A history of East Baltic through language contact

Jakob, A.

Part 1

CONTACTS WITH KNOWN LANGUAGES

1.1. Baltic–Slavic contacts

1.1.1. Early Slavic → Baltic loans

1.1.1.1. Preliminaries

The goal in this chapter is to establish the extent and nature of the earliest contacts between Baltic and East Slavic. While several studies have focused on the Slavic loanwords in Baltic (the pioneer articles being Brückner 1877; Endzelīns 1899; Būga 1911; Skardžius 1931), there have been few critical studies focusing on the earliest layer specifically: one often speaks of 'early' loanwords (e.g. Young 2009, Derksen 2020), but the actual material adduced encompasses ambiguous cases which could result in a skewed picture. The goal of this section is to identify the clearest and best quality data to substantiate the claim of early contacts.

The aforementioned ambiguities are in part the result of "traditional" substitution strategies (Stang 1957: 52–55). For instance, the examination of Lithuanian proper nouns transcribed into Cyrillic in 13th and 14th century documents, led Būga (1911: 18) to conclude that Proto-Slavic length was still contrastive at this time. The idea that such length contrasts were maintained well into the literary period is hardly tenable; however, it is clear that the substitutions Slavic *o* → Lt./Lv. *a* and Slavic *a* → Lt. *o*, Lv. *ā* have continued into recent times. This must at least in part be based on "traditional" equivalences extrapolated from earlier loanword strata (i.e. "etymological nativization"; see Aikio 2006b: 18–23 for a discussion of the concept). Stang notes Lt. dial. *notūrà* 'character, nature' from Polish *natura*, a recent Latin loanword;

and we can add here examples with Lt. *o* before a tautosyllabic resonant, which only became phonotactically possible in the last few centuries (see below), such as Lt. *kortà* ‘card’ (← Pl. *karta*; LEW 283), *gvõltas* ‘violence; uproar’ (← Pl. *gwałt* ← MoHG *Gewalt*; LEW 180).

Levin’s (2003: 141–142; cf. Derksen 2020: 44) reconstruction of a Proto-Lithuanian system with /aː/ and /ɒ/ seems to be more an attempt to force a phonetic explanation than something explicitly motivated by the data. True, a tendency for /ă/ to become rounded can be observed across the eastern edge of the Baltic territory, specifically in part of East Aukštaitian (but not on the Lithuanian–Belarusian border; cf. Zinkevičius 1966: 50–51 and Map No. 6), and as a conditioned change in Latgalian (Endzelīns 1923: 73–85). However, there is no reason to set this up as the most archaic system.[8]

Moreover, as the Finnic loanwords from Slavic also clearly show the preservation of inherited length (cf. Kalima 1956: 33–42), indicating an early date, a layer of loanwords in Baltic from the same chronological period would hardly be surprising. Note the following examples showing a reflection of length in unstressed syllables which are shared by Baltic and Finnic (cf. Stang 1957: 53):

Lt. *pyrãgas*, Lv. *pìrãgs* ‘pie’ (= F *piiraa*)	←	OR пирогъ (R *пирóг*, GEN.SG. *пирогá*)
Lt. *sopãgas*, Lv. *zàbaks* ‘boot’ (= K *soappoa*)	←	OR сапóгъ, GEN.SG. сапогá
Lv. *žìvats*[2] ‘animal’ (= F *siivatta* ‘livestock’)	←	OR живóтъ, GEN.SG. животá ‘life; animal’ (R dial. *животы́* ‘livestock; property’)

[8] On the substitution of Baltic **a* with Finnic **a* and **o*, see p. 119.

In Finnic, the smaller corpus of loanwords makes the suggestion of "traditional" substitution patterns less plausible, and the above examples must be accepted as early loanwords. However, this does not necessarily have any bearing on the age of these words within Baltic, and it cannot be ruled out that these too were borrowed at a later date following previously established nativization strategies. It is in fact highly difficult to identify the oldest layer of loanwords upon which the regular substitution patterns were originally based. The only unambiguous evidence of an early date of borrowing would be cases in which Baltic reflects phonemic contrasts subsequently lost in the history of East Slavic. I therefore limit myself to substitutions of this kind, and the resulting data will form the corpus for further analysis.

1.1.1.2. Reflection of yers

Finding unambiguous examples of the reflection of the Slavic reduced vowels in the Baltic loans is more complicated than usually recognized. Note, for instance, the following examples involving sequences of the type **CŭRC*:

- Lt. *kùrtas*, Lv. *kuŕts* 'greyhound' ← R, Uk. *хорт* 'greyhound' (= SCr. *hȓt*, Pl. *chart*); also → F *hurtta*, E *hurt*, Li. *ūrta-pi'ņ* 'greyhound' (Kalima 1952: 66)
- Lt. obs. *tùlkas*, Lv. *tuĺks* 'interpreter' ← OR *тълкъ, MR *молк* 'interpreter' (= OCS тлъкъ)
- Lt. *tuŕgus*, Lv. *tìrgus* 'market'[9] ?← OR, ONovg. търгъ, *u*-stem, 'market' (= OCS тръгъ); also → F *turku*, E *turg* (GEN.SG. *turu*) 'market' (Kalima 1952: 133)

[9] On the Latvian *-i-*, see fn. 23.

– Lt. *pulkas*, Lv. *pùlks* ‘crowd, troop’ ?← OR пълкъ ‘troop, regiment; crowd, throng’ (= OCS плъкъ)[10]

While these are generally acknowledged as early loans, it should not be overlooked that similar substitutions are also attested in relatively late loanwords. Historically, in the absence of a phoneme /ŏ/ in either Latvian or Lithuanian, a foreign /oRC/ has been substituted as /uRC/. Note, for instance, PrLt. *kuřbas*, dial. *kuřvis* ‘basket’ (← Prussian German *korb*, *korw*, Alminauskis 1934: 76; LEW 220) and Lt. dial. *guñtas* ‘roof shingle’ (← Pl. *gont*; LEW 176). A similar strategy is attested in Latvian loans from Estonian, e.g. *puĩka* ‘boy’ (← Võ. *poig*, E *poeg* ‘son’),[11] *kuĪda* ‘ash pit’ (← Võ. *kollõq*, GEN.SG. *koldõ*), cf. Thomsen (1890: 263–273). Most probably, the same applies for some dialectally isolated loans like Lv. dial. (Naukšēni) *burtenis* ‘empty beehive’ ← R dial. (Pskov, РЭС IV: 96) *бóртень* ‘hive used to attract bees’ (ME I: 354); Lv. dial. (Aloja) *turba* ‘knapsack’ ← R dial. *мóрба* (ME IV: 268); and High Latvian *pulna* ‘enough!’ ← R *пóлно* (cf. Būga 1925: 44).[12]

---

[10] The rare F *pulkka* ‘regiment, troop’, quoted by Ahlqvist (1871: 209), shows a limited dialectal distribution and occurs alongside the allegedly younger F-K *polkka*, E *polk* ‘regiment’. The antiquity of this loanword seems doubtful to me (cf. Mikkola 1938: 53). Note F dial. *hulkka* beside *holkka* ‘crowd; troop’, considered to be of Germanic origin (cf. LÄGLOS I: 119–120).

[11] Both formally and semantically, a more probable direct source seems to be Estonian Swedish *poik* ‘boy’ (Freudenthal/Vendell 1886: 165).

[12] Lv. dial. *tuĪpîties* ‘to crowd’ ← *молпи́ться* (Būga 1925: 43) may be based on the dial. 3SG.PRES. *мóлпится* (cf. СРНГ XLIV: 207), which must be the older form in view of the oxytone accentuation of *молпá* (Зализняк 2019: 208); cf. also Uk. *мóвпитися* (which must be analogical after 3SG.PRES. *мóвпиться*). Despite ME (IV: 260), it does not seem likely that the Latvian word is cognate with Lt. dial. *tùlpinti* ‘make room for, economize’. The latter seems somehow to have been formed secondarily from *tiĪpti* ‘fit, have enough room’ (compare

This observation raises suspicion with regard to the Aukštaitian dial. *bulvõnas* ‘idol; dummy’, Lv. dial. (Vidzeme) *buĪvāns* (ME I: 349), *bulens* (EH I: 251) ‘decoy bird’ ← Russian *болва́н* ‘idol; stuffed animal, dummy’ (Даль[2] I: 111), which have consistently been described as early loanwords (e.g. Būga 1925: 40; LEW 33; Derksen 2020: 40). While a source with preserved **o* is not probable in the Baltic area, it is conceivable that these were adopted from a dialect with dissimilative *akanje* typical of north-eastern Belarusian dialects (Wexler 1977: 79–80).[13] On the other hand, the distribution of the word within Baltic may not exactly favour such an interpretation; at least for Latvian, a certain role may here also have been played by Baltic German *Bolwan*, *Bulwan* ‘decoy bird’ (cf. Kiparsky 1936: 149).

The four cases cited at the start of this section could theoretically be dated to Proto-East-Baltic. In the case of Lt. *tuřgus*, Lv. *tìrgus*, cognancy with Slavic is also an option (Endzelīns 1899: 299; REW III: 123).[14] In

---

Smoczyński 2018: 1484), although the details are unclear. Incidentally, the Latvian accentuation might also speak against an old loan (see *1.1.1.7*).

[13] Similarly, Endzelīns (1899: 298) refers to the pronunciation of unstressed **o* as [u] in some Russian dialects (on this see ДАРЯ I: No. 2). This explanation seems possible for Lt. *tulkoczus* (Brodowski 923) ‘pestle’, cf. Bel. *маўка́ч* (Būga 1925: 763); perhaps also Lv. dial. (Endzelīns 1899: 299) *grumada* ‘assembly’, Lt. *grùmada* ‘crowd’ (Juška II: 478), *grummodas* ‘Haufe Fliegen’ (Ruhig II: 192) ?← Bel. dial. (*)*гр*[у]*ма́да*, if not simply ← Polish *gromada* ‘flock, crowd’.

[14] Further cf. Alb. *treg*, dial. *tregë* ‘market’ (cf. Meyer 1891: 323–324; Jokl 1924: 88). If we reconstruct **trg^{w}-*, we might blame the non-acute accentuation in Slavic on the *u*-stem (Stang 1957: 79–82; Pronk-Tiethoff 2012: 242–244); the circumflex in Baltic might support a Slavic intermediary. Note that a reconstruction **trg^{wh}-* would violate the IE root constraints. Note that Bjorvand/Lindemann (2018: 1328–1330) assume the Baltic and Slavic words are loans from Germanic, cf. ON *torg* ‘market’, but it is doubtful there are any Norse loans in Proto-Slavic or in East Baltic (see *1.2*). Significantly, their account fails to explain the Albanian data.

principle, the Slavic loan etymologies in all of these cases are encouraged primarily by circumstantial facts, for instance the existence of early parallel loans into Finnic from the same sources, although some phonetic details favour the loan etymologies.[15]

The Slavic word **pulka/u-* 'crowd; military regiment' is a loan from Germanic (cf. OHG *folc* 'people, crowd, troop').[16] It is possible that the Baltic words are parallel loans from a related Germanic source, possibly even West Germanic *folk*, rather than having been mediated through Slavic. In fact, several other suggested early Slavic loanwords in Baltic are ultimately of Germanic origin. Hirt (1898: 350) in this and other cases has assumed direct adoptions from Gothic (see also *1.2*), while Būga (1922: 71) has preferred to assume a Slavic intermediary:

- Lt. *ãsilas* 'donkey' ← OR осьлъ 'donkey' / ← Go. *asilus* 'donkey'
- Lv. *bruņas* F.PL. 'armour' ← OR бръня 'armour' (also бръня F.PL., cf. СРЯ 11–14 I: 321) / ← Go. *brunjo* 'breastplate'[17]

---

[15] For *kùrtas*, the correlation between Baltic *k-* and Sl. *x-* favours a Slavic → Baltic loan. For *tùlkas*, the acute accent in Baltic would be in disagreement with OIr. *do-tluchethar* 'seek, demand', Lat. *loquor* 'talk' (which rule out a root-internal laryngeal), which might favour a Slavic origin. See *1.1.1.7*.

[16] As the substitution Germanic **o* → Sl. **u* is unproblematic (Slavic had no **o*), attempts to track down West Germanic forms with /u/ (cf. Pronk-Tiethoff 2012: 112–113) are unnecessary.

[17] For the Slavic word, Pronk-Tiethoff (2012: 122–123) considers it impossible to decide between a Gothic or West Germanic source. However, on the basis of OR пѣнязь 'coin' (СДРЯ 11–14 IX: 407) ← OS *penning* 'penny' (where we must assume an original **pę̄nę̄ʒe-* by dissimilation; cf. the close parallel in OCS мѣсаць 'moon', dissimilated from **mę̄sę̄ce-* < IE **meh₁ns-* 'month'; Shevelov 1964: 320; Beekes 1982: 55), one would rather anticipate West Germanic **brunnjā* (> OHG *brunna*, MHG *brünne* 'chain mail, breastplate') to be borrowed as Slavic **brǭnjā-* (> OR **броуня). Therefore, a Gothic source seems preferable.

- Lt. *kãtilas*, Lv. *katls* ‘kettle’ ← OR котьлъ ‘kettle’ / ← Go. *katils** ‘kettle’
- Lt. *stìklas*, Lv. *stikls* ‘glass’ ← OR стькло ‘glass’ / ← Go. *stikls* ‘cup, chalice’[18]

It is curious that all of the above words have attested equivalents in Prussian, viz. Pr. E *asilis* ‘donkey’, *brunyos* ‘armour’, *catils* ‘kettle’, *sticlo* ‘drinking glass’. This suggests yet another possible route these words could have taken: they may have entered East Baltic through Prussian. In fact, the realia seem to speak against *ãsilas* ‘donkey’ being an old word in Lithuanian. According to G. Piličiauskienė (p.c. April 2023), there is so far no evidence for donkeys in the Lithuanian zooarchaeological record, and no documentary evidence for them having been bred or traded. It is therefore quite probable that the word was taken from Prussian by Bretke for the purpose of translating the bible.

On the other hand, it is not possible to rule out Bel. *асёл*, Pl. *osioł* ‘donkey’ as a source either, as this Slavic suffix appears as Lt. *-ilas* even in recent loanwords. Thus, South Aukštaitian *bùsilas* ‘stork’ (see Naktinienė et al. 1988: 48) ← Bel. *бу́сел* (GEN.SG. *бу́сла*) cannot be an old borrowing in view of its short first-syllable vowel. Likewise, Lt. *kazilaĩ* ‘saw-horse’ can hardly be from OR (*)козьлъı (*pace* Būga 1925: 41; Smoczyński 2018: 513), as this particular sense is a Polonism, and in turn a calque from German *Bock* (REW I: 590). In view of this, it is also possible that Lt. *kùbila*s ‘tub, barrel’ is a relatively

[18] Although the Slavic word seems to fit better semantically, it cannot be excluded that the word also meant ‘glass (material)’ in Gothic; cf. the polysemy exhibited by e.g. English *glass*, and also Lt. *stìklas*, Pr. E *sticlo* in the sense ‘drinking glass’. For further discussion of the semantics, see Kiparsky (1934: 210–211).

recent loan from Polish *kubeł* 'bucket' rather than representing an early loan from the source of RCS къбьлъ*, OPl. (hapax, 15th c.) *gbeł* 'bucket' (cf. Būga 1925: 38–39).[19]

As for the other words, no conclusions can be drawn as to their proximate source. Even if they were adopted through Prussian, it is unclear whether the Prussian words were themselves adopted directly from Gothic or via West Slavic. Given the multitude of possibilities, these words can hardly serve as evidence of direct early contact between Slavic and East Baltic.

Incidentally, Levin (1974: 88) has suggested that certain other "general Baltic" Slavicisms spread along a trajectory from west to east; the following seem to be decent candidates:

- Lt. *kùmetis* 'serf, peasant' in Suvalkia and Prussian Lithuania ?← Pr. E *kumetis* 'gebuer' ← Lechitic **kumeti-* (> OPl. *kmieć* 'serf, peasant').
- Lt. *krìkštas* 'baptism' ?← Pr. **kriksta-*, cf. Pr. III *crixti lāiskas* 'tauffbüchlein' ← Lechitic **krista-* (> OPl. *krzest* 'baptism'; SSP I: 257)

For Lt. *krìkštas*, an ultimately West Slavic origin is attractive for semantic reasons. Continuants of Proto-Slavic **krista-* mean 'baptism' in traditionally Catholic areas (i.e. West Slavic and Slovene), while in Eastern Christianity they have the meaning 'cross, crucifix'.

[19] Both Slavic words are of West Germanic origin, cf. MHG *kübel* 'bucket, tub; dry measure' (Pronk-Tiethoff 2012: 126–127). Latvian *kubls* 'tub, barrel', like *katls* 'kettle' are taken as instances of syncope by Endzelīns (1923: 47), but could just as easily be relatively recent loans from the Slavic oblique stems in Pl. GEN.SG. *kotła*, *kubła*; cf. similarly Lt. *pãslas* 'ambassador' ← Pl. *paseł*, GEN.SG. *pasła*. The High Latvian *kubyls* cited by Būga (1925: 38) as evidence of an unsyncopated form is considered dubious by ME (II: 297).

Accordingly, a sense ‘baptism’ is unattested in Old Russian (СДРЯ 11–14 IV: 315–317), which practically rules out an East Slavic origin for the Lithuanian form, only attested in the sense ‘baptism’ (*pace* Skardžius 1931: 110; ALEW 607).[20] Lt. *krikščiónis*, Lv. dial. (Zemgale) *krišķāns* ‘Christian’ may also have come via the same route (cf. Pr. III *crixtiānai** NOM.PL.). This has both linguistic and non-linguistic implications.

On the linguistic side, the *-š-* in Lt. *krìkštas* has sparked much discussion. Būga (1912: 3; 1925: 41) considered a possible German influence, while Endzelīns (1911: 60; 1937: 164) suggested contamination with Polish *chrzcić* /xš-/ ‘baptise’. However, the prevailing opinion is that that the shift to *-š-* is due to the intrusive *-k-* (Endzelīns 1911: 35; Skardžius 1931: 110; Stang 1966: 14; Smoczyński 2018: 606). Whether this is chronologically plausible is uncertain; it is in any case worth noting that several words which show a pervasive *-k-* do not exhibit a subsequent shift to *-š-* in Lithuanian (e.g. Lt. *áuksas* ‘gold’ = Pr. E *ausis*; Lt. *aĨksnis*, dial. *alìksnis* ‘alder’ = R *ольха́*; Lt. *úoksas* ‘tree hollow’ = Lat. *ōs* ‘mouth’).

A Prussian transmission gives us a new possible explanation. In transitional Prussian–Lithuanian dialect areas, there would undoubtedly have been a significant level of bilingualism, providing the prerequisites for “etymological nativization” to take place. The regular

[20] It is perhaps this semantic issue which encouraged Smoczyński (2018: 606) to interpret *krìkštas* as a back formation from the verb *krìkštyti*, 3 PRES. *-ija* ‘baptise’, which, in his opinion, is in turn derived from OR крьсти́ти (Зализняк 2019: 365) ‘baptise’. However, the root stress in the verb rather points towards a denominal formation. Importantly, the Lt. form corresponds precisely to Pr. III *crixti-*, PRES.SG. *crixtia* ‘baptise’ (whose second *-i-* never shows a macron, implying the root here was also accented). As a result, the Lithuanian verb could be explained as a Prussianism, as well.

correspondence between Lithuanian *š* and Prussian *s* may have been recognized by bilingual speakers, leading them to favour the seemingly counterintuitive substitution /s/ → /š/ over the phonetically more natural /s/ → /s/ (Mažiulis 1979: 147). A few other words suspect of being of Prussian origin may show a similar substitution strategy. At least the following can be cited:

- Lt. (W Žem.; cf. LKA I: No. 82) *bruñšė* ‘roach’ ← Pr. E *brunse* ‘roach’ (PKEŽ I: 161) beside Žem. *bruĩšė* (< **brų̃šė*, cf. Trautmann 1910: 145).
- Lt. (W Žem.) *jū́šė* ‘fish soup; slops, viscous liquid; mixture’ ← Pr. E *iuse* · juche.
- Lt. (W Žem.; PrLt.) *kriáušė* ‘pear’ ← Pr. E *crausy* ‘pear tree’, NOM.PL. *crausios* ‘pears’. The dialectal limitation of this and the previous word speaks in favour of a Prussian borrowing (Būga 1915: 342).[21]
- Lt. *šárvas*, PL. *šarvaĩ* ‘armour; arms’ ← Pr. E *sarwis* ‘weapons’, which is no doubt ultimately from Go. *sarwa* N.PL. ‘weapons, armour’ (= OE *searo*, DAT.SG. *searwe* ‘craft, wile’ and ‘weapon, armour’). Alternative etymologies (cf. LEW 965–966; PKEŽ IV: 65) derive from a need to explain the unexpected *š*-.
- ⟨Ƶukmistras⟩ ‘Fischmeister’ (ClG I: 663)[22] ← Pr. E *suckis*, III *suckans* ACC.PL. ‘fish’.

[21] Note that *kriáušė* is now the standard word for ‘pear’, but the more widespread dialectal terms are the Slavic loanwords *grūšià* and *dū̃lė*. Although this is not the place to discuss the full material, I agree with Būga that the application of the RUKI law cannot be considered regular after **u* and **i* in East Baltic.

[22] The stressed form *žùkmistras* given by Brugmann (1897: 104), and found abundantly elsewhere (even in LKŽ) is apparently not attested (cf. Būga RR II: 721). Mielcke (I: 341) and Nesselmann (1851: 552) have *źukmistras* without

From a non-linguistic perspective, this analysis removes one of the key arguments for an early informal Christianization from the east, which has been taken for granted since Būga (cf. Būga 1912: 11). Indeed, assuming a Prussian intermediary, one would hardly need to date such Christian terminology significantly earlier than Lithuania's official adoption of Christianity in 1387. In Latvian, on the other hand, one does find a trace of early eastern Christianity: Lv. *krusts*, dial. *krists* 'cross', whose meaning points to an Old Russian source.[23]

Despite these ambiguities, the following ten cases seem to provide compelling evidence of early contacts between Slavic and East Baltic:

- Lt. dial. *bìrkavas*, *bìrkuva* 'a weight of 10 pūdai', Lv. *biřkavs* 'ship-pound' ← ONovg. бърковьскъ, a weight measure, frequent in Novgorod gramotas (cf. Зализняк 2004: 713 and Thörnqvist 1948: 29–32, where the ultimate connection with ML *Birca*, OSw. *Biærkö* is discussed)
- Lv. *krusts*, dial. *krists* 'cross' ← OR крьстъ 'cross' (see above)
- ? Lt. *Póvilas*, Lv. *Pāvils* 'Paul' ← OR Павьлъ (Būga 1925: 44). However, see the discussion on p. 39

---

any stress mark (this form may ultimately derive from ClG). Kurschat (1883: 527) specifically notes that the word is "bei den Hafffischern ungebräuchlich".

[23] Latvian *-u-* is perhaps to be explained as the result of a contamination with Lat. *crux* 'crucifix' (Endzelīns 1899: 301). Būga (1925: 42–44) rather blamed the *-u-* in Lv. *rutks* 'radish' (see below) and *krusts* on the Belarusian depalatalization of /r'/, but this is chronologically implausible; cf. Wexler (1977: 153). Lv. *rutks* may be explained by assuming a dissimilation **i*–*i* > **u*–*i* which would find a partial parallel in *tìrgus* 'market' < **turgus*; however, it is usually assumed to have been influenced by *ruds* 'red-brown' (ME III: 565). Note the other examples of a hesitation between *-i-* and *-u-* listed in Young (2009: 187). For an explanation involving Latgalian /y/, see Seržant (2006: 99–100).

- Lt. *pipìras* ‘pepper’ ← OR пьпьрь, OPl. *pierz*, *pieprz* ‘pepper[24]
- Lv. *Pliskava* ‘Pskov’ (whence German *Pleskau*) ← OR Пльсковъ (Būga 1925: 42)
- Lt. *ridìkas*, dial. *rudìkas*, Lv. *rutks* ‘radish’ ← OR рьдькъı*, cf. редькови ACC.PL. (13th c.), R, Uk. *ре́дька* ‘radish’, of West Germanic origin, cf. MLG *reddik* (1492 in *Gaerde der Suntheit*; MoLG *Röddick*) ‘radish’ (Pronk-Tiethoff 2012: 136–137)
- HLv. dial. *pluts* ‘raft, barge’ (ME III: 359) ← cf. R *плот*, OPl. ⟨plty⟩ NOM.PL. (in Latin context; SSP VI: 167) ‘raft’. Probably of West Germanic origin; compare MDu. *vlotscip* ‘barge’, Du. *vlot* ‘raft’ (on which cf. Philippa et al. IV: 545–546; Kroonen 2013: 149)[25]
- Lt. obs. *smirdas* ‘peasant’ (cf. Būga 1925: 43), Lv. obs. *smirds* ‘poddany, subjectus’ (Elger 1683: 385; cf. ME III: 966) ← OR смьрдъ ‘peasant’, cf. OPl. *smard* (SSP VIII: 318)
- Lt. *šil̃kas* (also *plurale tantum*: *šilkaĩ*) ‘silk’ ← OR *шьлкъ, cf. ONovg. шолкоу GEN.SG. (Зализняк 2004: 541), R *шёлк*, Uk. *шовк* ‘silk’, most probably of Norse origin; cf. ON *silki* (REW III: 387)
- Lv. *zizlis*, also *zizls* ‘rod; spoke of a wheel’ (assimilated from **žizls*?) ← OR жьзлъ ‘stick, staff’

[24] REW (II: 341) has suggested the Slavic word may have been loaned through a Gothic **pipirs*, and we might equally assume a borrowing through Gothic in East Baltic. Yet if the Slavic word is directly ← Lat. *piper*, as usually thought (cf. M. Matasović 2011: 118), then reconstructing such a Gothic term seems superfluous. A Slavic origin would be favoured by the non-initial stress of Lt. *pipìras*. By contrast, Lv. *pipars* ‘pepper’ is possibly borrowed from Swedish via Est. *pipar* (cf. ME III: 221); compare Estonian Swedish *pippar* (Freudenthal/Vendell 1886: 163), OSw. *pipar* ‘pepper’ (EES s.v. *pipar*).

[25] A Germanic origin is rejected by Kiparsky (1934: 80), but without reference to the West Germanic data. Note that the similar MoHG *Floß* ‘raft’ goes back to **flōta-* rather than **fluta-*.

Since Погодин (1903: 161–162), Lv. *cìlvę̃ks*, dial. *cilę̃ks* ‘person’ has been viewed as an early loanword from OR человѣкъ ‘person’ (cf. ME I: 382–383; Young 2009: 183; Derksen 2020: 40). There are two phonological issues with this derivation. First, Погодин’s Slavic preform **čilvēka-* cannot be supported by any actual Slavic data. A slight improvement would be to start from a pre-Lv. **cilavēkas* with regular elision of the compounding vowel within Latvian (Endzelīnes 1923: 187). However, even a form *чьловѣкъ is unattested. Although the evidence of OCS чловѣкъ is perhaps inconclusive (the word was only rarely spelled out), the complete lack of vocalization of prepositions in West Slavic would speak against Pl. *człowiek* deriving from a variant **čilavēka-* (Havlová 1966: 80), and would favour an account starting from irregular syncope.[26] At any rate, in East Slavic, the only usual form is of the type человѣкъ, cf. also Uk. *чоловíк*, Bel. *чалавéк* ‘person’. As a result, a donor form with **i* is questionable.

Finally,[27] Lt. *grìkai*, *grìkiai*, Lv. *griķi* ‘buckwheat’ are often interpreted as early loans from the (unattested) Russian *грьк- (Rūķe-Draviņa

[26] Although seemingly a rather arbitrary suggestion, many of the Slavic reflexes of this word show irregular developments, which must result from frequency of use (to the lists in Berneker I: 140–141, and Havlová loc. cit. we can add R colloquial *чек* ‘person’). In addition, some Russian dialectal forms show an irregular raising, cf. Obojań *чилэк* (Шахматов 1915: 152), Rjazań *цылье́къ* (Даль³ IV: 1301), but so far, I have not identified any forms nearer to the Baltic territory.

[27] Other doubtful examples are the following: (1) Lv. dial. *buca* ‘barrel’ is more probably secondary for *muca* ‘barrel’ (as suggested by Mühlenbach in ME I: 344) under influence of R *бо́чка* ‘barrel’, rather than a direct loan from OR бъчи*, ACC.SG. бъчьвь (СДРЯ 11–14 I: 331; on the Slavic form, cf. REW I: 113–114); (2) Lt. dial. *cìrkva* ‘church’ is a form apparently only recorded by Būga (1925: 42); in a genuinely old loan we would expect **cirkuva*; (3) Lv. *dukurs* ‘polecat’ was probably adopted through E *tuhkur* (Kiparsky 1949: 65); (4) Lv. *siruobs* ‘notch at the end of a beam’ is hardly from OR *съроубъ (cf. R *сруб*

1964: 125; Smoczyński 2018: 387; РЭС XII: 91) or directly from *грьча (> MR *гре́ча*; Seržant 2008: 126). However, this appears to be excluded by the realia. Archaeobotanical evidence for buckwheat in the East Baltic region only certainly emerges in the 14th century, which coincides with the first documentary evidence (Sillasoo/Hiie 2007: 76; Grikpėdis/Motuzaitė Matuzevičiūtė 2020: 166). Moreover, the word is so far unattested in Old Novgorod-Pskov. Although an argument *ex silentio*, this fact is potentially significant, as terms for grains (пьшеница 'wheat', овьсе* 'oats', жито 'barley', ръжь 'rye') are abundantly present in the birchbark letters.

The Russian word is not in fact found until the turn of the 16th century (GEN.SG. *гречки* 1495, *гречи* 1498, *гречихи* 1500; СРЯ 11–17 IV: 132), where it emerges in Novgorodian trade books. In its German form, the word appears a century earlier in Riga (attested in 1383 as *Kricken*; Kiparsky 1936: 84; Rūķe-Draviņa 1964: 118), and crops up in the late 15th century in Prussian German *Greck*, *Grick* (PrWb II: 513) and Old Polish *grece*, *grice* (1487/1488 in Latin context; SSP II: 506). The forms with *-e-* are reminiscent of Middle German *Grecken* 'Greeks', *greckisch* 'Greek' (DWb IX: 256), opening up the possibility that the word was actually formed within German, and that R *гре́чка* as the 'Greek grain' is a calque coined in the context of Hanseatic trade. While this account remains quite uncertain, there is still no evidence that would allow us to backdate this word, or buckwheat cultivation, to the Old Russian period.

'log frame'; ME III: 848; Seržant 2006: 96); instead, the Lv. *-i-* may be epenthetic in the illegal cluster **sr-*; (5) Lv. dial. *timnica*, *timnice* 'dungeon', *timenîca* 'dark place; dungeon' (cf. OR тьмьница; Endzelīns 1899: 301; Būga 1925: 767) is perhaps rather built after *tìmsa*[2], dial. *tima* 'darkness' on the model of R arch. *темни́ца* 'dungeon' rather than being directly borrowed from it.

1.1.1.3. *TerT *and* *TarT

As the East Slavic pleophony clearly predates the earliest texts, one has often hesitated to accept loanwords predating this development. With regard to the Finnic evidence, one has either assumed the existence of dialects which evaded pleophony (Mikkola 1894: 45–47; Kalima 1929: 165) or that a phonetically weak anaptyctic vowel was lost in the borrowing process (Setälä 1929: 34; Kalima 1956: 31–33; Kiparsky 1963: 83). The scepticism of earlier scholars is tied to a conception that any form lacking pleophony must necessarily belong to Proto-Slavic proper.[28] However, this is certainly not the case; it has even been argued that the development spread through East Slavic as an areal feature (cf. Garde 1974: 112–115; Николаев 1988: 123–124; Крысько 1994: 18–19; Зализняк 2004: 39–41).

Bjørnflaten (2006: 66) has claimed that loanwords predating the East Slavic pleophony are only ascertained in Latvian. As far as Lithuanian is concerned, he refers to Zinkevičius (1987: 71), who argues that syncope cannot be excluded in any individual case. However, it is not clear why the same argument could not equally apply to Latvian: if we have Lv. *pẹ̀lni* 'ashes' (= Lt. *pelenaĩ*) and *èrglis* 'eagle' (= Lt. *erẽlis*) (Endzelīns 1923: 47), then why not suggest Lv. *kaĩps* 'farmhand' and dial. *kȁrms*[2] 'building' derive from an earlier **kalapas* and **karamas*, respectively? On the other hand, setting up hypothetical forms like

[28] The position is exemplified by the statement of Колесов (1980: 69): "Все славянские языки изменили исходное сочетание типа **tort*, но изменили по-разному. Следовательно, это изменение началось в праславянском языке" ["All Slavic languages modified the original sequences of the type **tort*, but in different ways. It follows, therefore, that this change started in Proto-Slavic"]. The second statement, however, does not logically follow from the first.

these to explain away any relevant evidence (cf. Mikkola 1938: 25–26) would be circular.

In the case of Lv. *žeřbiņš* 'lot' ← Bel. *жэ́рабя* (Būga 1925: 37), the lack of lengthening before **rC* proves that this sequence has arisen by syncope (Derksen 2020: 34, fn. 5; note also dial. *žerebiņš*, *žeberis*). Syncope is also quite imaginable in polysyllabic forms such as Lt. obs. *čerpyčia* ~ *čerepyčia* 'roof tile' ← Bel. *чарапі́ца*;[29] Lt. *karvõjus* 'wedding loaf', dial. *karavõjus* ← Bel. *каравáй*; Lt. *skavardà* 'frying pan' ← Bel. *скавараднá* (cf. Lv. dial. Vidriži *skavardnīca*, without lengthening of *-a-*). A case not explainable as a pre-pleophony borrowing is Lt. *kalmaškà* 'light carriage' ← Bel. *каламáжка* (= Pl. obs. *kolimaga* 'cart'; cf. Skardžius 1931: 99; Kiparsky 1948: 48; Zinkevičius 1966: 131–132).

On the other hand, there are words in both Lithuanian and Latvian for which the assumption of syncope would seem *ad hoc*:

- Lt. *čérpė* 'clay pot; roof tile' ← **čerpa-* (> R *черепóк*, Bel. *чэ́рап* 'potsherd', Pl. dial. *trzop* 'clay pot').
- Lv. *kaĩps* 'servant, farmhand' (= Lt. *kálpas*, if not from Latvian, cf. Derksen 2020: 34) ← **xalpa-* (> OR холопъ 'servant')
- Lv. *kȁrms*[2], i.e. HLv. *kùorms* 'building' ← **xarma-* (> OR хоромъ 'house, building'; СДРЯ 1387)
- Lt. *šálmas* 'helmet' ← **šalma-* (> OR шеломъ, шоломъ, OPl. in Latin context *szłom* 'helmet'), earlier **šelma-*.

[29] The first manuscript edition of Szyrwid has *ćierpićia*, which in the third edition is apparently corrected to *ćierepićia*. Both variants are also found in Bretke: *czerpjczios* NOM.PL. and *czerepijczes* ACC.PL. (see Skardžius 1931: 54–55; ALEW[2] s.v. *čerepyčia*).

The last example must have post-dated the backing **CelC* > **CalC*, which Holzer (2001: 42) has described as a "North Slavic" areal change. Note that a similar backing does not seem to be reflected in Lt. *šil̃kas* 'silk', which I have accepted above as an East Slavic loanword. This might suggest a chronological or dialectal difference in the loan source. The word for 'silk' does appear to have had a backed variant already in at least part of Old Russian judging by Finnish *sulkku*, Veps *šuuk* 'silk' ← OR *шълкъ (cf. also R *шёлк*, Uk. *шовк*).[30]

Occasionally, it is difficult to decide between cognancy and borrowing. Lt. dial. (N) *kařbas* 'basket', Lv. *kà̃rba* 'box; birchbark vessel, basket' could be borrowed from OR (Novg.) коробъ 'unit of measure', R *ко́роб* 'bast or birchbark vessel' (cf. Berneker I: 568; ME II: 194; REW I: 629), but a regular cognate cannot be ruled out.[31] A similar consideration applies to the derivative Lt. dial. (Szyrwid, S Aukšt.) *kařbija*, *karbijà* 'kind of woven basket' and its comparandum ONovg. коробьꙗ 'grain measure' (Зализняк 2004: 749).[32]

[30] Mikkola (1894: 117) cited a Russian dialectal "шулк", but it appears this is merely a hypothetical form based on Шахматов's (1893: 296) claim of a sporadic shift *o* > [o̯ ~ u] in Petrozavodsk Russian. This shift was later explained as a reflex of etymological **a* in accentually immobile words (see Л. Васильев 1929: 14). Since the vowel of R *шёлк* does not reflect an etymological **a*, the Finnic *u*-vocalism should rather be taken as a direct reflection of **u* (cf. Кулешов 2010: 349).

[31] *Contra* Berneker (cf. also LEW 220), it is not probable the Slavic word was borrowed from Germanic. First, the word is accentually mobile in Old Russian (Зализняк 2019: 527), which is atypical of Germanic loanwords (cf. Pronk-Tiethoff 2012: 242–244 with lit.). Secondly, from West Germanic **korb-* (OHG *korb*, OS *korf* 'basket, pannier'), I would anticipate Slavic **kurb-* (see fn. 16); a Germanic loan may well underlie Cz. dial. (Machek 1968: 291) *krb* 'dovecote', SCr. *kŕbulja* 'basket made of bark' (PCA X: 449), as suggested by Berneker.

1.1.1.4. Nasal vowels

The loss of nasal vowels appears to have happened very early in East Slavic; there is no trace of nasal vowels in any East Slavic dialect (cf. Shevelov 1979: 132).[33] Although certain Norse loanwords seem at face value to have predated the loss of the nasal vowels,[34] Шахматов (1915: 112–113) has suggested that this is illusory, and that syllable final nasals were simply omitted for phonotactical reasons. Indeed, omission plus compensatory lengthening better accounts for cases such as: OR Игорь (cf. MGr. Ἴγγωρ) ← OSw. *Inguar* (not *Ягорь); ижера 'Ingrians' ← Ingr. *Inkeroin*; ѩкорь 'anchor' ← OSw. *ankare* (not *оукорь; cf. Thörnqvist 1948: 99).[35,36] Similarly, for OR поудъ 'a weight

---

[32] Pr. E *tarbio* (for **carbio*) 'meal box' may, in turn, be analysed as a further Balto-Slavic cognate or as a loanword from Lechitic, cf. Pl. dial. *krobia* 'large woven basket' (PKEŽ II: 117–118).

[33] Николаев (1995: 111) has claimed that a distinction between *$\bar{o}$ and *$\bar{u}$ is preserved in some Carpathian Ukrainian dialects as /ʊ̯ʊ̯/ and /ʊ̯ʊ/, respectively, both of which supposedly had a number of allophones in free variation (idem: 107–108). I have been informed by S. Tarasovas that Николаев has rechecked the data using modern software, and (apparently) now rather considers these alleged reflexes to be phantoms.

[34] A key argument here is OR варѧгъ, MR *вáрязъ* (Зализняк 2019: 722) 'Varangian', cf. ON *væringjar* NOM.PL., pointing to **vārēga-*. On the other hand, if we set up a source form **várjængja-* (cf. Thomsen 1877: 121; Falk/Torp 1911: 1403), supported by MGr. βάραγγοι 'Norsemen in service of the Greek emperor', then we may also assume a direct substitution of *-æng-* with -яг-, without the need for an intermediary form with a nasal vowel. See below.

[35] Thörnqvist (1948: 105) ultimately settles on a reconstruction **ękorь*, which is unlikely given the Norse *a-*. Note that Lt. *iñkaras* 'anchor' and Lv. *ęñkurs* go back to Early MoHG *enker* 'anchor' (cf. ME I: 470) and do not support a Slavic nasal vowel, despite Berneker (I: 29).

[36] Perhaps this could account for OR грамота 'letter, literacy, written document' ← Gr. γράμματα (compare fn. 17). Although geminates were lost in Middle Greek, this was not universal, and this geminate occurs in the usual

measure' ← ON *pund*, one need not set up an intermediate form with **ǭ*; instead, we could posit a direct substitution of **-unC-* with **-ūC-* (Brückner 1929: 142; Kiparsky 1934: 157); cf. similarly the names Асмоудъ, Веремоудъ (cf. ON *Ásmundr*, *Vermundr*) mentioned in the chronicles (Thomsen 1877: 71–72; on these differently Николаев 2017: 28).

An earlier nasal vowel is supposedly proven by Lt. *pùndas* 'a weight measure', which would be from Old Russian (Būga 1925: 28; Thörnqvist 1948: 75). The key argument here is that *pùndas* cannot be separated from the other loaned weight measurement *bìrkavas* (cf. Derksen 2020: 37). On the other hand, Lithuanian also has a word *pū́das* (dial. also *pū̃das*), whose acute accentuation might suggest a relatively early loan (cf. *1.1.1.7*). Since the variant *pùndas* may well have been adopted directly from, or have been influenced by, Baltic German *pund* (Alminauskis 1934: 106–107; Kiparsky 1948: 37; LEW 667), it can hardly be treated as a certain example.

The only remaining plausible case is the ethnonym *lénkas* 'Pole',[37] which has been derived from an early **lę̄xa-* (> OR лѧхъ) (Brückner 1877: 103; Būga 1925: 33; LEW 356). Kiparsky (1948: 39; followed by

environment where 'spontaneous' secondary gemination is also attested (Holton et al. 2020: 135–136).

[37] Little faith can be given to the derivation of the ethnonym *Unguras** 'Hungarian', recorded only in Daukša, from Slavic **ǭgra-* (Būga 1925: 24; not **ǭgura-*: cf. Smoczyński 2018: 1561; Derksen 2020: 37). It is difficult to imagine what would motivate the early Lithuanians to borrow such an ethnonym, with there being no evidence of direct contact between Baltic and Hungarian-speaking groups, and with the latter having no particular folkloric significance in Lithuania. Instead, it seems obvious that this is a neologism created by Daukša on the basis of ML *ungarus* (Kiparsky 1948: 37), which would practically be proven by the fact that the Daukša himself uses the Polish loanword *veñgras* (viz. *Wę́grų* GEN.PL.) in another passage.

REW II: 84) has suggested the word may instead be taken from an unattested Lechitic source, yet there is no evidence this ethnonym was used as a self-designation by West Slavs.[38] A unique archaism would not be too surprising, as names of ethnic groups often figure among the earliest borrowings,[39] and the names of certain (presumably) Baltic tribes seem to have entered Slavic prior to the loss of nasal vowels: OR ѭтвѧгъı ACC.PL. (← **jātvingai*), голядь (← **galind-*); see p. 66.

*1.1.1.5. Reflection of Slavic *ȳ*

One might anticipate that Slavic **ȳ* would have been substituted with Baltic **ū*, yet examples do not present themselves (Derksen 2020: 40). While the usual substitution for Slavic **ȳ* in older East Baltic loanwords is long /ī/,[40] in a number of cases, we also find *-ui-*, which is often interpreted as archaic (e.g. Būga 1911: 25; 1912: 10–11; Kiparsky 1948: 31; Seržant 2006: 97, fn. 7). However, it is remarkable that several of the examples involve a preceding labial:

- Lt. *buĩlis* ‘chervil, chives’, cf. MR *быль* (СРЯ 11–17 I: 364), Bel. *былле̑* ‘weed’
- Lt. *muĩlas* ‘soap’ ← Bel. *мы́ла* ‘soap’
- Lt. *muĩtas*, Lv. *muĩta* ‘toll’ ← MBel. *мыто* ‘toll’

[38] Neither is the more primary **lę̑d-* reliably attested in West Slavic sources, although Hungarian *lengyel* (older *lengyen*) ‘Pole’, and the tribal name Λενζανηνοί mentioned in the 10th century *De administrando imperio* seem to imply a form **lę̑djān-* (cf. REW II: 84 for details and references).

[39] Thus OR литва ‘Lithuanians’ (see *1.1.2*), *вьсь (cf. Шахматов 1916a: 19) ‘Vepsians’ ← **vepsä* (REW I: 193), роусь ‘Rus’, probably ← Norse **rōþs-* (REW II: 551) and possibly Lv. *krìevs* ‘Russian’ (but see *1.1.1.6*).

[40] Examples include Lt. dial. *bagotỹrius* ‘rich man’, *kỹtras* ‘sly’ (LEW 29, 261), *tỹnas* ‘fencepost’ (LKŽ; cf. on *tuĩnas* below), Lv. *sìts*[2] (= HLv. *sèits*) ‘full’ (Endzelīns 1899: 310; ME III: 855).

- Lt. *smuĩkas*, dial. *smùikas* ‘violin’ ← MBel. *смыкъ* (16th c.; ГСБМ XXXII: 3) ‘a kind of instrument’, cf. modern Bel. *смык* ‘bow (for a string instrument)’
- Lv. dial. *vùikls*[2] ‘deft, able’, cf. R dial. *вы́клый* (Pskov, Novgorod; СРНГ V: 292) ‘experienced, able’ (ME IV: 676)

In these cases, the diphthong need not demonstrate any particular antiquity, but may simply be a result of the strong velarization of labials before /ɨ/. A similar representation is found in Richard James’ 17th century English-Russian notebook (*buic* ‘a bùll’, *muïla* ‘sope’, but *sït* ‘satisfied’, *yazïke* ‘a tounge’; Ларин 1959: 23–24) and the Russian manual of Tönnies Fonne (*buik* ‘bulle’, *muilo* ‘seepe’ but *iaszÿck* ‘tunge’; Касьян 2012: 73–78; Hendriks 2014: 94). In Finnic, the substitution *-ui-* is also only attested after labials: Karelian *vuitti* ‘portion, share’ (← R *вы́ть*); *muila* ‘soap’ (← R *мы́ло*), cf. Kalima (1956: 41). Incidentally, this group need not be archaic in Finnic either; on the contrary, the Karelian *buitto*, *puitto* ‘as if’ must be recent (← R dial. *бы́дто*, СГРС I: 242 ≪ *бу́дто*; Kalima 1956: 65; РЭС V: 53). In Finnic, it is also remarkably difficult to find examples of **ū* from Slavic **ȳ*.[41]

In Prussian, however, /ui/ substitutes Lechitic **ȳ* without exception: Pr. E *waldwico* ‘knight’ ← **valdȳkā*, cf. OPl. *włodyka* ‘nobleman’;

[41] In unstressed position, we find **u* after labials even in recent loanwords (F *populi*, E *pobul* ‘landless peasant; cottager’ ← R arch. *бобы́ль*), so that cases like Veps *kaput* ‘hoof’ (← *копы́то*) cannot serve as evidence. Even the **u* in F dial. *muula* ‘lye’, Võ. *mugõl*, *mukl* ‘soap suds; lye’ (← OR dial. **mȳgla-*, cf. R *мы́ло* ‘soap’; Ojansuu 1922: 139) is not necessarily probative, as it may as well be subsumed under the other examples of **ui*, PF ***muikla* being phonotactically impossible. Kallio (2008a: 155) cites only Vt. dial. *suura* ‘home-made curd cheese’ (← *сыр*), but this is found alongside many other dialectal forms (e.g. Luditsa *siira*, Jõgõperä *sõõra*, Mati *syyru*) and singling out the variant with *-uu-* would seem like cherry-picking.

**suiristio* (attested *sutristio*) ‘rennet’, cf. OPl. *syrzisko* ‘rennet’ (SSP VIII: 155); *wuysis* ‘guard dog’ ← **vȳž-*, cf. Pl. *wyżeł* ‘pointer’; Pr. III *zuit* ‘genug’ ← **sȳta*, cf. Pl. *syty* ‘full, satisfied’ (sceptically Levin 1974: 38–39). In East Baltic, there also remain a couple of examples of *-ui-* ← **ȳ* without a preceding labial, which may indeed represent archaisms:

- Lt. *kùila* ‘hernia’ ← OR *къыла, cf. MR *ки́ла* ‘hernia, outgrowth’ (Зализняк 2019: 192).[42]
- Lv. obs. *suits*, *suišs* ‘excessive’ (whence probably Lt. *suitus* ‘abundant’ in Daukantas)[43] ← OR съıтъ ‘full, satiated’, cf. R *сыто будетъ* ‘много, лишку, нескромны требованія твоя’ (Даль[2] IV: 386).
- Lt. *tuĩnas* ‘branch in a wattle fence; fence post; picket fence’ ← OR тънъ ‘fence; defensive wall’.[44]

---

[42] Since Endzelīns (1899: 310; ME II: 300), Lt. *kuilȳs*, Lv. *kuīlis*, Pr. E ⟨tuylis⟩ */kuilis/ ‘boar’ have been derived from an Old Russian *къıль (thus also Trautmann 1910: 451; LEW 305; Derksen 2020: 41). However, such a source form cannot be set up (cf. Sabaliauskas 1968: 176; PKEŽ II: 294). R dial. *киля́к*, *кило́н* ‘boar’ (are clearly derived from *ки́ла* ‘hernia’ (cf. *горбу́н* ‘person with a hump’ < *горб* ‘hump’); compare the dialectal senses *кило́н* ‘animal with a hernia; an animal (usu. piglet) with abnormal testicles; uncastrated boar’ (СРНГ XIII: 209). Perhaps, as Smoczyński (2018: 622) surmised, *kuilȳs* etc. was formed within Baltic from *kùila* ‘hernia’. In that case, HLv. *kèiļs* (= *ķìlis*[2], ME I: 388) ‘boar’ could have been similarly formed to *ķìla* ‘hernia, outgrowth’ (Bērzgale, EH I: 706), of the same origin. Alternatively, and perhaps more attractively, we could conceive of a relationship to Lt. *kiaũlė* ‘pig’ (see Sabaliauskas 1968: 175–177).

[43] The existence of the form *suitis*, attributed to Daukantas by Geitler (cf. LEW 937; also in Miežinis 1894: 232), is questioned by Būga (RR II: 724). In the LKŽ, all of the data from Daukantas is listed under *suitus*.

[44] In view of PrLt. *rùimas* ‘space’ (cf. MLG *rūm*), *šliùižė* ‘sluice’ (cf. MLG *sluse*; Prussian German *šlūse*; Alminauskis 1934: 129), *šiùilė* ‘school’ (cf. MHG

These tie into the debate regarding the phonetic value of Slavic **ȳ*. Already since Miklosich (1878: 149–152), a minority viewpoint has been that **ȳ* was pronounced as a diphthong /ui/, for which the evidence which has been adduced is so diverse that it seems difficult to dismiss out of hand (see Thomson 1927; Shevelov 1964: 377–379; Press 1986: 217–243; Касьян 2012: 84–88).[45] The theory that the development from **ū* to /ɨ/ in Slavic went through an intermediate stage with a diphthong **ui* (rather than through a direct delabialization as suggested by e.g. Kortlandt 1989a: 50) would explain the lack of loanword evidence for a stage with **ū*, despite evidence for pre-vowel-shift values for all other vowels in the Slavic vowel system.

According to Shevelov (1964: 378; and Levin 1974: 39), since *-ui-* still renders Belarusian *-ы-* ("which undoubtedly was and still is a monophthong"), the loanword evidence cannot be used. However, I see no reason to suspect that the above three loanwords should be late adoptions from Belarusian. As generally acknowledged (cf. Derksen 2020: 41), *kùila* cannot be a particularly young loanword, as it must have predated the change ъı > и after velars. Būga dates this to the 13th century in Belarus (1925: 52; cf. Колесов 1980: 155). Even though it might have taken place as late as the 14th century in Novgorod-Pskov (Зализняк 2004: 90–91), it is clear, in any case, that we are not dealing with a late Belarusian borrowing here. Based on this example, it would be reasonable to date the other two examples to an earlier date as well.

---

schuole; Prussian German *šoil*; idem: 126), Prellwitz (1891: 35) has suggested that *tuĩnas* might be derived directly from MLG *tūn* 'hedge, fence', like Lv. dial. *tûna*, *tũna* 'a fence of slanted planks' (ME IV: 282). The circumflex accent seems to tip the balance in favour of a Slavic origin, however.

[45] The most striking indication perhaps remains the fact that this sound is rendered with the digraph *ŭi* in OCS: Glagolitic ⟨ⰚⰈ ~ ⰚⰋ⟩, Cyrillic ⟨ъı ~ ъи⟩.

In any case, there is nothing in these words that specifically favours a younger dating.

*1.1.1.6. Reflection of Pre-Slavic *ǭ, *ę̄*

In Latvian, Slavic **ū* is frequently substituted as /uo/, for instance: *duõmât* 'think', *kàpuôsts* 'cabbage', *karuõgs* 'banner', *muõkas* F.PL. 'torment', *sùodît* 'punish, judge', *suõma* 'bag, satchel' (← R *ду́мать*, *кану́ста*, *хору́гвь*, *му́ки*, *суди́ть*, *сума́*, respectively). These have almost always been interpreted as archaic. Endzelīns (1899: 306) and Būga (1912: 14–15) suggested that they reflect a preserved Slavic diphthong **ou*, but the *communis opinio* is now that they represent a monophthong **ǭ*, predating the common Slavic raising to **ū* (McKenzie 1919: 170; ME I: 533; Kiparsky 1948: 33–34; Young 2009: 178; Derksen 2020: 45).

In Lithuanian, only one generally accepted example of this substitution is known: *kuodẽlis* 'flax prepared for spinning' ← R *куде́ль*. This state of affairs has led to the conclusion that the Latvian loanwords are generally more archaic than the Lithuanian ones (Bjørnflatten 2006: 67; Derksen 2020: 48–49). However, this is not consistent with the other lines of evidence for early loanwords, where the Latvian evidence does not significantly surpass the Lithuanian in any other category (see Table 2, overleaf).

It seems very unlikely that Latvian simply happened to borrow a large number of Slavic words containing **ū* at an earlier date than Lithuanian. We therefore must agree with McKenzie (1919: 171) that treating the different reflexes as representing distinct chronological layers is unwarranted.

Table 2. Archaic features in Slavic loanwords in Lithuanian and Latvian

	Lt.	Lv.
Reflection of *yers*	8 [+ 5]	10 [+ 4]
TerT groups	2 [+ 2]	2 [+ 1]
Nasal vowels	1	-
Slavic **y*	2	1
-uo- for **u*	1	**18+**

One possible solution, hinted at by Derksen (2020: 43), is that Lithuanian and Latvian were in contact with distinct varieties of East Slavic. It is conceivable that the reflex of Slavic **ū* was pronounced closer to */o̰/ in the East Slavic dialect with which Latvian was in contact. A similar contrast in the source dialects probably underlies the difference in outcome of Slavic **ē*, which has yielded *-ie-* in Lithuanian but *-ē-* = */æː/ in Latvian, cf. Lt. dial. *griẽkas* but Lv. *grę̀ks* 'sin' (cf. Bjørnflaten 2006: 68; Derksen 2020: 39). In this respect, we can note sporadic instances of /'a/ from **ē* recorded in modern north-western dialects, cf. NW dial. *кя́п* 'flail', Arxangeľsk *ря́па* 'turnip' (= R *цпь*, *рпьа*; Николаев 1990: 60), *ця́лой* 'whole', Vladimir *медвя́дь* 'bear' (= *цпьлый*, *медвпьдь*; Галинская 1993: 39–40), which indicates that the difference was at least partly dialectal rather than diachronic (cf. Derksen 2020: 47).[46]

On the other hand, Seržant (2006) has offered a plausible alternative account for Latvian /uo/: he suggests that it is the result of a dialectal diffusion from east to west. Latgalian (High Latvian) has undergone a chain shift **o̰* > *ū*; **ū* > *ou* ~ *yu* (Endzelīns 1923: 95–97). If this predated the influx of Slavic loanwords, one would expect East Slavic /ду́мать/ to have been adopted directly as High Latvian *dùmuôt'*. This

[46] Compare the substitution of this phoneme as **ā̆* in the Finnic loanwords (Kalima 1956: 37–38), matching Latvian.

would in turn be nativized in Low Latvian as *duõmât*, following the typical dialectal correspondences.[47] Since it is generally assumed that the early loanwords in Latvian are of East Slavic origin, such a trajectory would not be surprising.[48]

Interdialectal borrowing might also explain the Žemaitian forms *põ̜uks* 'down (of a bird)', *õ̜ustâ* NOM.PL. 'moustache', which ultimately derive from East Slavic *nyx* 'down', *усы́* 'moustache' (cf. Zinkevičius 1966: 79, fn. 27 and 84, fn. 30). These may be hypercorrections based on South Žemaitian dialects where **ō̜* has merged into **ū* (see the discussion in Būga 1912: 23–26). A similar explanation probably accounts for NW Žem. *viẽšnė* (= *vę̃išnę*) 'cherry', which Būga (1922: 177) has seen as an early loan from Slavic **vę̄šnjā* (cf. R *ви́шня*), but is more likely a hypercorrection based on South Žemaitian *vĩ·šnę* (cf. Aukšt. *vyšnià*).

While a number of examples supposedly reflecting a preserved Slavic **ę̄* (< **ei*) have been presented, the only convincing one is the Latvian ethnonym *krìevs* 'Russian', which is apparently loaned from OR кривичи (Endzelīns 1899: 285–286, 304–305, ME II: 284–285;

[47] In this respect, note that *prùods*^{2} 'pond' (← *пруд*; ME III: 400) seems to be an exclusively High Latvian word, so actually represents [prùds] (cf. the Latgalian form at vuordineica.lv, s.v. *dīķis*). The spellings in ME and Ulmann (*prohds* 'ein kleiner natürlicher Teich'; 1872: 212) are automatic transpositions of the dialectal form. From a High Latvian perspective, forms such as *kìukũļi* /kùkuļi^2/ 'corn cockle' (← *куколь*) and *pỳuka* (cf. *pùka*^{2}, ME III: 445) 'fluff' (← *nyx*), which have undergone diphthongization of original **ū*, might even be interpreted as more archaic (cf. Seržant 2006: 95).

