A History of East Baltic through Language Contact

Anthony Jakob

LEIDEN STUDIES IN INDO-EUROPEAN 24

BRILL

A History of East Baltic through Language Contact

Leiden Studies in Indo-European

Series editors

Alexander Lubotsky
Alwin Kloekhorst
Tijmen Pronk

VOLUME 24

The titles published in this series are listed at *brill.com/lsie*

A History of East Baltic through Language Contact

By

Anthony Jakob

BRILL

LEIDEN | BOSTON

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement nº 716732).

The Library of Congress Cataloging-in-Publication Data is available online at https://catalog.loc.gov
LC record available at https://lccn.loc.gov/2023052014

Typeface for the Latin, Greek, and Cyrillic scripts: "Brill". See and download: brill.com/brill-typeface.

ISSN 0926-5856
ISBN 978-90-04-68646-5 (hardback)
ISBN 978-90-04-68647-2 (e-book)
DOI 10.1163/9789004686472

Contents

PART 2
Contacts with Unknown Languages

Acknowledgements

The first seed for my dissertation, on which this book is based, was planted back in 2017 when I told Tijmen Pronk I'd be curious to study the whole subject of "substrate words" in Indo-European. This topic had interested me ever since I first flicked through Derksen's etymological dictionaries. The idea that some of the words in the modern languages could have been borrowed from undocumented sources was something that seemed very appealing to me, especially in northern Europe where the historical record is so shallow. Tijmen put me in touch with Guus Kroonen, and a couple of years on, I was accepted onto Guus' project ('The Linguistic Roots of Europe's Agricultural Transition').

It was amazing to have been able to conduct my research as part of a team, and the weekly group meetings, daily lunches and constant exchange of ideas (and idiocy) were a reliable source of inspiration and the perfect antidote to the mundane aspects of research. This team mentality led to three joint papers (Palmér et al. 2021; Kroonen et al. 2022; Thorsø et al. 2023), the workshop *Sub-Indo-European Europe*, which we hosted in August 2021, and a forthcoming volume based on the results of this workshop. Not only that, but the methodology expounded and applied in the second half of this work is as much a team effort as it is my own. I would like to sincerely thank Yvonne van Amerongen, Paulus van Sluis, Cid Swanenvleugel, Rasmus Thorsø, Andrew Wigman and honorary team member Axel Palmér for making these four years so special.

My individual project took me in a number of unexpected directions. Sergejus Tarasovas has taught me the ways of philological pedantry, and shown me that there is light at the end of most rabbit holes. Countless insights in this work, both big and small, ultimately derive from our discussions (some more are to be published as Jakob forthc. b.). Alexander Savelyev has shown incredible patience in giving me a crash course in Turkic linguistics. Practically all discussions of Turkic languages in this monograph have greatly benefited from his input.

Most significant, however, was my journey to becoming a Uralicist (which perhaps I could now call myself, having spoken at the Finno-Ugricist Conference in Vienna, August 2022) — something I'd never considered when starting out on this project. A significant catalyzer for me has been the work of Ante Aikio, whose accessible writing style was the gateway I needed, but I would also like to thank Petri Kallio, Sampsa Holopainen, Juho Pystynen, Abel Warries and Mikhail Zhivlov for giving me heaps of constructive criticism and ideas which have made this work considerably better, and especially Santeri Junttila,

not only for the same, but also for providing me with an enormous amount of literature I would not otherwise have been able to access.

Dozens of other people have, at various moments, given me answers, made me question the answers I had, or made me realize I was asking the wrong questions. Without them, this book would not be half what it is, and I will mention just one dozen who happen to spring to mind: Lucien van Beek, Chams Bernard, Eugen Hill, Martin Kümmel, Giedrė Motuzaitė Matuzevičiūtė, Michaël Peyrot, Tijmen Pronk, Peter Schrijver, Zsolt Simon, David Stifter and Rémy Viredaz.

The project of which this work forms a part would not have existed at all, of course, if it wasn't for its initiator Guus Kroonen and Professor (now Emeritus) Sasha Lubotsky. They supervised me throughout the five years it took to write this monograph, and I know all too well that wasn't the easiest of tasks. I am eternally grateful to them both for keeping their faith in me and helping me grow as a person. It has been an eventful few years, with the Covid-19 pandemic and war in Ukraine both massively impacting my personal life. The former — for the better, since it meant me and my wife could spend an unexpected and magical year together in Leipzig, the latter — decidedly for the worse, although for us it has been rather overshadowed by a serious struggle with my mother-in-law's health. I am incredibly thankful to Guus for being so understanding and allowing me to take the steps I needed.

Finally, let me mention my parents, Dan and Barbara, my sister, Lara, and my adoptive parent, Natalya, who are always categorically supportive and who I can always rely on to be on my side, no matter what kind of pickle I'm in. Last and most of all, I am thankful to my wife Sasha. I know she has had to put up with a lot, including a lot of my absence, but she has brightened up every single day and given me a reason to carry on. My only hope now is that we finally get a chance to find out what "normal" means!

Tables

Symbols and Abbreviations

For abbreviated titles, see the relevant subsection of the bibliography.

Symbols

*	1. (before lemma) reconstructed form; 2. (after lemma) unattested form of an attested lexeme
**	counterfactual form
#	(in reconstructions) word boundary
∅	zero
·	introduces a gloss
=	'(historically) identical to'
~	'belongs with (in unspecified capacity)', 'in variation with'
/	1. mutually exclusive alternative, 2. (in citations) demarcates co-authors
:	1. paradigmatically alternates with; 2. correlates with (providing an analogical model)
∞	'irregularly alternates with'
<	'developed from', 'derived from'
«	'analogically replaces' (also used of semantic developments)
←	'borrowed from'
/ /	phonological transcription
[ ]	1. phonetic transcription; 2. alteration to quoted text or form
⟨ ⟩	orthographic transcription
C	consonant
H	laryngeal
N	nasal
R	resonant
T	dental stop
V	vowel

General

arch.	archaic
BCE	before Common Era
c.	century
CE	Common Era

cf.	confer, see
comp.	compiler(s)
dial.	dialectal
diss.	dissertation
E	east
ed., eds.	editor(s)
e.g.	*exempli gratia*, for example
et al.	*et alii*, and others
etc.	*et cetera*
fn.	footnote
forthc.	forthcoming
FS	Festschrift
i.e.	*id est*, that is
in prep.	in preparation
lit.	literature
loc. cit.	*loco citato*, at the cited location
N	north
N.B.	*nota bene*, please note
NE	north-east
No.	number
NW	north-west
obs.	obsolete
op. cit.	*opere citato*, in the work cited
p., pp.	page, pages
p.c.	personal communication
poet.	poetic
refs.	references
S	south
sp.	(unspecified) species
s.v.	*sub verbo*, under the corresponding entry
SW	south-west
trad.	traditional
trans.	translator(s)
usu.	usually
vel sim.	*vel similia*, or the like
viz.	*videlicet*, namely, to wit
vol., vols.	volume(s)
W	west

Languages

N.B. Abbreviations which only occur as second elements are given in [square brackets].

Alb.	Albanian
Arm.	(Old) Armenian
Aukšt.	Aukštaitian
Av.	Avestan
[B]	Baltic
Bel.	Belarusian
Bg.	Bulgarian
Bret.	Breton
Čak.	Čakavian (dialect of Serbo-Croatian)
Celt.	Celtic
Chuv.	Chuvash
CLuw.	Cuneiform Luwian
[Co.]	Cornish
CS	Church Slavic
	~ Bes. *Besědy na evangelije*
	~ Ps. Sin. *Psalterium Sinaiticum*
	~ Supr. *Codex Suprasliensis*
Cz.	Czech
Du.	Dutch
E	Estonian
[E]	English
Eg.	Egyptian
F	Finnish
Fr.	French
[Fri.]	Frisian
Gaul.	Gaulish
Gm.	Germanic
Go.	Gothic
Gr.	(Ancient) Greek
H.	Hesychius
[HG]	High German
Hitt.	Hittite
HLuw.	Hieroglyphic Luwian
HLv.	High Latvian
Ic.	Icelandic
IE	Indo-European
Ingr.	Ingrian

[Ir.]	Irish
Iran.	Iranian
It.	Italic
K	Karelian
Kajk.	Kajkavian (dialect of Serbo-Croatian)
Kash.	Kashubian
Khot.	Khotanese
Lat.	Latin
[LG]	Low German
Li.	Livonian
LSrb.	Lower Sorbian
Lt.	Lithuanian
[Luw.]	Luwian
Lv.	Latvian
M…	Middle, Mediaeval
	[MBel., MCo., MDu., MGr., MIr., MR, MUk., MW see second element]
Ma.	Mari
	~ E Eastern (Meadow) Mari
	~ NW Northwestern Mari
	~ W Western (Hill) Mari
Mac.	Macedonian
Md.	Mordvin
	~ E Erzya
	~ M Moksha
ME	Middle English
MHG	Middle High German
ML	Mediaeval Latin
MLG	Middle Low German
Mo…	Modern
	[MoHG, MoGr., MoIr., MoLG, MoW see second element]
MoE	Modern English
MP	(Manichaean) Middle Persian
Myc.	Mycenaean
Ngan.	Nganasan
NP	New (= Classical) Persian
Nw.	Norwegian
O…	Old
	[OAv., OBret., OCo., OCz., ODu., OFr., OFri., OIr., OPl., OR, OSw., OTur., OW see second element]
OCS	Old Church Slavic

OE	Old English
OHG	Old High German
ON	Old Norse
ONovg.	Old Novgorodian (dialect of Old Russian)
OS	Old Saxon
Osc.	Oscan
Oss.	Ossetic
	~ I Iron
	~ D Digor
P…	Proto-
	[PGm., PMa., PMd., PSá see second element]
Parth.	Parthian
PIE	Proto-Indo-European
Pl.	Polish
Pr.	(Old) Prussian
	~ E Elbing Vocabulary
	~ Gr. Grunau's vocabulary
	~ III Third Catechism
	~ TC Trace of Crete
PrG	Prussian German
PrLt.	Prussian Lithuanian
PU	Proto-Uralic
R	Russian
RCS	Russian Church Slavic
Sá.	Sámi (Saami)
	~ I Inari (Aanaar) Sámi
	~ K Kildin Sámi
	~ L Lule Sámi
	~ N North Sámi
	~ S South Sámi
	~ Sk. Skolt Sámi
SCr.	Serbo-Croatian
SCS	Serbian Church Slavic
Skt.	Sanskrit
Sl.	Slavic
Slk.	Slovak
Sln.	Slovene
Sw.	Swedish
Tat.	Tatar
To.	Tocharian
Tur.	Turkic

Turk.	Turkish
Udm.	Udmurt
Uk.	Ukrainian
Umbr.	Umbrian
Ural.	Uralic
USrb.	Upper Sorbian
Võ.	Võro (South Estonian)
Vp.	Veps
Vt.	Votic
[W]	Welsh
YAv.	Young Avestan
Žem.	Žemaitian

Grammatical

1... (2..., 3...)	first (second, third) person
ABL.	ablative
ACC.	accusative
ACT.	active
ADJ.	adjective
ADV.	adverb
COLL.	collective
DAT.	dative
DU.	dual
ELAT.	elative
ESS.	essive
F.	feminine
GEN.	genitive
ILL.	illative
INST.	instrumental
LAT.	lative
M.	masculine
N.	neuter
OBL.	oblique
PART.	partitive
PL.	plural
PRES.	present
PRET.	preterite
PRT.	participle
SG.	singular

Data Sources and Conventions

Reconstructions

In general, following in part the practice adopted by OED[3], I have avoided providing reconstructions unless this is necessary for the argument. The rationale is to maximize the emphasis given to attested data, and also reduce the need for me as an author to make a clear stand with regard to theoretical aspects of reconstruction where these are not strictly relevant to an argument. For instance, while one might object to my reconstruction of Lithuanian *avìs* 'sheep' as PIE *$*h_3eui$-*, it is unlikely anyone will object to its equation with Latin *ovis* 'sheep'. I have made an exception in the case of evidence from Uralic languages, for which I have provided reconstructions quite systematically. This is partly a means to provide additional clarity for readers more familiar with Indo-European than with Uralic, but is also a reflection of my own process in dealing with these languages.

In the following cases, my reconstruction differs from the established norm and/or requires certain clarification:

East Baltic — Acute intonation is marked with the circumflex or caret ⟨ ˆ ⟩, following the use of this symbol as an indication of the broken tone in Latvian tonal orthography.

Slavic — My reconstruction of the Slavic vowel system differs from the OCS-based one conventionally used (for instance, in the dictionaries of Berneker, Vasmer, Derksen, and ЭССЯ). The use of the standard reconstruction leads to anachronisms, such as an apparent claim that the Early North Russian source of e.g. Finnish dial. *akkuna* 'window' is more archaic than its Proto-Slavic ancestor. In general, I consider the Slavic vowel shift to be a Common Slavic, not Proto-Slavic development (cf. Olander 2015: 59–67), and therefore use a reconstruction with pre-vowel shift values.[1] The reconstruction used in this work is as follows:

1 Differently from Olander, I do not operate with a Proto-Slavic predating the monophthongization of diphthongs. One reason for this is practicality: for instance, it is often impossible to decide whether Common Slavic *i* derives in any particular case from an earlier **ei* or **ī*. However, I also do not consider it likely that the monophthongization was a post-Proto-Slavic development: the absence of the second palatalization in North Russian (cf. also Holzer 2001: 39–40) does not necessarily imply that it branched off before the development **ai* > **ē* since there is no reason to exclude an intermediate stage */kæ:/.

This work			Traditional		
ī	ȳ	ū	i	y	u
i		u	ь		ъ
e		a	e		o
ē		ā	ě		a

However, for ease of cross-referencing with works that operate with the standard notation, I have supplied the traditional reconstruction in brackets after each Proto-Slavic form.

Finnic — In line with the "Leiden" tradition of not marking allophonic features, such as syllabic resonants, in reconstructions, I do not mark consonant gradation, as this is entirely predictable in Proto-Finnic (except for **s* between unstressed vowels, where I have preserved the alternation with **h*). Therefore, I reconstruct Finnish *hammas* (GEN.SG. *hampaan*) 'tooth' as **hampas* rather than **hambas*, **hamβas* or **hamp̆as*.

Data sources

In compiling this work, I have endeavoured to check all the forms cited in primary sources. I have generally avoided citing data which I was unable to independently verify, unless this is crucial to an argument. A selection of literature used to source the forms from the most important languages for this work is presented below. The orthography follows the cited sources unless otherwise indicated. References not provided in full can be found in the bibliography. All web links are valid as of 31 May 2023.

Lithuanian	DLKŽ; LKŽ; *Bendrinės lietuvių kalbos žodynas* (ekalba.lt)
Latvian	*Latviešu literārās valodas vārdnīca* (tezaurs.lv); ME; LVPPV (tonal orthography follows ME; differences with LVPPV have been noted)
Prussian	PKEŽ and facsimile copies hosted at www.prusistika.flf.vu.lt; Trautmann 1910
Russian	*Большой академический словарь русского языка* [*а-продел*]; *Толковый словарь русского языка* (ed. Дмитрий Н. Ушаков); СРНГ; СДРЯ 11–14; СРЯ 11–17; СДРЯ. I have followed pre-revolutionary orthography in the use of the symbol ⟨ѣ⟩ (italics ⟨*ѣ*⟩), where this is etymologically relevant.

Ukrainian	*Словник української мови* (sum.in.ua); ЕСУМ
Belarusian	*Тлумачальны слоўнік беларускай мовы* (1977–1984); ЭСБМ
Polish	*Słownik języka polskiego* (ed. Witold Doroszewski); SSP
Czech	*Příruční slovník jazyka českého* (psjc.ujc.cas.cz); Gebauer; digitized resources at vokabular.ujc.cas.cz
Slovak	*Slovník slovenského jazyka* (slovnik.juls.savba.sk/?d=peciar); SSN
Sorbian	Schuster-Šewc; dolnoserbski.de; hornjoserbsce.de
Slovene	Pleteršnik (tonal orthography follows Pleteršnik; additional data from SSKJ² is given as in the source)
Serbo-Croatian	RJA; PCA; Skok (additional dialect data cited after Derksen 2008, 2015)
Church Slavic	*Старославянский словарь* (*по рукописям X–XI веков*); SJS; Miklosich 1865
Bulgarian	*Речник на българ ски език* (ibl.bas.bg/rbe/); БЕР
Macedonian	*Дигитален речник на македонскиот јазик* (drmj.eu)
Gothic	Streitberg, Wilhelm. *Gotisch-Griechisch-Deutsches Wörter buch* (1910) and interlinear texts at wulfila.be/gothic/
Scandinavian	Cleasby/Vigfusson 1874; *Dictionary of Old Norse Prose* and other digitized resources at onp.ku.dk; Blöndal 1989, SAOB, *Norsk ordbok* (no2014.uib.no, last accessed 31 May 2023)
English	DOE; MED; OE. Middle English allophonic lengthening is not marked
Dutch	Philippa et al.; digitized resources at gtb.ivdnt.org
Low German	Tiefenbach 2010; MndWb; Schiller/Lübben. MLG orthography generally follows the latter: allophonic lengthening is not marked; however, I have, after MndWb, distinguished *ö* from *o*
High German	AWb; Schützeichel 2004; MWb; *Mittelhochdeutsches Handwörterbuch* (ed. Matthias Lexer); DWb
Celtic	eDIL; GPC; *Le dictionnaire diachronique du breton* (ed. Martial Menard, devri.bzh)
Latin	TLL; Lewis/Short; Walde/Hoffman; Ernout/Meillet
Albanian	Mann, Stuart E. *An Historical Albanian-English Dictionary* (1948); Demiraj 1997
Greek	LSJ; additional data from Frisk; Beekes 2010.
Armenian	Martirosyan 2009

Finnish	SMS [*a–mähistyä*]; VKS [*a–pitäytä*]; SSA; SKES
Karelian	KKS
Veps	Зайцева/Муллонен 1972; Зайцева 2010. Orthography follows the latter
Votic	VKS
Estonian	EMS [*a–puisklema*]; VMS; *Eesti keele seletav sõnaraamat* (2009); other digitized resources at portaal.eki.ee; South Estonian: synaq.org
Livonian	Kettunen 1938; Viitso/Ernštreits 2012. Courland Livonian orthography follows the latter. Salaca Livonian: Winkler/Pajusalu 2009
Western Sámi	Digitized resources at kaino.kotus.fi/algu/. Orthography after Bergsland & Magga Mattson (1993). *Sydsamisk-norsk ordbok* [South]; Korhonen (2007). *Lulesamisk-svensk ordbok* [Lule]; Sammallahti & Ocejohka (1989). *Saamelais-suomalainen sanakirja* [North]. Data from older sources is presented in updated orthography
Eastern Sámi	Digitized resources at kaino.kotus.fi/algu/. Orthography follows Sammallahti & Morottaja (1993). *Inarinsaamelais-suomalainen sanakirja* [Inari]; Sammallahti & Mosnikoff (1991). *Suomi-koltansaame sanakirja* [Skolt]. Kildin Sámi data is updated from Lehtiranta 2001 on the basis of Rießler 2022
Mordvin	MdWb; Серебренников et al. (eds.), *Эрзянско-русский словарь* (1993), *Мокшанско-русский словарь* (1998). Phonemic orthography simplified from MdWb, and updated to reflect the modern standard
Mari	TschWb. Phonemic orthography after e.g. Aikio 2014
Permic	Digitized resources at dict.fu-lab.ru, Лыткин/Гуляев 1970
Khanty	OstWb
Mansi	Artturi Kannisto (comp.), *Wogulisches Wörterbuch* (2014). Phonemic orthography follows secondary literature

Introduction

The aim of this monograph is to place the East Baltic languages in their prehistoric linguistic context through the analysis of lexical borrowings. The work will be divided into two sections: in the first, I will critically assess the evidence for the established prehistoric contact relationships with Slavic (Chapter 1), Germanic (Chapter 2) and Finnic (Chapter 3) and examine proposals of contact with other Uralic languages (Chapter 4). The second half of the work will be devoted to the question of contacts with unknown languages, a complex and no doubt controversial subject, which has not yet had an extensive treatment. One of the aims is to establish applicable methodological principles for analysing this kind of material, and this half of the book can be seen as a practical demonstration and evaluation of these new methodological tools.

The result will be a detailed catalogue of the contact relationships in which the East Baltic languages participated. In order to stratify these linguistic events, I will also attempt to incorporate evidence from other disciplines, specifically archaeology, archaeobotany, and genetics, to evaluate the context and nature of the individual contact situations. This will be particularly important in the analysis of contacts with unknown languages (Chapter 8), as we *a priori* have no other information about the other participants in these contact events.

The focus of this work will be on East Baltic specifically. This is in itself unusual. Sabaliauskas (1990), for instance, stratifies the Lithuanian lexicon into the layers "Indo-European", "Balto-Slavic", "Baltic" and "Lithuanian", without distinguishing a separate East Baltic layer. Discussions of vocabulary exclusive to the Baltic languages likewise often fail to demarcate East Baltic as a distinct unit (e.g. Zinkevičius 1984: 229–234 and Larsson 2018: 1687–1688 are only concerned with isoglosses involving Prussian). This reflects a wider tendency in the literature, where one can easily find grammars and handbooks on Baltic (such as Stang 1966; Endzelīns/Schmalstieg 1971; Dini 2014) and grammars and handbooks on individual East Baltic languages (e.g. Endzelīns 1923; Kazlauskas 1968; Zinkevičius 1980–1981; Forssman 2001), but very little discussion of the East Baltic languages together, and basically no systematic attempt at reconstructing a separate proto-language.

There are, however, clear arguments for the separate study of East Baltic. Firstly, while the status of "Baltic" as a branch of Balto-Slavic has been disputed (Kortlandt 1977: 323; Derksen 1996: 1; Andersen 1996a: 63; Kallio 2008: 265; Kim 2018: 1974),[1] the coherence of East Baltic as a subgroup appears to be univer-

1 Villanueva Svensson (2014: 164) mentions Иванов/Топоров (1958) and Harvey Mayer (e.g.

sally accepted (albeit often implicitly). This can be demonstrated by a small but robust set of innovations exclusive to East Baltic. Clear cases are the following:[2]

1. **ai* and **ei* merged into **ę̄* in certain environments (becoming further diphthongized to /ie/ in both standard languages), thus Lt. *líepa*, Lv. *liẽpa* 'linden' (= Pr. TC *leipen* ACC.SG., R *лúпа*) beside Lt. *sniẽgas*, Lv. *snìegs* 'snow' (= Pr. E *snaygis*, OCS снѣгъ) (Hirt 1892: 32–41; Stang 1966: 53–57; Hill 2016: 208).[3]
2. Probably related to this is the appearance of a prothetic *v-* in the word for 'one', cf. Lt. *víenas*, Lv. *viêns* (contrast Pr. III *ains*) (cf. Fraenkel 1950a: 26–27; Petit 2010: 14).
3. A stem with *-v-* has been generalized in the paradigms of the 2SG. and reflexive pronouns, and corresponding possessive adjectives. Thus ACC.SG. **ten* (> Pr. III *tien*, OCS тѧ) and DAT.SG. **tebVi* (> Pr. III *tebbei*, OCS тебѣ) have been replaced by Lt. *tavè*, Lv. *tevi* and Lt. *táu* (< *tãvi*), Lv. *tev*, respectively (Petit 2010: 14; Hill 2016: 209–210).
4. Initial *m-* has been generalized throughout the paradigm of the 1PL. pronoun: cf. Lt. *mū́sų*, Lv. *mũsu* GEN.SG. (against Pr. III *noūson*, OCS насъ) (Forssman 2001: 44; Petit 2010: 14).[4]

Aside from this, a number of convincing isoglosses can be found between East and West Baltic,[5] but also some seemingly non-trivial isoglosses between East

---

Mayer 1978) as sceptics of the Proto-Baltic theory, but the scepticism in both cases seems more directed at the *Stammbaumtheorie* in general and against Balto-Slavic unity in particular. Both use the term "Baltic" liberally in the traditional sense.

2 Most lists only note differences between East and West Baltic without distinguishing innovations from archaisms (Endzelīns 1944: 17–21; Forssman 2001: 42–46), or include isoglosses with which Slavic also participates (Petit 2010: 12–17).

3 Although the conditions of this merger are not fully resolved (see the discussions in Kuryłowicz 1956b; Stang 1966: 58–61; Mathiassen 1995; Petit 2003: 96–97), the high level of agreement between Lithuanian and Latvian shows that we cannot, at least, be dealing with a later areal development.

4 I exclude: (a) the change **-tl-* > *-kl-*, which is also shared by North Russian (Николаев 1989: 190–198; Зализняк 2004: 49), and is therefore to be considered an areal phenomenon which might have spread through an already diversified East Baltic; the development also seems to have taken place in the Prussian dialect of the Elbing Vocabulary (cf. in particular *sasin-tinklo* 'snare' where we cannot blame the graphical confusion between ⟨c⟩ and ⟨t⟩); (b) likewise, the loss of the neuter gender in nouns seems already to have been spreading to the dialect of the Prussian Third Catechism (cf. Endzelīns 1944: 84; Fraenkel 1950a: 28); cf. *unds* NOM.SG. 'water' against Pr. E *wundan*. On alleged traces of the neuter in Finnic loanwords, see 3.3.3.

5 See most recently Villanueva Svensson (2014) and Hill (2016), against which Kortlandt (2018). Here I would like to add another argument: the 1PL. and 2PL. pronouns, Lt. *mẽs*, *jū̃s*, GEN. *mū́sų*, *jū́sų* correspond exactly (except for Innovation 4, above) to Pr. III *mes*, *ioūs*, GEN. *noūson*, *ioūson*. In OCS, we find мъı, въı, GEN. насъ, васъ, where the oblique forms are old

Baltic and Slavic (Villanueva Svensson 2014: 163; Kortlandt 2018: 176). For the purpose of this work, an agnostic stance can be considered acceptable, as the internal structure of the Balto-Slavic family does not have any bearing on the validity of East Baltic as a subbranch.

While William Jones' famous idea that Germanic was "blended with a very different idiom" can be seen as foreshadowing a whole subfield within Germanic studies (cf. Kroonen 2012: 240), the reputation of Baltic has developed quite differently. As Antoine Meillet (1913: 205) famously put it, a person who wishes to hear an echo of what Indo-European sounded like "va écouter les paysans lituaniens d'aujourd'hui" (despite Dini 2014: 45, fn. 21, I have verified this quotation to be genuine). This continues a legend present in non-specialist literature since the 19th century. Thus, the *Encyclopædia Britannica* (9th edition, 1882; cited per Klimas 1957) claimed that "whole Sanskrit phrases are well understood by the peasants of the banks of Niemen", and one still often comes across claims that Lithuanian is "the oldest" (*Encyclopedia of Linguistics*, 2004, ed. Phillip Strazny, p. 119) or "most archaic Indo-European language still spoken" (as in the current online edition of the *Encyclopædia Britannica*, accessed 15 August 2023).

It is true, of course, that Lithuanian is remarkably archaic in certain aspects. In terms of phonology, it probably can indeed lay claim to being the "most archaic", and in nominal morphology its only serious competitor is Slavic (see the discussion in Erhart 1995). If we take the liberty of writing the Sanskrit *sandhi* variant *-s* (rather than usual *-ḥ*), then it is not difficult to assemble a collection of forms where Modern Lithuanian and Vedic Sanskrit appear almost identical (see Table 1, below).

TABLE 1 Words similar in Lithuanian and Vedic Sanskrit

Lt. *výras* 'man'	**Skt. *vīrás* 'man, hero'**
Lt. *sūnùs* 'son'	Skt. *sū́nus* 'son'
Lt. *ugnìs* 'fire'	Skt. *agnís* 'fire'
Lt. *šuõ*, GEN.SG. *šuñs* 'dog'	Skt. *śvā́*, GEN.SG. *śúnas* 'dog'

(cf. Lat. *nōs, vōs*). To explain the Baltic oblique forms, it seems we have to assume a two-stage development: first, the strong stem **jûs* spread throughout the 2PL. paradigm, yielding a new GEN. **jûs-un*; second, the vocalism of the 1PL. **nōsun* was modified after the 2PL., resulting in a new stem **nûs-*. These two non-trivial and consecutive developments seem to provide strong evidence of a common Baltic stage.

For context, compare the Modern Hindi *bīr* 'hero', *āg* 'fire', and *sūnā* 'dog',[6] or the continuants of these words in other modern languages: Irish *fear* /fʲaɾ/ 'man', Icelandic *sonur* 'son', Slovene *ógənj* 'fire'. The surface similarities in the above table are admittedly partly accidental, but mainly result from a phonological conservatism on the part of Lithuanian. This conservatism has no doubt led to the stereotype of Baltic as a 'pure' dialect which has had "little or no non-IE contact" (Nichols 1998: 254) and "has not mixed with any other Indo-European or non-Indo-European language" (Klimas 2002).

Finnish in many respects holds a position similar to that of Lithuanian: many words in the modern language "appear almost bizarrely archaic" (Aikio 2022: 5), being identical to their reconstructed Proto-Uralic predecessors; thus e.g. *muna* 'egg' (< PU **muna*), *pesä* 'nest' (< PU **pesä*). At the same time, we know that the Finnic languages did not develop in isolation. In the Proto-Finnic lexicon, we can identify layers of loanwords from Slavic (cf. Kalima 1956; Kallio 2006), Proto-Norse (collected in LÄGLOS I–III) and Baltic (see Chapter 3), while North Finnic also contains a significant lexical substrate from Sámi (Aikio 2009). Thus, a conservative phonology does not necessarily presuppose a conservative lexicon.

I hope that this study will go some way towards dispelling the myth about the 'purity' of Baltic, and East Baltic in particular, in demonstrating that this branch, like any other, has a complex history and has been subject to numerous external influences.

---

6 According to Turner's CDIAL. I cannot find the word for 'dog' in modern dictionaries, so it is perhaps obsolete, or at least dialectal (perhaps Turner's source was John D. Bate, *A Dictionary of the Hindee Language*, 1875, p. 724).

PART 1

Contacts with Known Languages

∴

CHAPTER 1

Baltic–Slavic contacts

1.1 Early Slavic → Baltic Loans

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.[1]

1 On the substitution of Baltic **a* with Finnic **a* and **o*, see p. 63.

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
HLv. *žìvats*[2] 'animal' (= F *siivatta* 'livestock')	←	OR живóтъ, GEN.SG. животá 'life; animal' (R dial. *животы́* 'livestock; property')

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.

1.1.2 *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'[2] ?← OR, ONovg. търгъ, *u*-stem, 'market' (= OCS тръгъ); also → F *turku*, E *turg* (GEN.SG. *turu*) 'market' (Kalima 1952: 133)
- Lt. *puĺkas*, Lv. *pùlks* 'crowd, troop' ?← OR пълкъ 'troop, regiment; crowd, throng' (= OCS плъкъ)[3]

2 On the Latvian *-i-*, see fn. 16.

3 The rare F *pulkka* 'regiment, troop', quoted by Ahlqvist (1871: 209), shows a limited dialectal

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, Prussian Lithuanian *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'),[4] *kuĺda* '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).[5]

This observation raises suspicion with regard to the Aukšt. dial. *bulvõnas* 'idol; dummy', Lv. dial. (Vidzeme) *buĺvāns* (ME I: 349), *bulens* (EH I: 251) 'decoy bird' ← Russian *болва́н* 'idol; stuffed animal, dummy' (Даль² 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).[6] 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).

---

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).

4 Both formally and semantically, a more probable direct source seems to be Estonian Swedish *poik* 'boy' (Freudenthal/Vendell 1886: 165).

5 Lv. dial. *tuĺpîties* 'to crowd' ← *толпи́ться* (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 *tiĺpti* 'fit, have enough room' (compare Smoczyński 2018: 1484), although the details are unclear. Incidentally, the Latvian accentuation might also speak against an old loan (see 1.1.7).

6 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'.

The four cases cited at the start of this section could theoretically be dated to Proto-East-Baltic. In the case of Lt. *tur̃gus*, Lv. *tìrgus*, cognancy with Slavic is also an option (Endzelīns 1899: 299; REW III: 123).[7] In 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.[8]

The Slavic word **pulka/u-* (trad. **pъlkъ*) 'crowd; military regiment' is a loan from Germanic (cf. OHG *folc* 'people, crowd, troop').[9] 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 Chapter 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 бръна 'armour' (also бръньꙗ F.PL., cf. СРЯ 11–14 I: 321) / ← Go. *brunjo* 'breastplate'[10]
– Lt. *kãtilas*, Lv. *katls* 'kettle' ← OR котьлъ 'kettle' / ← Go. *katils** 'kettle'
– Lt. *stìklas*, Lv. *stikls* 'glass' ← OR стькло 'glass' / ← Go. *stikls* 'cup, chalice'[11]

---

7 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. 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 Chapter 2). Significantly, their account fails to explain the Albanian data.

8 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 1.1.7.

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

10 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 пѣньзь 'coin' (СДРЯ 11–14 IX: 407) ← OS *penning* 'penny' (where we must assume an original **pēnēʒe-* (trad. **pęnędzь*) by dissimilation; cf. the close parallel in OCS мѣсаць 'moon', dissimilated from **mēsēce-* (trad. **męsęcь*) < IE **meh$_1$ns-* '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-* (trad. **brǫnja*) (> Old Russian ***броуня*). Therefore, a Gothic source seems preferable.

11 Although the Slavic word seems to fit better semantically, it cannot be excluded that the

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. Pilipčiauskienė (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 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).[12]

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-* (trad. **kъmetь*; > OPl. *kmieć* 'serf, peasant')
– Lt. *krìkštas* 'baptism' ?← Pr. **kriksta-*, cf. Pr. III *crixti lāiskas* 'tauffbüchlein' ← Lechitic **krista-* (trad. **krьstъ*; > OPl. *krzest* 'baptism'; SSP I: 257)

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).

12 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. *poseł*, GEN.SG. *posła*. The High Latvian *kubyls* cited by Būga (1925: 38) as evidence of an unsyncopated form is considered dubious by ME (II: 297).

For Lt. *krìkštas*, an ultimately West Slavic origin is attractive for semantic reasons. Continuants of Proto-Slavic **krista-* (trad. **krьstъ*) mean 'baptism' in traditionally Catholic areas (i.e. West Slavic and Slovene), while in Eastern Christianity they have the meaning 'cross, crucifix'. 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).[13] 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 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. *alk̃snis*, dial. *alìksnis* 'alder' = R *ольхá*; 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 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* (< **brū̃šė*, cf. Trautmann 1910: 145)
– Lt. (W Žem.) *jū́šė* 'fish soup; slops, viscous liquid; mixture' ← Pr. E *iuse* · juche

13 It is perhaps this semantic issue which encouraged Smoczyński (2018: 606) to interpret *krìkštas* as a back formation from the verb *krikštyti*, 3PRES. *-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.

– 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).[14]
– 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-*.
– ⟨Ź̃ukmistras⟩ 'Fischmeister' (ClG I: 663)[15] ← Pr. E *suckis*, III *suckans* ACC.PL. 'fish'

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.[16]

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. *bi̇̃rkavs* '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)

14 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.

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

16 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. *Póvilas*, Lv. *Pāvils* 'Paul' ← OR Павьлъ (Būga 1925: 44). However, see the discussion on p. 11
– Lt. *pipìras* 'pepper' ← OR пьпьрь, OPl. *pierz*, *pieprz* 'pepper'[17]
– 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)[18]
– 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, compare 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'

Since Погодин (1903: 161–162), Lv. *cìlvẽks*, dial. *cilẽks* '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-* (trad. **čьlvěkъ*) 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 (Endzelīns 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

17 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*).

18 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-*.

Slavic would speak against Pl. *człowiek* deriving from a variant **čilavēka-* (trad. **čьlověkъ*; Havlová 1966: 80), and would favour an account starting from irregular syncope.[19] At any rate, in East Slavic, the only usual form is of the type человѣкъ, cf. also Uk. *чоловік*, Bel. *чалавек* 'person'. As a result, a donor form with **i* (trad. **ь*) is questionable.

Finally,[20] Lt. *grìkai*, *grìkiai*, Lv. *griķi* 'buckwheat' are often interpreted as early loans from the (unattested) Russian **грьк-* (Rūķe-Draviņa 1964: 125; Smoczyński 2018: 387; РЭС XII: 91) or directly from **грьча* (> MR *греча*; 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

19 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 *цыльёкъ* (Даль³ IV: 1301), but so far, I have not identified any forms nearer to the Baltic territory.

20 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 *бóчка* '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* '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' (compare 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.

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.

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.[21] 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. *pẽlni* 'ashes' (= Lt. *pelenaĩ*) and *ẽrglis* 'eagle' (= Lt. *erẽlis*) (Endzelīns 1923: 47), then why not suggest Lv. *kaĺps* 'farm-hand' and HLv. *kàrms*[2] 'building' derive from an earlier **kalapas* and **karamas*, respectively? On the other hand, setting up hypothetical forms like these to explain away any relevant evidence (cf. Mikkola 1938: 25–26) would be circular.

In the case of Lv. *žeŕbiņš* '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*;[22] Lt. *karvõjus* 'wedding loaf', dial. *karavõjus* ← Bel. *каравáй*;

21 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.

22 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*).

Lt. *skavardà* 'frying pan' ← Bel. *скавараду́* (cf. Lv. dial. Vidriži *skavardnīca*, without lengthening of *-a-*). A case not explainable as a pre-pleophony borrowing is Lt. *kalmašk̀a* '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-* (trad. **čerpъ*; > 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-* (trad. **xolpъ*; > OR холопъ 'servant')
– HLv. *kȁrms*[2], i.e. *kùorms* 'building' ← **xarma-* (trad. **xormъ*; > OR хоромъ 'house, building'; СДРЯ 1387)
– Lt. *šálmas* 'helmet' ← **šalma-* (> OR шеломъ, шоломъ, OPl. in Latin context *szłom* 'helmet'), earlier **šelma-* (trad. **šelmъ*).

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*).[23]

Occasionally, it is difficult to decide between cognancy and borrowing. Lt. dial. (N) *kařbas* 'basket', Lv. *kȁrba* '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.[24] A similar consideration applies to the derivative Lt. dial.

23 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* (trad. **o*) 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* (trad. **ъ*; cf. Кулешов 2010: 349).

24 *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-* (trad. **kъrb-*; see fn. 9); a Germanic loan may well underlie Cz. dial. (Machek 1968: 291) *krb* 'dovecote', SCr. *křbulja* 'basket made of bark' (PCA X: 449), as suggested by Berneker.

(Szyrwid, S Aukšt.) *kařbija*, *karbijà* 'kind of woven basket' and its comparandum ONovg. коробьѧ 'grain measure' (Зализняк 2004: 749).[25]

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).[26] Although certain Norse loanwords seem at face value to have predated the loss of the nasal vowels,[27] 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. Ἰγγωρ) ← OSw. *Inguar* (not *Ягорь); ижера 'Ingrians' ← Ingr. *Inkeroin*; ѣкорь 'anchor' ← OSw. *ankare* (not *оукорь; cf. Thörnqvist 1948: 99).[28,29] Similarly, for OR поудъ 'a weight measure' ← ON *pund*, one need not set up an intermediate form with *$\bar{o}$; 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

---

25 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 11: 117–118).

