woman' > Finnish neito, Karelian neito(i), Votic neito, Est neiu, EstS näio (YSuS (https://sanat.csc.fi/wiki/Etymologiadata:YSuS)) and Udm nɨl, Komi nɨv 'girl, daughter'. The Permic forms can regularly reflect PU *näjδə (although based on the Permic forms alone several different Proto-Uralic forms are theoretically possible, including *nälə, *nälkə, *nülV, *nülkV, *nüjδV, *nulV, *nulkV, *nujδV etc.). One phonological reservation and the reason why the Saami words have also been thought to be Finnic loanwords is that while the reflexes of PU *jδ and *jt have largely coincided in Saami languages, the reflexes are thought to have been kept apart in Inari Saami, yet one finds nieiˈdâ, instead of **nieiˈđâ (Korhonen 1981 : 183). Although this could be taken as indicative of a Finnic loan, borrowing from North Saami nieida could easily well explain the unexpected reflex. It is also perhaps noteworthy that Finnic *-it- has also been substituted with PS *-jδ-, cf. SaaS daajr'edh, N dáidit, In täiđiđ < PS *tājδē- 'to know (how to); to probably do' ← LPF *taita- > Finnish taitaa (Lehtiranta 1989 : 130—131).

Saami and Mordvin from different Indo-European source languages. The Finnic and Mordvin words are likely to have been borrowed from Baltic. As there is no way of regularly reconciling the consonantism between Finnic and Mordvin, and as both branches have had independent contacts with Baltic, the words in all likelihood constitute two independent Baltic loanwords. A separate Baltic origin of the Finnic word is also supported by the fact that there are several other kinship terms of Baltic origin that have been borrowed into Middle Proto-Finnic, e.g. LFP *morcijan* 'bride, young wife' < MPF *mortijami* ← cf. Lith *martì*. LPF *tüťtär* could have been borrowed in the Middle Proto-Finnic period (MPF *tüttäri* ← Baltic) as well, although we can hardly exclude the possibility of an earlier borrowing given that the potential Pre-Proto-Finnic form, *tüttärə*, is almost identical. The Mordvin word could have been borrowed into Pre-Proto-Mordvin, as the Erzya word *ťejťeŕ* 'daughter' can be regularly derived from a Pre-Proto-Mordvin form *tüktärə*. A separate Scandinavian origin for the Saami word is supported by its Southwestern distribution, which coincides with separate nativisation patterns of Scandinavian loans, although admittedly the vowel correspondences remain problematic. A direct Baltic origin of the Saami words is speculative, as it seems that Pre-Proto-Saami was never in close contact with Baltic. Finnic mediation of the Baltic word to Saami remains speculation as well, and there is no clear indication of Finnic origin other than the general observation that Baltic loans have often been diffused to Pre-Proto-Saami through Pre-Proto-Finnic.

2.3. 'Brother'

The word is present throughout Finnic (Fi *veli* 'brother', Veps *vel'l'*, Est dial. *veli* 'brother; bride's brother in a wedding', Liv *ve'ļ*) and reflects LPF *velji* < EPF *weljə*. Likewise in Saami the distribution is pan-Saamic and the Saami words (SaaN *viellja* 'brother', S *vielle*, Sk *villj*) can be reconstructed as PS *vieljẹ* (Lehtiranta 1989 : 148—149). The Finnic and Saami words have traditionally been regarded as cognates whether or not they are treated as inherited from Proto-Uralic or Proto-Finno-Saamic or borrowed from Proto-Germanic. This is despite the fact that PS *-ie-* is irregular vis-à-vis *weljə*, which should regularly yield PS **velẹ*. It has been mentioned that based on this irregularity, a borrowing from Finnic is a possibility (Luobbal Sámmol Sámmol Ánte (Aikio) 2014b : 68). In the UEW (p. 567) the Hungarian instrumental suffix *-val/-vel* is also compared to the Finnic and Saami words, but this Gleichsetzung is considered uncertain and due to the uncertain semantics, it has been altogether rejected in later research, e.g. in SSA (3 : 424) only the Finnic and Saami words are considered cognates. SSA mentions that two loan etymologies have been suggested. Over the following two paragraphs, I shall give an outline of these two differing suggestions.