[48] Similar evidence of interdialectal diffusion is shown by the Žemaitian diphthong /ie/, which occurs instead of the usual reflexes of **ę̄* in Aukštaitian loanwords, including words of ultimately Slavic origin, e.g. *griẽks* instead of **grę̃iks* 'sin' (Būga 1912: 7–8; Zinkevičius 1966: 86; Derksen 2020: 40).

Būga 1922: 177; Skach 2010: 137; Derksen 2020: 38).[49] As this word is an ethnonym, a unique archaism is conceivable, but we may suspect that *krìevs* similarly results from interdialectal diffusion. This seems, at least, to be the best account for Lv. *siẽrs* 'cheese', an evident borrowing from East Slavic *сыр* 'cheese', but one whose vocalism has presented a problem (ME III: 859). This can be resolved if we assume the word was adopted first as HLv. *sìrs* (Seržant 2006: 97) and then subsequently borrowed into Low Latvian. Such an explanation also works for the Lv. agent suffix *-(e)niẽks* (cf. Latgalian *-inìks*; Seržant 2006: 96–97), provided this is indeed loaned from East Slavic -ьникъ (for an analysis as cognate, see Derksen 1996: 185–186).

In fact, the loanword evidence for a Slavic stage **ę̄* is very poor; almost all of the evidence traditionally adduced from Finnic is doubtful or demonstrably false. The Finnish and Karelian agent noun suffix *-niekka* (McKenzie 1918: 172; Mikkola 1938: 33–34; Kiparsky 1948: 31; Derksen 2020: 39) whatever its origin, cannot be old, as long vowels outside of initial syllables were not possible at the time of the earliest contacts with Slavic.[50] Aside from this, there is K *viehkuri* 'gust of wind' and *miero* 'the (outside) world; township, village council'.

That all of the examples are limited to Finnish and Karelian obviously speaks against this being a particularly archaic loanword stratum. However, Kallio (2006: 155) and Derksen (2020: 39) are both unconvinced by the suggestion of a Russian dialectal development *i* > *e*

[49] Other suggested examples such as Lt. obs. *mieras*, Lv. *miêrs* 'peace' are better interpreted as cognates with the Slavic forms (Derksen 2015: 316; ALEW 747).

[50] As witnessed by the reflection of *yať* in the loanwords K *netäli*, Võ. *nätäl'* *-i*, Li. *nädīļ* (< **nätäli*) 'week' ← *недѣля* 'week' and F *veräjä*, E *värav* (< **väräjä*) ← *верея́* 'gatepost' (< **verējā-*, cf. OCS верѣꙗ 'bar, bolt').

(Mikkola 1894: 57).[51] While it is true that no regular dialectal change can be set up, the fact remains that both words are actually attested with /e/ in North Russian dialects. R dial. *вéхорь* ‘strong gust of wind’ is relatively widespread (СРНГ IV: 208: Kem’, Petrozavodsk; СГРС II: 88–89: Arxangeľsk), and most probably results from a contamination with *вѣтер* ‘wind’.[52] Olonets (19th c.) *мѣръ* ‘community’ (Куликовскій 1898: 58; СРНГ XVIII: 112) appears isolated, and I cannot explain it within Russian (cf. Мызников 2019: 496), but the Karelian data remains highly dubious grounds for assuming a layer of loans in Finnic with preserved **ę̄*.[53]

On the other hand, Finnic does provide clear evidence of an earlier **ǭ* (cf. Kalima 1956: 42; Kallio 2006: 155). The main disadvantage of the ‘dialectal diffusion’ scenario outlined above is that this situation in Finnic would have to be divorced from the superficially similar situation in Latvian. However, this is not necessarily a problem, as we are dealing with two distinct contact zones. In theory, it is possible that both scenarios are correct, and that an earlier layer of loanwords with **ǭ* was bolstered by a later layer adopted with Latgalian *ū*. In this case, however, the Latvian evidence can only be used as indirect support of early contacts with Slavic.

---

[51] Skach’s (2010: 138) suggestion that these represent relics of a Russian dialect preserving **ę̄* seems completely gratuitous to me.

[52] A parallel development is found in MUk. (17th c.) *вѣхорь*, which is actually attested in collocation with *вѣтеръ* (РЭС VII: 270).

[53] According to Kallio (2006: 155) the *-hk-* and *-u-* in *viehkuri* would favour an early borrowing. However, *-hk-* is also found in some very recent loanwords, e.g. F *orehka* ‘gingerbread, cookie’ < R *орѣх* ‘nut’, and is not probative. The *-u-* is probably due to the analogical introduction of the suffix *-uri* as in F *tuhkuri* ‘mink’, E *tuhkur* ‘polecat’ (← OR **дъхорь*), F dial. *pippuri* ‘pepper’ (← Sw. dial. *pipar*), F *ankkuri* ‘anchor’ (← Sw. *ankare*) (Kiparsky 1949: 60). Note that *-u-* is also found in the younger variant *vihuri*.

1.1.1.7. Accentuation

As Derksen (2020: 41–42) has observed, there appears to be a correlation between the intonation of the oldest Lithuanian loanwords and their accentuation in Old Russian. While circumflex is generalized in later loanwords, the circumflex examples from the oldest layer appear to correspond to Old Russian oxytones. As noted by Young (2009: 184–185), the same group appear to show a falling tone in Latvian:

- Lt. *pulkas* 4 ‘crowd, troop’, Lv. *pùlks* — OR (14th c.) полкы NOM.PL., R *полка́* GEN.SG. (cf. Зализняк 1985: 134; 2019: 569; Pronk-Tiethoff 2012: 112; Николаев 2020: 290).
- ? Lt. *tuĩnas* 4/2 ‘branch in a wattle fence, etc.’ — MR (16th c.) *тыно́мъ* INST.SG. (Зализняк 2019: 601);[54] cf. Slk. dial. *tý́ň* ‘branch in a fence’.
- Lt. *tuřgus* 2 (rare dial. *turgùs* 4) ‘market’, Lv. *tìrgus* — MR *до торгу́* ‘to market’, *торго́выи* ADJ., etc. (Stang 1957: 81; Зализняк 1985: 134, 2019: 534).

On the other hand, many of the examples with a Lithuanian acute correspond to Old Russian barytones. Here we find a sustained tone in Latvian (Young 2009: 179–181):

[54] It should be admitted that modern Russian generally suggests accent paradigm (c); cf. early modern Russian *тыновы́й* ADJ., and *тыни́ть*, 3SG.PRES. *тыни́т* ‘to fence’ (Сл. Акад. VI [1794]: 344; also dial., cf. СРГК V: 543). The evidence for accent paradigm (b) supplied by Зализняк is very limited, but note also OCz. *o-týniti* ‘enclose, cover’, SCr. (Vuk) *tíniti*, 1SG.PRES. *tînīm* ‘partition’ (RJA XVIII: 333). Pronk-Tiethoff (2012: 90), apparently on the basis of SCr. dial. *tȉn* ‘partition wall’, analyses the word as having fixed initial stress, but does not take into account the evidence of the derived verb.

- Lt. *bìrkavas* ‘a weight measure’, Lv. *biȓkavs* (ME I: 298; but [bìrkàu̯s] in LVPPV 130) — R *бе́рковец*, cf. dial. (Pskov) *бе́рковец* ‘weight measure for flax’ (РЭС III: 132–133)
- Lv. *kaĪps*, (Lt. *kálpas* 1 ‘servant’) — OR (Merilo) холωпъ, MR *холо́пъ* (Зализняк 2019: 602); SCr. arch. *hlȁp*, GEN.SG. *hlȁpa* (cf. Skok I: 671)
- Lt. *kùila* 1 ‘hernia’ — MR *ки́ла*, Uk. *ки́ла*; cf. Cz. *kýla*, Slk. *kyla,* SCr. *kȉla* (Зализняк 1985: 132, 2019: 192; Derksen 2008: 265)
- Lt. *kùrtas* 1 ‘greyhound’, Lv. *kuȓts* (but LVPPV: *kùrts*) — R *хорт*, cf. dial. *хо́ртица* ‘(female) greyhound’ (СРНГ LI: 316); SCr. *hȓt*, Sln. *hŕt*, GEN.SG. *hŕta* (Derksen 2020: 41)
- Lt. *lénkas* 1 (although LKŽ reports variants with 2, 3 and 4) ‘Pole’ — MR *ля́хи* NOM.PL. (cf. Зализняк 2019: 752)
- Lt. *Póvilas*, Lv. *Pãvils* ‘Paul’ — OR Па́велъ (Young 2009: 180; Зализняк 2019: 842)
- Lv. (Janševskis) *suītums* ‘Menge’ = dial. *sùits*^{2}, *sùitâk*^{2} (cf. ME III: 1116) — MR *сы́та* NOM.SG.F. = SCr. *sȉta* ‘satiated’ (Derksen 2008: 484; Зализняк 2019: 494)
- Lt. *šálmas* 3 ‘helmet’ — Although synchronically oxytone in Middle Russian (cf. Зализняк 2019: 588; Николаев 2020: 313), the word must originally have been barytone; cf. SCr. *šljȅm*, Sln. *šlẹ́m*, GEN.SG. *šlẹ́ma* ‘helmet’ (Pronk-Tiethoff 2012: 87)

This correlation provides another argument in favour of the late origin of Lt. *muĩlas* 4 ‘soap’, which corresponds to the barytone MR *мы́ло*, cf. Slk. *mydlo*, Sln. *mílọ* (Derksen 2008: 336), and of Lv. dial. *buĪvāns* ‘decoy bird’, which Young (2009: 186) has noted as an exception to his accentological rules. Note that Young uses a much larger corpus of Latvian data, while I have limited myself to cases which unambiguously belong to the earliest period. However, the accentual rules seem generally to apply even within his larger data set. The same cannot be said of Lithuanian, where the circumflex clearly dominates in the

remainder of the material (cf. Derksen 2020: 41). A remaining exception is Lt. *šil̃kas* 4 ‘silk’, the source of which appears to have been barytone (cf. Зализняк 2019: 568). I do not have an explanation for this form at this time.

There is rather little evidence for borrowings of accentually mobile forms. The clearest examples, perhaps counter-intuitively, appear to have been adopted as acute in Lithuanian (note also Lt. *bíesas* ‘demon’, *svíetas* ‘world’, from accentually mobile бѣсъ, свѣтъ; Derksen 2020: 42):

- Lt. obs. *tùlkas* 1 ‘interpreter’, Lv. *tuĺks* — MR *мо́лка* GEN.SG., *толковы́и* ADJ. (Зализняк 2019: 569); cf. R *бе́з толку* ‘in vain’
- Lt. *čérpė* 1 (but dial. also *čeřpė* 2, *čerpė̃* 4) ‘clay pot; roof tile’ — MR *че́репъ* (Зализняк 2019: 602); cf. R *черепа́* NOM.PL.

However, Lt. dial. *kařbas* 4 ‘basket’, Lv. *kȁrba*, if it is loaned from Slavic rather than cognate, would represent an exception, cf. MR *ко́робъ*, R dial. *из́ короба* (Зализняк 2019: 527). All in all, the evidence is rather too scanty to draw any conclusions. Young (2009: 186, 187) in fact reaches the opposite conclusion (i.e. mobile nouns are borrowed in Latvian with falling tone) based on evidence which has not come into consideration here. The regular adoption of Slavic barytones with acute accentuation in East Baltic is, however, clear from my data set, and demonstrates that an inherited accentological contrast was still present in Slavic at the time of the earliest loanwords into Baltic.

1.1.1.8. Semantics, dating and context

Much of the above evidence seems to indicate contacts in the context of trade. These include the following (the loans which are not as certain are given here in square brackets):

- Words connected to the act of trade: ‘market’, ‘interpreter’ and perhaps ‘raft/barge’. Considering their general limitation to adverbial usage, it is possible that Lv. *suiti* ‘excessive’, *suitāk* ‘too much’; Pr. III *zuit* ‘genug’ originated as trade jargon. Compare similarly Italian *basta* ‘enough!’, which has been borrowed widely, in many cases, presumably, through trade (cf. e.g. Snoj 2003: 33, s.v. *bâsta*).
- The names of vessels: ‘clay pot’, [‘tub’, ‘kettle’]; and weight measurements: Lt. *bìrkavas*, [*pū́das*].
- The names of specific trade items: ‘pepper’, ‘radish’, ‘silk’, probably ‘greyhound’, [‘glass’].

In view of the borrowed term *bìrkavas*, it would seem obvious to associate these trade relations with the *Birka* trade network (thus explicitly Būga 1913: 34–35); however, it is disturbing that there is no unambiguous evidence of direct Norse loanwords in East Baltic (see *1.2*), and besides, the weight measure in question continued to be used after the collapse of Birka as a trade hub, being even recorded in the modern dialects of Pskov Region. At the same time, this term does place us in a rather narrow timeframe between the establishment of the eastern trade with Birka in the late 9th century (Ambrosiani 2005) and the loss of the reduced vowels in Novgorod Russian in the early 13th (Зализняк 2004: 60). Lt. *čérpė* ‘clay pot; roof tile’, if indeed transferred in a trade context, would place the contacts firmly in a preliterary context, the 10th century at the latest. We can therefore assume that most of these loans were adopted between the 10th and 12th centuries CE.

There are some words which suggest an unequal power dynamic between the two groups. First, there are the words for members of the lower echelons of society: ‘peasant’, ‘servant/farmhand’, [*kùmetis*

'serf'], and words connected with the military: 'helmet', ['armour'].[55] Finally, there is the Latvian word for 'cross', which suggests that, like Finnic-speaking populations (Kiparsky 1952: 70–71; Kallio 2006: 156), Latvians were subject to early attempts at Christianization on the part of the Slavs. The only possible evidence of this in Lithuanian is the Christian name *Póvilas* 'Paul' (but see the discussion on p. 39). All of these loanwords are suggestive of Slavic cultural imposition, and therefore can be classed as typical 'superstrate' words (cf. Vennemann 2011: 240).[56]

Of at least 20 certain loanwords, six (just under a third) are found in both languages. in most cases a Proto-East-Baltic reconstruction can be provided. Nevertheless, this fact can be attributed to the small number of phonological changes which have taken place, and need not compel us to assume these loanwords were already present in Proto-East-Baltic. The idea of loanwords into separate languages is supported by the rather large number of 'old loans' limited to one of the two, while many of the shared loanwords represent trade terms which have been borrowed into numerous other languages. Contact with different East Slavic dialects seems to be implied, at least, by the establishment of different substitution strategies for the phoneme **ē*.

In terms of the source dialect(s), the following can be said:

---

[55] Here also belongs Lt. *pulkas* 'regiment', but I suspect that the dialectally better represented sense 'crowd' is original and the military sense may be due to more recent Slavic influence.

[56] Finally, a small number of loanwords are too vague to be categorized: 'building', 'rod', 'fence post'. Since the original specific function of these borrowed terms cannot be determined, they may have been loaned in any number of contexts. Surprising is the word for 'hernia', as no other medical terms or terms for bodily defects appear in my corpus.

- The borrowing of the weight measurement *бе́рковец*, frequent in Novgorod-Pskov sources and preserved in this area in the modern dialects, suggests contact with the dialect of Novgorod-Pskov, which is also supported by the early adoption of Latvian *Pliskava* 'Pskov'.
- In addition, the adoption of OR търгъ as a *u*-stem might also favour a Novgorod-Pskov source, as the morphological distinction between *a*- and *u*-stems was much better preserved in this dialect than in the rest of Slavic (Николаев/Хелимский 1990; Зализняк 2004: 99–102, 112).

As a result, the general picture is that the earliest contacts between speakers of Lithuanian and Latvian with East Slavic took the form of relatively incidental adstrate contacts with similar but distinct dialects of Old North Russian, primarily in the context of trade. At the same time, there is some limited evidence of Slavic cultural imposition, which suggests a degree of Slavic dominance in these contacts.

1.1.2. Early Baltic → Slavic loans?

There are very few loanwords which can be plausibly dated to the same period as the early Slavic → Baltic loans, and reference works on the subject (e.g. Лаучюте 1982, Аникин 2005) tend to focus on the much more extensive later layer of loanwords. Most of the plausible early loanwords which can be identified are ethnonyms. In the Novgorod First Chronicle, we find examples like OR лит[ъ]ва 'Lithuanians' (= R *Литва́*, Pl. *Litwa* 'Lithuania'; ← Lt. *Lietuvà*; cf. Lv. *Lìetava*, ME II: 506); ѭтвѧгъı ACC.PL. 'Jatvingians', голѧдь 'Galindians' (cf. the region *Galindia* cited in *Chronicon terrae Prussiae* and the Γαλίνδοι in Ptolemy), корсь < *кърсь 'Curonians' (cf. Lt. *Kuršas*, Lv. *Kuŕsa*[2]

'Curonia'); see the overviews in Fraenkel (1950a: 60–73) and Dini (2014: 290–312).

A list of probable early loans has been given by Аникин (2014: 192), who divides them into several chronological layers. As 'Proto-Slavic' loanwords, he quotes R *дёготь* 'birch tar' and dial. *пертьь* 'cottage'.[57] As Baltic substrate words in early East Slavic, he cites *дере́вня* 'village; (dial.) arable field', dial. *а́лес* 'damp spot', *му́ма* 'bogey; (in children's language) louse' and *пу́сма* 'bundle'. Аникин evidently means to compare *му́ма* 'bogey' (Brjansk, СРНГ XVIII: 344) with Lt. *maũmas* in the same sense, where **au* > **ū* would suggest a very early date; but compare also Lt. dial. *mū̃mas* (LKŽ), which is evidently the origin of dial. *му́ма* 'louse' attested in Lithuania (Лаучюте 1982: 146). It seems far more probable that dial. *му́ма* is merely an arbitrary formation like Hungarian *mumus* 'bogey' (in children's language; note also Lt. *baũbas*, *bùbas* in the same sense), and treating it as an exceptionally early loanword is unwarranted.

Among the 'early' loanwords, Kiparsky (1973: 68–69) has mentioned *па́кля* '(flax or hemp) tow' (← Lt. *pãkulos*; cf. also Аникин 2005: 24) and *ковш* 'ladle, water scoop' (← *káušas*). There is no reason to consider the latter to have been adopted any earlier than the 14th century (cf. СРЯ 11–17 VII: 216); the same sound substitution is found in *Vytolth* 1386, *Витовт* 1396 ← *Výtautas* (Būga 1911: 36–37), and remained usual throughout the 15th century (op. cit. *passim*). As for *па́кля*, the substitution *a* → *a* clearly rules out an early date. The loss of

[57] Аникин also cites R *клѣть* 'storehouse', with widespread Slavic cognates, as a Baltic loanword (cf. Lt. *klė́tis*, Lv. *klẽts*). Here, he follows Eckert (1983: 86–87); however, Eckert's main argument, namely that the Baltic word is derivable from the verbal root *klóti* 'lay out' is already refuted by Аникин himself (2005: 170), and there seems no other reason to prefer a Baltic source over an inherited cognate.

the second syllable of the Russian word must be secondary; cf. dial. (N) *ná́кула*, Bel. *ná́кулле*, Pl. *pakuły* ‘tow’ (Лаучюте 1982: 18–19); compare the discussion of R dial. *ná́ккула* beside *ná́кля* ‘chaga (parasitic fungus)’ in Мызников (2019: 571–573).

The following cases deserve a more detailed discussion:

‘bath-house’. ONovg. *пьрть ← Lt. *pirtìs*, Lv. *pìrts* ‘bath-house’ — The word is attested in the Sermon of Ilya of Novgorod (*перьти* DAT.SG. in a 15th century copy, cf. СДРЯ 1772), where it probably referred to a kind of bath-house (Павловъ 1860: 19). This is supported by MR *перетьь* ‘bath-house’ (Pskov, 15th c.; Вахрос 1963: 157); *пере́дка* ‘hut’ (Pskov, 16th c.; СРЯ 11–17 XIV: 298) and the dial. (Novgorod, Karelia) derivative *при́перетоκ* ‘dressing room in a bath-house’ (Мызников 2019: 625; cf. СРЯ 11–17 IXX: 245). The Baltic source has an impeccable internal etymology: it is a derivative of Lt. *peřti*, Lv. *pȅrt* ‘beat (e.g. with a besom); bathe’.[58]

Existing etymological discussions make the mistake of conflating the above forms with R dial. (Kem’) *пертьь* ‘Karelian cottage’ (Подвысоцкій 1885: 120). The latter, however, in view of its meaning and geographical isolation, is most certainly a recent loan from Karelian *pertti* ‘hut, cottage’ and not a direct continuation of the Old Russian form. Incidentally, the Novgorodian word has itself also been derived from Finnic (Būga RR II: 516; Лаучюте 1982: 89; ALEW 899),

[58] The older meaning is ‘beat’ (cf. OCS пьрѣти сѧ ‘argue’, пьрꙗ ‘fight, dispute’). Lt. *pirtìs* is also attested as a verbal noun ‘bathing; flogging’, and it has been recorded as a root noun (Zinkevičius 1966: 265), which might make a direct connection with Skt. *pŕ̥t-* ‘battle, strife’ possible. On the other hand, the attestation as a root noun seems to be limited to areas where root noun inflection became productive (cf. Zinkevičius 1966: 263).

yet given the meaning ‘bath-house’ in the earliest attestations, a Baltic origin is semantically more attractive. For a further discussion, see p. 233–236.

Despite the former’s narrow distribution, the Russian and Baltic words have often been interpreted as cognates (Vasmer 1909: 142; Trautmann 1923: 215; REW II: 344–345; Nieminen 1953: 214–215; Derksen 2015: 358–359). The main argument for a native Slavic origin is the existence of R *на́перть* ‘church porch’, which has a much broader distribution within East Slavic, and has an OCS cognate, папрьтъ (SJS III: 14), in the same sense. However, that these contain the same root is not self-evident. Beside the semantic obstacle (Преображенский II: 47), we can add that the variant *напортъ* (СРЯ 11–17 XIV: 148) might imply an original **purt-* (with *на́перть* due to *yer* assimilation? cf. Соболевскій 1910: 116–117).

? **‘arable field’**. R *дере́вня* ‘village; (dial.) arable field’ ← Lt. *dirvà* ‘arable field’ (Schmid 1977: 51–53; Аникин 1998: 319, 2014: 192, РЭС XIII: 230) — Schmid argues in favour of a Baltic origin, noting that place-names containing the word are concentrated in the Upper Dniepr and north of Moscow, which in his opinion would be consistent with the area of Baltic influence.

The loanword etymology implies the so-called “second pleophony”. Curiously, this development has been assumed in this word even by scholars who do not favour a Baltic origin (e.g. REW I: 341; Sławski SP V: 57–58; Derksen 2008: 136), where it appears only to have been motivated as a means to more directly equate the Baltic and Slavic forms. However, if we assume the words are cognates, the older reconstruction **derv-* (Berneker I: 186), with a different ablaut grade, can hardly be ruled out (compare, with *o*-grade, ME *tare* ‘vetch seed; vetch’, MDu. *tarwe* ‘wheat’ < **tarwōn-*). We can note that the second pleophony is usually a dialectal phenomenon that rarely has a pan-

East-Slavic distribution. Николаев (2001: 88), besides *дере́вня*, cites only *верёвка* ‘rope’, but the latter could just as well be built analogically from *вервь* (in many places /ver’v’/) on the model of e.g. *селёдка* : *сельдь* ‘herring’.

The main argument in favour of a loan is the word’s narrow distribution. On the other hand, a suitable Baltic source is unattested. Schmid (1977: 52) assumes an original syntagm **dirvinē žemē* ‘arable land’ (cf. Lt. *dir̃vinis* ‘related to *dirvà*’) was subsequently substantivized in Russian. If the loan etymology is valid, it would be equally acceptable to start from an unattested nominal derivative **dirvinē* ‘arable field’ already in Baltic. Despite the doubts of Аникин (РЭС XIII: 231), I find it at least possible that *дере́вня* ‘cleared land; arable land’ is the same word as R dial. (W) *дере́вня* ‘(pile of) logs’, Uk. arch. *дере́вня* ‘timber’ and is therefore derived from *де́рево* ‘tree’ (cf. Jēgers 1969: 79; Vaillant 1974: 608). A possible semantic path could be ‘felled trees’ → ‘area where trees are felled’ → ‘cleared land’. Needless to say, this remains hypothetical.

? **‘carrot’**. R dial. *борка́н* (Novgorod, Pskov, Smolensk, Kaluga, etc.) ← Lv. *buř̃kãns* ‘carrot’ (Karulis I: 155; РЭС II: 222; Pronk/Pronk-Tiethoff 2018: 282) — Traditionally, the Latvian word has been derived from Slavic (Būga 1925: 48; REW I: 108); however, the limited distribution of the word within Russian implies the opposite directionality. The Russian vocalism would imply an early borrowing as OR **бъркань* (РЭС II: 222), which, while theoretically acceptable, is rendered slightly awkward by the late attestation of the word within Russian (since 1564 *apud* СРЯ 11–17 I: 294; cf. Bentlin 2008: 247).[59]

[59] Аникин interprets dial. *бурка́н* as a later Letticism, but the limitation of this form to Leningrad Region practically excludes such an interpretation. It is evidently the result of pretonic *o* > /u/ attested sporadically in the area

The picture is further complicated by Baltic German *Burkane*, *Borkane*, which could, phonologically speaking, just as well be the source of the Latvian and Russian words (Mikkola 1894: 91; РЭС II: 223). Иллич-Свитыч (1960: 17) has argued that the Baltic German words are instead Balticisms, but does not address Kiparsky's (1936: 201–202) argument that the stress — /burkánə/ — would speak against this. Kiparsky also argues against a Russian origin, stating that the form *Purkahne* (from 1577) is attested "lange vor Beginn der russischen Zeit".

Masing (1926: 80) connects the Baltic German forms to MLG *brackannyen* NOM.PL. appearing among a list of edible roots in the *Loccumer Historienbibel* (15th c.).[60] This is supplemented by Marzell (II: 62–64) with some Scandinavian dialect forms that appear to be of Low German origin, cf. Early Modern Danish *brekanne-rod* 1550, *barkena-roer* 1738 'carrot' (ODS s.v.), Sw. dial. (SW) *barkan-rot*, *barken-rot* 'carrot' (with *rod*, *rot*, etc. 'root'); cf. also Bentlin (2008: 248–249). Indeed, as Marzell states, it seems almost inconceivable that these forms are unrelated, yet the Low German *a*-vocalism is hardly reconcilable with Baltic German *-u-.*

Most probably through Russian, the word has spread to F *porkkana*, Vt. *borkkana* 'carrot', Võro *põrḱnas* 'carrot' (Mikkola 1894: 91; Kalima 1956: 107; SKES III: 604; Plöger 1973: 141; SSA II: 375).[61] It seems that

(ДАРЯ I, No. 1); compare СРГК I: 97, where forms of the type /burkán/ are listed under the headword *боркáн*.

[60] Masing cites the form as *brackannige* after Schiller/Lübben (I: 412), who set up this reading with a question mark. MndWb (I: 339) normalizes the form as *brakannie*, instead. Here, I have cited the actually attested spelling.

[61] The substitution of Russian pretonic **o* as Võro. *õ* before tautosyllabic /r/ is paralleled at least by Estonian *kõrts*, Võro *kõrtś* 'tavern' < R *корчмá* (Blokland 2005: 199–200).

Livonian *borkõn* must also be derived from East Slavic; at any rate, Kettunen (1938: 26) denies the possibility of a late Latvian loan. Beyond this, analysing the exact routes of borrowing is highly challenging, and the word can at best be characterized as a circum-Baltic term which has spread as a local *Wanderwort*. On Moksha *puŕ'kä* 'carrot', and for a discussion of the word's ultimate origin, see p. 361–363.

? **'drying barn'**. R *овúн* 'drying barn', Bel. *авíн* 'granary' ← Lt. *javaĩ* 'cereals' (Andersen 1996a: 154–155; Pronk/Pronk-Tiethoff 2018: 290) — The Slavic word is generally considered an inheritance from Proto-Indo-European (Berneker I: 455; REW II: 249; ЭССЯ XIII: 187–188; Трубачев 1994: 7). This cannot be ruled out, but in view of the geographical limitation, a Baltic loan etymology looks attractive. Although no precise Baltic source is attested, one could certainly imagine a formation **javýnas*, with the collective suffix *-ýnas* (Skardžius 1941: 266–267; Andersen 1996a: 155; thus 'a collection of grains' ≫ 'granary'). This would imply a development **jaw-* > **(j)ew-* > *ov-*. According to Andersen, we are rather dealing with a direct substitution **ja-* → **a-* due to the inadmissibility of an anlaut **jă-* in Proto-Slavic.[62] In either case, we would probably be dealing with an early loan, although an exact dating is uncertain.

---

[62] Primarily on the strength of R *ебáть* = Skt. *yábhati* 'to copulate' (1996a: 14, 155), Andersen assumes that original **je-* did not develop to *o-* in East Slavic. However, the evidence is not quite clear-cut: at least Uk. *орябок* ~ Lt. *jerubẽ* 'hazel grouse' would speak in favour of such a development (Andersen 1996a: 137–138 is misguided in disregarding the Baltic evidence for **j-* in this word; see the discussion on p. 282). Note also that Andersen is forced to assume an *ad hoc* early loss of **j-* in the words OR оже 'if, that' and оли 'when, if' (idem: 152–153; ~ Lt. *jéi*, Lv. dial. *jà*, Go. *jabai* 'if'; Gr. ὅτε 'when, as'; Dunkel 2014: 320–322).

? **'fish trap'**. R dial. *вя́терь*, *вя́тель* (and variants, cf. СРНГ VI: 79–80; РЭС IX: 255–256), Uk. dial. *в'я́мір*, Pl. *więcierz*, Kash. *wiącel* 'fyke net (kind of fish trap)' ← Lt. *vénteris* 'fyke net' (Būga 1922: 298; REW I: 245; Аникин 2005: 111–112) — This loan etymology is widely accepted. On the other hand, Brückner (1927: 620) has analysed the Lithuanian word as a loan from Polish, a position supported by Kiparsky (1948: 39, fn.).[63] On phonological grounds, this is difficult to rule out (even though the distribution within Slavic is suggestive of a Balticism), especially since the word's ultimate origin is uncertain.[64] At a later date, the Baltic word was certainly borrowed into Russian *ве́нтерь* and Prussian German *Wenter* (on the latter, see Frischbier II: 464, where an account of the realia is also given).

? **'marshy spot'**. R dial. (W) *алёс* 'damp, marshy spot'; Bel. dial. *алёс* 'alder forest in a swamp'; Pl. dial. (Lithuania) *olesie* 'swamp in a forest' (cf. Толстой 1969: 159; Черепанова 1973: 72; РЭС I: 158) ← Lt. *álksna*, Lv. dial. *àlksna* 'alder thicket; marshy spot' (Топоров/ Трубачев 1962: 199; Аникин 2005: 85–86; Мызников 2019: 45) — The *-k-* in Baltic is intrusive, and the original form can be set up as **álisnā* with syncope (Friedrich 1970: 70; Топоров ПЯ I: 53). A trisyllabic preform neatly accounts for the Lithuanian acute, and an unsyncopated variant is preserved in Szyrwid (SD[1]) *alixnis*, dial. (NE) *alìksnis* 'alder' (the reconstruction of two forms for Proto-Baltic as per Derksen 2015: 50–51 is unnecessary).[65]

---

[63] Contrast the more cautious wording in Kiparsky 1973: 69–70, 1975: 93–94.

[64] The derivation from Lt. *vánta* 'besom', Lv. dial. *viẽtêt*(EH II: 798) 'flog' (LEW 1223–1224; ALEW 1405), assuming an original meaning 'fish trap woven from twigs' does not seem compelling. Note that the Kaišiadorys Museum encyclopaedia (accessed at kaisiadoriumuziejus.lt/enciklopedija) specifically states that the distinguishing feature of a *vénteris* compared to other fish traps is the *absence* of a supporting frame.

The Slavic words have been viewed as cognate in some recent sources (Andersen 1996a: 127; Derksen 2015: 50–51; ALEW 23), although none of these authors attempt to account for the *-s-* in the above words as opposed to **-x-* in **alixā-* 'alder'.[66] This phonological difference could favour a Baltic origin. In view of the distribution, the loan etymology looks highly attractive, although it is hampered by the absence of an attested donor form.

† **'birch tar'**. R *дёготь*, Uk. *дьо́готь*, Pl. *dziegieć*, Cz. *dehet* 'birch tar' ← Lt. *degùtas*, Lv. *dęguts* 'birch tar' (Mikkola 1894: 111; Zubatý 1894: 423, fn. 4; Būga 1922: 141; Kiparsky 1973: 68; Лаучюте 1982: 12) — The advantage of the loan etymology is that the Balto-Slavic verb **deg-* 'to burn' is only found in the assimilated form **žeg-* (< **geg-*) in Slavic. However, this is hardly a decisive argument: if the formation were taken to be of Balto-Slavic age, the word's semantic specialization would make fertile grounds for a preserved archaism. The main counter-evidence to a Baltic origin is the word's existence in Czech-Slovak (Trautmann 1923: 49; Brückner 1927: 109), yet this is somewhat circular given that other loanwords of this potential age are so few.

Although the verb *deg-* is synchronically present in Lt. *dègti*, Lv. *degt* 'burn', the derivative **deguta-* can hardly be recent. In Lithuanian, the suffix *-uta-* is rare (Skardžius 1941: 361; Ambrazas 2000: 103–104)

---

[65] Schrijver's assertion (1991: 42) that syncope "did not occur in Lithuanian" is simply false, cf. Zinkevičius 1966: 131–135.

[66] East Slavic **s* would actually be the expected result of the progressive palatalization in this word, but in that case, we should expect Polish *-sz-*. On the other hand, the Polish variant is dialectally very limited and may be from East Slavic. It remains quite unclear (to me, at least) why the progressive palatalization did not occur in the word for 'alder' itself.

and nowhere else is it deverbal.[67] A comparable suffix is found only in Lt. *ãšutas* ‘horsehair’, where it looks old (= R *ocóm* ‘sow thistle’, Pl. *oset* ‘thistle’), and Lt. *ríešutas*, Lv. *riẽksts* ‘nut’. Since the word clearly cannot be young in Baltic, it may as well be of Balto-Slavic age, and there is also no particular reason to consider it the source of the Slavic words. As an argument against a loan etymology, one can also point to the archaic-looking athematic OCz. *dehet*, GEN.SG. *dehte* ‘turpentine tree’ (Gebauer I: 220; cf. Trautmann loc. cit.).

† **‘bundle’** R dial. *пу́сма* ‘bunch, bundle’ (Kursk, Voronež; СРНГ XXXIII: 142) ← Lv. *puõsms* ‘section, interval’ (Аникин 2005: 258) — Аникин claims that the older meaning of the Latvian word was ‘bundle’ (“связка”). However, this seems to derive from a misunderstanding of Karulis (1992 II: 74), who merely suggests that the older meaning might have been “mezgls” in the sense ‘node on a plant stem’ (= “der Knoten beim Schilf oder Getreidehalm”, Seewald 1865: 68). In my opinion, this is not likely to be the oldest meaning, but is rather the result of a metonymical shift from ‘internode; interval’. An original meaning ‘interval’ is also supported by the Slavic cognate in R dial. *на́смо* ‘length of yarn, lea’. All in all, Вершинин (IV: 384) is almost certainly correct in deriving the Russian words from Mordvin, cf. Md. E *pusmo*, M *pusma* ‘bunch, bundle’, meaning *пу́сма* has nothing to do with the Latvian word.

~ ~ ~

[67] There is a diminutive *-ùtas*, largely limited to Southern Lithuania (Ambrazas 1993: 56–57). However, it is almost entirely restricted to velar-final stems and is therefore the result of distant dissimilation from **-uka-* (Hasiuk 1970), which makes it unlikely that these represent an archaism (*contra* Ambrazas loc. cit.). The connection of these forms to the Lithuanian diminutive suffix *-ùtis* therefore remains unclear.

As can be seen from the above discussions, the main argument for analysing any word as a Baltic loan is its distribution. Only in the case of R /Bel. dial. *алёс* ‘marshy spot’ is a possible phonological argument available. In all the other cases, there is no phonological obstacle to treating the words as cognates. As a result of this and other ambiguities by way of loan sources, almost all of the examples must be considered uncertain.[68]

---

[68] I have attempted to identify other words with a limited distribution which might be interpreted as Baltic loans, but these have mainly turned out to be problematic:
(1) R dial. (N) *ля́га* ‘swampy area; low, damp place’ has been equated with Lt. ⟨léngė⟩ ‘ein Wieschen zwischen zweyen Anbergen’ (Ruhig I: 76) (cf. REW II: 65). However, this Lithuanian variant is only known from Ruhig, corresponding elsewhere to *lénkė* ‘swampy meadow; hollow’ (Smoczyński 2018: 688). The reliability of the Lithuanian form is therefore questionable. Compare, perhaps, a similar sporadic voicing in ⟨kengras⟩ ‘hager’ (Ruhig II: 188) = *keñkras* (Kupiškis *apud* Būga in Juška III: 76)? For the Russian word, Николаев (1988: 135), offers an alternative etymology, comparing Vologda *ля́жа* ‘damp, boggy place’, and deriving both from **lę̄djā-*, with a suggested (albeit controversial) Novgorodian development **dj* > *g*. Another account is given in Мызников (2019: 466).
(2) R dial. (Vjatka) *черв* (Даль[2] IV: 607) ‘sickle’ has been considered cognate to Lt. *kiř̃vis*, Lv. *cìrvis* ‘axe’ (Berneker I: 172; Trautmann 1923: 135; REW III: 317; ЭССЯ IV: 171; Derksen 2015: 248). The word is known only from Даль, where it is cited alongside *черп* ‘sickle’. The latter has been recorded in other dialects (e.g. Perm, Беляева 1973: 689 and Arkhangelsk, Левичкин/ Мызников 2014: 180), but *черв* does not seem to be. The form ⟨черва́къ⟩ · пила?, also cited here by Даль (followed by Zubatý 1894: 388, then Berneker and Vasmer) is not likely to belong here and is rather to be equated with dial. *червя́к* ‘cross-cut saw’ (СРГС V: 274) which is probably a semantic extension of *червя́к* ‘worm’. If we assume Даль’s *черв* was extrapolated from a phonetic [čerf], we might think of the facultative alternation /f ~ p/ reported in this dialect area (Сметанина/Иванова 2018: 208, cf. *подчефре́ниться* ~ *подчепре́ниться* ‘dress up’).

Nevertheless, I think that Old Novgorodian *пьрть ‘bath-house’, at least, is a highly probable loanword from Baltic. The distribution of this word would support the supposition made in *1.1.1.8* that the earliest contacts of the Balts and Slavs were with speakers of Old Novgorod-Pskov. Similarly, the dialectal distribution of R dial. *борка́н* ‘carrot’ could favour this interpretation, but its analysis as a Baltic loanword is uncertain. The latter is clearly a trade item, while the term ‘bath-house’ rather seems to be linked to a certain cultural diffusion. This might theoretically be an indication of a Baltic substrate in Old Novgorod.

The notion of a Baltic substrate reminds us of the evidence collected in a number of studies, chiefly by Топоров (Топоров/Трубачев 1962; Топоров 1972, 1988–1997), but going back to Būga (1923a) and Vasmer (1932), purporting to demonstrate a Baltic substrate in the hydronyms of the Upper Dnieper and Oka basins. The validity of this evidence has practically been taken for granted, and has remained absolutely central to discussions of the Baltic homeland (Zinkevičius 1984: 147–151; Gelumbeckaitė 2018: 1712; see also Grünthal 2012: 299–300 with lit.), soon also entering into archaeological discussions (Gimbutas 1963: 97; Rimantienė 1992: 137; Anthony 2007: 380).

It is beyond the scope of this dissertation to go through the evidence in any detail. However, Stang’s call for “tiefer gehende Sichtung und Diskussion” (1966: 2, fn.) seems to have largely remained unanswered, with later contributions rather looking to expand than critically assess the established material (cf. В. Васильев 2015 for a discussion of some of the issues).[69] In any case, the alleged pervasiveness of a Baltic

[69] Much of the evidence constitutes root etymologies, and these often permit alternative interpretations (see, for instance, the extended discussion of hydronyms of the type *Велья* in В. Васильев 2012: 545–550). The material is

substrate in the hydronymy of this area contrasts starkly with the almost complete absence of evidence of early substratal loans on a lexical level.[70]

Finally, contrary to the claim of Аникин (2014: 192), there is no reason on the basis of this data to assume any loanwords from Baltic into Proto-Slavic. It is possible that such unidentified loanword layers do exist; however, as in the case of the word for 'birch tar', there are few if any phonetic criteria that would allow us to distinguish Baltic loanwords in Proto-Slavic from inherited cognates. At the current stage of research, it can be said that no entirely convincing cases exist.

in need of a thorough critical review, and the results can certainly not be considered "hard facts" in the manner in which they are normally treated in the archaeological research.

[70] There is plenty of evidence for a later Baltic substrate in Belarus, Smolensk and the surrounding regions, as clearly visible in the material collected by Лаучюте (1982).

1.2. Early Germanic → Baltic loans

There have been few focused studies attempting to isolate the earliest layer of Germanic loanwords in Baltic, and we must largely be content with the collections of Hirt (1898: 349–351) and Būga (1922: 64–65; Senn 1925: 46–53 and Alminauskis 1934: 19–22 to a great extent repeat Būga's conclusions), as well as the later sceptical account of Otrębski (1966) and other comments found scattered in works of a more general character. This state of the field means that little has been done in terms of critically analysing and stratifying the material as a whole. The goal here is to present all of the (convincing) evidence for old Germanic loanwords in East Baltic, with a focus on the evidence for direct contacts with Gothic.

Būga (1922: 65) divided his material into two groups, the first supposedly deriving from Gothic and the second from North Germanic:

(a)

- Lt. *alùs*, Lv. *alus* 'beer' ← Go. **aluþ* (cf. OE *ealu*, GEN.SG. *ealoþ* 'ale', ON *ǫl* 'beer')
- Lv. *klàips* 'bread' ← Go. *hlaifs* 'bread'
- Lt. *midùs*, dial. *mìdus* 'mead' ← Go. **midus* (cf. OE *medu*, ON *mjǫðr* 'mead')
- Lt. *pẽkus* 'cattle, livestock' ← Go. *faihu* 'property, wealth'
- Lt. *šárvas*, PL. *šarvaĩ* 'armour; arms' ← Go. *sarwa* N.PL. 'weapons, armour'

(b)

- Lt. dial. *gãtvė* 'cattle way' (cf. Otrębski 1966: 63), Lv. *gatve*, *gatva* 'path between fences, cattle way; street' ← ON *gata* 'passage, street'

- Lt. *kviečiaĩ* (ACC. *kviečiùs*, dial. *kvíečius*), Lv. *kvìeši* ‘wheat’ ← ON *hveiti* ‘wheat’
- Lt. dial. *kliẽpas* ‘loaf’ ← ON *hleifr* ‘loaf’

Būga does not specify why he prefers to derive Lt. *gãtvė* from Norse. At first sight, Gothic *gatwo* ‘street’ appears phonologically more suitable (cf. ME I: 609; LEW 139). Senn (1925: 49), who follows Būga, observes that the word’s limitation to northwest Žemaitia and Curonia would favour a Nordic origin, but this is hardly decisive. Noting the Latvian variant *gate* ‘path between fences’, Zubatý (1892: 255) prefers to take the whole family from Low German (cf. ME I: 609, s.v. *gate*; Smoczyński 2018: 318–319), which seems possible; cf. Prussian German *Gatt* ‘opening; narrow passage’ (Frischbier I: 219). On the other hand, *gãtvė* is not easily analysed as an inner-Baltic derivative: the suffix *-vė* is rare and unproductive (Skardžius 1941: 379). Otrębski (1966: 63) takes Latvian *gatuve* as the original form, and assumes the more common forms with *-tv-* arose through syncope, but it is possible that *gatuve* is itself secondary (after the suffix *-tuve*; Endzelīns 1923: 280–282), and the coincidence with Go. *gatwo* ‘street’ is striking. For this variant at least, I believe an early Germanic origin should be preferred.

Fraenkel (LEW 271, 326) follows Būga in deriving Lt. dial. *kliẽpas* and *kviečiaĩ* from North Germanic (see also ME II: 356). This is motivated by the idea that Baltic **ę̄* must derive from **ei* (Būga RR III: 900–901; Senn 1925: 49–50; Alminauskis 1934: 21; on this, see also Endzelīns 1907). However, these authors’ denial that **ę̄* may derive from **ai* lacked a solid basis, as Stang (1966: 53–57) convincingly showed, and their formulation is now largely obsolete. As a result, there is no particular reason to posit a Norse origin for any of the loanwords.[71]

[71] Balaišis (1994) still maintains Būga’s view, but prefers to take the words in question from Gothic anyway. To do this, he is not only forced to assume a

Senn (1925: 50), who accepts Būga's reasoning, nevertheless presents a counter-argument: as Norse ⟨f⟩, outside of initial position, stood for the sound /ƀ/ (Noreen 1923: 40), we should expect ON *hleifr* to have given Lt. **kliẽbas*. To explain the *-p-*, Senn is forced to assume an *ad hoc* contamination with Lt. *kẽpalas* 'loaf'.[72]

Starting from Gothic, we can take the NOM.SG. *hlaifs* or ACC.SG. *hlaif* 'bread', with final devoicing, as the specific source. We may account for the vocalism of Lt. dial. *kliẽpas* 'loaf' and *kviečiaĩ* 'wheat' in two ways: (a) assume that in the Gothic dialect from which the word was taken, *ai* had monophthongized to */ę̄/, as probably in Wulfila's dialect (cf. Wrede 1891: 165; Bennett 1949), and that this monophthong was adopted directly as East Baltic **ę̄*; (b) assume that a preserved diphthong */ai/ was adopted directly as the Baltic diphthong **ai*, which only later developed to **ę̄*.

Potentially relevant for resolving this matter are two etymologies presented by Vasmer (1922) supposedly pointing to a Gothic source: Lt. *ýla*, Lv. *īlęns*, Pr. E *ylo* 'awl' ← Go. **ēla* and the Latvian hapax *glīsis* (ME I: 627)[73] 'amber' ← Go. **glēza-*. This would imply a narrowing of **ē*

Gothic sound shift */ai/ > */ei/ in order to derive the relevant words from Gothic, but also an *ad hoc* change back from **ei* > *ai* to explain Lv. *klàips* (idem: 11).

[72] To be precise, Senn actually derives **kliẽbas* from Slavic (cf. MBel. *хлѣбъ*) rather than Norse (the same explanation is taken up in Smoczyński 2018: 566). Berneker (I: 389) also derived the word from Slavic, but explained the *-p-* as the result of generalization from the NOM-ACC.SG. [xlěp]. However, this kind of phenomenon is unparalleled in the Slavic loanwords; see already Būga (1912: 31). Differently again (and implausibly), cf. Otrębski 1966: 53.

[73] I do not have access to the *Magazin der Lettisch-literärischen Gesellschaft* 20/3, cited by ME, but see Kregždys (2012: 330, fn. 470; and also idem: 330–336 for an attempted etymology). In Lange's dictionary (1773: 125), we find the following entry: "Glihſe (obſoletum) Börnſtein, Edelſtein[,] die Nordiſche Seefahrer nanten daher den Börnſtein Strandt, Glyswall." It seems unclear

towards */ī/ which can be witnessed in occasional spellings of ⟨ei⟩ in place of etymological **ē* and vice versa in the Gothic bible (suggesting they fell together in the language of later scribes; Marchand 1973: 51). Since this raising might well have been triggered by the monophthongization of **ai*, we might take these etymologies as indirect support for option (a).

However, it turns out that both of the reconstructed Gothic forms are problematic. While most sources seem to take the length of the first vowel in OHG *ala* 'awl' for granted (e.g. DWb² II: 73; Kluge/Götze 7; EWAhd I: 135 s.v. *âla*), the evidence of the other Germanic languages points to short *ă* (see Kluge/Seebold 22).[74] Similarly, for 'amber', OE *glær* 'amber' (cf. DOE s.v. *glæ̆r*) and MLG *glar* 'resin', *glar(r)en** 'smear with resin' (Schiller/Lübben II: 116), traditionally considered to contain a long vowel, are ambiguous and may just as well reflect short *ă*, as we do find in ON *gler* 'glass' and the Verner variant OE *glæs* (cf. ME *glas*) 'glass', OHG *glas* 'glass, amber' (see Meineke 1998: 141 with

whether the form **glīse* was merely inferred by Lange on the basis of the cited toponym.

[74] So, as universally agreed, ON *alr* 'awl'. OE *æl* is given a short vowel in DOE s.v., which is supported by Middle English *al* (the ME form *ēl* cited by e.g. EWAhd loc. cit. is a figment, the examples with ⟨e⟩ being Kentish or West Mercian for **æ̆*). Kluge/Seebold (loc. cit.) argue that High German might reflect **ă*, too. In support of this, we can note that DWb² (II: 73) cite a form *allen* DAT.SG. from Peter von Ulm's *Cirurgia* (c. 1430), which looks (at first sight at least) to point to a short vowel, and also the form *ale* 'awl' in the Elbing Vocabulary, where reflexes of MHG *ā* are regularly spelled ⟨o⟩ or ⟨oe⟩ (Braune 1876: 93–94; Trautmann 1910: XXV). A detailed study of the German dialectal evidence is obviously not possible here, but it is naturally far more straightforward to derive all the Germanic forms from a single ablaut grade than to set up a rare **ē/a* ablaut alternation. As Kluge/Seebold points out, Skt. *ā́rā-* 'awl' may just as well reflect IE **Hol-*, so there is no external evidence for Pokorny's **ēlā* (IEW 310). If we reject this variation, the only way to connect Lt. *ýla* (etc.), it seems, would be to assume a reduplicated **Hi-Hl-*.

lit.). Evidence for a Germanic variant with *-ē-* seems to be limited to Pliny's *glaesum* (note here the *varia lectio* ⟨glassū⟩). As a result, Vasmer's Gothic reconstructions are based on very uncertain evidence.

A similar assumption underlies Endzelīns' (ME IV: 277) suggestion that Lt. *tū́bas*, Lv. *tūba*, Pr. E *tubo* 'felt' derive from an equivalent of ON *þófi* 'felt' in a Gothic variety in which **ō* had become raised to **ū* (cf. Marchand 1973: 52). Previously, Trautmann (1910: 451) and Būga (1922: 294–295) had taken the word directly from Norse, but this fails to explain the vocalism. An alternative account would be to assume the word entered East Baltic through Prussian. As Prussian had no **ō* (cf. Smoczyński 2000: 66–70), it seems that a Gothic **þōb-* (or indeed ON *þóf-*) would most probably have been adopted here as **tūb-* directly.[75]

Thus, we return to option (b), namely that the Gothic diphthong */ai/ was adopted directly as a diphthong in Baltic. If this is the case, there is no reason to separate Lt. dial. *kliẽpas* 'loaf' from Lv. *klàips* 'bread'. Many doublets with and without monophthongization can be found, suggesting paradigmatic alternations may still have been present in Proto-East-Baltic: e.g. Lt. *eĩti* ~ Lv. *iêt* 'to go'; Lt. *žíedas*, dial. *žáidas* 'flower' (footnote 164); Lt. *saĩkas*, Lv. *sìeks* 'a dry measure', etc. (Hirt 1892: 37–40; Kuryłowicz 1956b: 234; Petit 2003: 97).

Another example of such an alternation is Pr. E *caymis*, Lt. *káimas* 'village' (and the derivative Lt. *kaimýnas* 'neighbour')[76] beside

[75] By contrast, the borrowing *broakay* 'breeches' ← MLG *brōk* or MHG *bruoch* (cf. Trautmann 1910: 314; PKEŽ I: 158) must have postdated the Pomesanian Prussian development **ā* > /ō/.

[76] Although the word belongs to the standard language, it is interesting that the LKŽ only cites Lt. *káimas* and *káima* from Žemaitia and Suvalkia. In view of this, we might suggest this is a borrowing from Prussian, which might potentially explain the acute accentuation, cf. Pr. III *kāimaluke* 'heimsucht'. On

Lt. *kiẽmas*, Lv. *cìems* ‘farmstead, village’. This word has often been understood as inherited and compared either with Lat. *quiēs* ‘rest, quiet’ (Uhlenbeck 1900: 69; Būga 1922: 70; LEW 251; Smoczyński 2018: 540) or Gr. κώμη ‘village’ (Zupitza 1896: 49; Trautmann 1910: 112). Both of these explanations encounter phonological issues, and far more attractive is the interpretation as a loanword from Go. *haims* ‘village’ (Hirt 1898: 347–348; Boisacq 1916: 544; ME I: 394; Derksen 1996: 215, 2015: 243–244; ALEW 565). This is favoured by the semantic closeness to the Gothic word; contrast the inherited cognate found in Lt. *šeimà*, Lv. *sàime* ‘family, household’.

Several forms involving the phoneme *-k-* have elsewhere been considered cognates, but the *centum* reflexes, as well as the close semantic and formal correspondence with Germanic, favour a loan origin:[77]

- Lt. *kiẽmas*, Lv. *cìems* ‘village, farmstead’ ← Go. *haims* ‘village’ — Contrast Lt. *šeimà* ‘family’ (and probably Skt. *śéva-* ‘dear’; see fn. 99)
- Lt. *kviečiaĩ*, Lv. *kvìeši* ‘wheat’ ← Go. *ƕaiteis** ‘wheat’ — Contrast, if related, Skt. *śvítna-* ‘white’[78]

---

the other hand, the derivative Lt. *kaimýnas* ‘neighbour’ does not show this dialectal limitation, and is probably a genuinely East Baltic word.

[77] Another possible example could be Lv. *kàuns* ‘disgrace, shame’, perhaps loaned from Go. *hauns* ‘lowly’ (Hirt 1898: 350; or rather from a corresponding noun, cf. MHG *hōn* ‘disgrace, shame’, MDu. *hoon* ‘humiliation’ < **hauna-*), rather than cognate in view of the absence of correspondences elsewhere in IE (cf. Stang 1972: 27).

[78] The exact correspondence with Germanic makes the assumption of a parallel formation based on Lv. *kvitêt* ‘flicker, glimmer’ (Otrębski 1966: 54; Sabaliauskas 1990: 41; ALEW 546–547; Pronk/Pronk-Tiethoff 2018: 302–303) entirely gratuitous.

- Lt. *pẽkus* ‘cattle, livestock’; Pr. I–III *pecku* ‘cattle’ ← Go. *faihu* /fęxu/ ‘property, wealth’ — Contrast Skt. *paśú-* ‘cattle’[79]
- ? Lt. *keřdžius* (secondary *skeřdžius*) ‘herdsman’ ← Go. *hairdeis* (NOM.PL. *hairdjos*) ‘shepherd’ (Hirt 1898: 332)

The word for ‘poppy’, Lv. *maguône*, Lt. *aguonà* is normally considered to have been loaned from Germanic (ME II: 547; Sehwers 1936: 312; Sabaliauskas 1960a: 72; Smoczyński 2018: 6). As the short first-syllable vowel makes a late German origin improbable, Endzelīns (ME loc. cit.) suggests an Old Saxon *mago** (attested in the compound *magonhouut* · papaver) as a proxy. However, given the absence of other evidence for Old Saxon loans in East Baltic, we might instead suggest a Gothic **mago*. The main issue with this explanation is that the earliest evidence for the opium poppy in the East Baltic region dates to the Middle Ages (Grikpėdis/Motuzaitė Matezuvičiūtė 2020: 167), which is too late to be reasonably associated with Gothic contacts. As a result, no fully satisfactory explanation is available for this word, although it is certainly borrowed. See p. 397–398 for further etymological discussion.

By contrast, I see no reason to assume that Lt. *alùs* ‘beer’ (etc.) is a Germanic loanword (cf. Machek 1930; Kiparsky 1934: 78–79; Stang 1972: 13; Топоров ПЯ I: 80; Mallory/Adams 1997: 60), despite frequent claims to the contrary (Hirt 1898: 346; Būga 1922: 64;

[79] Fraenkel (LEW 564–565) rejects this loan etymology due to the semantic distance. Yet since there does not appear to be any other word for ‘cattle’ attested in Gothic, it may well have been *faihu*; cf. the similar semantic range of the ON cognate *fé* ‘cattle; property, wealth’. Kortlandt (1978: 241) has attempted to explain the Baltic *-k-* as having spread from an oblique stem comparable to Skt. *paśvás* GEN.SG. with his rule *ḱ > *k* before *u̯* + back vowel. The back-vowel criterion for this rule cannot be fulfilled, however, as only *-*es* can be reconstructed as an athematic genitive ending in Balto-Slavic.

Kroonen 2013: 23; Derksen 2015: 53; ALEW 36; Smoczyński 2018: 22). Note that the word is not limited to northern Europe if Arm. *awłi* 'strong fermented drink' belongs here (Olsen 1999: 443; Martirosyan 2008: 155). Against a loan from Gothic speaks the absence of any trace of the stem-final dental, which ought to have been preserved there; cf. *miliþ* 'honey' (< **melit*-). In an inherited context, the loss of the final -*t* can be accounted for by regular sound law (cf. e.g. Kortlandt 1989a: 44).