26 Николаев (1995: 111) has claimed that a distinction between *$\bar{o}$ and *$\bar{u}$ (trad. *o* and *u*) is preserved in some Carpathian Ukrainian dialects as /ʊ̯ʊ̯/ and /u̯ʊ̯/, 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.

27 A key argument here is OR варѧгъ, MR *вáрягъ* (Зализняк 2019: 722) 'Varangian', cf. ON *véringjar* NOM.PL., pointing to **vārēga*- (trad. **varęgъ*). On the other hand, if we set up a source form **várjængja*- (cf. Thomsen 1877: 121; Falk/Torp 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.

28 Thörnqvist (1948: 105) ultimately settles on a reconstruction **ękorь*, which is unlikely given the Norse *a*-. Note that Lt. *iñkaras* 'anchor' and Lv. *ẽķurs* go back to Early MoHG *enker* 'anchor' (cf. ME 1: 470) and do not support a Slavic nasal vowel, despite Berneker (1: 29).

29 Perhaps this could account for OR грамота 'letter, literacy, written document' ← Greek γράμματα (compare fn. 10). Although geminates were lost in Middle Greek, this was not universal, and this geminate occurs in the usual environment where 'spontaneous' secondary gemination is also attested (Holton et al. 2020: 135–136).

loaned weight measurement *bìrkavas* (cf. Derksen 2020: 37). On the other hand, Lithuanian also has a word *pū́das* (dial. also *pũ̀das*), whose acute accentuation might suggest a relatively early loan (cf. 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',[30] which has been derived from an early **lę̄xa-* (trad. **lęxъ*; > OR лѧхъ) (Brückner 1877: 103; Būga 1925: 33; LEW 356). Kiparsky (1948: 39; followed by 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.[31] A unique archaism would not be too surprising, as names of ethnic groups often figure among the earliest borrowings,[32] 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*), голѧдь (← **galind-*); see p. 29.

1.1.5 *Reflection of Slavic *ȳ*

One might anticipate that Slavic **ȳ* (trad. *y*) would have been substituted with Baltic **ū*, yet examples do not present themselves (Derksen 2020: 40). While the usual substitution for Slavic **ȳ* (trad. *y*) in older East Baltic loanwords is long /ī/,[33] 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:

---

30 Little faith can be given to the derivation of the ethnonym *Unguras** 'Hungarian', recorded only in Daukša, from Slavic **ǭgra-* (trad. **ǫgrъ*, cf. Būga 1925: 24; not **ǭgura-*, i.e. **ǫgъrъ*: 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.

31 Neither is the more primary **lę̄d-* (trad. **lęd-*) 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 **lę̄djān-* (trad. **lędjan-*; cf. REW II: 84 for details and references).

32 Thus OR литва 'Lithuanians' (see 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 1.1.6).

33 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. *buĩlis* 'chervil, chives', cf. MR *быль* (СРЯ 11–17 I: 364), Bel. *быллё* 'weed'
– Lt. *muĩlas* 'soap' ← Bel. *мы́ла* 'soap'
– Lt. *muĩtas*, Lv. *muĩta* 'toll' ← MBel. *мыто* 'toll'
– 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 /i̯/. 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 *выть*); *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*.[34]

In Prussian, however, /ui/ substitutes Lechitic **ȳ* (trad. *y*) without exception: Pr. E *waldwico* 'knight' ← **valdȳkā* (trad. *voldyka*), cf. OPl. *włodyka* 'nobleman'; **suiristio* (attested *sutristio*) 'rennet', cf. OPl. *syrzisko* 'rennet' (SSP VIII: 155); *wuysis* 'guard dog' ← **vȳž-* (trad. *vyž-*), cf. Pl. *wyżeł* 'pointer'; Pr. III *zuit* 'genug' ← **sȳta* (trad. *syto*), 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 *къила, cf. MR *ки́ла* 'hernia, outgrowth' (Зализняк 2019: 192)[35]

34 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. *мъигло, 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.

35 Since Endzelīns (1899: 310; ME II: 300), Lt. *kuilỹs*, Lv. *kuĩlis*, Pr. E ⟨tuylis⟩ */kuilis/ 'boar' have been derived from an Old Russian *къиль (thus also Trautmann 1910: 451; LEW 305; Derksen 2020: 41). However, such a source form cannot be set up (cf. Sabaliauskas 1968:

– Lv. obs. *suits*, *suišs* 'excessive' (whence probably Lt. *suitus* 'abundant' in Daukantas)[36] ← OR сытъ 'full, satiated', cf. R *сыто будетъ* 'много, лишку, нескромны требованія твоя' (Даль² IV: 386)
– Lt. *tuĩnas* 'branch in a wattle fence; fence post; picket fence' ← OR тынъ 'fence; defensive wall'[37]

These tie into the debate regarding the phonetic value of Slavic **ȳ* (trad. *y*). 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).[38] 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:

176; PKEŽ II: 294). R dial. *киля́к*, *килу́н* 'boar' are clearly derived from *ки́ла* 'hernia' (cf. *горбу́н* 'person with a hump' < *горб* 'hump'); compare the dialectal senses *килу́н* 'animal with a hernia; an animal (usu. piglet) with abnormal testicles; uncastrated boar' (СРНГ XIII: 209). Perhaps, as Smoczyński (2018: 622) surmised, *kuilỹs* 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).

36 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*.

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

38 The most striking indication perhaps remains the fact that this sound is rendered with the digraph *ũi* in OCS: Glagolitic ⟨ⰗⰫ ~ ⰗⰬ⟩, Cyrillic ⟨ъı ~ ъи⟩.

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 **ȳ*	2	1
-uo- for **ū*	1	**18+**

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

1.1.6 *Reflection of Pre-Slavic *ǭ, *ę̄*

In Latvian, Slavic **ū* (trad. **u*) 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*, *капу́ста*, *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 1: 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, above).

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.

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 conceiv-

able that the reflex of Slavic *$\bar{u}$ was pronounced closer to */o̜:/ 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 *$\bar{e}$ (trad. *$\check{e}$), which has yielded *-ie-* in Lithuanian but *-ē-* = */æ:/ in Latvian, cf. Lt. dial. *griẽkas* but Lv. *grę̃ks* 'sin' (cf. Bjørnflaten 2006: 68; Derksen 2020: 39). In this respect, we can note sporadic instances of /'a/ from *$\bar{e}$ recorded in modern north-western dialects, cf. NW dial. *кя̀п* 'flail', Arxangel'sk *ря́па* 'turnip' (= R *цепь*, *рябá*; Николаев 1990: 60), *ця́лой* 'whole', Vladimir *медвя́дь* 'bear' (= *цѣльный*, *медвѣдь*; Галинская 1993: 39–40), which indicates that the difference was at least partly dialectal rather than diachronic (cf. Derksen 2020: 47).[39]

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 *$\bar{o}$ > $\bar{u}$; *$\bar{u}$ > *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 would in turn be nativized in Low Latvian as *duõmât*, following the typical dialectal correspondences.[40] Since it is generally assumed that the early loanwords in Latvian are of East Slavic origin, such a trajectory would not be surprising.[41]

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

---

39 Compare the substitution of this phoneme as *$\bar{a}$ in the Finnic loanwords (Kalima 1956: 37–38), matching Latvian.

40 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. http://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 *kìukûļi* /kùkuļi[2]/ 'corn cockle' (← *куколь*) and *pỳuka* (cf. *pùka*[2], ME III: 445) 'fluff' (← *пух*), which have undergone diphthongization of original *$\bar{u}$, might even be interpreted as more archaic (cf. Seržant 2006: 95).

41 Similar evidence of interdialectal diffusion is shown by the Žemaitian diphthong /ie/, which occurs instead of the usual reflexes of *$\bar{e}$ 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).

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; Būga 1922: 177; Skach 2010: 137; Derksen 2020: 38).[42] 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 (Olonetsian) agent noun suffix -*niekka* (McKenzie 1918: 172; Mikkola 1938: 33–34; Kiparsky 1948: 31; Derksen 2020: 39) is problematic, as long vowels outside of initial syllables were not possible at the time of the earliest contacts with Slavic[43] and the diphthong /ie/ is generally not permitted at all outside of initial syllables in most Finnish dialects. The vocalism must have something to do with the reanalysis as a compounding element (J. Pystynen p.c. June 2023), which is supported by its abundant attestation as a separate word in earlier Finnish (*Vanhan kirjasuomen sanakirja*, s.v. *niekka*). 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* (Mikkola 1894: 57).[44] 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

42 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).

43 As witnessed by the reflection of *yat'* in the loanwords K *netäli*, Võ. *nätäl' -i*, Li. *nädįl̓* (< **nätäli*) 'week' ← *недѣля* 'week' and F *veräjä*, E *värav* (< **väräjä*) ← *верея́* 'gatepost' (< **verējā-*, trad. **verěja*; cf. OCS верѣꙗ 'bar, bolt').

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

II: 88–89: Arxangel'sk), and most probably results from a contamination with *вѣтер* 'wind'.[45] Olonets (19th c.) *мьръ* '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*.[46]

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.

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);[47] cf. Slk. dial. *týň* 'branch in a fence'.

45 A parallel development is found in MUk. (17th c.) *вьхорь*, which is actually attested in collocation with *вьтеръ* (РЭС VII: 270).

46 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*.

47 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. *tuřgus* 2 (rare dial. *turgùs* 4) 'market', Lv. *tìrgus* — MR *до торгу́* 'to market', *торго́вый* 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):

– Lt. *bìrkavas* 'a weight measure', Lv. *biřkavs* (ME I: 298; but [bìrkàu̯s] in LVPPV 130) — R *бе́рковец*, cf. dial. (Pskov) *бе́рковец* 'weight measure for flax' (РЭС III: 132–133)
– Lv. *kal̃ps*, (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. *kuřts* (but LVPPV: *kùrts*) — R *хорт*, cf. dial. *хо́ртица* '(female) greyhound' (СРНГ LI: 316); SCr. *hȓt*, Sln. *hȓt*, 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) *suĩtums* '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. *buĩvāns* '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 adop-

ted 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. *tuĨks* — MR *мóлка* GEN.SG., *молковы́и* ADJ. (Зализняк 2019: 569); cf. R *бе́з толку* 'in vain'
– Lt. *čérpė* 1 (but dial. also *čeřpė* 2, *čerpẽ* 4) 'clay pot; roof tile' — MR *чéрепъ* (Зализняк 2019: 602); cf. R *черепá* NOM.PL.

However, Lt. dial. *kařbas* 4 'basket', Lv. *kàrba*, if it is loaned from Slavic rather than cognate, would represent an exception, cf. MR *кóробъ*, R dial. *из 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.

1.1.8 *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 Chapter 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'].[48] 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. 11). All of these loanwords are suggestive of Slavic cultural imposition, and therefore can be classed as typical 'superstrate' words (cf. Vennemann 2011: 240).[49]

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* (trad. **ě*).

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

– 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*- (trad. *o*-) and *u*-stems was much better preserved in this dialect than in the rest of Slavic (Николаев/Хелимский 1990; Зализняк 2004: 99–102, 112).

48 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.

49 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.

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.

1.2 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); ѭтвѧгы 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. *пертъ* 'cottage'.[50] 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* > /u/ would suggest a very early date; but compare also Lt. dial. *mũmas* (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,

50 Аникин also cites R *клѣть* 'storehouse', with widespread Slavic cognates, as a Baltic loanword (cf. Lt. *klétis*, 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.

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 *ня́кля*, the substitution *a* → *a* clearly rules out an early date. The loss of the second syllable of the Russian word must be secondary; cf. dial. (N) *ня́кула*, Bel. *ня́кулле*, Pl. *pakuły* 'tow' (Лаучюте 1982: 18–19); compare the discussion of R dial. *ня́ккула* beside *ня́кля* '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 1890: 19). This is supported by MR *перетъ* 'bath-house' (Pskov, 15th c.; Гальковскій 1913: 34; cf. Вахрос 1963: 157), *пере́дка* 'hut' (Pskov, 16th c.; СРЯ 11–17 XIV: 298) and the dial. (Novgorod, Karelia) derivative *при́перетюк* 'dressing room in a bath-house' (Мызников 2019: 625; cf. СРЯ 11–17 XIX: 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'.[51]

Existing etymological discussions make the mistake of conflating the above forms with R dial. (Kem') *пертъ* '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), yet given the meaning 'bath-house' in the earliest attestations, a Baltic origin is semantically more attractive. For a further discussion, see pp. 140–142.

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

51 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. *pṛ́t-* '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).

II: 47), we can add that the variant *напортъ* (СРЯ 11–17 XIV: 148) might imply an original **purt-* (trad. **pъrt-*; with *нáпертъ* 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. *diřvinis* '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. *боркáн* (Novgorod, Pskov, Smolensk, Kaluga, etc.) ← Lv. *buřkāns* '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).[52]

52 Аникин interprets dial. *буркáн* as a later Letticism, but the limitation of this form to

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.).[53] 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).[54] It seems that 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 pp. 229–231.

▸ **? '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:

Leningrad Region practically excludes such an interpretation. It is evidently the result of pretonic *o* > /u/ attested sporadically in the area (ДАРЯ I, No. 1); compare СРГК I: 97, where forms of the type /burkán/ are listed under the headword *боркáн*.

53 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.

54 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).

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-* (trad. **jo-*) in Proto-Slavic.[55] In either case, we would probably be dealing with an early loan, although an exact dating is uncertain.

▸ **? '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.).[56] 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.[57] At a later date, the Baltic word was certainly borrowed into Russian *вéнтерь* 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).[58]

---

55 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. *орáбок* ~ 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. 175). 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).

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

57 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 online at https://www.kaisiadoriumuziejus.lt/enciklopedija) specifically states that the distinguishing feature of a *vénteris* compared to other fish traps is the *absence* of a supporting frame.

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

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-* (trad. **olьxa*) 'alder'.[59] 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) and nowhere else is it deverbal.[60] 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 1923: 49).

▸ † **'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

59 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.

60 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.

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. *нáсмо* '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 *пýсма* has nothing to do with the Latvian word.

⁘

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.[61]

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 1.1.8 that the earliest contacts of the Balts

61 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; cf. LKŽ s.v. *kiñkras*). For the Russian word, Николаев (1988: 135), offers an alternative etymology, comparing Vologda *ля́жа* 'damp, boggy place', and deriving both from **lę̄djā-* (trad. **lędja*), with a suggested (albeit controversial) Novgorodian development **dj* > *g*. Another account is given in Мызников (2019: 466).

(2) R dial. (Vjatka) *черв* (Даль² IV: 607) 'sickle' has been considered cognate to Lt. *kir̃vis*, 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 ⟨червáкъ⟩ · 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').

and Slavs were with speakers of Old Novgorod-Pskov. Similarly, the dialectal distribution of R dial. *боркáн* '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, and 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 work 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).[62] In any case, the alleged pervasiveness of a Baltic substrate in the hydronymy of this area contrasts starkly with the almost complete absence of evidence of early substratal loans on a lexical level.[63]

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.

---

62 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 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.

63 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).

CHAPTER 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 unpro-

ductive (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.[1] Senn (1925: 50), who accepts Būga's reasoning, nevertheless presents a counter-argument: as Norse ⟨f⟩, outside of initial position, stood for the sound /þ/ (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'.[2]

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. *īļens*, Pr. E *ylo* 'awl' ← Go. **ēla* and the Latvian hapax *glīsis* (ME I: 627)[3] 'amber' ←

1 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 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).

2 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.

3 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).

Go. **glēza-*. This would imply a narrowing of **ē* 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).[4] 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 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. *tũba*, 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

---

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

4 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. *árā-* '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-*.

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.[5]

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' (cf. Chapter 3, fn. 81); 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')[6] beside 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; Smoczyńsky 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:[7]

– Lt. *kiẽmas*, Lv. *cìems* 'village, farmstead' ← Go. *haims* 'village' — Contrast Lt. *šeimà* 'family' (and probably Skt. *śéva-* 'dear'; see Chapter 3, fn. 16)
– Lt. *kviečiaĩ*, Lv. *kvìeši* 'wheat' ← Go. *hvaiteis** 'wheat' — Contrast, if related, Skt. *śvítna-* 'white'[8]

---

5 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/.

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

7 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).

8 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'[9]
– ? 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 pp. 253–254 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; 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. *milìp* '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

9 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.

argued above that Lt. *šárvas* 'armour' (p. 13) and *tū́bas* 'felt' (pp. 39–40) may well have been borrowed through Prussian. I have also mentioned some words which ultimately derive from Germanic but which may have equally 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 *bruŋyos*, 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 стъкло '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).[10]

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

10 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.

the reconstruction of such a root.[11] However, while the etymological connection with the name of the Goths is tantalizing,[12] 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-* → ∅-, a Low German source seems more probable (Smoczyński 2000: 35–36; admittedly, a formally or semantically suitable source appears to be lacking).[13]
– ? 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 (italicized 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

---

11 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.

12 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).

13 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.

It is remarkable that the majority of the direct loanwords are connected in some way to stockbreeding and agriculture (note also p. 194 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.

CHAPTER 3

Baltic → Finnic Loans

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 Balts 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 work 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,[1] 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):[2]

– 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).[3]
– 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* (< **kausta*) 'side beam on a sledge' ~ 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)' (< **kẹ̄tara*) ~ Lt. *keterà*, *sketerà* 'ridge, peak; crest (of the back)' (Būga 1908: 72; Posti 1977: 265–266).[4]
– 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).[5]

---

1 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.

2 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 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*.

3 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).

4 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.

5 Liukkonen points out that the word has in Finnish also been applied to other large fish, and

– F *uksi*, E *uks*, Li. *ukš* (< **uksi*) 'door' ~ Lt. *úoksas* 'hollow; cavity' (Koivulehto 1993a: 34; SSA III: 369 with "?").[6]

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".

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

attested in the general meanings 'big fish; fish god', but these are clearly metaphorical extensions of 'sturgeon'.

6 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).

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'[7] (= Ngan. *koŋhu* 'wave') ~ Lt. *gumbas*, Lv. dial. *gumba* 'bump, bulge, excrescence'
– PF **lanci* 'damp lowland' (= Ngan. *l̥ĭntə* 'plain, valley') ~ Lt. *sklandùs* 'slippery, smooth, flat' — The semantic connection would be 'flat' : 'flat land'
– PF **mene-* 'go' (= Ngan. *mi̯nsi̯*) ~ Lt. *mìnti* (PRES. *mẽna*), Lv. *mĩt* '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. *gn̦uta* '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,[8] it

7 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).

8 See Smoczyński (2018: 1498). Lv. dial. (Kurzeme) *tâst*[2] 'hew' probably shows secondary length after the preterite (ME IV: 151), cf. the derived dial. (SW) *tastît* 'hew', and the similar phenomenon in Lv. *tèst*, dial. *têst* (ME IV: 175–176) 'hew, chop, adze', where we indeed find a

is difficult to make any conclusion with 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 3.4), he concludes that a few cases could represent "parallel borrowings from a shared source, perhaps a lost substrate language" (2015: 31).[9]

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:

– F *olut*, Vt. *õlud*, Li. *vȯ̀'l* (< **olut*) 'beer' ~ Lt. *alùs*, Lv. *alus* 'beer'; Pr. E *alu* 'mead'; Sln. (dated) *ǫ̂t* 'beer' = ON *ǫl* 'beer', OE *ealu*, GEN.SG. *ealoþ* 'ale' (LÄGLOS II: 310; Junttila 2012: 273)[10]
– F *rastas*, E *rästas*, Li. *rastā* (< **rasta*(*s*); see p. 103) 'thrush' ~ Lt. *strãzdas*, Lv. *strazds* 'thrush, starling' = Ic. *þröstur* 'thrush' (Qvigstad 1893: 259; LÄGLOS III: 130–131)[11]
– F *terva*, E *tõrv*, Li. *tõ̃ra* 'tar' ~ Lt. *dervà* 'tarry log; tar, resin, pitch', Lv. *daŕva* 'tar, pitch' = ON *tjara*, OE *teoru* 'tar' (LÄGLOS III: 289–290)[12]

---

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.

9 In this connection, he cites e.g. F *kinnas*, E *kinnas*, Li. *kīndaz* 'mitten, glove' (?→ South Sámi *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.

10 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. 41.

11 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 pp. 203–205.

12 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).

– F *vaaja*, E *vai*, Li. *vaigā* (< **vakja*) 'wedge' = Sá. S *vuevjie*, L *vuoǰvve* '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. *akéčios*, 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. *kueiʹmm* 'companion'), for which, rather confusingly, a Germanic loan etymology is almost never 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-*).[13]

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

– Li. *kil* ~ *kiḷ* 'black woodpecker' and *palāndõks* 'pigeon' both predate developments specific to Latvian (viz. the palatalization in Lv. *dzîlna* '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.

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.

13 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; Теренть ев 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).

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.

3.2.1 *Kinship*

– F dial. *ativo*, *atima* 'visiting relatives; married woman visiting her parents', Vp. *adiv*, *-on* 'guest; marriageable woman' (< **ateįiva* / **ateįivo*)[14] ~ Lt. *ateĩvis*, dial. *atéiva* 'foreigner, newcomer' (to *at-eĩti* 'come, arrive', cf. Skt. *éti* 'go') (Kalima 1939–1940: 211–214)[15]
– F *heimo*, Võ. *hõim*, Li. *aim* 'tribe, kin' (?< **hęimo*; see 3.3.1.4) ~ Lt. *šeimà*, Lv. *sàime* 'family, household'; Pr. III *seimīns* 'gesinde' (= OIr. *cóim* 'dear')[16]
– F *morsian*, K *moršien*, E *mõrsja*, dial. *mõrsija* 'bride, newlywed' (< **morcijan*, OBL. **morcijamę-*) ~ Lt. *martì*, ACC.SG. *mařčią* 'son's wife; bride', Lv. *màrša* 'brother's wife'; Pr. III *mārtin* ACC.SG. 'bride' (cf. Lat. *marītus* 'married (man)', Gr. μεῖραξ 'girl' < **mer-ih$_2$-*)[17]

---

14 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 **ateįivo*. 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 *-ęi-* (which emerged due to the reduction of **-aj-* under certain conditions, see Kallio 2012: 32–34). The diagnostic (South Finnic) forms for this reconstruction are unfortunately unattested.

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

16 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.

17 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).

– F dial. *nepaa*, arch. *nevat* (< **nẹpat*) 'nephew, niece'; E *nõbu* 'cousin' (< **nẹpoi*)[18] ~ Lt. obs. *nepuotis* (GEN.SG. *-ies*) 'nephew, niece' (= Skt. *nápāt-* 'grandson')
– E *sõsar*, Li. *sõzār*, ?Võ. *sysaŕ* 'sister' (< **sẹsar*) ~ Lt. *sesuõ* (OBL. *sẽser-*) 'sister' (= Skt. *svásar-*)[19]
– F *tytär*, E *tütar*, Li. *tidār* 'daughter' (< **tüttär*) ~ Lt. *duktẽ* (OBL. *dùkter-*); Pr. III *duckti* 'daughter' (= Skt. *duhitár-*)

3.2.2 *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-*)[20]

3.2.3 *Adjectives*

– F *ahdas*, Vt. *ahaz*, Li. *ǭ'dõz* (< **ahtas*) 'narrow, cramped' ~ Lt. *añkštas* 'narrow' (**aNž-* + **-sta-*; = Skt. *aṃhú-* 'narrow', cf. LEW 11)[21]
– 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')
– 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')[22]

---

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

19 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 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 **sẹsar* after all. The assumption of two independent loans is not very economical, and does not help explain the North Finnic **i*.

20 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 *nøf*) means 'nave (of a wheel)' (Kalima 1936: 141).

21 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-*.

22 Traditionally, one has compared the noun F *kelta* 'the colour yellow', Vt. dial. *kõlta* 'egg

– Vp. *kurdeh*, Vt. dial. *kurrõ* (GEN.SG. *kurtõõ*), E dial. (Mulgi) *kurre* (< **kurtęh*),[23] 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)[24]
– F *tyhjä*, E *tühi*, Li. *tijā* (< **tühjä*) 'empty' ~ Lt. *tùščias*, Lv. *tukšs* 'empty' (= Skt. *tucchyá-*)

3.2.4 *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)[25]

---

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. *gelt̃as*, Lv. dial. (ME I: 543) *dzęlts* 'yellow' are not attested in early sources. I prefer to take **kęltainęn* 'yellow' directly from Baltic **geltâna-* (or even **geltaina-*? cf. Lv. dial. *dzęltains*, ME I: 542) with adaptation to the Finnic adjectival suffix **-inEn* (cf. Liukkonen 1999: 39). An apparently old derivative is Lv. *dzęl̂ta* 'ground cedar', which has also been borrowed into Finnic (see below).

23 If we assume analogical generalization of the weak grade, K (Olonets) *kuuris* 'deaf' also appears to be regular from **kurtęs* (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 **kurtęs* and **kurnęs*, an analogical origin of the latter seems more promising than the assumption of a second, independent loan from Baltic **kur̂nas* (Kalima 1936: 124; Junttila 2019: 42).

24 F dial. *tavea* replaces *tavia* (< **taɣja*) 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.

25 Thomsen (1890: 244) compared a different Baltic word, Lt. *žãgas* 'hayrick, heap', Lv. obs. *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.

– F *halla*, E *hall*, Li. *ǭla* (< **halla*) 'frost, hoarfrost' ~ Lt. *šalnà*, Lv. *saĺna* (LVPPV: *sàlna*) 'frost, hoarfrost' (~ Lt. *šálti* = Oss. ID *sœl-* 'freeze')[26]
– F *helle* (OBL. *helte-*) 'hot weather', K dial. (Olonets) *helleh* ADJ. 'sweltering' (< **helteh*) ~ Lt. *šil̃tis* 'heat', cf. Lt. *šil̃tas*, 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* 'clubmoss' ~ Lv. *dzęĺta* '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')[27]
– 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*, *daglis* 'tinder (fungus)', whose primary use is as fuel; to Lt. *dègti*, Lv. *degt* 'burn' (= Skt. *dáhati*)

3.2.5 *Wild Animals*

– F *ankerias*, E *angerjas*, Li. *aṇgõrz* (< **ankẹrjas*) 'eel' ~ Lt. *ungurỹs*; Pr. E *angurgis* */angurjəs/ 'eel' (~ Lat. *anguīlla*)[28]

---

26 Thomsen (1890: 220) compares F *halli*, E *hall* 'grey; grey animal', Li. *aĺ* '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).

27 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.

28 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; Otrębski 1955: 26; LEW 1163) seems less probable. An alternative account is that the *u*-vocalism originated in the zero-grade (**h₂ngʷʰ-*; cf. Smoczyński 2018: 1561), implying an old ablaut variant.

– F *herhiläinen* 'hornet', E *herilane* 'wasp' (< **herhiläinen*) ~ Lt. *šìršė*, Lv. *sir̂senis* 'hornet'; note Lt. dial. *širšilas*, Pr. E *sirsilis* (= Lat. *crābrō*)[29]
– 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')[30]
– 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/)[31]

3.2.6 *Animal Husbandry*

– F dial. *ehkonen* (dial. *hehvo*; standard *hieho*), E dial. *õhv*, Li. *õ'v* (< **ęhva*) 'heifer' ~ Lt. obs. *ašva, ešva* 'mare'; perhaps Lv. ⟨ôssa⟩ 'mare' (Elger 1683: 133; cf. Karulis 1992 I: 468) (= Lat. *equa*)[32]

---

29 Lt. dial. *širšilas* 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).

30 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. *siřms*) 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 Chapter 6, fn. 109.

31 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.

32 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 3.6.1.

- F *hanhi*, E *hani*, Võ. *haaḱ* (< **hanhi*, OBL. *hanhẹ*-) '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* (< **paimẹn*) 'shepherd' ~ Lt. *piemuõ*, OBL. *píemen*- 'shepherd' (= Gr. ποιμήν)
- F *villa*, E *vill*, Li. *vīla* (< **villa*) 'wool' ~ Lt. *vìlna*, Lv. *vil̃na* 'wool' (= Skt. *ū́rṇā*-)
- F *vuohi*, Vt. *voho* (< **vōhi*, OBL. **vōhẹ*-) 'goat' ~ Lt. *ožỹs*, Lv. *âzis* 'he-goat'; Pr. E *wosee* 'she-goat' (= Skt. *ajá*-)[33]
- F *vuona*, Vt. *võdna*, Võ. *vụụn*, Li. *ūoṇõz* (< **vōtna*) 'lamb' ~ not attested in Baltic; cf. OCS агнѧ 'lamb' (= Lat. *agnus*, Gr. ἀμνός)[34]

On Võ. *pahr* 'boar', see p. 73.

3.2.7 *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* (?< **pẹlut* NOM.PL.)[35] 'straw chaff' ~ Lt. arch., dial. *pẽlūs*, Lv. *pęlus* PL.; cf. Pr. E *pelwo* 'chaff' (= Skt. *palā́va*- 'chaff', Lat. *pulvis* 'dust' < **pel-ou*-)
- F *siemen*, E *seeme*, Li. *sīemt* (< **sēmen*) 'seed' ~ Lt. *sėmuõ*, 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)[36]

---

33 For the Votic form, cf. *toho* < **tōhi*, OBL. **tōhẹ*- 'birchbark'.

34 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. *īṇõ*, (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).

35 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 "*рели*" cited in SKES 516.

36 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 (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 Pr. E ⟨wagnis⟩ can stand for */wagnas/, a Baltic source does not raise any phonological issues, while the neuter *s*-

3.2.8 *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')[37]
– F *kirves*, E *kirves*, Li. *kīraz* (< **kirves*) 'axe' ~ Lt. *kiřvis*, Lv. *cìrvis* 'axe' (~ Gr. κείρω 'crop, shave'; cf. Trautmann 1923: 135)[38]
– 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. *snúid* '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)[39]
– 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* (< **rattaheț* PL.) 'cart' ~ Lt. *rãtai*, Lv. *rati* PL. 'cart' (= Skt. *rátha-* 'chariot'; see above)
– 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)[40]
– F *silta*, E *sild*, Li. *sīlda* (< **silta*) 'bridge' ~ Lt. *tìltas*, Lv. *tiĨts* '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)[41]

---

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).

37 A Germanic origin (Sammallahti 1998: 123), cf. ON *és* 'eyelet (in a shoe)' (< **ansjō-*) is formally less straightforward.

38 On R dial. *черв*, which hardly belongs here, see Chapter 1, fn. 61. It seems natural to compare Lt. *kiřsti*, Lv. *cìrst* 'chop, cut' (= Skt. *kṛ́ntáti* '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).

39 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', with the orthographic sequence *-Vh-* simply standing for **V̄* as in Krevinian ⟨pählin⟩ 'head, leader' (= Vt. *päälin*), ⟨śohla⟩ 'salt' (= Vt. *soola*).

40 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.

41 Further, OR тьло 'ground, bottom', Pr. E *talus* 'floor', OE *þel* 'plank (of wood); plate (of

– F *tempaa-*, E *tõmba-*, Li. *tõ̃mbõ-* (< **tęmpaita-*) 'pull, tug' ~ Lt. *tem̃pti* (3SG. *-ia*), Lv. *tìept* (1SG. *-ju*, *-u*) 'stretch, tighten' (= NP *tāb-* 'twist, spin')[42]
– F *torvi*, E dial. *tõri* (< **torvi*, OBL. **torvę-*) 'horn (for blowing)' ~ Lv. *tàure* 'hunting horn', Lt. *taurẽ̃* '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')

3.2.9 *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)[43]
– 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)' (?< **rę̄uneḱ*)[44] ~ Lt. *briaunà* 'brim, edge' (~ ON *brún* 'eyebrow; brow (of a hill); strip of cloth', Skt. *bhrū́-* 'eyebrow') (Būga 1908: 42)[45]

---

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.

42 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').

43 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 3.3.1.6).

44 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.

45 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 **h_3b^hru-n-*, while Lt. *bruvìs*, Skt. *bhrū́-* 'eyebrow', etc. would reflect a fossilized dual **h_3b^hru-h_1*.

– 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)[46]
– F *vuoro* 'turn, shift', E *voor -u* 'turn, time' (< **vōro*) ~ Lt. *vorà* 'line, row', dial. 'turn, shift' (from Lt. *vérti*, Lv. *veřt* 'pierce; thread, string together') (Koivulehto *apud* Häkkinen 2004: 1514)[47]

3.2.10 *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. *zìêds* 'flower'; cf. Pr. TC *zaidiantẽ* ACC.SG. 'blossoming' (~ Lt. *žydéti* 'to bloom') (Топоров/Трубачев 1962: 247; Mägiste 1970)[48]
– 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'[49] (to Lt. *žėréti* 'glow, sparkle')
– F *härkä* 'ox, bull', E *härg*, Li. *ārga* 'ox' ~ Lt. *žìrgas*, Lv. *zìrgs*; Pr. E *sirgis* 'horse, steed' (~ Lt. (*ap-*)*žeřgti* 'sit astride')
– F *härmä*, E *härm*, Li. *ārma* (< **härmä*) 'hoarfrost'; F *harmaa*, Vt. *harmaa* (< **harmaka*), Võ. *haŕm*, *-i* (< **harmi*) 'grey' ~ Lt. *šarmà*, Lv. *saŕma* 'hoarfrost'; Lt. *šìrmas*, Lv. *siřms* 'grey, dapple-grey'[50]
– F *luuta*, E *luud*, Li. *lūdõ* (< **lūta*) 'broom' ~ Lt. *šlúota*, Lv. *sluôta* 'broom' (to Lt. *šlúoti* 'sweep')[51]

---

46 On the Mordvin and Permic words for 'salt', see pp. 137–138.

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

48 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*.

49 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.

50 Baltic **šaŕmā-* 'hoarfrost' is a derivative of **šiŕma-* 'grey'. The semantics can be illustrated by several parallels: (1) Lt. *šeřkšnas* '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. 26), and not least (3) ME *hore-frost* 'hoarfrost', cf. OE *hār* 'grey, hoary; grey-haired'. See also Liukkonen (1999: 38).

51 According to Kortlandt (1995), *šlúo-* regularly reflects **ḱleh₃u-*, and the Baltic words are to

– F *puuro* 'porridge', E *puder* 'porridge; mash' (< **putro*) ~ Lv. *putra* 'porridge, mash; a kind of soup', Lt. *putrà* 'gruel, skilly; flour soup'. Compare Lt. dial. (E Aukšt.) *pùtelis* 'thickened soup of oatmeal and milk'[52]
– 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')[53]
– 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)[54]
– F *seiväs*, E *teivas*, Li. *tāibaz*, Võ. *saivas* (?< **stęipas*; see 3.3.1.4) 'post, stake' ~ Lt. *stíebas* 'stalk, trunk, pillar', ?Lv. *stìebrs* 'stalk, reed, rush' (~ R *стéбель* 'stalk', SCr. *stáblo* 'tree; trunk')[55]
– 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')[56]
– F *tuura*, E *tuur* (< **tūra*) 'ice chisel' ~ Lv. *dùre* 'fist', dial. 'ice chisel' (from the verb Lt. *dùrti*, Lv. *duŕt* 'stab, poke, prick')[57]

---

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).

52 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.

53 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 Chapter 6, fn. 82). In that case, the root is perhaps that of Lv. *gãju* 'went'.

54 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.)'.

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

56 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; but differently see Simon 2022: 71).

57 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).

– Vp. *värpita-*, dial. *värbita-* 'spin, rotate (a spindle)', Li. *vērbikšõ*, dial. *vārbõkš* 'spin (thread)' (?< **värpi-* ~ **värpe-*) ~ Lt. *veřpti* (PRES. *veřpia*) 'spin (thread)', Lv. *vȅrpt* 'spin; (REFL.) wind round' (Thomsen 1890: 240; Posti 1946: 386)[58]
– F *vielä*, E *veel*, Võ. *vi̯i̯l* (< **vēlä*)[59] '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)

3.3 Analysis of Sound Substitutions

3.3.1 *Vocalism*

3.3.1.1 *ǣ → *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 **ā̆* (e.g. PF **kēle-* 'tongue' < pre-PF **kā̆lə* < PU **kälə*). This raising must have predated the emergence of secondary **ā̆* resulting from contraction over PU **x* and **ŋ* (e.g. PF **pā̆* 'head' < PU **päŋə*, UEW 365), which was no longer subject to raising.[60] Therefore, we should

58 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', OR 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).

59 Livonian *vēl*, *ve'l* represents an independent loan from Latvian (cf. Suhonen 1973: 237).

60 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.

rather state that Baltic */æ:/ was borrowed as pre-PF *$\bar{a}$, and the subsequent raising in Finnic and Lithuanian must be considered parallel, unrelated developments.

A confirmation of this chronology is provided by the substitution of PB *$\bar{a}$ as PF *$\bar{o}$ in e.g. PF **võhi* ← PB **âžē* 'goat'. Although in fact entirely analogous to the case of */æ:/ → *$\bar{e}$, only this substitution has sparked any significant debate. One has either operated with a Proto-Baltic reconstruction *$\bar{o}$ (Kalima 1936: 66–67; cf. Mikkola 1930: 443, fn.) or assumed that the Finnic people came into contact with a Baltic dialect in which *$\bar{a}$ 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 *$\bar{o}$ has itself developed from an earlier *$\bar{a}$ (Pystynen 2018: 72–75).

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

a. Baltic *$\bar{a}$-stems are overwhelmingly adopted as Finnic **a*-stems, cf. **halla* 'frost', **villa* 'wool'. Likewise, Baltic *-$\bar{e}$ was adopted as Finnic *-*ä* in **mäntä* 'whisk' (← **mentē*).
b. The substitution *$\bar{o}$ → *$\bar{u}$, **ou* is naturally accounted for if Proto-Finnic lacked a phoneme *$\bar{o}$ at the time of borrowing (see pp. 68–69).

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

a. Baltic *$\bar{a}$-stems are occasionally adopted as Finnic **o*-stems, as in F *heimo* 'tribe, kin' ← Baltic **šeimā-* (on which see 3.3.3)
b. F *vohla*, dial. *vohli* 'kid', if loaned from Baltic (cf. Lt. *ožẽlis* 'kid'), implies an underlying **võhl-*, 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 **āšęla*, which subsequently developed to **võhęla* (with automatic **v*- before *$\bar{o}$-) 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.[61]

61 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.

3.3.1.2 *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. *cèrme* 'roundworm', Lt. *kermenaĩ* 'bee larvae' (Thomsen 1890: 98; Liukkonen 1999: 54).[62] Besides this, we find examples of the substitution **ę*: **sęsar* 'sister', **nępat* 'nephew', **kęlta* 'ground cedar', **kęltainęn* 'yellow', **pęlut* 'chaff', **tęmpaita-* 'pull, tug', **ęhka* '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 suggests **ę* could have been a sort of compromise between the back-vocalic stem-vowel and front vowel of the initial syllable.[63]

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', **morcijan* 'bride' and **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*, **krauš-*. 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.

62 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).

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

The substitution observed in PF **härmä* ← **šârmā-* 'hoarfrost' is quite unclear.[64] The expected back vocalism seems to be found in F *harmaa*, 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 the 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 3.5.3.

3.3.1.3 *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' (← **širš-*) 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.[65]

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' (← **žirga-*) might suggest that we are dealing with a genuine sound change in a Baltic dialect,[66] which would provide evidence that the source language was not a direct

64 Since Thomsen (1890: 221), one has generally referred to a Latvian *sęrma* (cf. ME III: 819; EH II: 478) to support the reconstruction of a Baltic source form **šerma*. However, this Latvian form is probably the result of a secondary dialectal development (Endzelīns 1923: 36–37) and cannot be projected back to Proto-Baltic.