According to the first loan etymology, EPF *weljə* 'brother' has been borrowed from Proto-Indo-European or Early Proto-Germanic. In the first and in the later iterations of the loan etymology as well, the loan original is thought to be a form that is identical to or closely resembles PIE *sweliyo-*, which can be reconstructed based on reflexes in Ancient Greek ἀέλιοι mpl. 'brothers-in-law' and Germanic, cf. ON *svilar* mpl. 'Schwäger von Schwestern' (Koivulehto 1993 : 34; 1994 : 5). In the "Lexikon der älteren germanischen Lehnwörter in den ostseefinnischen Sprachen" (LägLos 386—387) reference is made to Proto-

Germanic *sweljan- as the loan original. LägLoS finds the loan suggestion both phonologically and semantically possible and holds the view that EPF *weljə 'brother' is either a Germanic or an older loanword.

According to the second view, EPF *weljə 'brother' (*velje) was borrowed from *veljē, an earlier form underlying Lithuanian velẽ (supposedly an older variant of vélẽ) 'soul, ghost (of a dead person)' (Liukkonen 1999 : 152—154). According to Liukkonen, the words can be explained by assuming the following semantic changes: 'geisterhafte Gestalt des Verstorbenen' > 'Verstorbener als Doppelgänger' > 'Doppelgänger' > 'Bruder'. The semantic change from 'Doppel-gänger' to 'Bruder' is claimed to be easily understood especially in the context of identical twins, but also in general (Liukkonen 1999 : 154). Although the Baltic word would seem like a good match for the EPF word phonologically, the semantics leave a lot to desire (Koivulehto 2001 : 58). Liukkonen simply assumes four semantic changes, none of which is all that obvious, and more-over he does not attempt to provide any parallels for even one of the assumed steps. It has been remarked that such explanations can be constructed to bridge almost any semantic gap (Aikio 2009 : 20) and etymologies based on such dreamt up semantic strings are thus of little value, their only real value being perhaps in the critique they have sparked. The burden of proof is also not on the critics to show that the semantic steps are impossible. Proving a negative this way would be impossible. The Baltic etymology can be safely rejected, but the Proto-Indo-European or Early Proto-Germanic loan etymology has gener-ally found support. A closer look at the actual arguments, especially at the semantics is, however, in order.

It happens annoyingly often in Uralic etymology that the assumed semantic connection receives little to no scrutiny and is in many cases, like in LägLoS, only described as "possible". Koivulehto (1994 : 5) starts his semantic argument by claiming that men whose wives are sisters, share a relationship that is some-how special, perhaps especially cordial and brotherly. Assuming that such a relationship could be close is perhaps not a huge leap of logic, but attaching to this relationship attributes like 'cordial' and especially 'brotherly' is what in legalese would be called leading the witness. This is fortunately not the main bulk of the semantic argument provided. For the actual parallel of the semantic change 'brother-in-law' > 'brother' Koivulehto provides Latvian znuōts 'son-in-law; brother-in-law' that he supposes is cognate with Ancient Greek γνωτός '(blood) relative, (esp.) brother; (as a feminine) sister' (Koivulehto 1994 : 5). It is also mentioned by Koivulehto that in Old High German gi-lang 'relative, brother-in-law' was used when talking about the apostle Peter's brother.

The Latvian word is indeed a cognate with the Greek word, but based on the newest available research, it hardly serves as a convincing parallel for the semantic change 'brother-in-law' > 'brother' in the way Koivulehto envisioned, and it is not obvious why one of the Latvian meanings should have been regarded as primary in the first place. In any case, Latvian znuōts 'son-in-law; brother-in-law' reflects a PIE o-grade *ĝnoh$_3$-tó- that is paralleled by Vedic Sanskrit jñāta- 'known, recognised' and Latin nōtus 'recognised' (NIL 155; Mi-lanova 2020 : 148). Ancient Greek γνωτός 'relative, brother' and γνωτή 'sister' are described as zero-grade variants of PIE *ĝṇh$_3$-tó- together with homophonic γνωτός 'wahrgenommen, verstanden, bekannt' (NIL 155), although many researchers hold the view that Greek, too, reflects the above-mentioned o-grade. This detail is of no consequence, however, and the general consensus is that

both Latvian *znuōts* and Greek γνωτός 'relative, brother' are derivatives of PIE
*$\hat{g}neh_3$- 'erkennen'. According to Fraenkel, the semantic connection between the
base verb meaning 'erkennen' and 'son-in-law' can be explained by the fact
that it was the son-in-law who acted as the link between the wife's parents'
house and the young couple and became described as the 'Bekannt par excel-
lence' (LEW 1301). It is debatable how convincing this explanation is, but consid-
ering the meaning of the base verb and the much more ancient forms in Vedic
Sanskrit and Latin both meaning 'known' or 'recognised', there is little doubt
that the Latvian meaning 'son-in-law' represents a more recent development.
Alluding to the fact that the apostle Peter's brother is referred to as *gi-lang* in
Old High German is not really proof of actual semantic change and can thus
be dismissed.