According to Būga (1922: 66), all Gothic words entered East Baltic through Prussian, an opinion repeated by Senn (1925: 48: "weder Letten noch Kuren noch Litauer [hatten] an irgendeiner Stelle direkten Verkehr mit den Goten"; cf. also Senn 1943: 954). However, such conclusions must derive from the data, rather than from aprioristic assumptions. On phonological grounds, I have argued above that Lt. *šárvas* 'armour' (p. 42) and *tū́bas* 'felt' (p. 83) may well have been borrowed through Prussian. I have also mentioned some words which ultimately derive from Germanic but which may have equally have entered East Baltic indirectly, either through Slavic or Prussian. Due to the ambiguity, they cannot be used as evidence here:

- Lv. *bruņas* 'armour' = Pr. E *brunyos*, OR бръня 'armour', Go. *brunjo* 'breastplate'
- Lt. *kãtilas*, Lv. *katls* 'kettle' = Pr. E *catils*, OR котьлъ, Go. *katils** 'kettle'
- Lt. *stìklas*, Lv. *stikls* 'glass' = Pr. E *sticlo*, OR стькло 'glass', Go. *stikls* 'cup, chalice'

A specifically Gothic source must be assumed at least for Lt. *midùs* 'mead', where -*i*- for expected **e* can only reasonably be explained through the assumption of Gothic transmission (Hirt 1898: 346; Būga 1922: 65; unconvincing is Otrębski 1966: 55). Note also that German

mete 'mead' is glossed *alu* in the Elbing Vocabulary, which does not exclude the possibility of Prussian mediation (there may have been dialectal differences which are not reflected in the attested evidence), but it certainly does not favour it. Similarly, the attested Prussian words for 'wheat' (Pr. E *gaydis*, G *gaide*, *gayde*) and 'bread' (Pr. E *geytye*, III *geits*) differ from those attested in East Baltic and do not represent Gothic loans.

Another alleged piece of evidence for direct contact with the Goths is Lt. *gùdas* 'Belarusian; speaker of a different dialect', Lv. *guds* 'Belarusian raftsman; wandering merchant' (ME I: 675), which has been taken from Go. *guta** 'Goth' (Būga 1922: 67; LEW 174; Smoczyński 2018: 400; on the Gothic endonym, see Leumann 1986: 163–164 wih refs.), under the assumption that the word was used to refer to Slavs under Gothic rule. The medial *-d-* has been explained by assuming a pre-sound-shift loan from Germanic (Būga loc. cit.; Zinkevičius 1985: 73), which is hardly plausible, although the only other possibility is to assume an *ad hoc* contamination (cf. Karaliūnas 2004: 164).[80]

Karaliūnas (2004: 145–189) hypothesizes a native origin for Lt. *gùdas*. Noting the word's pejorative value in folk literature, he suggests a derivation from a root **gud-* 'small, of poor quality', which is set up on the basis of e.g. *(į-)gùsti* 'get used to', *gùd-obelė* 'hawthorn; crab apple' (*obelìs* 'apple tree'), Pr. E *gudde* 'bush' (with which LEW 174 already suspected contamination). Despite the detailed treatment, I am not convinced that the evidence, mainly plant names, warrants the

[80] One would like to see the missing link in the gloss *guti* · krzyrzacy 'Teutons' in the Narev vocabulary (although the reliability of this vocabulary remains uncertain). On this word differently, see Karaliūnas 2004: 164–165.

reconstruction of such a root.[81] However, while the etymological connection with the name of the Goths is tantalizing,[82] the *-d-* remains a significant stumbling block.

Given the general assumption that the East Baltic loanwords from Gothic were mediated by West Baltic, it is remarkable that all of the certain Gothic loanwords in Prussian are shared with East Baltic. Two others have been suggested as unique to Prussian (Būga 1922: 66, Senn 1925: 47), but neither of these are certain:

– ? Pr. E *ilmis* · bark 'hay shelter' ← Go. *hilms* 'helmet' (cf. ON *hjalmr* 'helmet; hay shelter'; Lidén 1906; Trautmann 1910: 346; Тороров ПЯ III: 42; PKEŽ II: 24). Yet in view of **h*- → *Ø*-, a Low German source seems more probable (Smoczyński 2000: 35–36; admittedly, a formally or semantically suitable source appears to be lacking).[83]

---

[81] Lt. *(į-)gùsti* 'get used to' may well be backformed from *gùd(r)inti* 'train, teach', which belongs with *gudrùs* 'smart, sly' (on this word familiy, see also ALEW 439; Smoczynski 2018: 401). Combined with *gùdė* 'whetstone', one might imagine an original meaning 'sharp' (although this is by no means the only option). As to *gùd-obelė*, etc. I would rather assume the first element means 'bush', as in Prussian. The dial. *gū́das* 'sad, gloomy', in view of its acute root, must also be separated.

[82] Bearing in mind the various parallels adduced by Karaliūnas (2004: 162) whereby terms for other peoples have been generalized in the meaning 'unchristened child': R dial. *лопь* 'Sámi; unchristened infant' (cf. Мызников 2019: 450–451), Lv. *krìevs* 'Russian', *krieviņš* 'unchristened child' (ME II: 284–285), it may be conceivable that the pre-Christian Balts, on the contrary, used the term **guda*- as a pejorative designation for their 'non-pagan' (i.e. Christianized) neighbours (cf. the juxtaposition of the *gudai* and the pagan deity *Perkū́nas* in folklore; Karaliūnas 2004: 159–160).

[83] The vocalism of Pr. E *kelmis* 'hat' (Sabaliauskas 1990: 257) shows that it cannot be from Gothic *hilms*. I will not enter into a discussion of this word here.

- ? Pr. E *lapinis* ‘spoon’ ← Go. **lapins* (cf. OHG *leffil*, MDu. *lepel* ‘spoon’; Kluge 1907: 361; Trautmann 1910: 368; Endzelīns 1943: 202; Sabaliauskas 1990: 257). While this etymology still seems possible, it is now widely rejected in favour of a native etymology (Falk *apud* Топоров ПЯ V: 90; PKEŽ III: 41–44).

Collecting together the evidence for direct Gothic loans into East and West Baltic, we obtain a rather interesting picture (underlined words are those shared by Prussian; those in brackets were possibly obtained indirectly):

- Agriculture: village, wheat, bread, ?poppy
- Stockbreeding: cattle, cattle way, ?herdsman
- Warfare: [armour (×2)]
- Trade: [felt], [kettle], [glass]
- Other: mead

It is remarkable that the majority of the direct loanwords are connected in some way to stockbreeding and agriculture (note also p. 310 for a discussion of the possibility that Lt. *rugiaĩ* ‘rye’ is a Gothic loanword). Borrowings in this semantic field seem more likely to be indicative of an inward migration of Germanic speakers rather than incidental trade. A possible proxy for this migration could be found in the appearance of grave artefacts in the 5th century in Eastern Lithuania showing a remarkable similarity to those popular in the Carpathian Basin. Even though these artefacts may rather attest to trade routes (Bliujienė/Curta 2011), this does not rule out a small-scale migration. Importantly, the loanwords do not indicate the assimilation of an elite class. It is remarkable that none of the words associated with trade can be considered unambiguous direct loanwords.

A second possible route for the incursion of Gothic-speaking populations could be a direct migration from the Lower Vistula region through Sembia and Žemaitia of "polyethnic warrior groups", bringing with them new kinds of weapons as well as new burial customs (cf. Kurila 2021: 21). This migration could explain certain originally Gothic words shared between East and West Baltic. Note particularly Lt. *šárvas* 'armour', which I have argued was most probably borrowed through Prussian.

1.3. Baltic → Finnic borrowings

1.3.1. Preliminaries

In this chapter, I have the following aims. The first task is to characterize the extent, chronology and nature of the contact relationship between Baltic and Finnic-speaking groups on the basis of the mutual loanwords. The second is to attempt to answer the question of whether some of the vocabulary shared between the two language families may in fact have originated in other, pre-Baltic languages spoken in the region before the arrival of the Baltic and Finns. Before doing so, it is important to define the corpus of loanwords I will use for my analysis.

In his 1890 magnum opus, Thomsen identified some 200 potential loans from Baltic into Finnic, of which he considered about 140 certain. To a large extent, Thomsen's work has stood the test of time, and there are comparatively few really solid etymologies that have been proposed since. Despite a rather impressive amount of research into the subject of Baltic-Finnic loans, Petri Kallio (2008a: 265), 118 years later, still states that only "about 200" certain Baltic loanwords can be found in Proto-Finnic. This is more or less in line with Vaba (1990a), who labels 189 loan etymologies as certain.

At the same time, Santeri Junttila's dissertation (2016a) covers a corpus of almost 1000 etymologies proposed up until 2009, which implies nearly 7 new loan proposals every year since Thomsen. It would be beyond the scope of this dissertation to discuss all the proposals, which would be a task of many years (Junttila in prep.). I have therefore limited myself to those which I have deemed reliable,

starting with those of Thomsen (1890) and Kalima (1936). The material presented here is certainly incomplete, but hopefully sufficiently representative to allow for valid conclusions to be drawn.

Insofar as the study of loan relationships cannot operate with strict sound laws in the Neogrammarian sense, the study of Baltic-Finnic loan relations has suffered from many of the same issues as long-range and pseudo-linguistic comparison. The first issue concerns semantics. To quote Robbeets (2004: 158): "The greater the semantic latitude permitted in external comparisons, the more likely it becomes that the apparent formal similarity is due to pure coincidence." This criticism can of course apply even to comparisons within the Neogrammarian framework, but without the constraint of exceptionless sound laws,[84] a low threshold for semantic similarity essentially leaves the scholar's own imagination as the only limiting factor (cf. Rédei 2000).

In the absence of any empirical approach to semantic shifts, such shifts ought to be approached with extreme caution. However, a cursive look at the state of the art in Baltic-Finnic loanword studies reveals that semantics have not been a primary consideration. Some representative cases of semantically questionable etymologies are given below. To avoid any risk of cherry-picking, I have limited myself to those assessed as "Relatively clear" in Junttila (2012):[85]

---

[84] Santeri Junttila (p.c. April 2023) has argued that substitution rules can be treated similarly to sound laws, and I would indeed recommend a strict approach. However, in practice, multiple substitutions for a single sound can and have been assumed. Even if we attribute such variation to different chronological stages or dialectal differences in the source language, this is rarely independently verifiable, with the result being that 'substitution laws' are more flexible than traditional sound laws.

[85] The choice of this article is merely dictated by convenience, and I do not mean to single out Junttila as a particular offender in this domain. Rather, the

- F *hiiva* ‘yeast; sediment; froth (on beer)’, Vt. *iiva* ‘yeast, leaven’ (< **hīva*); E dial. (Saaremaa) *iive* ‘froth’ ~ Lt. *šývas* ‘grey, whitish (usu. of horses)’ (Koivulehto *apud* Plöger 1982: 93; cf. Häkkinen 2004: 196; van Linde 2007: 35–37).[86]
- F *huone*, Võ. *hoonõh* (< **hōnẹ̄h*) ‘building; room’ ~ Lt. *šónas*, Lv. *sãns* ‘side’ (Koivulehto 1992b, supposedly in the sense ‘Nebenraum, Nebenhaus’; cf. the doubts in Häkkinen 2004: 221–222).
- F *kausta*, E *kaust*, Li. *kōsta* ‘side beam on a sledge’ (< **kausta*) ~ Lv. *skàusts* ‘withers; nape (of a person or animal)’ (Posti 1977: 264–265; SSA I: 333).
- F *ketara* ‘sledge stanchion’, E *kodar*, Li. *kõ’ddõrz* ‘sledge stanchion; spoke (of a wheel)’ (< **kę̣tara*) ~ Lt. *keterà*, *sketerà* ‘ridge, peak; crest (of the back)’ (Būga 1908: 72; Posti 1977: 265–266).[87]
- F *sakara* ‘point, protruding tip’, E *sagar* ‘(wooden) hinge’ (< **sakara*) ~ Lt. *stãgaras* ‘(dry) stalk, branch’ (Kalima 1936: 203 with “?”; SSA III: 144).
- F *sampi* (< **sampi*); E *samb*, GEN.SG. *samma* (Setälä 1902: 149–150; not in VMS — a Finnish loan?) ‘sturgeon’ ~ Lt. *stambùs* ‘big, beefy; coarse-grained’ (Liukkonen 1999: 124).[88]

---

lax approach to semantics exhibited by this article is characteristic of the subfield in general. Note that Junttila (in prep.) now rejects Koivulehto’s Baltic etymology for **hīva* and doubts the one for **hōnẹ̄h*.

[86] This suggestion “rescues” Thomsen’s (1890: 218) unsuccessful comparison with Lt. *sývas* (often pl. *sývai*) ‘sap’. The mention of *šỹvas* ‘Hausbier’, a hapax recorded in a daina (*šyvū* ACC.SG. ‘kvass’, Niemi/Sabaliauskas *apud* LEW 996), is hardly sufficient to make the comparison “relatively clear”. Note that “*šývis* ‘mould’”, cited by van Linde (2007: 35), is the result of his misunderstanding of German *Schimmel* ‘grey horse’ (cf. Nesselmann 1851: 520).

[87] The further comparison with Md. E *kodorks* ‘twining plant stem’, M *kodərks* ‘vegetable tops’ (cf. SSA I: 351; Grünthal 2012: 317) is also semantically implausible.

– F *uksi*, E *uks*, Li. *ukš* (< **uksi*) ‘door’ ~ Lt. *úoksas* ‘hollow; cavity’ (Koivulehto 1993a: 34; SSA III: 369 with “?”).[89]

The theoretical possibility of a semantic shift should not be considered sufficient grounds for an etymology to be accepted, as we always have to reckon with the possibility of chance resemblance. In the case of **hīva*, for instance, even if it is accepted that ‘yeast’ could plausibly be referred to as ‘grey matter’ (van Linde 2007: 35–37), the etymology cannot be considered by any means certain as there is no trace of a sense ‘yeast’ in Baltic, nor of a sense ‘grey’ in Finnic, and thus the proposal that a semantic shift occurred depends itself on the correctness of the borrowing proposal.

Semantic parallels must be specific and trivial: Posti (1977, see above) simultaneously presented two etymologies connecting Baltic words for body parts with Finnic words for parts of the sledge. But if Finnic **jalkas* ‘runner’ is derived from **jalka* ‘leg’, then surely the side beam would be the ‘arm’ and by no means the ‘withers’ or ‘neck’. Even if some of the above etymologies are actually correct, the sheer semantic distance means that they can never be regarded as “relatively clear”.

---

[88] Liukkonen points out that the word has in Finnish also been applied to other large fish, and attested in the general meanings ‘big fish; fish god’, but these are clearly metaphorical extensions of ‘sturgeon’.

[89] Koivulehto defends his comparison of the Finnic word for ‘door’ with the Baltic word for ‘tree hollow’ by referring to Lat. *ōstium* ‘door’, an indirect cognate of the Baltic word. But the Latin word represents a different formation, such a meaning is unknown in Baltic, and a development ‘tree hollow’ ≫ ‘door’ is itself almost inconceivable. Not to mention that, providing the comparison with Latin is correct, the Baltic *-k-* must be secondary (LEW 1165–1166). Its recent nature would apparently be supported by the rare form *úosvauti* (Jablonskis *apud* LKŽ) ‘search for a tree hollow (of scout bees)’, and the absence of the RUKI law. The root meaning is ‘mouth’ (IEW 821, without the Baltic word).

The possibility of chance resemblance between Baltic and Finnic words may also be increased by the simpler phonotactics of Finnic as compared to Baltic. Just focusing on word-initial position, Finnic **k-* can equally stand for Baltic **k-*, **g-* or **sk-* which results in a potentially significant increase in the 'hit rate' when searching for Baltic donors. Finnic **r-* would regularly substitute 11 phonotactically acceptable Proto-Baltic anlauts (**r-*, **sr-*, **pr-*, **br-*, **spr-*, **tr-*, **dr-*, **str-*, **kr-*, **gr-*, **skr-*).

As a brief illustration of the possibility of chance resemblance, I searched through the LKŽ for potential Baltic comparanda for Finnic words with cognates I was able to verify in Samoyed (some 140 items). These Finnic words can obviously not be classified as Baltic loanwords, although a few fairly good matches can be found. One such example has in fact been treated as a possible Baltic loanword in certain sources: F *lampi* 'pond' was hesitantly compared with Lt. dial. *klampà* 'swamp, muddy spot' by Kilian (1986: 494), and independently by Liukkonen (1999: 78–79). Junttila (2012: 282; 2016a: 103) categorizes this example as "dubious" rather than "erroneous". On the other hand, the word may equally be regularly cognate with Ngan. *l'üŋhə* 'boggy spot' (UEW 235; Aikio 2014c: 86). Since only one of the two etymologies may be correct, this a clear illustration of the risk of false positives in loanword research.

Allowing a rather small level of semantic flexibility, we can make a number of additional false comparisons, for example:

- PF **kaiho* 'grief, loss; yearning' (= Ngan. *koče* 'illness', Aikio 2014b: 3–5) ~ Lt. *gaižùs*, Lv. dial. *gàizs*[2] (EH I: 379) 'bitter, acidic' — cf. the derived verb Lt. *giẽžti* 'feel an unpleasant sensation (in the throat)', also 'long for, request insistently'
- PF **kaiva-* 'dig' (cf. Ngan. *kajbu* 'shovel') ~ Lt. *nu-káivinti* 'wear out (the soil); exhaust'

- F dial. *kumpu-* ‘well up’[90] (= Ngan. *koŋhu* ‘wave’) ~ Lt. *gumbas*, Lv. dial. *gumba* ‘bump, bulge, excrescence’
- PF **lanci* ‘damp lowland’ (= Ngan. *lį̄ntə* ‘plain, valley’) ~ Lt. *sklandùs* ‘slippery, smooth, flat’ — The semantic connection would be ‘flat’ : ‘flat land’
- PF **mene-* ‘go’ (= Ngan. *mį̄nsį̄*) ~ Lt. *mìnti* (PRES. *mẽna*), Lv. *mīt* ‘trample’ — the Lt. word is also attested in the meaning ‘go, tread’
- PF **nüci* ‘scythe handle’ (= Ngan. *ńir* ‘axe handle’) ~ Lt. dial. (Juška) *gniutìs*, Lv. dial. *gņuta* ‘thin plank used to attach straw to a roof’

In drawing up a corpus of etymologies upon which further conclusions can be based, only the clearest cases should be used. To this end, I have excluded all etymologies which involve speculative or non-trivial semantic shifts. That is not to say that I deem these etymologies impossible, but simply that it would be misguided to base any further conclusions on them. Their acceptance should rather be informed by the analysis of the clearer cases.

However, even etymologies which show perfect semantics cannot necessarily be regarded as certain loanwords from Baltic to Finnic. As an example, Lt. *tóšis*, Lv. *tãss* and F *tuohi*, E *toht*, Li. *tū’oigõz* (< **tōhi*) all mean ‘birchbark’; however, since neither the Baltic nor the Finnic word has a clear etymology,[91] it is difficult to make any conclusion with

[90] The Finnic word shows an irregular vowel, but I wonder if it could be explained by an (irregular) assimilation **o–u* > **u–u*, parallel to the recently proposed sound law **e–ü* > **ü–ü* (Aikio 2021: 171). Note that a generally high frequency of stems of the shape **u–u* in Finnic was already observed by E. Itkonen (1948: 133).

[91] See Smoczyński (2018: 1498). Lv. dial. (Kurzeme) *tâst*[2] ‘hew’ probably shows secondary length after the preterite (ME IV: 151), cf. the derived dial.

regard to directionality (Kalima 1936: 171; LEW 1107; cf. Bednarczuk 1976: 54). Although Junttila has argued that there are no Finnic loanwords in Proto-Baltic (see *1.3.4*), he concludes that a few cases could represent "parallel borrowings from a shared source, perhaps a lost substrate language" (2015: 31).[92]

In view of this possibility, only when the Baltic source word has a solid Indo-European pedigree can a specifically Baltic → Finnic directionality be proven. As a result, I think it worthwhile to limit my corpus of loanwords to those which have an Indo-European etymology. Ideally, this means regular, unambiguous cognates beyond the neighbouring Slavic and Germanic, although I have also included examples which require some additional (more or less trivial) assumptions with regard to word formation and semantics.

In a small number of cases, an etymology has been excluded because the derivation from Germanic is equally plausible:

(SW) *tastît* 'hew', and the similar phenomenon in Lv. *tèst*, dial. *têst* (ME IV: 175–176) 'hew, chop, adze', where we indeed find a vacillation between 1SG.PRES. *tèšu* and dial. *tešu*. These are to be equated with Lt. *tašýti* 'carve', which cannot be connected with *tóšis* on formal grounds.

92 In this connection, he cites e.g. F *kinnas*, E *kinnas*, Li. *kīndaz* 'mitten, glove' (?→ Sá. S *gamhtse*, 18th c. ⟨kamtes⟩ 'leather glove') ~ Lv. *cìmds* 'glove' (cf. Thomsen 1890: 187; Kalima 1936: 118; Posti 1953: 36–37; SSA I: 336). However, there are several generally accepted loanwords which could easily have been mentioned in the same context. Take, for instance, F *vuota* 'hide, pelt' ~ Lt. *óda*, Lv. *âda* 'skin; hide, leather' (Thomsen 1890: 205; Kalima 1936: 183). While the direction of borrowing has apparently never been doubted, and Koivulehto (2000: 104) has even explicitly ruled out a substrate word, the Baltic word remains unetymologized (cf. LEW 515–516; Smoczyński 2018: 883; ALEW 826). Due to the ambiguity in the analysis, these and similar cases have been excluded from the dataset.

- F *olut*, Vt. *õlud*, Li. *vó'l* (< **olut*) 'beer' ~ Lt. *alùs*, Lv. *alus* 'beer'; Pr. E *alu* 'mead'; Sln. (dated) *ǫ̂ł* 'beer' = ON *ǫl* 'beer', OE *ealu*, GEN.SG. *ealoþ* 'ale' (LÄGLOS II: 310; Junttila 2012: 273)[93]
- F *rastas*, E *rästas*, Li. *rastā* (< **rasta(s)*; see p. 178) 'thrush' ~ Lt. *strãzdas*, Lv. *strazds* 'thrush, starling' = Ic. *þröstur* 'thrush' (Qvigstad 1893: 259; LÄGLOS III: 130–131)[94]
- F *terva*, E *tõrv*, Li. *tõ̂ra* 'tar' ~ Lt. *dervà* 'tarry log; tar, resin, pitch', Lv. *dar̂va* 'tar, pitch' = ON *tjara*, OE *teoru* 'tar' (LÄGLOS III: 289–290)[95]
- F *vaaja*, E *vai*, Li. *vaigā* (< **vakja*) 'wedge' = Sá. S *vuevjie*, L *vuojvve* 'clothing insert' (< **vuovjē*) ~ Lt. *vãgis* 'peg, wedge', Lv. *vadzis* 'wall hook, wedge' = Nw. dial. *vegg*, OHG *weggi* 'wedge' (LÄGLOS III: 344)
- F *äes*, E *äke*, Li. *ä'ggõz* (< **äkes*) 'harrow' ~ Lt. *akė́čios*, Lv. *ecêšas*; Pr. E *aketes* PL. 'harrow' = OE *egeþe*, OHG *egida* 'harrow' (Koivulehto 1971: 591; LÄGLOS III: 429; Junttila 2012: 273)

The same goes for F *kaima*, E *kaim* 'namesake; relative, companion', Li. *kāima* 'neighbour' (= Sá. N *guoibmi*, Sk. *kuei´mm* 'companion'), for which, rather confusingly, a Germanic loan etymology is almost never

---

[93] While **o* ← **a* might favour a Baltic source, the final **t* seems rather to favour a Germanic one. Against deriving the Balto-Slavic words from Germanic, see p. 85–86.

[94] LÄGLOS favour a Baltic origin due to the final **-as* in Finnic, but the adoption of Germanic **-us* as Finnic **-as* does not seem impossible (see Koivulehto 1981: 193). For further discussion of the Indo-European background, see p. 324–327.

[95] Sá. N *darvi*, Sk. *tâ´rvv* 'tar' (< **tęrvē*) is hardly a loan from Finnic (Aikio 2006b: 32). Instead, it may be a Norse loan. For the vowel substitution, compare the example Sá. N *gavja* '(fine) dust', Sk. *kõbjj* 'dust, dandruff' (< **kępję*) ← Germanic **heuja-* (cf. Ic. *hý* 'down, fluff, dust'; Aikio 2006a: 24).

suggested. Semantically, the Finnic word is no closer to Lt. *kiẽmas*, Lv. *cìems* ‘farmstead, village’ than it is to Go. *haims* ‘village’, ON *heimr* ‘world, realm’ (in place names ‘village’); the existence of a derivative Lt. *kaimýnas*, Pr. III *kaimīnan* ACC.SG. ‘neighbour’, Lv. *kaĩmiņš* (different suffix!) ‘neighbour; resident of the same village’ (cf. Thomsen 1890: 177; Kalima 1936: 105) is hardly of any relevance, since such a derivative cannot have given the Finnic words directly. Perhaps closest to the attested Finnic sense comes the compound MDu. *oom* ‘uncle’, OE *ēam*, OHG *ōheim* ‘maternal uncle’ (< **awa-haima-*).[96]

I have also omitted words limited to Livonian, such as the following:

- Li. *kil* ~ *kiļ* ‘black woodpecker’ and *palāndõks* ‘pigeon’ both predate developments specific to Latvian (viz. the palatalization in Lv. *dziÎna* ‘woodpecker’ and loss of nasal in *baluôdis* ‘pigeon’). However, these loanwords need not date to Proto-Baltic, either. Endzelīns (1914b: 102) associates these cases with the so-called ‘Curonianisms’ in Latvian dialects, which is not implausible.
- Li. *kǭla* ‘sandbank’ < **kalla* < **kalva* (cf. Lv. *kaĨva*, Lt. *kalvà* ‘hill; sandbank’) has undergone a number of phonological developments within Livonian, but since these are specific Livonian changes, the word likewise need not be dated to Proto-Finnic.

[96] A slightly different case is F *ranta*, E *rand*, Li. *rānda* (< **ranta*) ‘coast, shore’, which has been etymologized both as a loan from Baltic (Lt. *krañtas*, *kránta* ‘shore; precipice’; Būga 1908: 30; Терентьев 1990: 30; Liukkonen 1999: 117–119) and from an unrelated Germanic source, cf. ON *strǫnd* (< **strandō-*) ‘coast, shore’. Both etymologies are formally and semantically unproblematic; it does not seem possible to choose between them (LÄGLOS III: 127; Junttila 2012: 282).

In general, I have erred on the side of caution, and taken the liberty of leaving out etymologies which seem problematic to me for any reason. No explicit attempt at exhaustivity has been made in this survey, but the following hopefully covers the most unambiguous evidence.

1.3.2. Baltic loanwords with an IE etymology

In total, I have identified 70 certain loanwords which have a strong Indo-European etymology. I present these below, organized by semantic category. Where no reference is given, the comparisons are already present in Thomsen 1890. On the Baltic side, I have prioritized East Baltic attestations, but have cited Prussian equivalents wherever these are available. For economy of presentation, I have usually presented the IE etymology by citing a single cognate from a non-contiguous branch. These etymologies are generally well-established and can be found in standard reference works. Any additional discussions have been confined to the footnotes.

Kinship

– F dial. *ativo*, *atima* ‘visiting relatives; married woman visiting her parents’, Vp. *adiv*, *-on* ‘guest; marriageable woman’ (< **atę̥iva* / **atę̥ivo*)[97] ~ Lt. *ateĩvis*, dial. *atéiva* ‘foreigner,

[97] To explain the absence of the assibilation **ti* > **ci* in the word *ativo* ‘visiting relatives’, Kalima (1939–1940: 212) has posited an Early Proto-Finnic **ateį̯vo*. Despite Koivulehto (1972: 628), the development here must be distinguished from pre-Proto-Finnic **ej* > **ij*, which, at least in non-initial syllables, clearly predated the assibilation of dentals, cf. **vecitä* (= F *vesiä*) ‘water’ PART.PL. (< **vete-j-tä*) (Kallio 2012: 35). Instead (also in view of vowel harmony), we must reconstruct the Proto-Finnic diphthong *-ę̥i-* (which emerged due to the reduction of **-aj-* under certain conditions, see Kallio 2012: 32–34). The

newcomer' (to *at-eĩti* 'come, arrive', cf. Skt. *éti* 'go') (Kalima 1939–1940: 211–214)[98]

- F *heimo*, Võ. *hõim*, Li. *aim* 'tribe, kin' (?< **hęimo*; see p. 122–129) ~ Lt. *šeimà*, Lv. *sàime* 'family, household'; Pr. III *seimīns* 'gesinde' (= OIr. *cóim* 'dear')[99]
- F *morsian*, K *moršien*, E *mõrsja*, dial. *mõrsija* 'bride, newlywed' (< **morcijan*, OBL. **morcijamę-*) ~ Lt. *martì*, ACC.SG. *mañčią* 'son's wife; bride', Lv. *mȁrša* 'brother's wife'; Pr. III *mārtin* ACC.SG. 'bride' (cf. Lat. *marītus* 'married (man)', Gr. μεῖραξ 'girl' < **mer-ih₂-*)[100]
- F dial. *nepaa*, arch. *nevat* (< **nępat*) 'nephew, niece'; E *nõbu* 'cousin' (< **nępoi*)[101] ~ Lt. obs. *nepuotis* (GEN.SG. *-ies*) 'nephew, niece' (= Skt. *nápāt-* 'grandson')
- E *sõsar*, Li. *sõzār*, ?Võ. *sysaŕ* 'sister' (< **sęsar*) ~ Lt. *sesuõ* (OBL. *sẽser-*) 'sister' (= Skt. *svásar-*)[102]

---

diagnostic (South Finnic) forms for this reconstruction are unfortunately unattested.

[98] Forms in *-eĩvis* are limited to Lithuanian, but as the suffix **-vīs* is rare and unproductive (Skardžius 1948: 379), they may represent an archaism.

[99] Stang (1972: 28) can see "keine einleuchtende Verbindung" with the Irish word, but it is semantically very close to the Baltic ones. The eDIL (s.v. *cáem*) glosses the Irish word as 'dear, precious, beloved; belonging to the family'. In Middle Irish, the word is often used substantively in the sense 'relation, comrade'. In addition, a semantic parallel can be found between Skt. *śéva-* 'dear' beside OE *hīwan* 'household, family' (< **ḱei(H)-uo-*), which is most probably also from the same root.

[100] This is, of course, merely a root equation, but since the word *martì* is one of only two Lithuanian nouns with a nominative in *-ì* (the other being *patì* 'wife'), it seems very likely it is inherited. For Latin *-a-*, see Vine (2011: 265–266).

[101] In Estonian, we are dealing with an affective derivation by clipping, compare F *tyttö*, K *tyttö*, *tytöi* 'girl' ≪ **tyttär*.

[102] In North Finnic, we find F *sisar*, Vp. *sizar*, with an irregular *-i-*. Kallio (2018: 225, fn. 6) also takes Võro *sysaŕ* from **sisar*, and considers two independent loans to have taken place. Indeed, the change **i* > *y* between two sibilants is

– F *tytär*, E *tütar*, Li. *tidār* ‘daughter’ (< **tüttär*) ~ Lt. *duktẽ* (OBL. *dùkter*-); Pr. III *duckti* ‘daughter’ (= Skt. *duhitár*-)

Body parts

– F *hammas*, E *hammas*, Li. *āmbaz* (< **hampas*) ‘tooth’ ~ Lv. *zùobs* ‘tooth’, Lt. *žam̃bas* ‘edge, hem; (dial.) blade’ (= Skt. *jámbha*- ‘tooth, jaw’)
– F *napa*, E *naba*, Li. *nabā* (< **napa*) ‘navel’ ~ Lv. *naba* ‘navel’; Pr. E *nabis* ‘navel’ (= YAv. *nāfa*-)[103]

Adjectives

– F *ahdas*, Vt. *ahaz*, Li. *ǭ’dõz* (< **ahtas*) ‘narrow, cramped’ ~ Lt. *añkštas* ‘narrow’ (**amź*- + **-sta*-; = Skt. *aṃhú*- ‘narrow’, cf. LEW 11)[104]
– Ingr. *haljas* ‘verdant’ (cf. F *haljakka* ‘pale, faded’), E *haljas*, Li. *ǭļaz* ‘green’ (< **haljas*) ~ Lt. *žãlias*, Lv. *zaļš* ‘green’; Pr. E *saligan* */zaljan/ (~ Skt. *hári*- ‘fallow, yellow-green’)

---

paralleled by Võ. *sysalik* ‘lizard’ (< **sisalikko*), and *sys* ~ *sis* ‘then’ (cf. E *siis*). Võro [y] elsewhere primarily occurs as an allophone of /õ/ before *n*. I still wonder whether our word could irregularly reflect **sę̆sar* after all. The assumption of two independent loans is not very economical, and does not help explain the North Finnic **i*.

[103] The word for ‘navel’ is more likely to be loaned from Baltic than from Germanic. In Germanic, the meaning ‘navel’ is usual for the suffixed **nablan*- (> ON *nafli*, etc.), while the more basic **nabō* (> ON *nǫf*) means ‘nave (of a wheel)’ (Kalima 1936: 141).

[104] The comparison has been considered uncertain (e.g. Kalima 1936: 86; SKES I: 4) because the verbal root seen in F *ahta*- ‘cram, stuff’ appears to be inherited; however, the semantic development of the latter is most probably the result of secondary convergence with the Baltic loanword (Koivulehto 1998: 244; Aikio in prep. 51). Furthermore, Võ. *atma*, 3SG. *ata* ‘cram’ implies PF **akta*- rather than **ahta*-.

- F *keltainen*, E *kollane*, Võ. *kõllanõ* (< **kẹltainẹn*) ‘yellow’ ~ Lt. *geltónas*, Lv. dial. *dzẹltãns* ‘yellow’, cf. Pr. E *gelatynan*, probably for */geltajnan/ (~ YAv. *zairita-* ‘pale yellow’)[105]
- Vp. *kurdeh*, Vt. dial. *kurrõ* (GEN.SG. *kurtõõ*), E dial. (Mulgi) *kurre* (< **kurtẹh*),[106] elsewhere *kurt -i* ‘deaf’ ~ Lt. *kuřčias*, dial. *kurtùs*, Žem. *kuřtas* ‘deaf’ (cf. Lv. *kuřls*, *kùrls*, dial. *kuřns*, YAv. *karəna-* ‘deaf’)
- F *taaja*, dial. *tavea*, K *toakie* (< **takja*) ‘dense, frequent’ ~ Lt. *tánkus*, ADV. *tánkiai* ‘dense, frequent’ (= Parthian *tng* ‘narrow, tight’) (Liukkonen 1999: 140–142)[107]

---

[105] Traditionally, one has compared the noun F *kelta* ‘the colour yellow’, Vt. dial. *kõlta* ‘egg yolk’ with Lt. *geltà* ‘jaundice; (dial.) the colour yellow’ (Thomsen 1890: 172; Kalima 1936: 115). However, F *kelta* may easily be a recent back formation based on pairs such as F *puna* ‘the colour red’ ~ *punainen* ‘red’, and even Lt. dial. *geĨtas*, Lv. dial. (ME I: 543) *dzẹlts* ‘yellow’ are not attested in early sources. I prefer take **kẹltainẹn* ‘yellow’ directly from Baltic **geltâna-* (or even **geltaina-*? cf. Lv. dial. *dzẹltains*, ME I: 542) with adaptation to the Finnic adjectival suffix **-inEn* (cf. Liukkonen 1999: 39). An apparently old derivative is Lv. *dzẹl̂ta* ‘ground cedar’, which has also been borrowed into Finnic (see below).

[106] If we assume analogical generalization of the weak grade, K (Olonets) *kuuris* ‘deaf’ also appears to be regular from **kurtẹs* (cf. e.g. *kieral* ‘at once’ < **kertalla*; Kalima 1924: 166–167). In this case, F dial. (SE) *kuurne* and K *kuurnis* may be borrowings from Olonets Karelian with hypercorrect *-rn-* (cf. in particular Olonets *kuuru* ~ F *kuurna* ‘chute’; see Kalima 1924: 164–166). Given the complementary distribution of **kurtẹs* and **kurnẹs*, an analogical origin of the latter seems more promising than the assumption of a second, independent loan from Baltic **kuŕnas* (Kalima 1936: 124; Junttila 2019: 42).

[107] F dial. *tavea* replaces *tavia* (< **tayja*) under the influence of the adjectival suffix *-ea* (cf. dial. *lavea*, older *lavia* < **lakja* ‘broad’; T. Itkonen 1982: 123). The surprising reflex in Karelian is paralleled by North Karelian *voakie* ‘peg’ (< **vakja*) and *roakie* ‘limb’ (< **rakja*) (idem: 124–125). The acute in Lt. *tánkus* remains unexplained, as admitted by ALEW (1072) and Smoczyński (2018: 1446), but the IE etymology seems difficult to reject.

– F *tyhjä*, E *tühi*, Li. *tijā* (< **tühjä*) ‘empty’ ~ Lt. *tùščias*, Lv. *tukšs* ‘empty’ (= Skt. *tucchyá-*)

Nature

– F dial. (W) *hako* ‘conifer branch; needle’, E *hagu*, dial. *haga* ‘fine lopped branches; branch, stick’, Li. dial. (W) *a’g* ‘conifer needle’ (< **hako* ~ **haka*) ~ Lt. *šakà*, Lv. dial. (ME III: 642) *saka* ‘branch’ (= Skt. *śā́khā-*) (Būga 1908: 30; Ojansuu 1921: 6)[108]
– F *halla*, E *hall*, Li. *ǭla* (< **halla*) ‘frost, hoarfrost’ ~ Lt. *šalnà*, Lv. *saĺna* (LVPPV: *sàlna*) ‘frost, hoarfrost’ (~ Lt. *šálti* = Oss. ID *sæl-* ‘freeze’)[109]
– F *helle* (OBL. *helte-*) ‘hot weather’, K dial. (Olonets) *helleh* ‘sweltering’ (< **helteh*) ~ Lt. *šiĺtis* ‘heat’, cf. Lt. *šiĺtas*, Lv. *sìlts* ‘hot’ (= MW *clyd* ‘sheltered, warm’) (Ojansuu 1921: 7; Kalima 1936: 100)
– F *kelta-lieko* ‘ground cedar’, Vp. dial. *küud* ‘smoke tree’, Vt. *kõlta* ‘clubmoss?; a plant used to die (eggs) yellow’ (cf. VKS), E *kold*

---

[108] Thomsen (1890: 244) compared a different Baltic word, Lt. *žãgas* ‘hayrick, heap’, Lv. *schaggas* F.PL. ‘Laub, feine belaubte Ruthen, dergleichen sie zu Badequäste brauchen’ (Lange 1773: 272). Junttila (2017: 139) has defended this etymology, assuming the senses in both Finnic and Baltic developed from an earlier ‘trunk’. In support of this, he adduces the Latvian hapax *zęga* ‘body’ (in Rucava *apud* ME IV: 702). The Finnic senses are more easily derived from ‘branch’, however (only the eastern F, K *hako* ‘rotten or submerged fallen tree’, Vp. *hago* ‘fallen tree; snag’ would be derivable from ‘trunk’); and an original sense ‘trunk’ would also be hypothetical in Baltic. Junttila (in prep.) instead assumes that the original meaning in Baltic may have been ‘branch’, but this still presupposes one additional hypothesis in comparison to the etymology suggested here.

[109] Thomsen (1890: 220) compares F *halli*, E *hall* ‘grey; grey animal’, Li. *aļ* ‘grey seal’ directly with Lt. *šaĺnis* ‘grey cattle’. However, these are more probably both productively formed from the respective words for ‘hoarfrost’ (cf. Kalima 1936: 95).

'clubmoss' ~ Lv. *dzęĵta* 'ground cedar'. Plants named for their use as dyes (cf. Lv. *dzęltãns* 'yellow')

- F *metsä*, E *mets*; Li. *mõtsā*, Võ. *mõts* (?< **męcca*) 'forest, wood' ~ Lt. *mẽdžias*, Lv. *mežs* 'forest'; Pr. E *median* 'wood' (= R *межá*, SCr. *mèđa* 'border between fields; boundary strip'; to Skt. *mádhya-* 'in the middle')[110]
- F *routa*, Vt. *rõuta* (< **routa*) 'frozen ground' ~ Lt. *grúodas* 'frozen ground' (= Lat. *grandō* 'hail', cf. Rasmussen 1999: 152–153)
- F *takiainen*, dial. *takkiainen*, E *takjas*, dial. *takijas* (< **takkijas*) 'burdock' ~ Lt. *dagỹs*, Lv. *dadzis* 'burdock, thistle' to Lt. *dègti*, Lv. *degt* 'burn' (= Skt. *dáhati*; cf. Lt. *dagùs* 'prickly, bitter, harsh', LEW 85–86)
- F *taula*, E *tael*, Li. *da'ggõl* (< **takla*) 'tinder (fungus)' ~ Lv. *dagla*, *dęgla* 'tinder (fungus)', whose primary use is as fuel; to Lt. *dègti*, Lv. *degt* 'burn' (= Skt. *dáhati*)

Wild animals

- F *ankerias*, E *angerjas*, Li. *aṇgõrz* (< **ankęrjas*) 'eel' ~ Lt. *ungurỹs*; Pr. E *angurgis* */angurjəs/ 'eel' (~ Lat. *anguīlla*)[111]

---

[110] Kalima (1936: 11) considers the comparison phonologically difficult, although the main reason he doubts the loan etymology is that he suspects the Finnic word to be cognate with Taz Selkup *mačį* 'forest, tundra'. However, the Selkup affricate derives from Proto-Samoyed **-j-* (cf. Janhunen 1977: 85) and the word shows no regular correspondence with the Finnic word. It is rather cognate with Md. EM *moda* 'earth, soil' and Finnish *muta* 'mud' (< PU **muďa*, Aikio 2002: 22–23). Note that the Selkup word is no longer mentioned in SKES (II: 343), who nevertheless follow Kalima and consider the Baltic etymology uncertain.

[111] The initial *u-* could be explained as an East Lithuanian dialecticism, in which case the word must have been borrowed from there into the other dialects (Derksen 2015: 479). Vowel assimilation (cf. Būga RR II: 509;

- F *herhiläinen* ‘hornet’, E *herilane* ‘wasp’ (< **herhiläinen*) ~ Lt. *šìršė*, Lv. *siȓsenis* ‘hornet’; note Lt. dial. *šiȓšilas*, Pr. E *sirsilis* (= Lat. *crābrō*)[112]
- F *hirvi* ‘elk’, E *hirv*, Li. *īra* ‘deer, roe-deer’ (< **hirvi*, OBL. **hirve-*) ~ not in East Baltic, cf. Pr. E *sirwis* ‘roe-deer’ (= MW *carw* ‘deer’)[113]
- F *karva* ‘animal hair; coat colour’, E *karv* ‘hair, fur; bristle’ ~ Lt. *gauraĩ* ‘fur; bodily hair’, ?Lv. *gauri* ‘pubic hair’ (Ulmann 1872: 73 “Scheint nicht sehr bek[annt]”) (= OIr. *gúaire* ‘animal hair, bristle’)
- F *kiiliäinen*, Võ. *kiińläne* ‘botfly’ (< **kīlijäinen* ~ **kīliläinen*); E *kiil*, *-i* ‘dragonfly; botfly’ (< **kīli*) ~ Lt. *gylỹs* ‘gadfly’, dial. ‘sting’ (~ *gélti* ‘to sting’ = Arm. *kełem* ‘torment, afflict’) (Mikkola 1906: 78)
- F *käärme* ‘snake’, Li. *kīermõz* ‘woodworm’ (< **kärmes* / **kärmeh*), E dial. (Saaremaa) *kärm*, *kärv -i* ‘snake’ (< **kärmi*) ~ Lt. dial. *kirmìs* ‘worm; snake’, Lv. *cìrmenis*, dial. (Kurzeme) *cirmis* ‘maggot’ (= Skt. *kr̥mí-* ‘worm’)
- F *vaapsainen*, Vp. *bapshaine*, Vt. *vaapsia*, E *vapsik* (?< **vapsas*) ‘wasp’ ~ Lt. *vapsvà*, dial. *vãpsas*, Lv. dial. *vapsene*; Pr. E *wobse* ‘wasp’ (= Pahlavi *wpc* /wabz/)[114]

Otrębski 1955: 26; LEW 1163) seems less probable. An alternative account is that the *u*-vocalism originated in the zero-grade (**h_2ng^{wh}-*; cf. Smoczyński 2018: 1561), implying an old ablaut variant.

[112] Lt. dial. *šiȓšilas* is extremely rare; the *l*-suffix may have been added secondarily within Finnic, cf. in particular **mehiläinen* ‘bee’ (Nieminen 1934: 32–35; Kalima 1936: 100).

[113] Lt. *šìrvas* ‘grey, dapple-grey’ could be cognate if it originally meant ‘roe-coloured’, but more likely it represents a contamination of Lt. *šìrmas* ‘grey, dapple-grey’ (= Lv. *siȓms*) and *šývas* ‘grey, whitish (usu. of horses)’ (= Pr. E *sywan*, SCr. *sîv* ‘grey’). Particularly note that the acute accentuation would be in conflict with MW *carw* (< **ḱr-uo-*; cf. Zair 2012: 94–95). On Lt. *kárvė* ‘cow’ etc., see fn. 463.

Animal husbandry

- F dial. *ehkonen* (dial. *hehvo*; standard *hieho*), E dial. *õhv*, Li. *õ'v* (< **ę̀hva*) 'heifer' ~ Lt. obs. *ašva*, *ešva* 'mare'; perhaps Lv. ⟨ôssa⟩ 'mare' (Elger 1683: 133; cf. Karulis 1992 I: 468) (= Lat. *equa*)[115]
- F *hanhi*, E *hani*, Võ. *haaḱ* (< **hanhi*, OBL. *hanhę̀-*) 'goose' ~ Lt. *žąsìs*, Lv. *zùoss* 'goose'; Pr. E *sansy* 'goose' (= Skt. *haṃsá-*)
- F *oinas*, E *oinas*, Võ. *oinas*, dial. *oonas* (< **oinas*) 'ram' ~ Lt. *ãvinas*, Lv. *àuns* 'ram'; Pr. E *awins* 'ram' (cf. Skt. *ávi-* 'sheep')
- F *paimen*, Vt. dial. *paimõõ*, Li. *paint* (< **paimę̀n*) 'shepherd' ~ Lt. *piemuõ*, OBL. *píemen-* 'shepherd' (= Gr. ποιμήν)
- F *villa*, E *vill*, Li. *vīla* (< **villa*) 'wool' ~ Lt. *vìlna*, Lv. *viĺna* 'wool' (= Skt. *ū́rṇā-*)
- F *vuohi*, Vt. *voho* (< **vōhi*, OBL. **vōhę̀-*) 'goat' ~ Lt. *ožỹs*, Lv. *âzis* 'he-goat'; Pr. E *wosee* 'she-goat' (= Skt. *ajá-*)[116]
- F *vuona*, Vt. *võdna*, Võ. *vu̯u̯n*, Li. *ūoṇõz* (< **vōtna*) 'lamb' ~ not attested in Baltic; cf. OCS arнa 'lamb' (= Lat. *agnus*, Gr. ἀμνός)[117]

---

[114] F *vaapsainen*, K dial. *vuapsahane*, Vp. *bapshaine* reflect a derivative **vapsahainen* (cf. F *muurahainen* 'ant'). Possibly, these are built on a more basic **vapsas* continued by Vt. *vaapsaz* (absent from VKS; cited after SKES 1580) 'wasp', E *vaps-ik* 'hornet', in any case, is the result of suffix substitution (Nieminen 1934: 35). The long vowel attested in these forms is the result of the sporadic but frequent secondary lengthening before **-Cs-* (T. Itkonen 1987: 195–196). Li. *vaps*, NOM.PL. *vapsūd* 'wasp' (< **vapso* or **vapsoi*, Nieminen 1934: 33–34), may well be a later loan from Latvian.

[115] The secondary nature of F *h-* is supported by the presence of *h*-less variants well outside of the area of Estonian influence (Junttila in prep.). Baltic **ašvā* was already moribund at the time of its earliest attestations, being replaced with *kumẽlė* in Lithuanian and *ķève* in Latvian: Both versions occur side by side in Szyrwid: "*kumełe*[,] *afzwa*". Likewise, Elger has "*kiêwa*, D *ôssa*". In Bretke, *efchwų* GEN.PL. only occurs as a marginal gloss to *kumelių* (see ALEW 60), and in Ruhig, the word is semantically specified: 'eine Stutte großer Art' (Ruhig I: 8). For a discussion of the semantics, see *1.3.6.1*.

On Võ. *pahr* ‘boar’, see p. 133.

Agriculture

- F *herne*, E *hernes*, Li. *jērnaz* (< **hernes* / **herneh*) ‘pea’ ~ Lt. *žìrnis*, Lv. *zir̃nis* ‘pea’ (= Lat. *grānum* ‘grain’)
- F *pelu*, usu. PL. *pelut*, Vp. *pelu* (?< **pe̯lut* NOM.PL.)[118] ‘straw chaff’ ~ Lt. arch., dial. *pẽlūs*, Lv. *pe̯lus* PL.; cf. Pr. E *pelwo* ‘chaff’ (= Skt. *palā́va-* ‘chaff’, Lat. *pulvis* ‘dust’ < **pel-ou-*)
- F *siemen*, E *seeme*, Li. *sīemt* (< **sēmen*) ‘seed’ ~ Lt. *sė́muõ*, OBL. *sė́men-* ‘linseed, seed; (obs.) sowing’; Pr. E *semen* ‘seed’ (= Lat. *sēmen*)
- F *vannas*, Vt. *vadnaz* ‘ploughshare’, E dial. (W) *vannas* ‘plough beam’ (< **vatnas*) ~ unattested in East Baltic; cf. Pr. E *wagnis* ‘coulter’ (= Gr. (H.) ὀφνίς ‘ploughshare’) (Paasonen 1909b)[119]

---

[116] For the Votic form, cf. *toho* < **tōhi*, OBL. **tōhe̯-* ‘birchbark’.

[117] The substitution **gn* → **tn* (see also **vatnas*, below) appears to suggest that the cluster **kn* was not yet licenced at the time of borrowing. According to Paasonen (1909b: 17), **kn* had developed into **nn* already in Early Proto-Finnic, cf. F *ynnä*, Li. *īn̦õ*, (Salaca) *ūnis* ‘together’ (< **ük(t)-nä* ESS.SG. from **ükci* ‘one’), Võ. *nännüt* ‘see ACT.PRT.’ (< **näk-nüt*). Kallio (2008b: 313–314), however, argues that **kn* was preserved in ‘Core Finnic’, but examples like Estonian *näinud* < **näk-nüt* might show restored **k*. If Paasonen is correct, the word **sakna* ‘sauna’ must have post-dated these loanwords (Kallio 2008b: 315; although Kallio’s pre-Germanic etymology can hardly be consistent with this chronology).

[118] The word is generally *plurale tantum* in Finnish. I cannot establish whether the situation is similar for Veps, as the form is absent from Зайцева/Муллонен 1972. I can trace it back as far as a Central Veps “*ре̑лu*” cited in SKES 516.

[119] LÄGLOS (III: 368–369) leave open the possibility of a Germanic origin. The Norse data (cf. ON hapax *vangsna* OBL.SG., Nw. dial. *vangsne*, (17th c.) *vagnsne*) seem to point towards a proto-form **wagnVsnan-*. According to Kroonen

Technology

- F *ansa* ‘trap, noose’, E *aas*, Li. *ǭz* ‘noose, loop; handle’ (< **ansa*) ~ Lt. *ąsà*, Lv. *ùosa* ‘handle, eyelet’; Pr. E *ansis* ‘(pot)hook’ (= Lat. *ānsa* ‘handle’)[120]
- F *kirves*, E *kirves*, Li. *kīraz* (< **kirves*) ‘axe’ ~ Lt. *kiřvis*, Lv. *cìrvis* ‘axe’ (~ Gr. κείρω ‘crop, shave’; cf. Trautmann 1923: 135)[121]
- F *mäntä*, E *mänd* (< **mäntä*) ‘stirring stick’ ~ Lt. *meñtė* ‘mixing stick; trowel, paddle’ (= Skt. *mánthā-* ‘churning stick’)
- F *niisi*, usu. PL. *niided*, E *niied*, Li. *nīdõd* (< **nīci*, PL. **nītet*) ‘heddle’ ~ Lt. *nýtys*, Lv. *nĩtis* F.PL. ‘heddle’, cf. Lt. *nýtis* ‘warp thread’ (~ Lv. dial. *nĩt*, *nìt*[2] ‘thread (a needle)’; OIr. *sníid* ‘twist, bind’)
- F dial. *pahla* ‘(fishing) rod’, E dial. (W) *pahl* ‘spit, skewer’ (< **pahla*) ~ Lt. *baslỹs* ‘fencepost, stake’ from Lt. *bèsti* ‘drive in, stick in’, Lv. *best* ‘dig, bury’ (~ Lat. *fodiō* ‘dig, pierce, thrust’) (Kalima 1928: 102–103)[122]

(2013: 565), the umlaut in Swiss *wägese* ‘ploughshare’ could favour a reconstruction **wagnisan-* (cf. the forms in *Schw. Id.* XV: 770–774, where folk etymology is instead suspected). Since Prussian ⟨wagnis⟩ can stand for */wagnas/, a Baltic source does not raise any phonological issues, while the neuter *s*-stem **wagnas*, OBL. **wagnis-* reconstructed for Germanic by Karsten (1915: 84–85) remains purely hypothetical. For a similar reason, it is difficult to derive the Prussian word from Germanic (*pace* Smoczyński 2000: 132–133).

[120] A Germanic origin (Sammallahti 1998: 123), cf. ON *æs* ‘eyelet (in a shoe)’ (< **ansjō-*) is formally less straightforward.

[121] On R dial. *черв*, which hardly belongs here, see fn. 68. It seems natural to compare Lt. *kiřsti*, Lv. *cìrst* ‘chop, cut’ (= Skt. *kr̥ntáti* ‘cut off’). However, the dental would not be lost in the formation **kirt-u̯i̯a-*, which means the root must be identified as **ker-*. The rare deverbal suffix **-vīs* appears to form agent nouns (Leskien 1891: 348; Skardžius 1948: 379).

[122] On *-hl-* < **-sl-*, see Aikio 2015a: 44. Krevinian ⟨pahlis⟩ ‘stake’ cited by VKS, on the other hand, like Livonian *pǭ’lõz* ‘stake’, is from Latvian *pàlis* ‘stake’,

- F *ratas*, E *ratas*, Li. (Salaca) *rat* (< **rattas*) ‘wheel, cartwheel’ ~ Lt. *rãtas*, Lv. *rats* ‘wheel’ (= Skt. *rátha-* ‘chariot’, Lat. *rota* ‘wheel’)
- F *rattaat*, E arch. *rattad*, Li. *rattõd* (< **rattahęt* PL.) ‘cart’ ~ Lt. *rãtai*, Lv. *rati* PL. ‘cart’ (= Skt. *rátha-* ‘chariot’; see above)
- F *silta*, E *sild*, Li. *sīlda* (< **silta*) ‘bridge’ ~ Lt. *tìltas*, Lv. *tiȉts* ‘bridge’ (to Lt. dial. *tìlės* ‘bottom of a boat; planks (as paving)’; for the semantics, cf. OR мостъ ‘bridge; pavement, floor’; СДРЯ 11–14 V: 25–26)[123]
- F *tempaa-*, E *tõmba-*, Li. *tõ̃mbõ-* (< **tęmpaita-*) ‘pull, tug’ ~ Lt. *teñpti* (3SG. *-ia*), Lv. *tìept* (1SG. *-ju*, *-u*) ‘stretch, tighten’ (= NP *tāb-* ‘twist, spin’)[124]
- F *siula*, K *šikla* (< **sikla*) ‘side net in a seine’ ~ Lt. *tiñklas*, Lv. *tìkls* ‘net’, cf. Pr. E *sasin-tinklo* ‘snare’ (~ Gr. τείνω ‘stretch, pull tight’) (Koivulehto 1979a: 267–269)[125]

---

with the orthographic sequence *-Vh-* simply standing for **V̄* as in Krevinian ⟨pählin⟩ ‘top’ (= Vt. *päälin*), ⟨śohla⟩ ‘salt’ (= Vt. *soola*).

[123] Further, OR тьло ‘ground, bottom’, Pr. E *talus* ‘floor’, OE *þel* ‘plank (of wood); plate (of metal)’, in compounds ‘floor’. Probably, all of these words are related to Lat. *tellūs* ‘ground, earth; the Earth’ (< **telH-nu-* with regular laryngeal loss? cf. van Beek 2011: 162–165) and OIr. *talam* ‘earth, ground’, although the formations are all different.

[124] Root cognates have been suggested in several branches, but most of these are semantically uncompelling (cf. IEW 1064–1065). The long *-ā-* in Iranian is unexpected (for a suggestion, see Cheung 2007: 389), but the etymological equation seems in principle attractive. Here probably, if reliably attested, Ic. obs. *þömb* ‘bowstring’ (which need not originally be from ‘gut’; Ic. *þömb* ‘belly’ is perhaps to be separated as a *Reimbildung* to *vömb* ‘belly, rumen’).

[125] Within Balto-Slavic, the same root is continued in Lv. *tît*, 1PRES. *tinu* ‘wrap, wind’ (IEW 1065–1066; hence probably the intonation *tîkls* given by LVPPV), and R dial. *менёта*, Slk. arch. *tenatá* N.PL. ‘net snare’ (whose suffix can be compared with that of R *решетó*, Slk. *rešeto* ‘sieve’; Vaillant 1974: 697); see ALEW 1280.