65 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. *kõlldd* '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. *skiltis* 'clove; slice; piece cut off' looks more tenuous.

66 This is far preferable to seeing the source in the deverbal noun *žargà* 'spread legs'

ancestor of the attested East Baltic languages. The fact that we find **ä–ä* as opposed to **ę–a* in this word (see above) possibly 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 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-*. 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 **ankęrjas* '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'.

3.3.1.4 ?*ę̄ → *ę̨i ~ *ī; *ō → *ou ~ *ū

A very interesting case as regards vocalism is PF **hęimo* ~ **haimo* 'tribe, kin'.[67] 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*)

(Liukkonen 1999: 55–56, taken over by Junttila in prep.), which lacks the required semantic specialization.

67 For a more detailed account of this problem, see now Jakob forthc. d.

– 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'

It is remarkable that the majority of these examples have Baltic comparanda.[68] Only the last is inherited from Proto-Uralic, where the cognates (e.g. Sá. N *čuožžut*, Eastern Mansi *tuńś-* 'stand') point to PU **saŋćа-*.[69] 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,[70] 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),[71] instead rather suggests that we 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

68 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. *rìẽkti* '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).

69 See Kallio (2007: 231–232; 2012: 35–36). Pystynen (2014a) rejects this reconstruction and prefers **sańćа-* (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 **лāńć-* ~ **лə̄ńć-* (Vakh *ļińt'-*, Kazym *лоńś-*) 'put, set' (< **saŋćа-*) and **kus-* (Vakh-Vasjugan *kŏs-*, Kazym *χŏs-*) 'urinate' (< **kuńćə-*). On this differently, see now Pystynen *apud* Живлов (2023: 144).

70 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).

71 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.

TABLE 3 i-diphthongs in Early Proto-Finnic

(*ī)	*ui
	*oi
*äi	*ai

BASED ON KALLIO 2018

i-diphthongs reconstructed for Early Proto-Finnic by Kallio (2018), the result is rather interesting (see Table 3, above).

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 words, cf. E *leibu*, Vt. *leipoi* PART.PL. < **-oita* < **-a-j-ta* (Kallio 2014: 159). As this 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*.[72]

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).[73]

72 The relevance of forms like Sá. N *suoidni* 'hay' < **šajna*, need not be overstated, as it is possible that the Sámi loans were adopted independently (see 4.1, particularly p. 124 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 **sańća-* > **sańśa-* > **sęjsa-* (cf. Ravila 1935: 32, fn. 1; Viitso 1978: 97). Although *ad hoc*, attributing a unique change to a unique environment is better than assuming a sporadic change with no conditioning factors. For more discussion of these points, see Jakob forthc. d.

73 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 11: 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'.

According to Kallio (2018: 262) **ei* has arisen secondarily in later Proto-Finnic due to contraction, cf. F *seimi*, Võ. *seiḿ* 'manger' < **sewi-mi*.[74] 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* (**pīmä* 'milk' < **pejmä*, of Iranian origin,[75] cf. Av. *paēman-* 'mother's milk', Holopainen 2019: 178–180). If the above account is correct, the Baltic loans must have postdated 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'[76]
– 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. 61). 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).[77]

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

---

In that case, Võro *hõim* would then need not have been borrowed from North Estonian (cf. Kallio 2021: 125).

74 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-*.

75 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 Chapter 4, fn. 6).

76 Note that, morphologically, Baltic **lę̄kas* 'surplus' more probably reflects an earlier **laikas* (= Gr. λοιπός 'left over').

77 This might explain the substitution of Baltic **ē* as **ęi* in some Livonian loanwords from pre-Latvian, provided these postdated the Livonian raising **ę̄* > **ī̆* (> Courland *ō̄*, Salaca *ū̄*; Kallio 2016: 49); compare Li. *kȭidaz* 'weaver's reed' (← Lv. *šķiets*, Lt. *skiẽtas*), *lȭiga* 'surplus' (← Lv. *lieks*, Lt. *liẽkas*).

**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.[78] 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).[79] At the same time, the Baltic loans predated the assibilation (see 3.3.2).
– If Finnic **?i* should be interpreted as **ęi*, the example F *reisi*, Võ. arch. *raiź* (< ?**rę̇ici*) 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 **i̇w* > **üw* witnessed in F *syvä*, Võ. *süvä* 'deep' < PU **ti̇wä* (Aikio 2015b: 9).[80] Importantly, the aforementioned example immediately confirms the suggested chronology, as the initial sibilant in F *syvä* implies an intermediate stage **civä*, whereby the change **i̇w* > **üw* must have post-dated the assibilation of **t*. As PU **e* and **ü* merged in Mari, this new

78 For **routa*, Junttila (2016b: 226) prefers an original **graudā*, reconstructed on the basis of R *ерýда* '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).

79 Compare PF **ūtin* 'mosquito curtain' < PU **owdəm(ə)* (= Eastern Mansi *åml*; Komi (Permjak) *ę̇n*, pointing to **o(-ə)*, see Aikio in prep. 81–82); PF **tūli* 'wind' < PU **towlə* (cf. Ma. W *tul* 'storm', Komi *tę̇v* 'wind'; see Aikio 2012b: 243); PF **kūsi* 'spruce' < PU **kowsə* (cf. Komi-Permjak *kę̇z*, Northern 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.

80 The change also has a potential parallel in Vp. *silöi*, E *siül*, 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 **kū̄* '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). For more discussion, and on the other possible exceptions, see Jakob forthc. d.

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 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,[81] and therefore an East Baltic origin of the loanwords (for a discussion, see 3.3.4). This, incidentally, can be seen as an argument for a Baltic origin of PF **hęina* 'hay' and **sęina* 'wall', even though the corresponding Baltic words do not have reliable Indo-European cognates beyond Slavic (cf. OCS сѣно 'grass, hay';[82] стѣна 'wall'; on the latter, see also pp. 219–220).

3.3.1.5 *eu → *ęu

In view of the etymology PF **ręuna* ~ Lt. *briaunà* 'edge', several scholars (Būga 1908: 42; Kalima 1936: 75) have argued that the Baltic loanwords predated the change **eu* > **jau*.[83] 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),[84] and

81 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. **pę́mōn*, 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 3.3.3.3). In the case of **häici* 'flower', a form with **âi* is actually attested in Lt. dial. *žáidas* 'flower'. A remaining question is how this conclusion can be reconciled with the evidence that the Gothic loans in East Baltic *predated* the monophthongization (see pp. 38–40).

82 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.

83 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).

84 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: 124), and would imply an earlier **driēud-* : **dreud-V́-*. This interpretation would be supported by

the idea that 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 /Crʲ/ 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 **rẹuna*. From a phonological perspective, a Baltic loan predating the change **eu* > **jau* seems preferable to the direct substitution **rʲau-* → **rẹu-* (Kulonen 1988).

3.3.1.6 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*), **ankẹrjas* (< **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* ← 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 **nẹpat* 'nephew' and possibly **sẹsar* 'sister' (see p. 82).[85]

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 **tíl̂ta-*) and **tj* > **cc* (in **męcca* 'forest' ← Baltic **medja-*),[86] as well as the metathesis

pievffen 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.

85 There are, however, a couple of examples which show non-initial **i* and **u* in Baltic loanwords, namely **pẹlu*(*t*) 'straw chaff' (adopted as **pẹlẹw* ~ **pẹlâw*?), and **oinas* 'ram' (adopted as **owẹjnas* ~ **owâjnas*?).

86 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*

*-*wR*- > *-*Rw*- (e.g. **torvi* 'horn' ← Baltic **taurē*) and the development *-*ln*- > *-*ll*- (e.g. **villa* 'wool' ← Baltic **vilnâ*). A few aspects deserve a more detailed discussion:

3.3.2.1 **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 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* (> -∅) 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 *kįrnįš* '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 4.4) would seem to contradict this sound law.[87]

'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' (< **fičču* ← Norse **fitjo* < Germanic **fetjō*-, cf. ON *fit* 'webbed foot; flipper').

87 In view of this, it is perhaps preferable to opt for Pystynen's alternative account that South

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.

3.3.2.2 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-*.[88] Moreover, in all other cases where an initial **c-* would be expected on etymological grounds, we find *s-* in South Estonian:

- 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,[89] 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.

Estonian originally preserved a form **hansi* (← Baltic **žans-*) which first developed to **hāsi* and only then was assimilated to **hāhi*. For more arguments for a late dating of Finnic **š* → **h*, see Pystynen (2023: 355–356).

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

89 Hosted at https://sanat.csc.fi/wiki/Luokka:Yhteissuomalainen_sanasto.

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

⁘

The Finnic cognates of F *seiväs* 'post, stick' are interesting in two respects. Not only do they show reflexes of the unclear diphthong **?i* (probably = **ẹi*, see 3.3.1.4), 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'[90] (~ Lv. obs. *stiepats* 'dace'; see p. 97)
– ? F *saparo* 'short tail' ~ Li. *tabār* 'tail' (?~ Lv. dial. *stebere* 'tail')[91]

The agreement between the words for 'wing' and 'post, stick' is striking: in both cases, the distribution between *s*- and *t*- is almost identical,[92] 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 con-

90 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).

91 But note F *sapa*, E *saba*(!) 'tail' (< **sapa*).

92 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.

stitute 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 **ę̆i* (see 3.3.1.4) 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 **stę̆ipas* in this chapter.[93]

3.3.2.3 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' (← **tanĥkjV-*) 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).[94] 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

93 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).

94 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.

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 3.2, only two contain a geminate: **rattas* 'wheel' and **tüttär* 'daughter', yet only a handful of others (e.g. **atẹivo* 'visiting relatives', **hako* 'branch' and **nẹpat* 'nephew') could have theoretically tolerated a geminate in Early Proto-Finnic. Steinitz (1964: 337) has proposed that the examples with 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,[95] 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,[96] 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.[97]

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):

– 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'

95 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 pp. 100–102). 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).

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

97 On the Sámi words for 'daughter', see Chapter 4, fn. 22.

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 pp. 140–142 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 **kāppä* 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).[98] The similarity with the Baltic word is therefore probably coincidental.

3.3.3 *Declinations*

3.3.3.1 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', **stᵩipas* '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: **kᵩltainᵩn* 'yellow' (fn. 22), **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*.[99]

98 Whatever reconstruction we use, it is about time we abandon the comparison with Mansi (South) *kᵩp*, (East) *käp* 'small hill' (still repeated in SSA I: 484; van Linde 2007: 84; Junttila 2017: 133); Mansi **ä* implies Proto-Uralic **i*, **e* or **ü* in the initial syllable. Futhermore, this word can hardly be separated from Mansi (South) *kᵩmp*, (East) *kämp* in the same sense, thus suggesting a Proto-Mansi **kämp*. In fact, the paradigm *käp*, OBL. *kämp-* is still recorded for the Middle Lozva dialect by Munkácsi/Kálmán (1986: 190).

99 Although the *u*-stem adjectives in Bretke appear to have been largely unspecified for

– Nouns: **hęina* 'hay', **hirvi* 'elk', **häici* 'flower', **härkä* 'ox', **męcca* 'forest', **pahla* 'rod, spit', **sikla* 'side net in a seine', **silta* 'bridge', **sōla* 'salt'.

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' < **terh$_2$-* (= 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.[100] 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.[101] Besides, an original masculine gender is secured for East Baltic **žirĝas* 'male horse' not only in view of the meaning, but also by Pr. E *sirgis*.[102]

---

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).

100 In Lithuanian, the originally pronominal ending *-a* occurs in predicative adjectives, but this does not imply that it was present in nouns, as it is logical that the 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).

101 A more extreme case is the mention of the isolated and surely secondary Veps dial. *kouvaz* (beside usual *kauh*) 'ladle, scoop' as evidence of a vacillation between neuter **kaûša* and masculine **kaûšas* within Baltic (Иллич-Свитыч 1963: 82; cf. Derksen 2015: 234).

102 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 3.5.3) and **kataka* 'juniper'

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 **hęinan* (← **šę̄nan* ACC.SG.) could easily have been apprehended as a genitive-accusative singular form, on which basis a new nominative such as **hęina* 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').

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. *zir̃nis* '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*.[103] Admittedly, the word for 'pea' may have arisen from an earlier *i*-stem (cf. Nieminen 1957: 206; Skardžius 1941: 53), although

(see pp. 84–85) have masculine cognates in Prussian, viz. Pr. E *attolis*, *kadegis*, and **vakja* 'wedge' (see p. 50) 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 pp. 217–218), 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).

103 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*; compare *wargien* 'copper', *kragis* (read **kargis*) 'army', *saligan* 'green'. Contrast Pr. E *tuylis* 'boar' (~ Lt. *kuilỹs*), *singuris* 'goldfinch' (?~ Lv. *žîguris* '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').

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 corresponds more or less exactly to that of Gothic *-jis* and *-eis* (Sievers' law; see Sommer 1914: 242 and *passim*):[104]

– **karjas* 'war, army' (> Lt. obs. *kãrias*, Lv. *karš*) = 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* 'wood', Go. *midjis*, Skt. *mádhya-* 'middle'
– **teljas* 'calf' (> Lt. *tẽlias*, Lv. *teļš*)
– **svetjas* 'guest' (> Lt. *svẽčias*, Lv. *svešs*)
– **varjas* 'copper' (> Lt. obs. *vãrias*, Lv. *varš*) = Pr. *wargien*

In view of this distribution, it is likely that **-jas* and **-īs* both reflect the same proto-form (i.e. **-ios*).[105] 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 **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

104 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. *zaļš* '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).

105 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.

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.

3.3.3.2 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', **rẹuna* 'edge', **tapa* 'way, custom', **villa* 'wool', **lūta* 'broom', possibly **takla* 'tinder'. On the other hand, there are a few examples which appear as *o*-stems, cf. **hako* (beside **haka*) 'branch', **hẹimo* 'tribe', **putro* 'porridge', **vōro* 'turn'.

As for **hẹimo*, 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*;[106] cf. the collocation *olla ativoissa* ~ *atimoissa* 'visit relatives', lit. 'to be in guests' (SMS), which is ambiguous between an *a*- and

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

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 **hẹimo*, 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. 63). Differently on the *o*-stems, see Holopainen, Kuokkala & Junttila (2022: 126–130).

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.

3.3.3.3 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.[107]

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

The words **sēmen* 'seed' and **paimęn* '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 **sā̄menen*, on which basis a new nominative singular **sā̄meni* > **sēmen* could have been backformed. In the case of **tüttär* 'daughter', both a nominative **duktēr* and an accusative **dukterin* could come into question.

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' (~ 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,[108] and Slavic continues a more archaic *u*-stem **ǭzu-ka-* (trad. **ǫzъkъ*) > OCS ѫзъкъ 'narrow, tight' (= Skt. *aṃhú-*).

108 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).

– 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 **hęina* '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.

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 pp. 64–65). 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.

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*, *kadęgs* (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:[109]

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

– Finnic **kataka* 'juniper' < **kaTaka* / **kęTaka* (where **T* = **t*, **d*, **d'* or **č*)
– Sámi N *goahcci*, Sk. *kuä´c'cev* 'conifer needle' < **koččawa* / **kaččəwa*
– Komi dial. *kač-pomel'* 'juniper' < ?**käččV* / **käčkV*
– Mansi (East) *köäsp* 'juniper' ~ (West) *käšäp* < **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 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),[110] 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. *sęsks* '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.[111] 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. 143). The Samoyed comparanda, Tym Selkup

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

111 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.

tō̈t, Kamas *ća'n* (= ⟨t́śa̡'n̥⟩, Donner/Joki 76) 'otter', are justifiably rejected by Aikio (2015a: 45): Selkup **ō̈* could perhaps reflect Proto-Samoyed **oj*, judging by Taz Selkup *tō̈ti̯-* '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 *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,[112] 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. *jäuʹrr* (< **jāvrē*); Md. E *eŕke*, M *äŕʿkä* (< **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).[113] The loss of the initial glide in Mordvin is paralleled by Md. E *ej*, M *äj* (< **jäŋə*) 'ice' and E *ezńe*, M *äźńä* 'joint' (< **jäsən*), cf. Bartens (1999: 46).[114] 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*).

112 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).

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

114 Bartens claims that the initial glide in Moksha dial. (Penza) *jäŕʿkä* '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'd'ä* 'mare' (cf. Sá. N *áldu* 'reindeer cow').

The only irregularity is the stem vowel: while Sámi and probably Mari point to **ä–ä*, Finnic unequivocally suggests **ä–ə* (Aikio 2015a: 41).[115] 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 *korị*, 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 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.[116] 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),[117] this would effectively exclude a Finnic origin.

(b) **'leather (strap)'**. F *hihna*, E dial. *ihn*, Li. (Kettunen) *nī̮ṇ* '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' (< **sęsnē*);[118] Md. E *kšna*, M *šna* (< **(šə)šna*) 'worked leather; leather strap'; Ma.

115 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).

116 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).

117 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. 84).

118 The Eastern Sámi languages (Sá. I *šišne*, Sk. *še´šnn*) reflect an irregular form **šišnē*. Accord-

E *šüštö*, W *šəštə* (< **šŭštə*) 'tanned leather (used for harnesses); harness, strap of such leather' (Thomsen 1890: 223; Kalima 1936: 101).

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'.[119] 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),[120] so that a loan from Finnic to 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).

---

ing 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'.

119 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 **šištə* could perhaps reflect PU **ćeštV vel sim*.

120 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.

(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 in 3.3.1.4, 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 *лúпа*, 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. *lǫ̃ja*) (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. *vóžā*) '(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).[121]

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

121 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 *skãrds* 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 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).

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 3.6).

3.5 Common loans from unknown sources?[122]

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),[123] 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 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.).

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

123 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.

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 5.1). 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).[124] Otherwise, Indo-European evid-

124 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 5.1, this particular branch of research was rather light on specific data.

ence 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 on p. 49, fn. 9).

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 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.

3.5.1 *Fish names*

Aside from an old inherited term for 'fish' (Lt. *žuvìs*, Lv. *zivs* = Gr. ἰχθύς, 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 **ankẹrjas*), Pr. E *angurgis* and — with divergent suffixal vocalism — R *у́горь*, Cz. *úhoř*, Sln. *ugọ́r* 'eel'.[125] Beyond this, just a couple of common Balto-Slavic terms can be cited, each having an obscure ultimate origin.[126] This situation can be explained in at least two ways. On the one

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

126 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. *līnj* 'tench' (note the mismatch in intonation!). See Pronk (2022: 270). Note my discussions of the words for

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.

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).[127] The question is whether there is any positive evidence that the word was adopted from a palaeo-Baltic source.

'ruffe' (p. 275), 'salmon' (pp. 258–259) and 'sturgeon' (pp. 218–219, 236–237), which show irregular correspondences between Baltic and Slavic.

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

Several sources have treated ON *síkr* as an inherited cognate of R *cuz* 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. *sil̃kė*, 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 awkward syncope[128] and mismatch in stem vowel, a direct equation with F *silakka* seems phonologically problematic. Most probably, Li. *siļk* is loaned from Lv. *siļ̃ķe*, 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. *sil̃kė* has arisen through syncope from **silekē* or **silikē* (Būga 1916: 143).[129] Trautmann (1910: 426) has assumed *svarabhakti* here, but there is simply no other evidence for such a phenomenon in Prussian.[130] This fact also rules out Brückner's (1877: 131) preform **sildkē* and derivation from Slavic.[131] Other etymologists

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

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

130 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.

131 The preform is itself anachronistic, as the R diminutive *селёдка* must derive from a virtual

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: 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 **sīþlō-* (Smith 1910: 141; Noreen 1923: 172).[132] In favour of the former clearly speak the early loanwords into Sámi (N *sallit*, L *sallet* 'herring' < **sęlētē*) and Slavic (R *сельдь*, Pl. *śledź* 'herring' < **silidi-*, trad. **sьlьdь*).[133]

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 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;

**silid-ikā-* or **-ukā-* (trad. **sьlьdь/ъka*), 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.

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

133 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-* (trad. **sьldь*), favoured by REW (II: 606–607), which ought to have yielded R ***солдь*, Pl. ***słudź*(?).

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).[134] 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:

– F *seipi*, E *teib* 'dace', Li. *teib* 'ide' (?< **stäipi*, *-e*-; see p. 74) ~ Lv. obs. *stiepats* 'chub', i.e. *Steepats* 'Alantsbleyer' (Lange 1773: 325; ME III: 1079) (Nuutinen 1987b)[135] — The Baltic stem **stẽp*- 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* (< **tokẹ̄h*) 'fishing weir' ~ Lt. *takišỹs*, Lv. *tacis* 'fishing weir'; Pr. E *takes* '(mill) weir' (Thomsen 1890: 226)[136] — Some connection with Lt. *tekéti* 'to flow' is often assumed (Miklosich 1886: 348; LEW 1052; PKEŽ IV: 181), but the formation has remained problematic (cf. the spec-

134 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.

135 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'.

136 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. *tākišas*); compare Pr. E *crays*, */kraiss/ 'hay' (= Pr. G *kraise* 'hay', cf, *craysewisse* 'a grain tax', on which see Chapter 7, fn. 13).

ulative 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,[137] 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*:

– 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. *łososiś*, 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 pp. 258–259). However, this evidence remains rather tenuous and open to interpretation.

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 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 conclu-

137 While in the most certain Baltic loanwords, **o* ← **a* is only found in the neighbourhood of a labial (cf. p. 63), the data is insufficient to rule out a chance correlation.

sion, 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 **ā̈*:

(a) Early	(b) Late	(c) Post-Baltic
ī ṻ ū	ī ṻ ū	ī ṻ 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 **ā̈* 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 60). 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*)[138] ~ Lt. *vė̀žỹs*, Lv. *vêzis* 'crayfish' (Thomsen 1890: 241) — The Baltic word has no clear etymology (cf. LEW 1235–1236; ALEW 1419).[139]
2. Li. *vägāli* 'burbot' ~ Lt. *vėgėlẽ* 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;

138 Estonian *-k* is a secondary excrescent consonant (cf. *mähk* 'sapwood' < **mäihä*). The *i*-stem may indicate a very recent origin (Junttila 2015a: 181), but it could also be secondary (cf. 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).

139 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).

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* (?< **apẹila* ~ **apẹlja*)[140] ~ 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 1: 235; LEW 99; ALEW 26–27). While Lv. (*d*)*âbuõls* does indeed appear to have been influenced by the word for 'apple',[141] 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 development.[142] If the word is not of Indo-European origin, the **d-* ∞ **∅-* 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 11: 115), Võ. *judõr*, (Hargla) *jutr*, Li. *ju'ddõr* 'false flax, *Camelina*'.[143]

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': cf. 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*, **varahiṇẹn*);[144] Sá. N *vuoras* 'old; old man', Sk. *vuõrâs* 'old man;

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

141 Note that e.g. Standard Latvian *âbuõliņš* 'clover' synchronically appears as if it is a diminutive of *âbuõls* 'apple'.

142 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.

143 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).

144 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

grown up (e.g. of a reindeer calf)' (< **vuoręs*) ~ Lt. obs. *voras*, Pr. III *urs*, ACC.PL. *urans* */ūrá-/ 'old' (Liukkonen 1999: 151–152) — The Baltic word is isolated. No cognates are suggested by LEW (1274), PKEŽ (IV: 211) or Smoczyński (2018: 1695).[145]

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.).[146] 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.[147] Note that if we start from Baltic **lâpetâ*, the Finnic second syllable vowel **i* is also unexpected, especially if we consider that the suffix **-etA* is frequent in Finnic, while **-itA* is otherwise unknown (cf. Koivulehto 2000: 110–111).[148,149]
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);

Sámi *vūras* 'large (of fish)'), with a subsequent development to 'timely' as in SCr. *dòspijeti* 'ripen, mature; be on time', and finally to 'early'.

145 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 **ū́rupė*?, cf. the river name *Ū̃rupiai* 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-*?

146 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*).

147 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.

148 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. *vędga* 'ice chisel'.

149 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).

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.[150]

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 *вѣко* 'eyelid', dial. (СРНГ IV: 101) 'lid of a basket or wooden vessel; grain basket', Cz. *víko* 'lid'; however, the vocalic relationship between the Slavic and Baltic words is irregular.[151,152]

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 **ē* 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.[153]

---

150 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.

151 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. 237 for a discussion.

152 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*.

153 For more length alterations, compare the examples collected in 7.5.1. 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. 89.

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

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. 65). 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'[154]
– F *ankerias*, E *angerjas*, Li. *aŋgõrz* ~ K (Olonets) *ängeriäs*, Vt. (Kukuzzi) *ängeriä*, E dial. (Vaivara) *änger*(*jas*) 'eel'

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.[155] 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

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

155 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).

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*,[156] 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. 232.

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 **atęla* rather than **atala*,[157] 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. 232), although it cannot be entirely excluded that the front vocalism in Finnic is merely secondary, as in the word for 'thrush'.[158]

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

156 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.

157 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).

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

of Lv. dial. *rẹgavas* (which is probably due to a secondary dialectal development, cf. Endzelīns 1923: 36–37) and toponymic evidence (*contra* Nieminen 1957: 202).[159]

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.

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. *tū'ontõ* (< **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 **tūšamti-* (Būga 1908: 138; Nieminen 1957: 190; Lühr 1993: 124; Liukkonen 1999: 15),[160] a reconstruction based primar-

159 A slightly different issue is posed by F *rieska*, E *rõõsk*, Li. *rõskõ* 'fresh, unleavened', which is compared to Lt. *prė́skas* in the same sense. These forms can be reconstructed to Proto-Finnic as **rę̄ska*, yet such a form would presuppose an Early Proto-Finnic **rǟska*, 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* '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. 271), 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.

160 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).

ily 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'.[161] 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 тысѧщи '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, 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-h$_1$s-nt-ih$_2$-* (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).[162]

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 6.3.2, there are possible parallels for an irregular alternation between **st* and **ts*, which might be an indication of parallel borrowing.[163] If the Indo-European word cannot be analysed as inherited, we

161 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.

162 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. *šim̃tas*; 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).

163 A somewhat similar irregularity is seen in the word for 'wax', F *vaha*, E *vaha*, Li. *vō̜*' '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 pp. 217–218), 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. 72), while

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).[164]
– 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).[165] 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 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,[166] 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

---

the development **kš* > **h* is regular in Finnic (Posti 1953: 7–9), cf. F *mehiläinen* 'bee' (< PU **mekšə*, UEW 271). Thus Germanic **wahsa-* → pre-PF **wakša* > **vaha* can be considered quite plausible (*contra* LÄGLOS III: 350 with further lit.).

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

165 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.

166 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.

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.[167] 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:

– F *tuulaalla*, Vp. dial. *tuľhuuda* (< **tūlahęla*-) '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,[168] 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*.

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

167 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".

168 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.

origin (Bartens 1999: 46). In view of Mari **tŭžem* (> E *tüžem*, W *təžem*) 'thousand', we might assume that Mordvin *'o* results from a 'breaking' from **ü* (cf. E. Itkonen 1946: 300–301; Mägiste 1959: 174–175; Keresztes 1986: 170). At first sight, a close parallel would appear to be found in Md. E dial. *śokś* 'autumn' < **sükćə*, but the initial palatal in the latter is evidently due to a secondary assimilation from the more usual form *sokś*, and cannot be associated with the palatal in the word for 'thousand'. At any rate, Mordvin *-ń* might be derivable from an earlier **-m*, which is strongly supported by the form *ťožəm*, GEN. *ťožmәń* recorded by Paasonen (MdWb 2411–2412) for the Erzya village of Seńkino.[169] 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*,[170] 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 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.

169 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 *uďem*, M dial. *uďəm* 'brain, marrow' (?< **wVdəm*; UEW 572–573).

170 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.

3.5.5 *The word for 'moss'*

Thomsen (1890: 214) compared F *sammal*, E *sammal*, Li. *sǭmal* (< **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 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' (< **se̮vmēl*)
d) Sá. K *sõvŋal* ⟨sɛuŋaл⟩ (T.I. Itkonen 1958: 487) 'hairmoss' (< **se̮vme̮l*)

The surface cluster *-vŋ-* in Eastern Sámi could reflect a number of possible preforms,[171] but *-vm-* seems to be the most suitable compromise with the Western forms. For **-vm-* > **-vŋ-*; compare Sá. S *saajmie* ~ I *sävŋi*, Sk. ⟨saŭ̄ŋ̄́ᵉ⟩, K ⟨sa̦ŭ̄ŋ̦ᵉ⟩ (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 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

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

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.

3.5.5.1 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 this subchapter 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 4, overleaf (the pre-forms correspond to the approximate time of Baltic-Finnic contacts).

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.

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- Gm. *tū(t)sn̥t-ī-	Md./Ma. *tüšäm-
'moss'	*saman-ā-	*sammal		Sá. *semol Sá. ?*siwmal

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 pp. 266–267)
– 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)[172]
– 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

172 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 **tül'ya*. 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.

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 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.

3.6 Analysis of contact relationship

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 **ẹ́hva* '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).[173] 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.[174]

173 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 1: 135–136). Note that this word is derived from the same Baltic **ešvā-* which was loaned into Finnic.

174 As another parallel for such a semantic shift, compare Ket *ku's* 'cow' as against Yugh

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 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', **paimẹn* '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).[175]

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

ku's '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).

175 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.

the domestic and wild goose diverged as 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.

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 pp. 239–240). 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 **mẽžjai* 'barley' are both of obscure origin.[176] 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.[177]

176 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.

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

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

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 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 **ateįva*, 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.[178]

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 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. *на́рма* '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 **pęlut* 'straw chaff', **nītet* 'heddle' and **rattahęt* '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

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

**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 3.3.4).

CHAPTER 4

Loanwords into Other Uralic Languages

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 *nap̣p̣e* (< **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).[1] 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 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).[2]

1 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. 50); see Pystynen 2014b.

2 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

However, not all of the etymologies originally accepted by Sammallahti are straightforward. The equation of Sá. N *gohčču̇t* 'call, order', Sk. *kåčča̮d* 'call' (< **kočču̇-*) 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).[3] Likewise, despite Nuutinen (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 жегъзоула, 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 *kǎ̆ɣʷi*) 'cuckoo' (< PU **käkə*).[4]

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*, Sk. *tue´l'lj* (< *tuoljē*) 'hide, skin rug' = 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

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.

3 According to Koivulehto **kuoit-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. *vęź-* '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; compare PF **suiccęt* > 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 (Aikio loc. cit.). As a result, there is no reason to suspect that a form **kuiccV-* should have been phonotactically impossible.

4 Compare Vakh-Vasjugan *wöɣ*, Surgut *wǎ̆ɣʷ*, *wǐ̆ɣʷ* 'strength' (< **wäkə*, UEW 563). Note, however, that Aikio (2015b: 2–3) has suggested that **ä* regularly yielded Khanty **ū* 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).

understood as parallel deverbal formations (cf. OE *sceran* 'cut, 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).[5]
– N *johtit*, Sk. *jåå´tted* (< **jotē-*) 'go, travel, migrate' ~ Lt. *judéti* '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' (< **riesęmē*) and the verb S *rïesedh*, N *riessat* (< *riesę-*) 'adorn' ~ Lt. *rìšti*, Lv. dial. *rist* 'tie (on, up)' — The semantics are not compelling.[6] Moreover, the Sámi vocalism is unexpected: **ie(–ę)* implies earlier **ä(–ə)*.
– 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. *seŕde* 'core, kernel' (Koivulehto 1990: 150) — The South Sámi form is cherry-picked. The

5 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*).

6 Thomsen (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* 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 *raśmấ* INST.SG. (or NOM.SG., Jamison *Commentary* VI.67.1) 'rein', Parth. *rsn* /rasan/ 'rope' (< **raćm̥n-o-*; Lubotsky 2001: 314), cf. Holopainen 2019: 207–208.

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'.[7]

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:[8]

- 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 *чепéц*, SCr. *čèpac* 'kind of cap', cf. REW III: 316, ALEW 552).[9]
- 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).[10]
- 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.[11]
- N *ruoida*, Sk. *ruõidd* (< **ruojtę*) 'shin, thigh' = F *reisi*, E *reis*, Võ. arch. *raiź* (< **rę̆ici*) 'thigh' ~ Lt. *ríetas*, Lv. *riẽta* 'thigh, loins; ham (of meat)' (Thomsen 1890:

---

7 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'.

8 Note, similarly, the above discussions of **luosę* 'salmon' (= PF **lohi*; p. 101), **luovē* 'raised platform' (= PF **lava*; pp. 101–102), **sę̆snē* 'tanned reindeer leather' (= PF **hihna*; pp. 87–88), **vuoręs* 'old' (~ PF **varas*; pp. 100–101). On Sámi **kuojmē* 'companion', see p. 50.

9 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.

10 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 воротитися 'return', СДРЯ 11–14 I: 477; MP *grdn* 'neck' from *grd-* 'revolve, turn', Durkin Meisterernst 2004: 163), the word **kʷekʷl-(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.

11 More plausible than the comparison with Sln. *molẹ́ti* '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 *męl* 'shingle, gravel bank' < **malō-* (but see de Vries 1962: 401).

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 *sʹjelaj* 'island' (< **suolōj*) = F *salo* 'dense forest; island; elevated spot in a swamp', E *salu* 'grove; area of woodland in an open landscape' (< **salo*)[12] ~ Lt. *salà*, Lv. *sala* 'island; elevated spot in a swamp'[13] (Thomsen 1890: 214; Kalima 1936: 158) — The Baltic word lacks a satisfactory etymology.[14]

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 3.2, five of them have a Sámi equivalent which cannot be explained as a recent loan from Finnic:[15]

– N *sarvva*, Sk. *sõrvv* (< **sʹerveʹ*) '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. *salâna* 'hoarfrost'
– N *suorri*, Sk. *sue´rr* (< **suorē*) 'branch, fork' = PF **hara* 'branch, fork' ~ Lv. *zars* 'branch, prong'

---

12 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).

13 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. *сяло́*. 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).

14 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 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).

15 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 Chapter 3, fn. 37).

– N *duovli*, Sk. *tu´vll* (< **tuovlē*) 'tinder (as a traditional remedy); tinder fungus' = PF **takla* 'tinder' ~ Lv. dial. *dagla*, *daglis* 'tinder, fire sponge' (ME I: 430, EH I: 313)
– S *vuessie*, N *vuosʹsi* (< **vuossē*) 'handle (of a cooking pot, bucket)' = PF **ansa* 'handle' ~ Lt. *ąsà*, Lv. *ùosa* 'handle, eyelet'[16]

Aikio lists **sęrvēs* (> 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 **sęrvēs* (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 **sęrvę* and **sęrvēs*. However, it appears just as probable that **sęrvēs* as an inner-Sámi derivative with the same suffix as in Sá. S *urries*, Sk. *åå´res* (< **orēs*) 'male (animal)'.[17]

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 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

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

17 The *Inarilappisches Wörterbuch* attests the form *meččin* /mečʹčin/ '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´cʹc* '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 *meččin* '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)'.

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*.[18] 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).[19]

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. *kvępi*, 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. *lȗko* '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. *vęŗga*, dial. *vędga* 'ice chisel'; Pr. E *wedigo* 'carpenter's axe'

The latter two etymologies do not involve any significant formal or semantic issues,[20] 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. 181), 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).

---

18 Whether or not they were transmitted through Finnic, it is possible that **sęrvę* 'elk' (← **širvas*) and **suolnē* (< **šalnâ*) '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).

19 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*, **d̓*, **č* > **t* in the Sámi material.

20 Admittedly, Sá. **vietkę* (< **wätkə*) is perhaps not quite expected from Baltic **vedegâ*. One might rather anticipate ***wätäkä* (> Sá. ***vātēkē* > Sá. N ***váhttit*).

Things look more positive in the case of 'soot'. Latvian *kvępi* 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,[21] the inner-Baltic etymology seems solid. In 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. 89), 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:[22]

– N *suoidni*, Sk. *sueiˊnn* (< **suojnē*) 'grass, hay' ≠ F *heinä*, E *hein*, Li. *āina* 'hay' ~ Lt. *šiẽnas*, Lv. *siens* 'hay'

---

21 Possibly here also R *кóпоть* 'soot', but the loss of **u̯* is irregular. Contrary to Schrijver (1991: 260–263) and Derksen (2015: 268), the Latvian acute hardly warrants the awkward reconstruction **kh$_2$uep-*, 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. *drẽbti* (cf. Chapter 7, fn. 78 and 79) and likewise Lv. *têst* 'carve; shave' vs. Lt. *tašýti* (cf. Chapter 3, fn. 8).

22 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 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).

While the Finnic word, partly on the basis of the Sámi evidence, has usually been reconstructed **šaina*, I have argued above extensively (see 3.3.1.4) that the word in question should be reconstructed **hẹina* 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* > **ẹ̄*.[23]

Nevertheless, it is possible to imagine an alternative account whereby pre-Sámi **šajna* was, after all, borrowed from pre-Finnic **šẹjna*. The Sámi change **i* > **ẹ* is a recent innovation (Sammallahti 1998: 106), postdating at least some of the Norse contacts (Aikio 2006a: 15) and, more importantly, the Baltic loanword **sẹrvẹ* '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 **ẹ*, meaning that the substitution **ẹ* → **a* would be perfectly intuitive.

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.

4.1.1 *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* (< **čuormẹs*) 'hail' ~ Lt. *šarmà*, Lv. *sar̂ma* '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. *щеть* 'bristle', Sln. dial. *ščệt* 'brush, thistle' (Pronk/Pronk-Tiethoff 2019: 285), the further ety-

23 Similar argumentation can be made in the case of **ruojtẹ* 'thigh' (cf. PF **rẹici*, 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 pp. 122–123.

mology 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* (< **šearęte*) 'clear (of the sky)' ~ Lt. *žėréti* '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').[24]
– Sá. S *sealma* 'threshold; pass, ridge', N *šielbmá* 'threshold (of a tent)' (< **šielmā*) ~ Lt. *šelmuõ* 'ridge (of a roof), eaves, gable' (= SCr. *sljȅme* '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 **śou̯on(i)-*, and Kallio (2009: 35) reconstructs Baltic **šāvā* on the basis of a Žemaitian form *šova* quoted in LEW (1023). The latter is evidently an untransposed Žem. *šǫva* = *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(u̯)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

24 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.

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 palaeo-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).[25]

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.

25 Reminiscent of **sāmē* is Lv. ⟨Sâhms⟩ '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.

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.[26] 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).[27]
– Md. E dial. *čonda* (?< **šondə*) 'bride price'[28] = F *hinta*, E *hind*, Li. *īnda* (< **hinta*) 'price' ~ Lt. *šiñtas* 'hundred' (Uotila 1990) — The fact that *šiñtas* (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. *inšť*), 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 *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. *penkť*), 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'.

---

26 Grünthal explicitly marks as uncertain the Baltic loan etymologies for PMd. **ärkə* 'lake' (on which see pp. 86–87), **kodər* 'twining plant stem' (see Chapter 3, fn. 4), **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.

27 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).

28 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).

The word is only known from lexical sources; cf. *bùźmas* 'eine Falte, Krause' (Kurschat 1883: 66).[29]

– 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).
– 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) *рудá* 'dirt, stain', *рудóй* 'dirty'.[30]
– Md. E *ťeŕďe-*, M *ťeŕďə-* (< **terďə-*) 'call over, invite' ~ Lt. *tìrdinti*, Lv. *tîrdît* '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.[31] 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').