Koivulehto (1994 : 5) mentions that it is more common for 'brother' to change
into 'brother-in-law' and cites English *brother-in-law* and French *beau-frère* as
examples. As this is a change opposite to what is required to explain the connec-
tion between EPF *$welj\partial$ 'brother' and IE *$sweliyo$- 'brother-in-law', it is not
really pertinent here. Besides, despite containing the word for 'brother' there
is no semantic change happening with either *brother-in-law* or *beau-frère*, they
are simply compounds formed on the word 'brother'. Although it could still
be argued that the semantic change from 'brother-in-law' to 'brother' is possible,
the arguments initially given in its favour are not tenable. Also it is not clear
what kind of evidence would suffice in order to conclusively show that any
given semantic change is impossible. The possibility needs to be demonstrated
with positive evidence, otherwise stating that something is possible has no
actual content. It is not a methodologically feasible practice to try and exhaust
the impossibility of a claim, but there are a few observations that can be made
based on both typological evidence, namely colexification (François 2008) and
the etymological materials collected for Kinura. Both of these seem to speak
against the semantic validity of the proposed loan etymology.

Polysemy and similarly colexification, i.e. which meanings are found
together in the world's languages, can provide important clues for what
semantic changes are actually plausibly reconstructable. It has been said
that synchronic polysemy and historical change of meaning supply the
same data in many ways, and that no historical shift of meaning can take
place without an intervening stage of polysemy (Sweetser 1990 : 9). If
one claims that a semantic change from 'brother-in-law' to 'brother' is
possible, one should be able to find an intervening stage of polysemy
'brother-in-law; brother' existing somewhere. Among the datasets in the
Database of Cross-Linguistic Colexifications (https://clics.clld.org/), no
polysemy between 'brother' and 'brother-in-law' exists. The reason for
this might, however, partly be due to mismatch between the datasets, i.e.
the concept 'brother' consists of 760 entries, while the concept of 'brother-
in-law' is made up of 161 entries, all of which are from Southeast Asia
and Australia, but it is perhaps still telling.

We can now turn our attention to Kinura's material that spans all main
branches of Uralic and see whether or not corroborating evidence for a semantic
change from 'brother-in-law' to 'brother' can be found in the form of polysemy.
As the meanings 'brother' and 'brother-in-law' are spread across several cate-
gories in our data, all the relevant categories are included. The meaning 'brother'
includes 'brother' (B), 'elder brother' (eB) and 'younger brother' (yB), likewise

'brother-in-law' includes three different categories, namely 'sister's husband' (ZH), 'wife's brother' (WB) and 'husband's brother' (HB).

**Words for 'brother' (B), 'elder brother' (eB), 'younger brother' (yB),
'sister's husband' (ZH), 'wife's brother' (WB) and 'husband's brother' (HB)
in selected Uralic languages**