- F *torvi*, E dial. *tõri* (< **torvi*, OBL. **torvę*-) ‘horn (for blowing)’ ~ Lv. *tàure* ‘hunting horn’, Lt. *taurẽ* ‘chalice, drinking horn’, cf. Lt. *tauras* ‘aurochs’ (= Gr. ταῦρος ‘bull’)
- F *tuulas* (< **tūlas*) ‘night fishing; fishing spear’; F *tuulaalla*, Vp. dial. *tuľhuuda* (< **tūlahęla*-) ‘spear-fish by torchlight’ ~ Lv. *dũlis* ‘torch for night fishing’, also ‘torch to fumigate beehives’, Lt. *dū́lis* ‘fog; smoke to fumigate beehives’ (cf. Hitt. *tuhhae-*^{zi} ‘produce smoke’)

Other

- F *jo* ‘already’, E *jo*, dial. *ju* ‘already, indeed’, Li. *jõ*, *ju* (< **jo*) ‘already’ ~ Lt. *jaũ*, Lv. *jàu* ‘already’; Pr. III *iau* ‘je’ (= OCS oy, ю ~ Go. *ju*; further Gr. (Hom.) αἰεί ‘always’; cf. Dunkel 2014: 352–353)[126]
- F *liika*, E *liig* (< **līka*) ‘surplus, extra; odd (number)’ ~ Lt. *liẽkas*, Lv. *lìeks* ‘surplus; odd (number)’ (= Gr. λοιπός ‘left over’)
- F *reuna*, Vp. *röun* ‘edge’ (< **rę̄una*), ?Võ. *rõõnõq* ‘strip (of fabric)’ (?< **rę̄unę̄k*)[127] ~ Lt. *briaunà* ‘brim, edge’ (~ ON *brún*

[126] The Baltic etymology, suggested by Thomsen (1890: 174), is considered by LÄGLOS (I: 140) to be “lautlich problematisch”, and a Germanic etymology is preferred (thus also SSA I: 238; Häkkinen 2004: 278). However, a Germanic origin is unattractive, as not only is the word unattested in Norse, the substitution Germanic **u* → Finnic **o* lacks convincing parallels. The Baltic etymology, on the other hand, does not pose any phonological issues (see p. 130).

[127] The Võro form may belong here if it originated in the eastern dialects showing *õu* : *õõ* gradation (cf. dial. *lõõnõq* ‘south’, GEN. *lõunõ*). The word is indeed recorded primarily in this dialect area, although VMS reports a couple of stray attestations from further west. If true, the Võro GEN. *ryynõ* would have to be analogical. Semantically, we can compare the sense ‘strip of fabric’ in Norse.

'eyebrow; brow (of a hill); strip of cloth', Skt. *bhrū́-* 'eyebrow') (Būga 1908: 42)[128]

- F *suola*, E *sool*, Li. *sūol* (< PF **sōla*) 'salt' ~ Lv. *sȁls* (GEN.SG. *sȁls*), dial. *sȁlis*[2] (GEN.SG. *sȁļa*; EH II: 470) 'salt' (= Gr. ἅλς) (Būga 1924b: 104)[129]
- F *vuoro* 'turn, shift', E *voor -u* 'turn, time' (< **vōro*) ~ Lt. *vorà* 'line, row', dial. 'turn, shift' (from Lt. *vérti*, Lv. *veĩt* 'pierce; thread, string together') (Koivulehto *apud* Häkkinen 2004: 1514)[130]

Inner-Baltic etymologies

There is a small group of words which can be analysed as derivatives within Baltic, even though their ultimate origin is unknown. In these cases, the direction of borrowing can nevertheless be considered certain:

- E *õis*, dial. *heis* 'flower' (< **häici*, OBL. *häite-*), whence F *heiti-* (< **häiti-*), E *õitse-*, Võ. *häitse-* (< **häiticce-*) 'to bloom' ~ Lt. *žíedas* (dial. *žáidas*), Lv. *ziêds* 'flower'; cf. Pr. TC *zaidiantẽ*

[128] For this polysemy, compare also Lt. dial. *brunìs* ~ *brùnė* 'eyebrow; dull edge' and further comparanda in LEW 57. The attractive analysis of Pronk (2015: 333) would see the forms with **-n-* as continuants of an original singulative **h₃bʰru-n-*, while Lt. *bruvìs*, Skt. *bhrū́-* 'eyebrow', etc. would reflect a fossilized dual **h₃bʰru-h₁*.

[129] On the Mordvin and Permic words for 'salt', see p. 229.

[130] For the development from 'row' to 'turn', compare R *о́чередь* 'line, row; turn'. The Baltic verb is related to OCS (Supr.) *проврѣти** 'thrust through', Bg. *ѳра* 'shove, thrust' and is generally considered to belong with Lt. *at-vérti*, Cz. *otevřít*i, Lat. *aperō* 'open' (e.g. LIV 227–228).

ACC.SG. ‘blossoming’ (~ Lt. *žydė́ti* ‘to bloom’) (Топоров/ Трубачев 1962: 247; Mägiste 1970)[131]

- F *haara*, E dial. *haar* (GEN.SG. *haara*), Võ. *haro*, Li. *a’r*, NOM.PL. *a’rūd* (< **hara* ~ **haro*) ‘branch, fork’ ~ Lv. *zars* (?→ Lt. dial. *žãras*) ‘branch; prong’ also ‘ray of light’[132] (to Lt. *žėrė́ti* ‘glow, sparkle’)
- F *härkä* ‘ox, bull’, E *härg*, Li. *ārga* ‘ox’ ~ Lt. *žìrgas*, Lv. *zir̂gs*; Pr. E *sirgis* ‘horse, steed’ (~ Lt. *(ap-)žeř̃gti* ‘sit astride’)
- F *härmä*, E *härm*, Li. *ārma* (< **härmä*) ‘hoarfrost’; F *harmaa*, Vt. *harmaa* (< **harmaka*), Võ. *haŕm*, *-i* (< **harmi*) ‘grey’ ~ Lt. *šarmà*, Lv. *sar̂ma* ‘hoarfrost’; Lt. *šìrmas*, Lv. *sir̃ms* ‘grey, dapple-grey’[133]
- F *luuta*, E *luud*, Li. *lūdõ* (< **lūta*) ‘broom’ ~ Lt. *šlúota*, Lv. *sluôta* ‘broom’ (to Lt. *šlúoti* ‘sweep’)[134]
- F *puuro* ‘porridge’, E *puder* ‘porridge; mash’ (< **putro*) ~ Lv. *putra* ‘porridge, mash; a kind of soup’, Lt. *putrà* ‘gruel,

---

[131] The etymological comparison with OHG *kīnan** ‘sprout, come forth’, OE *cīþ* ‘sprout, shoot’ (Walde/Pokorny I: 544; LIV 161–162; Kroonen 2013: 287) is not certain (cf. ALEW 1506), as the Baltic **d* is unexplained. Assuming an earlier present-tense formant (cf. Smoczyński 2018: 1735) remains *ad hoc*.

[132] Cf. *zarus zaruodama* ‘casting rays (of the sun)’ in folk songs, ME IV: 691–692; compare English *beam* or Lv. *stars* ‘ray of light; (dial.) branch’; see ME loc. cit.

[133] Baltic **šar̂mā-* ‘hoarfrost’ is a derivative of **šir̂ma-* ‘grey’. The semantics can be illustrated by several parallels: (1) Lt. *šeř̃kšnas* ‘hoarfrost, rime; grey (of animals)’ (= Sln. *srẹ̑n* ‘hoarfrost’ ~ CS срѣнъ ‘greyish-white’ < **ḱersno-*, cf. ME III; 722), (2) F *halli*, E *hall* ‘grey; grey animal’, Li. *aļ* ‘grey seal’ (< PF **halla* ‘hoarfrost’, see fn. 109), and not least (3) ME *hore-frost* ‘hoarfrost’, cf. OE *hār* ‘grey, hoary; grey-haired’. See also Liukkonen (1999: 38).

[134] According to Kortlandt (1995), *šlúo-* regularly reflects **ḱleh₃u-*, and the Baltic words are to be compared with Lat. *cluere*, *cloāre* ‘purify’. However, the implied phonological development in Baltic is doubtful (cf. Villanueva Svensson 2015). Furthermore, both Latin variants are hapaxes used to explain a divine name, and therefore raise suspicions (cf. Rix 1999: 519).

skilly; flour soup'. Compare Lt. dial. (E Aukšt.) *pùtelis* 'thickened soup of oatmeal and milk'[135]

- F *rako*, E *pragu* (< **rako*) 'crack, crevice; gap' ~ Lt. *spragà*, Lv. *spraga* 'gap (usu. in a fence), crack' (cf. Lt. *sprógti*, Lv. *sprâgt* 'burst, crack')[136]
- F *rouhi-*, Vp. *rouhi-* 'grind (coarsely), crush', E *rõhu-* 'press down, oppress' (< **rouhi-*) ~ Lv. *kràusêt* 'crush', Lt. dial. *kraušýti* 'barge, shove' (cf. Lt. *krùšti* 'pound in a mortar') (Kalima 1936: 156)[137]
- F *seiväs*, E *teivas*, Li. *tāibaz*, Võ. *saivas* (?< **stęipas*; see p. 122–129) 'post, stake' ~ Lt. *stíebas* 'stalk, trunk, pillar', ?Lv. *stìebrs* 'stalk, reed, rush' (~ R *сте́бель* 'stalk', SCr. *stáblo* 'tree; trunk')[138]

---

[135] Since the soup is generally thickened with flour, probably from Lt. *pùsti* 'swell up' (LEW 681–682), like *pùtos* 'foam, froth'. Other motivations are possible; compare the secondary sense 'eat sloppily', or even "Nuog putros tik pilvas išsipūtė" (Aukštadvaris, LKŽ) 'All I got from that *putrà* was a bloated stomach'. Note that Mikkola (1896a: 121) was unconvinced by this derivation, and derived the Baltic words instead from Finnic.

[136] The metatony is awkward, so the association with the verbal root may be secondary. If the initial *s-* is due to lexical convergence, the rare dial. *pragà* beside Lt. *próga* 'opportunity', dial. 'forest clearing' (cf. *progas* 'Lücke', ClG I: 1219) seem to suggest an analysis **pra-gā-* (Smoczyński 1998: 255–256); compare Lt. *próperša* 'thawed patch of ice; break in the clouds' beside *praparšas* (Szyrwid; see fn. 436). In that case, the root is perhaps that of Lv. *gãju* 'went'.

[137] Finnish *louhi-* 'chip away (stone), quarry', for which Thomsen (1890: 194–195) has suggested another Baltic etymology, most likely represents a secondary alternant of *rouhi-* (for similar cases, see Nikkilä 1999: 130–134), perhaps under the influence of *lohjeta* (*lohke-*) 'chip, break (INTR.)'.

[138] The Baltic acute is unexplained. Note the similarly obscure acute in Lt. *stámbas* 'stem, stalk' (= Skt. *stambha-* 'pillar').

- F *tapa*, E dial. *taba*, *tava* (< **tapa*) ‘custom, habit’ ~ Lt. dial. *dabà*, Lv. *daba* ‘way, custom’ (~ OCS по-добати, Go. *ga-daban* ‘befit, be suitable’)[139]
- F *tuura*, E *tuur* (< **tūra*) ‘ice chisel’ ~ Lv. *dùre* ‘fist’, dial. ‘ice chisel’[140] (from the verb Lt. *dùrti*, Lv. *duŕt* ‘stab, poke, prick’)
- Vp. *värpita-*, dial. *värbita-* ’spin, rotate (a spindle)’, Li. *vērbikšõ*, dial. *vā̆rbõkš* ’spin (thread)’ (?< **värpi-* ~ **värpe-*) ~ Lt. *veřpti* (PRES. *veřpia*) ’spin (thread)’, Lv. *vèrpt* ’spin; (REFL.) wind round’ (Thomsen 1890: 240; Posti 1946: 386)[141]
- F *vielä*, E *veel*, Võ. *vi̯i̯l* (< **vēlä*)[142] ‘still, yet’ ~ Lt. *vẽl*, dial. *vẽle*, *vẽlei* ‘again, still’, Lv. *vêl* ‘still, yet’ (?< a fossilized adverbial derivative of Lt. *vėlùs*, Lv. *vę̂ls* ‘late’ with a development ‘lately’ > ‘recently’ > ‘still’, cf. Būga 1923–1924: 95–96)

---

[139] The Baltic word must have developed from a verbal base meaning ‘be suitable’. Arm. *darbin* ‘blacksmith’ is rather to be derived from Urartian (Yakubovich 2009: 267–270), which also makes the appurtenance of Lat. *faber* ‘smith; artisan’ less certain (Pronk 2019a: 152).

[140] Although this meaning is limited to a small area in northern Latvia, so it cannot be entirely excluded that this sense arose under Finnic influence (cf. Thomsen 1890: 169), the derivation from the cited verbal base is semantically satisfactory; compare the parallel derivative in Žem. *durà* ‘ice chisel’ (LEW 113).

[141] According to Junttila/Holopainen (2022: 112–113), the connection to weaving is an East Baltic innovation, but the original semantics, even within Balto-Slavic, are difficult to establish (cf. LIV 691 s.v. **u̯erp-* fn. 1); therefore, it is unclear whether Pr. III *powiērpt* ‘leave, forsake’, ORu. вьрпати* (attested вьрпеши 2SG.PRES.) ‘tear, rob', CS на-врапити ‘invadere’ (Miklosich 1865: 399) even belong here. At least from a semantic point of view, it is tempting to compare OE *warp*, OHG *waraf* ‘warp (in a loom)’ (Persson 1912: 497–499; Trautmann 1923: 353), which could be connected by assuming Kluge’s law. Compare the OED’s definition of *warp*: “The threads which are extended lengthwise in the loom, usually *twisted* harder than the weft or woof" (emphasis mine).

[142] Livonian *vēl*, *ve’l* represents an independent loan from Latvian (cf. Suhonen 1973: 237).

1.3.3. Analysis of sound substitutions

1.3.3.1. Vocalism

***æ → *ē; *ā → *ō**

While it seems natural to interpret the first-syllable vowel in PF **sēmen* as an exact equivalent of Baltic **ē* (thus e.g. Kalima 1936: 68; cf. Lt. *sémen*-), this is rather a notational fallacy. The Proto-Baltic precursor of Lt. *ė* was almost certainly a low vowel */æː/. It remains low (in part) to this day in standard Latvian; likewise, the low vowel realizations /aː/ and /æː/ are still present in certain East Aukštaitian dialects (cf. Bacevičiūtė et al. 2004: 124–125), and one can still find the spelling ⟨a⟩ for Proto-Baltic **ā* in early Lithuanian texts deriving from Prussian Lithuania, such as the Wolfenbütteler Postille, the Mažvydas Catechism and sporadically elsewhere (Palionis 1995: 46).

As shown by Lehtinen (1967: 150–151) and Aikio (2012b: 232), Finnic **ē* has also developed from an earlier low vowel **ǟ* (e.g. PF *kēle*- 'tongue' < pre-PF **kǟlə* < PU **kälə*). This raising must have predated the emergence of secondary **ǟ* resulting from contraction over PU **x* and **ŋ* (e.g. PF **pǟ* 'head' < PU **päŋə*, UEW 365), which was no longer subject to raising.[143] Therefore, we should rather state that Baltic */æː/ was borrowed as pre-PF **ǟ*, and the subsequent raising in Finnic and Lithuanian must be considered parallel, unrelated developments.

[143] This chain of developments can be attractively analysed as a push shift. The fact that the Baltic loans underwent this raising in Finnic suggests, by extension, that they predated the loss of intervocalic **ŋ* and **x* (or at least the vowel contraction). The fact that these phonemes are not represented in the Baltic loanwords is not surprising, as no corresponding phonemes are present in Baltic.

A confirmation of this chronology is provided by the substitution of PB **ā* as PF **ō* in e.g. PF *vōhi* ← PB **âžē* 'goat'. Although in fact entirely analogous to the case of */æ:/ → **ē*, only this substitution has sparked any significant debate. One has either operated with a Proto-Baltic reconstruction **ō* (Kalima 1936: 66–67; cf. Mikkola 1930: 443, fn.) or assumed that the Finnic people came into contact with a Baltic dialect in which **ā* had become rounded, be it Curonian (Nieminen 1934: 59), High Latvian (Endzelīns 1932: 255), or "North Baltic" (Kallio 2008a: 272). However, all of these speculations are rendered unnecessary by the insight that Finnic **ō* has itself developed from an earlier **ā* (Pystynen 2018: 72–75).

The following additional pieces of support can be presented for this chronology:

a. Baltic **ā*-stems are overwhelmingly adopted as Finnic **a*-stems, cf. **halla* 'frost', **villa* 'wool'. Likewise, Baltic **-ē* was adopted as Finnic **-ä* in **mäntä* 'whisk' (← **mentē*).
b. The substitution **ō* → **ū*, **ou* is naturally accounted for if Proto-Finnic lacked a phoneme **ō* at the time of borrowing (see p. 126).

At the same time, a couple of arguments can be put forward in favour of a Baltic rounded vowel:

a. Baltic **ā*-stems are occasionally adopted as Finnic **o*-stems, as in F *heimo* 'tribe, kin' ← Baltic **šeimā-* (on which see *1.3.3.3*)
b. F *vohla*, dial. *vohli* 'kid', if loaned from Baltic (cf. Lt. *ožẽlis* 'kid'), implies an underlying **vōhl-*, with shortening of the vowel before a consonant cluster (Koivulehto 2000: 104).

The latter case can be explained easily provided the syncope of the medial **-ę-* is a late development (Kallio 2007: 241): Baltic **âžel-*

would be adopted into Finnic as **ā̆šę̄la*, which subsequently developed to **vōhę̄la* (with automatic **v-* before **ō-*) and finally to **vohla* by syncope. This incidentally nicely accounts for the absence of the development **wo-* > **o-* (Posti 1953: 72; Aikio 2014b: 10) in Finnic. The Proto-Finnic status of the word *vohla* remains doubtful, however, as the word is limited to the dialects of Western Finland, which seems suggestive of a local innovation.[144]

***e → *ä ~ *e; *a → *a ~ *o**

In some cases, short **e* is substituted as **ä*, cf. **mäntä* 'whisk', **värpi-* 'to spin' and perhaps the second syllable of **tüttär* 'daughter' (< **dukter-*). This is the expected substitution, as Proto-Baltic **e* was probably an open vowel */æ/, as it remains to this day in the modern languages. In my opinion, here also belongs **kärmes* 'snake', which is most probably derived from a full-grade variant **kerm-* still attested in Lv. *cȅrme* 'roundworm', Lt. *kermenaĩ* 'bee larvae' (Thomsen 1890: 98; Liukkonen 1999: 54).[145] Besides this, we find examples of the substitution **ę*: **sę̄sar* 'sister', **nę̄pat* 'nephew', **kę̄lta* 'ground cedar', **kę̄ltainę̄n* 'yellow', **pę̄lut* 'chaff', **tę̄mpaita-* 'pull, tug', **ę̄hka* 'heifer'. Kallio (2008a: 270) has argued that the conflicting reflexes might be explained if Baltic **e* was phonetically *[ɛ], standing somewhere in between Finnic **e* and **ä*. An interesting fact, however, is that all of the examples involve a Baltic back vowel in the second syllable, which

[144] The oft-quoted Estonian *vohl* is found only in the Kuusalu coastal dialect in the far north (according to VMS), and is probably a loan from Finnish.

[145] An ablauting **kerm-* : **kirm-* in Proto-Balto-Slavic might be required to account for the unexpected reflex **ir* as opposed to **ur* after a labiovelar, cf. MW *pryf* 'worm, maggot, fly' (see Kortlandt 1978: 240; Matasović 2004: 350).

suggests **ę* could have been a sort of compromise between the back-vocalic stem-vowel and front vowel of the initial syllable.[146]

The two substitution strategies for Baltic **e* are mirrored by the similar situation with regard to Baltic **a*: here the usual substitution is Finnic **a* (of which there are many examples), beside which examples of **o* can also be identified. Here, a similar solution could be proposed by suggesting that Baltic **a* was in fact **[ɔ]* (Kalima 1936: 64–65; Steinitz 1964: 338). Among the loanwords with a clear Indo-European background, four certain examples show an **o*: **oinas* 'ram', **torvi* 'horn' and **morcijan* 'bride', **rouhi-* 'crush'. These examples would be consistent with Nieminen's theory (1957: 199–201) that Baltic **o* reflects **a* where a front vowel follows in the next syllable. However, this theory encounters counter-evidence (e.g. **hanhi* 'goose'), and is typologically questionable (Steinitz 1964: 336).

A typologically more apt observation is that Baltic **a* in all three examples is found adjacent to a labial, viz. **avinas*, **taurē*, **martjan*. It is therefore possible that the substitution with Finnic **o* was a reflection of an allophonic rounding in a labial environment within the Baltic donor dialect. Nevertheless, we could only talk of a tendency here, as no rounding is found in **hampas* 'tooth', **karva* '(animal) hair' or **vapsas* 'wasp'. All in all, the evidence is rather too limited to convincingly identify conditioning factors.

The substitution observed in PF **härmä* ← **šârmā-* 'hoarfrost' is quite unclear.[147] The expected back vocalism seems to be found in F *harmaa*,

---

[146] This is an argument in favour of interpreting **męcca* 'forest' as archaic. See below.

[147] Since Thomsen (1890: 221), one has generally referred to a Latvian *sęr̂ma* (cf. ME III: 819; EH II: 478) to support the reconstruction of a Baltic source form **šer̂mâ*. However, this Latvian form is probably the result of a secondary

Võ. *haŕm* ‘grey’, which Liukkonen (1999: 38) has plausibly analysed as inner-Finnic derivatives of a noun **harma*; however, no regular derivational process can account for shift to front vocalism within Finnic. Although similar cases of secondary vocalism are sporadically observed in Finnic (Saukkonen 1962; Nilsson 1996: 186), these normally concern words of an expressive character. See also *1.3.5.3*.

***i → *i ~ *e; *u → *u ~ *ü**

Baltic **i* is normally reflected as Finnic **i* in loanwords (**villa* ‘wool’, **silta* ‘bridge’, etc.). In some cases, however, we find Finnic **e*, instead, viz. **herneh* ‘pea’ (← Baltic **žirnīs*), **herhiläinen* ‘hornet’ (← **šiŕš*-) and **helteh* ‘hot weather’ (← **šilta*-). This hesitation could simply be attributed to a more centralized pronunciation of /i/, as is found in modern Lithuanian (Pakerys 2003: 24–25). On the other hand, conditioning factors may be identified: Kalima (1936: 70) has attributed the lowering to the influence of a following resonant. Yet, as Ritter (1998) has pointed out, it is hardly a coincidence that all the examples feature an initial **š* or **ž* in Baltic.[148]

---

dialectal development (Endzelīns 1923: 36–37) and cannot be projected back to Proto-Baltic.

[148] Ritter in fact operates with a rule the **i* is lowered after both **k*- and **š*-. However, the examples with **k*- are unconvincing. For **kärmeh* ‘snake’, I posit an original *e*-grade; see above. In view of its distribution, F *kelles*, *kelle*, K dial. (N) *kelleš* ‘split log; large round chip; thick slice, chunk’ is more likely to be loaned from Sámi (cf. Sá. N *galda* ‘block of wood; tree stump’, Sk. *kõldd* ‘block; wooden lure’) than the opposite (*contra* Kalima 1936: 115). This is supported by the fact that the substitution of Finnic *e* → Sá. **ę* is practically unparalleled (Aikio 2006b: 32), while the opposite (i.e. Sá. **ę* → Finnic *e*) is known to have occurred (see Aikio 2009: 77). If true, then the association with ‘something split’ would have arisen secondarily within Finnic, and the connection to Lt. *skìltis* ‘clove; slice; piece cut off’ looks more tenuous.

This might be phonetically understood if we suggest that Baltic **š* and **ž* were realized as retroflex consonants, as usually assumed for Proto-Uralic **š*. Retroflexion tends to be disfavoured in the environment of front high vowels, and may be accompanied by concomitant vowel lowering (Hamann 2003: 94, 99–100). However, the substitution **ä* in **härkä* 'ox; bull' (← **žiŕga-*) might suggest that we are dealing with a genuine sound change in a Baltic dialect,[149] which would provide evidence that the source language was not a direct ancestor of the attested East Baltic languages. The fact that we find **ä–ä* as opposed to **ę–a* in this word (see above) probably implies that lowered **i* was still phonologically distinct from **e*.

As a substitution for Baltic **u*, we find both **u* (in **kurtęh* 'deaf' and **putro* 'porridge') and **ü* (in **tüttär* 'daughter' and **tühjä* 'empty'). This is noted by Thomsen (1890: 100) and Kalima (1936: 71), but not commented upon. Again, one might attribute this vacillation to a 'laxer' pronunciation of Baltic short vowels (cf. Pakerys 2003: 24–25), but Koivulehto (1971) has compared the front-vocalic forms to doublets such as F *rastas* ~ dial. *rästäs* 'thrush'. In light of this, Kallio (2008a: 269) writes "[t]here are no reasons to think that the substitution **u* → **ü* had anything to do with the actual pronunc[ia]tion of Proto-Balto-Slavic **u*". In my view, it is anachronistic to use post-Proto-Finnic vacillations such as that in the word for 'thrush' to explain phenomena in Early Proto-Finnic (see *1.3.5.3*).

In the word for 'daughter', the front vocalism can be explained in the context of Finnic vowel harmony: we may assume that the choice of front vocalism was triggered by the second syllable of Baltic **dukter-*.

[149] This is far preferable to seeing the source in the deverbal noun *žargà* 'spread legs' (Liukkonen 1999: 55–56, taken over by Junttila in prep.), which lacks the required semantic specialization.

Such an explanation does not really work for 'empty', however; although it is phonetically possible that the second-syllable vowel in Baltic **tuštja* was allophonically fronted in the neighbourhood of **j*, this is an *ad hoc* assumption, especially in view of the back-vocalic **haljas* 'green' and **ankẹrjas* 'eel'. It therefore seems that the explanation should be at least partly phonetic, although multiple factors may be at play.

The length of Baltic **ī* and **ū* is reflected in the Finnic loans, cf. **tūlas* 'spear for night-time fishing', **kīli-* 'gadfly', **nīci* 'heddle'.

?*ę̄ → *ęi ~ *ī; *ō → *ou ~ *ū

A very interesting case as regards vocalism is PF **hęimo* ~ **haimo* 'tribe, kin'.[150] Here one finds reflexes of a diphthong **ęi* throughout all of Finnic except in Livonian and South Estonian, where we instead find **ai* (cf. Li. *aim*, Leivu *aim*). The following words show a similar pattern, showing **ei* in "Core Finnic", and **ai* elsewhere (Kallio 2014: 159):

- F *heinä*, E *hein* ~ Li. *āina*, Võ. *hain* 'hay' (~ Lt. *šiẽnas*)
- K dial. *leinä*, E *lein* 'grief, sorrow' ~ Seto *lainalinõ* 'sorrowful'
- F *leipä*, E *leib* ~ Li. (Salaca) *laib** (Winkler/Pajusalu 2009: 107), Leivu *laib* 'bread' (~ Lt. dial. *kliẽpas*, ON *hleifr*)
- F *reikä* 'hole', E dial. (insular) *reig* 'wound' ~ Võ. *raig* 'scab'
- F *reisi*, E *reis* ~ Võ. arch. *raiź* 'thigh' (~ Lt. *ríetas*)
- F *seinä*, E *sein* ~ Li. *sāina*, Võ. *sain* 'wall' (~ Lt. *síena*)
- F *seiväs*, E *teivas* ~ Li. *tāibaz*, Võ. *saivas* 'post, stake' (~ Lt. *stíebas*)
- F *seiso-*, E *seisa-* ~ Võ. *saisa-* 'stand'

[150] For a more detailed account of this problem, see now Jakob forthc. d.

It is remarkable that the majority of these examples have Baltic comparanda.[151] Only the last is inherited from Proto-Uralic, where the cognates (e.g. Sá. N *čuožžut*, Eastern Mansi *tuńś-* 'stand') point to PU **saŋća-*.[152] Therefore, it is normally assumed that the *e*-vocalism shown by 'core Finnic' is an innovation, and that Livonian and Võro preserve an archaism (Thomsen 1890: 101–102; Koivulehto 1979b: 140). Since inherited **aj* is normally preserved as such throughout Finnic,[153] one must then speak of a "sporadic development" (Kallio 2014: 159).

In view of the systematic distribution of the reflexes, it is likewise unattractive to assume multiple layers of independent loanwords (e.g. Uotila 1983: 7–8; Viitso 1998: 12). The lack of any clear conditioning factors (cf. Kallio 2018: 258–259),[154] instead rather suggests that we

---

[151] Both **reikä* ~ **raika* and **leinä* ~ **laina* have been derived from Baltic, too. Liukkonen (1973: 17–25; cf. Sammallahti 1998: 127; SSA III: 60) compares the former with Lt. *riẽkti* 'slice (e.g. bread); plough for the first time', *riekẽ* 'slice', but this is nothing more than a (semantically weak) root etymology. Nirvi (1964: 153–154) has derived the latter from Lt. *klíenas*, Lv. *kliẽns*[2], *kliẽns* 'thin, lean', but this again requires unsubstantiated assumptions with regard to semantics (van Linde 2001: 291–293).

[152] See Kallio (2007: 231–232; 2012: 35–36). Pystynen (2014a) rejects this reconstruction and prefers **sańća-* (thus also Sammallahti 1988: 549); however, **ńć* does not normally develop into **js* in Finnic, instead simply becoming **s*, cf. **osa* 'part, share' (< **ońća*), **kusi* 'urine' (< **kuńćə*). At the same time, **-ŋć-* (> *-ŋś-*) > **-js-* would be a typologically similar development to **-ŋs-* > **-ws-* found in F *jousi* 'bow' (< **joŋsə*). I wonder whether such a PU reconstruction could also explain the difference between Khanty **ʌāńć-* ~ **ʌi̯ńć-* (Vakh *li̯ńť-*, Kazym *ʌɔńś-*) 'put, set' (< **saŋća-*) and **kus-* (Vakh-Vasjugan *kŏs-*, Kazym *χŏs-*) 'urinate' (< **kuńćə-*).

[153] For example, F *aivot*, Vp. *aivod*, Võ. *aivõq* 'brain' (< **ajŋə*, UEW 5); F *kaiva-*, E *kaeva-*, Li. *kōva-* (< **kauva-* < **kaiva-*; Kallio 2016: 55) 'dig' (< **kajwa-*, UEW 116–117); F *aita*, E *aed*, Võ. *aid* 'fence' (< **ajta*, Aikio 2014b: 1–2).

should reconstruct two diphthongs for Proto-Finnic, which I will provisionally notate **ai* (> F *ai*, Võ. *ai*) and **?i* (> F *ei*, Võ. *ai*).

Of course, the diphthong **?i* must somehow be part of the Proto-Finnic phonemic system, and so the number of options is limited. If we consider the *i*-diphthongs reconstructed for Early Proto-Finnic by Kallio (2018), the result is rather interesting:

Table 3. i-diphthongs in Early Proto-Finnic (based on Kallio 2018)

(*ī)	*ui
	*oi
*äi	*ai

Three possible diphthongs appear to be missing: **ei*, **ęi* and **üi*. It is attractive to assume that one of these corresponds to our diphthong **?i*. Our choice is narrowed down the fact that Estonian and Votic show a partial back-vocalic inflection for this group of word, cf. E *leibu*, Vt. *leipoi* PART.PL. < **-oita* < **-a-j-ta* (Kallio 2014: 159). As this is can hardly be analogical, it is a compelling argument in favour of original back vocalism, but not necessarily in favour of an original **ai*. I would therefore like to suggest the Proto-Finnic reconstruction **ęi*.[155]

---

[154] Kallio's own solution seems to be to assume a residual Baltic 'substratal tendency' to confuse **ei* and **ai*, but this is clearly anachronistic, not to mention that it is precisely Võro and Livonian, which have been subject to the most persistent Baltic substrate influence, that have preserved the supposedly more archaic form.

[155] The relevance of forms like Sá. N *suoidni* 'hay' < **šaina*, need not be overstated, as it is possible that the Sámi loans were adopted independently (see *1.4.1*, particularly p. 210 onwards). The change **a* > **ę* in the word for 'stand' could be explained as a raising due to the influence of the following palatal cluster **saŋća-* > **saŋśa-* > **sęjsa-* (cf. Ravila 1935: 32, fn. 1; Viitso 1978: 97). Although *ad hoc*, attributing a unique change to a unique

In this case, we can assume a regular fronting **ę* > **e* in "Core Finnic" conditioned by the following palatal resonant, triggering an automatic shift to front harmony (i.e. **lęjpa* > **leipä*). If we generalize this sound law to any tautosyllabic palatal, we could also explain the fronting of **męcca* 'forest' (Li. *mõtsā*, Võ. *mõts*) to **meccä* (F *metsä*, E *mets*) in "Core Finnic" (Santeri Junttila p.c. March 2022), as **c* must have still remained a palatal consonant in Early Proto-Finnic. At the same time, we can assume a regular lowering **ęi* > **ai* in Livonian and South Estonian (cf. Viitso 1978: 95–97).[156]

According to Kallio (2018: 262) **ei* has arisen secondarily in later Proto-Finnic due to contraction, cf. F *seimi*, Võ. *seiḿ* 'manger' < **sewi-mi*.[157] As a similar contraction took place in the loanword PF **oinas* 'ram' (← Baltic **awinas*), there seems to be no chronological objection to the assumption that the Baltic loans predated the emergence of Late Proto-Finnic **ei*. The absence of **ei* in Early Proto-Finnic may be explained by the development of earlier **ej* into **ij*

---

environment is better than assuming a sporadic change with no conditioning factors. For more discussion of these points, see Jakob forthc. d.

[156] The diphthong **ęi* seems only to be found in evidently late words like **lęikka-* 'to cut' (cf. Kallio 2018: 260; even here we find E dial. *leika-*), **pęippoi* (> F *peippo* 'finch', Vt. *põippõ* 'chick'). An exceptional case is F *leivo*, Võ. *lõiv* 'lark' (< **lęivo*), normally taken as a Germanic loan (LÄGLOS II: 190–191), although Schrijver (1997: 309) considers the possibility of a parallel substrate borrowing. It is possible that the preservation of **ęi* is due to the stem-vowel **o*, cf. Võ. *hõim* (but Leivu *aim* ← Livonian?; cf. Pajusalu, Krikmann & Winkler 2009: 293–294; Jakob forthc. d.). This would not only explain *põippõ*, but also the Votic back-vocalic forms *sõiso-* (in NE Estonian dial. also *sõisa-*) 'to stand', *õimo* 'kin, relatives'. In that case, Võro *hõim* would then need not have been borrowed from North Estonian (cf. Kallio 2021: 125).

[157] Apparent examples of **ei* often show irregularities. For 'manger', some languages show reflexes of **soimi* instead (> F dial. *soimi*, E *sõim*, dial. *soime*). For the verb **peittä-* 'cover, hide', containing the causative suffix **-ttA-*, South Estonian *pi̯i̯tä-* appears rather to suggest **peettä-*.

(**pīmä* ‘milk’ < **pejmä*, of Iranian origin,[158] cf. Av. *paēman-* ‘mother’s milk’, Holopainen 2019: 178–180). If the above account is correct, the Baltic loans must have post-dated this change. There are two Baltic loans, however, which are argued to have predated the change **ej* > **ij* (cf. Toivonen 1917: 27–28; Kallio 2008a: 273):

- F *liika*, E *liig* ‘surplus, extra’ ~ Lt. *liẽkas*, Lv. *lìeks* ‘surplus’[159]
- F *tiine*, Vp. *tineh*, Li. (Salaca) *tīn* ~ Žem. *dienì* NOM.SG.F. ‘pregnant (of animals)’ (Lõo 1911: 86; Kalima 1936: 169 with “?”).

If we assume a specifically East Baltic source for the Finnic loans, then the most natural solution is to assume that **ī* in these cases is a direct substitution of East Baltic **ę̄*. Phonologically, such a substitution would not be unexpected given the absence of long **ē* in Early Proto-Finnic (see p. 116). In this case, one might imagine a chronological difference, with **ęi* representing an earlier, still diphthongal, pronunciation of Baltic **ę̄* (Liukkonen 1973). However, this is not strictly necessary, as the substitution **ē* → **ęi* in the absence of a corresponding long monophthong is also quite conceivable. The same substitution is found for Swedish /ē/ in recent loanwords in Finnish, where inherited **ē* has developed into /ie/, e.g. F *kreivi* ‘count’ ← Sw. *greve* /gre:vɛ/ (Thomsen 1870: 56–57; Būga 1908: 23–24).[160]

---

[158] The substitution PU **e* for Iranian **a* is more or less regular in the position adjacent to a palatal, cf. **sejtə* ‘bridge’ (← **saitu-*, cf. YAv. *haētu-* ‘dam’), **rećmä* ‘rope’ (← **raćman-*, cf. further fn. 267).

[159] Note that, morphologically, Baltic **lę̄kas* ‘surplus’ more probably reflects an earlier **laikas* (= Gr. λοιπός ‘left over’).

[160] This might explain the substitution of Baltic **ē* as **ęi* in some Livonian loanwords from pre-Latvian, provided these postdated the Livonian raising **ę̄* > **ī̯* (> Courland *ō̃*, Salaca *ū̃*; Kallio 2016: 49); compare Li. *kõ̃idaz* ‘weaver’s reed’ (← Lv. *šķìets*, Lt. *skiẽtas*), *lõ̃iga* ‘surplus’ (← Lv. *lìeks*, Lt. *liẽkas*).

This analysis seems to be confirmed by the substitutions of Baltic **ō*, which is not diphthongal in origin. Here, we also find two Finnic equivalents: **ū* (in **lūta* 'broom') and **ou* (in **routa* 'frozen earth'), which can only be understood as two alternate substitution strategies for a foreign phoneme **ō*; compare similarly F *housut* 'trousers' (← Sw. arch. *hosor* 'leggings'), *ruusu* 'rose' (← Sw. *ros*); Thomsen 1870: 51.[161] The parallelism between **ę̄* → **?i* and **ō* → **ou* is, incidentally, another argument in favour of interpreting **?i* as **ęi*.

Against the direct substitution **ę̄* → **ī*, however, speaks Finnic **tīneh* 'pregnant (of animals)'. As already noted by Kalima (1936: 169), the absence of the change **ti* > **ci* in Finnic can only be understood if **ī* in this word is of secondary origin. Therefore, one has assumed an earlier **tejnəš* (Koivulehto 1972: 627–628; Liukkonen 1999: 144; Aikio 2014c: 90–91). The assumption that the change **ej* > **ij* postdated the assibilation **ti* > **ci* is potentially problematic. The following facts would speak against this chronology:

– The Baltic loans **routa* and **torvi* evidently postdated the symetrical change **ow* > **uw* (cf. Pystynen 2018: 53).[162] At the

[161] For **routa*, Junttila (2016b: 226) prefers an original **graudā*, reconstructed on the basis of R *ƨpýda* 'mass, heap', Pl. *gruda*, Sln. *grúda* 'clod (of earth)'. However, this is hardly necessary from a phonetic point of view, and using an actually attested Baltic form as a source is of course preferable; note that the Slavic words are probably unrelated to Baltic **grôdas* (see Villanueva Svensson 2015: 315).

[162] Compare PF **ūtin* 'mosquito curtain' < PU **owdəm(ə)* (= Eastern Mansi *åml*; Komi (Permjak) *ȩn*, pointing to **o(-ə)*, see Aikio in prep. 81–82); PF **tūli* 'wind' < PU **towlə* (cf. Ma. W *tul* 'storm', Komi *tȩv* 'wind'; see Aikio 2012b: 243); PF **kūsi* < PU **kowsə* (cf. Komi-Permjak *kȩz*, Eastern Mansi *χɔwt*, North Sámi *guossa* 'spruce'; Collinder 1960: 407; Живлов 2014: 139). On the principle of symmetry in sound changes, see now Jakob forthc. d.

same time, the Baltic loans predated the assibilation (see *1.3.3.2*).

– If Finnic **?i* should be interpreted as **ęi*, the example F *reisi*, Võ. arch. *raiź* (< ?**rę́ici*) would show assibilation but lack the raising to **ī*.

In my opinion, there is only one possible Uralic reconstruction which could safely account for Finnic **tīneh*, namely **tüjnəš*. The change **üj* > **ij* would run parallel to the established change **iw* > **üw* witnessed in F *syvä*, Võ. *süvä* 'deep' < PU **tiwä* (Aikio 2015b: 9).[163] Importantly, the aforementioned example immediately confirms the suggested chronology, as the initial sibilant in F *syvä* implies an intermediate stage **civä*, whereby the change **iw* > **üw* must have post-dated the assibilation of **t*. As PU **e* and **ü* merged in Mari, this new reconstruction still allows for an equation with Ma. E *tüž*, dial. *tü.üž* 'pregnant (of animals)' (E. Itkonen 1953: 183; Aikio 2014c: 90–91), but seems to speak against the comparison with Baltic. Note that the existence of Baltic loanwords in Mari is itself questionable (see *1.4.3*).

In conclusion, despite the opposite conclusion of Kallio (2008a: 273), the Finnic reflexes of Baltic **ę̄* and **ō* can be explained with the assumption of a monophthongal pronunciation in Baltic,[164] and

[163] The change also has a potential parallel in Vp. *silöi*, E *siil*, Li. *tsīļ* 'hedgehog' (< PF **sīli* < ?**ćüjələ*), cf. the cognates Ma. W *šülə* and Hungarian *sün*, older *szül* 'hedgehog', which suggest a rounded first-syllable vowel (Aikio in prep. 127). The reconstruction of PF **kū̆* 'adder' is also too uncertain for it to constitute a counter-example (< **küü* ?< **küjü* ?< PU **kejəw*, cf. Md. E dial. *kijov* 'snake'; see Pystynen 2017). On other possible exceptions, see Jakob forthc. d.

[164] That we find a diphthong in **paimęn* 'shepherd' (~ Lt. *piemuõ*) need not be an issue. It is known that Baltic **ę̄* was a conditioned development. If, for instance, it only arose under stress (cf. Hirt 1892: 37–40; Kortlandt 1977: 323), the word for 'shepherd' would have exhibited an alternation, viz. NOM.SG.

therefore an East Baltic origin of the loanwords (for a discussion, see *1.3.3.4*). This, incidentally, can be seen as an argument for a Baltic origin of PF **hę̇ina* 'hay' and **sę̇ina* 'wall', even though the corresponding Baltic words do not have reliable Indo-European cognates beyond Slavic (cf. OCS сѣно 'grass, hay';[165] стѣна 'wall'; on the latter, see also p. 347–348).

***eu → *ę̇u**

In view of the etymology PF **rę̇una* ~ Lt. *briaunà* 'edge', several scholars (Būga 1908: 42; Kalima 1936: 75) have argued that the Baltic loanwords predated the change **eu* > **jau*.[166] While it is often assumed that this is a common Baltic-Slavic change (Kortlandt 1989a: 48; Matasović 2008: 105), it does not appear to have been shared by Prussian (Levin 1974: 5, fn. 4; Derksen 2010: 38),[167] and the idea that

---

**pę̋mōn*, GEN.SG. **pâimenés*. It is possible that the allomorph **pâi-* was generalized in the dialect which donated the form to Finnic (note that the Finnic form must in any case be from an oblique form, see further p. 147–148). In the case of **häici* 'flower', a form with **âi* is actually attested in Lt. dial. *žáidas* 'flower'.

165 Guus Kroonen (p.c. August 2022) suggests Du. *heen* 'upright sedge, *Carex stricta*' as a possible cognate, noting that the plant is used as animal fodder.

166 The other example, F *leuka*, E *lõug*, Li. *lõga* 'chin' ~ Lt. *liaukà* 'gland' is highly doubtful for semantic reasons (cf. Nieminen 1945: 45; Junttila 2016b: 222–223, whose alternative does not fare much better).

167 The Elbing Vocabulary consistently shows ⟨eu⟩. Note that the glide in Pr. E *piuclan* 'sickle' was not adopted analogically from the full-grade (Arumaa 1964: 87), but is instead probably from inherited **-i̯-* (see Hackstein 1992). The Third Catechism offers very little evidence: for **jau* clearly speaks *iaukint* 'üben' (= Lt. *jaukìnti* 'tame, train', OCS оучити 'teach'). On the other hand, the PRET. *driāudai* 'furen (sie) an' beside IMP.PL. *draudieiti* seems to show a similar pattern of 'breaking' under stress otherwise observed only in *e*-diphthongs (cf. *tiēnstwei* 'reytzen' beside IMP.PL. *tenseiti* */tenséiti/ 'reitzet', *etwiērpt* 'vergeben' beside IMP.SG. *etwerpeis* */etwerpéis/ 'verlasse', cf. Kortlandt 1998:

the diphthong **eu* may have been preserved in Baltic until quite recently remains plausible (Kallio 2008a: 274). Thus, Nieminen's chronological concerns (1945: 53–55) can be disregarded.

A valid criticism of Nieminen (1945: 43–45) concerns the Baltic reconstruction: it is indeed true that the general hesitation between /Cr/ and /Cr^{j}/ in Lithuanian dialects (Zinkevičius 1966: 153–156) makes the Lithuanian anlaut non-probative. On the other hand, Lv. *braũna* 'shed skin, scale; husk' (cf. ME I: 327), which would support Nieminen's reconstruction **braunā-*, is semantically remote and better kept separate (LEW 57; ALEW 147). Therefore, I see no particular reason to doubt the Baltic etymology for Finnic **rę̆una*. From a phonological perspective, a Baltic loan predating the change **eu* > **jau* seems preferable to the direct substitution **r^{j}au-* → **rę̆u-* (Kulonen 1988).

Non-initial syllables

In Early Proto-Finnic, the vowel contrasts in non-initial syllables were very limited (cf. Kallio 2008b: 269). In fact, it seems possible that only the archiphonemes **A* and **E* existed at the time of the Baltic loans, and that later **i* and **u* can be interpreted instead as **Ej* and **Ew* (Kallio 2012: 31–32). This explains the adoption of both **i* and **u* as **E*, cf. **kärmes* (← **kermis*), **ankę̆rjas* (< **angurjas*; cf. Kallio *apud* Junttila 2015a: 19). The phoneme **o* in non-initial syllables may have synchronically still been **aw*. This would explain **jo* 'already' (< **jaw*

124), and would imply an earlier **driēud-* : **dreud-V́-*. This interpretation would be supported by *pievſſen* ACC.SG. 'pine' in the Trace of Crete (Lemeškin 2014: 142; in our interpretation: */piēusen/), provided this is correctly read (differently see Kaukienė/Jakulytė 2015: 46–47). If this is the case, *iaukint* must be understood as an East Baltic loanword.

← Baltic **jau*), which as a prosodically unstressed particle may show developments typical of unstressed syllables (similarly E dial. *ju*, showing the development **-o* > *-u*). Baltic **ō*, outside of initial syllables, seems to have been substituted as **a*, as in **nępat* 'nephew' and possibly **sęsar* 'sister' (see p. 147).[168]

1.3.3.2. Consonantism

Compared to the Slavic loans (see Kalima 1956), the Baltic loans predated several Early Proto-Finnic developments affecting the consonants: namely **š* > **h* (e.g. **haljas* 'green' ← Baltic **žaljas*), **ti* > **ci* (e.g. **silta* 'bridge' ← Baltic **tiÎta-*) and **tj* > **cc* (in **męcca* 'forest' ← Baltic **medja-*),[169] as well as the metathesis **-wR-* > **-Rw-* (e.g. **torvi* 'horn' ← Baltic **taurē*) and the development **-ln-* > **-ll-* (e.g. **villa* 'wool' ← Baltic **vîlnâ*). A few aspects deserve a more detailed discussion:

***š > *h; *s > *s**

It seems that Finnic speakers were able to reliably distinguish the two Baltic sibilants, consistently substituting Baltic **š* with **š*. The only

---

[168] There are, however, a couple of examples which show non-initial **i* and **u* in Baltic loanwords, namely **pęlu(t)* 'straw chaff' (adopted as **pęlęw* ~ **pęlâw*?), and **oinas* 'ram' (adopted as **owęjnas* ~ **owâjnas*?).

[169] Koivulehto 1986 (cf. also 1979a: 290 fn.) discusses a couple of convincing parallels among the Germanic loanwords: F *otsa* 'forehead', E *ots* 'end, front; forehead', Li. *võntsa* 'forehead' (← **anþja-,* cf. ON *enni* 'forehead') and probably F *maltsa*, E *malts* 'orache', Li. *mõltsõz* 'goosefoot, orache' (vocalism after *mõltsi* 'green'? Kettunen 1938: 222) (← **maldjō-,* cf. OSw. *mäld*; Ritter *apud* LÄGLOS II: 248), although admittedly the latter is of obscure origin (cf. Kroonen 2013: 351). Note also Sá. N *fihčču*, K *võhč* '(seal's) flipper' (< **fiččō* ← Norse **fitjo* < Germanic **fetjō-,* cf. ON *fit* 'webbed foot; flipper'.

apparent exception to this rule is PF **hanhi* ‘goose’, but this can be explained as the result of a rather trivial assimilation. In fact, P. Kallio points out to me (p.c. March 2023) that there are no old words with the combination **h–s* in Proto-Finnic, so that the assimilation might even be treated as regular. The final **h* (> *-Ø*) in F *herne* ‘pea’, *käärme* ‘snake’ can be of analogical origin. Due to the change **s* > **h* between unstressed vowels, stems in **-es* and **-eh* are indistinguishable in most oblique cases, e.g. GEN.SG. **-ehen*. Such a vacillation is also known in inherited words, e.g. F *kaarne* ‘raven’ beside Vp. dial. *karnįš* ‘crow’, Li. *kārnaz* ‘raven, crow’ (= Sá N. *gáranas*, Komi *kį̈rnįš* ‘raven’ ?< PU **karnəš* ~ **kärnəš*).

The exact dating of the change **š* > **h* is difficult, but the sibilant pronunciation must have been preserved until after the arrival of Finnic speakers in Fennoscandia. A layer of older Germanic loans show the substitution **s* → **h*, e.g. F *ahjo* ‘furnace, forge’ ← **asjō*, cf. Sw. *ässja*, OHG *essa* ‘furnace, forge’ (LÄGLOS I: 5–6); *keihäs* ‘spear’ ← **gaizas*, cf. ON *geirr*; cf. also Koivulehto 1984: 193–195. Furthermore, the earliest Sámi loans from Finnic still show **š* for Finnic **š* (cf. Sá. N *vašši* ‘hatred’ ~ F *viha* ‘hatred’ (< PU **wiša*; Aikio 2006a: 41).

Juho Pystynen (2016) has presented an argument which could show that this sound law even post-dated Proto-Finnic. The word *haaḱ* ‘goose’ in some peripheral South Estonian dialects (Seto, Lutsi, Kraasna) apparently shows a development **Vn* > **V̄* before **š*, which would seem to parallel the common South Finnic change **Vn* > *V̄* before **s* (cf. E *maasikas*, Võ. *maaśk* ‘strawberry’ < **mansikka*; Kallio 2014: 162). This might suggest that South Estonian originally preserved a sibilant **š* longer than the rest of Finnic, and the change to **h* only diffused into this dialect area at a later date. This remains highly tentative, however, especially since the equally peripheral Leivu and

Kraasna *vahn* 'old' (= F *vanha* < PU **wanša*, on which see *1.4.4*) would seem to contradict this sound law.[170]

Nevertheless, this observation might explain the occurrence of a couple of exclusively South Estonian loanwords which have undergone the sound change **š* > **h*, in particular Võ. *pahr*, (Hargla) *parh* 'boar' ~ Lt. *paršas* 'piglet, castrated boar' (= Lat. *porcus* 'pig') (Kalima 1936: 145) and perhaps Võ. *eherüs*, Mulgi *eerus* 'trout' ~ Lt. *ešerỹs* 'perch' (Ojansuu 1921: 5–6). Still, it remains possible that these indeed belong to the earliest layer of loanwords, and were merely lost after South Estonian split off from the rest of Finnic.

Initial *c- and *st-

Kallio (2007: 235, 241–242; 2014: 157) has argued in favour of reconstructing a phoneme **c* for Proto-Finnic on the basis of the South Estonian evidence. While there indeed do appear to be some compelling examples of South Estonian *-dś* (sporadically) reflecting Proto-Finnic **-ci* (< **-tə*), the status of this phoneme in initial position is less certain. Kallio's only example is the word for 'pig': F *sika*, E *siga* ~ Võ. *tsiga*, yet this word's etymology is uncertain, the traditional comparison with Mordvin **tuvə* (> E *tuvo*, M *tuva*) 'pig' (UEW 796) being phonologically irregular (Aikio 2015a: 46). Therefore, it cannot be proven that the South Estonian *ts-* goes back to an earlier **c-*.[171] Moreover, in all other cases where an initial **c-* would be expected on etymological grounds, we find *s-* in South Estonian:

---

[170] In view of this, it is perhaps preferable to opt for Pystynen's alternative account that South Estonian originally preserved a form **hansi* (← Baltic **žansî*) which first developed to **hāsi* and only then was assimilated to **hāhi*.

[171] It is possible that Võ. *ts-* is due to the secondary influence of Latvian *cūka* 'pig'.

- F *sinä*, E *sina* = Võ. *sina* ‘you (SG.)’ (cf. Kallio 2007: 242) < **tinä* (UEW 539)
- F *syvä* = Võ. *süvä* ‘deep’ < **tiwä* (UEW 525–526) — According to VMS, the word is practically limited to South Estonian and adjacent Tartu dialects, while in North Estonian, it is only found as a relic in the western periphery.
- F *sitkeä*, E *sitke*, Võ. *sikkõ* ‘tough, durable’ ~ Sá. N *dađgat* ‘firm (of body parts)’ (Sammallahti 1999: 74–75) — As the Võro term shows regular **tk* > *kk*, a loan from North Estonian is improbable.
- F *silta*, E *sild*, Võ. *sild* ‘bridge’ ← Baltic **tiĺta-* — Attested throughout all of South Estonian, including the language islands.

In the *Yhteissuomalainen sanasto* (YSuS) online database,[172] Kallio has adduced several other examples of **c-* based on correspondences between initial **ts-* in Võro and affricates in Karelian and Veps; however, the data encompasses at least seven distinct correspondence patterns between the three languages. Furthermore, the majority of the words are clearly onomatopoeic (e.g. K dial. *čivissä* ‘rustle’, Võ. *tsibisemä* ‘whisper’; Vp. *čiraita* ‘sizzle’, Võ. *tsirisemä* ‘buzz’) or belong to semantic fields where expressivization could be anticipated (K *čirkku* ‘small bird’, Võ. *tsirk* ‘bird’; Olonets *čongie*, Võ. *tsunǵma* ‘root about’). Kallio (2007: 242) has himself acknowledged that the Karelian and Veps data are largely irrelevant.

It is clear that the above data provide no evidence for a contrast between **s-* and **c-* in initial position in reconstructable Proto-Finnic. Given that it remains possible that **c-* > **s-* took place in initial position earlier than it did in intervocalic position, I prefer to

[172] Hosted at https://sanat.csc.fi/wiki/Luokka:Yhteissuomalainen_sanasto.

reconstruct the word for ‘bridge’ as **silta* (and not **cilta*) for Proto-Finnic.

~ ~ ~

The Finnish cognates of F *seiväs* ‘post, stick’ are interesting in two respects. Not only do they show reflexes of the unclear diphthong **?i* (probably = **ęi*, see p. 122–129), but also an unclear alternation between initial *t-* and *s-*. This correspondence has some clear parallels. Compare the following:

- F *seiväs*, Vt. *seiväz*, Võ. *saivas* ~ E *teivas*, Li. *tāibaz* ‘post, stake’ (~ Lt. *stíebas*)
- F *siipi*, Vt. *siipi*, Võ. *siib* ~ E *tiib*, Li. *tībõz* ‘wing’
- F *seipi* ~ E *teib* ‘dace’, Li. *teib* ‘ide’[173] (~ Lv. obs. *stiepats* ‘dace’; see p. 170)
- ? F *saparo* ‘short tail’ ~ Li. *tabār* ‘tail’ (?~ Lv. dial. *stebere* ‘tail’)[174]

The agreement between the words for ‘wing’ and ‘post, stick’ is striking: in both cases, the distribution between *s-* and *t-* is almost

[173] Nirvi (1961: 152) and Heikkilä (2013: 583) have adduced E dial. *taivikas* to support a Proto-Finnic reconstruction with **?i*. This form apparently derives from Wiedemann’s dictionary (*non vidi*; cf. Nuutinen 1987b: 109) where it occurs alongside numerous other variants (among which *teivikas* and *täivikas*). P. Kallio (p.c. March 2023) informs me that the form first appears in the second (posthumous) edition of his dictionary and is perhaps the result of a mere printing error; note also that Ariste (1975: 471–472) leaves out the variant with *-a-*. In any case, none of these forms are likely to be South Estonian, as VMS only records *teib* and variants across the north and on the islands, Although the variant in *-äi-* must be somehow secondary, the Proto-Finnic vocalism is quite possibly to be reconstructed as **äi*, anyway (cf. Kallio 2018: 261), even though **?i* cannot be excluded.

[174] But note F *sapa*, E *saba*(!) ‘tail’ (< **sapa*).

identical,[175] yet not geographically contiguous. If the reason was 'unstable' substitution strategies (Kalima 1936: 160) or independent loans (Kallio 2018: 258–259), we should expect a more or less random distribution. Since North Estonian and Livonian do not constitute a subgroup of Finnic languages, the fact that the same two languages happened to 'reborrow' the words in question is a remarkable coincidence. Nuutinen (1987c: 61) and Vaba (1997: 177) have suggested that the loanwords were adopted into an already dialectally diverse Proto-Finnic, but the fact that the vocalic reflexes of **ę̇i* (see p. 122–129) straddle the two groups makes this very awkward.