---

29 In Ruhig (II: 53), we find *buźmas* 'Bauchbruch am Reße'; however, this is presumably a printing error for **bůźmas*; Mielcke (I: 31) lists the same word under *boźmas* '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).

30 But admittedly, Smolensk is geographically far removed from Mordvinia. Curious is the Russian dialectal form *рýдос* '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.

31 According to P. Kallio (p.c. March 2023), Wälchli himself is now unenthusiastic about his older proposals.

– 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'

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. *giminẽ* '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. *regéti*, Lv. *redzêt* 'see, discern', REFL. 'seem, be evident' (Wälchli 1997: 319–320; Junttila 2018: 79–80). Wälchli's opinion is that **d'* 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 **räďa-* alone is hardly enough to postulate such a feature for a hypothetical Baltic dialect.

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. *korȳs*, 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. *kŏ̈rpe* 'grain', Tat. *körpä* 'bran' ← R *кpyná* 'grain' (see p. 256). 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 *teŋ̂z*, W *taŋ̂ž* '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 pp. 248–249.

'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'[32] (Grünthal 2012: 333) — Grünthal suggests this Baltic source as an alternative to the older Turkic etymology (MdWb 2258–2259), which compared

32 The Lt. word generally refers to distance, but may also have a temporal reference (e.g. *tolì priẽš* 'long before').

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 Volga region". However, a glance in the Tatar and Chuvash dictionaries reveals that the word is indeed present there: Tat. *talaj* 'quite a lot' (TPC 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'. Md. E *kardaz* 'yard, stable', M *kaldaz* 'stable, pen' (?< **kardas*)[33] ~ Lt. *gařdas* 'enclosure, stall', Lv. dial. (SW) *gãrds* '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* (< **närgas*) 'badger', which is etymologically related to Mari E *nerɣe*, W *nerɣə* (< **nirgə*) 'badger'.[34] There are 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).

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

33 For a discussion of the Moksha *-l-*, see van Pareren (2008: 89–90).

34 The relationship between Volgaic **närkä* and Finnic **mäkrä* (> F *mäyrä*, E *mäger*, *määr*, Li. *mäggõrz*) 'badger' is unclear, but a relationship looks possible: the irregular correspondence perhaps suggests a shared substrate word.

been discussed elsewhere in this work. As a result, 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 pp. 87–88)
– Md. E *l'epe*, M *l'epä* (< **lepə*) 'alder' ~ Lt. *líepa*, Lv. *liẽpa* 'lime tree' (see p. 89)
– Md. E dial. *ťožań*, M *ťožäń* (< ?**ťožən* : **ťožəm-*) 'thousand' ~ Lt. *tū́kstantis*, Lv. *tũkstuôtis* 'thousand' (see 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. 122)

The remaining cases are unique to Mordvin, and must be discussed separately. In doing so, it is important to evaluate not only the 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.[35] The Baltic word is of unclear origin (see pp. 181–182).

'bridle'. Md. E *panct*, *panst*, M *pandəz* (< **pandəs*) 'bridle' ~ Lt. *pántis* 'hobble, fetter' (= Pr. E *panto*; OCS пѫта 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 *kalopeilis* '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 etymology; the older connections with Lt. *pielà*, R *нилá* 'saw' are abandoned in recent sources (ALEW 862; Smoczyński 2018: 954).

35 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.

'***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 pp. 261–263.

'**other**'. Md. E *l'ija*, M *l'ijä* (< **lija*)[36] '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 **lẽka-*. 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 *sâts* 'soot, coal' nor Komi *sa*, Udm. *su* 'soot' (UEW 769) represent a phonological match,[37] and they cannot be accepted as 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 *сáжа*, Sln. *sáje* 'soot' and further OE *sōt* 'soot'.[38]

'***thunder***'. Md. E *puŕgińe*, dial. *piŕgińe*, M dial. *puŕgəńä* (< **puŕgəńə*) 'thunder' ~ Lt. *perkū́nas*, Lv. *pę̀rkuôns* '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

---

36 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*.

37 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 **ćüd'ə* 'coal'; for **d*/**d'* > 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 **šüć* ~ **šü*(*j*) is the result a 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).

38 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.).

be attributed to assimilation to the Mordvin diminutive suffix *-*ə́ńə*.[39] Despite REW (II: 345–346), the Baltic words cannot be separated from OR Перунъ 'thunder god', Pl. *piorun* 'lightning', which show an irregular correspondence with Baltic, most probably pointing to a foreign origin.

? '***forest***'. Md. EM *vińŕ* (< **vińŕ*) 'forest' ~ Lv. dial. (Vidzeme) *vẽris* 'spruce forest', dial. 'riverside meadow' (Grünthal 2012: 336) — While the comparison seems attractive at first sight, the vowel substitution 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),[40] 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 *мрубá*). 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*)[41] ~ Lv. *sàls* '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 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 *so̭l*, Udm. *si̭lal* 'salt' < Proto-

39 Van Pareren posits a Baltic source **perku-*, citing Narevian *pjarkuʃ* (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 *garf* 'stork' ~ Lt. *garnỹs*.

40 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 **vińŕ* (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'.

41 The reconstruction **sal* may be preferred over **salə* in view of *Sal · Zout* in Witsen 1785; cf. Pystynen 2020b.

Permic **sǫl*), 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'.[42]

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 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 **žans(i)-* 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.[43] 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.

---

42 The development **ę(–ə)* > Komi/Udmurt *ę* is possibly regular before a resonant (Aikio 2012b: 241), while the loss of the nasal is regular. Aikio (2015a: 57) instead compares the Permic data with the Ob-Ugric words for 'mallard', reconstructing **čęčə*. Of these, Khanty **čāč* (> Vakh-Vasjugan *čač*, Kazym *šɔš*) could potentially also reflect **čęnčə* (the development **-nč#* > **-č* is not regular, but paralleled by **poč* (> Vakh-Vasjugan *poč*, Surgut *pŏč*) 'back (of the head)' < PU **ponča*). However, Mansi **št̮šə* (> West *šę̄š*, South *sās*) 'mallard' does indeed appear to rule out a nasal. While the simplification **-nš* > **-š* is regular in syllable coda (Pystynen 2020c: 256–257), the preserved long vowel in Western Mansi *šę̄š* implies a Proto-Mansi open syllable, and is thus not consistent with a nasal.

43 Incidentally, a Permic word has been taken as a loanword from the Indo-European word for 'goose': Komi *ʒǫʒęg*, (Jaźva) *ʒü·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 **ǫ* ~ Udm. **a* is extremely rare (we expect Udmurt **u*). Moreover, the Permic forms, if taken from **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-*".

† **'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 **-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 3.4 and 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 1.1.8). 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.

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,[44] 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 3.4 and 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; Вахрос 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 pp. 30–31).[45] Starting from a 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).

А. 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'[46] (usu. in the collocation *ev barq* 'house and home'; Clauson 1972: 359–360) and Yakut *bïrt* (Пекарский 625) 'wellbeing, wealth'.[47]

---

44 Most of his explanations are unsuccessful. Even his claim that Ma. E *šukerte*, W *šukerδə* 'for a long time' must be segmented *šuk-ertə* (cf. Ma. E *šuk ertak* 'for a long time', TschWb 729) and therefore not contain a cognate of F *kerta*, Md. E *kirda* 'time, -fold' (cf. also Grünthal 2012: 317) is perhaps put into doubt by the compound Ma. W *pülä-γerδə* 'quite a while ago' (TschWb 573) which would appear to imply the former existence of a word **kirdə*.

45 Nieminen and Вахрос 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 unattested in Russian. Note that R dial. (Vetluga) *перт* 'cottage' (СРНГ 26: 294; Мызников 2019: 599) is a loan from Mari.

46 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.

47 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,

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.[48] The main issue with 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. *pert'* '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 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

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. *ïrt*- 'load (onto an animal)' (Пекарский 3822) (< **ārt*-, ЭСТЯ I: 180–181), *kïrt* 'hawk' (< **kārt*-, ЭСТЯ V: 317–319); note also the derivative *sïrdā*- 'grow light' beside arch. *sarā*- 'to dawn'.

48 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' (< **mə́tə* < **medu*-) as against *śak* 'ten' (< **ćəkə* < **deḱm*). This is apparently a result of the fact that palatalized labials are generally disfavoured cross-linguistically (Ohala 1978).

attested; cf. F *kelo* 'snag; dead tree' ← Sámi **čęlę* (> Sá. S *tjalle* 'tree stump'; Aikio 2009: 77). On the other hand, it is possible that the Finnic variant with /e/ vocalism has spread from Ingrian, where **ir* > /er/ is regular, cf. Ingr. *kerves* 'axe' < **kirves* (J. Pystynen p.c. June 2023). On balance, the latter explanation appears more likely, as otherwise the origin of the Sámi word remains unclear. Even if we assume a direct Baltic → Sámi loan, it is awkward that the Sámi word does not usually mean 'bath-house'.

To summarize, Baltic **pirt(ı̇)-* 'bath-house', a word of native origin, was borrowed into ONovg. *пьрть (see pp. 30–31), whence also F *pirtti* 'house, cottage' and (perhaps via Ingrian) Karelian *pertti*. Sámi **pęrttę* is most likely from Finnic. On the other hand, there is no clear way to connect Mari *pört* 'house, cottage' to the Baltic and Finnic data, and it has an alternative Turkic etymology which seems just as promising. As a result, this word cannot be considered to offer evidence of direct Baltic loanwords in Mari.

? 'lynx'. Ma. E *šurmaŋše*, (Upša) *šŭrmŏ*, W *sârmâ*[49] (< **šŭrmə*) 'lynx' ← Lt. obs. *šermuõ* (modern *šermuonẽlis*), Lv. *sęĝrmulis* '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.[50] 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 reconstructing PU **ćirma* (cf. Ma. E *užar*, W *âžar*, Komi *vež*, Udm. *vož* < **wiša* 'green', UEW 823; Aikio 2014a: 156).[51]

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

49 On Western Mari *s-*, see Wichmann 1906: 23–25.

50 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 *sarńikᵊ* 'wolf' (< **sårmå*, Janhunen 1977: 136), on the other, do not match each other, or any of the other forms, in terms of vocalism.

51 Whether **-rm-* > **-r-* in Permic is regular is uncertain. A parallel could be Komi *jir* (< **jı̯r*, 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 *jǭr*) 'river bed', also dial. 'deep spot in water' (OstWb 400), which would presuppose a Uralic **jurə*.

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.[52]

? **'mink'**. Ma. E *šaške*, W *šäškə* (< **šäškə*) 'mink' ~ Lt. *šẽškas* 'polecat' — As discussed on p. 85, 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'.[53]

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-*.[54] 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

52 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.

53 The "Kyrgyz" (more properly Kazakh) *šeške* cited in these works stems from Ильминский 1860–1861. Since Ильминский gathered his Kazakh materials in Orenburg and Bashkiria (cf. I: 109), we are probably dealing with a localized Bashkir loanword.

54 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).

always equated with the cited Mari forms (Thomsen 1890: 237; E. Itkonen 1953: 159; SKES 1660), 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.[55] 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. *vȯ́rž* (< **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.

† **'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. 59) and Sá. N *suorri*, Sk. *sue´rr* 'branch, fork' (p. 123). 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.

---

55 Even within Mari, some of the dialects have reflexes of **u* rather than **ŭ*, e.g. Ma. W *wurðô*. A similar situation is found in the near synonym Ma. E *wurɣo*, Volga *pŭrɣo*, W *wurɣô* '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. А. Савельев (p.c. July 2023) sees in Mari **wŭrgə* ~ **pŭrgə* 'shaft' a loan from Turkic, cf. Chuvash *părăx* 'tube, pipe'.

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.

4.4 Permic

The situation with regard to Permic is even less promising than with Mari. Here I will leave aside Koivulehto's (1983: 122–127) 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*).[56] 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).[57]

56 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 **sętə* for both forms, however, we will have to explain the difference between *za*, *zu* 'stem' and Komi *vo*, dial. (Upper Sysola) *o̧* 'year', Udm. *wa-pum* 'time, period' (< **ędə*; e.g. was the lowering to Proto-Permic **å* blocked by the *w*-prothesis?).

57 As there appear to be no monosyllabic nouns in Komi *-o̧*, we might entertain a regular development of **e*(–*ə*) > Permic **o̧* in monosyllables followed by a further lowering **o̧* > **a* in Komi. This two-stage analysis is supported by the fact that two verbal stems of the shape **Co̧-* have been suggested to derive from **e*(–*ə*); viz. Komi *lo̧-* 'be, become' (?< **lexə*; cf. Metsäranta 2020: 327) and *vo̧-* 'come, arrive' (< **wexə-*; Metsäranta 2020: 146–147).

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 surely incorrect, as in other cases internal **-Cr-* has given **-rC-* in Permic.[58] As the Finnic word has been considered a Baltic loan (Lt. *tetervà* 'black grouse hen', Lv. *teteris* 'black grouse'),[59] 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. *ttara-*, *ttatara-* '(Tibetan) partridge' (?< **tataru*, cf. Khot. *pasa-* 'sheep' < **paću*).[60]

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).[61] The Selkup word seems to be 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.[62]

---

58 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).

59 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'.

60 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. *ź́ić́į* 'fox' (← Iran. **ropāća-*, cf. Parth. *rwb's* /rōbās/, Oss. I *ruvas* 'fox'; Palmér et al. 2021: 247).

61 I thank Abel Warries for helping me track down this word.

62 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.

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.

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.

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.

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’
		*šaL̂nā-	→	*halla	=	*suolnē	‘hoarfrost’
		*žara-	→	*hara	=	*suorē	‘fork’
		*šę̄na-	→	*hęina	≠	***suojnē**	‘hay’
		*širvē-	→	*hirvi	=	*sęrvę	‘elk’
		*daglā-	→	*takla	=	*tuovlē	‘tinder’

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. 36). 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 language. 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.

PART 2

Contacts with Unknown Languages

∴

CHAPTER 5

Introduction

5.1 Research History

As soon as it became apparent that the Indo-European languages were intrusive to Central Europe, the question arose as to the region's pre-Indo-European inhabitants. Schrader (1883: 161–162) admitted that a "vor- und nichtindogermanisch" lexical layer should probably be present in all Indo-European languages, yet conceded that it may never be possible to recognize it. In early research, words with a narrow geographical spread were typically explained as chain loanwords, their ultimate source being commented on only vaguely. For instance, Hehn (1870: 177–179) in treating the family of Gr. ἐρέβινθος 'chickpea', takes Lat. *ervum* 'bitter vetch' and OHG *arawīz* 'pea' as loanwords from Greek, and the latter as a "Fremdwort aus Kleinasien". Words for other cultivated plants, like rye and hemp, are similarly taken back to unspecificied "eastern sources" (e.g. Schrader/Nehring I: 440, II: 226).

A more specific hypothesis of contact with an autochthonous European population emerged in the form of the Mediterranean Substrate Theory, which became commonplace in Romance linguistics during the early 20th century (for a prehistory of the concept, see Craddock 1969: 18–22). While the theory originated in Italy, it gained traction after garnering the support of Antoine Meillet, who in an influential article (1908–1909), suggested that a number of words common to Greek and Latin may represent parallel loanwords from another source (see also Hirt 1907: 568). Although Meillet still felt his hypothesis was "nécessairement une part d'arbitraire", his implicit methodology was clear: if we can exclude cognation or a direct loanword relationship, our only option, aside from rejecting the relationship altogether, is to assume an unidentified source language.

Around the same time, another theory was developed in Northern Europe by Sigmund Feist (1910: 350).[1] Noting that a large proportion of the Germanic

1 Since the purpose of this chapter is to investigate contact with unknown non-Indo-European languages, theories of unattested Indo-European languages such as "Frühitalisch" (Haas 1960), *Alteuropäisch* (Krahe 1963; Schmid 1968) and Temematic (Holzer 1989), will remain outside of the scope of this work. Furthermore, I will not discuss theories of contact with other languages of "known" affiliation, such as the Vasconic substrate and Semitic superstrate theories of Theo Vennemann (e.g. Vennemann 2003).

lexicon had not been etymologized, he hypothesized that the Germanic people were an "autochthone rasse" that became Indo-Europeanized secondarily. He famously estimated that 30 % of the Germanic lexicon is of pre-Indo-European origin, a figure which has been much repeated (see Witczak 1996: 71 fn. 5 and 72 with lit.; Bichlmeier 2016: 319–324). It appears that Feist set the tone, as discussions of the Germanic Substrate Theory have continuously revolved around statistical measures of Indo-Europeanness with an emphasis on negative evidence (i.e. words without an etymology; see for instance Polomé 1986; Salmons 2004; Mailhammer 2008: 152–198; and the sceptical overviews in Bichlmeier 2016; Schuhmann 2016). More concretely, Hirt (1909: 69–70) drew attention to the large amount of seafaring terminology in Germanic lacking an etymology and Feist (1913: 187) noted the absence of widely-distributed fish names reconstructible for Proto-Indo-European (see also Schrader/Nehring 1: 321). As a result of these early works, *Seewörter* have remained central to the discussion of the non-IE element in the Germanic lexicon (e.g. Sausverde 1996; Witczak 1996; Schuhmann 2014).

Meanwhile, building more on the ideas of Meillet than those of Feist, the Slovene linguist Karel Oštir developed an eccentric theory which he termed "Alarodian", following the orientalist Fritz Hommel. Adducing evidence from a dizzying array of languages, he proposed an equally complex system of "uralar[odischer] Stufenwechsel", based on parallels from Uralic (1921: 24–33). Craddock's assessment of Oštir's work as "hopelessly obtuse" (1969: 32) may sound harsh, but as Oštir's theories clearly did not stand the test of time, it is arguably fair.[2] Nevertheless, Oštir holds an important position in the research history, in that he provides one of the most comprehensive catalogues of potential non-Indo-European components in Europe (including — what is relevant for our purposes — material from Balto-Slavic), as well as systematizing the alternations upon which this hypothesis was built in a way which has perhaps not been paralleled since (see Jakob forthc. a.).

The Mediterranean Substrate took a methodological step forward with Bertoldi (1932), who approached the issue with a cautious hopefulness: while admitting the impossibility of achieving absolute certainty, in Bertoldi's view, the way forward was methodological rigour.[3] Yet Bertoldi's austerity was short-

2 Oštir still gets a mention in the bibliographical notes to various Slavic etymological dictionaries (not only that of his compatriot, France Bezlaj (ESSJ), but also ЭССЯ and ESJS).

3 He warns "Ne pouvant presque jamais atteindre une certitude absolue, la nécessité à plus forte raison s'impose de ne jamais perdre de vue du moins les limites du possible" (Bertoldi 1932: 175). Contrast Schuchardt's (1922: 21) criticism of Oštir: "er gibt sich keine Rechenschaft über die Grenzen der Erkenntnismöglichkeit".

lived, as his successor, Giovanni Alessio, once again embarked on a kind of "substrate expansionism", speculating that the so-called "Mediterranean" substrate may have left traces as far afield as India (Alessio 1946: 142). With Alessio, the Mediterranean Substrate once again meets the Baltic, although mainly on the level of impressionistic equivalences between toponyms (such as Lt. *Lietuvà* 'Lithuania' ~ Gallo-Latin *Letavia*);[4] and the "Veneti" theory (cf. also Feist 1932; for a discussion of this question, see Priestly 1997). True, he also adduced some more concrete lexical evidence (e.g. Gr. κηρός ~ Lt. *korỹs* 'honeycomb'; see pp. 248–249).

The first to apply the Mediterranean Substrate theory to Slavic, at least to any great extent, was the Czech etymologist Václav Machek. While he had already shown a willingness to push the Neogrammarian boundaries in the 30s (see Machek 1934), it was in a series of articles on Czech plant names in the mid-40s (Machek 1944–1946) that he began to refer specifically to a pre-Indo-European substrate.[5] A few years later, Machek summarized his ideas on the subject (1950b), incorporating the existing views of the Italian school.[6] He emphasizes the importance of comparing entire words, rather than resorting to vague root etymologies, and refers to several kinds of irregular correspondences which could point to a foreign origin (Machek 1950b: 148–151).

He would later put his ideas to paper in a monographic treatment of plant names (1954),[7] many of which he described as non-Indo-European. However, Machek did not limit himself to plants. He also, like Feist, commented that terms for fish tended to be "undurchsichtige, isolierte Wörter" (1947: 66). In fact, in his posthumously published etymological dictionary (1968), one finds the phrase "asi „praevropské"" ("probably pre-European")[8] so often that one might even be surprised at Kiparsky's cautious optimism (1959a: 224–225; also 1975: 19), granted that the latter still considers Machek's work a "kühner Flug der Phantasie". It seems that the "Czech School" both started and ended with

4 *Letavia* is a Latinization of Old Breton *Letau*; for details, see Delamarre 2003: 204–205.

5 As Machek repeatedly stated (1944: 179; 1950b: 160; 1968: 10; Boček/Malčík 2011: 122, 304), he took his term "praevropský" from Josef Janko, who indeed did use the term significantly earlier to denote the pre-Indo-European inhabitants of Europe (e.g. Janko 1912: 140), but without reference to language.

6 This was his first publication on the subject outside of a Czech journal, and it apparently had some impact, drawing the attention of the Romance scholar Johannes Hubschmid and Indologist Manfred Mayrhofer (Boček/Malčík 2011: 303, 486).

7 According to his letters, Machek actually completed this book in 1944 (Boček/Malčík 2011: 485), although considering the relative caution of his contemporary articles on the subject, it seems likely that many of his appeals to substrate origin were added after this date.

8 This phrase also sometimes appears in Holub and Kopečný's slightly earlier 1952 dictionary.

Machek. While Havlová's views on Machek's work are largely positive (1994: 392), her own ESJS takes a considerably more sober stance on the issue, with Machek's substrate proposals more often relegated to footnotes (as also in Трубачев's ЭССЯ).

At this point, our investigation runs into yet another strand of substrate research, namely the Pre-Greek Hypothesis. Although some wider connections for supposedly Pre-Greek words had been proposed long ago (cf. Kretschmer 1896: 405),[9] Kuiper (1956: 221–225) was the first to draw direct parallels between pre-Greek and the Germanic Substrate, and in doing so approached the latter from a new angle. Kuiper's key innovation was to explain the variation in stem-final consonants often found in Germanic — such as that between voiced and voiceless stops, and between geminates and singletons — to a substrate language. However, it would take his student Edzard Furnée to produce a monographic explication of this "consonant variation" theory, albeit not in connection to Germanic (Furnée 1972).

The work of Furnée can in many respects be compared to that of Oštir, particularly the latter's later work on bird names (Oštir 1930). Furnée's monograph essentially constitutes a catalogue of consonant alternations in Pre-Greek, a mammoth task of twenty years, the results of which certainly have value in themselves (Dressler 1974: 736). However, his conclusions are marred by the frequent appeal to "exotic" comparanda like Basque, Berber and Caucasian languages (here, we often have Hubschmid to thank),[10] even though he did not further develop Kuiper's North European connections. Also like Oštir, Furnée's work was generally ignored by later research (see, for instance, the negative reception in e.g. Georgiev 1971; Dressler 1974).

The key exception was Beekes, another of Kuiper's students, who reviewed Furnée's work favourably (Beekes 1975), and cited the former systematically in his later dictionary (2010), characterizing the scholarly neglect for the author as "a major mistake in Greek scholarship" (idem: xiv). At the same time, he fundamentally disagreed with Furnée in the interpretation of these alternations. While Furnée preferred to see all the variation in the pre-Greek lexicon as the result of expressive alternations within the source language (1972: 89–90), Beekes interpreted this variation as the result of different substitutions of foreign phonemes (1975: 71; see the similar reasoning already in Kuiper 1968;

9 For an extensive bibliographical treatment of the Pre-Greek Hypothesis, I refer to Furnée (1972: 29–79).

10 For instance, following Hubschmid (FEW v: 173), Furnée (1972: 223, 285) connects Greek λάπη 'scum, phlegm' with forms in Basque and Berber, and even adds in Finnish *lampi* 'pond'.

Beekes 1969: 193–195). One can of course not help but agree that Furnée's explanation is unsatisfactory: what is gained by positing an unattested source language when all variation is nevertheless deemed "expressive"? Yet Beekes' own approach to the issue (as exemplified by Beekes 2010: xiii–xlii and *passim*; Beekes 2014) is certainly not immune to criticism, either (cf. e.g. Meissner 2013: 6–15; Garnier 2015).

In the late 80s and early 90s, the substrate theory suddenly drew a lot of attention from American scholars. Although the most prominent voice was clearly that of Edgar Polomé (for example Polomé 1986, 1990, 1992), this wave of interest was apparently sparked by Eric Hamp's article about the word for 'apple' published several years earlier (Hamp 1979). This word stood at the centre of the debate, largely revolving around the phonological features of the palaeo-European Substrate, with barely a single paper appearing on the subject that did not refer to it (cf. Markey 1989: 591; Hamp 1990: 296; Huld 1990: 398–400; Polomé 1992: 77–78; Salmons 1992: 268–271). A return to Northern Europe once again represented a return to broad theoretical discussions with little data presentation, and to a large extent, the interest appears to have waned rather quickly.

It was around the same time that Kuiper (1995) — without referring to any of the above authors — returned to the debate with a reiteration of his "consonant alternation" theory. Variations in stem-final consonantism (i.e. differences in voicing and gemination) are presented as important recurring features of European substrate words. It is after this publication that we start to see a new "Leiden school" emerge (although see already Schrijver 1991 *passim*). Kuiper's "language of the geminates", as Schrijver (2001: 420) would later christen it, has fed directly into the studies of Beekes (1996: 223–227) and Boutkan (1998, 2003a, 2003b) and the dictionary of Boutkan/Siebinga (2005).[11] The key result is perhaps not so much a methodological shift, but more a normalization of the "substrate" concept within Leiden (see also Derksen 1999, 2000; Beekes 2000).

In more recent years, several attempts have been made to formulate criteria by which substrate words might be identified. Polomé's list (1989: 54–55), paraphrased by Salmons (1992: 267, 2004: 315), formed the basis for Aikio's (2004: 8–9; 2012a: 83), while Schrijver (1997: 294–296) can be considered to have established the Dutch school of thought on the issue (cf. e.g. Lubotsky 2001: 301;

11 After Boutkan's untimely death, Siebinga continued to pursue the former's methodology as a "substrate word specialist" in the Amsterdam *Etymologisch woordenboek van het Nederlands* (see Philippa et al. 1: 13).

TABLE 6 Suggested criteria for identifying substrate borrowings

	Salmons	Schrijver	Aikio	Beekes
Absence of an etymology	✓		[✓]	✓
Limited geographical distribution		✓		
Particular semantic fields	✓	(✓)	✓	(✓)
Irregular correspondences		✓	✓	✓
Remarkable word formation	✓	✓	✓	✓
Onomastic parallels			✓	

BASED ON SALMONS (1992), SCHRIJVER (1997), AIKIO (2012A) AND BEEKES (2010)

Beekes 2010: xxiii). These attempts to formulate sets of criteria imply a fundamental recognition that the assumption of a loanword from an unknown source is better supported if it is backed up by multiple lines of evidence. Most explicit on this point was Schrijver (1997: 296): "If the IE origin of a word is rendered suspicious by a number of criteria, it is usually the cumulative evidence rather than an individual criterium that tips the balance."

However, Schrijver's most important contribution, and something which forms a great part of this work, is his identification of *recurring alternations*, most significantly the "a-prefix" (see 7.1).[12] It is Schrijver's work that can be seen as having directly inspired the more recent studies by Kroonen (2012; see also Iversen/Kroonen 2017) and Matasović (2013; 2020).

Many similarities between these lists can be observed, which are presented in Table 6, above. The differences in criteria partially derive from differences in the scope and research goals of the respective authors. For instance, Aikio's criteria are designed as a test for the presence of a substrate layer within a language overall, whereas the other authors attempted to identify characteristics applicable to individual lexemes. The absence of a compelling etymology is, of course, a prerequisite for considering a word non-inherited, and therefore this criterion is implicitly present in the methodology of all authors, and Salmons (1992: 267) is explicit that the absence of an etymology in itself is the weakest criterion. These criteria will all be explored more deeply in the following section.

12 Although I value Schrijver's methodological rigour, the extra-Indo-European comparisons he has drawn, for instance with Uralic (2001: 422–423), Hattic, Sumerian and Linear A (2018: 361–363), are rather too speculative for my taste.

5.2 Methodological Considerations

The task for much of the remainder of this monograph will be to produce a corpus of likely non-IE borrowings in and around the Baltic branch. It is therefore of vital importance to build a robust and consistent methodology to identify and reject potential data. Above, I have given an overview of the criteria identified by various authors for identifying loanwords from non-Indo-European sources. These are as follows:

– Absence of an etymology
– Limited geographical distribution
– Belonging to particular semantic fields
– Irregular correspondences
– Remarkable word formation

Not all of these criteria are equally strong, however. As discussed above, the absence of a compelling etymology, either as an inherited word or loanword, is a necessary prerequisite for a word to be considered a borrowing from a non-IE source, and this criterion need not be expressed explicitly in our methodology. A similar thing can be said of geographical distribution: words with comparanda on the eastern edge of the Indo-European language family (that is, in Indo-Aryan or Tocharian) can hardly come into question as non-IE loanwords in Baltic. While a geographically limited distribution does not prove a borrowing, as generally acknowledged (e.g. Schrijver 1997: 294), a broader distribution would essentially disprove it. Thus, geography constitutes a "negative criterion". The formulation of what constitutes "broad" must remain vague, as any strict criterion not deriving itself from the data would be circular; however, it can be stated that the broader the distribution, the less probable it is that we are dealing with a non-IE borrowing.

It also goes without saying that we cannot argue for a non-IE origin on the basis of semantics alone. Even words for local plants and animals which cannot have been known to Proto-Indo-European speakers may have native designations. A classic example is the application of the native term 'elk' to the indigenous American species *Cervus canadensis* (cf. Mallory/Adams 2006: 133). On the other hand, the chance of a word for a local species being borrowed is naturally significantly higher than for a basic vocabulary item.[13]

13 For instance, I have argued above (p. 112) and below (pp. 266–267) that Germanic, Finnic, Sámi, Baltic and Celtic have all borrowed their respective words for 'seal' from foreign sources. In addition, Russian *нéрпа* 'ringed seal' is borrowed from Finnic (cf. North Karelian *ńorppi*; REW II: 214), and numerous other Sámi words for 'seal' have been suspected to be of Palaeo-Laplandic origin (Aikio 2004: 11). This is hardly surprising, given the

Unusual word-formation has also consistently been identified as a criterion for identifying non-Indo-European loanwords; according to Salmons (1992: 267) it is perhaps the "most powerful". Yet as Schrijver (1997: 294) points out, it is often difficult to identify loanwords on the basis of affixes alone. Thus, while OCS сапогъ 'shoe, boot' does not have a compelling etymology (cf. REW II: 578; ESJS 795) and contains a relatively infrequent suffix -огъ whose Indo-European background is uncertain, it would be circular to assume the word is of substrate origin purely on the basis of this suffix. After all, such a suffix (whether ultimately borrowed or inherited) has also been applied to native roots in Slavic.[14] Similarly, Beekes has regarded οὐρά 'tail' as possibly pre-Greek in view of the "typically pre-Greek suffix" in the derivative οὐραχός 'a foetal organ; apex of the heart, etc.' (Beekes 2010: 1127). However, the suffix seems to have had some limited productivity: cf. στόμαχος 'throat, gullet' to στόμα 'mouth' (Chantraine 1933: 403), which Beekes himself accepts as Indo-European (2010: 1408).

As a result, unusual affixation and specific semantics must both be considered insufficient indications of non-Indo-European origin. However, both may be used as an additional argument where this hypothesis is supported by other evidence.

The only remaining criterion identified by multiple authors is that of irregular correspondence. It is clear that the presence of entirely plausible comparanda which do not regularly correspond to each other remains the most certain indication that a lexeme is of a non-Indo-European origin. Thus, in order to argue that a word is loaned from a non-IE source, we must identify comparanda that are both (a) plausible and (b) irregular. To this end, I have devised the following five-point test:

1. Is the data reliable?

A word which is not reliably attested cannot be used as a basis for further analysis. This much is self-evident, but the question is not often explicitly asked, and it is remarkable how often big claims are made on the basis of doubtful data. To take a random example, the reconstruction of the IE word for 'fire' as **n̥gnis* might never have happened (or at least not as early) if it were not for the alleged Old Lithuanian "*ungnis*" (cf. Pedersen 1905: 395; Walde 1910: 377; Walde/Pokorny I: 323; for the form, see Bezzenberger 1877: 42). Yet since this

semantics, yet still, Dutch *rob* and MoLG *Rubbe* have been thought to represent language-internal innovations (Philippa et al. III: 671–672; Kluge/Seebold 770).

14 Thus OR пирогъ 'fine bread' (СДРЯ II–14: 391), Slk. *piroh* 'dumpling' is apparently derived from the root of OR пиръ 'feast' (cf. Vaillant 1947: 496–497); R *острога́* 'trident', Sln. *ostŕóga* 'spur; bramble' evidently belong with R *о́стрый* 'sharp' (REW II: 287).

word occurs in Bretke only once as against dozens of examples of *ugnis* (Būga 1923b: 399), there can hardly be any doubt that it was a simple slip of the pen.

One of the fundamental flaws of Furnée's work (and of Beekes' continuation of it) is the perpetuation of doubtful forms which may represent scribal errors or late variants (for examples see Georgiev 1971; also cf. Nikolaev 2018: 2–4, 19–20). Furnée makes a point to "resurrect" forms long discarded by philologists as evidence for a particular consonant alternation; the form χέλυμνα 'tortoise', attested once in Babrius' fable *The Tortoise and the Eagle*, is now regarded a "wohl zu Unrecht angezweifelte Lesart" (Furnée 1972: 247; followed by Beekes 2010: 1623). But so long as an error is equally possible (see e.g. LSJ s.v. χέλυμνα), it is methodologically questionable to use the form as evidence. All things being equal, it is a far bolder claim that a form uniquely reflects a genuine dialect variant than to write it off as a simple error.

As a general rule, the greater the importance of a form for the validity of a hypothesis, the more I have endeavoured to check its reliability. While a time-consuming task, it is undoubtedly a fundamental requirement for any empirical investigation that the raw data used is of a good quality.

2. Do the words belong together?

A potential weak spot in any etymological equation which is not kept in check by exceptionless sound laws is that what constitutes a "similar enough" comparandum is necessarily somewhat arbitrary (Schrijver 1997: 296). However, there are a couple of constraints which may be applied here to maximize objectivity.

First, the comparisons should be semantically perfect — or at least almost perfect. An increase in semantic latitude leads to an increase in potential comparanda. If we apply the strictest semantic criteria, we essentially compare a group of synonyms in Language 1 with a corresponding group of synonyms in Language 2. In this context, the statistical significance of a potential "match" will vary depending on the level of synonymy exhibited by a particular seme. Matching terms for more specific concepts (such as 'nose' or 'oak'), where synonymy tends to be minimal, will be more significant than those in more abstract semantic domains ('strike', 'sad', etc.). Moreover, since a relaxation of semantic criteria is likely to be accompanied by an increase in semantic abstraction, any loosening of semantic requirements will cause a disproportionate increase in our corpus of potential comparanda.

Several scholars have compared the family of SCr. *lûb* 'outer bark' with that of R *лупúть* 'strip (bark)' in a substrate context (Beekes 1971, 1996: 221; Derksen 2008: 289, 2015: 296–297; Matasović 2013: 96; Šorgo 2020: 444–445). However, it should be noted that this verbal root in Slavic does not only refer to bark; cf. Slk.

dial. *lúpiť*, Bg. *лю̀пя* 'shell, peel (of nuts, fruit, eggs)'. The two words also have plausible Indo-European comparanda: the word for 'bark' perfectly matches Go. *laufs** 'foliage' (cf. OHG *louft* 'bast') and Alb. dial. *labë* 'bark',[15] and also shows Indo-European ablaut: cf. Lat. *liber* 'bark' (< **luber*, cf. Leumann 1977: 89–90) and Lv. *luba* 'linden or fir-tree bark; roof shingle, board' (ME II: 509). On the other hand, the verb has plausible Indo-Iranian comparanda in Skt. *lumpati* (MED. *lupyáte*) 'tear', Khot. *rrv-* 'remove' (Emmerick 1968: 117), MP *rb'y-* /rubāy-/ 'rob, snatch' (cf. Lv. *làupît*, Pl. *łupić* in the sense 'rob, snatch'), whose semantics do not support an original connection to 'bark'.[16] As a result, treating the two words as variants of each other would seem unwarranted, and the partial semantic convergence within Slavic can be interpreted as secondary.

A second constraint concerns what I will term "string length". Not only is it important to compare entire words rather than abstracted "roots" (Machek 1950b: 148), the more linguistic material compared, the less likely it is that the similarity is coincidental (Holzer 1989: 22–26). This may provide an answer to what Simon (forthc.) has termed the "*deus* / θεός fallacy": how can a methodology built around irregularity elevate itself above a pre-scientific collection of chance lookalikes? Although it is impossible to exclude chance entirely (as even in the traditional method), we might reduce the risk by applying a "string length" constraint. The following is what can be reconstructed for the words for 'god' based on internal evidence:

TABLE 7 Comparison of *deus* and θεός

deus	*d	*e	*i̯	*u̯
θεός	*dʰ	*e	*?	∅
	≈	=	?	≠

The correspondence between Latin **i̯* and Greek -∅- could be regular, but since Greek -∅- can equally reflect IE **s*, this can be labelled an *ambiguous cor-*

15 Çabej (1976: 307; cf. Demiraj 1997: 229) considers *labë* to be a variant of *lapë* 'flap of skin, lobe', but the assumption of irregular voicing is clearly *ad hoc* and not supported by the different meanings of the two words.

16 While this etymology is formally and semantically flawless, it is hampered by the existence of another, equally acceptable, etymology for Indo-Iranian, namely the comparison with Lat. *rumpō* 'break, burst', ON *reyfa* 'tear, rob', which is in fact the more generally accepted one (cf. IEW 870; LIV 420).

respondence. The initial consonant matches in place of articulation, but not manner, so we will call this an *irregular correspondence*. The only exact equation concerns the vowel **e*. As a result, the comparison can be expressed in the form **De*(*i*)-. As we know this comparison to be false, let us suppose that this is too little data to prove a relationship. What would be sufficient? Ideally, it would be desirable to mathematically quantify the similarities between words, but simply counting the number of correspondences could potentially give misleading results. It is important to factor in, for instance, the relative frequency of particular phonemes: as the number of possible vowels is far lower than the number of possible consonants, a correspondence in consonantism will in most cases be more significant. As a general, and somewhat arbitrary, guide, however, I would suggest that a comparison can be considered acceptable if at least three segments are equivalents *or* irregular equivalents, and of course, the more material compared, the more robust the etymology.

3. Is the correspondence irregular?

The only positive linguistic evidence for cognancy is the existence of regular sound correspondences between phonemes, and therefore the possibility of reconstructing a common proto-form. Likewise, the only positive linguistic evidence for a non-IE origin must be considered the impossibility of reconstructing a common proto-form, which in most cases presupposes the presence of *irregular* sound correspondences. Such irregularities are the most important indication that a word could be of non-IE origin. Thus, the central pillar of my methodology can be called the "principle of irregularity".