	B	eB	yB	ZH	WB	HB
EnF	*kasa*	*ina, aga, kasa*	*ńabi, ṕɛbi, sooku*	–	–	–
Est	*vend, veli*	–	–	*õemees*	*nääl*	*küdi*
Fi	*veli*	–	–	*lanko*	*lanko*	*lanko*
Hung	*fivér*	*bátya*	*öcs*	*sógor*	*sógor*	*sógor*
KhKaz	*jŭw-pŏχ*	*jaj*	*apəḷi, apśi*	*wɛŋ*	*ŭp, aj-ŭp*	*aki, ʌɛ-χọ*
Komi	*vok*	–	–	*źaťˊ*	*šurin*	*piver*
Liv	*ve'ḷ*	–	–	*sõzārmīez*	–	–
MariH	*əzäk-šoľak, šümbel*	*əzä*	*šoľa, šoľŝ*	*kŝrska*	*onŝska, pörŝž*	*onŝska, pörŝž*
MariM	*izak-šoľak*	*iza*	*šoľo*	*kurska, βeŋe*	*onŝska, pörŝž*	*onŝska, pörŝž*
MdE	*brat*	*ľeľa*	*jalaks, duga*	*čiče*	*šurin, aľa, baľźa*	*aľa*
MdM	*brat*	*paťä*, **aľńaka**	*dugan, päľńä*	*äzna, šičä*	*šuŕin, šuŕəń, aľǵä,* **aľńaka** *baźä, päľźä, paźä*	*aľǵä,* **aľńaka**
MsSo	–	*jaγ-piγ, kaŋk*	*āpśi, kaš*	*kil, wāps*	*up*	*aki, pānt*
NenT	*ńa*	*ńeəka, ńaəka, ńińeka*	*papa, ṕeb́a*	*ńińadi, jij*	*ŋinabə, nado*	*ŋinabə, jiŕi, nado*
Ngan	–	*əˀə, ńini*	*ŋaaŋku, ŋaďa*	*biŋi*	–	–
SaaN	*viellja*	–	–	*máhka*	*máhka*	*sivjjot*
SaaS	*vielle*	–	–	*maake, vïjve*	*maake, maaketje*	*vööhpedimmie, sïbjege*
SaaSk	*villj*	–	–	*maakk, vivv*	*vuõppbie'll, maakkâv*	–
SlkTaz	*tiɱńa, tipyńa*	–	–	*kərmä, qənyŋ*	*kərmä, mōqqy*	*iľca, mōqqy*
Udm	*brat, agaj-vịn*	**agaj, ńuńu**	*vịn*	*kịṛśi*	*varmajpi*	**agajzị, ńuńuzị,** *šịdnar*
Veps	*vel'ľ*	*veik*	-	*vävu*	–	*küdu*

I have highlighted those cases where there is lexical overlap between the categories for 'brother' and 'brother-in-law', and I will examine those cases more closely to see whether or not they do in fact constitute valid parallels for the change 'brother-in-law' > 'brother'. In Udmurt we find *agajzị* 'husband's

elder brother', which is formed on *agaj* 'elder brother' (dialectally also 'uncle; male cousin older than ego' (WotjWsch 3)). The Udmurt word *agaj* was, in turn, borrowed from Tatar *agaj* 'uncle (honorific term used of older men)' (Csúcs 1990 : 95). The *-zi̮* is not a derivative suffix but the 3rd person plural possessive suffix. The form *agajzi̮* literally meaning 'their elder brother' has emerged through lexicalisation. In Tatar, the word *agaj* was originally used as an honorific term to address male relatives that were older than the speaker. In Udmurt much of the same usage has been preserved, even if there are some dialectal differences in terms of which kinship categories the word covers exactly. All in all, *agajzi̮* and *agaj* do not offer an example of semantic change from 'brother-in-law' to 'brother', as the former has gained its affinal meaning 'brother-in-law' only through lexicalisation, and the order of the semantic change would be opposite to what we are looking for, even if this was not the case. Udmurt *ńuńu* 'elder brother' (dial. also *ńuń, ńuńa* etc.) (WotjWsch 179) and *ńuńuzi̮* 'husband's elder brother' follow a similar pattern to *agaj* and *agajzi̮*.

According to Mészáros (2001 : 171, 175, 176) the Moksha Mordvin word *aľńaka* is used for 'elder brother' ('a bátya valakinek'), 'husband's younger brother' ('a férj öccse'), 'wife's younger brother' ('a feleség öccse') and 'husband's younger sister's husband' ('a férj húgának a férje'). The derivational analysis offered is somewhat varyingly formulated in different places of Mészáros' article; *aľńaka* 'elder brother; husband's younger brother; husband's younger sister's husband' is analysed as consisting of two diminutive suffixes, *-ńa* and *-ka*, the base word being *aľä* that is glossed as 'uncle' ('bácsi'), although the semantics of the word are considerably more complex and the word is used for 'father' as well (Гришунина 2000 : 55; Mészáros 2001 : 169). *Aľńaka* 'wife's younger brother' is analysed as a diminutive derivation of *aľnä* 'brother-in-law' ('sógor'), which is elsewhere in Mészáros' table also analysed as being ultimately a derivation of *aľä* 'uncle'. Whatever the formulation, the result is the same: *aľńaka* is derived from *aľä* 'uncle' with two diminutive suffixes. This analysis is unproblematic, as Mordvin is known to employ a lot of diminutive suffixes in derivation (Bartens 1999 : 106). The distribution of *aľńaka* is scarce, as it has been collected from one dialect of Erzya (Kaljajewo) and two dialects of Moksha (Tschembar, Selischtsche). Based on the derivative alone, there seems to be no way of further assessing which of the meanings found should be considered primary, so we will turn our attention to the base word itself, E *aľa*, M *aľä*.