Heikkilä (2013: 586–587) has argued that the above evidence would prove that a cluster **st-* was licensed in Early Proto-Finnic. While a phonological solution would be welcome, the assumption that a phonotactical restriction against initial consonant clusters could have been relaxed in Early Proto-Finnic before being reinforced again later on, though not impossible, is certainly uncomfortable, especially given that there are no examples of the alleged Finnic **st-* among the Germanic loanwords. Nevertheless, in lieu of an alternative solution, I have used the notation **stę̇ipas* in this chapter.[176]

---

[175] While *saivas* is purely South Estonian, *siib* has apparently spread into neighbouring Central Estonian dialects (see VMS), and is also attested in northeast coastal Estonian, which must probably be attributed to influence from Votic and/or Finnish. Nevertheless, I think it is possible that the distributions of the words were originally identical.

[176] Considering that I do not reconstruct an initial **c-* in Late Proto-Finnic, one might consider that this is what underlies the correspondence **t-* ~ **s-*. However, this is chronologically problematic, as in my model **c-* (or rather **ć-*) would still have been present in Early Proto-Finnic, at the time of the contacts with Baltic. Furthermore, a palatal affricate **ć-* would be a phonetically unlikely substitution for a foreign **st-* (I thank Santeri Junttila for pointing these issues out to me).

Syllable structure

At the time of the contacts with Baltic, Finnic still seems to have had a fairly strict maximum syllable structure **CVC*. The avoidance of heavy clusters can be observed in **ahtas* 'narrow' (← **aNštas*), **takja* 'dense' (← **tańkjV*-) and **sikla* 'side net in a seine' (← **tinkla*-), which show the regular loss of a nasal before two consonants (cf. PF **kanci* 'cover' ~ **kat-ta*- 'to cover'; Posti 1953: 56–59).[177] In the case of **morcijan* 'bride', which reconstruction seems to be confirmed by North Karelian *moršien* (contrast *hoaśśa* 'hayrack' < **hāsja* ?← Sw. *hässja*, LÄGLOS I: 62), an epenthetic vowel appears to have broken up the heavy cluster **-rtj*-. In PF **tühjä* 'empty', a similar cluster **-štj*- was resolved to **-šj*-. The single example of **vōtna* 'lamb' is problematic, because it appears that *CR*-type clusters could not occur after long vowels even in relatively recent loanwords. Although a convincing explanation is lacking, it is potentially relevant that the essive form **vōt-na* 'year ESS.SG.' (> F *vuonna* 'in the year'), where the long vowel might have been restored early on due to analogical pressure from other case forms, would have been formally identical to the word for 'lamb'.

Phonotactic constraints also explain the rarity of geminates, which are typically found as substitutions for voiceless stops in Germanic and Slavic loanwords. In the material collected in *1.3.2*, only two contain a geminate: **rattas* 'wheel' and **tüttär* 'daughter', yet only a handful of others (e.g. **atęivo* 'visiting relatives', **hako* 'branch' and **nępat* 'nephew') could have theoretically tolerated a geminate in Early Proto-Finnic. Steinitz (1964: 337) has proposed that the examples with

[177] Note that Posti does not adduce this Baltic evidence and considers the possibility of a very early dating for this change. An early dating is not excluded by the Baltic evidence, as the sound change may have been productive over a long period. Aikio (2022: 11) even reconstructs this rule for Proto-Uralic.

geminates represent a younger layer, while Junttila's (2017) explanation is that geminates were restricted to disyllabic stems. On the basis of such limited data,[178] it is difficult to draw any firm conclusions.

A substitution *t → Finnic *tt seems necessary to account for PF *$tüttär$ 'daughter' ← Baltic *$duktēr$. As Proto-Finnic possessed a cluster *-kt- (> Võ. -tt- and -ht- elsewhere; Posti 1953: 38–43; Sammallahti 1977: 133; Kallio 2014: 156), it is unclear why we do not find reflexes of *-kt- in this word. Posti (1953: 45) has suggested that the substitution strategy was conditioned by the position of the stress in Baltic,[179] but Kallio (2007: 237) sticks to the view that *$tüttär$ shows an "exceptional" development from earlier *$tüktäri$. Since the evidence does not permit the reconstruction of *-kt- at any stage in Proto-Finnic, it seems necessary to assume that Baltic *$dukter$- was perceived as */tü(k)ttär/ by Finnic speakers, and realized as *$tüttär$ when subjected to Early Proto-Finnic phonotactics.[180]

A number of loan etymologies have been proposed in the literature which show a geminate after a heavy syllable, such as the following examples in Thomsen (1890: 74), and Kalima (1936: 53):

---

[178] Junttila, of course, uses a larger corpus, but besides two new proposals, the only other example with a geminate he classed as certain (cf. 2017a: 142) is *$vakka$ 'wooden container' (on this, as well as *$hakkaita$-, see p. 174–176). The "probable" etymology F-K *huttu* 'flour porridge' ← Lt. *šùsti* (3PRET. *šùto*) 'stew, steam, sweat' is a mere root etymology, as a word of appropriate meaning is not attested in Baltic. He also (2017a: 141–142) proposes to compare F obs. (18th century hapax?) *lappa* 'thin plate', Vp. *lapak* 'flat, shallow' with Lt. *lãpas* 'leaf'; on F *kukka* and variants, I refer to his discussion (idem: 134–137).

[179] Note in this context that *-kt- > *-tt- is apparently regular in Finnic after unstressed syllables; cf. F *sädettä* 'ray, beam PART.SG.' (< *$sädek$ + *-tA).

[180] On the Sámi words for 'daughter', see fn. 283.

- F *laukki* (dial. *laukas*, *laukko*; K dial. *laukka*), Li. *laik* (< **laukki*) ‘blaze; blaze-faced animal’ ?← Lt. *laũkas*, Lv. *làuks* ‘blaze-faced’
- F *pirtti* ‘cabin’, Vp. *pert’* ‘house, cottage’ ?← Lt. *pirtìs*, Lv. *pìrts* ‘bath-house’
- F dial. *kääppä*, Vt. *tšääppä*, E *kääbas* ‘burial mound’ ?← Lt. *kãpas*, Lv. *kaps* ‘grave, burial mound’

While these etymologies are semantically strong, Nikkilä (1982: 254) emphasizes the rarity of such a syllable structure among the Baltic loan material, an observation which has recently found support in an extended treatment by Junttila (2019), who has argued that in fact no such words can be counted among the early Baltic loans. Since Nieminen (1953), *pirtti* has generally been regarded as a Slavic loan (although see p. 233–236 for a detailed discussion). As for *laukki* (etc.), Junttila (2019: 61–62) argues that the relevant forms should be seen as inner-Finnic derivatives of a more primary **lauka*, cf. F dial. *laukama* ‘bare patch (of land, fur)’.

With regard to *kääppä*, Junttila (2017: 133; 2019: 55) has stated that the vowel results from a secondary lengthening. This cannot be accepted: ‘sporadic’ secondary lengthening is only observed under specific conditions (cf. T. Itkonen 1987), and cannot simply be invoked as a license. If Finnic **kãppä* goes back to Early Proto-Finnic, it would have to reflect a trisyllabic preform **käŋəppä* or **käxəppä* with contraction of the vowel sequence, as in **kāri* ‘curve; rib of a boat’ < **kᵩŋərə* (UEW 126; Aikio 2015a: 58).[181] The similarity with the Baltic word is therefore probably coincidental.

[181] Whatever reconstruction we use, it is about time we abandon the comparison with Mansi (North) *kᵩp*, (East) *käp* ‘small hill’ (still repeated in SSA I: 484; van Linde 2007: 84; Junttila 2017: 133); Mansi **ä* implies Proto-

1.3.3.3. Declinations

Reflection of Baltic *-*s*

The nominative ending *-*as* of the Baltic *a*-stems is abundantly reflected in Finnic, e.g. **hampas* 'tooth', **oinas* 'ram', **rattas* 'wheel', **stę̆ipas* 'post' and perhaps **vapsas* 'wasp'. This is also the case for the adjectives **ahtas* 'narrow' and **haljas* 'green', which are evidently based on Baltic masculine nominative singular forms. There are, however, several words which show no trace of *-*s*. These fall into the following categories:

- Words with suffix replacement: **kę̆ltainę̆n* 'yellow' (fn. 105), **herhiläinen* 'hornet' (cf. **mehiläinen* 'bee', **kimalainen* 'bumblebee', Nieminen 1934: 32–35), E *vapsik* 'hornet', and the *i*-stems E *kurt* 'deaf', dial. *kärv* 'snake' (cf. Nieminen 1944: 249)
- The adjectives **tühjä* 'empty', **līka* 'surplus', which could equally be based on feminine or, more probably, neuter (~ predicative) forms. Behind **takja* 'dense' perhaps lies a ACC.SG.F. **tan̂kjan*, or an adverbial form **tan̂kjai*.[182]
- Nouns: **sikla* 'side net in a seine', **silta* 'bridge', **hę̆ina* 'hay', **hirvi* 'elk', **häici* 'flower', **härkä* 'ox', **mę̆cca* 'forest', **pahla* 'rod, spit', **sōla* 'salt'.

Uralic **i*, **e* or **ü* in the initial syllable. Futhermore, this word can hardly be separated from Mansi (North) *kę̆mp*, (East) *kämp* in the same sense, thus suggesting a Proto-Mansi **kämp*.

[182] Although the *u*-stem adjectives in Bretke appear to have been largely unspecified for gender, even here there is some level of syncretism with the *jă̆*-stems (Specht 1932: 276–279), and due to the overall transfer of original *u*-stems to *ja*-stems in Latvian, this tendency is probably to be dated to Proto-East-Baltic, at least. On the adverbial suffix *-*jai* applied to old *u*-stems in Latvian, see Endzelīns (1923: 461–462).

It has often been suggested (Thomsen 1890: 112; Būga 1924b: 104, fn. 4; Nieminen 1944: 243–248; Иллич-Свитыч 1963: 42–44; Kortlandt 1993: 47) that the last group of words reflect Balto-Slavic neuters. There is indeed independent support for a neuter in two cases: Baltic **šę̄nas* 'hay' is cognate with the neuter OCS сѣно 'grass, hay' and **tinklas* 'net' contains the neuter instrument suffix *-*klas* found in Pr. E -*clan* (*piuclan* 'sickle') and Slavic (e.g. Czech) -*dlo*. In addition, several scholars (Иллич-Свитыч 1963: 78; Kortlandt 1993: 47; Derksen 2015: 466) have compared **silta* 'bridge' with the neuter Skt. *tīrthá-* 'ford; descent to the water'. However, this is a false comparison, as the Sanskrit word rather belongs with *tárati* 'pass, cross; overcome' < **terh₂-* (= Hitt. *tarah-*^{zi} 'overcome'); see EWA I: 650; ALEW 1277.

It is difficult to evaluate this evidence. First and foremost, there is no independent evidence that the East Baltic nominal neuter ending was originally *-*a*, matching Slavic, and not *-*an*, matching Prussian.[183] Furthermore, in some cases, the absence of the *-*s* in Finnic is the only evidence adduced in favour of an original neuter, which runs the risk of circularity. For instance, of the unsuffixed cognates to Lt. *paršas* 'piglet, castrated boar', only OHG *farah* (NOM.PL. *farhir*) is neuter, where we might consider analogy after *lamb* 'lamb', *kalb* 'calf', while OE *fearh* (PL. *fearas*) 'young pig' and Lat. *porcus* are masculine. Given this ambiguity, it can hardly be stated (with Иллич-Свитыч 1963: 48) that Võro *pahr* would prove an original neuter for Baltic.[104] Besides, an original

[183] In Lithuanian, the originally pronominal ending -*a* occurs in predicate adjectives, but this does not imply that it was present in nouns, as it is logical that ending would have first spread to adjectives; cf. the secondary spread in Pr. III *sta wissa* 'das alles' (cf. also *zuit* 'genug', with expected apocope; PKEŽ IV: 273). In Prussian, there is also some evidence for this ending in participles in predicate function, e.g. Pr. III *israṅkīt postāt* 'erlöset werden' (Endzelīns 1944: 199).

masculine gender is secured for East Baltic *žirĝas* ‘male horse’ not only in view of the meaning, but also by Pr. E *sirgis*.[185]

Moreover, there is an alternative available. Already Thomsen (1890: 112), followed by Kalima (1936: 78), has suggested that Finnic forms lacking *-*s* may have been abstracted from other Baltic case forms. The most obvious option that comes to mind is the genitive singular *-*ā*, although one has generally taken the accusative *-*an* as the most probable basis. Иллич-Свитыч (1963: 42) denies the latter possibility, stating that final *-*n* would not have been lost in Finnic by sound law, but he does not consider the possibility of analogy. A Finnic form such as **hę̄inan* (← **šę̄nan* ACC.SG.) could easily have been apprehended as a genitive-accusative singular form, on which basis a new nominative such as **hę̄ina* could have been backformed. An accusative source form must be assumed at least for PF **morcijan* ← Baltic **martjan* ACC.SG., where the final *-*n* has not undergone reanalysis as an oblique form in Finnic, but has instead assimilated into the **me*-stems, the only common category with nominatives in *-*An* (e.g. **sütän*, **sütäme*- ‘heart’).

---

between neuter **kaûša* and masculine **kaûšas* within Baltic (Иллич-Свитыч 1963: 82; cf. Derksen 2015: 234).

[185] More counter-evidence could be retrieved from other widely accepted loan etymologies in *-*a*, which have not been mentioned here due to the lack of an Indo-European etymology. The proposed Baltic sources of Finnic **ätälä* ‘aftermath’ (see *1.3.5.3*) and **kataka* ‘juniper’ (see p. 150–152) have masculine cognates in Prussian, viz. Pr. E *attolis*, *kadegis*, and **vakja* ‘wedge’ (see p. 98) is masculine in Germanic (cf. OHG *weggi*). Иллич-Свитыч (1963: 128–129) assumes an original neuter for **vaha* ‘wax’ due to the Germanic evidence (which is not regularly cognate, see p. 344–345), but the evidence he adduces for accent paradigm (b) — which he would predict in Slavic in the case of an original neuter — is marginal; almost all the evidence points to accent paradigm (c) and therefore an original masculine (cf. Зализняк 1985: 137).

Finnic **hirvi* ‘elk’ is not easily explained on the basis of a Baltic *a*-stem, and may instead reflect a Baltic feminine **širvē* as in Lt. *vìlkė* ‘she-wolf’ to *viĨkas* ‘wolf’ (Nieminen 1940: 378); similarly, Finnic **vōhi* ‘goat’ may well be from a feminine **âžē* as attested in Pr. E *wosee* (Thomsen 1890: 205). The form **häici* ‘flower’ perhaps likewise presupposes a different formation (such as **žaîdē* instead of **žaîdas*), but this cannot be supported by any Baltic-internal data.

The Finnic reflexes *-*jas* (**ankęrjas* ‘eel’, **haljas* ‘green’) and *-*es* (**hernes* ‘pea’, **kirves* ‘axe’) seem to echo the dichotomy between the East Baltic nominative *-*īs* (e.g. **žirnīs* > Lt. *žìrnis*, Lv. *ziřnis* ‘pea’) and *-*jas* (e.g. **žaljas* > Lt. *žãlias*, Lv. *zaļš* ‘green’; cf. Thomsen 1890: 114–117; Kalima 1936: 79–80). While Lt. *ungurỹs* ‘eel’ belongs to the former category, Pr. E *angurgis* might presuppose the existence of an earlier **angurjas*.[186] Admittedly, the word for ‘pea’ may have arisen from an earlier *i*-stem (cf. Nieminen 1957: 206; Skardžius 1941: 53), although no trace of this is found in the Baltic languages. A probable original *i*-stem is **kärmes* ‘snake’.

It appears that the distribution between *-*jas* and *-*īs* is essentially related to syllable weight. While the variant *-*īs* is clearly the productive type and occurs with after all syllable structures, almost all nouns with a nominative *-*jas* which can be reconstructed back to Proto-East-Baltic have a light first syllable. The distribution thus

[186] The grapheme ⟨g⟩ in the Elbing Vocabulary, in its function as representing a glide, only occurs after stem-final resonants in cases where East Baltic shows NOM.SG. *-*jas*; cf. *wargien* ‘copper’, *kragis* (read **kargis*) ‘army’, *saligan* ‘green’. Contrast Pr. E *tuylis* ‘boar’ (~ Lt. *kuilỹs*), *singuris* ‘goldfinch’ (?~ Lv. *žīguris* ‘sparrow’). Lithuanian *ungurỹs* may be the result of a general preferrance for *-*īs* in polysyllabic words (cf. Lt. *kumelỹs* ‘colt’ beside Lv. *kumeļš* ‘colt; male horse’).

corresponds more or less exactly to that of Gothic *-jis* and *-eis* (Sievers' law; see Sommer 1914: 242 and *passim*):[187]

- **karjas* 'war, army' (> Lt. obs. *kãrias*, Lv. *kaŗš*)
  = Pr. E **kargis*, Go. *harjis* 'army
- **keljas* 'way' (> Lt. *kẽlias*, Lv. *ceļš*)
- **kraujas* = */kravjas/ 'blood' (> Lt. *kraũjas*)
  = Pr. III *krawia* 'blood', Skt. *kravyá-* 'bloody'
- **medjas* 'forest' (> Lt. *mẽdžias*, Lv. *mežs*)
  = Pr. E *median*, Go. *midjis*, Skt. *mádhya-*
- **teljas* 'calf' (> Lt. *tẽlias*, Lv. *teļš*)
- **svetjas* 'guest' (> Lt. *svẽčias*, Lv. *svešs*)
- **varjas* 'copper' (> Lt. obs. *vãrias*, Lv. *vaŗš*) = Pr. *wargien*

In view of this distribution, it is likely that **-jas* and **-īs* both reflect the same proto-form (i.e. **-ios*).[188] Usually, one has assumed that **-īs* went through an intermediate stage **-ijas* (Sommer 1914: 227), and proof of this has been seen in Estonian dial. *takijas* 'burdock' ← Baltic

[187] Lt. *-ias* has become somewhat productive in adjectives (particularly after dentals, apparently to avoid consonant alternations such as *t* : *č*, occurring in *u*-stems?). Nevertheless, a similar tendency can be observed here as well; note Lt. *šlãpias*, Lv. *slapjš* 'wet', Lt. *žãlias*, Lv. *žaļš* 'green', Lt. *naũjas*, Lv. *naujš* 'new' (= Go. *niujis*). Forms such as *mẽdis* 'tree' beside Lt. dial. *mẽdžias* 'forest' and Žem. *svetỹs* 'guest' beside *svẽčias* must result from analogy (Būga RR II: 509).

[188] Taking **-īs* from **-iHo-* per Hill 2016: 214 is unnecessary, and moreover, the contraction **-ijV-* > **-jV-* in the oblique cases would be irregular (compare uncontracted Lt. *eldijà* 'dugout canoe' which corresponds exactly to OCS ладии 'boat'). It seems more likely that a satisfactory solution can be found starting from an model based on syllable weight. At first sight, the correspondence between Lv. *âzis*, beside GEN.SG. *âža* with Gothic *hairdeis* (< **-ijas*) beside NOM.PL. *hairdjos* is remarkable. For Baltic, we may suggest that the development **-j-* > **-ij-* only occurred after a heavy syllable and before a short vowel, with later generalization of **-j-* in the oblique cases.

**dagīs* (Stang 1966: 190; Zinkevičius 1980: 217–218; Kortlandt 1977: 324; 2018: 182). However, the analysis of the Finnic form is somewhat problematic.

First of all, it appears that Finnic **-ja-* and **-ija-* were likewise in complementary distribution, whereby the disyllabic reflex was automatic after a heavy syllable (Ritter 1977; see **morcijan* 'bride', discussed above). The preserved *-k-* in Estonian *takjas*, dial. *takijas* 'burdock' shows we are dealing with an original geminate, which can also be seen in F dial. *takkiainen*, Vt. dial. *takkiaz*, *takkiain*, K *takkis-heinä* 'burdock'. The disyllabic suffix in Finnic **takkijas* can thus be viewed as a symptom of the weight of the initial syllable, and does not directly have any bearing on the reconstruction of the Baltic word.

The question remains, however, as to why the Finnic word has a geminate in the first place. This would certainly not be expected on the basis of Baltic **-g-*, which could only usually be substituted by a single **-k-*. The usual explanation (Thomsen 1890: 231; SSA III: 258) is that the Finnic word has been influenced by the verb **takkista-* 'stick, hinder'; compare OE *clīfe*, OHG *klība* 'burdock' (< *clīfan*, *klīban* 'adhere, stick'). In this case, the introduction of the stem **takk-* into a borrowed form **takjas* would have automatically resulted in **takk-ijas* due to the aforementioned phonotactic rules. This would allow us to assume Baltic **dagīs* is in fact secondary for **dagjas* (with the expected suffix variant after a light syllable).

On the other hand, the assumption of contamination is never exactly compelling. A possible alternative solution presents itself if we indeed start from **dagijas*, namely that the introduction of a geminate in Finnic was necessitated by the inadmissability of the sequence **-ija-* after a light syllable. However, we must admit that other examples of Baltic **-īs* do not show any evidence of an earlier **-ijas* (Sommer 1914:

228; Kalima 1936: 79–80); cf. **hernes* ‘pea’ and **kirves* ‘axe’ noted at the start of this section.

Vocalic stems

As touched on above, the Baltic feminine *ā*-stems were generally adopted in Finnic as *a*-stems, cf. **ansa* ‘loop’, **halla* ‘frost’, **karva* ‘(animal) hair’, **rę̄una* ‘edge’, **tapa* ‘way, custom’, **villa* ‘wool’, **lūta* ‘broom’, **takla* ‘tinder’. On the other hand, there are a few examples which appear as *o*-stems, cf. **hako* (beside **haka*) ‘branch’, **hę̄imo* ‘tribe’, **putro* ‘porridge’, **vōro* ‘turn’.

As for **hę̄imo*, it has been suggested its stem vowel represents an inner-Finnic development. Since the Finnic *a*- and *o*-stems coincide in the oblique plural, Nieminen (1934: 19) has suggested that an analogical shift to an *o*-stem might have been encouraged by the frequent plural use of the word in the sense ‘relative’. He supports this with some alleged traces of the original *a*-stem in Karelian dial. *heima-kunda* ‘tribe’ and Võ. dial. *hõimanõ* ‘relative’ (absent from VMS).

A similar account seems to be required to explain *ativo* ‘visiting relative’ beside the *a*-stem *atima*;[189] cf. the collocation *olla ativoissa* ~ *atimoissa* ‘visit relatives’, lit. ‘to be in guests’ (SMS), which is ambiguous between an *a*- and *o*-stem. A similar explanation could perhaps account for the co-ocurrence of *haga* and *hagu* ‘branch, stick’ in Estonian (but cf. Junttila in prep. s.v. **hako*). For **hę̄imo*, another account may be to assume a Baltic accusative form **šeiman* as a source, with the otherwise attested substitution **a* → **o* (note also the labial environment; see p. 119). Differently on the *o*-stems, see Holopainan, Kuokkala & Junttila (2022: 126–130).

[189] With sporadic dialectal *-*v*- > -*m*-, cf. Nikkilä 1999: 14–17.

The Baltic *ē*-stems appear to have been adopted either as **A*-stems (**mäntä* 'whisk'; perhaps **tūra* 'ice chisel', cf. Lv. *dùre*), or as **E*-stems (cf. **torvi* 'hunting horn' ← Baltic **taurē*). The latter substitution may also be accounted for by assuming a loan based on the ACC.SG. **-en*.

The category of feminine *jā*-stems with a NOM.SG. **-ī* is now only represented by two common Lithuanian nouns, namely *patì* 'wife' and *martì* 'son's wife; bride'. This group was originally a larger, however. In the Elbing Vocabulary, twenty-five words are attested with a nominative in *-i* (Levin 1973; 1974: 48–49). The only other loanword for which such a nominative in attested is **žansī* 'goose' (= Pr. E *sansy*), which form might directly account for Finnic **hanhi*. Note that the Finnic *i*-stems apparently did not exist at the time of the Baltic loans, so that an *e*-stem would be the closest match (Junttila 2015: 18–19). However, the word for 'goose' shows good evidence for an earlier consonant stem (Nieminen 1957: 200–201; Zinkevičius 1966: 266), and in either scenario, it is difficult to rule out an *i*-stem accusative **žansin* as the basis for borrowing.

Consonant stems

F obs. *nevat* 'nephew, niece' (whence Sá. N *neahpát*, S *neapede* 'sister's son or daughter') apparently belonged to the same inflectional class as F *kevät* 'spring', GEN.SG. *kevään*. It must have been loaned from Baltic **nepōt-s* (most probably on the basis of ACC.SG. **nepōti-n* → GEN.SG. **nepatę-n*; see below). For **sęsar* 'sister', a bolder solution is required. We could start from an earlier oblique form **sesari-n* (compare secondary Lt. *sẽserį*, after *dùkterį* 'daughter'?), matching Skt. *svásāraṃ* ACC.SG. On the other hand, it would also be possible to start from a Baltic nominative singular **sesōr*. The loss of final resonants has often been dated very early (Schmalstieg 1983: 152–154; Jasanoff 2002: 34–35), but the Slavic evidence suggests a fairly recent loss (Kortlandt

1979b: 264, 1983; Pronk 2018: 301), and there is no clear argument as to why it should be early in Baltic, either. Note that some forms such as *pirmuonìs* ‘forebear’ (in Daukša a consonant stem, cf. *pirmůnés* GEN.SG.), *schirfchonis* ‘hornet’ (Bretke; see ALEW^2), and others, look to be built on nominatives in *-*ōn*, suggesting the loss of final resonants in fact occurred not long before the historical period.[190]

The words **sēmen* ‘seed’ and **paimę̄n* ‘shepherd’ are nevertheless better derived from an oblique stem in *-*en*-. As with the examples described in the previous section, a plausible source may be the accusative **sēmenin*, which would be adopted as Early Proto-Finnic **sā̆menen*, on which basis a new nominative singular **sā̆meni* > **sēmen* could have been backformed. In the case of **tüttär* ‘daughter’, both a nominative **duktēr* and an accusative **dukterin* could come into question.

1.3.3.4. The dialectal origin of the Baltic loans

It has been noted that the Baltic loans in Finnic seem in certain aspects more akin to West Baltic than East Baltic (cf. Nieminen 1957: 188; Vaba 1998: 182–184; Kallio 2008a: 275), and it has been suggested the loans were adopted from some other unknown Balto-Slavic dialect (Junttila 2016b), or at least partly adopted from Proto-Balto-Slavic itself (Kallio 1998: 212, 2008a: 265; Koivulehto 1999: 9–11).

The evidence of a particular connection to Prussian is not strong. I have accepted two etymologies where the source form is only found in Prussian: **hirvi* ‘elk; deer’ (~ Pr. E *sirwis*) and **vatnas* ‘ploughshare’ (~

[190] Note, however, that the southern Žemaitian forms entered in the LKŽ under *šuõn*, *vanduõn*, *piemuõn* show a secondary development (Zinkevičius 1966: 196–197).

Pr. E *wagnis*); however, these do not represent West Baltic innovations, and may well once have existed in East Baltic, too. Based on the inflection, PF **vōhi* 'goat' also stands somewhat closer to Prussian (~ Pr. E *wosee*), but the Prussian form represents an archaism (in East Baltic we find innovative forms: Lt. *ožkà* < **âž-(i)kā-*; Lv. *kaza* ← R *коза́*), and it cannot be excluded that suitable forms were previously present in East Baltic (cf. Endzelīns 1933: 80–81). A similar argument can be put forward with regard to Pr. E *angurgis* ~ Lt. *ungurỹs* 'eel'.

On the other hand, there are some forms which betray innovations that are limited to East Baltic:

- The form **ahtas* 'narrow' reflects an innovative form with the adjective suffix **-stas*. This suffix has been somewhat productive in East Baltic (Skardžius 1941: 324–325), but not elsewhere in Balto-Slavic,[191] and Slavic continues a more archaic *u*-stem **ǭzu-ka-* > OCS жзъкъ 'narrow, tight' (= Skt. *aṃhú-*).
- Both **kärmes* 'snake' and **hernes* 'pea' reflect semantic shifts which are only attested in East Baltic, cf. the more archaic meanings of Pr. E *girmis* · made 'worm, maggot', and Pr. E *syrne*, OCS зрьно 'grain' (= Lat. *grānum*).
- If my analysis of **līka* 'surplus' and **hę̄ina* 'hay' as showing direct substitutions for Baltic **ę̄* can be upheld, this would be a strong argument in favour of a specifically East Baltic origin for the Finnic loanwords.

[191] In Prussian it is found in one form, III *angstainai* 'in the morning' (cf. Lt. *ankstì* 'early'). However, according to Petit (2005), this word is derived from the verb attested in Lt. dial. *ant-stóti* 'to begin', cf. Lt. *apstùs* 'abundant' to *ap-stóti* 'surround', *atstùs* 'distant' to *at-stóti* '(obs.) move away'. The suffix may therefore not be akin to that of Lt. *añkštas* 'narrow'. The form III *auckstimiskan* 'Obrigkeit' is an error: all 8 other attestations show *au(c)kt-* (PKEŽ I: 113).

I therefore consider the most likely source of the Finnic loanwords to be an East Baltic dialect. It still remains probable that the source of the Finnic loanwords was not a direct ancestor of the extant Baltic languages. One possible argument for this is the evidence for a dialectal lowering **i* > **e* after **š*, **ž* and before **R* (see p. 120–121). A further indication is the lack of any evidence for early Finnic loanwords in the attested East Baltic languages, as will be argued in the following section.

1.3.4. Loans from Proto-Finnic to Proto-Baltic?

As with loans from Baltic to Finnic, the only reverse loans which could be considered certain are those with regular Uralic cognates, particularly in non-adjacent branches. In most cases where a Finnic to Baltic loan has been suggested, it has been done so on this basis, although for the most part the suggested comparanda predate our modern understanding of Uralic sound changes, and cannot be upheld.

An exemplary case is the word for 'juniper', attested in Lt. *kadagỹs*, Lv. dial. *kadags*, *kadẹgs* (ME II: 131), Pr. E *kadegis* and F *kataja*, E *kadakas*. Setälä (1909) connected the Finnic words with a plethora of Uralic material, which led Kalima (1936, cf. p. 12) to exclude the word from his treatment of the Baltic loanwords. The idea that the Baltic word should be derived from Finnic became quite pervasive in the literature, at least among Uralicists (SKES 170; Rūķe-Draviņa 1955: 404–409; Kiparsky 1959b: 424; Bednarczuk 1976: 48; UEW 165; cf. SSA I: 326–327). Already Collinder (1955: 79) noted that the Finnic vocalism was problematic, and was sceptical towards the etymology; however, UEW still accepted a link with the Sámi and Mansi material. In reality, there are clear phonological problems with all of the Uralic

comparanda (see also van Linde 2001: 288–290). Here I present the data along with the possible PU reconstructions:[192]

- Finnic **kataka* 'juniper' < **kaTaka* / **kęTaka* (where **T* = **t*, **d*, **d̓* or **č*)
- Sámi N *goahcci*, Sk. *kuä´c'cev* 'conifer needle < **kočča̮wa* / **kačča̮wa*
- Komi dial. *kač-pomeľ* 'juniper' < ?**käččV* / **käčkV*
- Mansi (East) *kööasp* 'juniper' ~ (West) *kā̈šäp* < **käČ(k)V-* ~ **ke/iČ(k)V-* (where **Č* = **ć* or **č*; the Mansi forms do not regularly correspond to each other)

The words indeed bear a certain similarity, but they cannot be related by sound law. Only the Komi and Eastern Mansi forms could theoretically be cognate, but since the word is irregular within Mansi, and the development **ä* > Komi *a* (cf. Aikio 2021: 167–168) is somewhat dubious, this is most probably due to chance. In any case, the Finnic word cannot be related to any of the others.

Mikkola (1930: 442) presented another argument in support of the word being native to Finnic, namely the suffix **-aka*. This suffix is present in other tree names, e.g. F *pihlaja*, E *pihlakas* 'rowan tree' (< **pićlä* ~ **pećlä*, UEW 376), F dial. *petäjä*, Li. *piedāg* 'pine' (< **pečä*, UEW 727). According to Mikkola, the suffix stands quite alone in Baltic. However, he overlooked an important example. Lt. *mẽdžiaga*, which now means 'material', is preserved in older texts and Belarusian language islands in the sense 'tree; wood'. The original form is probably **medaga*, cf. Lv. dial. *mędaga* 'timber', while *mẽdžiaga* shows

[192] I omit Sámi **kęsŋęs* (> S *gasnges*, N *gaskkas*) 'juniper', already considered doubtful by Setälä, and Mari E *lume-ɣož*, W *lə̑me-kož* 'juniper' (TschWb 352) in which the second element is simply *kož* 'pine' (UEW 165), cognate with Finnish *kuusi* 'pine'.

the influence of the root word *mẽdis* (GEN.SG. *mẽdžio*) ‘tree; wood’. Since, in each case, the suffix has clearly been added within Finnic (being absent from the other Uralic comparanda),[193] one may ask whether this ‘tree suffix’ *-*aka* was actually imported from Baltic (or from somewhere else).

Another word for which a Finnic → Baltic loan is often assumed is Lt. *šẽškas*, Lv. *sẹsks* ‘polecat’ ~ K dial. (Olonets) *hiähky*, Vp. *hähk* ‘mink’ (Wichmann 1911: 253; Kalima 1936: 102–103; Kiparsky 1949: 46–47, cf. Kiparsky 1972; Mägiste 1959: 171; ALEW 1179). This was the only loanword of this type positively assessed by Junttila (2015a: 27), who stated “the sound correspondences between the Uralic words are flawless”.

However, this is clearly not the case.[194] Mari E *šaške*, W *šäškə* ‘mink; otter’ reconstructs to PMa. **šäškə*, while PMa. **ä* has no regular origin and is not usually found in inherited words (E. Itkonen 1953: 203–207; for a more detailed discussion of the Mari word, see p. 237–238). The Samoyed comparanda, Tym Selkup *tȭt*, Kamas *ća’n* (= ⟨ťśą̧’ǹ̥⟩, Donner/Joki 76) ‘otter’, are justifiably rejected by Aikio (2015a: 45): Selkup **ȭ* could perhaps reflect Proto-Samoyed **oj*, judging by Taz Selkup *tȭti̯-* ‘vomit’ (< **tojtə̂ apud* Janhunen 1977: 164–165), but cannot be squared with the other Uralic data, nor can such a reconstruction even account for the Kamas form. The addition of Mator

[193] In the case of **pihlaka*, the unsuffixed form is widely preserved: Vp. *pihľ* (GEN.SG. -*än*), Vt. (Цветков) *pihl-puu*, E dial. (insular) *pihl*, Võ. *pihl*.

[194] Junttila still defended the Baltic origin with the argument “there are no less than three possible Baltic derivational explanations for Lith. *šeškas*”. This would rather speak for the opposite: if scholars cannot agree on the origin of the Baltic word, then probably none of the proposals are fully satisfactory. This is indeed the case: all proposals mentioned present semantic and phonological issues, cf. ALEW loc. cit.

tit ‘otter’ to this cognate set by Helimski (1997a: 362) only complicates matters, as the development **ä* > Mator *-i-* is only supported by dubious examples (cf. idem: 99). The invalidity of these cognates was later also admitted by Junttila (in prep. s.v. *hähkä*).

Bednarczuk (1976) has suggested that a whole host of other comparisons represent loans from Uralic into Baltic. As Junttila (2015a) has already written an extensive article criticizing Bednarczuk’s views, and it seems that his conclusions can generally be upheld,[195] I will limit myself to the examples which have plausible cognates in other West Uralic branches:

(a) **‘lake’.** F *järvi*, E *järv* ‘lake’ ~ Lt. *jáura* ‘boggy soil which cracks and dries out in the summer’ (LKŽ); cf. Sá. N *jávri*, Sk. *jäu´rr* (< **jāvrē*); Md. E *eŕke*, M *äŕʻkä* (< **ärkə*; **-kə* is a diminutive suffix); Ma. E *jer*, W *jär* ‘lake’ (< **jer*, cf. Aikio 2014b: 135–137) — This loan was first suggested by Būga (1908: 95; 1922: 238–241), although it was not until its independent discovery by Nuutinen (1989) that it received widespread acceptance among Uralicists (Sammallahti 1998: 249; van Linde 2007: 45–46; Junttila 2012: 281; Aikio 2012a: 107).

Most reference works (SKES 132; UEW 633; SSA I: 259) have considered *järvi* to be a native Uralic word. Indeed, a reconstruction **jäwrä* (e.g. Sammallahti 1998: 249) can account for most of the data. The metathesis **wr* > **rv* in Finnic is regular (cf. Koivulehto 1979a: 279).[196] The loss of the initial glide in Mordvin is paralleled by Md. E *ej*,

[195] I would like to point out that the claim that “a Finnic two-syllable *a*-stem cannot be dated [to] PU if it has a long vowel in the first syllable” (2015: 20) is accurate only for pre-Proto-Finnic, but not for reconstructible Proto-Finnic, in which long vowels can occur in such an environment if they result from contraction, as in e.g. F *pyörä* ‘wheel’ < **pi/eŋärä* (cf. also Plöger 1982).

[196] Prior to this, the standard reconstruction was **järwä*, but the assumed metathesis in Sámi would be *ad hoc*.

M *äj* (< **jäŋə*) ‘ice’ and E *ezńe*, M *äźńä* ‘joint’ (< **jäsən*), cf. Bartens (1999: 46).[197] The loss of **w* in Mordvin is probably paralleled by Md. M dial. (Penza) *śeńi* ‘a kind of fish, ?ide’ (< **sewnə* ~ **säwnə*, UEW 437–438), while the same development can potentially be posited for Mari, cf. *tić* ‘full’ (< **täwdə*).

The only irregularity is the stem vowel: while Sámi and probably Mari point to **ä–ä*, Finnic unequivocally suggests **ä–ə* (Aikio 2015a: 41).[198] Despite this irregularity, Ante Aikio (in a discussion forum) has recently suggested the revival of Wichmann’s (1902: 165) old comparison with Samoyed **jörä* ‘deep’ (> Tundra Nenets *joŕa*, Taz Selkup *korį*, Alatalo 2004: 327; cf. Janhunen 1977: 47; reconstruction given per A. Aikio). If this comparison is correct, then the word can certainly not be a Baltic loanword in Uralic, although some details admittedly need to be worked out.

The question now is whether a Finnic → Baltic loan can be proposed (cf. Senn 1943: 953; Bednarczuk 1976: 48). In my opinion, we must probably answer here in the negative, primarily for semantic reasons. In East Lithuania, whence the majority of the attestations in the LKŽ derive, *jáura* clearly refers specifically to a kind of boggy, infertile soil that dries out and hardens in the summer. The meaning seems to have

---

[197] Bartens claims that the initial glide in Moksha dial. (Penza) *jäŕ‘kä* ‘lake’, *jäj* ‘ice’, (etc.) shows the preservation of **j*, but it is rather a secondary prothetic glide as proven by its appearance in words with no etymological **j*: cf. Md. M dial. *jäl’* ‘hem’ (< PU **älä*), *jäl’ďä* ‘mare’ (cf. Sá. N *áldu* ‘reindeer cow’).

[198] The expected Finnic **jarvi* would appear to be found in Vt. *jarvi* and Li. *jǭra*; however, Salaca Livonian *järu* seems to prove a Proto-Livonian **järru* < **järvi* (Grünthal 2012: 313; Kallio 2016: 46); compare likewise Salaca *jämde*, but Courland Livonian *ja’mdõ* ‘thick’ (< **jämetä*). Also, Votic *jarvi* (dial. *järvi*, cf. VKS: 306) must be recent in view of Krevinian *järvi* (Kettunen 1930: 125–126, cf. the 17th century toponym Järfwenkylä).

broadened to ‘bog’ in Žemaitia, but nowhere does the word refer to a water body. Therefore, a Finnic origin is semantically unattractive.

I would also question whether this word really can be compared with Lt. *jū́ra*, Lv. *jūra*, Pr. E *luriay* */jūrjai/ ‘sea’ (as in Trautmann 1923: 335, etc.). From a semantic perspective, ON *aurr* ‘mud, mire’ seems a closer match.[199] Lt. *jū́ra* ‘sea’, while corresponding with the Uralic forms semantically, cannot be compared formally; moreover, if it is related to Arm. *ǰowr* ‘water’ (Meillet 1920: 251–252; Olsen 1999: 787),[200] this would effectively exclude a Finnic origin.

(b) **‘leather (strap)’**. F *hihna*, E dial. *ihn*, Li. (Kettunen) *nī’ņ* ‘leather strap or belt’ ~ Lt. *šikšnà*, Lv. *siksna* ‘untanned leather; leather strap or belt’; cf. Sá. S *sesnie* ‘untanned hide left to moult’, L *sassne* ‘tanned reindeer leather’ (< **sę̄snē*);[201] Md. E *kšna*, M *šna* (< **(šə)šna*) ‘worked leather; leather strap’; Ma. E *šüštö*, W *šəštə* (< **šŭštə*) ‘tanned leather (used for harnesses); harness, strap of such leather’ (Thomsen 1890: 223; Kalima 1936: 101).

---

[199] True, *aurr* and *jū́ra* are often combined under a single etymology (e.g. IEW 78–81), which would appear to be supported by OE *ēar* ‘sea’. However, it still remains difficult to explain the initial glide in the Baltic form (see the following footnote).

[200] The outcome of initial **i̯-* in Armenian remains controversial (see Martirosyan 2008: 706–707 with lit.; Olsen/Thorsø 2022: 203–204), but this etymology seems quite compelling to me. It is preferable to the comparison of the Baltic term with Skt. *vā́r* ‘water’, Lat. *ūrīna* ‘urine’ (e.g. Derksen 2015: 215), as this leaves the Baltic **j-* unaccounted for; an analogical **j-* from the full-grade, postdating **eu* > **jau-* is hardly possible for Prussian at least, since the latter development does not appear to have occurred there (see fn. 167).

[201] The Eastern Sámi languages (Sá. I *šišne*, Sk. *še´šnn*) reflect an irregular form **šišnē*. According to Aikio (2009: 151), these are later loans from Finnic. While this is probably true, note that Aikio has later characterized West **s* ~ East **š* as a common feature of palaeo-Laplandic words (2012a: 85); compare Sá. N *siekkis* ~ K *ši´ŋŋg* ‘dewclaw’, N *sáhppasat* ~ K *šaahpreš* ‘small intestine’.

Aikio (2009: 151) stresses that the correspondences within Uralic are irregular (cf. already UEW 786). He therefore assumes that the Mordvin and Mari words represent independent loans from Baltic. Grünthal (2012: 318) agrees, stating that the expected Mordvin cognate of Finnish *hihna* would be **šokšna*. However, the loss of pre-Mordvin high-vowels in unstressed syllables is a well-described phenomenon (Надькин 1988: 7); cf. similarly E dial. *kšta-*, M dial. *šta-* 'wash' ~ F *huuhto-* 'rinse, wash' (< **šušta-*?) and Md. E *šta* (dial. *kšta*), M dial. *šta* ~ Ma. W *šəštə* 'wax'.[202] The initial *kš-* (< **č-*) in Erzya appears to be regular (cf. even Md. E dial. *gžniva* 'stubble' ← R dial. *жни́во*, Juho Pystynen p.c. October 2021). The Mari form, on the other hand, is probably indeed irregular, as the expected reflex of Proto-Uralic vowel combination **i(-a)* in Mari is **ŭ* (Aikio 2014a: 156). As with the word for 'lake', the irregularities here are quite modest.

Due to the existence of apparent cognates in West Uralic, the direction of loaning has occasionally been questioned (Mikkola 1930: 440–441; Mägiste 1959: 171; Bednarczuk 1976: 53; cf. Karulis 1992 II: 180). Indeed, the Baltic word does not have an acceptable etymology (ALEW 1183; cf. Holopainen 2019: 249, fn. 43),[203] so that a loan from Finnic to

[202] Holopainen (ibid.; cf. also Pystynen 2020a: 83) reconstructs **śišta* for this word, but it seems only the Mordvin form might be able to reflect such a preform: we would expect Mari **šŭštə* and Komi **śeš(t)* instead of the attested *śiś* 'candle' (cf. *ež* 'surface' < PU **iša* 'skin'). If the Komi **-i-* shows a special development (or is unrelated), then Udm. *śuś* 'wax, honeycomb' and Mari **šĭštə* could perhaps reflect PU **ćeštV vel sim*.

[203] A promising suggestion has been made in van Sluis et al. (2023: 226) who compare the Baltic words with MW *cen*, Bret. *kenn* 'skin, hide; scales', providing a Proto-Celtic reconstruction **kisnā-*. This Celtic form is traditionally compared instead to ON *hinna* 'membrane' (LEIA C–55 with lit.; IEW 929; Kroonen 2013: 226), which still, however, cannot be ruled out on formal grounds.

Baltic would seem more probable than the opposite. On the other hand, the irregular Mari form, non-Uralic phonotactics (medial *CR*-cluster), and the occurrence of the phoneme **š* make it unlikely we are dealing with a genuinely inherited word in Uralic (cf. J. Häkkinen 2009: 47; Aikio 2015a: 44–46). It therefore cannot be excluded that the word was adopted into Baltic and the West Uralic languages from some other source (Junttila 2015a: 31).

(c) **'alder'.** F *leppä*, E *lepp*, Li. *liepā* (< **leppä*) 'alder' ~ Lt. *líepa*, Lv. *liẽpa*; Pr. TC *leipen* ACC.SG. 'lime tree'; cf. Sá. N *leaibi*, S *liejpie* '(grey) alder' (< **leajpē*; Sk. *leä´p'p*, K *lie´hp* ← Finnic); Md. E *l'epe*, M *l'epä* (< **lepə*) 'alder' (Sammallahti 1977: 139) — To my knowledge, a Finnic → Baltic loan has never been suggested, although the Uralic words have traditionally been treated as cognates (e.g. UEW 689). Sammallahti's Baltic → Uralic loan etymology was accepted by Koivulehto (1992a: 173–174) and Aikio (2012a: 74) although it has often been qualified as uncertain (Suhonen 1988: 611; SSA II: 64–65; Häkkinen 2004: 595; van Linde 2007: 107–109).

The Uralic words do not show regular sound correspondences, as has long been recognized (E. Itkonen 1946: 306 attributes the irregularities to "dem allgemein bei den Baumnamen zu beobachtenden lautlichen Schwanken"). The Mordvin form has been explained as a loanword from Finnic (Sammallahti 1977: 139; Aikio 2012a: 108). This would explain the irregular vocalism, but the existence of Finnic loanwords in Mordvin requires further substantiation. On the surface, the Mordvin forms imply **lippä* or **lüppä*, while Sámi suggests **leipä*.

Sammallahti assumes that Finnic and Sámi borrowed the word from Baltic independently, and that Finnic **leppä* "was better suited to the sound system". This is rather a strange claim, since we know that Baltic **ę̄* is regularly substituted by Finnic **ęi* or **ī* in loanwords, as discussed

on p. 122–129, cf. **hęina* ‘hay’ ← Baltic **šę̄na-*, while the substitution PF **e* ← PB **ę̄* is completely unparalleled. Another issue with assuming independent Baltic loans is the semantics. The Uralic words all refer to the ‘alder’, while in Baltic, the word means ‘lime tree’. As noted by van Linde (2007: 109), these trees are not very similar to each other, so if a semantic shift can be assumed at all, it would be difficult to imagine it occurring twice. Grünthal’s (2012: 321) proposal to assume a third independent borrowing into Mordvin exacerbates the issue.

Aside from equivalents in Slavic (R *лúна*, Slk. *lipa*, SCr. *lȉpa* ‘lime tree’), the Baltic word has no other Indo-European cognates. The traditional etymology comparing Lt. *lìpti* ‘to stick’ (Trautmann 1923: 155; REW II: 44; Smoczyński 2018: 697) fails to explain the acute attested throughout Balto-Slavic (Kortlandt 1985: 121; Derksen 2008: 279; ALEW 669). Thus, if there is any relationship between the West Uralic words for ‘alder’ and the Balto-Slavic word for ‘lime tree’, then it would have to be indirect. This seems a fairly decent candidate for a shared substrate word (cf. Matasović forthc.), although in view of the difference in meaning, the possibility remains that the similarity is coincidental.

~ ~ ~

The only Lithuanian words for which a Finnic origin can be said to have gained general acceptance are late loanwords mediated through Latvian (cf. Thomsen 1890: 68–71). The most widely accepted example is Lt. *laĩvas*, Lv. *laĩva* ‘boat’ ← Finnic **laiva* (> F *laiva*, E *laev*, Li. *lǭja*) (Mikkola 1930: 443; Kalima 1936: 129; LEW 335; Smoczyński 2018: 660). As noted by Junttila (2015a: 24), the accentual relationship between Lithuanian and Latvian implies a post-Proto-Baltic diffusion. The direction of borrowing was confirmed by the discovery of a

convincing Germanic etymology, cf. ON poet. *fley* 'ship' (< **flauja*-, Koivulehto 1970; LÄGLOS II: 159–160; SSA II: 39).

In a similar semantic field, note Lt. *bùrė*, Lv. *buŗa* 'sail', which Kalima (1936: 148) considered to be most probably from Finnic **purjeh* (> F *purje*, E *puri*, Li. *pūŗaz*; cf. also Mikkola 1930; Bednarczuk 1976: 47; SSA II: 435). In an extended treatment, Nieminen (1955) has argued that the Lithuanian word was borrowed from Latvian, and that the word is indeed a Finnic loan (cf. LEW 65; Smoczyński 2018: 165). Incidentally, Koivulehto (1970: 182, fn. 27) has suggested a Germanic origin here, too (← Norse **buri*- > ON *byrr* 'sailing wind, favourable wind'). The same route was taken by Lt. dial. *aĩrė*, *áirė* (Būga 1924a: 24; LKŽ *kartoteka*) 'oar' ← Lv. *aĩris*, dial. *aĩre* 'oar', ultimately from Germanic **airō*-, cf. ON *ár*, OE *ār* 'oar', for which Endzelīns (ME I: 13; Zeps 1962: 100), probably correctly, assumes a Finnic intermediary (F *airo*, E *aer* 'oar').

Another plausible case is Lt. dial. *asiaĩ* 'rough horsetail', Lv. *aši* (secondary *ašķi*, cf. ME I: 146–147) 'horsetail, *Equisetum*', which might be analysed as a loan from Finnic **hosja* (> F *hosia*, E *osi*, Li. *vỏžā*) '(rough) horsetail, *Equisetum*'; cf. Thomsen (1890: 253). If this etymology is correct, however, it would have to postdate the change **š* > **h* and therefore cannot be interpreted as contemporaneous with the Proto-Finnic loans from Baltic (Junttila 2015a: 25).[204]

[204] Frankel (LEW 797, followed by Smoczyński 2018: 1185) has suggested the same for Žem. *skárda* 'sheet metal'. In his opinion, this was borrowed through Latvian *skārds* from E *kard*, Li. *kārda* (< **karta*) 'sheet metal' (cf. Endzelīns 1924: 120–121). However, a pan-Baltic distribution is implied by an attestation from Pelesa (Belarus) provided in the LKŽ. In North Finnic, the word appears to be limited to Ingrian, where it might have been borrowed from Votic (cf. Thomsen 1890: 138, fn. 1), and it is possible that this word

These words are of little interest for our purposes. As the above discussion has shown, there are no cases in which Finnic can be conclusively shown to have been the donor language into Proto-East-Baltic, even if this cannot always be excluded. It seems quite possible that there are no Finnic loans in Proto-East-Baltic at all, despite the relatively large number of loans in the opposite direction. While this could imply something about the power balance within the contact relationship, the most probable analysis appears to be that the Baltic language which donated loanwords to Finnic is not the direct ancestor of any extant Baltic language. This could be attractively interpreted in the context of a Baltic speech community being absorbed by a Finnic one (in other words, a Baltic substrate in Finnic; cf. Kalima 1936: 190), a hypothesis which is also supported by other lexical data (Kallio forthc.; see *1.3.6*).

1.3.5. Common loans from unknown sources?[205]

The theory that certain words within Finnic derive from an unknown 'autochthonous' or substratal language is chiefly associated with the Estonian linguist Paul Ariste (1962, 1971),[206] whose views on the subject seem to be regarded as synonymous with the theory itself (cf. Kendla/Viikberg 2015). Essentially, Ariste observed that words of

---

diffused through South Finnic fairly late. This word must be considered in the context of evidence for metal production in the Baltic region (Būga 1923: 3).

[205] This sub-chapter will be published, in a slightly modified form, as Jakob forthc. c.

[206] I have unfortunately been unable to access Ariste's monographic treatment (*Keelekontaktid: eesti keele kontakte teiste keeltega*. Tallinn: Valgus, 1981), although judging by the discussions in Kendla/Viikberg 2015, it appears most of the relevant material was already discussed in his earlier articles.

unknown etymology tended to cluster in certain semantic fields, particularly geographical terminology, "somatic words" (1962: 17) and fish names (1971: 10–11, 1975). As the only criterion for identifying substrate words was the absence of an etymology, it is not surprising that the theory failed to achieve widespread acceptance (Saarikivi 2004: 188): the clustering of etymologically obscure words in particular semantic fields may be a statistical argument in favour of a linguistic substrate (cf. Aikio 2004, Saarikivi 2004), but the suggestion becomes circular when applied on the level of an individual lexeme.

Thus, when the Finnic cognates of *saari* 'island' are reduced to a reconstructed Proto-Finnic **sāri*, what we are left with is a single, isolated data point — a single witness. In the absence of comparative data, we may speculate that the word is of foreign origin, but this cannot be substantiated with any positive evidence. A proposal built on the absence of an etymology alone is naturally very vulnerable. For instance, Ariste suggested that Estonian *aed* 'fence' was a substrate word (1962: 17), but this has since turned out to have an impeccable cognate in Khanty (Aikio 2014b: 1–2), and there are competing etymologies for many other suggested substrate words, some of which are now widely accepted (Kendla/Viikberg 2015: 143–147; Kallio forthc.).

Rather little evidence from Finnic has been adduced beyond the lists presented by Ariste. For instance, K. Häkkinen (2004) discusses the possibility of a "proto-European" origin where Ariste had previously suggested it (thus s.v. *helmi* 'pearl', *liha* 'meat', *saari* 'island') but does not expand the corpus, even though many words are otherwise labelled as lacking an etymology and could, at least as far as the semantics are concerned, be decent candidates (*hiki* 'sweat', *mahla* 'sap', *tavi* 'teal', etc.).

Support for a substrate loan origin has been furnished in specific cases by internal evidence, such as unusual phonotactics or morphology (J. Häkkinen 2009: 37–38; Aikio 2012a: 84; Живлов 2015), but even suggestions of this kind may be vulnerable and run the risk of circularity. For instance, both J. Häkkinen and Живлов (op. cit.) cite the internal cluster *-mm-* as evidence of non-Uralic origin, yet Aikio (in prep. 12) has argued in favour of such a cluster in native vocabulary. Furthermore, even if a word's phonotactics would indeed rule out an inherited origin, we can still not in principle exclude that the word's source will be later identified in an attested language.

Since Ariste, some attempts have been made to elevate the West Uralic substrate theory both on a general theoretical level (Напольских 1990, 1997; Wiik 1992; Helimski 2001), and with reference to new linguistic data (Aikio 2004; 2012a: 80–88; Saarikivi 2004), but it is only in the last decade that we have seen a real surge of interest in the area (cf. Живлов 2015; Kendla/Viikberg 2015; Aikio 2015a: 45–47; Soosaar 2021). These studies show an increased focus on phonological and phonotactic criteria for identifying substrate words. Aikio also identifies cases (and later patterns) of irregular correspondence between Sámi varieties (2004: 14–16; 2012a: 85). This is important, as it allows us to move beyond the "single witness" problem, allowing multiple proto-forms to be treated as independent comparanda in support of a substratal origin.

What can be remarked upon is that the results achieved in this area by Uralicists seem to have been largely independent of those achieved by Indo-Europeanists (on which see *2.1.1*). Of the cited authors, only Soosaar draws on any Indo-European evidence previously mentioned in this connection, noting the suggestion that F *leivo*, E *lõoke* and OE *lāwerce* 'lark' may be parallel loans from an unknown language

(Schrijver 1997: 309).[207] Otherwise, Indo-European evidence has rarely figured in the discussion of possible palaeo-Baltic borrowings in Finnic. Besides a brief comment by Junttila (2015a: 31) that certain lexical isoglosses between Finnic and Baltic may represent "parallel borrowings from a shared source, perhaps a lost substrate language", the potential relevance of the Baltic evidence to this debate has not been recognized.

Combining Baltic and Finnic evidence could be a further way to resolve the "single witness" problem, and allow us to substantiate proposals of substrate origin based on positive comparative evidence. However, the Baltic evidence can only be considered an independent witness of a shared substrate word where a direct loan relationship with Finnic can be ruled out. Where a Finnic word can, on phonological grounds, be treated as a Baltic loanword, it cannot constitute independent evidence, and while the possibility that the word was loaned into Finnic and Baltic from a third unattested source remains a theoretical possibility, it cannot be substantiated (compare, for instance, the examples in fn. 92).

Thus, in order to find reliable evidence for a shared substrate in Finnic and Baltic, which I will refer to here as the "palaeo-Baltic" substrate, we will need to identify words which are clearly related but which cannot be considered direct borrowings from one attested language to the other, thus presupposing the involvement of some third source. In

[207] Напольских (1990: 129; 1997: 200, fn. 5) does refer to some literature from the first half of the 20th century, namely Feist's theory of a lexical substrate in Germanic and Pokorny's theory of a non-IE substrate in Celtic. Kallio (1997: 126–128) can be considered responsible for bringing the American school of thought to the attention of Uralicists (Aikio 2004, Saarikivi 2004), although as discussed in *2.1.1*, this particular branch of research was rather light on specific data.

this section, I will try to identify cases in which the Baltic and Finnic evidence complement each other and support the supposition of a palaeo-Baltic lexical layer in both language families. After a case study on fish names, I will attempt to identify phonological criteria which might allow us to distinguish substrate lexemes, and finally present a couple of good candidates.