Although irregularity has often been considered a criterion for identifying substrate words, in practice, it has not always been viewed as compulsory. Particularly in the context of the Germanic Substrate Theory, the absence of a plausible etymology has often been viewed as sufficient to substantiate a hypothesis of non-Indo-European origin (see the discussion in 5.1), This is exemplified, for instance, by the work of Boutkan/Siebinga, where we frequently encounter phrases such as "[t]he word has no outer-Gmc. cognates and must be of substratum origin." (2005: 439, s.v. *wepin*). It seems clear, however, that a positive conclusion cannot be based only on negative evidence.

The European word for 'henbane', represented by R *белена́*, Cz. *blín*, Sln. *blȅn* and OE *beolone*, OS *bilina*, OHG *bilisa* 'henbane', has come up several times in discussions of possible substrate words (cf. Polomé 1990: 334–335; Philippa et al. 1: 316; Matasović 2013: 83). The same idea is also touched upon by Schrijver (1999: 25–26), before concluding that "the matter cannot be decided at present" (idem: 28). In my view, Schrijver's ambivalence is indeed justified, as all of the evidence in the relevant languages can be explained in terms of IE morpho-

logy (Derksen 2019). That does not mean that the word must necessarily be IE, but since it is circular to assume non-IE origin based only on the word's limited geography, such examples will not be considered in this work.

4. Can the irregularity be explained?
Naturally, wherever a potential irregularity is detected, it must first be excluded that we are actually dealing with a regular conditioned development. If this is not the case, then competing hypotheses are likely to involve analogy, contamination, or sporadic sound changes. Of course, such developments do occur, and ideally they should be excluded. In reality, with enough creativity, any kind of irregularity can be explained by such means, and only in exceptional cases will such an account be objectively superior to a loanword hypothesis.

Proposing a loanword from an unattested source presupposes the presence of non-IE languages in the vicinity which became extinct before being written down. The more time we assume to have passed between this supposed language death and the start of historical records, the more plausible such a claim becomes. Irregular correspondences between reconstructed proto-forms, which necessarily imply a certain time depth, are therefore more likely to point to non-IE loanwords than irregular correspondences between modern dialects. While seemingly intuitive, this has not been a major consideration in earlier works. Indeed, the methodology of Kuiper and Boutkan essentially relies on the uncritical back-projection of modern dialect forms.

Thus, Boutkan (1998: 109–110) derives Middle Dutch *dorpel*, *dreppel*, *drempel*, and *drumpel*, all meaning 'threshold', from four distinct proto-forms, assuming these to be parallel loanwords from an unattested source. Whether these are assumed to have been borrowed into the individual Dutch dialects, thus suggesting the unattested source was still spoken during the historical period, or whether they are supposed to have been borrowed already into Proto-Germanic (coincidentally all being preserved into Middle Dutch), the flaw in this reasoning is obvious: whatever the explanation for these variants, it is unlikely to exclusively involve an unattested source language.

Words of unclear derivation and unusual structure are particularly often subject to irregular "deformations" through folk etymology. This kind of development can affect both inherited words and loanwords: to take a random example, Uk. *горобе́ць*, Bel. dial. (Polesia) *шворобе́й*, *шурабе́й* (Журавлев 1980: 57) 'sparrow' irregularly continue Old East Slavic воробии (~ Pl. *wróbel*, Sln. *vrábəc* 'sparrow'), yet such distortions, belonging to the historical period, can hardly be used as evidence of borrowing from an unattested language.

In semantic domains such as bird names, one must also reckon with the influence of sound symbolism (Matasović 2020: 332–333). Irregular alterna-

tions are generally common in the domain of expressive vocabulary, cf. the voiced-voiceless pairs R *бры́згать* ~ *пры́скать* 'splash, sprinkle', Pl. *deptać* ~ R *монта́ть* 'stamp, tread', Lt. *pam̃pti* ~ dial. *bam̃bti* 'swell, grow fat' (cf. Liewehr 1956; Kiparsky 1968: 74). Similar 'expressivization' is presumably responsible for cases like Lv. *šļaka* beside *slaka* 'drop (of liquid)', MLG *slagge* 'drizzle' (*pace* Boutkan 2003b; cf. Endzelīns 1923: 137)[17] and Lt. *šmagóti* 'whip' beside *smõgti* 'strike' (see LEW 647–648 and Fraenkel 1955: 12–13 with further lit.).

An interesting case from a methodological point of view is the word for 'lip' attested in Lt. *burnà* 'mouth, face', Bg. dial. *бъ̀рна* 'lip' (only South Slavic),[18] with a variant in *p-* limited to Latvian *pur̂ns* 'face, snout'.[19] In view of the Latvian evidence, this word has been considered a loan from an unknown source (Matasović 2013: 91; Derksen 2015: 106). However, we are faced with a similar question: does this mean the word was borrowed independently into Lithuanian and Latvian? Does this mean that non-IE groups were still present in the Baltic region after the break-up of Proto-East-Baltic? A more plausible explanation was provided by Kiparsky (1968), who attributes the Latvian word to a Finnic substrate. Indeed numerous examples of voiced-voiceless pairs occur in Latvian, and examples like Lv. *pàtaga* beside Lt. *botãgas* 'whip, goad' ← MBel. *батогъ* (ГСБМ I: 202) 'cane (for punishment)' certainly do lend themselves to such an explanation (cf. Li. *pǭtõg* 'whip'; similarly ME III: 190; note also Endzelīns 1923: 183). A similar alternation is found in Lv. dial. *teĩba* 'chub, dace(?)' beside dial. (Talsi) *deĩbiņa* 'brown trout', itself a back-loan from Li. *teib* 'ide' (< PF **stäipi*, from a predecessor of Lv. obs. *stiepats* 'dace', see p. 97).

As for *pur̂ns*, no Livonian equivalent is recorded, but an actual lexical loan in Livonian is not necessary for the assumption of substrate influence on a phonological level. This particular word belongs to a category of affective and

17 Although, as Boutkan (2003b: 246) himself admits, a German influence is difficult to exclude; cf. MoHG dial. (DWb XV: 254–255) *Schlack* 'damp mass; heavy raindrop; mix of rain and snow'; cf. Lv. *ſchłahka* 'Regen und Schnee' (Ulmann 1872: 296).

18 Bg. *бъ̀рна*, dial. *бъ̀рла*, Mac. *брна* 'lip (of an animal)', SCr. *bȓnjica* 'muzzle', dial. 'ring inserted into an animal's snout', (Čak.) *brnjȕse* F.PL. 'moustache' (cf. Boryś 1977), Sln. *bŕna* 'a kind of carnival mask', see ЭССЯ III: 129–130.

19 Slovak poet. *perna* 'lip', adduced by Machek (1961: 356), was accepted with enthusiasm by ЭССЯ (III: 130; see also Derksen 2015: 106), but cannot belong here. The development **-r̥-* > *-er-* before non-palatalized consonants is limited to a narrow group of East Slovak dialects (Krajčovič 1975: 129), while *perna* is only attested in western Slovakia (cf. SSN s.v.). Furthermore, one has to assume an additional *ad hoc* irregular development **-rn-* > **-r-* to get the standard Slovak term *pera* 'lip'. A derivation of the latter directly from **pъrna*, per ЭССЯ, remains entirely fantastical.

low-status vocabulary where substrate influence is common (cf. Aikio 2009: 47).[20] In view of this, it is questionable whether the voicing alternation in this word can support the hypothesis of a non-IE source. However, such explanations are only rarely possible: the influence of a Uralic substrate may work for Latvian, but it can hardly apply to examples with a broader distribution (*contra* Schrijver 2001: 420–424; Andersen 2003: 68–71).

5. Is the irregularity paralleled?

A final and very important step brings us back to the work of Oštir and Furnée: any kind of irregular correspondence is rendered considerably stronger if it can be supported by the existence of parallels. An important distinction between my approach and that of my predecessors, however, is a focus on the geography of an irregularity.[21] Where a geographical distribution for the various reflexes can be identified, this strongly supports both the validity of this alternation, and the notion that it could reflect genuine dialectal variation in the source language. Considering that the various sub-branches of Indo-European cannot have been situated in the same time or place, we should also not expect the various non-IE substrates underlying them to be identical either (cf. Meissner 2013).

In the Leiden substrate school, irregular correspondences have usually been explained as the result of different adaptations of a foreign phoneme (Kuiper 1968; Beekes 1969: 193–195, 1975; Schrijver 1997). For instance, in discussing examples of an alternation **ai* ∞ **a* in Celtic and Germanic, Schrijver (1997: 306–307) sets up a substrate phoneme */aə/, which is essentially a compromise between the two attested reflexes. The fact that the only other possibility considered by Schrijver is that **ai* ∞ **a* could represent a "morphophonemic alternation" within the source language illustrates that he took the homogeneity of the supposed substrate language for granted. However, so long as we are dealing with parallel loanwords, it is highly improbable that the source language in both cases was identical.

---

20 For a case study of Quechua loanwords in a variety of Bolivian Spanish, see Babel 2016. Compare also the Yiddish substrate words in (American) English: *klutz* 'clumsy person', *schlep* 'haul, carry', *schmuck* 'contemptible person', and, relevant here — *schnozz* 'nose'. The role of affective words in the context of linguistic substrates is unfortunately not discussed in the recent handbook by Haspelmath and Tadmor (2009). This is partially dictated by the methodology of the *World Loanword Database*, which focuses on a fixed set of basic meanings, generally not extending to the realm of affective words.

21 Attention has been paid to the geography of irregular alternations in works attempting to prove specifically Indo-European substrates, such as Holzer (1989).

I would favour a more pragmatic approach. Since we cannot precisely identify the cause of irregular correspondences attributable to parallel borrowing, we can merely refer to the alternation itself and attempt to identify patterns in the material. A geographical distribution may favour the "dialectal" interpretation, but the reality may in fact be more complex, as we have next to no knowledge of the linguistic landscape of pre-Indo-European Europe. It is possible, for instance, that a loanword was mediated by yet another unattested language. Since these discussions will always remain on the level of speculation, they need not be pursued here any further.

Due to the potential complexity, I would not consider the absence of a geographical distribution to disprove the validity of an alternation, but it may cause us to doubt the coherence of the material. Schrijver identifies a non-IE *a*-prefix in the Germanic words **amslōn-* 'blackbird' and **arut-* 'ore'. These two examples fit together very well (see further 7.1), but it does not follow from this that all unexpected **a-*'s in Indo-European should automatically be considered related. Schrijver's further comparison of Greek (Cretan *apud* H.) ἄκαρα 'legs' with MW *gar* (PL. *garreu*) 'leg, shank' (1997: 310; 2018: 362) shows a very different geography, and it would be very risky to draw a direct parallel — this at least should not be our default assumption.

5.3 Excursus: Illegal Root Structures

Although the impossibility of reconstructing a word for Proto-Indo-European normally implies the correspondences are irregular, in a few cases, this might be implied by the root structure itself. In this small excursus, I will discuss two structural issues which could serve as additional evidence of a non-Indo-European origin in certain cases.

5.3.1 *T_D^{h} *Root*

It is generally accepted that Proto-Indo-European had a phonotactic limitation against roots containing both a voiced aspirate and a voiceless stop (e.g. Meillet 1912: 60; de Vaan 1999). Due to the merger of the voiced and voiced aspirate stops in Balto-Slavic, external evidence is sometimes required to demonstrate such a root structure, such as in the case of Lat. *fax* 'torch' ~ Lt. *žvãkė* 'candle' (whose vocalism is also problematic; cf. de Vaan 2008: 207–208 and 7.6).[22]

22 Similarly, Lat. *fracēs* 'olive pomace' and *falx* 'sickle, scythe' imply an illegal root struc-

If we do not accept a phoneme **b* (see the discussion on p. 269), it follows that no Balto-Slavic root containing both a voiceless consonant and **b* can be inherited. For instance, Pr. III *kaāubri* (for **kiāubrin*?) ACC.SG. 'thorn' has been compared with OSw. *hiupon* PL.? 'rosehip', OHG *hiufo* 'thornbush' (cf. Stang 1972: 27). If **b* is not reconstructed, the only possibility would be to reconstruct **keub^{h}-nV-* for Germanic (with Kluge's law), but **keub^{h}-* is an illegal root. In this case, however, we must concede that the Prussian form, a hapax containing at least one obvious misprint, is hardly reliable enough to use.

One possible case is the comparison between Pl. *kobuz* 'hobby, *Falco subbuteo*', USrb. *kobušk* 'red-footed falcon'[23] and ON *haukr* 'hawk, falcon', OHG *habuh* 'hawk', the first syllable of which implies **kob^{h}-*, which does indeed imply an illegal root structure. The suffix syllable is also curious. While almost all the Slavic forms continue a form **kab-ice-* (trad. **kobьcь*): Slk. *kobec*, Sln. *skóbəc*, dial. *kóbəc* 'sparrowhawk', R *кóбчик* 'red-footed falcon', this could be explained as the result of suffix replacement; compare for instance Slk. *vrabec*, Sln. *vrábəc* 'sparrow' as against the (probably older) Pl. *wróbel*.[24] On the other hand, Polish *kobuz* seems difficult to explain as secondary. In theory, the *-z* could be seen as a direct reflex of **ǵ* and be compared directly with the Germanic **-k-*, but the implied ablaut pattern **kob^{h}ouǵ-* : **kob^{h}uǵ-* does not look particularly Indo-European. As a result, even though a paper reconstruction is possible in Indo-European terms, both the root structure and suffix make it probable that we are dealing with parallel loanwords into Slavic and Germanic.

5.3.2 *Clusters of Three Consonants in Roots*

It may also be put forward that Indo-European had a constraint against roots ending in three consonants (e.g. Schmidt-Brandt 1967: 14–15; Byrd 2010: 107). Beekes, in a discussion of non-IE vocabulary, states that "a root ending in three

---

ture. See the discussion on pp. 190–191. See also the discussion of OCS кржгъ 'circle' (?< **kreng^{h}-*) on p. 249.

23 Uk. *кóбуз* (Желеховский I: 353) is poorly attested and may well be a Polonism (Berneker I: 536). According to Schuster-Šewc (579), the Sorbian word might itself be loaned from Polish. ЭССЯ (X: 92) cite a variant "**kobъzъ*" (= **kabuza-*) on the basis of the Russian dialectal hapax *кобéз* 'a kind of small falcon' (СРНГ XIII: 355; but I could not trace this form — the source given in СРНГ appears to be incorrect!) and the Polish hapax(?) ⟨kobzy⟩ INST.PL. in Mikołaj Rej (see SEJP II: 303). This data is clearly too unreliable, not to mention that ЭССЯ's reconstruction fails to account for the Russian form (and the latter could, incidentally, be **кобьзъ*).

24 "Probably older" because it is more difficult to explain as secondary. I consider the similarity to Gr. (H) ῥόβιλλος · βασιλίσκος ὄρνις coincidental.

consonants […] is very rare, but there are a few examples; so it is not a certain indication [of a] non-IE word" (Beekes 2000a: 22). The evidence for roots of this shape is indeed very slim. The following examples can be mentioned:[25,26]

– ***$b^herHǵ$-**. Lt. *béržas*, R *берёза*; ON *bjǫrk*: Skt. *bhūrjá-*, Oss. I *bærz*, D *bærzæ* 'birch'

  This example seems fairly clear, but the widely accepted link with Skt. *bhrā́jate* 'shine, beam' is only possible if the full-grade in Balto-Slavic and Germanic is secondary.[27] A zero-grade is indeed attested in Lt. dial. *bìržis* 'birch grove'.

– ?***b^hr[*e*]*uHg*-**. Lat. *fruor* 'enjoy', Go. *brukjan* 'need', OE *brūcan* 'use, enjoy; partake' (LIV 96).

  The long *-ū-* in the Lat. participle *frūctum* is not probative, as **u* would have been lengthened anyway by Lachmann's law (cf. Weiss 1994: 39–40). The Germanic *-ū-* is most probably a secondary full-grade common in class-two strong verbs (cf. Kroonen 2011b: 112–117). Thus, we can equally reconstruct **b^hreug-**.

– ***$delh_1g^h$-**. Gr. ἐν-δελεχής 'perpetual' ~ Skt. *dīrghá-*, YAv. *darəγa-*, Alb. *gjatë*, Lt. *ìlgas*, OCS *длъгъ* 'long', possibly Go. *tulgus* 'firm, sure'

  The expected full-grade **$dleh_1g^h$-* is found in Skt. *drā́ghīyas-*, YAv. *drājiiō* 'further'. Furthermore, it cannot be entirely excluded that the Greek form

---

25 In nominal roots, suffixation can often not be ruled out. Thus de Vaan (2003: 136) reconstructs ON *ǫnd* 'vestibule, entrance hall', YAv. *qiθiiā-* /anθi̯ā-/ 'door posts' as **h_2enHt-* in view of Skt. *átā-* 'door post', but we may in theory be dealing with a *t*-stem. Compare similarly Gr. σκῦτος 'leather', MW *eskit* 'boot, shoe', OHG *hūt* 'skin, hide' (< **kuH-to-*) beside Pr. E *keuto* 'skin, leather' (< **keh$_1$u-t-* / **keuH-t-*); cf. Lt. *kẽvalas* 'shell', and YAv. *vaēiti-* 'willow', Gr. οἶσος 'chaste tree; osier' (< **uoiH-t-*) beside Gr. ἰτέα, Lv. *vîtuõls*, OHG *wīda* 'willow' (< **uiH-t-*), which may be derived from the root of Lt. *výti* 'weave, twine', Lat. *vieō* 'plait, weave' (IEW 1120–1122); Skt. *yū́ṣ-*, Lat. *jūs* 'broth, sauce', Pr. E *iuse* 'soup' beside full-grade R *yxá* 'fish soup', SCr. *júha* 'soup' (< **ieuH-s-*), cf. Lt. *jáuti* 'throw together, mix', Skt. *yuváti* 'bind' < **ieuH-* (LIV 314).

26 Rejectable examples are: 1. Gr. ῥαιβός 'crooked, bandy (of legs)', Go. *wraiqs** 'crooked' (< **ureh$_2$ig^w-*), but there are plausible alternatives for Germanic (Kroonen 2013: 593); 2. OIr. *cairem*, MW *cryd* 'shoemaker' (< **kerh$_1$p-io-*?, Matasović 2009: 189–190), on which see Chapter 7, fn. 61; 3. Skt. *úpa-valhati* 'puzzle by riddles', Gr. (Hom.) ἐλεφαίρομαι 'deceive (*vel sim.*)', Lt. *vìlbinti* 'allure' (< **uelh$_1$b^h-* per LIV 678), but Skt. *-h-* from *-bh-* is exceptional (cf. Lubotsky 1995: 127–128), the appurtenance of the Greek form might be disproven by the Myc. personal name *erepa*(*i*)*ro* (Beekes 2010: 409), and the Lt. form may well be of onomatopoeic origin, cf. *ulbéti* = *vilbéti* 'warble, coo; flatter'; 4. Skt. *úrj-* 'vigour', Gr. ὀργή 'disposition; anger' do not reflect **uorHǵ-* (*pace* Beekes 1969: 241) in view of YAv. *vərəzuuaṇt-* 'invigorating', OIr. *ferg* 'anger' (< **uerǵ-*). The Sanskrit anlaut is probably regular as in Skt. *ūrdhvá-* = Gr. ὀρθός 'upright' (see van Beek 2011: 150–152).

27 Go. *bairhts* 'manifest, bright' is unlikely to belong here, but is instead to be compared with MW *berth* 'beautiful, rich' (< **b^{h}erǵ$^{(h)}$-to-*), which most probably rules out a laryngeal, and Alb. (*i*) *bardhë* 'white' (< **b^{h}orǵ$^{(h)}$-*).

directly reflects *-*dlh₁gʰ*- (see Rix 1976: 73–74; van Beek 2013: 561–563). The historical development of Gr. δολιχός 'long' is too obscure for us to base anything on it.

– ?**h₁*[*e*]*uHdʰ*-. Skt. *údhar/n-*, Lat. *ūber*, OHG *ūter*, R *вы́мя* 'udder'
  Van Beek has tentatively suggested that Gr. οὖθαρ 'udder' is regular from **uHdʰ*- (2011: 153–154, fn. 48). If so, this would leave us only the initial glide of ON *júgr* 'udder' as evidence of an e-grade. If this could be secondary, the root may be **(H)ueHdʰ*-.[28]
– ?**kehₓik*-. Lat. *caecus* 'blind; invisible', OIr. *cáech* 'blind in one eye', Go. *haihs* 'one-eyed'; Skt. *kekara-* (late) 'cross-eyed'
  Mayrhofer (KEWA I: 264) considers the appurtenance of Skt. *kekara-* "keineswegs sicher". Without it, a reconstruction **kh₂eik*- would be equally possible (Pronk 2019a: 139).
– ?*(*H*)*r*[*e*]*uHḱ*-. Skt. *rūkṣá-* 'rough, dry', ?OAv. *uruša-* 'meagre, emaciated', OE *rūh*, GEN.SG. *rūwes* (see Heidermanns 1993: 454–455) 'rough'
  A convincing explanation is not available. Compare, however, the proposed development **ur* > **ru* before a consonant (cf. Mayrhofer 1986: 161–162; Lubotsky 1994: 98–100). Could we start from a root **ureHḱ*- with zero grade **urHḱ*- > **ruHḱ*-?

Supporting evidence for a ban on roots ending in three-consonants, at least in pre-PIE, seems to be furnished by the *Schwebeablaut* in *s*-extensions to certain roots, a process which seems designed to avoid three-consonant clusters (Schindler 1970: 152; Ozoliņš 2015: 86–135):

– **h₂eug*- (Lt. *áugti*, Lat. *augeō* 'grow, increase', Go. *aukan* 'multiply') ~ **h₂ueg-s-*[29] (Skt. *vavákṣa* PF., Gr. ἀέξω, OHG *wahsan*) 'grow, increase'
– **h₂elk*- (Gr. ἀλκή 'boldness, defence', OE *ealgian* 'defend, protect') ~ **h₂lek-s-* (Skt. *rákṣati* 'protect', Gr. ἀλέξω 'ward off, assist')
– **meiḱ*- (Gr. μείγνυμι, Lt. *miẽšti*) 'mix' ~ **mieḱ-s-* (Skt. *myákṣati* 'sich festhalten; sich vereinigen'; Kümmel 2000: 388–389)

---

28 It is possible we are dealing with a compound; for instance, Garnier (2014: 149–150) has suggested a derivation involving the preverb **ud* and the verbal root **dʰeh₁*- 'to suckle' (with the 'Kortlandt effect', **ud-dʰh₁*- > **uh₁dʰh₁*-).

29 A palatovelar might be implied by Lt. *vešéti* 'grow lush, thrive'. In view of the extreme rarity of the sequence **uḰ* in reconstructed IE words, it is possible that there was a neutralization after **u* (Meillet 1894: 292–293; Kortlandt 1979a: 58). If this is the case, we would expect this word to show an alternation **h₂ueǵs-* : **h₂ugs-*, and we could assume that the latter became generalized in Indo-Iranian and the former in Baltic. On the other hand, Smoczyński (2018: 1644; cf. also p. 1617 s.v. *vãškas*) sees this word as evidence that IE **-ks-* regularly gave **š* in East Baltic, rather than **kš* as is usually assumed (cf. Stang 1966: 96).

In conclusion, although there are a couple of unresolved issues, it seems highly probable that Proto-Indo-European did indeed prohibit roots ending in three consonants. Therefore, an implied root of this shape could again be used as an argument in favour of a non-IE origin. For instance, any root in Balto-Slavic containing a diphthong root with (a) acute accentuation not attributable to Winter's law and (b) a final stop not analysable as a suffix (particularly **p*, **b*, **ś*, **ź* or **k*)[30] can be suspected to be of non-inherited origin. This applies to some of the examples discussed elsewhere in this work:

– ? Lt. *líepa*, Lv. *liẽpa*; R *лúна*, SCr. *lȋpa* 'lime tree' (< **leiHp-*); see p. 89.
– Cz. *labuť*, SCr. *làbūd* 'swan' (< **HolHbʰ-*); see pp. 176–177 and 234.
– Lt. dial. *lùnkas*, Lv. *lûks*, Pr. E *lunkan*; R *лы́ко*, SCr. *lȋko* 'bast' (< **l(u)nHk⁽ʷ⁾-*); see pp. 181–182.
– Lt. *ríešutas*, Lv. *riẽksts*, R *opѣх* 'nut' (< **(H)roiHs-*); see pp. 238–239.

It is not particularly difficult to find other potential examples. For instance Lt. *slíekas*, Lv. *sliêka*; Pr. E *slayx* 'earthworm' ~ Sw. dial. *slå*, Nw. dial. *slo* 'slow-worm', OE *slā-wyrm* (translating Latin words for various kinds of serpent) (< **slaih(w)ō-*, cf. Falk/Torp 1065; Stang 1972: 50). Yet I would hesitate to use the Balto-Slavic intonation alone as an argument to support a non-IE origin. There still remain a number of words containing an unexpected acute which is probably of non-laryngeal origin,[31] and as long as this is the case, such evidence must be treated with care. As these instances remain very few, however, we may still consider intonation as supporting evidence for non-IE origin in cases where other evidence is available.

5.4 Preliminaries

In the next three subsections, I will treat in detail all of the material which I consider to provide potential evidence for contact with pre-Baltic languages. I have restricted my material by the following criteria: (1) "pre-Baltic" words will be defined as those which are attested either in Baltic, or in both Slavic and one other "North European" branch (Germanic or Celtic); (2) the substrate proposal

30 Although there are some unambiguous examples of a deverbal suffix **-ka-* (trad. **-kъ*) in Slavic, cf. CS зна-къ 'sign' < OCS знати 'to know'; зра-къ 'sight, appearance' < зьрѣти 'to see', it does not appear that there are any reliable examples of plain *k*-suffixes of Balto-Slavic age; in any case, the examples cited here are not readily analysable as containing a suffix.

31 See for instance Lt. *tánkus* 'dense' (Chapter 3, fn. 24) and *stíebas* 'stalk, trunk', *stámbas* 'stem, stalk' (p. 60 and Chapter 3, fn. 55).

must involve some kind of irregular correspondence and be trivial semantically (i.e. it should not contradict the criteria set forth in 5.2). I have not made any attempt to discuss every substrate word proposed where I do not consider there to be sufficient evidence of substrate origin. However, I do discuss certain words which have frequently been suggested in this context, or which require a more detailed rebuttal. These examples are marked with "†" and will not contribute to any further analysis.

As my primary criteria for identifying substrate words is irregularity, and one of my goals in collecting the material is to identify geographical patterns, I have organized the lemmata according to the type of irregular alternation identified. The alternations have been organized into two main chapters — consonantism and vocalism — and each of these is divided into a number of subchapters.

Each comparison is introduced by a word in bold, which normally represents the most frequent meaning present in the comparanda. Where two lemmata are discussed with the same meaning, these are disambiguated by a number in brackets. After this, forms are adduced, with "~" demarcating the forms showing the relevant alternation. After this, I have adduced any literature in which it is suggested that the given forms are of non-Indo-European origin. Where no literature is adduced, I am not aware of any existing proposals of that nature (although the comparison itself will usually have been made in an Indo-European context). I then go on to discuss issues concerning individual branches and reject incorrect comparanda, before making a judgement as to whether the given irregularity can be viewed as evidence that the word is of non-Indo-European origin.

CHAPTER 6

Consonantism

6.1 'Nasalization', *-*VNT*- ∞ *-*VT*-

Alternations of the type *-*VNT*- ∞ *-*VT*- have been noted particularly in the Greek material, where there are numerous compelling examples: φάρυγξ ~ φάρυξ 'throat', κόχλος 'sea snail, *Murex*' ~ κόγχη 'mussel', τέρμινθος ~ (Nicander) τρέμιθος 'turpentine' (cf. Kretschmer 1896: 403; Kuiper 1956: 213–215; Furnée 1972: 275–291). Words exhibiting such alternations have been used by the above authors to support theories of language contact with "pre-Greek". This is supported by the obscure root etymology and suffixation of the relevant words.

The interpretation of such alternations has varied. In the rendition of suspected Etruscan (Fiesel 1928: 60–61) and Thracian (Schrader/Nehring II: 532) words, one has referred to 'nasal vowels' (cf. Huld 1990: 394; Kroonen 2012: 243), while in more recent literature, 'nasal insertion' has been the preferred option (see Furnée 1972: 269–270, with lit.). Kuiper (1956: 213; 1995: 68–69) suggested the term 'prenasalization' based on parallels he saw in the Munda languages. This has become the generally accepted term among Leiden scholars (see Kuiper 1956: 219–221; 1995: 68–72; Beekes 1996: 223–226; Boutkan 1998: 108–109; Schrijver 2001: 420–421). Beekes (2014: 14), albeit with hesitation, refers specifically to pre-nasalized stops.

I would rather avoid the term 'prenasalization', particularly in the narrow sense of Beekes, as in theory, other interpretations of these alternations are possible. The above accounts, whether starting from nasal vowels and prenasalized consonants, both assume that the irregularities lie in synchronic phonological features of the donor language. However, it is not certain (or even likely) that the donor language was homogenous, and it would not be far-fetched to suppose the co-existence of sister languages or dialects where one has historically undergone a loss of syllable-final nasals. As discussed above (see p. 167), I find an agnostic approach most appropriate here.

Outside of Greek, already Kretschmer (idem: 405) pointed out that Greek. ἐρέβινθος 'chickpea', and its irregular comparandum Lat. *ervum* 'bitter vetch' are quite possibly of non-IE origin. An equivalent to the Greek form without the nasal is OHG *arawīz* 'pea' (Oštir 1930: 14; Furnée 1972: 273; Kroonen 2012: 243;

Thorsø forthc.).[1] However, several more examples can be found in Northern Europe, and these will be the focus of our discussion here.

6.1.1 *Alternation between *-VNT- and Short Vowel*

▸ '**grouse**'. RCS ерѧбь (СДРЯ 11–14 III: 219) 'partridge', Uk. *оря́бок*, dial. *о́рябка*, Pl. *jarząbek* 'hazel grouse', Sln. *jerẹ̑b* 'partridge' ~ Lt. *jerubẽ* 'hazel grouse'; Lv. *iřbe*, dial. (ME II: 59) *ìerube*[2] 'partridge' (Derksen 2000, 2015: 212) — This bird name is characterized by what Derksen (2000: 80) has described as "spectacular stem variation". The large number of variants has encouraged several solutions, e.g. Endzelīns (ME I: 708–709) tries to assume reduplication with dissimilatory loss of the *r*-, but abandons the connection with Slavic. It seems clear, however, that the Baltic and Slavic data should be kept together (cf. Fraenkel 1936a: 231; ALEW 481). Andersen, in a special article on the subject, attempts to reduce the material to four basic pre-forms, which he derives from two different roots (1996b: 75; 84–85). According to him, the forms containing a nasal are derived from the *n*-stem underlying Gr. ὄρνις 'bird' with a suffix **-b*^{h}- (on this pattern of derivation, see pp. 187–188), while those lacking the nasal should, in his opinion, be compared with Nw. *jerpe*, Sw. *järpe* 'hazel grouse'.[2]

The most fundamental flaw in Andersen's account is the failure to account for the standard Lithuanian form *jerubẽ*, which shows a disyllabic stem but no nasal (Derksen 2000: 81–83). In fact, all the forms which would supposedly correspond to those in Norse probably result from syncope. Thus, Lv. *iřbe* can be explained from an older **ierube* (cf. High Latvian *ìerube*[2], ME II: 59; EH I: 537), for which a convincing parallel may be found in *ìl̃kss* 'carriage pole' as against High Latvian *ìelukši* NOM.PL. 'carriage pole' (Bezzenberger 1885: 169: ⟨ëļukschi⟩ Zvirgzdene; cf. Endzelīns 1923: 47, EH I: 528).[3] Variants in Baltic with a nasal are very rare. Juška (II: 684) cites *jerumbẽ* as a variant of *jerubẽ*. In addition to this,

---

1 Another example mentioned by Furnée is the word for 'lynx', which will be discussed in section 6.1.2.

2 These words must be derived from ON *jarpr* 'chestnut brown (usu. of hair)'. The Norse adjective corresponds to OE *eorp* 'dark, swarthy', OHG *erpfer* · *fus*[*c*]*us*. Unlike with R *рябо́й* (see fn. 7, below), it is by no means evident that the bird name is primary.

3 As already acknowledged by Andersen (1996a: 73), Bg. dial. (БЕР I: 73) *èрбица* is most likely an irregular reduction of *èребица*, and other alleged Slavic evidence is to be explained similarly (see Derksen 2000: 78). In Lithuanian, dial. *jérbė* is also most probably from *jerubẽ* (cf. the place name *Jeřbiškiai* < *Jerùbiškiai* cited in Zinkevičius 1966: 132). The Lithuanian evidence for a stem *irb*- possibly all stems ultimately from Latvian. Thus, ⟨Ýrbenis⟩ 'Viburnum' (Pabrėža 1834: 49) seems to be based on Latvian *iřbene* (cited by the author). Lt. *vìrbė* 'hazel grouse' (cf. HLv. dial. *virbe*, *vìrba*[2] 'Rebhuhn', ME IV: 603, EH II: 786), for which LKŽ provides no dialectal attestations, was perhaps popularized by Ivanauskas' *Lietuvos paukščiai* (the form is attributed to Ivanauskas in Elisonas' *Zoologijos sistematikos terminų žodynėlis*, 1920, p. 90, although

LKŽ cites only the isolated *arumbẽlė* (Palėvenė) and *vierumbẽlė* (Marcinkonys). The explanation of these forms is rather unclear, but the limitation to some isolated dialects suggests they are secondary. Perhaps, in some areas, a certain role may have been played by Polish *jarząb*(*ek*). In fact, the Žem. **jérumbė* (in dialect notation ⟨jiêrǫ̃mbę⟩), recorded in Šateikiai (*Papildymų kartoteka*) may be a direct loanword from Polish, showing /ė/ regularly for Slavic /a/ after a palatal.[4]

Another fact left unaccounted for by Andersen is the initial *je-* in Lithuanian. As this variant is concentrated in Kauniškiai dialects, and not in dialects which show *je-* < **e-*, the *j-* is most straightforwardly interpreted as original and cannot be taken back to an original **e-*.[5] While it is true that the correspondence between Lv. (dial.) (*)*ie-* and Lithuanian *je-* is not regular (cf. Derksen 2000: 78–79),[6] it still seems most parsimonious to assume that all the East Baltic data derives from a single proto-form, most probably **jerubē* (with **ęrubē* remaining a possibility).

As for Slavic, the East Slavic forms with *o-* as against *je-* elsewhere suggest a Proto-Slavic form in **e-* (Derksen 2000: 78); the variants with *ja-* attested in several Slavic languages may be secondary (*pace* Meillet/Vaillant 1933: 101; see Andersen 1996a: 74–76 for numerous parallels). Forms without an initial vowel are basically limited to East Slavic: e.g. MR *рябь* 'partridge; ?hazel grouse' (СРЯ 11–17 XXII: 281), dial. N *рябь*, *ряб* (cf. СРГК IV: 601), Bel. *рáбчык*, dial. *рабóк* 'hazel grouse'. Beside this, they are marginally attested in Slovene: *rę̑b*, *rebíca* (Caf *apud* Pleteršnik II: 412). The most likely solution is that we are dealing with instances of aphaeresis. Note that no such forms are found in West Slavic, where initial stress was generalized. As a parallel in a similar environment, compare R dial. *литóка*, *литóнья* (and variants; СРНГ XVII: 73–74) 'third stomach of ruminants' ~ Pl. *jelito* 'intestine' (see also, in particular, ДАРЯ I: No. 33).[7] It therefore seems that the Slavic words can probably be combined under a single preform **erę̄bi-* (trad. *(*j*)*ereębъ*).

I have not found its original source). Note another Latvianism attributable to Ivanauskas: *lestė* 'flounder' (= Lv. dial. *leste*, see LKŽ). The variant *ìrbė* 'hazel grouse' is only known from Šlapelis' dictionary (*apud* LKŽ).

4 A potential parallel may be Žem. dial. *munkà* 'suffering', which has been analysed as a modification of *mūkà* 'torment' (← Bel.) under the influence of Pl. *męka* 'torment' (Zinkevičius 1966: 198). The Aukštaitian variants *vẽrūbė*, *jerũbė* asserted by Būga (1923b: 402 and RR II: 537), with a long medial syllable, seem otherwise to be unattested.

5 I consider the variants with initial *ja-* and *a-* to be insignificant; cf. *ãknos* for *jẽknos* 'liver' (in Veliuona; see Juška I: 9), *ái* = *jéi* 'if' (LKŽ; see Zinkevičius 1966: 121–124). On the interchange of initial *j-* and *v-*, see Grinaveckis 1972: 74.

6 Note also Lv. *rubenis* 'black grouse', which will be discussed on p. 241.

7 The same distribution is found in the Slavic words for 'rowan' derived from the bird name

As a result, the modern dialects indeed show a great amount of variation, but the vast majority of this can be shown to be secondary. However, the second syllables of East Baltic **jerub-* and Slavic **erę̄b-* (trad. **(j)erębъ*) are not comparable in an Indo-European context, and the presence of a nasal in Slavic as against its absence in Baltic remains strong evidence of a non-IE origin.

▸ '**swan** (1)'. Pl. *łabędź*, Sln. *labǫ́d* 'swan' ~ R *лебедь*, Bg. *лѐбед*; ON *ǫlpt*, OE *ielfetu*, OHG *albiz*, *elbiz* 'swan' (Oštir 1930: 14; Machek 1968: 316; Derksen 2000: 84; Kroonen 2013: 20) — The reconstruction of a single Proto-Slavic form seems impossible, but two widespread forms can be reconstructed: (1) Cz. *labuť*, Pl. *łabędź*, SCr. *lȁbūd*, Sln. *labǫ́d* 'swan', which regularly reflect an acute **albǭ̄di-* (trad. **olbǫdъ*; or **lābǭ̄di-*, trad. **labǫdъ*);[8] and (2) R *лебедь*, Uk. *лебідь* (GEN.SG. *-едя*), CS *лебедь (attested дебель), Bg. *лѐбед* which reflect **lebedi-* (trad. **lebedъ*). The forms are almost in complementary distribution, although Pleteršnik (I: 503) cites a rather doubtful looking Sln. *lebed* from the dictionaries of Jarnik and Janežič,[9] and some other forms in South Slavic, e.g. Mac. *лабед* and SCr. obs. *lȅbūt* (RJA V: 944) seem to show a confusion between the two forms.[10] The mismatch between the second syllables **-bǭ̄d-* and **-bed-* is difficult to account for in Indo-European terms.

In Germanic, one has traditionally interpreted ON *ǫlpt*, OHG *albiz* beside OHG *elbiz*, OE *ielfetu* as reflecting two by-forms, **albut-* beside **albit-* (Noreen 1923: 151; Specht 1947: 114; IEW 30; de Vries 1962: 101; EWAhd 1033). The form **albut-* would come close to Slavic **albǭ̄di-* (trad. **olbǫdъ*), save for the nasal (cf. Meillet 1907: 377; Булаховский 1948: 118–119; Derksen 2008: 365). However, positing unmotivated by-forms is not an attractive solution. Since the *u*-umlaut in ON *ǫlpt* (GEN.SG. *alptar*) can be attributed to the analogical extension of *u*-umlaut to all feminine consonant stems (cf. Noreen 1923: 284–285; Kroonen

---

(the hazel grouse eats rowan berries in autumn; Cramp *apud* Andersen 1996b: 79; see the partial parallels adduced in Derksen 2000: 79–80 to which we may add German *Vogelbeere* 'rowan'), as well as in the word for 'mottled' in East/West Slavic: R *рябо́й*, Slk. *jarabý* 'mottled' derive from MR *рябь*, Slk. *jarabica* 'partridge' just as R *голубо́й* 'pale blue' derives from *го́лубь* 'pigeon' (cf. Andersen 1996b: 78).