In her treatment of Mordvin kinship terms Mészáros (2001 : 169) lists Erzya *aľa*, M *aľä* as perhaps being Tatar loanwords. This assumption seems to stem from the fact that Feoktistov (1965 : 339) proposed a Qaratay etymology for the words: Qaratay *aľaj* 'husband's older brother' → E *aľa* 'man, uncle, husband's older brother', M *aľä* 'father, man'. According to Feoktistov, judging by the semantics, the Erzya word bears a closer affinity to the Qaratay word than Moksha. The Qaratays represent a group of formerly Mordvin-speaking people that became linguistically and socially Tatarised during the 17th century (Bartens 1999 : 10), so a borrowing from Tatar is *a priori* a possibility. No mention of any Tatar word is made in this connection, however, so it is not readily obvious why a Tatar origin should be preferred. It is perhaps equally possible that the word is a Mordvinic substrate item in Qaratay. Apparently both Erzya and Moksha speakers were Tatarised, so the Qaratay word being semantically closer to the Erzya word is of little consequence and does not necessarily imply that 'husband's older brother' is the most ancient meaning of the word.

The "Mordwinisches Wörterbuch" lists several meanings for E *aľa*, M *aľä*, of varying distribution (MdWb 35—38). The most prominent meanings confined mostly to Erzya dialects are 'elder brother' and 'husband's elder brother', and in individual Erzya dialects also the following meanings are found: 'husband's younger brother', 'wife's older brother', 'wife's brother and all mother's male relatives', 'husband's sister's husband', 'sister-in-law's husband in case the sister is older than ego's husband', 'father's brother', 'mother's brother', 'father-in-law's brother'. The meaning 'father' is found only in Moksha dialects, as well as 'husband'. The most common meaning for both Erzya and Moksha dialects is 'man, peasant' ('Mann, Kerl, Bauer'). The common thread for most meanings found in Erzya and Moksha dialects is that they denote a male relative, both consanguineal and affinal, older than ego. The amount of variation would seem to indicate that the word originally had a rather general meaning and could thus be used to refer to and address almost any older male relative. In conclusion, we indeed find an overlap between 'brother-in-law' and 'brother' in this group of words in Mordvin, but it is certainly not a clear-cut parallel for the semantic change from 'brother-in-law' to 'brother'. It is not, however, easy to conclusively say which of the several meanings found in Mordvin is the primary one, but judging by its prevalence in both Erzya and Moksha dialects, the most probable candidate is '(common) man'.

Other than the few examples presented above, there is no polysemic overlap between 'brother-in-law' and 'brother' in the Uralic languages based on the materials gathered for Kinura. In general, a kinship term only rarely seems to cross the boundary from consanguineal to affinal relative and vice versa, although single counter-examples can undoubtedly be presented, e.g. PU **čečä* 'uncle' > MdE *čiče* 'sister's husband older than ego' (UEW 34). What does all this mean, then, for the proposed loan etymology of EPF **weljə* 'brother'? One must conclude that although the impossibility of the semantic change the etymology relies upon cannot be ruled out, comparative lexical and etymological evidence from elsewhere in Uralic does not offer any corroborating evidence, which in my mind casts serious doubt on the validity of the loan etymology. I would even go as far as to say that etymologies which lack positive evidence in their favour should be regarded as deficient and rejected until actual semantic parallels are presented or the semantic connection is otherwise demonstrated to be true. I am very uneasy about treating semantics as an inconsequential factor that can be brushed aside by simply stating that whatever connection one is trying to prove is possible. I do not understand how this possibility is defined, it certainly cannot just be something that one making the claim is personally capable of imagining to be true. One would hope to see more thorough treatment of semantics in Uralic etymology and loanword research in the future.