1.3.5.1. Fish names

Aside from an old inherited term for 'fish' (Lt. *žuvìs*, Lv. *zivs* = Gr. ἰχθύ̓ς, Arm. *jukn* 'fish'), very little of the fishing-related vocabulary in Baltic can be traced even as far as Proto-Balto-Slavic. A common term for 'eel' can be reconstructed on the basis of Lt. *ungurỹs* (→ Finnic **ankẽrjas*), Pr. E *angurgis* and — with divergent suffixal vocalism — R *у́горь*, Cz. *úhoř*, Sln. *ugọ́r* 'eel'.[208] Beyond this, just a couple of common Balto-Slavic terms can be cited, each having an obscure ultimate origin.[209] This situation can be explained in at least two ways. On the one hand, we might simply assume that early Balto-Slavic speakers did not engage much in fishing and did not distinguish many kinds of fish. Alternatively, and more probably, we can assume that an originally richer fishing terminology has largely been replaced, likely the result of changing subsistence practices and language contact.

---

[208] Based on the inherited word for 'snake': Lt. *angìs* 'adder', Lv. *uôdze* 'viper', Pl. *wąż*, Lat. *anguis* 'snake', etc. (LEW 1163).

[209] The best example is Lt. *šãmas*, Lv. *sams* ~ R *сом*, Pl. *sum*, SCr. *sȍm* 'wels catfish'; beside this, we find Lt. *lýnas*, Lv. *līnis* ~ Pr. E *linis*, R *линь* (GEN.SG. *линя́*), Cz. *lín*, Sln. *lînj* 'tench' (note the mismatch in intonation!). See Pronk (2022: 270). Note my discussions of the words for 'ruffe' (p. 429–430), 'salmon' (p. 404–406) and 'sturgeon' (p. 345–346, 372–373), which show irregular correspondences between Baltic and Slavic.

Unlike Baltic, Finnic has inherited a relatively rich range of fishing terms from its linguistic parent. As well as the generic noun F *kala*, E *kala*, Li. *kalā* 'fish', there are inherited words for specific fish species (F *särki* 'roach', *säyne* 'ide', E dial. *tötkes* 'tench'), and vocabulary related to fish (e.g. F *kute-* 'spawn (of a fish)', *suomus* 'scale') and fishing (F *pato* 'fishing weir; dam'); cf. Aikio 2022: 24. It therefore cannot be stated that Finnic has undergone massive lexical replacement in this semantic field to the same extent as Baltic.

Nevertheless, loanwords in this semantic domain would be unsurprising: the Baltic Sea represents a particular ecosystem featuring species that would not have been familiar to speakers of Indo-European or Uralic languages before they reached the Baltic coast, such as the whitefish, Baltic herring, Atlantic salmon, and sea mammals like whales and seals. The Latvian ethnologist Pēteris Šmits (see P. Schmidt 1930: 87) already noted that a substantial number of fish names in the region were of unclear origin, which he associated with an ancient autochthonous fishing population. This idea was repeated in Benita Laumane's monograph on Latvian fish names (1973: 14; cf. Лаумане/Непокупный 1968: 76; Ariste 1975: 468), and the same semantic field has been the focus of a number of devoted studies (Герд 1970, 1981; Ariste 1975; Sausverde 1996).

Although most of the fish names mentioned by Šmits are also present in Latvian, the majority of these are transparent loanwords from southern Finnic. On the other hand, a couple of the fish names he cites do have a wider distribution. I will treat these here in more detail:

(a) 'whitefish'. F *siika*, E *siig*, Li. *sīgõz* (< **sīka*) 'whitefish' ~ Lt. *sỹkas*; ON *síkr* (attested in kennings), whence Nw./Sw. *sik* 'whitefish' — Already before Šmits, the word for 'whitefish' had been labelled as a possible loanword from 'an aboriginal people' by Būga (RR II: 561). The word also featured among Ariste's lists of substrate words (1971:

11; 1975: 470–471), and was treated as such in a separate article by Герд (1981: 52).[210] The question is whether there is any positive evidence that the word was adopted from a palaeo-Baltic source.

Several sources have treated ON *síkr* as an inherited cognate of R *cuг* and Lv. *sīga* 'whitefish' (Falk/Torp 965; REW II: 621; de Vries 1962: 475), implying Finnic **sīka* was borrowed from Norse. However, the dialectal distribution of the word within Russian clearly favours its interpretation as a Finnic loanword (Kalima 1919: 217; Thörnqvist 1948: 247–248; Герд 1981: 52) and the Latvian word is also generally explained from Finnic, which indeed seems likely (Thomsen 1890: 279; ME III: 851). Thomsen (loc. cit.) has considered Lithuanian *sỹkas* a loan from Baltic German *Siek* 'whitefish' (with voiceless /s-/; cf. Kiparsky 1936: 181–182), which is itself probably from Estonian *siig* (Anderson 1938: 148), and SKES (p. 1013) would even take the Norse word from Finnic, which LÄGLOS (III: 231) acknowledge as a possibility.

As a result, depending on our analysis, all of the evidence can be explained as ultimately deriving from Finnic, or from Norse. In other words, we return to the "single witness" problem, and no positive data can be presented in favour of the substrate hypothesis. In this particular case, the Baltic evidence is furthermore most probably irrelevant to the word's ultimate origin. Although the word remains without a convincing etymology, that fact alone is insufficient to substantiate a hypothesis of palaeo-Baltic origin.

(b) 'herring'. F *silakka*, E dial. (rare) *silakas* 'Baltic herring; salted herring' ~ Pr. E *sylecke*, Lt. *siĨkė*, Lv. dial. (?) *silce* (cited for Rēzekne, see ME III: 840) 'herring' — E *silk* (GEN.SG. *silgu*) '(salted) herring' and Li. *siļk* (NOM.PL. *sīlkõd*) 'herring' are usually quoted here, but due to the

[210] Janne Saarikivi has made the same suggestion at the 13th Finno-Ugricist Conference in Vienna, August 2022.

awkward syncope[211] and mismatch in stem vowel, a direct equation with F *silakka* seems phonologically problematic. Most probably, Li. *siļk* is loaned from Lv. *siĪ̦ke*, which is itself from Lithuanian (ME III: 840), but E *silk* is not well accounted for.

In view of the trisyllabic Pr. E *sylecke*, it is attractive to assume that Lt. *siĪ̃kė* has arisen through syncope from **silekē* or **silikē* (Būga 1916: 143).[212] Trautmann (1910: 426) has assumed *svarabhakti* here, but there is simply no other evidence for such a phenomenon in Prussian.[213] This fact also rules out Brückner's (1877: 131) preform **sildkē* and derivation from Slavic.[214] Other etymologists have derived the Baltic words from ON *síld*, OSw. *sīldh* 'herring' (e.g. Solmsen 1904: 585; Smith 1910: 141; Falk/Torp 966), but this requires an entirely unparalleled and phonetically unexpected development **ld* > *lk* (Smoczyński 2018: 1168). In sum, all existing loan etymologies require unjustified assumptions, and cannot be upheld.

Šmits (P. Schmidt 1930: 87) also noted the similarity of the Baltic and Finnic words with ON *síld*, and assumed they were independently borrowed from a substrate source (cf. similarly Преображенский II:

---

[211] Contrast E *harakas*, dial. *arak*, Võ. *harak*, Li. *arāgõz* (= F *harakka*) 'magpie'.

[212] Alternatively, we could directly compare Estonian *silk* and assume a variant **silk-*, which may further support the non-IE etymology (see below).

[213] Trautmann cites J. Schmidt (1875: 209), but accepts neither of Schmidt's supposed parallels (*gelatynan* and *salowis*, cf. Trautmann 1910: 336, 417). The fact that *svarabhakti* is reported by Becker (1904: 262–263) to be frequent in Pervalkas (as also in South Kurzeme dialects, Endzelīns 1923: 106; Becker is the source of the 'Curonian' form ⟨fīlĕke⟩ cited by Trautmann) has little bearing on our understanding of a Prussian dialect some 600 years and a hundred miles removed from it.

[214] The preform is itself anachronistic, as the R diminutive *селёдка* must derive from a virtual **silid-ikā-* or **-ukā-*, which should have turned up in Baltic as **silidukē*, or the like. Mažiulis (PKEŽ IV: 107) starts with a Baltic preform **sildikē*, but in that case, the loss of **d* is completely unmotivated.

274; Būga RR II: 561; Герд 1980; Кузьменко 2013: 514–515, fn. 4). As lengthening is not regular before **ld* (Noreen 1894: 320–322), the long *-í-* either implies a disyllabic preform **silid-* or **siled-* (cf. Falk/Torp 966; Kroonen 2013: 436) or a metathesis from **sīƀlō-* (Smith 1910: 141; Noreen 1923: 172).[215] In favour of the former clearly speak the early loanwords into Sámi (N *sallit*, L *sallet* 'herring' < **sę̄lētē*) and Slavic (R *сельдь*, Pl. *śledź* 'herring' < **silidi-*).[216]

The disagreement between Baltic **sile/ik-* and Norse **sile/id-* would certainly favour the interpretation of these words as parallel loans from an unknown source. The irregularity is reminiscent of that between ON *hnot*, OHG *nuz* (< **knud-*) 'nut' and Lat. *nux* (?< **knuk-*) 'nut' discussed by Kroonen (2012: 248) and van Sluis (forthc.). One possible explanation for such a phenomenon could be a word-final neutralization of stops in the source language, such as we find in North Sámi (cf. *mádjit*, GEN.SG. *mádjiga* 'beaver'). However, this is merely a typological parallel. Other possible explanations can no doubt be suggested, and as we have no criteria to decide between them, we may limit ourselves to the observation that the correspondence is irregular.

Likewise, the Finnic words are not easy to explain as loans from Baltic, primarily because of their back vocalism. Already Mikkola (1903: 28) compared the Finnic and Baltic words, but stated that the direction of

---

[215] Note the parallels in Ic. *bíldur* (since 17th c.) 'lancet, device for bloodletting' = OHG *bīhal* 'axe' < **bīƀla-* (cf. EWAhd II: 36–37 with lit.), and ON *sáld* 'sieve' < **sēƀla-*, cf. OCS сѣти* 'sift' (see Kroonen 2017: 105, fn. 1 and 108, fn. 8).

[216] The connection with Du. *zeelt* 'tench', which would support this reconstruction, is uncertain on semantic grounds. For the Slavic reconstruction, cf. Mikkola (1903: 28), Būga (1916: 143), Thörnqvist (1948: 78). I fail to understand the alternative reconstruction **sildi-*, favoured by REW (II: 606–607), which ought to have yielded R ***солдь*, Pl. ***słudź*(?).

loaning is unclear. Since Posti (1962), however, the Finnic words have generally been derived from Middle Swedish **sill-laka* (cf. *sill-lake* 1700) 'herring brine' (cf. SSA III: 180; LÄGLOS III: 237). There are serious problems with this explanation, the most important being the single **-l-* in Finnic. If even Sw. *sill* has been loaned into Finnish with a geminate (cf. F *silli*), it is difficult to conceive of how *sill-lake*, where the geminate is further reinforced by a morpheme boundary, could turn up with a singleton /l/. There is no reason to suspect an original geminate would have been shortened in Swedish or Finnic (*pace* Posti 1962: 285).[217] Thus, we are only left with a rescue solution such as the assumption of a contamination with F *salakka*, E dial. *salak* 'bleak (type of fish)', itself of unclear origin (SSA loc. cit.).

We are faced, therefore, with three similar preforms — Baltic **sile/ik-*, Norse **sili/ed-* and Finnic **silakka* — whose relationship cannot adequately be accounted for either by cognancy or by borrowing. I would therefore argue that this is a good candidate for parallel borrowing from a palaeo-Baltic source language.

~ ~ ~

Some additional terms relating to fishing are shared between Baltic and Finnic and lack a plausible Indo-European etymology. At least the following can be cited:

---

[217] The Swedish compound does not appear to have ever been very frequent, and was probably never fully conventionalized, while the occasional spelling with *-ll-* in older Finnish sources could be due to Swedish *sill*. Secondly, the semantics are possible, but awkward, as a two-stage metonymical shift must be assumed from 'herring brine' (unattested in Finnic) to 'salted herring' (unattested in Swedish), followed, in several languages, by a further generalization to 'Baltic herring'. However, see Posti (1962: 286) for a possible parallel.

- F *seipi*, E *teib* ‘dace’, Li. *teib* ‘ide’ (?< **stäipi*, *-e-*; see p. 135) ~ Lv. obs. *stiepats* ‘chub’, i.e. *Steepats* ‘Alantsbleyer’ (Lange 1773: 325; ME III: 1079) (Nuutinen 1987b)[218] — The Baltic stem **stę̄p-* has no apparent further etymology (no attempt is made in ME IV: 1079; Laumane 1973: 79 speculates on a connection with Lv. dial. *stipt* ‘to become rigid’).
- F *toe*, Vt. *tõgõ*, Li. *tǭgõz* (< **tokęh*) ‘fishing weir’ ~ Lt. *takišỹs*, Lv. *tacis* ‘fishing weir’; Pr. E *takes* ‘(mill) weir’ (Thomsen 1890: 226)[219] — Some connection with Lt. *tekė́ti* ‘to flow’ is often assumed (Miklosich 1886: 348; LEW 1052; PKEŽ IV: 181), but the formation has remained problematic (cf. the speculative analysis as **tak-kiš-* with the root of Lt. *kìšti* ‘to stuff’ in Smoczyński 2018: 1441).

There is nothing in these comparisons that would rule out a transmission of the word through Baltic into Finnic,[220] meaning that we have no positive evidence for a palaeo-Baltic origin, although there might potentially be some statistical significance if numerous shared fishing terms turn out to be of unclear origin. We may tentatively add the word for ‘salmon’ to this list (Laumane 1973 *apud* Ariste 1975: 468), whose semantics would make a loanword very probable *a priori*:

---

[218] Nuutinen (op. cit. 109–110) points out that the suffix *-ats* has had some productivity in fish names, e.g. dial. *šķaunats* (ME IV: 22) ‘carp’.

[219] The Latvian word is much more easily explained from **tacsis* < **tacisīs* with syncope than, as often suggested, through reanalysis of a NOM.SG. **taciss*. Prussian *takes* must, however, be taken for */takiss/ (= Lt. dial. *tãkišas*); compare Pr. E *crays*, */kraiss/ ‘hay’ (= Pr. G *kraise* ‘hay’, cf, *craysewisse* ‘a grain tax’, on which see fn. 495).

[220] While in the most certain Baltic loanwords, **o* ← **a* is only found in the neighbourhood of a labial (cf. p. 119), the data is insufficient to rule out a chance correlation.

– F *lohi*, E *lõhe*, *lõhi* (< **lohi*, *-ę-*); Sá. N *luossa*, Sk. *luõss* (< **luosę*) ‘salmon’ ~ Lt. *lašišà*, Lv. *lasis* ‘salmon’ (Thomsen 1890: 194)

The Baltic word has further comparanda in Pr. E *lalasso* */lasasā/, R *лосóсь*, Pl. *łosoś*, and ON *lax*, OHG *lahs* ‘salmon’, which cannot strictly be combined under a shared proto-form. As I suspect that Lv. *lasis* and Lt. *lãšis* have resulted from syncope from an earlier **lašišīs*, a potential irregularity in the Finnic transmission could be the absence of any reflection of the second **š* (the existence of a Proto-Baltic form with syncope is questionable; see the detailed discussion on p. 404–406). However, this evidence remains rather tenuous and open to interpretation.

1.3.5.2. Finnic short vowel vs. Baltic long vowel

Even if the word for ‘herring’ seems to be a reasonable candidate for a palaeo-Baltic substrate word, it would be nice to find some patterns that would help to identify such parallel borrowings in Finnic and Baltic, for example correspondence patterns which do not occur in direct loanwords. In this context, I would like to examine the Baltic vowels **ē* and **ā*. The usual substitutions we find for Baltic **ē* (= **ǣ*) and **ā* in words with a clear Indo-European pedigree are Finnic **ē* and **ō* (see *1.3.3.1*).

On the other hand, several examples of short **a* as a substitution of Baltic long **ā* were collected by Koivulehto (1990: 152, 2000: 105–106 and *passim*; cf. also Kallio 2008a: 207). In his opinion, these loanwords must belong to an earlier layer predating the rounding of Proto-Baltic **ā*, a development he assumes to explain the supposedly later substitution with Finnic **ō.* However, it has now been shown that Finnic **ō* developed from an earlier **ā*, and so the innovation took place on the Finnic side (Lehtinen 1967: 150–151; Aikio 2012: 232). As

noted by Pystynen (2018: 72–75), this points to the opposite conclusion, namely that the loanwords showing short **a* must be later, postdating the raising of Early Proto-Finnic **ā* (> **ō*) but predating the emergence of a new phoneme **ā*:

(a) Early		(b) Late		(c) Post-Baltic	
ī ǖ	ū	ī ǖ	ū	ī ǖ	ū
		ē	ō	ē	ō
ǟ	ē			ǟ	ā
PB *ǣ → PF ǟ		**PB *ǣ → PF ä**			

While Pystynen's account does indeed explain the facts, it seems unattractive to view the raising of original **ā* and emergence of a new **ā* as unrelated phenomena. The two developments seem to be interpretable as a push shift caused by the loss of intervocalic **ŋ* and **x*. The resulting contractions (e.g. **kaŋərə* > **ka.ərə* > **kāri* 'curve; rib of a boat') can be seen as having motivated the raising of the earlier low vowels (cf. footnote 143). In this context, it is unnecessary to assume that Proto-Finnic went through a stage in which **ā* was absent, as in system (b).

If we examine the examples which supposedly show short reflexes of Baltic **ē* and **ā*, it is notable that none of them have a completely evident Indo-European etymology. In five cases, the Baltic word lacks any plausible comparanda entirely:

1. E *vähk*, GEN.SG. *vähi*, Li. *vē'jõz* (< **vähi*)[221] ~ Lt. *vėžỹs*, Lv. *vêzis* 'crayfish' (Thomsen 1890: 241) — The Baltic word has no clear etymology (cf. LEW 1235–1236; ALEW 1419).[222]

[221] Estonian *-k* is a secondary excrescent consonant (cf. *mähk* 'sapwood' < **mäihä*). The *i*-stem may indicate a very recent origin (Junttila 2015a: 181),

2. Li. *vägāli* ‘burbot’ ~ Lt. *vėgėlė̃* 3ª, Lv. *vêdzele* ‘burbot’ (Thomsen 1890: 77) — Although limited to Livonian, the assumption of a later loan from Žemaitian (Thomsen 1890: 141–142) or Curonian (Endzelin 1914: 102; Nieminen 1957: 199) does not help to explain the short first-syllable vowel. The Baltic word has no clear etymology (cf. LEW 1212; ALEW 1392).

3. F *apila*, dial. *apelias* (?< **apęila* ~ **apęlja*)[223] ~ Lv. *âbuõls*, cf. Pr. E *wobilis* ‘clover’ (Thomsen 1890: 156; Kalima 1936: 94 with “?”) — The Baltic forms cannot be separated from Lv. *dâbuõls*, Lt. *dóbilas* ‘clover’, with an unclear initial *d*-. It is generally assumed that the *d*- was lost due to contamination with the word for ‘apple’ (Lv. *âbuõls*; Pr. E *woble*) and/or influence of Lv. *ãmulis* ‘mistletoe’ (cf. dial. *amuols* ‘mistletoe; clover, wood sorrel; daisy’; ME I: 235; LEW 99; ALEW 26–27). While Lv. *(d)âbuõls* does indeed appear to have been influenced by the word for ‘apple’,[224] the similarity of Baltic **dâbila*- and **âbōla*- seems hardly sufficient motivation for the former to have lost its initial stop, which is a typologically unusual

---

but it could also be secondary (cf. *1.3.3.3* on E *kurt*, dial. *kärv*). The Livonian form appears on paper to suggest something like **vähjes*, which could suggest an originally different inflectional type. Salaca Livonian *väji** ‘crayfish’ may rather represent a loan from Leivu *väi* (cf. Pajusalu, Krikmann & Winkler 2009: 293) or Estonian *vähi* (P. Kallio p.c. February 2022).

[222] The connection with NP *gazīdan* ‘bite, sting’ is formally impossible (Cheung 2007: 117–118) and that with Skt. *vāhaka*- ‘a kind of insect’ very uncertain (KEWA III: 198).

[223] VKS cites Vt. *apila* only from the botanical notes of Gustav Vilbaste. Perhaps this is a Finnish loan.

[224] Note that e.g. Standard Latvian *âbuõliņš* ‘clover’ synchronically appears as if it is a diminutive of *âbuõls* ‘apple’.

development.[225] If the word is not of Indo-European origin, the **d*- ∞ **Ø*- alternation might be attributed to the source language(s). A potential parallel is found in the plant name ME *doder*, MHG *toter*, *totter* 'dodder' beside Lt. *jùdros*, Lv. *idra*, dial. *judras* (ME II: 115) , Võ. *judõr*, (Hargla) *jutr*, Li. *ju'ddõr* 'false flax, *Camelina*'.[226]

4. E *hakkama* 'begin; grasp', ?Li. *akkõ* 'grasp, catch' (cf. Junttila 2017a: 131) ~ Lv. *sâkt* 'begin', Lt. *šókti* 'jump, spring (into action)' also dial. 'start suddenly (esp. of weather phenomena)' (Vaba 1992: 222; Holopainen/Junttila 2022: 97) — The original meaning is probably 'jump': compare ME *sterten* 'jump, spring (up, forth); come suddenly into a state or condition' > modern *start* (16th century) 'begin'. The connection with Gr. κηκίς 'ooze, viscous liquid (of blood, pitch, fat, etc.)' (LIV 319; ALEW 1213) is semantically unconvincing.
5. F *varhainen*, dial. *varas*, E *varane*, Võ. *varahinõ*, Li. *varāz*, *va'rri* 'early' (< **varas*, **varahinęn*);[227] Sá. N *vuoras* 'old; old man', Sk. *vuõrâs* 'old man; grown up (e.g. of a reindeer calf)' (< **vuorę̄s*) ~ Lt. obs. *voras*, Pr. III *urs*, ACC.PL. *urans* */ūrá-/ 'old' (Liukkonen 1999: 151–152) — The Baltic word is isolated. No

[225] Koivulehto (2000: 107) suggests that the *d*-forms could instead be secondary, but since he does not provide any explanation for the *d-*, this cannot be considered a fully-formed hypothesis.

[226] The relationship between the Baltic and Võro/Livonian words is unclear (LEW 196). A loanword is conceivable in either direction (cf. Sommer 1914: 197), as well as in a relatively recent timeframe (Junttila 2012: 273).

[227] Liukkonen (1999: 152) suggests a semantic shift 'old' ≫ 'long ago' ≫ 'early', citing as a parallel Hungarian *rég* 'long ago' and its derivative *régi* 'old' (but this shows the opposite development). Another possibility could be to start from the sense 'fully grown, ripe' (cf. Kildin Sámi *vūras* 'large (of fish)'), with a subsequent development to 'timely' as in SCr. *dòspijeti* 'ripen, mature; be on time', and finally to 'early'.

cognates are suggested by LEW (1274), PKEŽ (IV: 211) or Smoczyński (2018: 1695).[228]

Even though the latter example has an equivalent in Sámi, the second syllable vowels do not correspond, meaning that no common proto-form can be set up (cf. Liukkonen loc. cit.).[229] Likewise, the other examples have comparanda in Slavic, but in two of three cases, the comparison is phonologically irregular, suggesting the words in question postdate Proto-Balto-Slavic:

6. F *lapio*, dial. *lapia*, E *labidas*, Li. *lä'bḍi* (< **lapita*) 'spade' ~ Lt. *lópeta*, Pr. E *lopto* 'shovel, spade' (Thomsen 1890: 197 with "?") — The Baltic forms are clearly related to OCS лопата '(winnowing) shovel', but the correspondence is irregular.[230] Note that if we start from Baltic **lâpetâ*, the Finnic second syllable vowel **i* is also unexpected, especially if we consider

[228] As the word is only attested in older lexical sources, the circumflex given by Trautmann (1910: 127), Fraenkel (LEW 1274), and other authors, does not appear to have any basis (cf. Būga RR II: 720). The word is essentially limited to Prussian Lithuanian, and may be a Prussianism (cf. Smoczyński 1983: 171, fn. 15), but the derivative *vorùšis* 'frail person' reported from Linkmenys implies a broader distribution. The form *ùrupė* (rather **ŭ́rupė*?, cf. the river name *Ũ̃rupiai* in Luokė) (= *vórupė*) 'old river bed', cited by Juška (*apud* LKŽ), is, *contra* Smoczyński (2018: 695), hardly reliable evidence for ablaut. Could it be a Sembian Prussianism with regular */ūr-/ < **wār-*?

[229] It is unlikely that Sámi shows suffix replacement. On the contrary, we would expect retention of the suffix **-ēs* to have been encouraged by the more usual synonym **poarēs* 'old' (> Sá. N *boaris*, Sk. *puä´res*).

[230] There is no indication that the Baltic word represents a derivative with lengthened grade (Fraenkel 339–340; Smoczyński 2018: 724), and the comparison with Lt. *lãpas* 'leaf' is better abandoned.

that the suffix **-etA* is frequent in Finnic, while **-itA* is otherwise unknown (cf. Koivulehto 2000: 110–111).[231,232]

7. F *lava* ‘platform, deck’, E *lava* ‘(sleeping) platform; bench (in a sauna)’, Li. *lovā* ‘bed; bench (in a sauna)’ (< **lava*); Sá. N *luovvi*, Sk. *lue´vv* ‘raised platform (for storing meat)’ (< **luovē*) ~ Lt. *lóva*, Lv. *lâva* ‘bunk (for sleeping); bench in a sauna’ (Wiklund 1896: 45–46; Kalima 1936: 131) — The Baltic word is cognate with R dial. *лáвa* ‘bench; platform for washing clothes’, Cz. dial. *lava* ‘bench (along a wall)’, but has no further etymology.[233]
8. F *vakka*, E *vakk* ‘oval container made of bark; dry measure’, Li. *vakā* ‘woven basket; dry measure’ ~ Lt. *vókas* ‘(eye)lid; woven grain basket’, Lv. *vâks* ‘lid’, DIM. *vâcele* ‘woven grain basket; dry measure’ (Koivulehto 2000: 114–115) — The Baltic forms must be connected to R *вéко* ‘eyelid’, dial. (СРНГ IV: 101) ‘lid of a basket or wooden vessel; grain basket’, Cz. *víko* ‘lid’; however,

---

[231] On the other hand, it is possible that a variant with **-i-* existed in Baltic, as in Lt. dial. *vedigà* ‘adze’ (LKA I: 87), *mẽdiga* ‘material’ (for *vedegà*, *mẽdžiaga*), and this might underly Prussian *lopto*, cf. Pr. E *wedigo* ‘Carpenter’s axe’, Lv. dial. *vẹdga* ‘ice chisel’.

[232] Koivulehto (2000: 114) also discusses F *lapa* ‘shoulder blade’, but this is rather an inherited word and cognate with Inari Sámi *lyepi* and Eastern Mansi *lūp* ‘shoulder blade’ (Aikio 2015b: 13).

[233] Fraenkel (LEW 387) suggests a derivation from the root of Lt. *liáutis* ‘cease’ (note this verb in the sense ‘abgeschnittet, verstümmelt werden’ appears to be unattested); however, the semantic connection between this verb and ‘raised platform or deck’ is by no means trivial. Furthermore, one would anticipate the palatal onset of the verb to be preserved in such a derivation, as in *paliovà* ‘break’ < *pa-liáuti*. The derivational chain set up by Smoczyński (2018: 726), involving an unattested verbal form **lóvyti*, involves too many hypothetical stages to be taken seriously.

> the vocalic relationship between the Slavic and Baltic words is irregular.[234,235]

As a result, we are faced with a situation where all of the Baltic loanwords whose Indo-European background is certain show long reflexes of Baltic **ā* and **ǣ* in Finnic, which is actually what we should expect in the case of direct loanwords, while all the plausible examples in which Finnic shows short vowels lack an Indo-European etymology, being at best common Balto-Slavic. In this context, we may venture the conclusion that the two different substitution patterns do not represent different chronological layers, as was previously assumed, but rather betray a distinction between direct and indirect contacts. A possible explanation for this could be that a substrate language underlying Baltic had undergone a sound change (such as open-syllable lengthening) which resulted in phonetically long vowels, while a related substrate underlying Finnic retained short reflexes.[236]

It must be acknowledged, however, that this theory is to a large extent built on a theoretical postulate ("Proto-Finnic always had a phoneme

---

[234] From an o-grade **uoh₁k-*, I would anticipate Lt. **úoka-*; cf. the discussions in РЭС VI: 196 and Derksen 2015: 509. A potential parallel is the word for 'turnip', Lt. *rópė* ~ R *рѣпа*, which is, however, almost certainly of non-IE origin; see p. 373–374 for a discussion.

[235] Md. E *vakan* 'vessel, bowl', as already noted by Paasonen 1896: 36, is hardly from R dial. (СРНГ IV: 9) *вагáн* 'wooden trough'. *Contra* van Linde's (2007: 177) claim that **-k-* is a usual substitution for foreign **-g-*, this substitution actually seems to be highly exceptional. The only generally comparable example listed in Paasonen (1903: 17) is Moksha dial. *avkâs* 'August'. The Erzya word could instead be seen as cognate with the Finnic word, with a suffix as in Md. *kućkan* 'eagle' < PU **kočka*.

[236] For more length alterations, compare the examples collected in *2.3.5.1*. A similar example could perhaps be F *leppä*, E *lepp*, Li. *liepā* 'alder' ~ Lt. *líepa*, Lv. *liẽpa* 'lime tree'. On this word in detail, see p. 157–158.

*$\bar{a}$") and can be viewed as a potential house of cards. Should a convincing Indo-European etymology be discovered for any one of the Baltic source words, we would be forced to accept a Baltic → Finnic loanword, and with it, the possibility of a substitution *$\bar{a}$ → *a. In that case, we would be compelled to accept an alternative solution, such as Pystynen's chronological one, and we might as well apply that explanation to all of the examples. Thus, although the theory potentially carries more weight than Ariste's in that it identifies a linguistic pattern in the data, its vulnerability is only exacerbated, as it depends not only on a single word lacking an etymology but on a whole set of words lacking one.

1.3.5.3. Irregular front vocalism

Koivulehto (1971) collects some material which would show Finnic front vowels as substitutes for back vowels in loanwords, but does not concern himself with any explanation of this phenomenon. I will not address the Germanic evidence, which is beyond the scope of my study. As for the Baltic evidence, **tüttär* 'daughter' and **tühjä* 'empty' are open to interpretation (see p. 121). Two other frequently cited examples (e.g. Kalima 1936: 66; Koivulehto 1971: 577; Nuutinen 1989: 498) show front and back variants within Finnic:

- F *rastas*, Vt. dial. *rassa* 'thrush', E dial. *raastas*, Li. *rastā* 'starling' ~ F dial. *rästäs*, Vt. dial. *räsäz*, E *rästas*, Võ. *rästäs* 'thrush'[237]
- F *ankerias*, E *angerjas*, Li. *aŋgõrz* ~ K (Olonets) *ängeriäs*, Vt. (Kukuzzi) *ängeriä*, E dial. (Vaivara) *änger(jas)* 'eel'

[237] Compare similarly the bird name F *varpunen*, E *varblane* ~ E dial. (E) *värb*, *värblane*, Vt. *värpo* 'sparrow', of Slavic origin.

In both cases, the front-vocalic form appears to be secondary. This is shown by the lack of clear dialectal patterning: E dial. *änger(jas)* is rare and marginal, while *rästas* is attested throughout Estonia (see VMS s.v.). In North Finnic, the fronted variants are in principle infrequent. It is clearly anachronistic to blame these dialectal effects on a borrowing event many centuries prior.[238] The transfer of back-vocalic words to front harmony is a typical expressivization mechanism in Finnic (cf. Saukkonen 1962; Nikkilä 2002: 132; Vaba 2011: 749), and both words show other signs of expressivization, e.g. introduction of the primarily non-native phoneme /č/ in Karelian *račoi* 'thrush' or irregular suffix substitution in F dial. *angerva* (SMS s.v. *ankerias*).

In this context, we can examine the following case:

– Vt. *ätälä*,[239] E *ädal* (secondary *hädal*) (→ F dial. *ätälä*, Ojansuu 1916: 202), Võ. *ätäl* 'aftermath' ~ Lt. *atólas*, Lv. *atãls*, Pr. E *attolis* 'aftermath' (Thomsen 1890: 159) — For various etymological analyses, none of which are convincing, cf. ME I: 149; Witcak 2001; Kabašinskaitė/Klingenschmitt 2004: 89–95. See also p. 365.

The consistent front vocalism shown in Finnic is difficult to explain starting from the attested Baltic forms. While Li. (Kettunen) *a'ddõl* 'aftermath' does indeed suggest a variant with back vocalism, according to Kettunen (1938: 2), the word should be reconstructed

---

[238] Uotila (1986: 213) and Vaba (2011: 749) suggest that the words in question were originally disharmonic, with this discrepancy only being resolved in the individual languages, but it is hardly believable that the violation of vowel harmony was permitted in Early Proto-Finnic only to be reinforced again in Late Proto-Finnic (compare Pystynen 2018: 70–72).

[239] Ojansuu (cf. SSA III: 499; Junttila 2012: 272) assumes the Votic word was adopted from Estonian, but apparently only because he takes the latter as a late Latvian loan, which is hardly necessitated by the data.

**atę̨la* rather than **atala*,[240] therefore neither representing a back-vocalic equivalent to the Estonian forms, nor being straightforwardly derivable from Latvian (see also Gāters 1953: 155, who offers an unconvincing solution). As a result, this example is not directly comparable with those cited above, where equivalent front and back variants were attested dialectally. Furthermore, there are no other indications of 'expressivization' in this word.

To resolve this problem, we might suppose that the irregularity is the result of an indirect loanword relationship. There is otherwise possible evidence that the Baltic word was borrowed from a non-IE source in its irregular comparandum in Slavic (see p. 365), although it cannot be entirely excluded that the front vocalism in Finnic is merely secondary, as in the word for 'thrush'.[241]

We can also note the word for 'sleigh': F *reki*, E *regi*, Li. *re'ggõz*, whose *e*-vocalism is unexpected on the basis of Lt. *rãgės*, Lv. *ragus*, *ragavas* PL. 'sleigh, sledge'. The traditional etymology for Baltic connects these to Lt. *rãgas*, Lv. *rags* 'horn', based on the "horn-like" shape of the sledge's runners (thus ME III: 465, LEW 685; ALEW 964; Smoczyński 2018: 1105). Needless to say, this is merely guesswork, and does not account for the Finnic evidence (cf. Kalima 1936: 66). A Proto-Baltic variant **regē* can hardly be posited on the strength of Lv. dial. *ręgavas* (which is probably due to a secondary dialectal development, cf. Endzelīns 1923: 36–37) and toponymic evidence (*contra* Nieminen 1957: 202).[242]

---

[240] Compare Li. *vie'ddõl* 'liquid' (< **vetelä*) as against *madāl* 'low' (< **matala*). In a footnote, Vaba (loc. cit.) notes a form ⟨addal⟩ from Hupel's dictionary, but this must be a printing error: the German-Estonian part of the dictionary has ⟨åddal⟩ (Hupel 1818 II: 417).

[241] In North Finnic, there is yet another suspiciously similar word: F *odelma*, Ingrian *oelma* 'aftermath' (< **otę̨lma*). The derivation from F *ota* 'spear, thorn' (SSA II: 258) does not seem particularly convincing.

On balance, while the vocalism of the examples adduced here is indeed problematic and has not yet found a satisfactory solution, the evidence is rather limited. While the proposal of parallel loanwords from a palaeo-Baltic source might provide a possible explanation, it is uncertain whether there are sufficient examples for such a proposal to be justified.

We may conclude that the search for phonological criteria to distinguish direct and indirect loanwords between Baltic and Finnic has yielded only modest results. In the following, I will tackle the question from a slightly different perspective, and treat two case studies in detail.

1.3.5.4. The word for 'thousand'

First, we will examine the word for 'thousand', which is generally accepted to be a Baltic loanword in Finnic (Thomsen 1890: 232–233; Kalima 1936: 170–171; SSA III: 318). The data are as follows:

– F *tuhat* (OBL. *tuhante-*), E *tuhat*, Li. *tūontõ* (< **tuhat*, OBL. **tuhantę-*) 'thousand' ~ Lt. *tū́kstantis*, Lv. *tũkstuôtis* 'thousand'

Despite the consensus, it has always been clear that the East Baltic forms do not represent a suitable source as attested. As a result, one has instead operated with a hypothetical Baltic source such as

[242] A slightly different issue is posed by F *rieska*, E *rõõsk*, Li. *rõskõ* 'fresh, unleavened', which is compared to Lt. *pré̇skas* in the same sense. These forms can be reconstructed to Proto-Finnic as **rę̄ska*, yet such a form would presuppose an Early Proto-Finnic **rāska*, in violation of vowel harmony (Pystynen 2018: 71–72; a similar issue faces Vaba's derivation of E *lõõts* from Lv. *plẽšas* 'bellows', on which see Holopainen/Junttila 2022: 64). There are two possible solutions. First, the back vocalism could be secondary, an unusual development which, however, does have a parallel in F *mela*, E *mõla*

**tūšamti*- (Būga 1908: 138; Nieminen 1957: 190; Lühr 1993: 124; Liukkonen 1999: 15),[243] a reconstruction based primarily on the Finnic form, and unsupported by the comparative data. If the alleged Baltic **š* reflects IE **ḱ*, then it remains to be explained why no trace of this phoneme is found in Go. *þusundi*, ON *þúsund* 'thousand'.[244] If we assume **š* reflects IE **s* with RUKI law, then it remains to be explained why we do not find a RUKI reflex in Slavic (cf. OCS тъıсѧщи 'thousand'). Moreover, in both scenarios, the actually attested East Baltic data is unexplained.

The only way to reconcile the Germanic and Slavic evidence is to reconstruct a medial cluster **-ts-*: the **t* would be lost in Germanic, and would block the effects of RUKI law in Slavic. From this starting point,

---

'paddle' (< **melä*, cf. Sá. N *mealli*, Md. M *miľä* 'oar'; Kallio 2014: 161). The alternative solution is to assume a younger loan, which would also be supported by the young syllable structure **CV̄CC*- (cf. Junttila 2019: 36). However, none of the other loan evidence can support the existence of Baltic loanwords in Late Proto-Finnic. It is perhaps of relevance that the Baltic word has an irregular cognate in OHG *frisc* 'fresh' (see p. 423–423), although since the issue with this loanword is mainly chronological, it is uncertain whether the unexpected Finnic vocalism can be explained away by positing a loanword from an unknown source.

[243] Kalima (1936: 57, 86–87) sees a parallel for the substitution **kst* → **š* in F dial. *ahingas* (?← Estonian, Junttila 2016b: 226), E *ahing*, Li. *a'ŋgõz* (< **ahinka* ~ **ahinkas*) 'fishing spear' ~ Lt. *ãkstinas*, HLv. obs. (Bezzenberger 1882: 275) *akstyns* 'thorn, goad' (Thomsen 1890: 157). However, this comparison is best abandoned, as the Finnic stem-final velar is also unexplained (cf. Liukkonen 1999: 15).

[244] A reconstruction of the type **tuHs-(d)ḱmt*- (Bugge 1888: 327; Leumann 1942: 126–128; Kroonen 2013: 554) has usually been suggested based on a notion (in my view misguided) that this word contains the Indo-European word for 'hundred'. The development of **-s(d)ḱ*- to **-s*- in Germanic is implausible (Hirt 1896: 343; Pijnenberg 1989: 101; Gorbachov 2006: 8) and not supported by any other evidence.

there is no room for a Baltic form with **-š-*. Instead, the Baltic evidence can only be accounted for by assuming an irregular metathesis to *-st-*. As a result, Pijnenberg (1989: 104–105) has reconstructed an underlying **tuHt-h₁s-nt-ih₂-* (in his notation **tūt-sn̥t-ī*) 'eine große Quantität bildend'. However, the root **tuHt-* (a supposed extension of the root of Skt. *tavás-* 'strong, powerful') lacks external parallels, meaning his semantic reconstruction is *ad hoc*, and moreover, the Baltic metathesis remains irregular (see also Lühr 1993: 118).[245]

In view of the problems in reconstructing a common proto-form, Stang (1966: 282; 1972: 49) has suspected that the word for 'thousand' is in fact of non-IE origin. Indeed, as discussed in *2.2.3.2*, there are possible parallels for an irregular alternation between **st* and **ts*, which might be an indication of parallel borrowing.[246] If the Indo-European word

---

[245] The **m* is usually reconstructed on the basis of Pr. III *tūsimtons* ACC.PL. 'thousand', but this, like ON *þús-hund* 'thousand', is more probably a folk-etymological distortion after the word for 'hundred' (cf. Lt. *šimtas*; Hirt 1896: 345–347; Vaillant 1958: 647). As Hirt pointed out, the word-internal **-sḱ-* should have given Germanic **-sk-* by sound law, so any sequence *-sh-* must necessarily be of secondary origin. Indeed, we would expect an old **-m-* to have been preserved in East Baltic (Stang 1966: 100).

[246] A somewhat similar irregularity is seen in the word for 'wax', F *vaha*, E *vaha*, Li. *vǭ'* 'wax', which cannot be regularly derived from Baltic **vaškas* (> Lt. *vãškas*, Lv. *vasks* 'wax'). Here, as in the examples discussed below, one has assumed the generalization of a weak consonant grade (Thomsen 1890: 76; Kalima 1936: 171). Since the irregular cognate in OHG *wahs*, ON *vax* 'wax' can be seen as an indication that the Baltic word is of non-Indo-European origin (see p. 344–345), one may suggest the same for the irregular Finnic form. It must be admitted, however, that the Finnic word could be of Germanic origin, after all: the substitutions Germanic **h* → Finnic **k* and **s* → (**š* >) **h* are known from other early loanwords, (e.g. **kaltas* 'bank, shore' ← Gm. **halþaz*, cf. ON *hallr* 'slope, hill'; PF **kana* 'chicken' ← Gm. **hanan-*, cf. ON *hani* 'rooster'; see LÄGLOS II: 20, 35. On **s* → **š*, see p. 132), while the development **kš* > **h* is regular in Finnic (Posti 1953: 7–9), cf. F *mehiläinen* 'bee' (<

cannot be analysed as inherited, we may suggest that the Finnic word likewise represents a borrowing from a non-Indo-European source, rather than a direct loan from an Indo-European one. Note that there are a couple of other cases where Finnic shows **h* as against Baltic **s*, neither of which have a watertight IE etymology:

- F *laiha*, E *lahja*, Li. *lajā* (< **laiha*) 'thin, lean' ~ Lt. *líesas*, Lv. *liẽss* 'thin, lean' — Regarded as etymologically obscure (ALEW 670; Smoczyński 2018: 698–699).[247]
- F *lahto*, Vp. dial. *lahk*, *-on* (< **lahto*) 'bird trap' ~ Lt. *slãstai* PL., Lv. *slasts*, usu. *slazds* 'trap, snare' — Etymology uncertain (LEW 827; Smoczyński 2018: 1219).

The former has also been explained as showing a reflex of Baltic **š* due to RUKI law (Kallio 2008a: 267).[248] While it appears likely that the RUKI law must have applied after **u* and **i* at some point in pre-Proto-Baltic, the more typical attested reflex is *-s-* (cf. Endzelīns 1911: 29–60; Stang 1966: 99). As the exact chronology of these developments is difficult to establish, it cannot be ruled out that Finnic reflects an earlier Baltic **laiša-*. In this context, we might favour the comparison of the Baltic word with OE *lǣs*, OS *lēs* ADV. 'less' (Kroonen 2013: 324) and

---

PU **mekšə*, UEW 271). Thus Germanic **wahsa-* → pre-PF **wakša* > **vaha* can be considered quite plausible (*contra* LÄGLOS III: 350 with further lit.).

[247] Lt. *láibas* (?→ Lv. dial. *laĩbs*) 'thin, lean' cannot be linked by any known derivational process (*contra* LEW 329–330; Derksen 2015: 268–269).

[248] The traditional explanation has been to assume the generalization of a weak consonant grade (Kalima 1936: 58–59; Posti 1953: 61–62), but such a theory applied to Proto-Finnic is in principle problematic, as the phonologization of consonant gradation postdated Proto-Finnic (see Viitso 1981; Nahkola 1995). Not only that, but **sC*-type clusters did not undergo gradation in Proto-Finnic at all (cf. Posti 1953: 9), meaning that such an explanation is excluded for *lahto*. In any case, **s* was only ever weakened to **h* between unstressed vowels.

further with OHG *līso* ‘mild, soft’, ?Gr. λιαρός ‘mild, warm’ (Osthoff 1910: 325–326; Heidermanns 1993: 370), which I think cannot be ruled out.

However, the **h* in Finnic **lahto* can hardly be blamed on the RUKI law,[249] and the irregularity in this word might be compared with that found in the word for ‘thousand’, and assumed to be an indication of shared substrate origin. Still, given that the substitution **s* → **h* is well known from Germanic loanwords (Koivulehto 1984: 193–195), an alternative way out might be to suggest that the word for ‘thousand’ is of Germanic origin, a solution which has almost never been suggested. True, the Baltic suffixal syllable **-ant-* does come closer to the Finnic data than Germanic **-und-* (but see Koivulehto 1981: 193).

An obstacle to both the Germanic and Baltic etymologies could be the short **u* in Finnic. This has not usually been viewed as a problem, or even remarked upon.[250] Such short reflexes have been attributed to the fact that long vowels were originally only possible in *e*-stems (Koivulehto 1981: 193). However, if such a phonotactic limitation did once exist, there is plenty of evidence the Baltic loanwords postdated it (cf. Plöger 1982: 93). Compare the following etymologies:

---

[249] Nieminen (1934: 28) has in fact suggested that the RUKI law may be responsible in the case of **lahto* by positing a donor form **slakštā-* or **slagždā-* (cf. Lv. obs. *slagzds*; ME III: 912) with an intrusive velar. The dating of a dialectal by-form in Latvian to Proto-Baltic does, however, feel anachronistic.

[250] Thomsen (1890: 99) simply remarks that both long and short **u* are substituted as short **u* in Finnic, while Kalima (1936: 71) passes over the short reflexes in silence (similarly Kallio 2008a: 272). Nieminen (1957: 190) writes dryly: “Das *ū* der ersten Silbe wurde bei der Entlehnung durch *ŭ* ersetzt”.

- F *tuulaalla*, Vp. dial. *tuľhuuda* (< **tūlahe̯la-*) ‘spear-fish by torchlight’ ~ Lv. *dũlis* ‘torch for night fishing’
- F *tuura*, E *tuur* (< **tūra*) ‘ice chisel’ ~ Lv. dial. *dùre* ‘ice chisel’
- F *luuta*, E *luud*, Li. *lūdõ* (< **lūta*) ‘broom’ ~ Lt. *šlúota*, Lv. *sluôta* ‘broom’

The substitution of Baltic **ō* as Finnic **ū* in the last example can only be understood if this loanword predated the raising of early Proto-Finnic **ā* to **ō*, demonstrating that this must belong to a chronologically earlier period (see above). We might suggest that Finnic **tuhat* belongs to an even earlier layer, but this feels *ad hoc* without other supporting evidence. Aside from ‘thousand’, there is one more possible example of the substitution of **ū* as **u* among the Baltic loanwords:

- F *kulo* ‘wildfire; last year’s grass’, E *kulu*, Li. *ku’l* ‘last year’s grass’ ~ Lv. *kũla* ‘last year’s grass; old hair of an animal’, Lt. dial. *kū́lymas* ‘last year’s grass’

Here again, the Baltic source word is of uncertain origin,[251] and the direction of loaning has often been declared uncertain (Thomsen 1890: 190; Kalima 1936: 121–122; SKES II: 234–235). Therefore, there is no solid evidence that would support the substitution **ū* → **u* among the Baltic loanwords, but even if such a substitution is accepted, we are still left with the awkward Finnic **h*.

---

[251] The Lithuanian word looks deverbal, which suggests a comparison with West Aukštaitian *iš-kūlýti* ‘dry up, deteriorate’, yet the latter itself looks be denominal (cf. 3PRES. *-ija*). Nieminen (1934: 26) connects Lv. *kàlst* (1SG.PRET. *kàltu*) ‘dry out, wither’, but the vocalism and intonation are prohibitive. The further connection with Gr. (Hom.) κήλεος* ‘burning (of fire)’, καίω ‘kindle, set on fire’ (Walde/Pokorny I: 376; ALEW 617; Smoczyński 2018: 625) is formally possible but not compelling.

Next, we have to address the words for 'thousand' in Mordvin and Mari. While the vocalism in Md. E *ťožań*, M *ťožäń* 'thousand' seems to match that of Finnic, Mordvin **ť-* normally only occurs in words of affective or obscure origin (Bartens 1999: 46). In view of Mari **tü̆žem* (> E *tüžem*, W *təžem*) 'thousand', we might assume that Mordvin *'o* results from a 'breaking' **ü* > *'o* potentially paralleled by Md. E *śokś* 'autumn' < **sükća* (E. Itkonen 1946: 300–301; Mägiste 1959: 174–175; Keresztes 1986: 170). Mordvin *-ń* might also be derivable from an earlier **-m*, which is strongly supported by the form *ťožəm*, GEN. *ťožmań* recorded by Paasonen (MdWb 2411–2412) for the Erzya village of Seńkino.[252] The result is that the Volgaic forms could possibly go back to a common proto-form **tüžäm(ə)*, but cannot be compared directly with the Finnic forms. Since a derivation directly from Baltic involves a similar issue with regard to the medial **š* and an additional issue by way of the final **-m*,[253] these forms can be adduced as further support for an unknown source language.

To summarize, there are several indications that the word for 'thousand' has been loaned independently into the individual Indo-European (and Balto-Slavic) branches, and the Finnic and other Uralic forms cannot be derived either from a common preform, or be explained as direct loanwords from Indo-European sources without accepting a number of awkward and poorly paralleled substitutions. As a result, it would seem that this word cannot be satisfactorily

---

[252] The regular outcome of word final **-m* is apparently **-n*, as shown by the 1SG.PRES. ending EM *-an* (< **-Vm*) (Bartens 1999: 50). In other instances, *-m* has been restored from oblique case forms, e.g. E *uď'em*, M dial. *uď'əm* 'brain, marrow' (?< **wVdəm*; UEW 572–573).

[253] A development **-ńď-* > **-ń-* occurs in some grammatical morphemes in Erzya dialects (Paasonen 1903: 41), but is not common-Mordvin; therefore, the reconstruction **tušaNtə* (Grünthal 2012: 335) cannot be correct.

explained without assuming the involvement of an unknown language or unknown languages, and the word might have entered the Indo-European and Uralic languages independently from an unattested source.

Given the distribution, we are perhaps dealing with a *Wanderwort* whose trajectory and original source are difficult to identify. However, we might also suggest some kind of connection with the so-called "West Uralic substrate". In support of this idea, we can note that the phoneme **š* has been considered characteristic of West Uralic words showing morphological and phonological irregularities (Живлов 2015; Aikio 2015a: 45–47). On the other hand, as the word is present already in Proto-Germanic, it must have spread into Europe fairly early, and drawing any conclusions on the basis of a single phoneme would be premature.

1.3.5.5. The word for 'moss'

Thomsen (1890: 214) compared F *sammal*, E *sammal*, Li. *sǫ̃mal* (< **sammal*) 'moss' with Lt. *sãmanos* PL. 'moss', but considered the equation questionable. Although Vaba (2011: 757) still labels the comparison as possible, it has rarely featured in discussions of Baltic loanwords, being omitted from Kalima's treatment (1936). Later on, without reference to the Baltic data, Ariste (1971: 10) labelled the Estonian word as a probable loan from an unknown substrate. The most obvious problem is that the geminate *-mm-* in Finnic cannot be explained on the basis of the Baltic evidence. A loanword in the opposite direction would in theory be possible, but the very existence of such loanwords has been considered doubtful (see *1.3.4*). At any rate, there is no clear IE etymology (LEW 761; Smoczyński 2018: 1130).

However the relationship with Baltic is interpreted, it is clear that the Finnic data cannot be divorced from a group of similar Sámi words referring to various mosses (cf. SSA III: 151). Not only do none of these correspond regularly to the Finnic word, they also show irregular correspondences within Sámi. As many as four different groups must be distinguished:

a) Sá. N *seamul* ‘spikemoss; house moss’, L *sämol* ‘(a kind of) peatmoss’ (< **seamōl*)
b) Sá. I *siävŋul* ‘a kind of peatmoss’ (< **seavmōl*)
c) Sá. Sk. *sââu´ŋel* ‘hairmoss’ (< **se̯vmēl*)
d) Sá. K *sõvŋal* ⟨sɛuŋaл⟩ (T.I. Itkonen 1958: 487) ‘hairmoss’ (< **se̯vme̯l*)

The surface cluster -*vŋ*- in Eastern Sámi could reflect a number of possible preforms,[254] but -*vm*- seems to be the most suitable compromise with the Western forms. For *-*vm*- > *-*vŋ*-; compare Sá. S *saajmie* ~ I *sävŋi*, Sk. ⟨saṷ̄ŋ̄́ᵉ⟩, K ⟨sa̦ṷ̄ŋ̦ᵉ⟩ (T.I. Itkonen 1958: 478; modern Sk. *säu´nnj*, K *saa´vvn*), ‘seam’, cf. Ic. *saumur* ‘seam’ (Kallio 2008b, fn. 3).

This is a very interesting case, as the high level of irregularity within Sámi clearly suggests that our word belongs to a relatively recent palaeo-Laplandic substrate layer, entering the individual Sámi dialects independently (cf. Aikio 2004: 14–16; 2012a: 85). On the other hand, the word’s robust presence in Finnic and even as far south as Lithuanian brings the centre of gravity far away from the Arctic Circle. As a possible solution, we could speculate that the word was loaned into palaeo-Laplandic from further south (palaeo-Baltic?), and only from there into Sámi. On the other hand, as Sámi represents a centre of

---

[254] Other possibilities are *-*vŋ*-, *-*vń*- or probably *-*mŋ*- (Eino Koponen p.c. May 2022).

diversity, we might assume an ultimately Laplandic origin, in which case we would have to assume that the word was carried south. Given that we are hardly dealing with a trade word, this would probably imply an actual (southward) migration, presumably by speakers of another unattested language, took place prior to the arrival of Finnic-speakers in the Baltic region. This could potentially suggest a genetic relationship between palaeo-Laplandic and at least one palaeo-Baltic language.

Conclusion

In the above, I have attempted to substantiate the hypothesis that a certain proportion of the shared vocabulary between Baltic and Finnic may not represent mutual loanwords, but rather parallel borrowings taken by the two language families from an unattested source. For the most part, evidence adduced in favour of this hypothesis in the past cannot be further substantiated, as it depends primarily on the absence of an etymology. In theory, unusual morphology or phonology could favour a non-native etymology, but it is difficult to use this evidence to support a specifically non-Indo-European source. Nevertheless, in the course of the above chapter, I have gathered some material which could provide some concrete linguistic support for the hypothesis.

While I have tried to identify substitution patterns which could betray such parallel loanwords, a more robust argument can be built on etymologies for which there are simultaneously multiple indications of palaeo-Baltic origin. In this section, I have discussed three such cases, which I present in the table below (the pre-forms correspond to the approximate time of Baltic-Finnic contacts):

Table 4. Probable shared substrate words in Baltic and Finnic

	Baltic	Finnic	Other comparanda	
			Indo-European	Uralic
'herring'	*sile/ik-ē-	*silakka	Gm. *sile/iT-	
'thousand'	*tûstant-(i)-	*tušaNt(ə)	Sl. *tū(t)sant-ī-	Md./Ma. *tüšäm-
			Gm. *tū(t)sn̥t-ī-	
'moss'	*saman-ā-	*sammal		Sá. *semol
				Sá. ?*siwmal

It is interesting to note that the three words point to a rather different contexts of borrowing. The word for 'moss' must be connected to the palaeo-Laplandic substrate and with some kind of physical migration either into or out of Lapland, but the word for 'herring' shows a more localized distribution, and perhaps points to an autochthonous fishing community around the Baltic coast, similar to the one surmised already by Šmits. Finally, the word for 'thousand' is widely distributed, and must either be considered an old *Wanderwort*, or perhaps be associated with a group of other widespread loanwords identified in West Uralic.

Although we should hesitate before drawing far-reaching conclusions on the basis of just a handful of words, the overall impression is of a rather complex language contact situation involving multiple donor languages. It seems unlikely that the pre-Indo-European and pre-Uralic languages of north-eastern Europe represented a monolith, and it is probable that multiple source languages contributed to the substratal lexicon of the attested languages.