8 On the final **t* in some of the reflexes, which must be secondary, see the discussion in 6.2.

9 This variant does not appear to be known dialectally (Tijmen Pronk p.c. October 2022).

10 Despite ЭССЯ (VI: 19) and Николаев (2020: 39, fn. 6; cf. Зализняк 2019: 640), it seems incorrect to take the East Slavic forms from **lebēdi-* (trad. **lebędъ*). All of the Old Russian evidence suggests **-bed-* (СДРЯ II: 13–14), as does Ukrainian *лебідь* (GEN.SG. *лебедя*). The modern Russian adjective *лебя́жий*, is by all appearances a late creation, replacing earlier *лебежий* in the 17th century (СРЯ 11–17 VIII: 183; cf. Булаховский 1968: 103). It can be considered a hypercorrection due to the widespread merger of /'a/ and /e/ in unstressed syllables (ДАРЯ I, No. 3).

2013: 26), the most straightforward solution would be to posit a *t*-stem **albet-* (slightly differently cf. Orel 2003: 13). In this case, the suffix would be more closely aligned with that of Slavic **lebedi-* (trad. **lebedь*; although note that Slavic requires **d*^{h} as against Germanic **d*).

The acute accent implied by the reflex *la-* throughout West Slavic would alone be sufficient reason to abandon the traditional comparison with Lat. *albus* 'white' (Miklosich 1886: 162; Osthoff 1898: 64–65; ЭССЯ VI: 19; and elsewhere; see Derksen 2000: 84), and when combined with the irregular alternation between **-eD-* and **-onD-* in the second syllable, the case for a loanword from a non-IE source appears very strong. For further discussion, see p. 234.

▸ **'goosefoot'**. Lt. *balánda*, Lv. dial. *baluôda* 'goosefoot, *Chenopodium*' ~ R *лебедá* 'orache, *Atriplex*', SCr. *lobòda* 'goosefoot'; ?OS *maldia*, OHG *melta* 'orache' (Mikkola 1903: 46; Machek 1947: 66–67, 1950b: 149) — The prevailing view is that the Slavic term is somehow related to the word for 'swan' (above; cf. e.g. REW II: 21–22; ЭССЯ VI: 18, XXXII: 50; Derksen 2000: 84, 2008: 366); however, as Vasmer and Derksen both admit, the alleged proto-form **albadā-* (trad. **olboda*) could not possibly yield the attested forms. Practically all of the relevant evidence points instead to **labadā-* (trad. **loboda*): cf. unambiguously Slk. *loboda*, SCr. *lobòda*, Sln. *lóboda* (in SSKJ^{2} stressed *lobóda*), Bg. *лòбода* (and further R dial. *лободá*) 'orache'. Beside this, we find a variant **lebedā-* (trad. **lebeda*): R *лебедá*, Cz. *lebeda*, Sln. *lebę́da* 'orache', Pl. *lebioda* 'goosefoot'. Some forms like SCr. *labòda* 'goosefoot' (PCA XI: 146) apparently show the secondary influence of the word for 'swan' (Derksen 2008: 366).[11]

The semantic relationship between 'swan' and 'goosefoot' is ostensibly paralleled by the English name for the plant,[12] but Mikkola (1903: 46) has instead suggested we compare Lt. *balánda*, Lv. dial. *baluôda*, assuming that Slavic **labadā-* (trad. **loboda*) was derived via metathesis from **baladā-* (due to the influence of 'swan'?). He describes this as a "Kulturwort" and additionally adduces Greek βλίτον 'purple amaranth'. Machek (1947: 66–67; 1950b: 149) mentions the same Balto-Slavic combination, but compares instead OHG *melta* 'orache' (< **maldjō-*, Kroonen 2013: 251), which I consider more promising. In this case, we have to assume an additional alternation **b* ∞ **m* (see 6.4.2). On

11 The opposite direction of influence might explain the confusing variants in Bg. dial. *лòбод* 'swan' (БЭР III: 448), Sln. obs. ⟨lobòt⟩ 'swan' (17th c.; see Pleteršnik I: 526).

12 The term seems first to be attested in the works of 16th century botanists (thus Philippa et al. II: 167 quote Dodonaeus, dated 1554; OED cite W. Turner's *Names of Herbes* from 1548). It is therefore, as stated in OED, most probably based on the form *Chenopus*, itself attributed to Pliny (see also Marzell I: 933; G. Hegi *apud* Kroll 1990: 46).

the loss of the second syllable, see below on 'oriole'. While the extra-Balto-Slavic comparanda are less certain, the comparison between Slavic **labadā-* (trad. **loboda*) and Baltic **balaⁿdā-* looks tempting, and would be another example of the alternation **-VNT-* ∞ **-V̌T-*.

▸ '**pigeon** (1)'. OCS голѧбь; R *го́лубь*, Pl. *gołąb*; Lat. *columba* 'pigeon' ~ OE *culfre*, *culufre* 'pigeon' (for refs. and a more detailed discussion, see p. 187) — Both Old English variants have been analysed as primary: Campbell (1959: 159, and already Pogatscher 1898: 98) considers *culfre* an example of syncope (cf. OE *siolfor* beside *siolufr-* 'silver'), while Hogg (1992: 231–232) treats *culufre* as an instance of vowel epenthesis. I am inclined to side with Campbell on this issue,[13] and reconstruct the preform as **kulubrō(n)-*. Skeat (1882: 146) saw this as a 'corrupted' Latin *columba*, while Pogatscher (1898: 97) suggested the source could be found in a diminutive **columbula* (cf. Old Occitan *colombla* 'dove', FEW II: 930). An alternative was suggested by Holthausen (1899), who analysed the English word as cognate to Slavic **galǭbi-* (trad. **golǫbь*).

The obvious issue with connecting the words either through borrowing or cognancy is the absence of a nasal in English.[14] At the same time, it would be unattractive to separate **kulubrō-* from Lat. *columba*. The correlation between Germanic **u* and Italic **o* is paralleled by OHG *hulis* (< **kulis-*) against MW *celyn* (< **kolisno-*) 'holly' (Kroonen 2013: 253; van Sluis et al. 2023: 216). This would be another example of an alternation involving nasals, and give support to the non-Indo-European origin of the word (see further pp. 187–189).

▸ '**oriole**'. Lt. *volungẽ* 3ª, Lv. *vãluôdze* 'oriole' ~ Pl. *wilga*, Sln. *vólga*; R *и́волга*, Bg. *авлѝга*; ME *wode-wale*, MHG *wite-wal* 'oriole' (Oštir 1930: 101; Machek 1968: 694; Derksen 2008: 216–217; Kroonen 2013: 571; Matasović 2013: 87) — The Latvian form suggests an underlying **-ang-* in the second syllable, which does not match the Lithuanian data. By way of a solution, ALEW 1469 suggests that the standard Latvian form is a hypercorrection based on a High Latvian dialect where **uo* and **ū* have merged. However, the typical development in

13 Judging by the examples provided in these sources, the epenthesis almost exclusively occurs before word-final *_RC#* (where *C* is usually a velar) or before the clusters *-ht-* or *-gd-*. In this context, the form *culufre* stands out as exceptional. Furthermore, as Hogg states, the epenthesis is typical of Northumbrian, while this form (according to the data in the *Dictionary of Old English Web Corpus*) appears to be more widespread. Compare *siolufres* GEN.SG., attested in a West Saxon source, where the vowel is old.

14 Pogatscher's solution, involving a novel sound law **-mr-* > *-fr-* has evidently not stood the test of time (see the alternative etymologies already in Holthausen 1934, s.v. *ċealortún*, *hæf-ern*, etc.). Paulus van Sluis (p.c. August 2021) pointed out to me that **kulumfrōn-* would also be a possible preform, with regular loss of **m* before **f*, although in this case the syncope would be unexpected (Campbell 1959: 49; Hogg 1992: 230).

High Latvian is in fact a chain shift, so that no merger takes place. Note in this context the South Aukštaitian hapax *ulangėlė*, cited in LKŽ, which might favour a reconstruction **-ang-* and imply that Lt. *volungẽ* is originally an East Lithuanian form.

ЭССЯ (XIII: 251–252) unites all of the Slavic forms under the reconstruction **ivilgā-* (trad. **jьvьlga*), but such a reconstruction is hardly possible, at least, for Sln. *vółga* and SCr. *vȕga* 'oriole', as initial **i-* (trad. **jь-*) is always preserved in these languages, while here no trace of the vowel can be identified. The status of Pl. *wilga* and Slk. *vlha* is less certain, as **i-* (trad. **jь-*) > ∅- is frequent here (see Derksen 2003 for the data). In any case, East Slavic clearly demands a reconstruction with **i-*, as does Bulgarian *авлѝга*, a form which is most easily explained by metathesis from CS (*)ивлъга.[15] The significance of this 'prefix' is unclear. It is hardly, with Трубачев (1972: 19–20), an irregular reduction of the prefix **iz-* (trad. **jьz-*; **z* would not be lost before **v*); neither is the parallel with R *изю́брь* 'Manchurian wapiti' watertight; see p. 242.

The Balto-Slavic comparison goes back at least to Miklosich (1865: 68; 1886: 379), but attempts to account for the relationship between the words in Indo-European terms (e.g. Mikkola 1897: 247) cannot be viewed with optimism. Moreover, treating the Baltic second syllable as a suffix (Endzelīns 1924: 123, citing the river name *Bebrunga*) does little to elucidate the relationship with Slavic.[16] As a result, some recent works have rejected the relationship altogether (Smoczyński 2018: 1693; ALEW 1469). Nevertheless, the Baltic and Slavic words are semantically identical and share a consonantal structure:

Baltic:	**v**	*â*	**l**	*an*	**g-**
Slavic:	**v**	*ĭ*	**l**	∅	**g-**

15 This CS form is attested among a list of birds in the Hexameron of John the Exarch; however, it is not entirely certain how it is to be read. The actual manuscript has "косыже · йсоїе · йвлъгъı · йжлъны · щурыже". Since the sequences ⟨йсоїе⟩ 'jays' and ⟨йжлъны⟩ 'woodpeckers' clearly both contain the word и 'and', it is natural to suspect that ⟨йвлъгъı⟩ does, too (thus Miklosich 1865: 68, and thence the CS form влъга usually encountered in the literature, e.g. ЭССЯ VIII: 251). Aitzetmüller (1958: 38), on the other hand, reads "ивлъга" here, citing a variant ⟨и ивлъга⟩ and the modern Bulgarian evidence. This theory is supported by Bg. *и́волга* attested in Геров (II: 171; a dialectal form with **lъ* > /ol/ like others recorded in Геров, e.g. *мо́лзѫ* 'to milk', *мо́рковъ* 'carrot', III: 78, 82).

16 The etymological comparison (cf. Endzelīns 1914a: 126; LEW 1273–1274; REW I: 469) with YAv. *vārənjana-*, *vārəγna-* 'a bird of prey' (cf. Sogdian *w'rγn'k*, Khwarezmian *w'rγnyk* 'falcon', Hintze 1994: 198–199) is semantically weak. Note that Endzelīns and followers operate with Bartholomae's non-specific translation 'Name eines Vogels', which might explain their enthusiasm.

As far as the nasal alternation is concerned, it is clear that this word does not behave in quite the same manner as most of the above examples. Instead of an alternation between *-*VNT*- and *-*V̌T*-, Slavic lacks the second syllable altogether. A potential parallel for this is found between Lt. *balánda* 'goosefoot' and OHG *melta* 'orache' (see above under 'goosefoot'), provided a comparison between these forms is warranted.[17] For the vocalic alternation, Oštir (1930: 22) has adduced Lat. *taxus* ~ R *muc*, Sln. *tîsa* 'yew tree' as a parallel. While the latter is probably indeed of non-IE origin (see pp. 265–266), the parallel is imperfect due to differences in vowel length. For some other potential parallels, see 7.3.1.

Endzelīns (1924: 123; similarly Machek 1950a: 49–50; Derksen 2008: 216–217; Kroonen 2013: 571, and others) compares this word with Germanic forms like ME *wode-wale*, MHG *wite-wal* 'oriole'. A trace of the velar of the Balto-Slavic forms could be found in Swiss and Bavarian dial. *Wiedewalch* (attested since the 15th century, cf. Suolahti 1909: 170). Machek (loc. cit.) suggests the Germanic reconstruction *-*walka* in order to unite the material, but the loss of *-*k*- elsewhere would be irregular. A Germanic reconstruction *-*walhō*- might just work, however. The loss of **h* in Low German and Dutch would be regular, cf. MDu., MLG *male* 'bag' (< **malhō*-, Kroonen 2013: 351). While this development is more sporadic in Middle English and High German, the simplification of the cluster may have been supported by the word's unstressed position as the second element of a compound. This would imply an additional alternation **g*$^{(h)}$ ∞ **k* (see 6.2.1).

6.1.2 *Alternation between* *-VNT- *and Long Vowel*

▸ **'lynx'**. Gr. λύγξ ~ Lt. *lúšis*, Lv. *lũsis*, Pr. E *luysis*; R *рысь*, Sln. *rîs*;[18] OE *lox*, OHG *luhs* 'lynx'; Arm. (hapax) *lusann** 'lynx'[19] (Furnée 1972: 121–122; Martirosyan 2008: 317; Kroonen 2013: 342) — Strictly speaking, the East Baltic form for 'lynx' does not rule out an older nasal, and the word could therefore be identical,

17 The comparison of Lithuanian *jerumbẽ* : *ìrbė* 'hazel-grouse' (Derksen 2015: 510) is unlikely to be valid, as both are probably secondary variants of Baltic **jerub*-. See above on this word.

18 The most convincing explanation of the Slavic *r*- is contamination with the adjective in Cz. obs. *rysý* (Kott III: 239), LSrb. obs. *rysy* 'red-haired' (see Śmieszek 1909: 408). One might argue that this adjective is itself derived from the name of the lynx, but certainly old is R *ры́жий* 'red-haired', Pl. dial. *rydzy* 'copper-red', SCr. *rîđ* 'reddish, rust-coloured' < **rȳdja*- (trad. **rydjь*), from IE **h$_1$reud^h*-; see REW II: 557–558.

19 OIr. *lug*, translated as 'lynx' by Pedersen (1909: 186), apparently mainly on the basis of the formal similarity, is doubtful. For the interpretation as 'warrior, hero', see eDIL s.v.

aside from the acute intonation, to Greek λύγξ (see LEW 392; ALEW 696).[20] However, at least Pr. E *luysis* 'lynx',[21] Elfdalian *luo* 'lynx' (cf. Kroonen 2013: 342), and the West Germanic material are inconsistent with a nasal preform (Armenian is ambiguous; cf. Martirosyan 2008: 317). This nasal has been referred to as an 'infix' (e.g. Smoczyński 2018: 734; cf. Pedersen 1909: 188), but this remains *ad hoc* as there is no generally accepted morphological process of nasal infixation in nouns. Even granted this, the acute long vowel in Balto-Slavic as against the short **u* in Germanic are still suggestive of parallel loanwords, as they preclude the reconstruction of a common proto-form (see 7.5.1).

▸ **? 'bast'.** Lt. dial. *lùnkas*, Lv. *lûks*, Pr. E *lunkan* ~ R *лы́ко*, SCr. *lȋko* 'bast'[22] — The validity of this example depends on whether the loss of the nasal in Slavic is regular. It has been suggested based on pairs such as OCS исто* (NOM.-ACC.DU. истесѣ) ~ Lt. *ìnkstas* 'kidney' and OCS вꙑкнѫти 'learn, become accustomed' ~ Lt. (*pri-*)*jùnkti* 'get used to' that high nasal vowels were denasalized in early Slavic under acute intonation (see Mikkola 1897: 246–247; Meillet 1907: 362; Arumaa 1964: 129–130; Kortlandt 1979b: 269). If the loss of the nasal can be considered regular in Slavic, then this example does not belong here.

On the other hand, this Slavic sound law is far from certain (see in detail Pronk 2013). Lt. *jùnkti* itself contains a nasal infix (there is clearly no nasal in the causative OCS оучити 'teach'). The nasal in Lt. *ìnkstas* has also been suspected to be secondary (LEW 188; Pronk 2013: 120).[23] The clearest counter-evidence

20 In this connection, the Žem. variant *lúnšis* is usually mentioned (cf. also Specht 1947: 171–172; Chantraine DELG III: 648), but this form has no etymological value, resulting from a general sporadic nasalization of high vowels before sibilants (Būga 1922: 42; Trautmann 1923: 164; Zinkevičius 1966: 196–197).

21 The Prussian form has long been problematized. Endzelīns (1943: 206) is undecided as to whether we are dealing with *i*-epenthesis or a spelling variant for /ū/ (see similarly Топоров ПЯ V: 389). Būga (1911: 41), on the other hand, read **lunsis*. An important form is Lt. dial. *luišỹs* (Bartninkai), which supports the reality of the Prussian /ui/. Here, as Trautmann (1910: 145) already surmised, we are dealing with an epenthesis of *-i-* as also found sporadically in Western Žemaitia, particularly in *ja*-stems after rounded vowels (more examples in Bezzenberger 1887: 36, 1911: 311; Endzelīns 1914b: 102; Būga 1924a: CXXI–CXXII). Note similarly Pr. E *girnoywis* (where ⟨oy⟩ probably = */ui/) 'quernstone' < **girnuwĭs*, cf. OCS жрънъı*.

22 Older sources (cf. LEW 390–391; REW II: 75) connect Pāli *luñcati* 'pull out, pluck (a bird), tear, peel' (CDIAL 642, KEWA III: 105). This must be rejected on accentological grounds. The Pāli verb, provided it is inherited, could rather be connected with Lat. *runcō* 'grub up, weed', Gr. ὀρύσσω 'dig (up)'.

23 Deriving Cz. *výheň* 'forge' from **Hngʷni-o-* (Hamp 1970a: 77; Kortlandt 1988: 388; Derksen 2008: 534) is not very satisfactory, especially since the difference in vocalism with Cz. *oheň* 'fire' is not well accounted for (compare Pronk 2013: 124–125).

is the verb Pl. *dąć* (1SG.PRES. *dmę*), SCr. dial. *dȕti* (1SG.PRES. *dmēm*), Lt. *dùmti* 'to blow' where the reconstruction of **domH-* for the Slavic infinitive (Derksen 2008: 114) is *ad hoc* (see already Meillet 1907: 366). I therefore see the word for 'bast' as a potential example of alternation between a sequence **-VNC-* and **-V̄C-*. The foreign origin of this lemma is supported by the root-final cluster **-NHK-*, implied by the acute accentuation (see 5.3.2). Note that this word has also been borrowed into Mordvin, although whether Baltic was the proximate source is uncertain (see p. 135).

▸ **? 'elm** (1)'. Lt. *vìnkšna*, Lv. *vîksna*; R *вяз* 'elm', SCr. dial. (PCA II: 459) *vêz* 'field elm' ~ OE *wice* 'wych-elm'; Alb. *vidh* 'elm' (OED³ s.v. *wych*) — The Baltic forms reflect **vînž-* + **-snā-* > **vîn(k)šnā*.[24] OE *wice* is often assumed to contain a long vowel (Holthausen 1934: 392; IEW 1177), but OED³ (s.v. *wych*) argues that forms such as *wiech* (15th c.) would show the effects of northern lengthening in an open syllable, implying an original short vowel. On the other hand, a long vowel must be reconstructed for continental Germanic, cf. MoLG (obs.?) *Wieke* (= Prussian German *Wieken* 'white elm; small-leaved lime', Frischbier II: 468), MoHG obs. *Weiche* 'elm' (< **wīkō(n)-*).

It is uncertain whether the Albanian form is consistent with a nasalized pre-form **uiṇ́-*. If Geg *ãnkth* 'incubus, nightmare' is derived from **h₂emǵʰ-* 'narrow', it would imply satemization was blocked after a nasal (Demiraj 1997: 79; de Vaan 2018: 1745). On the other hand, this etymology is uncertain, and Huld (1981: 305) has pointed out a nasalized *vĩdh* in an early 20th century grammar, which, if reliable, would align Albanian with Balto-Slavic.

There also appears to be some related Iranian data: Gorani *wiz*, Talysh *vɪzm*, *vezm* (Пирейко 1976: 46), Khunsari *vizvā*, Bakhtiari *gzəm*, Zaboli *ɣuzbe* (Henning 1963a: 71–72; Цаболов 2001: 214; В. Дыбо 2002: 469), all in the sense 'elm', are reconstructed by Henning as **u̯izu̯ă-*, i.e. a virtual **uiǵ⁽ʰ⁾-uV-*. Based on this reconstruction, the Iranian words could be cognate with the European forms, and confirm a broader distribution (cf. Polomé 1990: 334; Mallory/Adams 2006: 159). In a footnote, Henning (op. cit. 72) also admits the possibility of a reconstruction **u̯inz-*, bringing the Iranian words in line with Balto-Slavic (see again Henning 1965: 43). There are indeed potential examples of a nasal being lost

24 Since the *-k-* can be intrusive, the claim (in ALEW 1444) that the suffixation must predate the assibilation of **ǵ* seems completely gratuitous; cf. the similar comment under Lt. *añkštas* 'narrow' (ALEW 34) which is rather < **aNž-* (= OCS жзъкъ) + **-sta-* (cf. *áukštas* 'tall' < *áugti* 'to grow' and Skardžius 1941: 324–325; LEW II; Stang 1966: 108; Smoczyński 2018: 1671). The secondary nature of the velar might be proven by Zietela *vyšnẽ* 'cross beam on a sledge' (cf. Lt. dial. *vìnkšna* in the same sense), which might well stand for **vį́šnẽ*.

before a reflex of an IE palatovelar in Iranian,[25] but also exceptions.[26] The safer option, therefore, is to identify the Iranian with the Germanic forms.

A radical solution is taken by ALEW (1444–1445), who do not mention the other Indo-European comparanda, and consider the Baltic and Slavic words independent derivatives of a root **ueiǵ-* seen also in OCS вѧзати 'bind, fetter' and Lt. *výža* or *vyžà* 'bast shoe' (see already Būga 1922: 301). This can hardly be maintained, first of all, because the Slavic verbal forms show evidence of oxytone accentuation not consistent with Winter's law (В. Дыбо 2000: 388; Derksen 2008: 521; РЭС IX: 235–236). More generally, it hardly seems attractive to separate the Balto-Slavic words from the synonyms in the other languages.

OED[3] (s.v. *wych*) have suggested that the formal problems could be accounted for by assuming the word originated in a non-Indo-European substrate language. Considering the parallels collected above, this possibility should be reckoned with. The question is whether the existence of Iranian cognates would rule out a non-Indo-European loanword (cf. 'hemp' on pp. 206–207). In this connection, we can remark that the Iranian cognates are late-attested and limited to a group of West Iranian languages spoken in a relatively compact geographical area. This might suggest the word is intrusive to Iranian; on the other hand, the fact that the word has apparently undergone satemization there would imply it is indeed very ancient. The only way out would be to assume the palatalization took place in the donor language (an IE satem language?). On balance, while a non-IE origin might help to explain the nasalized forms, it is difficult to account for all of the facts convincingly.

▸ † **'nit'**. Lt. *glìnda* (< **gnìnda*?) 'nit' ~ R *гни́да*, Sln. *gnída*; OE *hnitu*, OHG *niz* 'nit' (Beekes 1969: 290; Kroonen 2012: 247; van Sluis forthc.) — Kroonen has suggested this as an example of non-Indo-European nasalization. A nasal infix is also allegedly found in Latin *lēns* (usu. PL. *lendēs*) 'nit', but this form has so little in common with the other cognates (only the *-d-* poses no issue) that it is uncertain it belongs here (cf. van Sluis forthc.). Puhvel (1990: 366) posits a com-

25 As discussed by Martin Kümmel at the 2021 *Österreichische Linguistik-Tagung*. A partial parallel is the word for 'twenty', whereby against the remarkable parallelism of Oss. D *insæj* and Skt. *viṃśatí* 'twenty' (cf. Henning 1965: 43), the rest of Iranian shows **ī* (YAv. *vīsaiti*, MP *wyst* /wīst/). However, the vowel turns up long here. As another possible example, note Parth. *bz-* 'receive help' as against YAv. *bązaiti* 'support' (cf. Cheung 2007: 72; however, Khot. *baś-* (Emmerick 1968: 94), Oss. ID *bæzz-* 'be suitable' reflect **bazi̯a-* with probable zero-grade).

26 Most notably MP *hnzwg* 'narrow' (→ Arm. *anjuk*), hardly to be separated from Skt. *aṃhú-* 'narrow' (Henning 1963b: 196–197).

mon preform **lind-* (< **nind-*) for Lat. *lēns* and Lt. *glìnda*, assuming the initial guttural in the other languages is "moveable" (i.e. of secondary origin), which is clearly *ad hoc*. Note that even **lind-* does not explain the Latin vocalism (de Vaan 2008: 334).

As for the Lithuanian form, one is reminded of cases of secondary *-ninC-* < **-nīC-*, which are particularly common in Žemaitian dialects: cf. *bažnìnčia* < *bažnýčia* 'church' (← Bel. *бажні́ца*), dial. *kningà* < *knygà* 'book' (← Bel. *кні́га*). The main issue here is that it is precisely in Žemaitian where we actually find a form without a nasal: dial. *gnýda*.[27] However, this is not fatal, as such nasalized forms are only sporadic in Žemaitian. Furthermore, forms with a secondary nasal are also occasionally recorded in Aukštaitian; note in particular the agent-noun suffix *-ininkas* (beside dial. *-inykas*), which even belongs to the literary standard.

The ablaut relationship between OE *hnitu*, OHG *niz* 'nit' (< **knid-*) and Alb. *thëri*, Gr. κονίς, PL. κονίδες 'nit' (< **konid-*) looks highly archaic, and is easier to explain in an IE context than through independent borrowings.[28] It seems impossible to get away from the notion of taboo distortions here (cf. IEW 608): at least the initial *gn-* in Balto-Slavic must be explained in this way;[29] in this context, we can note that many Slavic words starting in **gn-* have a negative connotation, e.g. R *гнус* 'gnats', SCr. *gnôj* 'manure, pus'. It is possible that taboo also played a part in the replacement of earlier (*)*gnýda* with *glìnda* in Lithuanian. On balance, due to the many difficulties with this word and its alleged existence in almost every European branch, I will leave it out of consideration here.

6.1.2.1 Conclusion

The certain and possible evidence for nasal alternations is collected in Table 8, overleaf. The forms are presented as approximate quasi-Indo-European recon-

27 See LKŽ, where the word is marked as a Polish loanword. While this loan etymology cannot be ruled out, there is nothing in particular to suggest that the Žemaitian form is not simply regularly cognate with the Slavic forms. Latvian *gnīda* 'nit' is of course ambiguous, and could reflect a preform with or without a nasal, or also be loaned from Slavic.

28 The ablaut **konid-* : **knid-* seems to belong to a rather rare type, but compare **melit-* (Hitt. *milit* NOM.SG., Gr. μέλι, Go. *miliþ*, Alb. *mjaltë* 'honey') : **mlit-* (Hitt. *maliddu-* 'sweet, pleasant', Gr. βλίττω 'cut out honeycomb', ?Alb. (*m*)*bletë* 'honeybee'). This is not the place to go into a discussion of Armenian *anic* (for **nic* < **knid-s*?, Martirosyan 2008: 86–89) and Celtic forms pointing to **snida* (why **s-*?), although they may somehow belong here.

29 Also note the voiced anlaut of Lt. *blusà*, R *блоха́* 'flea' as opposed to Skt. *plúṣi-* 'flea'.

TABLE 8 Possible examples of nasal alternations

	Baltic	Slavic	Germanic	Elsewhere
'grouse'	*i̯erub^{h}-	*i̯eremb^{h}-	? *reb^{h}-	
'swan (1)'	–	*alband^{h}-	*alb^{h}ed-	
		*leb^{h}ed^{h}		
'goosefoot'	*b^{h}aland-	*lab^{h}ad^{h}-	*malT-	
'pigeon (1)'	–	*Golomb^{h}-	*guluBr-	Lat. *kolomb^{h}-
'oriole'	*u̯âlanG-	*u̯(i)lg$^{(w)}$-	*u̯alk-	
'lynx'	*lû(n)ḱ-	*rûḱ-	*luḱ-	Gr. *lunḱ-
? 'elm (1)'	*u̯inǵ-sn-	*u̯inǵ-	*u̯(e)iǵ-	Iran. *u̯iǵ-u̯-
				Alb. *u̯i(n)ǵ-
? 'bast'	*lûnk-	*lûk-	–	

structions, but without the use of laryngeals. Long vowels which turn up as acute in Balto-Slavic are written with the caret ⟨ ˆ ⟩. Where the Indo-European reconstruction is ambiguous, cover symbols are used (e.g. *G in Slavic = *$g^{(h)}$ or *$g^{w(h)}$). Forms containing a nasal are presented in shaded cells. Where the presence of a nasal is ambiguous, the cell is shaded in a lighter grey.

Several bird names occurring in Baltic, Slavic and Germanic show a conspicuous alternation in the suffixal syllable. It seems quite probable that these can be attributed to a related source. All of them show a morphologically similar structure involving a second syllable of the shape *VND alternating with *$V\check{D}$. The distribution is fairly consistent, with the nasal being absent in Germanic, and Baltic and Slavic adopting an intermediate position.

A couple of other European bird names can be noted with a similar structure, where irregularities also support the notion of borrowing. First, there is Lat. *hirundō* 'swallow', which should not be separated from Gr. χελῑδών 'swallow' (cf. Chantraine DELG IV: 1253), or from Alb. *dallëndyshe* 'swallow' (cf. Meyer

1891: 59),[30] although they cannot go back to a common proto-form. Here, we find both a disagreement in terms of vocalism and between *-r-* and *-l-* (see also fn. 37, below). The variant without the nasal in Gr. χελῑδών strongly recalls the similar phenomenon in our northern European bird names. Another bird name with a similar structure is Lt. *balañdis* 'pigeon' (?~ Lat. *palumbēs* 'wood pigeon'); for a detailed discussion of this word, see pp. 209–210.

The word for 'lynx' is different in that the variant without a nasal occurs in the initial syllable and alternates with a long, acute vowel in Balto-Slavic. Although the word for 'bast' is superficially similar, as the nasal in the word for 'lynx' may be due to a phenomenon peculiar to (pre-)Greek, and the presence of a nasal in the Baltic word for 'lynx' is doubtful, there is no reason to suppose that these two words belong to the same loanword stratum.

6.2 Voicing Alternations

Based on examples such as Lat. *habeō* (< **g^{h}-*) ~ Go. *haban* (< **k-*) 'to have', the existence of *Konsonantenwechsel* or alternations between different consonant series in the Indo-European proto-language has repeatedly been suggested (e.g. Zupitza 1904: 387–391; Hirt 1927: 297–303; Machek 1934: 7–36; Otrębski 1939: 156–171). These proposals can be seen as reactions against rigid Neogrammarianism, with alternations invoked as an unexplained "mysterious force" awaiting later elucidation. Despite this, some of the comparisons were so tantalizing that the notion has not disappeared from the literature. Yet as the mechanism behind this alleged phenomenon has never properly been explained, it has never quite entered the mainstream, and remains incompatible with a strict application of the comparative method. Comparanda such as those collected by the above authors have also inspired other theories. Both Haas (1960) and Holzer (1989) have assumed the existence of a lost Indo-European language, which has undergone a consonant shift, underlying Latin and Slavic, respectively. While this remains a theoretical possibility, the heavy reliance on root etymologies, many of which often do not fare better than the traditional solutions (Аникин 1992 and in particular Matasović 2013: 77–82), has meant they have had little resonance among comparativists.

30 Alb. *d-* regularly corresponds to Lat. *h-* (Alb. *dimër* ~ Lat. *hiems* 'winter'). It must be admitted that the alternative comparison with the Illyrian tribal name Ταυλάντιοι, reported by Hecateus of Miletus to have neighboured the Χελιδόνιοι(!), is tempting (Çabej 1976: 105–106).

As noted above in 5.1 (see point 4), irregular voicing alternations are relatively frequent in words of an expressive character. Other alternations, particularly in final position, might be explained as the result of *sandhi* phenomena. Such an account is probably necessary for cases like Cz. *labuť* as against Pl. *łabędź* 'swan' (see above on pp. 176–177). The alternation between OCS (Euch.) дрѫгъ* 'club, cudgel', Slk. dial. *drúh* 'thick branch', SCr. dial. *drûg* 'pole, long sick', against the dialectal variants Slk. *drúk*, SCr. *druk* (cf. RJA II: 807) has been explained by positing a substrate origin (e.g. Derksen 2008: 121; Matasović 2013: 83–84), but given the existence of both variants side by side in the individual languages, this would imply the existence of non-IE groups in Europe practically until the modern period. Considering the improbability of this scenario, we are better off seeking an irregular motivation such as contamination or expressivization (REW I: 374; Liewehr 1956: 20; the Serbo-Croatian form might well originate in a dialect with word-final devoicing, T. Pronk p.c. March 2023). Despite this caveat, there are numerous examples of voicing alternations which, in my view, constitute plausible evidence for non-Indo-European origin. The examples below are organized into five groups based on the consonants involved.

6.2.1 *Baltic* *g$^{(h)}$ ∞ *k *Elsewhere*

▸ **'pigeon** (1)'. OCS голѫбь; OE *culfre*, *culufre* 'pigeon' ~ Lat. *columba* 'pigeon' (Oštir 1921: 49, 1930: 39; Machek 1951a: 103–104; Treimer 1954: 70; Machek 1968: 175; Bezlaj I [1977]: 159; Kleyner 2015: 53–54; ERHJ I [2016]: 284) — For the Old English word, see the discussion on p. 178. The identity of the Slavic and Latin words has long been recognized (already Bopp 1833: 336), but as the comparison is clearly irregular, it is generally rejected, having already been omitted from the fourth edition of Fick's comparative dictionary (Stokes 1894: 92). Nevertheless, the similarity of the words has remained obvious. Leaving aside the *ad hoc* notion that the Slavic **g-* is simply secondary (Shevelov 1964: 365; Lockwood 1990: 262), this word has been used to bolster theories of contact with unidentified Indo-European languages (cf. Haas 1960: 34; Holzer 1989: 161–162). Соболевскій (1914: 441) proposed that an unknown language had mediated a Latin loanword, while Szemerényi (1967: 20–21) insists on a Latin origin; however, only on the basis that the word cannot be explained within Slavic.

What unites all these theories is the assumption that Lat. *columba* is inherited, of which there is no solid indication. Morphologically, the word is isolated in Latin, aside from the near synonym *palumbēs* 'wood pigeon'. The traditional analysis sees these words as containing a compound suffix **-n-b^{h}-* (Brugmann 1906: 386, Meillet/Vaillant 1933; ЭССЯ VI: 216; Sławski SP VIII: 46; Аникин

РЭС XI: 146), in which case Lat. *columba* would be derived from an underlying *n*-stem continued by Gr. κελαινός 'black, dark' (Prellwitz 1897: 102–104; Persson 1912: 169–171; IEW 547–548; Batisti 2021: 206–207).[31] For Slavic, Derksen (2008: 175) comments that "the suffix *-(*V*)*mbʰ*- is frequent in bird-names", noting the parallel in Slavic **erębi*- (trad. *(*j*)*erębь*) 'grouse' (on which see pp. 174–176).[32]

This morphological analysis is based primarily on the co-occurrence of Skt. *vṛ́ṣan*- and *vr̥ṣabhá*-, both appearing in the senses 'manly' and as a substantive 'bull'. Except for the synonym *r̥ṣabhá*- 'bull' (belonging with Av. *aršan*- 'man, male'), other examples of this pattern of derivation within Indo-Aryan are quite uncertain.[33] A close parallel to *vṛ́ṣan*- beside *vr̥ṣabhá*- is nevertheless found in Gr. ἔλαφος 'deer', which beside OCS елень 'deer', Arm. *eln* 'doe, hind', has traditionally been segmented ἔλ-α-φος (e.g. Prellwitz 1897: 100; Osthoff 1901: 305–308; Chantraine 1933: 263; Beekes 2008: 402).

Despite this, other examples of this supposed compound suffix *-*n*-*bʰ*- are sparse, and appear to be limited to European bird names: Lat. *palumbēs* 'wood pigeon' and Arm. *salamb* 'francolin'.[34] While some productive suffixes in individual branches contain reflexes of **bʰ*, their semantic function is not aligned (e.g. the Gr. diminutive -άφιον, the deadjectival OCS зълоба 'evil' < зълъ 'bad, wicked' and the Gothic adverbial suffix -*ba*). It therefore remains uncertain whether a suffix **bʰ* can be reconstructed (most of the evidence adduced in Hyllested 2009: 202–205 is open to interpretation).

Despite the potential derivational parallel in the words for 'deer', the separation of Lat. *columba* from OCS голѫбь feels artificial: the words mean exactly the same thing, and aside from the voicing of the initial stop, show an identical stem. It is *a priori* questionable that two branches would have used the same inherited suffix only in words for 'pigeon', and have independently innovated a word for 'pigeon' which happens to be virtually identical. The invalidity of the traditional morphological analysis would seem to be confirmed by other

31 See Batisti (2021: 207, fn. 4 with lit.) for other root etymologies, none of which are any more convincing. As for κόλυμβος 'grebe', I fully agree with Batisti that the word should be kept separate.

32 Walde/Hofmann (I: 249) insist that the Slavic word must be native because of the colour term in R *голубо́й* 'light blue', but this is rather a derivative of the word for 'pigeon' (Loewenthal 1901: 31–32; Machek 1951b: 103; Herne 1954: 91).

33 Skt. *śarabhá*-, a kind of game animal, continued in Dardic and Nuristani in the senses 'markhor, ibex, mountain goat' (CDIAL 714) is supposedly connected to Lat. *cornū* 'horn' (EWA II: 616; Nussbaum 1986: 6), but this is far from certain. Two words for 'donkey' — *rāsabhá*- and *gardabhá*- — are not well explained; the latter is probably not of Indo-European origin (EWA I: 473; cf. Pinault 2008: 393–394).

34 Nothing can be said of Gr. κόραφος (H.), an unidentified bird.

irregularities, such as the missing nasal in the Old English form (see p. 178). It therefore seems entirely reasonable to explore the possibility of a non-Indo-European origin.