Address

Niklas Metsäranta
University of Helsinki
E-mail: niklas.metsaranta@helsinki.fi

Abbreviations

EnF — Forest Enets; **EPF** — Early Proto-Finnic; **Est** — Estonian; **EstS** — South Estonian; **Fi** — Finnish; **Hung** — Hungarian; **J** — Jaźva Komi; **KhKaz** — Kazym Khanty; **Lith** — Lithuanian; **Liv** — Livonian; **LPF** — Late Proto-Finnic; **MariH** — Hill Mari; **MariM** — Meadow Mari; **MdE** — Erzya Mordvin; **MdM** — Moksha Mordvin; **MPF** — Middle Proto-Finnic; **MsSo** — Sosva Mansi; **NenT** — Tundra Nenets; **Ngan** — Nganasan; **ON** — Old Norse; **OPruss** — Old Prussian; **Oss** — Ossetic; **PF** — Proto-Finnic; **PI** — Proto-Iranian; **PIE** — Proto-Indo-European; **PII** — Proto-Indo-Iranian; **PM** — Proto-Mari; **PP** — Proto-Permic; **PS** — Proto-Saamic; **PSam** — Proto-Samoyedic; **PScand** — Proto-Scandinavian; **PU** — Proto-Uralic; **PUdm** — Proto-Udmurt; **Ru** — Russian; **SaaIn** — Inari Saami; **SaaK** — Kildin Saami; **SaaLu** — Luleå Saami; **SaaN** — North Saami; **SaaPi** — Pite Saami; **SaaS** — South Saami; **SaaSk** — Skolt Saami; **SaaT** — Ter Saami; **SaaU** — Ume Saami; **SlkTaz** — Taz Selkup; **Udm** — Udmurt**.**

LEW — E. F r a e n k e l, Litauisches etymologisches Wörterbuch I—II, Heidelberg 1962—1965; **LägLoS** — A. D. K y l s t r a, S.-L. H a h m o, T. H o f s t r a, O. N i k k i l ä, Lexikon der älteren germanischen Lehnwörter in den ostseefinnischen Sprachen I—III, Amsterdam 1991—2012; **MdWb** — H. Paasonens Mordwinisches Wörterbuch I—IV. Zusammengestellt von K. Heikkilä. Bearbeitet und herausgegeben von H.-H. Bartens, A. Feoktistow, G. Jermuschkin, M. Kahla, Helsinki 1990—1996 (LSFU XXIII. Kotimaisten kielten tutkimuskeskuksen julkaisuja 59); **MNySz** — Ö. B e k e, Mari nyelvjárási szótár (Tscheremissisches Dialektwörterbuch) I—IX, Szombathely 1997—2001 (Bibliotheca Ceremissica. Tomus IV); **NIL** — D. S. W o d t k o, B. I r s l i n g e r, C. S c h n e i d e r, Nomina im Indogermanischen Lexikon, Heidelberg 2008; **TschWb** — A. M o i s i o, S. S a a r i n e n, Tscheremissisches Wörterbuch. Aufgezeichnet von V. Porkka, A. Genetz, Y. Wichmann, M. Räsänen, T. E. Uotila und E. Itkonen, Helsinki 2008 (LSFU XXXII. Kotimaisten kielten tutkimuskeskuksen julkaisuja 151); **WotjWsch** — Wotjakischer Wortschatz. Aufgezeichnet von Yrjö Wichmann. Bearbeitet von T. E. Uotila & M. Korhonen. Herausgegeben von M. Korhonen, Helsinki 1987 (LSFU 21); **ÅDB** — K. B e r g s l a n d, L. M a t t s o n M a g g a, Åarjelsaemien-daaroen baakoegærja / Sydsamisk-norsk ordbok, Alta 1993.

B I B L I O G R A P H Y

A i k i o, A. 2007, Etymological Nativization of Loanwords. A Case Study of Saami and Finnish. — Saami Linguistics, Amsterdam—Philadelphia (Current Issues in Linguistic Theory 288), 17—52.

— — 2009, The Saami Loanwords in Finnish and Karelian, Oulu [Doctoral dissertation].

— — 2015, The Finnic 'Secondary *e*-Stems' and Proto-Uralic Vocalism. — JSFOu 95, 25—66.

B a r t e n s, R. 1999, Mordvalaiskielten rakenne ja kehitys, Helsinki (MSFOu 232).