Some support for this argument could be the words for 'seal' in Baltic, Finnic and Sámi, which all appear to derive from different foreign sources:

- Lt. *rúonis*, Lv. *ruônis* ‘seal’, which is clearly related to, but not regularly cognate with OIr. *rón*, Breton *reunig* ‘seal’ (see p. 416–417)
- F *hylje*, E *hüljes*, Li. *īlgaz* (< **hülkes*) ‘seal’, which seems to be connected to, but is hardly loaned from, ON *selr*, OHG *selah* (cf. Suolahti 1899: 64) (< **selha*-; Iversen/Kroonen 2017: 519)[255]
- Sá. S *nåervie* (< **noarvē*) ~ Sk. *nue´rjj* (< **nuorjē*) ‘seal’, which are perhaps irregularly connected to Finnish *norppa* ‘ringed seal’ (Aikio 2004: 15)

We could argue that these terms originally referred to different kinds of seal, but there is no indication that this should be the case, as they represent neutral terms in all of the languages where they are attested. On this basis, we might assume that Finnic and Baltic interacted with distinct fishing populations speaking potentially unrelated languages. Such a scenario can certainly not be ruled out, and perhaps more such cases could be identified with further research.

As a final note, I would like to point out that the dearth of evidence adduced here cannot be taken as an indication that Finnic and Baltic have been only minimally affected by palaeo-Baltic languages, but simply that very little can be identified. Given that my methodology demands both the survival of the word in multiple branches, and the demonstration of identifiable irregularities, we cannot expect a huge amount of data to be available to us. Furthermore, it is naturally more

[255] Sadziński/Witczak (2016: 58–59) have additionally compared Norwegian Sámi (19th c.) *dullja* ‘(a kind of) seal’ (Stockfleth 1852: 694), for which they provide an arbitrary, and entirely erroneous, Proto-Sámi reconstruction **tüľɣa*. This Sámi word is confined to older lexical sources, and looks to be an unexplained variant of Sá. N *dealljá* ‘harp seal’ (< PSá. **tealjā*). Any kind of connection with Finnic **hülkes* is more or less excluded on phonological grounds.

difficult to substantiate a substrate origin based on words in unrelated languages, because apparently irregular substitutions in loanwords can often be accounted for by assuming different chronologies or dialectal differences, while such options are usually unavailable when dealing with exceptionless sound laws. It is merely a matter of fortune that enough material has survived in these three cases to allow us to make a case for a palaeo-Baltic origin. In fact, many more of the suggested Baltic loanwords in Finnic are of unclear ultimate origin, but with the tools currently available to us, this can only serve as a statistical argument. If this area of research continues to be pursued, I am confident that more hard evidence will be uncovered.

1.3.6. Analysis of contact relationship

1.3.6.1. Animal husbandry

While no Baltic words related to cattle appear to have been loaned into Finnic, it is highly remarkable that two loanwords related to horse breeding seem to surface in Finnic as cattle terminology. Thus, Finnic **ẹhva* 'heifer' and **härkä* 'ox' can plausibly be analysed as loanwords from the Baltic words for 'mare' and 'male horse', respectively. The application of terminology for one domesticate to refer to another is trivial; a parallel can be seen in the adoption of the same Finnish *härkä* in North Sámi as *heargi* 'draught reindeer'. However, as with the Sámi example, such a shift does most probably point to a difference in animal husbandry practices. The 9th century traveller Wulfstan of Hedeby remarked that the Balts consumed mare's milk and ate the

meat of their draught animals (Gimbutas 1963: 25–26).[256] The milking of horses was potentially already practiced by early Indo-Europeans, as evidenced by *Equus* milk peptides identified in the dental calculus of two Yamnaya individuals from the western Steppe (Wilkin et al. 2021: 630). A possible analysis would be to associate the semantic shift from 'horse' to 'cow' with a transition from horse to cattle as milk animals. Remarkably, Proto-Finnic **lehmä* 'cow' is the phonetically regular equivalent of the Mordvin word **lišmə* (> E *l'išme*, M *l'išmä*) 'horse' (cf. Ojansuu 1908: 32), which might be understood in a similar mixed Finnic-Baltic cultural context.[257]

In this context, it is worth noting the remarkable absence of horse remains in Baltic Corded Ware material (Piličiauskas 2018: 186). This is typical of the Corded Ware culture in general, where the few extant horse remains belong to local wild populations (Librado et al. 2021). This is problematic to the identification of the Balts with Corded Ware. On the other hand, horse teeth are prevalent in barrow cemeteries from the Late Bronze/Early Iron age associated with the hillfort phenomenon (Merkevičius/Muradian 2016; Аллмяэ et al. 2018: 350; Legzdiņa et al. 2020: 1846). This must indicate a certain cultural significance of horses in the Baltic region, but admittedly does not directly inform us of their domestic status. In the Fatjanovo-derived

[256] This tradition appears to have been continued by the Prussians until at least the 15th century, as shown by the gloss *aswinan* 'kobilmilch' in the Elbing Vocabulary (see further Топоров ПЯ I: 135–136). Note that this word is derived from the same Baltic **ešvā-* which was loaned into Finnic.

[257] As another parallel for such a semantic shift, compare Ket *ku'ś* 'cow' as against Yugh *ku'ś* 'horse' (cf. Fortescue/Vajda 2022: 268). Apparently, Proto-Finnic speakers were not introduced to milking by Indo-Europeans, as the word **lüpsä-* 'to milk' appears to have been adopted from an unidentified source, from where it also entered Mordvin, Mari and Permic (Aikio 2015a: 46).

Djakovo Culture, horse becomes the dominant domestic species during the Iron Age, contemporaneously to many cultural changes in the Eastern Baltic, while osteological evidence points to horse as a primary meat source (Кренке 2019: 43, 58).

Most of the loanwords associated with animal husbandry concern sheep and goat. In this domain we can count **vōhi* 'goat', **oinas* 'ram', **vōtnas* 'lamb', **villa* 'wool', **paimę̄n* 'shepherd', and probably **karva* '(animal) hair'. The earliest directly dated remains of domestic livestock in the Eastern Baltic date to the Middle Bronze Age, including a sheep/goat mandible from the mid-2nd millennium BCE in central Žemaitia (Piličiauskas et al. 2016: 186; Motuzaitė Matuzevičiūtė 2018: 152). While similar chronologies have also been suggested for Estonia (Lõugas, Kriiska & Maldre 2007: 25), this dating is not certain as none of the finds have been radiocarbon dated. Evidence for large-scale sheep and goat farming is not found until the Late Bronze Age, or the mid-1st millennium BCE (Rannamäe 2016: 23).[258]

Gimbutas (1963: 35) includes **hanhi* 'goose' among her list of domestic species. According to Lang (2016: 17), the word must have referred to a wild species as, in his view, goose domestication took place no earlier than the 1st millennium BCE in Southern Europe. However, recent research has established that the domestic and wild goose diverged as

[258] Unworked bone remains may have been misdated due to layer mixing, while worked remains found in grave sites might be trade items (Lõugas, Kriiska & Maldre 2007; Rannamäe 2016: 23). While Rannamäe et al. claim that the earliest sheep bones date from 1200 BCE, i.e. the Bronze Age, only one sample has been dated so early (1200–800 BCE) by archaeological context. Furthermore, two bones from the same site which have been radiocarbon dated belong to the Late Iron Age and Modern Period, respectively, suggesting the possibility that the third bone has also been misdated. The oldest directly dated sheep remains from Estonia are found in Asva on Saaremaa, dating to 786–522 BCE.

long as 5000 years ago (Heikkinen et al. 2015). In the Baltic, it has been claimed that the domestic goose emerged in the Middle Ages, but a recent study based on isotope analysis has identified potential domestic specimens in Estonia from the Late Iron Age (Ehrlich et al. 2021). The evidence is therefore not as conclusive as Lang would imply, but it must be admitted that concrete indications of domestic geese at a sufficiently early date appear to be lacking.

1.3.6.2. Agriculture

Many agricultural loanwords from Baltic into Finnic constitute generic terms: **sēmen* 'seed', **hęina* 'hay', **pęlut* 'straw chaff'. More notable is the word **hernes* 'pea', a plant which is first recorded in the Eastern Baltic in the mid-1st millennium BCE (Pollmann 2014; Grikpėdis/ Motuzaitė Matuzevičiūtė 2017: 6; Minkevičius et al. 2020). This coincides with a general diversification of cultivated crops in the Late Bronze Age/Early Iron Age, again associated with the hillfort phenomenon (Lang 2007; Motuzaitė Matuzevičiūtė 2018: 156).

The loaning of the word for 'pea' specifically is striking, as while the plant is present throughout Northern Europe by the Late Bronze Age, it is comparatively infrequent, implying a low economic significance (Grabowski 2011: 488; Stika/Heiss 2012: 192). Etymologically, the Baltic word is a specialization of an inherited generic term for 'grain', which might point to the crop becoming a staple among Balts. This is not supported by the existing evidence from the East Baltic, however, where the pea is recorded with the lowest frequency of all crops, matching the situation in the rest of Europe (Pollmann 2014: 409). The Finnic words for other specific crops are not Baltic loanwords; **vehnä* 'wheat' may only indirectly be connected with the Baltic word **avižā*- 'oats' (for a discussion, see p. 376–378). Considering that the first

small-scale agriculture in the East Baltic appears to have been exclusively barley-based (Motuzaitė Matuzevičiūtė 2018), it is interesting to note that Finnic **osra* (~ ?**ocra*) and Baltic **mę̄žjai* 'barley' are both of obscure origin.[259] The above evidence appears to suggest that the Finnic speakers became acquainted with diversified agriculture by other means than through contacts with Baltic-speaking populations.[260]

1.3.6.3. The wheel

Interesting from a cultural perspective is **rattas* 'wheel (of a cart)', in the plural **rattahẹt* 'cart'. Archaeological evidence for wheeled vehicles in the Bronze and Iron Age Baltic appears to be largely lacking, and the introduction of wheeled vehicles in the Late Bronze Age has only been inferred by indirect evidence (Viires *apud* Lang 2007: 252). There also appears to be a general dearth of evidence for wheels in Central Russia throughout the Bronze Age, aside from a pair of pottery discs discovered in a child's grave in Balanovo, which has been interpreted as belonging to a model wagon (Piggot 1969: 302). According to Lang (2007: 252), **silta* 'bridge' may also have been loaned in connection with wheeled vehicles (cf. von Hertzen 1973: 85), and may originally have referred to trackways across swampy areas, traces of which can be identified since the Roman Iron Age. Note the etymological connection of the Baltic source with Lt. dial. *tìlės* 'planks

[259] For a discussion of various attempts to etymologize the Baltic word, see Kroonen et al. (2022: 15). The Finnic word has been derived from Indo-Iranian (Holopainen 2019: 155–156), but since the proposed source does not correspond in sense, the etymology is doubtful.

[260] Finnic **rukis* 'rye' and **kakra* 'oats' have been adopted from Germanic, cf. Häkkinen/Lempiäinen 1996: 167–173.

(as paving)' (cf. also F *silta* 'wooden floor') might further support such an original meaning.

1.3.6.4. Context

Many scholars have characterized the Baltic–Finnic contact relationship as long-term, in some cases as having lasted millennia (Kallio 2008; Vaba 2011: 756; Lang 2016). In this context, it has been claimed that Finnic would have come close to being fully assimilated by Baltic, before eventually becoming dominant (Lang 2018a: 29). This scenario seems unnecessarily complex; furthermore, the structural influence of Baltic on the Proto-Finnic phonemic system appears to have been minimal, which contrasts strongly with other cases of intense language contact eventually leading to language replacement, such as Latvian and Livonian (Suhonen 1973: 53–66) or Veps and Russian (cf. Зайцева 2008: 79). The conservative phonology of Finnic from a Uralic standpoint makes it unlikely that it was almost replaced by Baltic, and rather speaks in favour of the assimilation of a Baltic dialect into Proto-Finnic (cf. Kallio 2015: 90; Kallio forthc.). Moreover, the most important linguistic evidence for a long-term contact relationship concerns the substitution of the Proto-Baltic long vowels **ē* and **ā* (Junttila 2012: 266); however, as argued in *1.3.5.2*, the different substitution patterns need not necessarily be analysed as evidence of chronological differentiation. Even if they are, this would not necessarily imply continuous, long-term contact.

A relatively large proportion of the Baltic loanwords constitute what Lang (2016: 17) has referred to as "luxury borrowings", i.e. loanwords which cannot be connected with the transfer of cultural practices or material goods. It is highly remarkable that the Baltic loans in Finnic include several kinship terms, in particular **sęsar* 'sister', **tüttar*

'daughter', **nępat* 'nephew, niece', **morcijan* 'bride'. The term **atęiva*, which is recorded in the sense 'marriageable woman' in Veps (Зайцева 2010: 18) and as 'married woman visiting her parents' in Finnish, in combination with other borrowed words for female family members, is likely to suggest exogamous marriage practices (Gimbutas 1963: 36; Lang 2015: 72). Genetic studies of European populations have repeatedly referred to female exogamy as a driver of intercultural contact in the Corded Ware up until the Bronze Age (Knipper et al. 2017; Mittnik et al. 2018, 2019), although there have not been any studies investigating this phenomenon in more recent periods, or further east, where the contacts are most likely to have taken place, so it cannot yet be confirmed whether such a hypothesis is supported by the genetic evidence.

As "luxury" loanwords, we can also consider terms connected with topography and nature, such as **męcca* 'forest' and **halla* 'frost', and the names of animals of low economic significance — here, we are largely dealing with those that have a negative connotation — **herhiläinen* 'hornet', **kīli(l)äinen* 'botfly', **vapsas* 'wasp', **kärmes* 'snake'. In addition, we find the basic adjectives **ahtas* 'narrow', **tühjä* 'empty', **haljas* 'green', **kęltainęn* 'yellow', and the body part terms **hammas* 'tooth' and **napa* 'navel'. From a typological perspective, the last two are particularly remarkable: according to the WOLD database, both 'tooth' and 'navel' rank among the 400 least likely words to be borrowed.[261]

The above semantic clustering seems most coherent with a scenario involving a Baltic substrate in Finnic. Geographical terminology and words related to natural phenomena are frequently identified as

[261] Note the Romance substrate word *imlīq* 'navel' (< **imbilicus*, cf. Galician *embigo*) in Andalusian Arabic (Griffin 1959: 347).

characteristic of borrowings from linguistic substrates (e.g. Kalima 1919: 257–258; Bertoldi 1932: 94; Ariste 1971: 9–10; Saarikivi 2004; Aikio 2009: 41). Close semantic parallels for many of the borrowed animal names can be identified among the Finnic substrate words in Russian dialects, cf. R dial. *па́рма* 'botfly', *ки́гачи* 'gnats' (cf. Мызников 2019: 295), *товка́ч* (Шахматов *apud* Куликовскій 1898: 119) 'a kind of woodworm', *ши́жлик* 'lizard' (cf. Kalima 1919: 257; Мызников 2004: 113–116).

The strongest linguistic evidence for a mixed group involving bilingualism can be seen in the *plurale tantum* nouns **pę̄lut* 'straw chaff', **nītet* 'heddle' and **rattahę̄t* 'cart', which correspond to Baltic nouns also used exclusively in the plural (in the relevant meanings). This implies that the Baltic words were identified as plural upon borrowing, which can only be understood if we assume a certain level of bilingualism. This is particularly remarkable in the case of Baltic **nîtīs* 'heddle', where the ending is morphologically ambiguous, and could only be understood as plural by a person well acquainted with Baltic grammar. The hypothesis of a Baltic substrate that was ultimately absorbed by Finnic would further be supported by the evidence that the source of the Baltic loanwords was not the direct ancestor of any attested Baltic language (see *1.3.3.4*).

1.4. Loanwords into other Uralic languages

1.4.1. Sámi

Many of the originally Baltic loanwords have been loaned into Sámi through North Finnic, in several cases early enough be distributed throughout the entire family; compare the following:

- Sá. S *lijkie*, N *liigi*, K *li´jjg* 'surplus' (< **lijkē*) ← F *liika* (~ Lt. *liẽkas*)
- Sá. S *naepie*, N *náhpi*, Sk. *nää´pp*, Ter *nap̦p̦e* (< **nāpē*) 'navel' ← F *napa* (~ Lv. *naba*)
- Sá. S *daajvaj*, N *dávjá* , Sk. *täujja*, Ter *tajva* (< **tāvjā*) 'often' ← F *taaja* (~ Lt. *tánkus*)
- Sá. N *šaldi* 'bridge', Sk. *šâ´ldd* 'floor' (< **šę̄ltē*) ← F *silta*, K *šilta* 'bridge' (~ Lt. *tìltas*)

In the last case, a Finnic intermediary is proven by the initial consonant, which must be the result of the specifically Finnic change **ti* > *si* (and further North Karelian > *ši*). The other cases also show vocalic substitutions indicative of borrowing rather than common inheritance; note that stem-final **ā* is typical of younger loanwords (Aikio 2006b: 36).[262] Since distribution is not a decisive factor, it is occasionally difficult to rule out a common Finno-Sámic proto-form. For instance, both F *siemen* 'seed' and Sá. N *siepman* could theoretically

[262] A younger age of Sámi **tāvjā* might also be shown by the metathesis **vj* > **jv* in South Sámi, as South Sámi appears to have kept **vj* and **kj* distinct: cf. Sá. S *vuevjie* 'clothing insert' (= Finnish *vaaja* 'wedge' ← Baltic or Germanic; see p. 98); see Pystynen 2014b.

reflect a PU **säməⁿ* (compare F *kieli* tongue' = Sá. N *giella* 'language' < PU **kälə*), but the principle of parsimony speaks rather in favour of a Finnish transmission.

Comparing the list of Baltic loans in Sámi given by Sammallahti (1999: 410–411) with those accepted by Aikio (2012a: 107), it would appear that the latter's revisions mainly involve the removal of words which could equally be borrowed through Finnic. Thus, examples which show the correlation F *e*, *o* ~ Sámi **ea*, **oa* have been omitted, since although such correspondences are found in inherited words, they are also common in Finnic borrowings of all ages (Aikio 2006b: 31–34).[263]

However, not all of the etymologies originally accepted by Sammallahti are straightforward. The equation of Sá. N *gohčču̥t* 'call, order', Sk. *kåčča̮d* 'call' (< **kočču̥-*) and F *kutsu-*, E *kutsu-*, Li. *kutsõ*, 3PRES. *kutsūb* (< **kuccu-*) 'invite, call' with Lt. *kviẽsti* (3PRES. *kviẽčia*) 'invite' (Mägiste 1923: 35–36; Koivulehto 1986: 272–274) is doubtful, as the substitution of **vę̄* (or earlier **wei*/**wai*) with **u* lacks phonological plausibility (cf. Junttila 2011: 107).[264] Likewise, despite Nuutinen

[263] Specifically, PSá. **keartē* (> N *geardi*) 'time, layer, strand', **seaprē* (> L *siebrre*) 'company, society' and **loamē* (> N *loapmi*) 'gap, cleft' could just as well be loans from F *kerta*, *seura* and *loma* (cf. SSA I: 348, II: 90, III: 172, respectively). In addition, the etymology PSá. **piemmę-* (N *biebmat*, Sk. *peâm'mad*) 'feed, rear' ~ Lt. *penìmis* 'fattening pig' is explicitly rejected as phonologically problematic.

[264] According to Koivulehto **kų̮oit-ja-* was adopted as **kut-ja-* because "ein /j/ konnte [vor **cc*] nicht bestehen". However, a sequence **-jcc-* seems to have been possible even in inherited vocabulary: F arch. *seitsen* 'seven' (< **säiccen* < PU **ćäjćəmä*, Aikio in prep. 109–110) and *veitsi* 'knife' (< **väicci* 'knife' ?< **väjćə* ~ Ko. dial. *vę́-* 'cut slantwise', Hungarian *vés* 'chisel, cut', cf. UEW 565), where it results from a fortition **-jć-* > **-jcc-*. Koivulehto is led astray by the notion that **j* in these stems derives from an earlier **ŋ* (Koivulehto 1981: 169; cf. PF **suiccę̮t* > F *suitset*, Võ. *suidsõq*, Li. Salaca *suiksud* 'bridle', from virtual **ćuwə-ŋćə-*, cf. **suu* > F, E *suu*, Li. *sū* 'mouth'), for which there is no evidence

(1992), Sá. N *bievla*, Sk. *piõull* (< **pievlę*) = F *pälvi* ‘snow-free patch (in spring)’ cannot be compared with Lt. *paĨvė* ‘wind-levelled plain among dunes’ (← Pr., cf. Sabaliauskas 1974), as the assumed metathesis in Sámi is *ad hoc*. The Sámi and Finnish words rather presuppose an earlier **päwlə*.

The old derivation of Sá. N *giehka*, Sk. *ǩiõkk* (< **kiekę*) and F *käki*, E dial. *kägi* (usu. *kägu*), Li. *ke'g* (< **käki*) ‘cuckoo’ from Lt. *gegẽ* ‘cuckoo’ (Thomsen 1890: 172) can also not be accepted. As recognized by LEW (142–143), the Lithuanian form is a recent clipping of the inherited Lt. dial. *gegužẽ* (= Lv. *dzęguze*; OR жегъзоулѧ, cf. СДРЯ 11–14 III: 238; Николаев 2020: 593), and can hardly be dated to Proto-Baltic. Furthermore, the Finnic and Sámic words are most likely regularly cognate with Khanty **käɣ-əj* (Vakh-Vasjugan *köɣi*, Surgut *kə̆ɣʷi*) ‘cuckoo’ (< PU **käkə*).[265]

In addition, many of the etymologies involve serious obstacles on the semantic side. In the following cases, the incompatibility in meaning makes the equation highly improbable:

– N *duollji* ‘hide, skin rug’, Sk. *tue´l'lj* (< *tuoljē*) = F *talja* ‘animal hide’ ~ Lt. *dalià* ‘fate’, Lv. *daļa* ‘portion, share’ — Koivulehto (1984: 12) attempts to bridge the semantic gap by comparing OE *scearu** (attested *scaru*, OBL.SG. *sceare*) ‘division’ (> MoE *share*) and R arch. *скорá* ‘hide’, but both words must be understood as parallel deverbal formations (cf. OE *sceran* ‘cut,

(Aikio loc. cit.). As a result, there is no reason to suspect that a form **kuiccV*- should have been phonotactically impossible.

[265] Compare Vakh-Vasjugan *wöɣ*, Surgut *wə̆ɣʷ*, *wə̆̈ɣʷ* ‘strength’ (< **wäkə*, UEW 563). Note, however, that Aikio (2015b: 2–3) has suggested that **ä* regularly yielded Khanty **ŭ* before **k* in Uralic *ə*-stems. He does not mention **käkə* as a counter-example, presumably because he considers the Finno-Sámic word to be a Baltic loan (Aikio 2012a: 107).

divide'), and do not constitute a parallel for the semantic shift 'share' > '(animal) hide'.

- N *faggi* '(wooden) hook', Sk. *vâ´ǧǧ* 'wooden hook, pothook' (< **vęŋkē*) ~ Lt. *vìngis* 'bend, turn, bypass' — A root etymology; the meaning 'hook' is unattested in Baltic (cf. Kalima 1936: 178–179).[266]
- N *johtit*, Sk. *jåå´tted* (< **jotē*-) 'go, travel, migrate' ~ Lt. *judé´ti* 'move (about), be restless' — The original meaning in Baltic is certainly not 'move', but rather 'be restless', cf. the glosses 'arguo, obiurgo' in Szyrwid (cf. ALEW 491), Lv. dial. *jũdît* 'unruhig machen' (ME II: 120) and the cognates Skt. *yúdhyate* 'fight', To. A *yutk*- (< **(H)ieud*^{h}*-sk*-) 'worry' (IEW 511–512; LIV 225–226).
- N *luokta*, Sk. *luhtt* (< **luoktę*) = F *lahti*, E *laht* (< **lakti*) 'bay, inlet' ~ Lt. *lañktis* 'yarn winder' (Posti 1977: 267–268) — Rather a root etymology, comparing Lt. *leñkti* 'bend', from which other words for 'bay' have been derived, e.g. Lv. *lìcis* 'bay, inlet'. However, the etymology is suspect since the right combination of form and meaning is unattested in Baltic (Saarikivi 2004: 200).
- N *riessan* 'decorative fringe', Sk. *riõzzâm* 'collar band' (< **rieseęmē*) and the verb S *rïesedh*, N *riessat* (< *riesę*-) 'adorn' ~ Lt. *rìšti*, Lv. dial. *rist* 'tie (on, up)' — The semantics are not compelling.[267] Moreover, the Sámi vocalism is unexpected: **ie(–ę)* implies earlier **ä(–ə)*.

---

[266] Aikio (2009: 176–178) has previously suggested that this Sámi word is a palaeo-Laplandic substrate word in view of the variants **vęŋę* (L *vagŋa* 'hook, barb', Sk. *võŋŋ* 'snag, submerged tree stump') and N *vievgŋa* 'snag' (< **vievŋę*).

[267] Thomsen 'thread, rope; snare' (1890: 212, cf. SSA III; 72–73) takes PF **rihma* 'thread, rope; snare' from an *m*-derivative of Baltic **riš*-, citing Lt. *rišìmas* '(the process of) tying', which is a productive derivative which cannot be blindly projected to Proto-Baltic and the Latvian obs. hapax (Valle *apud*

- Sá. I *ruodâs*, Sk. *ruõddâs* (< **ruontęs*) ‘wrist of a glove’; F *ranne*, E *ranne* (< **rantęh*) ‘wrist’ ~ Lt. *grandìs* ‘Armband’ (Ruhig II: 31) (Liukkonen 1999: 116–117) — LÄGLOS (III: 125) has already questioned the plausibility of the semantic shift ‘bracelet’ > ‘wrist’, but the situation is in fact worse, since the sense ‘bracelet’ is limited to a single lexicographical source, while the usual meaning in Lithuanian is ‘(metal) link, ring’, cf. also Pr. E *grandis* · rincke ‘beam link on a plough’ (Trautmann 1910: 342).
- Sá. S *saertie* ‘reindeer heart (as food)’ (< **sārtē*) ~ Lt. *šerdìs*, Lv. *serde* ‘core, kernel’ (Koivulehto 1990: 150) — The South Sámi form is cherry-picked. The other Sámi cognates: N *sárdi* ‘rib without meat, strip of reindeer liver’, Sk. *sä´rdd* ‘small piece of meat’, K *saa´rrd* ‘broad cut of meat’ clearly show an original meaning ‘piece of meat’.[268]

Next, there are a number of seemingly unproblematic Baltic loanwords with Finnic equivalents which did not feature in the above discussions due to their uncertain Indo-European background. Note, for instance, the following:[269]

Mancelius) *riffamais* ‘band’ (ME III: 531; the definite form of the gerundive adjective). Liukkonen (1987: 9) has assumed an unattested source **rišma-*. Grünthal (2012: 328–329) also analyses Md. E *ŕiśme*, M dial. *ŕiśmä* ‘chain’ as a Baltic loan, but this is more convincingly derived from Indo-Iranian, cf. Skt. RV *raśmā́* INST.SG. (or NOM.SG., Jamison *Commentary* VI: 103–104) ‘rein’, Parth. *rsn* /rasan/ ‘rope’ (< **raćm̥n-o-*; Lubotsky 2001: 314), cf. Holopainen 2019: 207–208.

[268] In a more geographically limited area, we also find a meaning ‘piece of fabric’, cf. N *sárdi* ‘piece of a tent’, L *sárdde* ‘strip of canvas’, which probably suggests a basic meaning ‘piece, strip’.

[269] Note, similarly, the above discussions of **luosę* ‘salmon’ (= PF **lohi*; p. 171), **luovē* ‘raised platform’ (= PF **lava*; p. 176), **sęsnē* ‘tanned reindeer

- Sá. N *gahpir*, Sk. *keä´p'per* (< **kępērē*) ‘hat, cap’ = F *kypärä* ‘helmet’, E *kübar* ‘hat (with a brim)’, Li. *kibār* (< **küpärä*) ‘hat’ ~ Lt. *kepùrė*, Lv. *cępure* ‘hat’ (Thomsen 1890: 185) — The Baltic word has no plausible comparanda beyond Slavic (R *ченéц*, SCr. *čèpac* ‘kind of cap’, cf. REW III: 316, ALEW 552)[270]
- N *guovllas*, Sk. *kuvlâs* (< **kuovlęs*) ‘wooden collar band’ = F *kaula*, E *kael*, Li. *ka'ggõl* (< **kakla*) ‘neck’ ~ Lt. *kãklas*, Lv. *kakls* ‘neck’ (Thomsen 1890: 177) — The Baltic word is of uncertain etymology (cf. ALEW 502–503)[271]
- Sá. L *muolos*, Sk. *muâlas* (< **muolōs*) ‘shore lead, i.e. strip of ice melt along the shoreline’ = F dial. *malo* ‘edge, flank’, K dial. (N) *malo* ‘shallows, shoreline’ ~ Lv. *mala* ‘edge, shore, boundary’, Lt. dial. *lýg-malis* ‘filled to the brim’ (Loorits *apud* Mägiste 1939: 68–69; Nuutinen 1987a) — The Baltic word is etymologically ambiguous.[272]

---

leather’ (= PF **hihna*; p. 155–157), **vuoręs* ‘old’ (~ PF **varas*; p. 174). On Sámi **kuojmē* ‘companion’, see p. 98.

[270] The loan etymology requires the assumption of a metathesis **käpürä* > **küpärä* (Thomsen 1890: 96; Kalima 1936: 124). The alternative interpretations of the Finnic word as a native formation (Mikkola 1930: 442; Nilsson 1996) are unconvincing.

[271] The traditional etymology (Mikkola 1896b: 218; Trautmann 1923: 125; Derksen 2015: 220) compares Skt. *cakrá-* ‘wheel’, but this is semantically problematic, and the etymology is not taken up by Smoczyński (2018: 469–470). While Walde/Pokorny’s (I: 515) “‘Hals’ als ‘Dreher’” would have semantic parallels (cf. OR воротъ ‘neck’ beside воротитиса ‘return’, СДРЯ 11–14 I: 477; MP *grdn* ‘neck’ from *grd-* ‘revolve, turn’, Durkin Meisterernst 2004: 163), the word **k^{w}ek^{w}l-(o)-* was specialized in the meaning ‘wheel’ already in PIE, and the notion that it could have uniquely preserved an abstract meaning ‘turner’ in Proto-Baltic is far from trivial. Note that Grinaveckienė/Mackevič (1989: 74) have even suggested the Baltic word was borrowed from Finnic.

- N *ruoida*, Sk. *ruõidd* (< **ruojtę*) ‘shin, thigh’ = F *reisi*, E *reis*, Võ. *rais* (< **rę̄ici*) ‘thigh’ ~ Lt. *ríetas*, Lv. *riẽta* ‘thigh, loins; ham (of meat)’ (Thomsen 1890: 212) — The Baltic word has no certain comparanda beyond Slavic (SCS рить, Cz. *řiť* ‘anus’). The equation with Arm. *eri* ‘shoulder of animals’ (IEW 863) is uncertain (Martirosyan 2008: 263)
- N *suolu* ‘island; isolated patch of forest’, Ter *sįelaj* ‘island’ (< **suolōj*) = F *salo* ‘dense forest; island; elevated spot in a swamp’, E *salu* ‘grove; area of woodland in an open landscape’ (< **salo*)[273] ~ Lt. *salà*, Lv. *sala* ‘island; elevated spot in a swamp’[274] (Thomsen 1890: 214; Kalima 1936: 158) — The Baltic word lacks a satisfactory etymology[275]

---

[272] More plausible than the comparison with Sln. *molę́ti* ‘jut, protrude’ (IEW 721–722) is the connection with OIr. *mala* ‘eyebrow’ < **mlH-* (Pedersen 1913: 99; as a semantic parallel, OIr. *brú* ‘edge, shore’ < ‘brow’, see eDIL s.v.). Alternatively, compare ON *mǫl* ‘shingle, gravel bank’ < **malō* (but see de Vries 1962: 401).

[273] The Finnic and Sámi words reconstruct to a common proto-form **salaw* (Kuokkala 2012: 78) with which we may compare Lt. obs. *salavà* ‘island, river island (German *Werder*)’ (Bezzenberger 1877: 320; Ruhig II: 399), adduced already by Thomsen. In terms of word formation, *salavà* stands quite apart from other words with a suffix *-ava*, which usually have a collective meaning (cf. Skardžius 1941: 379–380).

[274] Despite the *communis opinio* (ME III: 64; LEW 758; Smoczyński 2018: 1126–1127), it seems obvious that Lt. *salà*, Lv. *sala* ‘village’ is borrowed from Bel. *сялó*. It is hardly a coincidence that the Lithuanian word is practically limited to Vilniškiai dialects where *s* is regularly depalatalized (Zinkevičius 1966: 165 and Map 74; cf. Smalinskienė 1994: 178).

[275] Latin *īnsula* ‘island’ is hardly to be separated from OIr. *inis* ‘island’ (Ernout/Meillet 319–320; de Vaan 2008: 306). Endzelīns (ME III: 664) proposed that *salà* was abstracted from **api-sala* ‘that which [water] flows around’, but such a form is unattested, and the verbal root **sal-* ‘to flow’ is itself supported by doubtful evidence (Jakob forthc. b.). While a Baltic source

It cannot be excluded that the above words originated in Baltic, but depending on one's assessment of the existing etymologies, alternatives cannot be excluded. Where the Baltic word is isolated within Indo-European, an early loan from Finnic into Baltic may still be on the table. Where Slavic equivalents are attested, this becomes far less likely; however, the possibility remains that the words in question are parallel loans from unattested source languages.

This brings us to the unambiguous cases. Of the Finnic etymologies accepted in *1.3.2*, four of them have a Sámi equivalent which cannot be explained as a recent loan from Finnic:[276]

- N *sarvva*, Sk. *sõrvv* (< **sę̄rvę̄*) 'elk' = PF **hirvi* 'deer; elk' ~ Pr. E *sirwis* 'roe-deer'
- N *suoldni*, Sk. *sue´lnn* (< **suolnē*) 'mist over water in late summer; hoarfrost' = PF **halla* 'frost, hoarfrost' ~ Lt. *šalnà*, Lv. *saĺna* 'hoarfrost'
- N *suorri*, Sk. *sue´rr* (< **suorē*) 'branch, fork' = PF **hara* 'branch, fork' ~ Lv. *zars* 'branch, prong'
- N *duovli*, Sk. *tu´vll* (< **tuovlē*) 'tinder (as a traditional remedy); tinder fungus' = PF **takla* 'tinder' ~ Lv. dial. *dagla*, *dęgla* 'tinder, fire sponge' (ME I: 430, EH I: 313)
- S *vuessie*, N *vuos'si* (< **vuossē*) 'handle (of a cooking pot, bucket)' = PF **ansa* 'handle' ~ Lt. *ąsà*, Lv. *ùosa* 'handle, eyelet'[277]

---

is usually assumed (LEW 758; Sammallahti 2001: 411; Aikio 2012a: 107), Thomsen and Kalima both admit the possibility of a Finnic → Baltic loan (cf. also Bednarczuk 1976: 52), and others have suggested a loanword from an unknown source (Saarikivi 2004: 208; Aikio 2004: 24; J. Häkkinen 2009: 48; Holopainen, Kuokkala & Junttila 2017: 129).

[276] I can only imagine that **suolnē* was an accidental omission in Aikio (2012a: 107). As for **vuossē*, the omission perhaps follows from the fact that Sammallahti considered a Germanic etymology equally possible (see fn. 120).

Aikio lists **sǫ̊rvēs* (> N *sarvvis*, Sk. *sââ´rves*) 'uncastrated reindeer buck' as a separate loanword. In reference works (SKES 77–78; SSA I: 167), the Sámi word has been equated with F *hirvas*, K *hirvaš* 'uncastrated reindeer buck'. However, given that this word is only known in the northern dialects of Finland and Karelia, combined with the exact semantic correspondence with Sámi, it seems much more probable that it is a partial calque resulting from a crossing of the native *hirvi* with Sámi **sǫ̊rvēs* (Junttila in prep. s.v. *hirvas*; cf. also the direct loanword *sarvas*; Aikio 2009: 276). Note that the interpretation of F *hirvas* and *hirvi* as independent loanwords from the Baltic masculine **širvas* (= Pr. *sirwis*) and the feminine **širvē*, respectively (Nieminen 1940: 378), could also be applied to Saami **sǫ̊rvǫ̊* and **sǫ̊rvēs*. However, it appears just as probable that **sǫ̊rvēs* as an inner-Sámi derivative with the same suffix as in Sá. S *urries*, Sk. *åå´res* (< **orēs*) 'male (animal)'.[278]

All the cited words in Finnic and Sámi can be given a common Uralic proto-form. This can be interpreted in at least three ways: (a) the loans

---

[277] For the development **-ns-* > **-ss-*, compare Sá. N *guos'si*, K *kū´ss* 'guest' (< **kuossē* = F *kansa* 'guest'); cf. Sammallahti 1998: 54.

[278] The *Inarilappisches Wörterbuch* attests the form *mečči*n /meč'čin/ 'im Walde' which would reflect the inessive singular of a word **meaččē*. The latter has been interpreted as an early loan from Baltic (Aikio 2012a: 107). The other Sámi languages attest a similar but irreconcilable **meaccē* (> Sá. S *miehtsie*, Sk. *meä´c'c* 'forest'), pointing to a later loan. The Inari form corresponds formally to Lule *miehttjen* 'against the wall of the tent, as far as possible from the hearth', which is semantically aligned with the other West Sámi languages, e.g. South Sámi *miehtjiedidh* 'move away, put by the wall', Ume *miehttjiedit* 'remove the pot from the fire'. South Sámi *meahtsanidh* 'withdraw oneself to the wall (of the tent)' = N *meahccánit* 'stray too far (of cattle)' might show the confusion of the two word families. The question, then, is whether Inari *mečči*n 'im Walde' should also be explained as the result of a contamination of the Finnic loan *mecci* 'forest' and **meaččē* 'far away (from the hearth)'.

were adopted into a single Finno-Sámic proto-language; (b) the Sámi forms are very early adoptions from Finnic, predating most of the sound changes; or (c) the words were adopted independently by Sámi and Finnic, and the fact that they go back to identical proto-forms is due to coincidence.

The main issue with option (a) is that the reconstruction of a Finno-Sámic branch is nowadays increasingly disfavoured, with the shared features being explained as the result of secondary areal diffusion (T. Itkonen 1997; Salminen 1999: 20–23; 2002: 47–48; Saarikivi 2011: 106–109; Aikio 2022: 3–4). In view of this, Aikio (2012b: 73) opts for option (b), assuming that all of the Baltic loanwords in Sámi were mediated through Finnic, an opinion that was already held by Thomsen (1890: 28–29). A number of criticisms can be raised against this.

First, there is no unambiguous evidence for Finnic loanwords in Sámi of a sufficiently early date. Very few identifiably Finnic loanwords predate the Sámi vowel shift, and even Sámi **puošē* (> N *buošši*) 'angry', apparently an early loan from PF **paha* 'bad, evil' (Sammallahti 1998: 183), evidently postdated **š* > **s*.[279] It is unclear, however, exactly by what criteria such early mutual loanwords could be distinguished from common inheritances, and the existence of unidentifiable borrowings from this period cannot be excluded (Aikio 2012b: 72).[280]

[279] Whether or not they were transmitted through Finnic, it is possible that **sę̄rvę̄* 'elk' (← **širvas*) and **suolnē* (< **šaĺnâ*) 'mist over water' did not in fact predate the pre-Sámi change **š* > **s*, but merely predated the innovation of a new phoneme **š*, as in the absence of such a phoneme, a substitution **š* → **s* would be in line with expectations (cf. Kallio 2009: 34).

[280] Although we might expect more such traces in the case of intense early language contact. For instance, one might anticipate evidence of the Finnic merger **t*, **d*, **d'*, **č* > **t* in the Sámi material.

Next, there are a number Baltic loanwords in Sámi which are unattested in Finnic. To my mind, there are three plausible examples:

- N *giehpa*, Sk. *ǩiõpp* (< **kiepę*) 'soot' ~ Lv. *kvę̂pi*, dial. *kvêpji* PL. 'soot' (cf. Sammallahti 1998: 127).
- N *loggut*, K *lå´ŋŋge* (< **loŋkō-* ~ **loŋkē-*) 'strip (birch bark); peel' ~ Lt. dial. *lùnkas*, Lv. *lûks* 'bast'; cf. further Pr. E *lunkan*; R *лы́ко*, SCr. *lȉko* 'bast'. A nominal form is attested in the Sámi loanword F *lunka* 'bark which flakes off easily' (Aikio 2009: 115–116)
- N *vietka*, Sk. *viõtkk* (< **vietkę*) 'adze' ~ Lt. *vedegà* 'adze', Lv. *vę̂rga*, dial. *vędga* 'ice chisel'; Pr. E *wedigo* 'carpenter's axe'

The latter two etymologies do not involve any significant formal or semantic issues,[281] but cannot be considered unambiguous evidence of direct contacts between Sámi and Baltic, as the Baltic words themselves do not have reliable Indo-European etymologies. The word for 'bast' has a potentially irregular comparandum in Slavic (see p. 291), while the word for 'adze' contains an opaque suffix **-eg-* otherwise found only in the equally obscure Lt. *uodegà* 'tail' (the derivation of the latter from *úodas* 'mosquito' *per* ALEW 1328 and Smoczyński 2018: 1563 wants semantic parallels).

Things look more positive in the case of 'soot'. Latvian *kvę̂pi* is clearly related to the verb *kvêpt* 'smoke, smell; get covered in soot', Lt. *kvẽpti* 'breathe, blow; smell', *kvãpas* 'breath, smell'. Although further connections are difficult,[282] the inner-Baltic etymology seems solid. In

[281] Admittedly, Sá. **vietkę* (< **wätkə*) is perhaps not quite expected from Baltic **vedegâ*. One might rather anticipate ***wätäkä* (> Sá. ***vātēkē* > Sá. N ***váhttit*).

[282] Possibly here also R *кóномь* 'soot', but the loss of **u̯* is irregular. Contrary to Schrijver (1991: 260–263) and Derksen (2015: 268), the Latvian acute

Sámi, the substitution of **kv-* with **k-* is paralleled by Torne Sámi *gierdnas* 'mill, grinder' ← Nw. *kvern* and South Sámi *gearhka* 'throat' ← Nw. *kverk* (Quigstad 1893: 14; Koivulehto 1992a: 92). The question still stands as to whether this word for 'soot' might once have existed in Finnic but was subsequently simply lost (Aikio 2012a: 74). The Proto-Finnic word for soot (F *noki*, E *nõgi* < **noki*) is itself of obscure origin, but it is possible that it replaced an earlier Baltic loan **kēpi*. While this cannot be excluded, reconstructing unattested Finnic words to explain away evidence of direct contact is, of course, circular.

Moreover, Aikio (loc. cit.) and Saarikivi (2022: 33) do still admit the possibility of direct contact on the basis of one example, namely Sá. S *liejpie*, N *leaibi* (< **leajpē*) '(grey) alder', which for phonological reasons could not have been adopted through Finnic **leppä* 'alder', and whose preform **lejpä* actually more closely resembles pre-Baltic **leîpâ* than the Finnic word does. As I have discussed in detail above (see p. 157–158), there are several problems with this word family which make a simple Baltic loan hypothesis unsatisfactory; if there is any relation at all, it is most probably a shared substrate word in Balto-Slavic and West Uralic. Although the word for 'alder' may not be reliable, there is another word which provides evidence of direct, independent contact between Baltic and Sámi:[283]

---

hardly warrants the awkward reconstruction **kh₂uep-*, which cannot in any case account for Gr. καπνός 'smoke', Lat. *vapor* 'steam, heat' (IEW 596–597). We are probably dealing with metatony, as in Lv. dial. *drêbt* 'beat; sleet' vs. Lt. *drẽ̃bti* (cf. footnotes 560 and 561) and likewise Lv. *têst* 'carve; shave' vs. Lt. *tašýti* (cf. footnote 91).

[283] South Sámi *daktere* 'daughter (by marriage)', Pite ⟨taktier⟩ (Lehtiranta 1989: 130) is presented by Sammallahti (1998: 127) as an example of a word which could not have been mediated by Finnic (where we find **tüttär*). It is, however, more probable that the Sámi word is of Norse origin, especially in view of the limitation to the western edge of the family (cf. already Qvigstad

– N *suoidni*, Sk. *suei´nn* (< **suojnē*) ‘grass, hay’ ≠ F *heinä*, E *hein*, Li. *āina* ‘hay’ ~ Lt. *šiẽnas*, Lv. *siens* ‘hay’

While the Finnic word, partly on the basis of the Sámi evidence, has usually been reconstructed **šaina*, I have argued above extensively (see p. 122–129) that the word in question should be reconstructed **hęina* for Proto-Finnic. If this is correct, then the Sámi equivalent cannot have been adopted through Finnic, but must rather represent an independent loanword. The difference can be explained by assuming an earlier chronology, namely a date before the Baltic monophthongization **ai* > **ę̄*.[284]

Nevertheless, it is possible to imagine an alternative account whereby pre-Sámi **šajna* was, after all, borrowed from pre-Finnic **šęjna*. The Sámi change **i* > **ę* is a recent innovation (Sammallahti 1998: 106), postdating at least some of the Norse contacts (Aikio 2006a: 15) and, more importantly, the Baltic loanword **sęrvę* ‘elk’ (?← pre-Finnic **širvə*) ← Baltic **širvē* (see above). Prior to this change, it is possible that pre-Sámi **a* in fact represented the phonetically closest match to Finnic **ę*, meaning that the substitution **ę* → **a* would be perfectly intuitive.

1893: 125). Although the word for ‘daughter’ in Old Norse is an assimilated *dóttir*, the older cluster *-ht-* is reflected as **-kt-* in several loans, cf. Sá. S *slikte*, N *livttis* ‘smooth’ (< **liktēs* ← **slihtaz*; cf. ON *sléttr*); S *raaktse* ‘harness trace’ (**rāktęs* < **drahtuz*; cf. ON *dráttr* ‘dragging’, Nw. dial. *drått* ‘trace’); N *divttis* ‘tight, watertight’ (< **tiktēs* < **þinhtaz*; cf. ON *þéttr*). See Posti (1953: 45).

[284] Similar argumentation can be made in the case of **ruojtę* ‘thigh’ (cf. PF **rę̄ici*, Lt. *ríetas*), however, in this case there is no evidence for an original **-ai-* diphthong in Baltic, and the word is of obscure origin. See above on p. 207.

Therefore, at least the word for ‘soot’ and perhaps also ‘hay’ might offer some evidence for direct contact between Sámi and Baltic. Even so, the fact that five out of six of the Sámi loans from Baltic are shared with Finnic can hardly be considered coincidental. It seems, therefore, that some of the relevant material must have percolated through a Finno-Sámic dialect continuum, but that does not exclude a small level of direct contact taking place between Sámi and Baltic.

Earlier and later loanwords

All of the examples mentioned above involving Baltic **š* show a reflex **s* in Sámi. Two other Sámi substitutions for this phoneme have been suggested in the literature: **č* and **š*. These examples have been used to support the idea of an older and younger layer of Baltic loanwords in Sámi, respectively. The former is supported by two etymologies (Kallio 2009: 32–33):

- Sá. N *čuorpmas*, Sk. *čuõrmâs* (< **čuormęs*) ‘hail’ ~ Lt. *šarmà*, Lv. *saŕma* ‘hoarfrost’ (Koivulehto 1983: 188–189)
- Sá. N *čohkut*, Sk. *čååkkad* (< **čokō-*) ‘to comb, currycomb’ = F *suka*, E *suga*, Li. *sugā* ‘currycomb; heckling comb’ ~ Lt. *šùkos* F.PL., Lv. *suka* or *sukas* F.PL. ‘comb, heckling comb’

Although the Baltic word for ‘comb’ has a probable cognate in Uk. *щетьо* ‘bristle’, Sln. dial. *ščę̑t* ‘brush, thistle’ (Pronk/Pronk-Tiethoff 2019: 285), the further etymology is uncertain (ALEW 1216) so that the exact relationship between the words cannot be determined. In the other case, there is a semantic difference. Although not prohibitive (compare Lt. *grúodas* ‘frozen ground’ vs. OCS градъ ‘hail’), it is questionable whether this single etymology is of sufficient calibre to carry the weight of an entire loanword stratum, especially given that a

term associated with a natural phenomenon such as 'hail' is an unlikely candidate for borrowing during incidental early contacts.

As for the supposed 'late' loans, Sammallahti (2001: 401) has suggested two which would show Sámi **š-* for Baltic **š-* (or **ž-*), and a third was later adduced by Koivulehto (*apud* Aikio 2009: 200). While Sammallahti assumed a Proto-Balto-Slavic source for these loanwords, the Sámi phoneme **š* is of recent, probably post-Proto-Sámi origin (cf. Kallio 2009: 35), which suggests a later date. Even though the notion of Baltic loanwords in an already disintegrating Proto-Sámi seems *a priori* unlikely, we must keep our minds open at this point:

- Sá. N *šearrat*, L *sjerrat* (< **šearętē̮*) 'clear (of the sky)' ~ Lt. *žėrė́ti* 'glow (e.g. of coals); shine, glitter' — There is no precise formal or semantic match, so it is essentially a root etymology. Junttila (2015b: 477) has suggested that the Sámi word is rather a loan from Finnic **heretä* (> E *ere*, dial. *here* 'bright').[285]
- Sá. S *sealma* 'threshold; pass, ridge', N *šielbmá* 'threshold (of a tent)' (< **šielmā*) ~ Lt. *šelmuõ* 'ridge (of a roof), eaves, gable' (= SCr. *sljème* 'mountain ridge; (dial.) roof ridge') — There is no exact semantic correspondence. Aikio (2012a: 107) has tentatively suggested a loan etymology from Finnic **hęlma* (> F *helma*, E *hõlm*) 'hem'.
- Sá. S *sjåavonje*, N *šūvon* (< **šuovuńę*) 'well-trained shepherd dog' ~ Lt. *šuõ* (OBL. *šun-*) 'dog' — The comparison is phonologically problematic. Sammallahti (2001: 400) erroneously derives Lt. *šuõ* from **śoų̯on(i)-*, and Kallio (2009: 35) reconstructs Baltic **šāvā* on the basis of a Žemaitian form

[285] Note that Holopainen/Junttila (2022: 103–105) and Junttila (in prep. s.v. **heräiťtäk*) have suggested that this Finnic stem is after all ultimately of Baltic origin.

> *šova* quoted in LEW (1023). The latter is evidently an untransposed Žem. *šọva* = *šuvà*, which is not a derivative, but a special development of the nominative singular (see Zinkevičius 1966: 256–257). All in all, we can reconstruct a Proto-East-Baltic **š(ų)ōn*, OBL. **šun-*, neither of which can explain the Sámi word.

None of the etymologies are convincing, and they can hardly serve as a basis for drawing the far-reaching conclusions which they would imply, namely that the Balts would have been in contact with the Sámi already after their migration into Fennoscandia. Of course, it is in principle possible that certain Baltic populations crossed the Gulf of Finland, where enclaves could have interacted with Sámi populations until a relatively recent date. However, this is merely a speculation unsupported by any other evidence.

Note in this connection that it has been suggested that the Sámi endonym PSá. **sāmē* (> N *sápmi*) is derived from a Baltic source, cf. Lt. *žẽmė* 'earth, land' (Koivulehto 1993b). Koivulehto analyses **sāmē* as a cognate to F *Häme*, the name of a historical region in Finland. The notion that a Baltic loanword would be used as a local toponym in Finland, which would later serve as the self-identification of the Sámi people, would probably imply that the Balts settled in Finland, and most probably before the Sámi arrived there. This is an extremely bold claim. Furthermore, even though it seems attractive to compare F *Häme* with the self-designation of the Sámi, this encounters an important issue: it is hardly possible to separate Sámi **sāmē* from Finnish *Suomi* 'Finland'. These two forms together point to a common preform **sämä* rather than **šämä* (Pystynen 2018: 83; Holopainen 2021: 207–208).

All in all, the idea that Balts should once have been present in the area of modern-day Finland is too bold a claim to base on a single toponym,

especially as it remains possible, or even probable, that the ethnonym **sāmē* was adopted from a paleo-Lakelandic contact language of the type we know must have been spoken in this area before the arrival of the Sámi (Aikio 2012: 80–88).[286]

1.4.2. Mordvin

The possibility that Baltic and Mordvin were in contact was already recognized in the 1880s in two articles by Wilhelm Tomaschek (1883: 704–705, 1889: 11–12). Of the 15 comparisons made in these works, many reappear in Thomsen's work on Finnic-Baltic loanwords. In the latter's opinion (1890: 154–155), these must mostly have passed through a dialect continuum from Finnic into Mordvin, although Thomsen does admit that a small number may have been borrowed directly. A similar conclusion was reached by Kalima (1936: 191–192). By contrast, as many as four of the seven loanwords accepted in the recent comprehensive study by van Pareren (2008) were classified as direct. The goal of this subchapter is to establish the degree of direct and indirect contacts between Baltic and Mordvin.

1.4.2.1. Rejectable comparisons

The most recent detailed study (Grünthal 2012) shows a manifold increase in the number of accepted loanwords, with a total of 36.[287]

[286] Reminiscent of **sāmē* is Lv. ⟨S̆áhms⟩ 'Finne; Oesulaner' cited by Ulmann (1872: 244), but Endzelīns (ME III: 803) attractively derives this form from Li. *sārmā* 'Saaremaa' (note Oesel = Saaremaa). Several parallels for the loss of /r/ after long vowels in Courland Latvian are provided in Endzelīns 1923: 159–160.

[287] Grünthal explicitly marks as uncertain the Baltic loan etymologies for PMd. **ärkə* 'lake' (on which see p. 153–155), **kodər* 'twining plant stem' (see

Unfortunately, a large proportion of the additional etymologies accepted and proposed by Grünthal involve hypothetical semantic shifts or anachronisms, and in my opinion should certainly be rejected:

- Md. EM *al* (< **al* / **alə*) ‘egg’ ~ Lv. *uõla* ‘egg’, dial. ‘pebble’ (Joki 1973: 294) — The sense ‘egg’ is an extension of ‘pebble’ in only part of the Latvian dialects, displacing older *pàuts* (EH II: 186 = Lt. dial. (S Aukšt.) *paũtas*, Pr. G *paute* ‘egg’), and is hardly to be reconstructed for Proto-Baltic (Lanszweert 1984: 38); compare Lt. *uolà* ‘whetstone, rock; cliff’. Furthermore, the substitution **(w)ō-* → **a* is phonologically unlikely (van Pareren 2008: 86).[288]
- Md. E dial. *čonda* (?< **šondə*) ‘bride price’[289] = F *hinta*, E *hind*, Li. *īnda* (< **hinta*) ‘price’ ~ Lt. *šimtas* ‘hundred’ (Uotila 1990) — The fact that *šimtas* (usu. as PL. *šimtaĩ*) can be used hyperbolically to mean ‘a lot’ (like English *hundreds*) cannot be considered a sufficient semantic bridge.
- Md. E *inže* (PL. *inšt*), M *inži* (< **inžə* : **inž-*) ‘guest’ ?= F obs. *inhiminen* ‘person’, Vp. *inehmoi* ‘lazy or sickly person’, Võ. *inemine*, Li. (Salaca) *imi* ‘person’ (?< **inehminen*) ~ Lt. *įžymùs* ‘notable, famous’ (Liukkonen 1999: 61–62) — The Lithuanian word is a productive deverbal adjective from

---

fn. 87), **mukərə* ‘rump, rear’ (~ Lv. *mugura* ‘back’; against which see van Pareren 2008: 109–111), and Md. EM *luv* ‘space between the fingers’ ~ Lt. *lomà* ‘hollow, valley’ (the Md. word rather belongs with F *lovi* ‘cleft, notch’, Pystynen 2020b). These will be ignored in the following discussion.

[288] The alternative view is that EM *al* ‘egg’ is a semantic extension of EM *al-* ‘area under or below’ (Rédei 1968: 160; Keresztes 1986: 33), which is itself of Uralic origin (< **ęla*, Aikio in prep. 52–53).

[289] If *čonda* (Velikij Vrag) is the most archaic form, the metathetic variant *čando* could perhaps be explained as the result of contamination with *čana* ‘price’ ← R *цѣна́* (cf. Вершинин I: 486).

į-žymė́ti ‘note, mark’, and cannot be reconstructed for Proto-Baltic. Furthermore, the semantic development is speculative.

– Md. EM *karks* ‘belt, waistband; rope used to bind a sheaf’ ~ Lt. *kárti* ‘hang (up)’ (Grünthal 2012: 315) — EM *-ks* is a deverbal suffix, but the connection with ‘hanging’ is not evident.
– Md. E *penge* (PL. *penkt’*), M *pengä* (< **peŋgə* : *peŋg-*) ‘(piece of) firewood’ ~ Lt. *spiñgis* ‘forest aisle’ (Grünthal 2012: 324) — The two words are semantically distant.
– Md. E *pusmo*, M *pusma* (< **pusmə*) ‘bunch, bundle’ ~ Lt. *bùžmas* (Grünthal 2012: 326) — Grünthal incorrectly glosses the Lithuanian word as ‘bunch’. The word is only known from lexical sources; cf. *bùźmas* ‘eine Falte, Krause’ (Kurschat 1883: 66).[290]
– Md. E *raško*, M *raška* (< **raškə*) ‘crotch, fork’ = F *rahko* ‘fork-shaped torch holder’ ~ Lt. dial. *raškà* ‘skeltu galu kartis obuoliams raškyti [a device with a forked end used to pick apples]’ (Skirsnemunė; Skardžius 1941: 41), *rãškės* ‘prietaisas obuoliams raškyti [a device used to pick apples]’ (Daugėliškis; LKŽ) (Liukkonen 1999: 114–115) — The basic sense of these rare Lithuanian words must be ‘picker’ (cf. *rẽkšti*, *raškýti* ‘to pick’), while the Finnic and Mordvin words would imply an original sense ‘fork, crotch’ (cf. Nilsson 2001: 185).