Very curious is the Coptic comparandum adduced already by Oštir (1921: 49), cf. Sahidic ϭⲣⲟⲟⲙⲡⲉ, Bohairic ϭⲣⲟⲙⲡⲓ, Lycopolitan ϭⲣⲁⲙⲡⲉ 'pigeon', deriving from a Late Egyptian (~ 12th c. BCE) form  *gr-(n)-p.t* */kʰV̆rámpV̆/ (cf. Allen 2020: 115).[35] This form is written as *'gr*-bird of the sky', and as a result has been viewed as a native formation by Egyptologists. Peust (1999: 280) has suggested that the Egyptian form may be the source of the Indo-European words (similarly Иванов 2002). On the other hand, Vycichl (1990: 249) has argued that the Egyptian spelling is folk-etymological ("la colombe n'est pas un «oiseau du ciel» comme l'aigle ou le faucon"), and supported the earlier suggestion of Worrel (1934: 67) that we are dealing with a borrowing from an unknown source. It is in any case clear that Egyptian cannot be the direct source of the European words, due to both a mismatch in vocalism (Latin *-umb-* requires a labial vowel, cf. Leumann 1977: 81), and Egyptian *-r-* vs. European *-l-*.[36] The latter alternation is paralleled in the Mediterranean by Lat. *hirundō* ~ Gr. χελῑδών 'swallow'.[37]

In principle, a North African source for a word for 'pigeon' would not be in contradiction to the facts of the bird's domestication history (cf. Batisti 2021: 210), but it must be stressed that little is certain, except for the fact that the pigeon was domesticated extremely early (Gilbert/Shapiro 2013).

▸ **'swan** (2)'. Lt. *gulĩbė*, (Szyrwid, E dial.) *gulbìs*, Lv. *gùlbis* 'swan'; Pr. E *gulbis* ~ MR *колпь* (СРЯ 11–17 VII: 254; R *ко́лпица*) 'spoonbill'; Kash. *kôłp*, USrb. *kołp* (GEN.SG. *kołpja*), SCr. dial. *kûf* and *kûp* 'swan' (Oštir 1930: 66; Derksen 1999; ALEW 432–433) — As to the rare SCr. *gûb* 'swan', scholars are divided. Some reject it as an irrelevance (Vaillant 1929: 270 "douteux et sans intérêt"; Sławski 1960: 40), while others accept it at face value (Топоров ПЯ II: 332; ЭССЯ VII: 190; Andersen 1996a: 124, 2003: 68; Derksen 1999: 72, 2008: 97). The SCr. form is indeed very poorly attested, going back to a form ⟨gūb⟩ in J. Stulli's diction-

35 Allen actually reconstructs a final */-nipV̆/, but apparently only because the Egyptian genitive marker ⟨n⟩ is reconstructed as */ni/. This might be anachronistic, as spellings with ⟨m⟩ are already attested in Late Egyptian (Allen op. cit.; see Erman/Grapow V: 181), suggesting that no vowel was present in at least some Late Egyptian varieties. The spellings with ⟨n⟩ may be etymological, or, as follows from the discussion below, folk etymological.

36 On the nature of Egyptian ⟨r⟩, see Peust (1999: 127–129).

37 And perhaps — if not a mere dissimilation — by Lat. *līlium* ~ Gr. λείριον 'lily'. The latter are frequently also connected with Coptic ϩⲣⲏⲣⲉ 'flower' < Egyptian *ḥrr.t* */harīra.t/ (Worrel 1934: 67, Beekes 2010: 845; on the reconstruction, see Vycichl 1990: 94), but this etymology is suspect due to the absence of any reflection of the first syllable in the European languages, and the imprecise semantic match (cf. Vycichl 1983: 310).

ary, where it is attributed to "Gjorg." (apparently Ignazio Giorgi; cf. RJA III: 484, where the form is explicitly labelled as doubtful) and a form *guf* in the 17th century dictionary of J. Mikalja (idem: 495). Such forms otherwise only appear to be attested in lexicographical sources.[38]

The difference between Baltic **gulb-* and Slavic **kulp-* (trad. **kъlp-*) is already sufficient to suggest a non-IE origin (cf. Derksen 1999: 73 and *passim*).[39] The distribution of the word in Slavic is remarkable, being limited to the peripheral dialects of West Slavic, an isolated pocket in South Slavic, and East Slavic in a secondary meaning. All of this seems suggestive of an archaism: this might be the older Slavic word for 'swan', which was later displaced by **albǭdi-* (trad. **olbǫdь*) in the West and **lebedi-* (trad. **lebedь*) in the East — a theory supported by the fact that no common Proto-Slavic form for the latter can be reconstructed (see pp. 176–177).

▸ **'dregs'**. Lv. (Kurzeme) *dradži* 'dregs of melted fat',[40] Pr. E *dragios* PL. 'dregs'; ON *dregg* 'yeast, (PL.) dregs' ~ Lat. *fracēs* F.PL. 'olive pomace' (Ernout/Meillet [1951]: 251; Schrijver 1991: 486; Derksen 2008: 121) — The traditional explanation (Walde/Hofmann I: 539; de Vaan 2008: 238; ALEW 248) that the stem-final /k/ in Latin *fracēs* was carried over from the nominative singular does not hold water, as the word was *plurale tantum* in Latin, the singular *frax* only being attested in glosses. Moreover, neutralizations on the basis of nominative forms are generally suspect, as the nominative usually occupies a weak position in analogical processes (see Niedermann 1918: 22–23; and note the discussion in Decaux 1966). As a result, Latin implies an illegal root structure **D*ʰ*_k-* (see 5.3.1). The remainder of the words traditionally adduced here are uncertain.

In Slavic, OCS (Ps. Sin.) дрождьꙗ 'dregs (Gr. τρυγίας)', Pl. *drożdże* F.PL. 'yeast, leaven' suggests an underlying **drazg-* or **drazdj-*. The old explanation has been to posit **dʰragʰ-sk-* for Slavic with the subsequent development to **-gsk-*

38 I do not have access to all of the sources cited by PCA (III: 721), but I suspect that most or all of the forms trace back to these two sources. One wonders whether there might have been a confusion among the lexicographers with the Latin loanword attested as dial. *gûb* 'goby' (PCA loc. cit.) ← Latin *gōbius* (on this loan and variants, see M. Matasović 2011: 163–164 against ЭССЯ and Derksen loc. cit.). The Sln. dial. (Gorizia) *golbica*, which Bezlaj (I: 157) adduces in this connection, refers to the 'skylark', a tiny passerine bird which has absolutely nothing in common with the swan.

39 Derksen's inclusion of Pl. *kiełb* 'gudgeon' and OIr. *gulban* 'sting; beak' in a substrate context looks like an unnecessary stretch to me, as the semantic link between the three word families is not self-evident.

40 Prussian Lithuanian *drãgės* (attested as *dragges* in Bretke) is most likely a loanword from Prussian (Žulys 1966: 151–152).

> *-*zg*- (e.g. Berneker I: 228; REW I: 371), for which cf. OCS про-брѣзгъ 'dawn' beside Skt. *bhrā́jate*. Regardless of whether this development is phonologically regular, the explanation is inadequate as there is insufficient evidence for a nominal suffix *-*sk*-. If we posit an underlying form **drazdj*-,[41] another reasonable etymology presents itself, namely a comparison with OE *dræst*, *dærste* 'dregs, leaven', OHG *trestir* 'pomace' < **dʰrosd*-. Per tradition, the Germanic forms are derived from **drahstu*- (i.e. **drag-stu*-), to the root of ON *dregg* (Holthausen 1934: 69; IEW 251–252); cf. the semantically similar Go. *maihstus**, OHG *mist* 'dung' < **mihstu*-. However, the latter is no example of a suffix *-*stu*-, as it derives from a more primary **mihsa*-, cf. OE *meox* 'dung, filth' (Kroonen 2013: 369), and is ultimately deverbal, cf. OE *mīgan* 'urinate'. For **drag*-, we have neither a verbal root nor a primary *-*sa*- derivative, which means that the alternative reconstruction **drastV*- remains a clear possibility.

Contra Meyer (1891: 72) and Demiraj (1997: 141), Alb. *drā* 'dregs of melted fat' cannot derive from **dragā*, as **g* was not lost intervocalically (cf. Schumacher 2013: 240). A possibility would be to posit a preform **drasā*- < **dʰrHs*-, and compare OE *drōsna*, Du. *droes*, *droesem* 'dregs, sediment' < **dʰroHs*-.

▸ **'scythe'.** Lt. *dal̃gis*, *dal̃gė*, Pr. E *doalgis* 'scythe' ~ Lat. *falx -cis* 'sickle, scythe' (Alessio 1946: 165) — This rather self-evident comparison (cf. Mikkola 1899: 74; Hirt 1927: 299) has generally been disfavoured in view of the irregular Latin /k/ and irregular vocalic correspondence (Walde/Hofmann I: 449–450; LEW 81; against an analogical origin of the /k/, see above on 'dregs'). The Latin word has been suspected to be of foreign origin, but the Baltic equivalent is rarely mentioned in this connection (e.g. Ernout/Meillet 214; Boutkan/Siebinga 2005: 75; de Vaan 2008: 200). For Latin, a non-IE origin is supported by the illegal root structure **Dʰ_k*- (see 5.3.1). Walde/Hofmann (loc. cit.) claim that *dal̃gis* is "aus semasiologischen Gründen" better compared with OIr. *dluigid* 'split, cleave', ON *telgja* 'carve, hew (wood or stone)' (also Trautmann 1923: 44; IEW 196). This is rather a strange argument, since *falx* and *dal̃gis* mean exactly the same thing, and the cited verbs are semantically rather remote, belonging to the sphere of artisanry rather than agriculture.[42]

▸ **'rye'.** Lt. *rugiaĩ*, Lv. *rudzi*, Pr. E *rugis*; R *рожь*, Sln. *rȓ̂ž*; ON *rugr*, OE *ryge* 'rye' ~ MW *ryc* 'rye' (Hoops 1915–1916: 509–510; Walde/Pokorny II [1927]: 375;

41 Sln. *drȏzga* 'pulp, dregs of lard' does not disprove the reconstruction with **zd*, cf. Sln. *drȏzg* 'thrush' < **drazda*-; see p. 204. On the other hand, this Sln. word may not belong here, as it is highly reminiscent of the synonym *trȏska* (cf. SCr. *trȍska*, arch. *trŭska* (RJA XVIII: 829) 'slag').

42 Note that, according to Schumacher (2004: 284–285), the Irish verb is rather to be reconstructed **dlug*-, and a connection to the Germanic root is thus impossible.

Schrader/Nehring II [1929]: 265; Charpentier 1930: 71; Porzig 1954: 143; Markey 1989: 595; Polomé 1992: 70; OED³ s.v. *rye*)[43] — All of the cited forms show a formant *-i-*. Divergent forms are attested in Continental Germanic, where we find OS *rokko** (attested *rogko*), OHG *rocko* as against Old Frisian *rogga*, MDu. *rogge*. This vacillation in voicing is to be explained from an old *n*-stem (Kroonen 2011: 23). This is probably a localized innovation, however, and Kroonen (2013: 417) points out some possible West Germanic traces of **rugi-*.

MW *ryc* has generally been derived from Old English *ryge* (Schrader/Nehring II: 265; GPC III: 3136; Witczak 2003: 110), but this is chronologically difficult, as Welsh /k/ could hardly be a substitute for the OE spirant /j̯/, cf. MW *pabi* 'poppy' ← OE (Ælfric) *papig* (early ME *papig*), MW *llidiat* 'gate' ← OE *hlidgeat* 'swing gate' (cf. Parry-Williams 1923: 41–42).[44] On the other hand, if the loan were of Proto-Celtic age, one would expect Celtic **g* (> Welsh **∅).[45] At face value, the Welsh data points to **rukį̯o-* or **rukī-*, showing a mismatch compared to the **g*ʰ elsewhere.

Beyond Indo-European, a similar word is found in several Uralic and Turkic languages. Already Paasonen (1906: 2–3) recognized that the Mordvin and Permic words for 'rye' cannot be derived directly from Russian, as had previously been thought (e.g. Thomsen 1890: 213). For Mordvin *roź*, the problem is the final *ź*, which otherwise does not substitute Russian *ž* (*pace* Häkkinen/Lempiäinen 1996: 169). In Permic, we have Komi *ruźeg*, Udmurt *źeg*, dial. *źiźeg*. Already the development **r-* > *ź-* in Udmurt speaks against a Russian loan, but the palatal affricate and suffix solidify this impression. The initial vowel in Udmurt reflects earlier **į* (< **u*, Лыткин 1964: 215–218) which has become fronted by the following palatal, cf. Komi *ruć*, Udm. *źićį* 'fox'. The syncope in the standard language is attested dialectally in other lexemes, e.g. dial. *slal*

43 Here one has often included a form βρίζα (e.g. IEW 1183), a crop which according to Galen was grown in Thrace and Macedonia, resembling τίφη 'einkorn', and from which a black and malodorous bread was made (cf. Schrader/Nehring II: 265). The word is found in the sense 'rye' in modern Greek dialects, first resurfacing in a 16th century Macedonian–Greek glossary as ἄρκυ · βρίζα (cf. Mac. *ржс* 'rye'; Giannelli/Vaillant 1958: 32). Despite its meaning, it is perhaps better connected to a different *Wanderwort*, represented by Gr. ὄρυζα, Pashto *wríže*, Skt. *vrīhí-* 'rice' (Георгиев 1957: 55); for the shift to another kind of grain, cf. Kati *wrić* 'barley' (CDIAL 708; Kümmel 2017b: 281).

44 These could both be from Middle English according to GPC III: 2663 (s.v. *pabi*) and GPC I: 1297 (s.v. *fflodiart*), but this is of little relevance if the spirantization of /g/ is dated to the continental Old English period (Campbell 1959: 173).

45 Cf. MW *meu-dwy* 'hermit' (*duw* 'God'), MCo. *maw* 'boy, servant' ← Germanic **magu* > OE *magu* 'boy, young man' (van Sluis et al. 2023: 201, 212).

'salt' beside literary *sɨlal* (Перевощиков 1962: 37–38). Thus, we can confidently reconstruct a Proto-Permic **ruʒ́ɛg* 'rye'.[46]

In Turkic, we find rather a similar situation, although little recognized. Ахметьянов (1981: 48–49) has argued that Tatar and Bashkir *arïš* 'rye' cannot have been adopted directly from Russian (thus e.g. Joki 1973: 162), as neither the prothetic *a-* nor root vowel can be accounted for (one would rather anticipate ***ïruš* or ***erüš*). Despite Федотов (II: 474), a similar conclusion must be drawn with respect to Chuvash *ĭraš* 'rye', as *ĭ-* never occurs as a prothetic vowel, nor is *-a-* for Russian stressed *-o-* the usual substitution.[47] The correspondence of Chuv. *ĭ-* as against *a-* elsewhere is in fact more typical of inherited vocabulary, where it would reflect a Proto-Turkic phoneme notated **ạ* (in the Russian school) or **ë* (e.g. Doerfer 1971: 340–341). Reconstructing a word for 'rye' back to Proto-Turkic is suspect, however, as early rye cultivation is normally associated with Central Europe (cf. Hillman 1978: 157–158; Напольских 2006: 5–6; 2010: 56).

Paasonen (op. cit.) assumes the word for 'rye' was adopted into the Uralic languages from Iranian, or more specifically, Scythian. Of course, as long as no Iranian equivalent is attested (cf. Kümmel 2017b: 283 on the alleged Pamir words), this remains purely hypothetical. Slightly better is Guus Kroonen's suggestion (see Kroonen et al. 2022: 22) of an early Slavic loanword mediated by "steppe Iranian". Although still hypothetical, this would obviate the need to reconstruct the word for Proto-Indo-Iranian. Furthermore, a couple of other agricultural *Wanderwörter* seem to have passed into Scythian from a Balto-Slavic dialect, most notably Oss. I *ӕxsyrf*, D *æxsirf* 'sickle' ← Lv. *sìrpis*; R *cepn*, etc. (Abaev 1965: 8–9; Gołąb 1992: 333).[48]

While the Slavic → Scythian route perhaps makes the most sense, a theoretical pre-Oss. **ruʒ*(*-æg*) would hardly account for the Turkic evidence. If the Turkic forms belong here at all, then perhaps we can assume the initial **ạ*/*ë-* was some kind of prefixal element or the like (see 7.1.2), but in the absence of parallels, the idea must be approached with caution. The Uralic and Turkic pal-

46 Mari E *urža*, W *ârža*, *râža* 'rye' is indeed probably loaned from Russian. As a precise source, the final *-a* is best accounted for starting from GEN.SG. *ржа* from R dial. *рож* (M.) (Orenburg, etc., see СРНГ XXXV: 146).

47 Exceptions like Chuv. *salat* < *со́лод* 'malt' are rather to be explained from end-stressed forms (*со́лод* is originally accentually mobile, cf. Зализняк 2019: 541).

48 The suffix **-ɛg* seems to almost call for an Iranian origin and comparison with Oss. *-æg* (cf. Paasonen 1906: 4–5), as in Old Permic ⟨идог⟩ */idẹg/ 'angel' (cf. Лыткин 1952: 65, line 27; also idem: 70, fn. 4 and idem: 130) ← Oss. D *idawæg* 'angel, spiritual guardian' (Абаев 1958: 348–349; Rédei 1986: 70). Not all cases of the suffix can be explained as Iranisms, however.

atal affricates, by the way, do not necessitate a Slavic origin, as palatalization is a trivial change which could have occurred independently before a following **i* or **į* in another (hypothetical) source language.

It has long been suspected that the word for 'rye' is of non-Indo-European origin, although primarily on the basis of non-linguistic facts. The Celtic form (see above) can now provide some more concrete linguistic evidence in favour of this analysis. It is clear we are dealing with a cultural *Wanderwort* whose spread is difficult to precisely trace. Rye was first domesticated in Eastern Turkey and Armenia, but already arrived in Northern Italy in the Neolithic (Zohary/Hopf 2012: 63–66); however, the sporadic finds in later Polish sites are probably more consistent with the plant being tolerated as a weed than intentionally cultivated (Behre 1992: 142–143).

Rye cultivation only really took off in Northern Europe during the Iron Age, and does not appear to have reached the Eastern Baltic until the common era (Grikpėdis/Motuzaitė Matuzevičiūtė 2016). On this basis, it would be tempting to interpret the East Baltic forms as Germanic loanwords, which would be phonologically unproblematic; compare similarly Lt. *kviečiaĩ* 'wheat', which I have interpreted as a Gothic loanword (see p. 40). A Germanic loanword cannot be ruled out on phonetic grounds for Slavic, although such an assumption would not be necessitated by the realia.

▸ **'hornbeam'**. R *грабъ*, Cz. *habr*, dial. *hrab*, SCr. *grȁb* 'hornbeam' ~ Lat. *carpinus* 'hornbeam'. Here also belong Lt. *skrõblas*, *skróblas*, (S Aukšt.) *skrúoblas* 'hornbeam'[49] (Machek 1950b: 152; Holub/Kopečný 1952: 118; Matasović 2013: 84; ERHJ I [2016]: 291; Matasović forthc.) — The original Slavic form can be reconstructed **grābra-* (trad. **grabrъ*) with various dissimilations (Berneker I: 343; Skok I: 598), cf. SCr. dial. *grȁbar*, Sln. dial. *grâbər*.[50] Perhaps here also belong

49 The literary standard is circumflex, although *-ó-* seems better supported dialectally. The variant with *-úo-* is in any case irregular (contamination with *gúoba* 'elm', *úosis* 'ash', *glúosnis* 'willow'?). The initial *s-* in Baltic is also unclear. Otrębski (1955: 29; cf. 1939: 167) finds a parallel in Lt. *strãzdas* ~ R *дрозд* 'thrush', yet here we are probably dealing with anticipation of the second **s* (see the discussion on p. 204).

50 The generalization of /grab/ in East Slavic is not surprising considering the partial parallel in R *брат* ~ Cz. *bratr*, Sln. dial. *brâtər* 'brother' (Holub/Kopečný 1952: 118). The two variants must clearly not be separated (despite Būga 1922: 82; LEW 176–177). Būga, followed by Boryś (2008: 176), has also adduced Pl. dial. *gab*, *gabina* 'elm', attested in transitional Polish-Belarusian dialects (= Bel. dial. *габ*, *габіна*). In view of their distribution, these words must no doubt be considered Balticisms (Лаучюте 1982: 43–44). Note that a loan directly from Lt. *gúoba* 'elm' is prohibited by the Slavic vocalism, so it would be preferable to posit a Prussian **gābas* as the immediate source. The different form and semantics imply it should be separated from our words for 'hornbeam'.

Lv. *skābardis* (*skãbârdis*[2] Dunika, EH II: 503, *skãbàrdis* LVPPV) 'beech', Pr. E *stoberwis* (corrected to **sc*-) 'hornbeam', assuming an original **skrābar*- with dissimilatory loss of the first **r* (Trautmann 1910: 439), although the formation of these words is obscure (cf. ME III: 878; IEW 945). The comparison between Slavic **grābra*- (trad. **grabrъ*) and Lat. *carpinus* is obviously semantically attractive. The Latin word has no satisfactory etymology. The connection to *carpō* 'pluck' (supposedly < *'cut') based on the hornbeam's crenated leaves (Walde/Hofmann I: 171; Schrijver 1991: 430) is hardly logical, as the plant's leaves are neither sharp nor capable of cutting, nor for that matter, strikingly different to those of the elm or beech.[51]

The etymological equation involves multiple irregularities. First of all, the labial stop alternates in voicing along with the velar in a similar way to Lt. *gulbė* ~ Kash. *kôłp* 'swan', above. In addition, there is a metathesis of **r*, which does not appear to have reliable parallels in my corpus. An alternative analysis is to assume that the Latin form goes back to an earlier **crarp*- by dissimilation, as probably in *prōcērus* 'lofty' (~ *crēscō* 'grow'; Leumann 1977: 315; de Vaan 2008: 491), *prō portione* 'in proportion' (with *portione* < **prō ratione*, Ernout/Meillet 524). In this case, we would have a potentially more trivial metathetic relationship between **-Pr*- and **-rP*-.[52]

▸ [**'oriole'**. Lt. *volungẽ* 3ª, Lv. *vãluôdze* 'oriole'; Pl. *wilga*, Sln. *vólga*; R *úволга*, Bg. *авлùга* ~ ME *wode-wale*, MHG *wite-wal* 'oriole' — See the discussion on pp. 178–180.]

▸ † **'many'**. OCS мъногъ 'many, numerous'; Go. *managṣ*, OHG *manag* 'many' ~ OIr. *meinic*, MW *mynych* 'frequent' (< **menekki*-) (Boutkan 1998: 124; Schrijver 2001: 422; Boutkan/Siebinga 2005: 256–257; Philippa et al. III [2007]: 334; Kroonen 2013: 352; Matasović 2013: 265; ERHJ I [2016]: 625; van Sluis forthc.) — The main question is whether the Slavic word can be interpreted as a Germanic loan (thus Hirt 1898: 355). The cost of this assumption would be an irregular raising **a* > **u* (trad. **ъ*) in an unstressed syllable within Slavic (Младенов 1909:

51 A relationship with Umbrian krapuvi DAT.SG., an epithet of Mars and Jupiter (Kretschmer 1921) cannot be demonstrated (cf. Untermann 2000: 309–310 with lit.). Comparing Hitt. *karpina*- 'a kind of tree or bush' (IEW 944; Puhvel 1997: 99) is also precarious in view of its uncertain meaning. I would also like to keep Lt. *skiřpstas* 'elm; alder buckthorn', Pr. E *skerptus* 'elm' apart due to the semantic and formal difference (note also the Lithuanian circumflex; but it must be admitted that *skiřpstas* has also been recorded in the sense 'hornbeam').

52 Compare OSw. *hagre* as against MIr. *corca* 'oats' (van Sluis et al. 2023: 219; however, more sceptically: Kroonen et al. 2022: 20).

85) for which a couple of parallels may be found.[53] As the Germanic *-*g*- could just as well reflect *-*k*- in this position, this would then not be a certain example of a voicing alternation.[54]

Kortlandt (2007: 9) claims that мъногъ has developed from **min-aga*- (trad. **mьnogъ*) as a result of vowel assimilation, comparing Lt. *minià* 'crowd'. However, it is doubtful that the Lithuanian word belongs here (LEW 453; ALEW 753 and Smoczyński 2018: 804–805 all accept a derivation from *mìnti* 'tread, trample'; compare French *foule* 'crowd' < *fouler* 'trample', FEW III: 846). Moreover, although the "normalized" form мъногъ is found in dictionaries, the OCS word is very frequently written мног-, showing an early reduction; it is uncertain how much the spelling мъног- can be relied on.[55]

While the geminate in Celtic is very difficult to explain in an IE context and could very well point to a non-IE borrowing, the difficulty in analysing the word within Slavic and the possibility of a Germanic mediation means that this word cannot be used here as an example of a voicing alternation.

▸ † **'naked'**. OCS (Supr.) голъ 'naked'; OE *calu*, OHG *kalo** 'bald' ~ Lat. *calvus* 'bald' (Philippa et al. II [2005]: 593–594) — Despite the striking correspondence between the substantivized Lat. *calva* 'bald head' and OCS глава, Lt. *galvà* 'head' (Derksen 2008: 176), the comparison is probably false. West Germanic **kalwa*- matches Latin *calvus* formally and semantically, and is therefore most easily explained as a loan from Latin (Senn 1933: 521; FEW II: 106; cf. Philippa et al. loc. cit.).[56] The Latin word must reflect **kalawo*- < **klH-eu*- (Schrijver 1991: 299) and can hardly be separated from Skt. *kulva*-, YAv. *kauruua*- 'thin-haired' (< **klH-uo*-, on which see Lubotsky 1997: 144). In view of the Indo-Iranian com-

53 The only relatively clear example is the verb 'to want': cf. Pl. *chcieć* 'to want' vs. *chociaż* 'although' (= R *хотя́ть*, *хотя́*). A similar change has also been suggested in OCS сто 'hundred' if a loan ← Iranian **sata* (cf. Vasmer 1913: 176; Шахматов 1916b: 29, Arumaa 1964: 130) and perhaps Pl. *młyn*, Cz. *mlýn* 'mill' if cognate with Pr. E *malunis* 'mill' (cf. Meillet 1907: 373–374; but compare Fraenkel 1951: 129). Suffice to say that neither of these parallels are uncontroversial. On a similar sporadic raising **e* > **i* (trad. *ь) before a palatal, cf. Kortlandt 1984–1985.

54 It is unclear to me why Младенов (loc. cit.) and Kiparsky (1934: 75) after him insist that the Slavic and Germanic words must be cognate (cf. Viredaz 2020: 413–415).

55 Compare the similar situation with regard to the inst.sg. мъноѭ, also seen as an example of this assimilation by Kortlandt, but essentially representing a "traditional" OCS form, not based explicitly on facts of the language (cf. Vaillant 1958: 446; Lunt 2001: 77). Leskien (1922: 109) has even considered the dative variant мьнѣ, on the contrary, to have arisen by assimilation from мънѣ. See also the discussion in Kapović (2006: 39–41), with lit., who problematizes the dative, but remarkably takes the form of the instrumental for granted.

56 This possibility is denied by EWAhd V: 353, but without any argumentation.

paranda, it is more likely that the Balto-Slavic words are unrelated. They may instead be cognate with Arm. *čeł* 'bald' < **g(ʷ)el-* (Olsen 1999: 206).

6.2.2 *Baltic* *k ∞ *g(h) *Elsewhere*

▸ **'pear'**. Pr. E *crausy* 'pear tree', *crausios* 'pears' ~ R *ерýша*, Cz. *hruška* 'pear' (Hehn 1870: 454; Trautmann 1923: 140; Machek 1954: 114; LEW [1962]: 296; Matasović 2013: 92) — According to Būga (1915: 342), Žem. *kriáušė* and Lv. dial. (Nīca) *kraûsis*[2] (ME II: 264) are loaned from Prussian, which is supported by their marginal distribution (see also Топоров ПЯ III: 168–169; Žulys 1966: 152; Аникин 2004: 380). The Slavic reflexes are as follows. In East Slavic, we find only *ерýша*,[57] a form which is otherwise only known in West Slavic (e.g. Pl. *grusza*, Cz. *hruška*). Pleteršnik (I: 258) cites a Sln. *grûška*, but this form is actually a normalization of dialectal /hrù:ška/ (Karničar 1986: 153) and might represent a localized borrowing from a dialect with a realization /hr-/ < **xr-*, cf. standard Sln. *hrûška* (T. Pronk p.c. March 2023). South Slavic more typically shows **k-* (Bg. *крýша*, SCr. *krûška*),[58] and such a variant occurs as a relic in West Slavic, cf. USrb. arch. *krušej* (GEN.SG. *krušwje*), LSrb. *kšuška*; Kash. *krëszka* (also Pl. dial., cf. Popowska-Taborska 1996: 154). As a result, the word for 'pear' in Slavic shows a partial complementary distribution. At first sight, the peripheral attestation of **k-* in West Slavic would point to an archaism, suggesting that **g-* has spread through this territory secondarily (although still in the preliterary period). If true, this would allow us to draw an earlier isogloss between East Slavic **g-* and West/South Slavic **k-*, which would be somewhat reminiscent of the situation with 'oriole' and 'swan' (see under 6.1.1).

Since the 19th century, the consonantal alternation has been considered evidence that the word for 'pear' is a loanword from an unknown language.[59]

---

57 It is tempting to see RCS хроуша as an early reflection of the Ruthenian 'spirant g', even though Shevelov (1979: 351) simply dismisses it as a scribal error. Curious is the form кроуша, which glosses the Greek name Апии (corrupt for Σινάπης!) in the Chronicle of George Hamartolos (СДРЯ 1338). A Bulgarian form? (cf. Пичхадзе 2002).

58 Sln. *hrûška*, SCr. dial. (NW) *hruška* presumably show a secondary spirantization (dissimilation?).

59 Schrader (1901: 93; Schrader/Nehring I: 148) refers to a Kurdish "*korêši*, *kurêši*", which has been routinely mentioned in later works (e.g. Berneker I: 358; LEW 296; REW I: 314; Sławski SP VIII: 256). Schrader's immediate source appears to be Rhea (1872: 145: "*korēshī* or *kurēshī*, *n.* pear"). However, I am unable to trace this form elsewhere. Noting that Rhea fails to distinguish /q/ and /k/ (see the editor's note op. cit.: 120), Patrick Taylor (p.c. June 2022) has attractively suggested that the Kurdish word could represent the common surname *Qureyşî*, and that this would have referred to a cultivar associated with someone of that name. As a typological parallel, he notes the grape variety *Kureyş* (*üzüm*) found in eastern Turkey.

This seems quite probable, and could perhaps explain the divergent forms within Slavic. Smoczyński (2018: 603) would rather posit an irregular "sonorization" in part of Slavic, which is *ad hoc*. No better is the native etymology in ЭССЯ (VII: 156–157) comparing MR *крушити* ~ R dial. *грушúть* 'break up, crumble; destroy' — an implausible suggestion from a semantic point of view (cf. Matasović 2013: 92).

▸ † '**meadow**'. Lt. *lénkė* 'depression, marshy spot';[60] Pl. *łąka* 'meadow', Sln. *lǫ́ka* 'damp meadow by a river' ~ R *луг* 'meadow', Pl. dial. (Sł. Warsz. II: 805) *łąg* 'flood meadow by a river' (Derksen 2008: 288; 2015: 279–280) — Derksen further adduces Lt. *líeknas*, Lv. *liẽkna* 'depression; marsh, swamp' and a number of other forms.[61] It can be difficult to tease the *k*-forms apart from derivatives of the root **lenk-* 'to bend', cf. Lv. *lañka* 'low-lying meadow; bend in a river', R *лука́*, Bg. *лъ́кà* 'river bend; meadow in a river bend', and if *lénkė* is metatonical (Derksen 1996: 200), there really seems to be no decisive argument against this internal etymology. Similarly, Slavic **lǭga-* (trad. **lǫgъ*) might be explained as an inner-Slavic derivative based on the present stem **lę̄g-* (trad. **lęg-*) 'to lie' (OCS 1SG.PRES. лѧгѫ; see Loma 2012: 84); for the semantics, compare R *лог* 'broad valley; (dial.) low-lying, damp spot; water meadow' (cf. СРНГ XVII: 103), an undisputed derivative of *лечь* 'lie' (e.g. REW II: 51).

6.2.2.1 Conclusion

The certain and possible evidence for voicing alternations involving velars is collected in Table 9, overleaf. The principles used in this table are the same as for Table 8 (p. 185). In addition, forms which do not provide relevant data are presented in light grey. Shaded cells indicate forms containing a voiceless velar.

The parallelism between these examples is quite striking. In all of the examples except 'pear', Balto-Slavic almost consistently shows a voiced velar, while Italo-Celtic shows a voiceless one. Seven examples not only showing a similar alternation, but also a largely matching distribution, can hardly be a coincidence. This correlation is most straightforwardly explained as a reflection of a genuine dialectal difference in the underlying source language. The

60 On *léngė*, see Chapter 1, fn. 61.

61 I think it is going too far to include e.g. Sln. *lúža* 'puddle' and Lt. dial. *liū̃gas* (beside *liū́gas*) 'puddle, marsh', which lack the nasal. The Lithuanian word is always analysed as cognate, but I wonder whether it is rather a loan from Belarusian *луг* 'meadow; swampy area' — the initial /lʲ/ could be explained through contamination with the semantically similar *liū́nas* 'swamp'.

TABLE 9 Possible examples of *$g^{(h)}$ ∞ *k alternations

	Baltic	Slavic	Germanic	Elsewhere
'pigeon (1)'	–	*Golamb^{h}-	*gulu[b^{h}]-	Lat. *kolomb^{h}-
				Eg. *k^{h}Vramp-
'swan (2)'	*Gulb^{h}-	*kulp-	–	
'dregs'	*Drag^{h}-i̯-	–	*d^{h}ra[g^{h}]-i̯-	Lat. *d^{h}rak-
'scythe'	*Dalg^{h}-	–	–	Lat. *d^{h}alk-
'rye'	[*rug^{h}-i̯-]	*rug^{h}-i̯-	*ru[g^{h}]-i̯/n-	Celt. *rukī-
'hornbeam'	*skrâB-r-	*GrâB-r-	–	Lat. *k(r)arp-
'oriole'	*u̯âlanG-	*u̯(i)lg$^{(w)}$-	*u̯alk-	
'pear'	*kraus-i̯-	*graus-i̯-	–	
		*kraus-i̯-		

Slavic word for 'swan' is at first sight an exception, but we may consider this a function of the intermediate position of Slavic, which would enable contacts both with the Mediterranean and with Northern Europe. The Germanic evidence in this section is largely obscured by Verner's Law: the words for 'dregs' and 'rye' could equally be taken back to an earlier *k and final stress. However, the word for 'oriole' appears unambiguously to imply *k.

As will emerge from the following sections, examples of voicing alternations involving stops other than velars are relatively few. This might potentially be connected to the cross-linguistic tendency of /g/ towards lenition (Foley 1977: 25–35), exemplified by the Central European areal change *g > /ɣ ~ ɦ/. In languages which lack a phoneme /g/, such as Czech, foreign /g/ may be substituted with /k/ in loanwords, e.g. Czech dial. *kuláš* 'goulash', *brikáda* 'brigade' (ČJA v: 317). Thus, one possible explanation for a general trend towards the devoicing of velars in the south might be the mediation of an unattested language which lacked a phoneme */g/. Of course, this remains purely speculative, and one

might object that the word for 'pigeon' is already attested with a "fortis" */kʰ-/ in Late Egyptian, some two millennia before it emerges with a */g-/ in Northern Europe.

6.2.3 *Alternations Involving Dentals*

▸ **'drone'**. Lt. *trānas* 'drone', R *тру́тень*, Pl. arch. (Sł. Warsz. VII: 130) *trąd*, Sln. *trȏt* 'drone' ~ Gr. (Nicander) τενθρήνη 'wasp'; OE *dran*, OHG *treno* 'drone' (Kuiper 1956: 221–222; Beekes 2010: 105; Matasović 2013: 96; van Sluis 2022: 12–15, forthc.) — Kuiper and Beekes cite a plethora of variants for the Greek word, but it is all but impossible that each of these is equally old, and their sheer number would only support the notion of secondary developments. Most likely, we are dealing with multiple lemmata which have influenced each other (cf. Chantraine DELG I: 90). If we accept the derivation of ἀνθηδών 'bee' from ἄνθος 'flower' (Chantraine 1933: 361; cf. Frisk I: 108), then a collision with τενθρήνη 'wasp' would explain variants such as ἀνθρήνη (Aristotle) 'a kind of hornet' (differently Chantraine DELG loc. cit.). The apparent reduplication in τενθρήνη is reminiscent of (Epic) δένδρεον 'tree' (< **der-drew-*, cf. Chantraine DELG I: 263). The only other form relevant for our purposes is the Hesychian gloss θρώναξ · κηφήν 'drone (Laconian)'.[62]

OE *dran* and OS *dran*, *drano* (> MoLG *Drahn*) 'drone', most probably with short vowel (OED³ s.v. *drone* n.[1]), differ in vocalism from OHG *treno* 'drone'. As the OS variant *drenon* (ACC.PL.) may be the result of a secondary development in the neighbourhood of /r/ (Cordes 1973: 137), the form with **e* seems essentially to be limited to High German. MoE *drone*, attested since the 15th century, does not represent a regular continuation of the OE form and Kroonen (2013: 101) has argued that this, like MDu. *dorne*, could represent an additional ablaut variant **drunan-*. While it is possible that the vocalic alternations could be explained by positing various ablaut grades in Germanic and Greek, the number of variants which have to be assumed makes this quite unattractive.

Šorgo (2020: 437) rejects a non-Indo-European origin, prefering the traditional explanation that the whole family is of sound-symbolic origin (Walde/Pokorny I: 861; Frisk I: 681–682, etc.). Certainly, some of the variants may be explained in this way; for instance the variant **drunan-* could plausibly have

62 Unreliable is ἀθρήνη, only attested by Byzantine-period lexicographers. The forms θρήνη and θρηνῶδες are additionally cited by Beekes (2010: 105; evidently taken over from Winter 1950: 45). The former is a hapax in Eustathius (12th c. CE) and is probably a corrupt form, while the latter is a manuscript variant of τενθρηνιῶδες 'honey-combed' ("in der Überlieferung stark entstellt, z.T. zweifelhaft", Frisk II: 877). None of this evidence can be used to support pre-Greek origin.

arisen under the influence of MoE *drone* (since 16th c.), MDu. *dronen* 'hum, buzz' (OED³ loc. cit.). However, it is difficult to justify this analysis in detail. Other forms cited in this connection, e.g. Gr. θρῆνος 'lament, dirge' (: τενθρήνη), Pr. III *trinie* 3PRES. 'threaten' (: Lt. *trãnas*) (Frisk loc. cit.; Endzelīns 1943: 266; LEW 1110–1111), are semantically ambiguous and the assumption of an underlying *Schallverbum* remains without direct support.

▸ '**reed**'. Lt. dial. *trùšis* 'reed', Lv. dial. *trusis* 'rush, bullrush' (ME IV: 248; EH II: 699); OCS тръсть, R arch. *трость*, OPl. *treść* (SSP IX: 184) 'reed' ~ Gr. θρύον 'reed' (Kuiper 1956: 224; Furnée 1972: 135; Beekes 2000a: 28; Matasović 2013: 88) — The OCS variant трьсть, along with R dial. (N) *тресть* are to be explained as cases of *yer* assimilation (Соболевскій 1910: 116–117).[63] The Lithuanian variant *triùšis* has been seen as paralleling the Slavic variants (Булаховский 1958: 91), but is rather to be explained as a result of the frequent but sporadic dialectal change /Cr/ > /Crʲ/ (Zinkevičius 1966: 153–156). An inner-Baltic derivative with ablaut is Lt. (Žem.) *triaũšiai* 'horsetail, *Equisetum*', Lv. (Stender *apud* ME IV: 227) *traušļi* 'Flusskannenkraut', i.e. *Equisetum fluviatile*(?); the further comparison with Lt. *triáušėti* 'crack, split (usu. of hairs)' (cf. Būga 1922: 288; LEW 1133) is semantically unattractive.