B e r e c z k i, G. 1992, Grundzüge der tscheremissischen Sprachgeschichte II, Szeged (Studia Uralo-Altaica 34).

C a n t e r a, A. 2017, The Phonology of Iranian. — Handbook of Comparative and Historical Indo-European Linguistics. Volume 1, Berlin, 566—599.

C s ú c s, S. 1990, Die tatarischen Lehnwörter des Wotjakischen, Budapest (Bibliotheca Uralica 10).

D e r k s e n, R. 2015, Etymological Dictionary of the Baltic Inherited Lexicon, Leiden (Leiden Indo-European Etymological Dictionary Series. Volume 13).

F r a n ç o i s, A. 2008, Semantic Maps and the Typology of Colexification. Intertwining Polysemous Networks across Languages. — From Polysemy to Semantic Change. Towards a Typology of Lexical Semantic Associations, Amsterdam (Studies in Language Companion Series 106), 163—215.

G r ü n t h a l, R. 2012, Baltic Loanwords in Mordvin. — A Linguistic Map of Prehistoric Northern Europe, Helsinki (MSFOu 266), 297—343.

H a s p e l m a t h, M., T a d m o r, U. 2009, The Loanword Typology Project and the World Loanword Database. — Loanwords in the World's Languages. A Comparative Handbook, Berlin, 1—35.

H o l o p a i n e n, S. 2019, Indo-Iranian Borrowings in Uralic. Critical Overview of the Sound Substitutions and Distribution Criterion, Helsinki [Doctoral dissertation].

J o k i, A. J. 1973, Uralier und Indogermanen. Die älteren Berührungen zwischen den uralischen und indogermanischen Sprachen, Helsinki (MSFOu 151).

J u n t t i l a, S. 2015, Tiedon kumuloituminen ja trendit lainasanatutkimuksessa. Kantasuomen balttilaislainojen tutkimushistoria, Helsinki [Doctoral dissertation].

K a l i m a, J. 1936, Itämerensuomalaisten kielten balttilaiset lainasanat, Helsinki (SKST 202).

K a l l i o, P. 2014, The Diversification of Proto-Finnic. — Fibula, Fabula, Fact. The Viking Age in Finland, Helsinki (Studia Fennica Historica 18), 155—168.

—— 2018, Ensitavun diftongit kantasuomessa. — Περὶ ὀρθότητος ἐτύμων. Uusiutuva uralilainen etymologia, Helsinki (Uralica Helsingiensia 11), 251—268.

K o i v u l e h t o, J. 1993, Zur Etymologie von fi. *susi*, *tosi* und *kesi* — eine Entgegnung. — LU XXIX, 21—37.

—— 1994, Kuuluvatko *lanko* ja *lanka* yhteen? — Vir., 3—24.

—— 2001, Etymologie und Lehnwortforschung: ein Überblick um 2000. — FUF 56, 42—78.

K o r h o n e n, M. 1981, Johdatus lapin kielen historiaan, Helsinki (SKST 370).

L e h t i r a n t a, J. 1989 [2001], Yhteissaamelainen sanasto, Helsinki (MSFOu 200).

L i u k k o n e n, K. 1999, Baltisches im Finnischen, Helsinki (MSFOu 235).

L u o b b a l S á m m o l S á m m o l Á n t e (A i k i o, A.) 2012, An Essay on Saami Ethnolinguistic Prehistory. — A Linguistic Map of Prehistoric Northern Europe, Helsinki (MSFOu 266), 63—117.

—— 2014a, On the Reconstruction of Proto-Mari Vocalism. — Journal of Language Relationship / Вопросы языкового родства 11, 125—157.

—— 2014b, The Uralic-Yukaghir Lexical Correspondences: Genetic Inheritance, Language Contact or Chance Resemblance? — FUF 62, 7—76.

—— 2015, On Regular, Irregular and "Sporadic" Sound Change (Handout from a presentation at the XII International Congress for Finno-Ugric Studies. University of Oulu, Aug 19, 2015). https://www.academia.edu/51152618/On_regular_irregular_and_sporadic_sound_change.

—— 2020, Loanwords from Unattested Nordic Source Forms in Saami. — FUF 65, 5—24.

M a y r h o f e r, M. 1989, Vorgeschichte der iranischen Sprachen: Uriranisch. — R. Schmitt, Compendium Linguarum Iranicarum, Wiesbaden, 4—24.