[290] In Ruhig (II: 53), we find *buʒmas* ‘Bauchbruch am Reße’; however, this is presumably a printing error for **bůʒmas*; Mielcke (I: 31) lists the same word under *boʒmas* ‘das Bauchreß, der Bauch vom Reße’ and Nesselmann (1851: 333) has *bůźmas* = *boźmas* ‘das Eingeweidenetz, Bauchnetz’ (the word was not familiar to Kurschat 1883: 54). According to Nesselmann (but no-one else?), *boźmas* also = *baźmas* ‘eine große Menge, eine Masse von Menschen, Thieren, Körnern’, which must be where Grünthal’s ‘bunch’ ultimately originates (but note that all of the example sentences in LKŽ s.v. *bãžmas* refer to people, unlike the Mordvin words).

- Md. E *rudaz*, M *ərdaz* (< **rudas*) ‘dirt, faeces’ (Grünthal 2012: 329) ~ Lt. *rùdas* ‘chestnut brown’, Lv. *ruds* ‘red-brown’ — From a semantic perspective, the Russian data comes far closer, cf. Ru. dial. (Smolensk) *pyðá* ‘dirt, stain’, *pyðóй* ‘dirty’.[291]
- Md. E *ťeŕďe-*, M *ťeŕďə-* (< **terďə-*) ‘call over, invite’ ~ Lt. *tìrdinti*, Lv. *tiȓdît* ‘badger with questions, torment’ (Grünthal 2012: 335) — The Baltic words are frequentative derivatives of Lt. *tìrti* ‘question, examine’, Lv. obs. *tirt* ‘Ausfragen’ (Lange 1773: 351). The semantics are unconvincing.

Wälchli (1997: 312–319) has suggested Baltic etymologies for a number of grammaticalized relational nouns. All of these are rejected by van Pareren, but accepted by Grünthal.[292] Again, the semantic developments stretch the imagination:

- Md. E *lango*, M *langa* (< **laŋgə*) relational noun ‘on’, ‘surface’ ~ Lt. *lankà* ‘water meadow; swamp, valley’, Lv. *lañka* ‘low-lying meadow; river bend’. The basic sense in Baltic appears to be ‘river bend’, cf. R *лукá*, Bg. *лъкà* ‘river bend; meadow in a river bend’ (~ Lt. *leñkti* ‘to bend’).
- Md. E *potmo*, M *potma* ‘insides, stomach, bosom’, E *potso*, M *potsa* INESS.SG. ‘inside’ (< **potmə* : **pot(m)-*) ~ Lt. *putmuõ* ‘swelling (as an ailment)’
- Md. E *turtov*, dial. *turtoŋ* (< **turtəŋ* LAT.) ‘for’ ~ Lt. *tuřtas*, ‘wealth, property’

---

[291] But admittedly, Smolensk is geographically far removed from Mordvinia. Curious is the Russian dialectal form *pýðoc* ‘swampy area where rusted water comes to the surface’, attested in the Komi Republic (СРНГ XXXV 235), cf. Komi *rodęg* ‘dirt, stain; rust in standing water’, Mari E *rüδaŋa-,* W *ərδäŋge-* ‘to rust’ (Лыткин/Гуляев 1970: 241–242). For Komi and Mari, a common preform **rentV-* could perhaps be reconstructed.

[292] According to P. Kallio (p.c. March 2023), Wälchli himself is now unenthusiastic about his older proposals.

In addition, a few etymologies must be rejected on formal grounds. Although Grünthal himself notes that sibilants in Erzya are not subject to palatal harmony (2012: 330; cf. Bartens 1999: 43), he still resorts to it in two cases where we find an unexpected sibilant reflex:

- Md. E *raśke* (< **raśkə*) 'relative, kin' ~ Lv. *rads* 'relative, lineage' — Grünthal posits a preform **radəs-kə*, contradicting the Erzya evidence for **ś*. Additionally, despite ME (II: 463), it seems quite possible that the Latvian word is loaned from Russian; cf. OR родъ 'lineage; birth, origin; relative (etc.)' (СДРЯ X: 408–415)
- Md. E dial. *simeń*, M dial. *śimәń* (< **simәń*) 'tribe, family' ~ Lt. *giminė̃* 'relative, tribe, family' — The Erzya form proves an initial **s*.

In the latter case, it is assumed that **ś* would substitute a palatalized allophone of **g* in Baltic. Junttila (2018: 78), who also provides the erroneous reconstruction **śimәń*, specifies this Baltic dialect as "Altlettgallisch". He suggests two parallels: E dial. *śive*, M dial. *śivä* 'salary, pay' ~ Lv. *dzîvuôt* 'live', dial. +ACC. 'work, be occupied with' and E *śiŕe*, occurring in collocation with *paro* 'good' in curses ~ Lt. *gìrti* 'praise', *gẽras* 'good'. The evidence of these two words alone, both of which require additional assumptions, seems insufficient to support a substitution **g* → **ś*.

Another supposed piece of evidence for an "Old Latgalian" source is the verb Md. E *ŕeďa-*, M dial. *ŕäďa-* (< **räďa-*) 'see, notice' ~ Lt. *regė́ti*, Lv. *redzêt* 'see, discern', REFL. 'seem, be evident' (Wälchli 1997: 319–320; Junttila 2018: 79–80). Wälchli's opinion is that **ď* may have directly substituted **g*, as Proto-Mordvin lacked a phoneme */g/. However, there is no reason to consider the loss of **g* particularly ancient (Grünthal 2012: 328), and the possible Baltic loanword **lija* ?< **läjkä* 'other' (see below) must have predated it. In the opinion of van

Pareren (2006: 49; cf. 2008: 120), we should expect *g → k in such a late loanword, which is indeed what we find in some borrowings from Tatar and Russian (Paasonen 1903: 17; Keresztes 1987: 67–68).

According to Junttila, Mordvin *d' could directly substitute a Baltic palatalized *[g'], and a realization [d'] for /g'/ is indeed attested in South Aukštaitian (Zinkevičius 1966: 140–141). However, the hypothesis that a dialect in which the velars were palatalized was spoken in the necessary time and place remains unproven. As Junttila (2018: 80) himself admits, no evidence for a Latvian-type palatalization has been identified in Baltic substratal hydronymy. The evidence of Mordvin *$räd'a$- alone is hardly enough to postulate such a feature for a hypothetical Baltic dialect.

1.4.2.2. Turkic or Baltic?

In a couple of cases, a Turkic origin appears just as likely or more probable than a Baltic origin:

'honeycomb'. Md. E *keŕaz* (also as PL. *keŕazt*), M *käŕaz* (< *$käŕas$*) 'honeycomb' ~ Lt. *korỹs*, Lv. *kāre* 'honeycomb' (Tomaschek 1883) — The Mordvin word cannot be separated from a wider group of Volgaic terms. On the one hand, we find Ma. E *karaš*, W *käräš* and Udm. *karas* (which is not regularly cognate with the Mari forms), and on the other Tatar *käräz*, Bashkir *käräδ* (< *$käräz$*), all in the same sense. Due to its final sibilant, Chuvash *karas* 'honeycomb' cannot be cognate with the Volga Kipchak forms. Räsänen (1920: 245) has derived the Turkic words from Uralic, and these from an Iranian *$kāras$. As such a word is unattested in Iranian, this can hardly be accepted (cf. Joki 1973: 226–227; Holopainen 2019: 127). Disregarding the language-specific phonotactic limitations, Mordvin *$käŕas$ is phonologically identical to

Tatar *käräz*, and indeed already Paasonen (1897: 37) suggested that Mordvin borrowed the word from Tatar.

If the Volgaic and Baltic words are indeed connected, one might speculate whether it was the Turkic words that were in fact loaned from Baltic. Indeed, this could potentially explain the unexpected front vocalism in Mordvin. In Turkic, **k* was allophonically rendered as *[q] (> Chuv. *x*) in back-vocalic environments, resulting in an association of foreign /k/ with front vocalism, and leading to cases such as Chuv. *kĕrpe*, dial. *kŏ̆rpe* 'grain', Tat. *körpä* 'bran' ← R *крупá* 'grain' (see p. 401). Thus, we might anticipate a front-vocalic substitution in the case of a direct loan from Baltic. On the other hand, Volga Kipchak **-z* would be difficult to explain starting from a Baltic NOM.SG. **kârjas*. Furthermore, Mari **käräš* cannot be understood as a Volga Kipchak loan, as **-z* should have been preserved in Mari, cf. e.g. Ma. E *teŋâz*, W *taŋâž* 'sea' (cf. Tat. *diŋgez*, Kyrgyz *deŋiz*).

As a result, the relationship between the various Volgaic forms is difficult to establish, and if there is any connection with Baltic at all, the exact route of borrowing cannot be recovered. However, it seems quite evident that Mordvin adopted this word specifically from Tatar. On the further relationship with Gr. κηρός, see p. 390–391.

'far'. Md. E *talaj* 'quite a while (ago)', cf. *talajs* ILL.SG. 'for long', *talajste* ELAT.SG. 'from a distance', M *talaj* 'quite a while, quite far' ~ Lt. *tolì*, *toliẽ*, obs. *tõl* 'far, distant'[293] (Grünthal 2012: 333) — Grünthal suggests this Baltic source as an alternative to the older Turkic etymology (MdWb 2258–2259), which compared Kazakh *talaj*, Kyrgyz *dalaj* 'a few, quite a lot; often'. Grünthal's main criticism is that SSA (I: 138) does not mention any "corresponding words in the Turkic languages of the

[293] The Lt. word generally refers to distance, but may also have a temporal reference (e.g. *tolì priẽš* 'long before').

Volga region". However, a glance in the Tatar and Chuvash dictionaries reveals that the word is indeed present there: Tat. *talaj* 'quite a lot' (ТРС II: 302), Chuv. poet. ⟨талай хирне⟩ 'to distant lands' (Скворцов 440, s.v. талай II; see also Федотов II: 167).

With regard to the semantics, we can note that the Turkic words can be used in certain case forms with a temporal and spatial reference; compare Tat. *talajga* 'for (too) long' (*-ga* DAT), Kaz. *talajdan beri* 'for a long time' (*-dan* ABL, *beri* 'to here'), *talaj žer*, 'far away' (*žer* 'space'). As the Turkic etymology is phonologically trivial and raises no serious semantic issues, it should be preferred over Grünthal's Baltic etymology.

'yard'. E *kardaz* 'yard, stable', M *kaldaz* 'stable, pen' (?< **kardas*)[294] ~ Lt. *gařdas* 'enclosure, stall', Lv. dial. (SW) *gārds* 'pigpen' (Tomaschek 1883) — The Mordvin word cannot be separated from Md. E *kardo*, M *karda* 'stable, pen', which Paasonen (MdWb 619), no doubt correctly, derives from Chuvash *karta* 'stockyard, stable; fence'. The word is also found in the Volga Kipchak languages, viz. Tatar *kirtä*, dial. *kärtä* 'pole; fence, enclosure', Bashkir *kärtä* 'pole, fence; stockyard', and the Chuvash word was also borrowed into Komi *karta* 'stable, barn'. These cannot be separated from a group of similar words in the Caucasus, cf. Oss. I *kært* 'yard, estate', D *kært(æ)* 'stockyard', Ingush *kart* 'fence' (Абаев 1958: 586–587).

The question is whether Mordvin **-as* can be seen as a suffix. While it is not a productive derivational element, such a suffix must be present in Md. E *ńeŕgaz*, M *ńäŕgaz* (< **näŕgas*) 'badger', which is etymologically related to Mari E *neryе*, W *neryə* (< **nirgə*) 'badger'.[295] There are

[294] For a discussion of the Moksha *-l-*, see van Pareren (2008: 89–90).

[295] The relationship between Volgaic **närkä* and Finnic **mäkrä* (> F *mäyrä*, E *mäger*, *määr*, Li. *mä'ggõrz*) 'badger' is unclear, but a relationship looks

numerous other Mordvin nouns ending in *-*as*, but very few can be reliably analysed (see Maticsák 2014). Nevertheless, as the word for 'badger' shows, the presence of final *-*as* is not sufficient to guarantee an Indo-European origin (*pace* Wälchli 1997: 307).

1.4.2.3. Acceptable comparisons

Despite the large number of rejected or doubtful comparisons, we are still left with a corpus of formally and semantically acceptable loan etymologies. A couple of these examples are also present in Finnic, and have therefore already been discussed elsewhere in this dissertation. Therefore, a simple list will suffice. Note that none of these examples can be considered certain evidence of direct contact between Baltic and Mordvin as the Baltic words themselves are of uncertain origin:

- Md. E *kšna*, M *šna* (< **(šə)šna*) 'worked leather; leather strap' ~ Lt. *šikšnà*, Lv. *siksna* 'untanned leather; leather strap or belt' (see p. 155–157)
- Md. E *ľepe*, M *ľepä* (< **lepə*) 'alder' ~ Lt. *líepa*, Lv. *liẽpa* 'lime tree' (see p. 157–158)
- Md. E dial. *ťožań*, M *ťožäń* 'thousand' (< ?**ťožən* : **ťožəm-*) 'thousand' ~ Lt. *tū́kstantis*, Lv. *tũkstuôtis* 'thousand' (see *1.3.5.4*)
- Md. E *malaso*, M *malasa* INESS.SG. (< **mala-*) 'near' ~ Lv. *mala* 'edge, shore, boundary' = F dial. *malo* 'edge, flank', Sá. L *muolos* 'shore lead' (see p. 206)

The remaining cases are unique to Mordvin, and must be discussed separately. In doing so, it is important to evaluate not only the

---

possible: the irregular correspondence perhaps suggests a shared substrate word.

plausibility of the comparison, but also the etymological background of the suggested Baltic source. Only those with a clear Indo-European etymology can provide objective evidence in favour of a loanword from Baltic into Mordvin. Those of unclear ultimate origin are presented here in *italics*.

'*bast*'. Md. E *l'enge* (PL. *l'engt*), M *l'engä* (< **leŋgə*, OBL. *leŋg-*) 'bast' ~ Lt. dial. *lùnkas*, Lv. *lûks* 'bast' — The Mordvin form could reflect an earlier **lüŋkV* (cf. Aikio 2009: 116). Surprisingly, this semantically and formally attractive etymology is rejected by both van Pareren and Grünthal.[296] The Baltic word is of unclear origin (see p. 291).

'bridle'. Md. E *panct*, *panst*, M *pandəz* (< **pandəs*) 'bridle' ~ Lt. *pántis* 'hobble, fetter' (= Pr. E *panto*; OCS пѧта PL. 'fetters') (Tomaschek 1889: 11) — While the semantic match is not exact, both bridles and hobbles are tools used to restrict a horse's movement. The Lithuanian word is probably derived from the verbal root seen in Lt. *pìnti*, Arm. *henum* 'weave' (LIV 578–579).

'*knife*'. Md. E *pejel'*, M *pejəl'* (< **pejəl'*) 'knife' ~ Lt. *peĩlis* 'knife'; Pr. E *kalo-peilis* 'cleaver' (Tomaschek 1883) — The scepticism of van Pareren (2008: 113–114) is hardly justified, as the etymology appears formally and semantically straightforward. The Baltic word lacks an

[296] Both suggest a native origin. Van Pareren (2008: 103–104) assumes a derivational relationship with Md. E *l'ejks* 'young alder (whose bark has been stripped)' and *l'evš* 'bast', although a detailed morphological analysis is wanting. Grünthal (2012: 321) follows Mägiste (1962) in equating **l'engə* with F *niini*, Komi *ńin* 'bast' (< **nijnə*); this, however, leaves the stem-final velar unexplained. As an alternative, Grünthal adduces E *luvoď e-*, M *luŋgâď ə-* (< **luŋəď ə-*) 'flake off; fade' as a comparandum for the same Baltic word, but the unexpected substitution **nk* → **ŋ* and less obvious semantics makes this comparison less attractive.

etymology; the older connections with Lt. *pielà*, R *nuлá* ‘saw’ are abandoned in recent sources (ALEW 862; Smoczyński 2018: 954).

‘*millet*’. Md. E *suro*, M *sura* (< **surə*) ‘millet’ ~ Lt. *sóros*, Lv. dial. *sûra*[2], (17th c.) *sāre* ‘millet’ (Tomaschek 1883) — The Baltic word lacks a clear etymology. For a detailed discussion of this comparison, see p. 409–411.

‘other’. Md. E *l’ija*, M *l’ijä* (< **lija*)[297] ‘other’ ~ Lt. *liẽkas*, Lv. *lìeks* ‘surplus’ (Paasonen 1909a: 89) — In Uralic terms, the Mordvin words could reflect **lekä* or possibly **lejkä*, which would both be reasonable substitutions of an East Baltic **lę̃ka-*. An earlier form **laika-* is more difficult, but I would not rule out a pre-Mordvin reconstruction **läjkä* (compare Md. E *śiśem*, M *śiśəm* ‘seven’ < **ćäjćəmä*; Aikio in prep. 119). The semantic difference seems to be bridged by the derived verb E *l’ijado-*, M dial. *l’ijadə-* ‘stay behind, remain’, which precisely corresponds in meaning to the Baltic verb seen in Lt. *lìkti* (3PRES. *liẽka*) ‘remain, be left over’ (van Pareren 2006: 36, but sceptically 2008: 105).

‘soot’. Md. EM *sod* (< **sod*) ‘soot’ ~ Lt. *súodžiai*, dial. *súodys* PL., Lv. obs. *suods* (EH II: 610), dial. *suôdri*[2] ‘soot’ (Paasonen 1909a: 127) — A Baltic origin is rejected by van Pareren (2008: 122) and Grünthal (2012: 308) due to the existence of a native etymology. However, neither Mari E *šüć*, W *sə̑ts* ‘soot, coal’ nor Komi *sa*, Udm. *su* ‘soot’ (UEW 769) represent a phonological match,[298] and they cannot be accepted as

---

[297] The final *-ä* in Moksha is due to a secondary fronting of final *-a* after a palatal consonant (Bartens 1999: 63), cf. M *pŕä* (*pŕa-* in inflected forms) ‘head’ < PMd. **piŕa*.

[298] Initial *š-* in the Malmyž dialect points to PMa. **š-* (Wichmann 1906: 21; TschWb 740), suggesting the Mari word is instead cognate with Md. *śed’* ‘coal’ (< PU **ćüďə* ‘coal’; for **d/*ď* > PMa. **ć* compare *tić* < **täwdə*; Metsäranta 2020: 43; however, Aikio in prep. 147 adduces a different Mari cognate in this dataset: E *šüj*, W *šü* ‘charcoal’; perhaps **šŭ̑ć* ~ **šü(j)* is the result a

cognates. The substitution **ō* → Md. *o* can be considered reasonable so long as the loanword postdated the pre-Mordvin change **o* > **u*. The Baltic word is cognate with R *са́жа*, Sln. *sáje* 'soot' and further OE *sōt* 'soot'.[299]

'*thunder*'. Md. E *puŕgińe*, dial. *piŕgińe*, M dial. *puŕgəńä* (< **puŕgəńə*) 'thunder' ~ Lt. *perkū́nas*, Lv. *pę̂rkuôns* 'thunder', also a theonym; Pr. E *percunis* 'thunder' (Tomaschek 1883) — The Mordvin vocalism must result from a metathesis, which could be motivated by the lack of rounded vowels in non-initial syllables in Proto-Mordvin (van Pareren 2008: 119). The palatalized suffix is probably to be attributed to assimilation to the Mordvin diminutive suffix **-əńə*.[300] Despite REW (II: 345–346), the Baltic words cannot be separated from OR Пероунъ 'thunder god', Pl. *piorun* 'lightning', which show an irregular correspondence with Baltic, most probably pointing to a foreign origin.

? **'*forest*'**. Md. EM *viŕ* (< **viŕ*) 'forest' ~ Lv. dial. (Vidzeme) *vēris* 'spruce forest', dial. 'riverside meadow' (Grünthal 2012: 336) — While the comparison seems attractive at first sight, the vowel substitution

---

paradigmatic split?). The correlation Komi *a* ~ Udmurt *u* does not usually occur in inherited words except where it is a reflection of **-eCə#*, cf. Живлов 2013. Metsäranta (2020: 140–141) has attempted to substantiate a preform **setə* by comparing the verb Komi *sǫt-*, Udm. *sutį-* 'burn, set on fire', allegedly < **set-tä-* (differently on this verb see Aikio 2021: 169–173).

[299] Since I do not think that a lengthened grade yielded acute, the Balto-Slavic form (**sod-i-*) cannot be directly equated with Germanic **sōta-* (< **sōdo-*), but both words probably derive from the root **sed-* 'to sit' via the sense 'sediment'. OIr. *suide** 'soot', is to be derived from **sūdi̯ā-* in view of Modern Irish *súiche*, Catalan *sutge* 'soot' and cannot be directly related (see Walde/Pokorny II: 485; Zair 2012: 125, with lit.).

[300] Van Pareren posits a Baltic source **perku-*, citing Narevian *pjarkuſ* (Zinkevičius 1985: 77). The controversies around the Narev glossary aside, this cannot be considered evidence of a shorter form; the loss of **n* before final **-s* is paralleled by *garſ* 'stork' ~ Lt. *garnỹs*.

seems suspect if we assume Latvian *ē* continues Proto-Baltic */æ:/. Some cases of Md. **i* deriving from **ä* do occur, but this appears to be conditioned by a preceding palatal (Aikio in prep. 114). No Indo-European etymology is suggested by Endzelīns (ME IV: 562),[301] while Karulis' (II: 508) comparison with Gr. εὐρύς 'broad, wide' is neither semantically nor phonologically convincing.

? **'left'**. Md. E *kerš*, M *kerži* (< **kerš*) 'left' ~ Lv. *krèiss* 'left' (Viitso 1990: 141; van Pareren 2008: 93; Grünthal 2012: 316) — Although semantically attractive, this etymology requires some assumptions on the phonological side. First, it must be assumed that the inadmissibility of initial consonant clusters resulted in a metathesis. While imaginable, reliable parallels are few (cf. Md. EM *turba*, dial. *truba* 'horn' ← R *мруба́*). The second assumption is that Lv. *-s-* reflects Baltic *-*š*-. True, the traditional equation with Lt. *kréisva* 'flaw' (LEW 203) would imply Baltic **s*, but due to the mismatch in accentuation, it is uncertain that the Lithuanian word belongs here. The Latvian word is apparently related (with metatony?; Derksen 1996: 190, 196–197) to *kreĩlis*, *ķeĩris* 'left-hander' (< **kreirīs*) and further Lt. *kreĩvas*, R *кривóй* 'crooked' (LEW 203; ALEW 523; Smoczyński 2018: 598).

? **'salt'**. Md. EM *sal* 'salt' (< **sal*)[302] ~ Lv. *sā̀ls* 'salt' (= Gr. ἅλς) — The Mordvin form cannot be directly equated with Finnic **sōla* (despite Напольских 2015: 163–164), as the Finnic stem type **ō–a* is of recent and secondary origin (cf. Plöger 1982), but it may be analysed as an

[301] Endzelīns suggests a loan from Estonian *veer* 'edge', but only for the sense 'riverside meadow'. Incidentally, this Estonian word has been considered cognate with Mordvin **viŕ* (UEW 820–821). However, this is not phonologically acceptable; Aikio (2012b: 234) reconstructs the former as **wärə* and equates instead Md. E *veŕe*, M *väŕä* 'above, over'.

[302] The reconstruction **sal* may be preferred over **salə* in view of *Sal · Zout* in Witsen 1785; cf. Pystynen 2020b.

independent loanword from Baltic with the vocalic substitution **ā* → **a* (Holopainen 2019: 215). On the other hand, the analysis as a direct Baltic loan is rendered somewhat uncertain by the Permic evidence (Komi *sov*, Upper Sysola *sǫl*, Udm. *sʲilal* 'salt' < Proto-Permic **sǫ̈l*), which must also be related, but lacks a clear source. Apparently, we are dealing with an ancient *Wanderwort* of ultimately IE origin. For Mordvin, a proximate Baltic source is possible, but other possibilities are imaginable.

† **'duck'**. Md. E dial. *šenže*, *šenš* 'duck' ~ Lt. *žąsìs*, Lv. *zùoss* 'goose' — Due to its limited attestation, a Proto-Mordvin reconstruction is difficult. The only reliable attestations provided by Paasonen (MdWb 2227) derive from the Kadom and Kaljaevo dialects, which happen to be the same dialects which show a raising **a* > *e* in *šenžej*, *šenžij* 'spider' (< **šanžəŋ*; cf. Paasonen 1903: 81). Therefore, as well as **šenž-* or **šänž-*, a reconstruction **šanž-* can also be considered. The latter would allow for a direct equation with Komi dial. (Udora) *čęž*, Udm. *čęž* 'duck', as well as Khanty **čāč* (> Vakh-Vasjugan *čač*, Kazym *šɔš*) and Mansi **šī̮š* (> West *šę̄š*, South *sās*) 'mallard' (?< **čęnčə*; cf. UEW 58).[303]

By contrast, Sammallahti (2001: 398) has reconstructed **šänšä* and treated the Mordvin word as a regular cognate of PF **hanhi* 'goose', which is a Baltic loanword. This preform was later substantiated by the

[303] The development **ę*(–*ə*) > Komi/Udmurt *ę* is possibly regular before a resonant (Aikio 2012b: 241). For the Permic and Ob-Ugric words, Aikio (2015a: 57) reconstructs **čęčə*. The loss of the nasal would be regular in Permic. In Khanty, the development **-nč#* > **-č* is not regular, but paralleled by **poč* (> Vakh-Vasjugan *poč*, Surgut *pŏč*) 'back (of the head)' (< PU **ponča*). In Mansi, the simplification **-nš* > **-š* is regular in syllable coda (Pystynen 2020c: 256–257). While the nasal is usually preserved in inflected forms, the form without the nasal has occasionally been generalized, as in Mansi (North) *šį̮š*, (South) *sis* 'back' (cf. e.g. Vakh-Vasjugan Khanty *čŏ̮ṇč*).

sound law PU **ä(–ä)* > PF **a(–e)* (cf. Heikkilä 2014: 86). If this is correct, this loanword would have to predate the other Baltic loanwords in Finnic (cf. the preservation of **ä–ä* in **mäntä* 'stirring stick', **härkä* 'ox'). The most awkward aspect of this is that pre-Finnic **šanši* is closer to Baltic **žansi-* than the suggested West Uralic **šänšä*. If we instead assume that Erzya *šenže* was an independent loan from Baltic (Grünthal 2012: 331), then we would have to assume that an identical assimilation **š–s* > **š–š* took place independently in Finnic and Mordvin.[304] However, as with Finnic, this assimilation is potentially regular, as there do not seem to be any Proto-Mordvin words with **š–s*. Nevertheless, the native etymology, on balance, seems more convincing.

† **'daughter'**. Md. E *ťejťeŕ* 'girl, daughter' beside E *śťiŕ*, dial. *stiŕ*, M *śťiŕ* 'girl, daughter' ~ Lt. *duktẽ* 'daughter' — These two Mordvin forms appear to be found in almost complementary distribution across the dialects (in MdWb 2384, only the Gorodišče dialect attests both variants). It is generally assumed that the latter represents an irregular reduction of the former, although exactly how this works is unclear to me: particularly problematic is the Erzya form with unpalatalized *s-*. Since **-k-* became **-v-* or **-j-* in Mordvin depending on vowel harmony, and there are examples of **-kt-* > **-vt-* in back-vocalic words (e.g. Md. E *kavto* 'two', *avto-* 'set a trap' < PU **kakta*, **ękta*), one might anticipate

[304] Incidentally, a Permic word has been taken as a loanword from the Indo-European word for 'goose': Komi *ʒo̯ʒęg*, (Jaźva) *ʒu̇·ʒok*, Udm. *ʒaźeg* 'goose' (cf. Holopainen 2019: 377–378, where either an Indo-Iranian or Baltic etymology are considered). The Permic forms show an irregular vowel correspondence: Komi **o̯* ~ Udm. **a* is extremely rare (we expect Udmurt **u*). Moreover, the Permic forms, if taken from IE **źans-*, would presuppose yet another assimilation. Rejectable is Koivulehto's (2001: 244) derivation of PSá. **ćuońēk* (> N *čuonjá*) 'goose' from a hypothetical PIE "**ǵʰan-əd-*".

**-kt-* > **-jt-* in front vocalic words, and reconstruct **tüktär* for Mordvin. Unfortunately, this development is contradicted by Md. E *ńevťa-*, M *ńefťə-* 'pluck, tear' < PU **ńüktä* (cf. F *nyhtää* 'pluck', Ma. E *ńəkta-* 'skin'). Therefore, the relationship of this word to the Baltic data remains uncertain.

~ ~ ~

In the above, we have identified 7 plausible and 3 possible loanwords from Baltic into Mordvin. Three of the plausible examples also have a clear Indo-European background ('other', 'bridle' and 'soot'), which would appear to demonstrate direct, independent contacts between Mordvin and Baltic, and it is possible that some of the other words were also adopted from Baltic directly. Contrary to the conclusion of previous works on the subject, none of the Mordvin words could plausibly have been borrowed through Finnic. While Baltic **lę̄ka-* 'surplus' has been borrowed into both Finnic and Mordvin, the two forms cannot be traced back to a common proto-form, with Mordvin pointing to front vocalism, and Finnic to back vocalism. The words for 'belt', 'alder' and 'thunder', noted at the start of this chapter, encounter similar issues (see the discussions in *1.3.4* and *1.3.5.4*).

With regard to semantics, words in the sense 'knife' or 'bridle' might well be understood as technological loans and be regarded as characteristic of an adstrate loan context, and this analysis seems most convincing given the small number of loanwords overall. In such a context, however, the words 'soot', and in particular 'other', are rather unsettling. Specifically, according to the *World Loanword Database*, 'other' ranks among the 300 least likely words to be borrowed. Of course, the loanword proposal presupposes that the word was borrowed in the sense 'surplus', with only a secondary shift to 'other'. In this context, we could compare the Latvian *suitāk* 'too much', Pr. III *zuit* 'genug', which I have suggested may have been borrowed from

Slavic as trade jargon (see *1.1.1.8*). On the other hand, given that Md. **lija* could reflect a number of possible proto-forms, one may ask whether the loan etymology is even correct.

1.4.3. Mari

By contrast to Mordvin, whose contacts with Baltic have never been doubted, the idea that there are Baltic loanwords in Mari has not been universally accepted. The staunchest opponent to the idea was Mägiste, who in a 1959 article provided alternative analyses for all suggested Baltic loanwords. However, despite his efforts,[305] there still remain a small number of potential loanwords which have not been explained away by previous studies. Among those shared by other West Uralic branches, we can note the words for 'belt' and 'thousand' which have already been discussed (see *1.3.4* and *1.3.5.4*). The following are exclusive to Mari:

? **'house'**. Ma. EW *pört* 'house, cottage' ← Lt. *pirtìs* 'bath-house' (Thomsen 1890: 208; Kalima 1936: 148) — Several scholars have assumed a Russian origin, instead (Nieminen 1953: 213; Mägiste 1959: 170; Baxpoc 1963: 159; Bereczki 1994: 117; TschWb 541). However, this remains problematic, as there is no evidence that the Russian word was ever in use in the Volga region (see p. 68–69).[306] Starting from a

[305] For instance, he pointed out that Ma. *šukertə* 'for a long time' must be segmented *šuk-ertə* (cf. Ma. W *šuk* 'much', E *ertak* 'constantly') and does not contain a cognate of F *kerta*, Md. E *kirda* 'time, -fold' (cf. also Grünthal 2012: 317).

[306] Nieminen and Baxpoc claim that the Mari word would prove that the Russian word used to be more widespread, but the sheer geographical distance from the actual Russian attestations makes this argument quite circular. Moreover, the Mari word is a general term for 'house', a sense

Baltic source, the vocalism is not quite clear: note that the Mari vowel would be the usual reflex of PU **ę* (cf. Aikio 2014a: 131–135).

A. Дыбо (2008: 231–232) has suggested an alternative etymology. Contrary to the *communis opinio* (Räsänen 1920: 259; Федотов I: 462), which takes Chuv. *pürt*, dial. *pört* 'house, cottage' as a Mari loanword, she assumes a borrowing in the opposite direction, and compares the Chuvash word with Old Turkic *barq* 'shrine, temple'[307] (usu. in the collocation *ev barq* 'house and home'; Clauson 1972: 359–360) and Yakut *bïrt* (Пекарский 625) 'wellbeing, wealth'.[308] As for the vocalism, the correlation Chuv. *ü* ~ common Turkic *a* is found in other words after **b*-, cf. Turk. *parmak* ~ Chuv. *pürńe*, dial. *porńa* 'finger', Old Turkic *bāš* (cf. Clauson 1972: 376) ~ Chuv. *püśek* 'wound'. Мудрак (1993: 113) has plausibly analysed this as a reflex of the diphthong **ia*, which elsewhere has a palatalizing effect in Bulghar.[309] The main issue with

---

unattested in Russian. Note that R dial. (Vetluga) *nepm* 'cottage' (СРНГ 26: 294; Мызников 2019: 599) is a loan from Mari.

[307] For the translation, see Hao (2019), who points to a Chinese parallel text which would apparently prove the meaning 'shrine' for Old Turkic. As Hao points out, early texts show that a *barq* is something which can be built, so Clauson's translation 'moveable property' (1972: 359) must be false, but Дыбо's own gloss 'здание, постройка' also appears too general.

[308] The same Turkic comparison was also briefly mentioned in a slightly earlier contribution by Мудрак (2007). Yakut *-rt* is regular from **-rk* (СИГТЯ V: 662); compare Yakut *kïrt-* 'shear, trim' < **kïrk-* (ЭСТЯ VI: 238). Yakut *ï* (< **a*) is a much-discussed issue that I will not enter into here, but I will note some more occurrences before **rt*: Yakut dial. *ïr̄t-* 'load (onto an animal)' (Пекарский 3822) (< **ārt-*, ЭСТЯ I: 180–181), *kïr̄t* 'hawk' (< **kārt-*, ЭСТЯ V: 317–319); note also the derivative *sïrdā-* 'grow light' beside arch. *sarā-* 'to dawn'.

[309] As a couple of typological parallels for palatalization of a labial being expressed on the vowel, cf. Livonian *käpā* 'hoof' (< **kapja*) as against *paḑā* 'pillow' (< **patja*) (Kallio 2016: 45) and Tocharian B *mit* 'honey' (< **mətə* < **medu-*) as against *śak* 'ten' (< **ćəkə* < **deḱm*). This is apparently a result of

the etymology is the final *-t* in Chuvash, which is not regular, but would have to be explained as due to the influence of the synonym *śurt*, dial. *śort* 'house, building'; cf. the compound *pürt-śurt* 'household' (А. Савельев p.c. September 2021). If this etymology is accepted, the similarity with the Baltic forms must be considered coincidental.

The picture is further complicated by a similar word in Sámi (N *barta* 'hut, cabin', Sk. *põrtt* 'house, cottage, room'), which appears to show regular sound correspondences and is attested in all Sámi languages except South Sámi (Lehtiranta 2001: 96–97, who reconstructs **pęrttę*). The word is also found in Finnish and in the north-eastern dialects of Karelian in the form *pirtti* 'cabin, cottage'. The sense 'bath-house' is limited to some western Finnish dialects (cf. Вахрос 1963: 159) and is also found in Ume Sámi and in the Swedish loanword *pörte* (< obs. *pyrte* ← Finnish).

In the more eastern North Finnic languages, we find an irregular *e*-vowel, cf. K *pertti*, Vp. *perť* 'house, cottage'. This form was at first written off as secondary (Kalima 1936: 70), but later explained as due to Russian influence (Nieminen 1953: 216–217; SSA II: 350). However, this explanation is quite uncertain, since the underived word is very rare in Russian, and is only recorded in the area of Novgorod and Pskov, which is too far south to have been in recent contact with Veps and Karelian. Furthermore, the usual Russian sense 'bath-house' is apparently not recorded for the form **pertti* in Finnic.

It is universally acknowledged that Sámi **pęrttę* is borrowed from Finnic (Thomsen 1890: 208; SKES 576; SSA III: 350; Aikio 2006b: 29). But since the Finnic forms are so narrowly distributed and do not even reflect a common proto-form, one might even suggest that they were

---

the fact that palatalized labials are generally disfavoured cross-linguistically (Ohala 1978).

loaned from Sámi. The substitution Sámi **ę* → Finnic **i* is a well-attested form of 'etymological nativization' (Aikio 2009: 15–16). The substitution**ę* → **e* is less frequent, but also attested; cf. F *kelo* 'snag; dead tree' ← Sámi **čęlę* (> Sá. S *tjalle* 'tree stump'; Aikio 2009: 77).

In that case, we would have to distinguish three groups: (a) Baltic **pirt(i)-* 'bath-house', of native origin → ONovg. *пьрть (see p. 68–69); (b) Sámi **pęrttę* 'house, cottage' ?→ F *pirtti*, K *pertti* and (c) Mari *pört* 'house, cottage' ?← Turkic. On the other hand, it is quite difficult to imagine that all of these words are unrelated, especially since the Sámi word is left without an etymology. One might be inclined to assume that Sámi borrowed the word from Baltic, thus limiting us to two families, but we are still left with the uncomfortable fact that the Sámi word does not usually mean 'bath-house'. It feels unlikely that a convincing solution to this puzzle can be proposed at this time.

? **'lynx'**. Ma. E *šurmaŋše*, (Upša) *šŭrmŏ*, W *sə̂rmə̂*[310] (< **šŭrmə*) 'lynx' ← Lt. obs. *šermuõ* (modern *šermuonẽlis*), Lv. *sę̂rmulis* 'stoat' (Топоров/Трубачев 1962: 248; Bednarczuk 1976: 46; Breidaks 1983: 47) — The Mari word is usually viewed as a Uralic inheritance (Collinder 1955: 8; UEW 490–491; Bereczki 2013: 258–259). However, the suggested cognates are mostly to be rejected.[311] If we reconstruct **ćurmə* for Mari, we might compare Khanty **ćōrəm* (Irtysh *ťurəm*, Nizjamer *śurəm*) 'weasel, marten, stoat' (for **u(–ə)* > **ō*, cf. Aikio in prep. 141), although the Khanty affricate remains irregular. On the other hand, Komi dial. *śer*, Udm. *śor* 'marten' can be combined with the Mari word by

[310] On Western Mari *s-*, see Wichmann 1906: 23–25.

[311] Sá. Sk. *čõrmm*, K *čirrm* 'evil spirit; wolf' (which seem to be irregular even among themselves), on the one hand, and Forest Enets *same*, Tundra Nenets *sarḿikᵊ* 'wolf' (< **sårmå*, Janhunen 1977: 136), on the other, do not match each other, or any of the other forms, in terms of vocalism.

reconstructing PU **ćirma* (cf. Ma. E *užar*, W *ôžar*, Komi *vež*, Udm. *vož* < **wiša* 'green', UEW 823; Aikio 2014a: 156).[312]

The reconstruction **ćirma* does indeed bring us close to the Baltic forms. We may get even closer if we compare the apparent "zero-grade" formation *širmuonėlis* (Baranauskas, Ivanauskas), although since this variant is late and rare, it more likely represents a secondary development (e.g. contamination with *šìrmas* 'grey, dapple-grey'). If the loan etymology is accepted, the directionality would have to be from Baltic to Mari: Lt. *šermuõ* has an almost perfect cognate in OHG *harmo* 'stoat'. However, the potential Khanty or Permic comparanda mean that this can only be seen as one possibility among several.[313]

? **'mink'**. Ma. E *šaške*, W *šäškə* (< **šäškə*) 'mink' ~ Lt. *šẽškas* 'polecat' — As discussed on p. 152, the Mari word cannot be considered cognate with Finnic **hähkä* 'mink'; therefore, one might assume an independent loan from Baltic. Chuvash *šaškĕ* has been taken from Mari (Wichmann 1911: 25; Räsänen 1920: 264), but E. Itkonen (1953: 204; UEW 498) has suggested the opposite direction of borrowing in view of the existence of comparanda in Volga Kipchak, cf. Tatar *čäške*, Bashkir dial. *šäške* 'mink'.[314]

---

[312] Whether **-rm-* > **-r-* in Permic is regular is uncertain. A parallel could be Komi *jir* (< **jįr*, cf. Jaźva *jər*) ~ Sá. N *jorbmi* 'deep spot in water' (< **jurma*, UEW 105). However, this etymology is (implicitly) rejected by Aikio (2002: 47). М. Живлов (p.c. October 2021) has suggested an alternative, and equally acceptable, etymology for the Komi word, comparing Khanty **jə̄r* (> Nizjamer *jur*, Kazym *jǫr*) 'river bed', also dial. 'deep spot in water' (OstWb 400), which would presuppose a Uralic **jurə*.

[313] There is also a difference in semantics. Admittedly, Ruhig (I: 148) cites a meaning 'eine wilde Katze' for Lithuanian, but the reliability of this gloss is questionable.

[314] The "Kyrgyz" (more properly Kazakh) *šeške* cited in these works stems from Ильминский 1860–1861. Since Ильминский gathered his Kazakh

Tatar *č-* is unexpected based on the Baltic original. One might assume it arose by dissimilation as in Tatar *šešä* beside dial. *čiša* ‘bottle; glass’ ← NP *šīša* (Ахметьянов 2015 II: 442), dial. *šišta* ~ *čišta* ‘pole for climbing competitions’ ← R *шест*, GEN.SG. *шестá* ‘pole’ (idem: 488). On the other hand, these parallels are inexact, as the variants with *č-* are in each case purely dialectal, while *čäške* belongs to the standard language. Moreover, there are instances of an assimilation **č–š* > **š–š* in Bashkir, including in the homonym *šäške* ‘cup’ (← R *чáшка*; cf. Ишкильдина 2018: 35). Further support for an initial affricate could be provided by Komi dial. (Udora) *ćuš* ‘mink’, which could reflect an earlier **ćaškV-*.[315] In addition, there is a clear resemblance with the narrowly distributed Sámi lexeme Sá. S *tjetskie*, Ume *tjaskie* ‘stoat’ (< **ćęckē*) (cf. Wichmann 1911: 25; Лыткин/Гуляев 1970: 314). In Uralic terms, the Sámi word could reflect **ći/üčkä* (or **ćička*). While the vocalism clearly rules out that the Sámi, Permic and Mari words are cognates, some kind of relationship is conceivable in the context of a shared *Wanderwort* or substrate word (cf. Junttila 2015a: 31).

? **‘stem’**. E *wurδo*, Volga *wŭrδo* (?< **wŭrdə*) ‘stem, handle’ ~ Lv. *vãrde*; Lt. *vìrdis* ‘cross beam for hanging or drying’ — The Baltic words have been compared since Būga (1908: 139) and Ojansuu (1921: 63) with F *varsi*, E *vars*, Li. *vaŗž* (< **varci*, OBL. *vartę-*) ‘stem, handle’. At the same time, the Finnic words are almost always equated with the cited Mari forms (Thomsen 1890: 237; E. Itkonen 1953: 159; SKES 1660),

---

materials in Orenburg and Bashkiria (cf. I: 109), we are probably dealing with a localized Bashkir loanword.

[315] The difference between Komi *ć* /tɕ/ and Tatar *č* /ɕ ~ tɕ/ is purely notational. Note that here, Proto-Komi **ć* should be reconstructed. While a regular development **č–Š* > **ć–Š* has affected most Komi dialects, Udora has generally preserved *č-* in these words (Сорвачева/Безносикова 1990: 18).

implying that the Baltic word has been adopted into Mari as well (thus explicitly Koivulehto 1979b: 142).

There are several issues with this theory. First of all, the vocalic correlation between Finnic and Mari is not regular, so direct cognancy between these words is probably to be rejected.[316] Secondly, Finnic **varci* 'stem, handle' could alternatively be cognate with Sá. N *veardi* 'mouthpiece of a pipe, handle of a rake', I *verdi* 'shaft' < **wärtä* (cf. E. Itkonen 1977: 6). The suggested Baltic source is semantically rather remote; this same Baltic word could rather be seen as the source of F *orsi*, E *õrs*, Li. *vȯ́ŗž* (< **orci*, OBL. *ortę-*) 'beam; perch', which could regularly derive from an earlier **wortə* (Nieminen 1963: 238–240; Ritter 1993: 105–106).

As the development **wo-* > **wŭ-* in Mari is regular, cf. Ma. E *wurɣem*, W *wârɣem* 'clothes' (= Komi *vur-*, Hungarian *varr* 'to sew' < **worka-*; Sammallahti 1988: 551), Ma. **wŭrdə* 'stem, handle' — provided the reconstruction is correct — could be cognate with Finnic **orci*. Thus, we are faced with the awkward situation that the Mari word corresponds phonologically to Finnic **orci* 'beam', but semantically to Finnic **varci* 'stem, handle'. Since it is unlikely that both of these derive from the same Baltic word, the Mari word cannot be considered a certain Baltic loanword.

---

[316] Even within Mari, some of the dialects have reflexes of **u* rather than **ŭ*, e.g. Ma. W *wurδâ*. A similar situation is found in the near synonym Ma. E *wurɣo*, Volga *pŭrɣo*, W *wurɣâ* 'shaft'. In both words, we also find an irregular alternation of *p-* beside *w-*. This must be the result of decompounding (both words are frequent as second members of compounds; see the lists in TschWb 60–61). Either **p-* or **w-* could be primary: in the latter case, *w-* would be generalized from intervocalic position, and in the former, dial. *p-* would result from hypercorrection.

† **'rake'**. Ma. E *šor-wondo* (cf. *wondo* ~ *pondo* 'stem, stick') 'rake' ~ Lv. *zars* 'branch; prong' (Aikio 2009: 149) — Above, I have accepted the Baltic loan origin of F *haara* 'branch, fork' (see p. 113) and Sá. N *suorri*, Sk. *sue´rr* 'branch, fork' (p. 208). At the same time, reference works have further equated the Finnic word with Mari *šor-wondo* (SKES 57; UEW 783). Semantically, there is no issue; the sense 'rake' is even attested in the Finnish derivative *harava*, and Aikio has previously accepted both the Baltic loan etymology and the Mari cognate. However, Bereczki (2013: 247) has pointed out forms with *s-* from the Malmyž dialect which would suggest a Proto-Mari **s-* and rule out the etymology. Aikio (2015a: 56) agrees with Bereczki and instead proposes a comparison to Sá. N *suorgi* 'fork, branch'. Therefore, this Mari word cannot be considered a Baltic loan.

As a result there is not a single Baltic etymology in Mari which does not have alternative explanations. In the case of the word for 'mink', a connection seems probable, but the nature of the relationship is far from clear. Here one might be inclined to side with Mägiste (1959: 176): "Wenn die Anzahl der evtl. balt. Lehnwörter im Tscher[emissischen] nur auf einen einzigen Fall begrenzt ist, dürfte kein Anlaß vorliegen, von balt. Lehnwörtern im Tscher. zu sprechen." There certainly does not at this stage appear to be any solid evidence that would prove the existence of Baltic loanwords in Mari.

1.4.4. Permic

The situation with regard to Permic is even less promising that with Mari. Here I will leave aside Koivulehto's proposal that certain 'pre-Baltic' loans (where the source forms are back-projections of Baltic data into Proto-Indo-European) may have been adopted into a common

'Finno-Permic' language. These loanwords, if reliable, would simply be too early to describe them as 'Baltic' *per se*. On the other hand, Живлов (2008) has suggested one direct Baltic loanword in Permic. According to him, Komi *važ* ~ Udm. *vuž* 'old' are derived from Baltic **vetuša*- (> Lt. obs. *vetušas*, Lv. *vęcs*) 'old'. This seemingly attractive etymology has generally been well received (e.g. Pystynen 2016; Nikulin 2016). As Живлов notes, in inherited words, the correlation Komi *a* ~ Udmurt *u* is otherwise only observed as a reflex of the sequence **-etə*-. Komi *va*, Udm. *vu* 'water' (< **wetə*), Komi *ma*, Udm. dial. *mu* 'honey' (< **metə*).[317] It does not necessarily follow from these examples, however, that the conditioning factor was the lost **t*; we might, for instance, rather be dealing with a special vocalic development in **CV*-type roots (Лыткин 1964: 172).[318]

True, one of Живлов's examples is indeed word internal: Komi *tar*, Udm. *tur* ~ F *teeri*, dial. *tetri*, E *teder*, Li. *te'ddõr* (< **tetri*, OBL. **tetre*-) 'black grouse' (UEW 794); however, the Finnic–Permic equation is

[317] Живлов's third example *za* 'stem, stalk, shaft', Udm. *zu* 'stem of a pipe; axle of a cart' < **setV* is based on an equation with Ma. E *šüδür*, W *šəδər* 'axle; spindle' (UEW 757–758). However, since the Mari word has *š*- in the Malmyž dialect, the comparison is most probably incorrect. UEW reject the older comparison of the Permic word with Erzya dial. *sad* 'stalk (of the hop plant or cucumber)', yet this might be more promising. If we reconstruct **sę̣tə* for both forms, however, we will have to explain the difference between *za*, *zu* 'stem' and Komi *vo*, dial. (Upper Sysola) *o̤* 'year', Udm. *wa-pum* 'time, period' (< **ę̣də*) and (e.g. was the lowering to Proto-Permic **å* blocked by the *w*-prothesis?).

[318] As there appear to be no monosyllabic nouns in Komi *-o̤*, we might entertain a regular development of **e(–ə)* > Permic **o̤* in monosyllables followed by a further lowering **o̤* > **a* in Komi. This two-stage analysis is supported by the fact that two verbal stems of the shape **Co̤*- have been suggested to derive from **e(–ə)*; viz. Komi *lo̤*- 'be, become' (?< **lexə*; cf. Metsäranta 2020: 327) and *vo̤*- 'come, arrive' (< **wexə*-; Metsäranta 2020: 146–147).

surely incorrect, as in other cases internal **-Cr-* has given **-rC-* in Permic.[319] As the Finnic word has been considered a Baltic loan (Lt. *tetervà* 'black grouse hen', Lv. *teteris* 'black grouse'),[320] Nikulin (2016) interprets this Permic word as a loan from Baltic, too, setting up a pre-Permic **teδərə*. However, it is more likely that the actual source is Iranian, cf. NP *taδarv* 'pheasant', Khot. *ttarä*, *ttatara-* '(Tibetan) partridge' (?< **tataru*, cf. Khot. *pasä* 'sheep' < **paću*).[321]

As Metsäranta (2020: 245) notes, the main weakness of this etymology is that the contact relationship depends on a single comparison. The Permic word has traditionally been etymologized as a cognate to F *vanha*, E *vana*, Li. *vanā* 'old' (< **wanša*: UEW 813; Sammallahti 1988: 544), despite the irregular vocalic relationship. In passing, Aikio (2015a: 33) has mentioned a Samoyed **wåntå* 'old' as a cognate to West Uralic **wanša*. This form does not appear in the appendix to that article, but is apparently based on the Selkup stem **kuântə-* attested in derivatives in Ket Selkup, viz. *kwândəj* 'old', *kwândəga* 'old man or woman' (Alatalo 1998: 20, 2004: 293).[322] The Selkup word seems to be

---

[319] Komi *bǫrd*, Udm. *burd* 'wing' ?← Iranian **patra-*; cf. Skt. *pátra-* 'wing'; Holopainen 2019: 180; Komi *ćęrs*, Udm. *ćers* 'spindle; axis' ← Iran. **častra-,* cf. Pashto *cā́x̌ay* 'spindle'; Holopainen 2019: 378.

[320] The Baltic loan etymology is phonologically problematic. The Finnic word is rather of echoic origin like Eastern Mari *küδər*, Obdorsk Khanty *kutər* 'black grouse' (**kütrV*?) and Turkic **kürtük* (> Shor *kürtük*, Khakas *kürtkü*) 'black grouse'.

[321] The same vowel correspondence from Iranian **a* is found in Komi *dar*, Udm. *durį* 'ladle' (~ Skt. *dárvi-* 'spoon'), and Komi *taśti*, Udm. *tuśtį* 'cup, bowl' (~ YAv. *tašta-*, MP (Pahlavi) *tšt'* 'bowl'; Rédei 1986: 68, 78). In addition, certain Iranian loans in Permic have predated the loss of intervocalic stops: Komi dial. *gu-* 'steal' (← **gada*; cf. YAv. *gaδa-*, Pashto *ɣal* 'thief'; Rédei 1986: 69); Komi *ruć*, Udm. *ʒićį* 'fox' (← Iran. **ropāća-*, cf. Parth. *rwb's* /rōbās/, Oss. I *ruvas* 'fox'; Palmér et al. 2021: 15).

[322] I thank Abel Warries for helping me track down this word.

a phonologically regular equivalent of Finnic **vanha* (compare Ket Selkup *kwâdəgej* 'left' ~ Estonian *vasak* 'left' < **wasa*; Aikio 2015a: 66) and the semantics are ideal, so that a Uralic form **wanša* can indeed be postulated. In this light, it becomes even more difficult to separate the Permic word for 'old', even if the Komi *-a-* remains unexplained.[323]

Thus, as with Mari, I am led to a pessimistic conclusion as to whether there are any Baltic loanwords in Permic. In both cases, the evidence is very limited and alternative accounts are possible. It therefore does not certainly surpass the threshold of coincidence.

1.4.5. Conclusion

The contacts between Baltic and the other West Uralic branches were by no means of the same calibre as those with Finnic. The evidence as regards Mari and Permic is inconclusive: all of the suggested examples have competing etymologies, and we cannot state with any confidence that any direct contact has taken place. In the case of Sámi and Mordvin, many of the etymologies previously proposed are formally or semantically dubious, and must be rejected. However, even if we limit ourselves to cases where the Baltic source has a clear etymology, there still remain a handful of convincing cases which cannot be rejected.

Table 5 overleaf illustrates the contact situation. Certain, direct loanwords are highlighted in bold.

---

[323] It is tempting to consider it a borrowing from another branch, probably Finnic (Saarikivi 2018: 312). However, the existence of Finnic loans already in Proto-Permic is doubtful, and Metsäranta has considered this proposal "anachronistic" (Metsäranta 2020: 245). From the point of view of vocalism it is possible to assume that Udmurt *vuž* is inherited, in which case we might limit the loanword proposal to Komi.

Table 5. Baltic loanwords in Sámi and Mordvin

Mordvin		Baltic		Finnic		Sámi	
		*ansā-	→	*ansa	=	*vuossē	‘handle’
		*kvēpV-	→			***kiepę***	‘soot’
*kerš	?←	*kreiš-					‘left’
lija	←	*lę̄ka-	→	*līka			‘surplus’
pandəs	←	*pantīs					‘bridle’
*sal	?←	*sāl-	→	*sōla			‘salt’
sod	←	*sôd(i)-					‘soot’
		*šalnā-	→	*halla	=	*suolnē	‘hoarfrost’
		*žara-	→	*hara	=	*suorē	‘fork’
		*šę̄na-	→	*hęina	≠	***suojnē***	‘hay’
		*širvi	→	*hirvi	=	*sęrvę	‘elk’
		*daglā-	→	*takla	=	*tuovlē	‘tinder’

The majority of the loanwords in Sámi are shared with Finnic, and this appears to suggest that the contacts largely took place through Finnic mediation. Nevertheless, at least two direct loanwords have to be accepted. The situation with regard to Mordvin is quite different. In both cases where Mordvin shares a loanword with Finnic, the reconstructed proto-forms cannot be reconciled. Therefore, contrary to the claims of previous research, it does not seem helpful to assume that any of the words entered Mordvin through Finnic mediation.

Given the small number of loanwords, we would expect the contacts to have been brief and incidental. However, as I have noted above with regard to Mordvin, the semantics are only partially consistent with this interpretation. Particularly remarkable is the loaning of a word for

'soot' into both Mordvin and Sámi, which is difficult to understand in an adstratal trade context.

As there is no positive evidence for the presence of the Balts in Fennoscandia, it seems most parsimonious to assume that the Balts came into contact with pre-Proto-Sámi speakers before the latter migrated into the region (contrast the illustration in Aikio 2006a: 45). Kallio (2009: 39) has suggested that the Sámi had already arrived in the peninsula in the late 2nd millennium BCE. Similarly, Lang (2018a: 26) has suggested that the Sámi may have begun their migration from the Upper and Middle Volga regions in the latter half of the 2nd millennium.

However, there does not seem to be any certain evidence against a comparatively late migration; the earliest loanword evidence from Germanic can be dated as late as the first centuries CE (Aikio 2006a: 39–40; Kallio loc. cit.), and there is no other linguistic evidence that would necessitate such an early arrival of pre-Proto-Sámi speakers. Lamnidis (et al. 2018) have noted that an individual showing Siberian ancestry in Finland (dated 300–800 CE) correlates with modern Sámi populations, but there so far does not appear to be any genetic evidence which would support an earlier arrival of Uralic populations, the first individuals in the Baltic region showing Siberian ancestry being dated to the Final Bronze Age (Saag et al. 2019). There is currently very little ancient DNA evidence from Fennoscandia, however, so it is possible that such ancestry will later turn up.

Linguistically, the single example of 'hay', if analysed correctly, would show that the independent contacts with Sámi took place at an earlier date than the contacts with Finnic, as the former would have predated the East Baltic monophthongization of inherited **ai*. While this is an extremely tentative conclusion, it is possible that the contacts took place further east, closer to the Middle Volga region. Indeed, the contacts between the Balts and pre-Mordvin speaking populations have

normally been located in the Volga-Oka region (cf. Grünthal 2012: 299–302), a proposal which has been encouraged primarily by hydronymic evidence (see the discussion on p. 77). The evidence of loanwords in itself is arguably a far stronger argument for a more eastern spread of the Baltic languages. However, we should note that, in the absence of any back loans, there is no necessity in assuming that this source language was the direct ancestor of any modern or attested Baltic languages. Rather we may be dealing with an eastern offshoot, which would permit us to place the ultimate Baltic homeland somewhere between this contact zone and the Baltic Sea region.