M é s z á r o s, E. 2001, A mordvin rokonságnevek eredet szerinti vizsgálata. — Acta Universitatis Szegediensis 41, 169—184.

M e t s ä r a n t a, N. 2020, Periytyminen ja lainautuminen. Marin ja permiläisten kielten sanastontutkimusta, Helsinki [Doctoral dissertation].

M i l a n o v a, V. 2020, Kinship and Affinity in Indo-European: Universal, Inherited, Contact, and Area Specific Features of Indo-European Kinship Terms (with a Special Focus on the Iranian Branch), Vienna [Doctoral dissertation].

P o s t i, L. 1953—1954, From Pre-Finnic to Late Proto-Finnic. — FUF 31, 1—91.

Q v i g s t a d, J. 1893, Nordische Lehnwörter im Lappischen, Christiania.

R ä s ä n e n, M. 1920, Die tschuwassischen Lehnwörter im Tscheremissischen, Helsinki (MSFOu 48).

S a l m i n e n, T. 2002, Problems in the Taxonomy of the Uralic Languages in the Light of Modern Comparative Studies. — Лингвистический беспредел. Сборник статей к 70-летию А. И. Кузнецовой, Москва, 44—55.

S a m m a l l a h t i, P. 1984, Saamelaisten esihistoriallinen tausta kielitieteen valossa. — Suomen väestön esihistorialliset juuret. Tvärminnen symposiumi 17.—19.1.1980, Helsinki (Bidrag till kännedom av Finlands natur och folk 131), 137—156.

S w e e t s e r, E. 1990, From Etymology to Pragmatics. Metaphorical and Cultural Aspects of Semantic Structure, Cambridge.

U o t i l a, E. 1985, *Kohta* paikalleen. — Vir., 373—380.

W i c h m a n n, Y. 1903, Die tschuwassischen Lehnwörter in den permischen Sprachen, Helsinki (MSFOu 21).

W i k l u n d, K. B. 1896, Entwurf einer urlappischen Lautlehre I. Einleitung, Quantitätsgesetze, Accent, Geschichte der hauptbetonten Vokale, Helsinki (MSFOu X).

Г р и ш у н и н а, В. П. 2000, О терминах родства и свойства в мордовских языках. — Финно-угристика 4, 55—58.

Л ы т к и н, В. И. 1964, Исторический вокализм пермских языков, Москва.

Ф е о к т и с т о в, А. П. 1965, К проблеме мордовско-тюркских языковых контактов. — Этногенез мордовского народа, Саранск, 334—337.

НИКЛАС МЕТСЯРАНТА (Хельсинки)

'СЕСТРА', 'ДОЧЬ' И 'БРАТ': ОБ ЭТИМОЛОГИИ ЗАИМСТВОВАННЫХ ТЕРМИНОВ РОДСТВА В УРАЛЬСКИХ ЯЗЫКАХ

В статье рассматривается индоевропейская этимология трех уральских терминов родства ('сестра', 'дочь', 'брат'). Приводится обстоятельный этимологический анализ слов на базе результатов научных исследований последних лет. Автор попытался проследить связи слов, исходя из множества возможностей толкования, и предложить наиболее вероятный сценарий попадания их в лексику уральских языков. Анализировались как фонологическая сторона этимологии заимствований, так и семантическая, чему до сих пор при исследовании происхождения уральских слов часто внимания не уделялось.

NIKLAS METSÄRANTA (Helsinki)

'ÕDE', 'TÜTAR' JA 'VEND': UURALI KEELTE LAENATUD SUGULUSTERMINITE ETÜMOLOOGIAST

Artiklis käsitletakse kolme uurali sugulustermini ('õde', 'tütar', 'vend') indoeuroopa laenuetümoloogiat. Esitatakse sõnade põhjalik etümoloogiline analüüs uuemate uurimistulemuste põhjal. On püütud selgitada sõnade seoseid tõlgendamisvõimaluste paljususest lähtudes ja pakkuda välja kõige tõenäolisem stsenaarium, kuidas need sõnad on võinud uurali keelte sõnavarra jõuda. Analüüsitud on nii laenuetümoloogiate fonoloogilist kui ka semantilist külge, mis on uurali sõnade päritolu uurimisel sageli tähelepanuta jäänud.