The initial aspirate in Greek is not consistent with Balto-Slavic **t-*. Smoczyński (2018: 1530) is willing to accept anticipatory aspiration due to **s* (cf. Sommer 1905: 46–82; Chantraine DELG II: 443). However, this development is assessed as highly doubtful by Frisk (I: 688), and is probably to be rejected. The Greek word is no longer mentioned by ALEW (1303), who leave the Balto-Slavic word without an etymology. It seems the Greek form can hardly be separated, but in view of the incongruent initial stops, the words cannot be directly cognate. Therefore, the suggestion of independent loanwords from an unknown source can be considered attractive. For a suggestion regarding Slavic **-st-*, see under 'furrow' (p. 224).

▸ ? '**lentil**'. RCS лача, SCr. *léća*, Bg. *лѣща* (< **lę̄tjā-*); Lat. *lēns -tis* 'lentil' ~ Gr. λάθυρος 'grass pea' (Hoops 1905: 463; Walde/Hofmann I: 783; LEW [1962]: 359; ЭССЯ XV [1988]: 63–65; etc.) — If the suffix **-jā-* is an inner-Slavic innovation, it cannot entirely be excluded that the word was borrowed from Latin. However, it is difficult to explain OHG *linsī* 'lentil' as a Latin loanword (EWAhd V: 1323; Kluge/Seebold: 580).[64] MDu. (15th c.) *lins* 'lentil' could phono-

63 Note that *contra* Соболевскій, Pl. *trzcina*, Cz. *třtina* 'reed' do not show a reflex of a front vowel, but have /r̥/ < **rs* as in Cz. *křtíti* (OCz. *krstiti*), Pl. *chrzcić* 'baptize' < **kristītī* (trad. **krьstiti*), cf. Lamprecht, Šlosar & Bauer 1977: 71.

64 A possible parallel for a borrowed nominative form is OS *pavos*, OHG *bābest* 'pope' (for the long *ā*, cf. *bâbes* in Notker, and also the loanword OCS папежь 'pope', ESJS 625). However,

logically be cognate, but has also been interpreted as a loan from German (de Vries 1971: 404–405, s.v. *linze*). As a result, a Proto-Germanic age is not ascertained, but the absence of the word in the other branches of Germanic may simply be due to the absence of the crop in northern Europe. A German origin is hardly possible for Lithuanian *lę̃šis* 'lentil' (cf. the recent loanword *lìnzė* 'lens').[65,66] Due to the nasal vowel, a Slavic origin is also implausible (see 1.1.4). Thus, this form is rather a conundrum: as lentils do not emerge in the archaeological record for Lithuania until the Middle Ages (Grikpėdis/Motuzaitė Matuzevičiūtė 2020: 167), it is difficult to accept a non-Indo-European origin, but an exact source cannot be established with confidence. In Latvian, *lę̃ca* 'lentil' must be explained as a loan from East Slavic **ля́ча* (cf. ME II: 455; REW II: 84),[67] whence it has been adopted into Estonian as *lääts* 'lentil'.

The comparison with Greek has not been universally accepted (cf. e.g. Berneker I: 708; REW II: 84; Frisk II: 71), and indeed it is based on rather little linguistic material and depends on the ultimately unprovable assumption that Greek *-ά-* goes back to an earlier nasal vowel. Since the word also refers to an edible legume, it may well belong here, but the evidence remains uncertain. While it is probable that our word for 'lentil' is of non-IE origin, the clearest irregularity is between the Germanic sibilant on the one hand and the dental in Latin on the other. It is uncertain to what extent the Balto-Slavic evidence is relevant here.

▸ **? 'lightning'**. Pr. E *mealde* 'lightning' ~ ON poet. *mjǫllnir* 'Thor's hammer'; MW *mellt* PL. 'lightning' — ON *mjǫllnir* must reflect **melƀuni-* (cf. Noreen 1923: 199, 258; *contra* IEW 722). In view of the ambiguity of Lv. dial. *milna* 'hammer of Pērkons' (ME II: 627) and OCS млъниı 'lightning', where the dental has been lost before **-n-* (cf. Endzelīns 1923: 162; Vaillant 1950: 90–91), the evidence for

---

the borrowing context is quite different; in the case of a title, the adoption of a nominative form is to be expected, cf. similarly Turkish *papaz* ← MGr. παπάς 'priest'.

65 Lt. /š/ is a poor phonological match for German /z/; furthermore, a computer-assisted search of the LKŽ did not yield any Germanic loanwords containing Lithuanian nasal vowels.

66 West Žemaitian *lẽ·iš^{a}s* 'lentil' and *łeyśiey* 'lens' in Szyrwid (ALEW^{2} s.v. *láišis*) apparently show sporadic dial. **ę > ei* (Zinkevičius 1966: 137). The forms cited under *láišis* in LKŽ (the factual basis for the acute set up here is unclear) must partially reflect the same form with regular hardening of /l'/ as in dial. (Zietela) *lãšis* 'lentil' .

67 ЭССЯ (XV: 64) claims that the word is limited to South Slavic, apparently interpreting the Old Russian examples (cf. СДРЯ 100; СДРЯ 11–14 IV: 489) as Church Slavic loans. To my mind, it is very unlikely that an East Slavic scribe would 'nativize' CS лаща as ⟨лача⟩ without actually being familiar with the word. Sergejus Tarasovas suggests to me that the dial. *ля́ща* (Orjol, Kaluga) cited by Даль^{2} (II: 292) may be an incorrect transposition of a local **ля́*[ɛ]*а* (= */ля́ча/, cf. ДАРЯ I: No. 48) influenced by the Church Slavic spelling.

a voiced dental is limited to the Prussian *Elbing Vocabulary*.[68] Since there are some other examples of unexplained voiced stops, such as *girmis* · made 'maggot' (= Lt. dial. *kirmìs* 'worm'), this evidence must be treated with care. Taking it at face value, we may see it as evidence for a voicing alternation.

▸ ? '**nettle**'. Lt. *notrẽ* (ACC.SG. *nõtrę*), dial. *noterẽ* (ACC.SG. *nóterę*), Lv. *nâtre*; Pr. E *noatis* 'nettle'; OIr. *nenaid*, MW *dynat*, *danat* 'nettle' (< **ninati-*)[69] ~ OSw. *nätla*, *nätsla*, OE *netele*, OHG *nezzila* 'nettle' (cf. underived Fårö Gutnish *nate*, *nata*) (Philippa et al. III [2007]: 418; Matasović 2009: 291; Derksen 2015: 337) — Despite the difference in meaning, it is probable that Pl. *nać*, Sln. *nât* 'vegetable tops' also belong here.[70] This semantic shift would imply that nettles were either eaten or given as fodder. On the basis of the East Baltic forms, Specht (1935: 253; followed by REW II: 201) has reconstructed an archaic *r*-stem, but as the Slavic and Prussian *i*-stems cannot be explained on this basis (cf. ALEW 815), it is preferable to view the East Baltic forms as innovative (on the suffixation, see Skardžius 1941: 305–306).

In principle, the Baltic forms could reflect a root **neh$_2$t-*, while Celtic would be consistent with **nh$_2$t-* (Zair 2012: 197). Parallels may also be found for the reduplication (see the discussion under 'sedge' on pp. 240–241). The Germanic dental is difficult to explain. Resorting to Kluge's law would be *ad hoc*, since most of the evidence points to an original singleton **t*. Kroonen (2013: 384) has suggested the Balto-Slavic forms were borrowed from Germanic, but this is unlikely in view of the formal discrepancy. Furthermore, Celtic clearly points towards an original **t* (cf. Derksen 2015: 337). If the example is accepted as non-IE, the long vowel in Baltic can be compared with the other examples in 7.5.1.

▸ † '**thrush**'. R *дрозд*, Pl. *drozd*, SCr. *drȍzd* 'thrush' ~ Pr. E *tresde*; ON *þrǫstr* (attested in Þul *Fugla*, cf. Ic. *þröstur* 'thrush'); Lat. *turdus* 'thrush'; OIr. *truit*

---

68 Note that the Slavic reconstruction **muldnijā-* (trad. **mъldni*; Derksen 2008: 333 following ЭССЯ XX: 220) should be corrected to either **milnijā-* (trad. **mьlni*; Mikkola 1908: 123; Matasović 2008: 200) or **mulnijā-* (trad. **mъlni*) — the two are difficult to distinguish. The cluster **-dn-* is based only on East Bel. dial. *маладня́* 'lightning' (thus explicitly Мартынов 1985: 7), a form which is most certainly a hypercorrection in dialects with *-dn-* > *-nn-*, cf. Bel. dial. *малання́* (ДАБМ No. 311; see Касаткин 1999: 124 and somewhat differently Wexler 1977: 149).

69 The alternative reconstruction **nenati-* (Pedersen 1909: 186; Schrijver 1995a: 49) is less probable, as this should have become ***nanati* by Joseph's law.

70 In East Slavic only the derived R dial. *нати́на* (СРНГ XX: 219), Bel. *наці́на*, Uk. dial. *нати́ня*. The usually cited Uk. dial. *нать* appears to be confined to the easternmost Carpathian dialects (АУМ II: No. 324), so it is plausible that it represents a loanword from Slovak dial. *nať*. For other, less convincing, accounts of the Slavic word, see ЭССЯ XXIII: 186–187.

'starling'[71] (Ernout/Meillet [1951]: 708; Matasović 2009: 392; ERHJ I [2016]: 200; Matasović 2020: 335; Stifter forthc.) — With regard to the nature of the dental, Lt. *strãzdas*, Lv. *strazds* 'thrush, starling' are ambiguous, as they show an additional *s*-, possibly due to anticipation of the second sibilant (Walde/Pokorny I: 761; LEW 920).[72] Apart from the initial **d*- in Slavic, the correspondence with Norse is precise. The Prussian vocalism is surprising; there is an outside possibility that it nevertheless represents **trasdḗ* (cf. Trautmann 1923: 327).[73]

Latin *turdus* is most straightforwardly explained as the reflex of the zero-grade **trsd*-. This zero-grade could also be continued in the Germanic diminutive OE *þrostle* 'turdella', MHG *trostel* 'merula' (< **trust-lō*-; Kluge/Seebold 218), providing the position of the *-r*- has been restored on the basis of the full grade (Kroonen 2013: 545). Sln. *drȇzg* (dial. *drȇzd*) and SCr. dial. *drȍzg* 'thrush' (Skok I: 443) result from of a semi-regular dissimilation (Solmsen 1904: 578–579; Endzelīns 1911: 54–55, fn. 3). Dissimilation has also been suggested to account for the variant **trusk*- attested in OE *þrysce** (attested *þryssce*), OHG *drosca* (EWAhd II: 803), which might be preferable to the suggestion of an inner-Germanic suffixal formation (Kluge/Seebold; Kroonen loc. cit.).[74] Problematic

71 Arm. *tordik* (Hamp 1978: 188, 1981: 88; de Vaan 2008: 634; Kroonen 2013: 545) is evidently a learned creation based on Italian *tordo* (V. Petrosyan on en.wiktionary.org, s.v. unnnhl [8 April 2020]; Thorsø forthc.).

72 As a parallel, note the Old English by-form *strosle* 'blackbird' (Kitson 1997: 485; OED³ s.v. *throstle*).

73 The grapheme ⟨e⟩ in the Elbing Vocabulary only rarely stands for /a/, and usually in non-initial syllables (e.g. Pr. E *pepelis* ~ III *pippalins* ACC.PL. 'bird', E *pirsten* 'finger', cf. III *pīrstans* ACC.PL.). However, a potential parallel is found in Pr. E *wessis* ~ Lt. *vãžis*, dial. *važỹs* 'one-horse sleigh' (PKEŽ IV: 232; for the translation, see Trautmann 1910: 460). This is uncertain, however, as this word may also show the reintroduction of the vowel from the verbal root seen in Lt. *vèžti* 'transport'.

74 Old English *þræsce* is normally cited here, but as a hapax in the Corpus Glossary finding no concrete support in either later English or elsewhere in Germanic, its reliability is questionable. It seems more probable that the dialectal *thresh* (Oxfordshire, Berkshire), in which OED (s.v. *thrush* n.[1]) would see a continuation of this **þræsce*, contains a regional continuation of OE *þrysce**. Perhaps it is a Kentish form (with *y* > *e*, Campbell 1959: 122–124) which has spread beyond its original geographical zone; compare similarly dial. (Sussex, Essex) *sherve*, *sharve* 'service tree' (< OE *syrfe**, attested OBL.SG. *syrfan*; OED³ s.v. *serve* n.[1]) and perhaps SW dial. *rex* 'rush' (differently see OED³ s.v. *rush* n.[1]). Note that the Old English word is glossed as *truitius* (cf. also the similar gloss *þrisce · trutius*) which Kitson (1997: 484) would see as a "corruption" of Latin *turdus*. Far more likely, this is a Latinization of Irish *truit* 'starling' (Suolahti 1909: 52, fn. 1). As for the semantics, it is worth mentioning that Lat. *turdus* is twice glossed as OE *stær* 'starling' (Lacey 2013: 66).

are the British forms OBret. *tra*[*s*]*cl* (modern *drask*, *draskl*), MoW *tresglen* 'thrush'. Stifter (forthc.) equates the apparent suffix *-*sk*- in the above forms with that found in MW *mwyalch* 'blackbird' and *alarch* 'swan'. In his opinion, this would favour a non-IE origin.

As has long been recognized, OIr. *truit* 'starling' can reflect an earlier **trozdi*- (Zupitza 1900: 233; cf. Brugmann 1897: 691). However, the British equivalents MW *trydw*, MBret. *tret* 'starling' cannot, which has led them to be analysed as Goidelic loanwords (Walde/Pokorny I: 761; Stifter forthc.). The Welsh and Breton vocalism is consistent with the reflex of **u* with *i*-affection (after the plural), while the unaffected vowel is preserved in the Old Breton gloss *trot* · strution.[75] In any case, a Goidelic loanword seems preferable to assuming an *ad hoc* "expressive gemination" in Old Irish (de Bernardo Stempel 1999: 274; implicitly Matasović 2009: 392).

Ernout/Meillet (p. 708) claim that it is futile to reconstruct the original form of this word. However, an ablauting **trosd*- : **trsd*- accounts for the Baltic, Germanic and Latin and Old Irish data without any serious problems.[76] The remaining evidence for irregularity is the initial *d*- in Slavic, but it is possible that this has arisen through assimilation, as has undoubtedly occurred in MW *drydw* and MoIr. *druid* 'starling'.[77] As the Latin form is more easily explained starting from an Indo-European ablaut variant, while all of the irregular developments can be accounted for within the individual branches, I do not think there is any truly compelling evidence for a non-IE borrowing.

6.2.3.1 Conclusion

The certain and possible evidence for voicing alternations involving dentals is collected in Table 10, overleaf (see p. 185 for help reading the table). Forms which do not provide relevant data are presented in light grey. Shaded cells indicate reflexes of a voiced or voiced aspirated dental.

---

75 The vocalism of OCo. *troet* is unclear as ⟨oe⟩ in other cases represents an inherited vowel sequence (e.g. *hoern* 'iron' = MW *haearn*; *moelh* 'blackbird' = MW *mwyalch*); we apparently have to assume contamination with e.g. OCo. *hoet* 'duck' (= MW *hwyat*).

76 Hamp's (1981: 88) insistence on **d*^{h} is not necessary, as Winter's Law was blocked by an intervening **s* (Kortlandt 1988: 394); and such a reconstruction is contradicted by Germanic.

77 It has often been claimed that the word for 'thrush' is ultimately onomatopoeic (Suolahti 1909: 53; Булаховский 1948: 112; EWAhd II: 803; Kluge/Seebold 218), but this does not seem certain to me. РЭС (XIV: 363) notes SCr. *drsk!*, representing the sound of the mistlethrush, but it is possible that this onomatopoeia partly derives from the name of the bird itself.

TABLE 10 Possible examples of alternations involving dentals

	Baltic	Slavic	Germanic	Elsewhere
'drone'	*tran-	*trant-	*dʰrV̆n-	Gr. *tʰrV̄n-
'reed'	*truḱ-	*trust-	–	Gr. *tʰrus-
? 'lentil'	?*lenḱ-	*lent-i̯-	?*lens-	?Gr. *lntʰ- Lat. *l(e)nt-
? 'lightning'	? *meld(ʰ)-	*mlT-ni-	*melt-uni-	Celt. *melt-
? 'nettle'	*nât-	*nât-	*nad-	Celt. *ninat-

It is interesting that the examples in this section do not show a similar behaviour to the examples of *$k \infty *g^{(h)}$ given in 6.2.1. Only the word for 'lightning' possibly shows the same distribution, with Baltic voiced *D contrasting with Celtic *t; however, as discussed above, this is based on rather tenuous evidence. The clearest examples here involve Greek; specifically, in two or three cases, we find Greek *t^h as opposed to *t elsewhere. Since in Greek we actually find a voiceless stop, it is unclear whether back-projecting it to IE *d^h would be anachronistic: perhaps, rather than a 'voicing' alternation, we are dealing with an 'aspiration' alternation. Such alternations are well-known in Greek words of presumed foreign origin, e.g. ἄνηθον ~ Aeol. ἄνητον 'dill' (Furnée 1972: 187–193). Against this conclusion, we can note that Germanic indeed does show a reflex of *d^h in the word for 'drone'. On the other hand, note the word for 'turnip', discussed in the following section, which might show a comparable 'aspiration alternation'.

6.2.4 *Alternations Involving Labials*

In two of the words discussed above (see 6.2.1), we have observed an alternation *$b^{(h)} \infty *p$ occurring alongside *$g^{(h)} \infty *k$, cf. Lt. *gul̃bė* ~ Kash. *kôłp* 'swan' and R *ерaб* ~ Lat. *carpinus* 'hornbeam'. Examples of an independent alternation *$b^{(h)} \infty *p$ not associated with a parallel velar alternation are in fact comparatively few, and the only certain cases constitute rather widespread *Wanderwörter*:

▸ **'hemp'**. R *конопля́*, Pl. *konopie*, SCr. *kònoplja* 'hemp' ~ OE *hænep*, OHG *hanaf*; Gr. κάνναβις 'hemp' (Schrader/Nehring I [1923]: 441; Huld 1990: 406–407;

Matasović 2013: 89; Kroonen 2013: 209; etc.) — In Baltic, we have Lt. *kanãpės*, Lv. *kaņepes* and Pr. E *knapios*, which are usually considered to be Slavic loans (e.g. Berneker I: 361; ME II: 156–157; LEW 214; Levin 1974: 96; Smoczyński 2018: 482). On formal grounds, cognancy is equally possible (Būga 1913: 255–256; PKEŽ II: 231). While there is some evidence for hemp having been used in Lithuania during the 1st millennium CE (Gimbutas 1963: 117; Grikpėdis/Motuzaitė Matuzevičiūtė 2020: 165), it is uncertain whether this evidence is early enough to rule out a Slavic origin for local hemp production. A form with **p* is likewise widespread in Romance, cf. Italian *canapa*, Romanian *cânepă*, attested since Late Latin (cf. FEW II: 213–214).

On the authority of Herodotus, Greek κάνναβις is traditionally considered a loanword from Scythian or Thracian (cf. Schrader/Nehring I: 441; Frisk I: 779), although this has no concrete linguistic basis. In Ossetic, which would be closest to the supposed Scythian donor language, we find Oss. I *gæn*, D *gænæ* 'hemp', which probably implies **kanā-* without the labial (cf. Abaev 1958: 513).[78] Elsewhere in Iranian, a form **kanafa-* seems to be suggested by Khotanese *kaṃha-* 'hemp' and NP *kanaf* 'flax cord' (Steingass 1892: 1055),[79] while the NP variant *kanab* 'hemp (seed); hempen rope' (idem: 1052) would imply **kanapa-*.[80]

The word for 'hemp' is widely recognized as a *Wanderwort* of indeterminate origin, and the precise source of the various *p*-forms in Europe is difficult to establish. The ultimate origin of the word has been seen in the Near East, cf. Syriac *qnp'* /qenpā/ 'hemp (for making ropes)', and Akkadian (Neo-Assyrian) *qunnabu*, '(possibly) the flower or seed of the hemp'. The latter would predate the Greek attestations, although it is hardly the original source (note that Sumerian **kunibu* is a ghost, cf. Barber 1991: 38).

▸ [**'turnip'**. Lt. *ropė́*, Lat. *rāpum* 'turnip' ~ Gr. ῥάφανος 'cabbage, radish';[81] ?MW *erfin*, Bret. *irvin* 'turnip' — See the discussion on p. 237.]

78 It seems at least possible that this could have developed via **kanapā-* > **kanaba* > **kanba* (syncope, cf. Cheung 2002: 55–56), then by (irregular?) metathesis to **kabna* > **kan(n)a* (cf. Oss. I *kʷynæg*, D *kunæg* 'meagre, small' < **kabna*, Cheung 2002: 30). In any case, the initial *g-* is irregular, and has no regular origin.

79 The vocalism of Northern Kurdish (Kurmanji) *kinif* 'hemp' (cf. Цаболов 2001: 554) is unclear, but the form must be borrowed, as intervocalic **f* has regularly given *-v-* in Kurdish, cf. *nāv* 'navel' < **nāfa-* (Цаболов 2010: 32; M. Kümmel p.c. December 2022).

80 A Proto-Iranian **p* would also be suggested by Buddhist Sogdian *kynp'* (Gharib 1995: 203), perhaps meaning 'hemp' or 'flax', provided this is not an independent loan from Syriac (Henning 1946: 724). Bailey (1979: 51–52) quotes a MP (Pahlavi) *kʾnb* that I have been unable to verify. If reliable, it would seem to suggest **-b-* (cf. Peyrot 2018: 270). Arm. *kanepʿ*, *kanapʿ* 'hemp' appears to be an Iranian loan, but its exact source is unclear.

81 The Greek variant with -π- (cf. Beekes 2014: 61) rests on extremely doubtful evidence: (a)

▸ ? **'furrow'**. Lt. *biržė* 'row, furrow; timber tract; border mark', Lv. *bìrze* 'furrow, strip of a field' ~ Lat. *porca*, OHG *furh*, *furuh*, MW *rych* 'furrow' (on the Slavic equivalent, see p. 224) — The similarity of these forms has been noted by Machek (1968: 65) and Holzer (1989: 51–54), and the example remains one of the most attractive of Holzer's "Tememe" etymologies, as the formal correspondence **prḱ-* ∞ **bʰrǵʰ-* is precise aside from the difference in voicing. What further speaks in favour of Holzer's interpretation is that **prḱ-* has some potential IE comparanda.

LIV (475) sets up a verbal root **perḱ-* 'graben, aufreißen'. On further inspection, however, it turns out that the reconstructed semantics are based almost entirely on the word for 'furrow'. The only comparandum attesting to a verbal root is Lt. *peršéti* 'to itch', while the other nominal formations are uncertain. The Rigvedic *párśāna-* (3×) is of uncertain meaning: it probably refers to a low place, but may mean 'valley' or 'plateau' (cf. KEWA 228–229; Jamison *Commentary* VII.140.5). Aside from this, the only evidence is Lt. *pró-perša* (*pra-peršà*) 'thawed patch in ice; break in the clouds; etc.',[82] but this, like *núo-perša* 'infertile patch of land', *ìš-perša* (Kupiškis) 'deep rut in a road' might well be derived from *peršéti* in a secondary sense, cf. *nu-*, *iš-peršéti* 'go bad, spoil'. We may conclude that the evidence for the verbal root rests on the Lithuanian word for 'to itch', which is semantically remote.

If the IE etymology can be abandoned, we may consider a non-IE origin for the whole group, which would eschew the need for a "Tememe" source or other IE substrate. In this case, 'furrow' can be considered an example of a **p* ∞ **b*$^{(h)}$ alternation. It is, however, a little troubling that none of the examples of a **k* ∞ **g*$^{(h)}$ alternation discussed in 6.2.1 were affected by satemization. If the word is non-IE, it must have been loaned extremely early, which may also be seen in the regular reflexes of syllabic **r̥*.

⟨ῥαπάνια⟩, attested in a Hellenistic period papyrus. This, and other examples of confusion between stops in Egyptian papyri, can plausibly be attributed to Coptic first-language interference (Holton et al. 2020: 187); (b) Athenaeus (*Deipnosophists* IX, Chapter 8) tells us that Glaucus, apparently the author of a cookery book, wrote ῥάπυς for ῥάφυς (meaning βουνιάς 'rapeseed'). Neither form is otherwise reliably attested (LSJ s.v.).

82 Lt. *praparšas*, known only from Szyrwid, is typically adduced here (Walde/Pokorny II: 46; Fraenkel 578; IEW 821; LIV 475). The gloss 'Graben' in all these sources (thus supporting the sense 'to dig'), is based on Szyrwid's *row*, *fossa* (SD 268$^{b}_{35}$). In the first edition of the dictionary, however, the word glosses Polish *iaskinia*, *odchłan*, *przepáść* (ALEW 102), suggesting a sense 'chasm, abyss'. As none of these senses appear to have been recorded elsewhere, I am led to wonder whether Szyrwid was unsuccessfully attempting to render a sense such as 'gap in the ice' in Polish.

▸ ? **'pigeon** (2)'. Lt. *balañdis*, Lv. *baluôdis* 'pigeon' ~ Lat. *palumbēs* 'wood pigeon'[83] — Both words are traditionally explained as derivatives of colour terms (cf. Schulze 1910: 799–800). In Baltic, the root is assumed to be that of Lt. *bálti* 'whiten' (Skardžius 1941: 101; LEW 31);[84] Karaliūnas (1993: 110) assumes an original colour adjective **balandas* 'whitish' (cf. 'white-marked one', Levin 1992: 86). Derksen (2015: 78) questions the derivation from 'white' on semantic grounds as, according to him (after Levin loc. cit.), "whiteness is not a natural colouring in pigeons". More importantly, an adjectival suffix **-anda-* would be completely unparalleled and therefore *ad hoc*. Lat. *palumbēs* is usually derived from the root of *palleō* 'be or grow pale' (e.g. Walde/Hofmann II: 242), cf. Gr. πέλεια 'wild pigeon' ~ πελιός 'black and blue, livid'. The first syllable is also reminiscent of Pr. E *poalis* 'pigeon'. If we start with 'grey', the semantic motivation makes some sense;[85] the Latin second syllable could have been influenced by *columba* (Lockwood 1990: 262–263; de Vaan 2008: 126).

Naturally, if we compare Lt. *balañdis* with Lat. *palumbēs*, both root etymologies would need to be abandoned. Due to the lack of morphological transparency on both sides of the equation, this might be justified. However, the irregularities are not limited to the initial stop; there is also a mismatch between the stem-final *-b-* in Latin as opposed to Baltic *-d-*. One way out is to assume, again, that the Latin word has been influenced by *columba*, although then one could question how exactly this etymology is preferable to the traditional explanation, which also demands the assumption of such a contamination. In defence of the new etymology, it seems more straightforward to assume contamination starting from a disyllabic **palond-* rather than, with Lockwood, from a more basic **palēs*.[86]

Klingenschmitt (1982: 165) compares Lat. *palumbēs* with Arm. *aławni* 'pigeon', reconstructing **plH-bʰ-nih₂-* (in his notation **pl̥h-bʰ-niə₂*), implying the

83 Apparently here also Oss. I *bælon*, D *bælæw* 'domestic pigeon' (Абаев 1965: 17; Weber 1997). Due to the *-l-*, the Ossetic word is likely to be a loanword. It is unclear whether Baltic could plausibly be the source, as there is no other clear evidence of contact, and no obvious historical scenario. According to Sasha Lubotsky (p.c. April 2021), the Iron suffix *-on* (< **-ān-*) might be equated with Baltic **-and-* through regular **a* > **ā* before a consonant cluster and subsequent loss of **d*. The Digor variant is of unclear formation.

84 The existence of the frequently cited *bãlas* 'white' (known only from Juška) is perhaps questionable, see Jakob forthc. b.

85 Cf. Russian *сизя́к* 'feral pigeon' < *си́зый* 'dark bluish-grey'; Oss. ID *æxsīnæg* 'wild pigeon' < (Digor) *æxsīn* 'dark grey' (Абаев 1958: 220–221).

86 Alternative, but no less *ad hoc* explanations would be to assume a dissimilation **b–b* > **b–d* in Baltic, or a suffixed Latin **palond-u̯-*.

application of Thurneysen's law of nasal metathesis in Latin (cf. Martirosyan 2008: 29). Such a preform would yield Latin **plamb-* rather than **palomb-*, and more importantly, would force us to disassociate the suffixes of *columba* and *palumbēs*, which seems quite unsatisfactory. Assuming Armenian has a derived *n*-stem, we may instead start from **palab-*, i.e. a variant without a nasal (cf. section 6.1). This is speculative, however, as the Armenian word is open to interpretation (see Batisti 2021: 208–210 with lit.).

▸ **† 'hollow'.** Slk. *dúpä* 'den, burrow', Sln. obs. (Pleteršnik I: 184) *dúpa* 'die Erdhöhle' (= Pl. *dupa*, Bg. *дупе* 'arse') ~ Lt. *daubà* 'ravine; (PrLt.) den, burrow' (Kuiper 1956: 223; 1995: 71–72; Schrijver 2001: 420; Philippa et al. I [2003]: 569 s.v. *diep*; Matasović 2013: 96; Derksen 2015: 144) — This word family is routinely quoted, mainly by members of the "Leiden school", as an example of a substrate word. Kuiper's main line of argument was built on the presence of numerous variants within Germanic, where root final **-b-* seems to alternate with **-p-*, **-bb-*, **-pp-* and **-mp-* (thus ON *dúfa*, *dýfa* 'dip (at a christening)'; Go. *diups* 'deep'; MDu. *dobbe* 'water pit, pool'; Nw. *duppe* and MLG *dumpeln* 'dip', respectively). This approach has been criticized by Kroonen (2011a: 255; 2011b: 127–129), who has convincingly argued that the variation can be more plausibly explained as a result of various analogies after Kluge's Law.

It seems likely that the 'nasal infix' supposed for Lt. *duñblas*, Lv. *duñbla* 'mud, sludge' (LEW 108–109; Smoczyński 2018: 263; ALEW 276) is also an illusion.[87] Rather, the *-b-* in these forms is epenthetic. This is possibly suggested by the forms *dumłas* (SD 64^{b}_{17})[88] and *dumlelus* (ACC.PL., Daukantas 1846: 67; see LKŽ s.v. *dumlas*), and certainly by Lv. dial. (Vidzeme) *dumûksnis* 'marsh' (cf. Prellwitz 1909: 387; Schulze 1910: 791; ME I: 514). In general, there is a fairly consistent semantic distinction between the two word-groups. Almost all words containing a nasal mean 'mud' or 'marsh', while words lacking the nasal mean 'valley, hollow'.[89] The latter group are transparently derived from the verbal

87 The latter two sources point specifically to the Lt. 3PRES. *duñba* as the source of the forms. The antiquity of this presentic formation cannot be proven, as nasal presents are productive in Lithuanian denominal verbs of the shape **TVT-* (where *T* = any stop, see Villanueva Svensson 2010: 206–208), and moreover, ME (I: 509) reports a plain thematic *dubu* for Latvian.

88 But note that Szyrwid also has ⟨dumbłas⟩ (SD 120^{b}_{19}).

89 Compare, on the one hand, Lt. *duñblas*, Lv. dial. *duñbla* 'mud', Lv. *duñbrs* 'boggy; marsh' (the suffix in Lv. dial. *duñbêris* 'muddy pit; puddle' is probably secondary), and on the other hand Lt. *dubùs* 'hollow, concave', Lv. dial. (Vārkava, ME I: 509) *dubums* 'tree hollow', Lt. *daubà*, Lv. dial. (ME I: 443) *daũba* 'ravine'. The two roots do seem to have influenced each other, however, cf. Lt. dial. *dumbrà* 'deep point in a river; pond' vs. Lv. dial.

root in Lt. *dùbti* 'sink down, become concave', which is further cognate with Go. *diups* (< **dʰeubʰ-nó-*) 'deep' and OIr. *domain* 'deep'.[90]

In view of the large number of derivatives and extensive IE-like ablaut in the root **dʰeubʰ-*, it seems more probable to me that it is inherited, despite the limited distribution. As a result of this, the Slavic forms with *-p-*, must either be unrelated or explained as the result of a secondary deformation. As I have identified no motivation for the latter, I would prefer to simply separate the forms.

▸ † '**post** (1)'. Lt. *stulpas* 'post, pillar', Lv. dial. *stùlps* 'pillar, leg of a boot'; OCS стлъпъ 'pillar, tower' ~ R *столб* 'post, pillar', Sln. obs. (Caf *apud* Pleteršnik II: 578) *stółb* 'Pfahl'; ON *stolpi* 'post, pillar' (> ME *stulpe* 'stake, post', MDu. *stolpe* 'small beam') — Vasmer (REW III: 18) rejected earlier proposals (Meringer 1909: 200; Stender-Petersen 1927: 279–281) to derive the Balto-Slavic words with *-p-* from Germanic, although he does not present any arguments. A point in favour of the loan etymology is that the Balto-Slavic *p*-forms are largely limited to the meaning 'post, pillar', while with *-b-* one finds archaic-looking derivatives such as Lt. *stulbti* 'be stunned', and Bg. *стѝлба* 'staircase, ladder', SCr. *stùba* 'step, stair'. On the other hand, the word is scarcely attested in Germanic, and one could seriously consider deriving the Norse word from Slavic (Tamm 1881: 31; dismissed, again without argumentation, by de Vries 1962: 551).[91] The complexity of the analysis makes it difficult to draw any clear conclusions.

6.2.4.1 Conclusion

The certain and possible evidence for voicing alternations involving labials is collected in Table 11, overleaf (see p. 185 for help reading the table). Shaded cells indicate reflexes of a voiceless labial. The cover symbol **B* stands for **b*$^{(h)}$. Forms which do not necessarily provide relevant data are presented in light grey.

---

duburs 'deep and wide spot in a river'. Additionally, some Latvian words seem to belong with the former root, but lack an *-m-*: *dubļi* 'muck; mud', *dubra* 'swamp, bog'. Might these reflect **dūb-* < **dumb-* with shortening before a *-CR-*cluster (cf. Derksen 2007: 44)?

90 Often adduced are To. A *tpär*, B *tapre* 'high'. However, the 'Tocharian Grassmann's law' (Winter 1962), if valid, would predict To. B **tsapre*. The original meaning 'deep' has been supported by the translation of To. A *top*, B *tewpe* as 'mine' (Adams 2013: 330). However, Imberciadori (2022) has argued that this word should instead be translated 'heap', which makes the comparison unattractive.

91 In any case, the root connections with Nw. *stelpe*, MDu. *stelpen* 'hinder' or with Lt. *stelbti* 'overshadow' are not compelling.

TABLE 11 Possible examples of labial alternations

	Baltic	Slavic	Germanic	Elsewhere
'swan (2)'	*Gulb^{h}-	*kulp-	–	
'hornbeam'	*skrâB-l-	*GrâB-r-	–	Lat. *k(r)arp-
'hemp'	[*kanap-]	*kanap-	*kanab-	Gr. *kannab-
'turnip'	*râp-	*rêp-	*rāP-	Gr. *rab^{h}- ?Celt. *arB-
? 'furrow'	*Brǵ^{h}-	?*BorsD-	*prk-	It.-Celt. *prk-
? 'pigeon'	*Baland^{h}-	–	–	Lat. *palomB-

Beside the words for 'swan' and 'hornbeam', which show $*p \infty *b^{(h)}$ alongside $*k \infty *g^{(h)}$ in the same word, there are two other potential examples showing a similar distribution, although neither of these are certain. The word for 'furrow', if loaned from an unknown source, would be the only example of a voicing alternation predating satemization. As a result, whether it represents a manifestation of the same voicing alternation cannot be considered certain. The remaining words appear to show the opposite pattern: it is notable that both 'hemp' and 'turnip' are widespread words associated with agriculture, and it is likely that they spread as *Wanderwörter*. The word for 'turnip' might constitute an example of the 'aspiration' alternation observed in 6.2.3. On the other hand, the Celtic, and potentially also Germanic, comparanda point to an underlying $*b^{(h)}$, yet it is by no means certain that the divergent stops in Celtic and Greek can be equated with one another (as virtual $*b^{h}$), and it is possible that they represent two unrelated phenomena — a specifically (pre-)Greek 'aspiration alternation' and a specifically (pre-)Celtic voicing.

6.2.5 *Baltic* *ž ∞ *Slavic* *s

▸ **'oats'.** Lt. *ãvižos*, Lv. *àuzas* 'oats' ~ R *овёс*, Sln. *óvəs*; Lat. *avēna* 'oats' (Ernout/Meillet [1951]: 56; ?Pisani 1968: 14; Huld 1990: 404; FEW XXV [2002]: 1213; Oettinger 2003: 189; de Vaan 2008: 64–65) — The relationship between the Baltic and Slavic words is irregular, suggesting the word entered the two branches

independently. Reconstructing a suffix *-*s*- in Slavic (thus Derksen 2008: 384) is *ad hoc*, as there not appear to be any other plausible cases of *-*s*- as a denominal suffix (cf. Vaillant 1974: 659).[92] Moreover, the Latin vocalism also precludes the reconstruction of a common pre-form. De Vaan's assumption of an underlying palatovelar and **aweksnā*- for Latin is potentially anachronistic. As Huld points out, if we are dealing with a non-IE loanword, "a spirant of indeterminate voicing" would account for the facts. Lat. *avēna* 'oats' could equally reflect **au̯e*(*T*)*s-n*- (where **T* can be essentially any stop, although **ts* or **s* would be most probable for our purposes). For further discussion, and on the question of Prussian *wyse* 'oats', see pp. 239–240.

▸ **? 'fishing trap'**. Lt. *várža* 'fishing basket', Lv. *varzi* 'Setzkörbe' (Lange 1773: 378), dial. *vaŕza*[2] 'fishing weir' (ME IV: 481) ~ R *вéршa*, Sln. *vŕša* 'fishing basket' (< **virs*- + **-jā*-; trad. **vьrs* + **-ja*) (cf. Pronk/Pronk-Tiethoff 2018: 295) — Existing etymological solutions either separate the two words, linking the Slavic forms with OCS врьхъ 'top' (REW I: 109), or assume a suffix **-sį̄*- for Slavic (Persson 1912: 505; Trautmann 1923: 355). The Baltic forms look related to the verb Lt. *veržti* 'tighten, tie up', Lv. *vîržît* 'direct, steer' (cf. *at-virzīt* 'untie', ME I: 211), but this has not been generally accepted (cf. РЭС VI: 351; ALEW 1384).[93] In view of the parallelism with the word for 'oats', above, it is tempting to derive these words from a non-Indo-European source. On the other hand, it is unclear to what extent it is justified to separate the words for 'fishing basket' from Latvian senses such as *vaŕza*, *važa* 'tangle, confusion' (ME IV: 481–482), which clearly belong with the verbal root (cf. Lv. *vaŕzât* 'plait together, tangle'). In addition, the difference in vocalism is striking; this sort of vowel alternation is perhaps more easily explained as the result of Indo-European ablaut than through parallel borrowing (compare, with the opposite distribution, Lt. *biržė* ~ OCS бразда 'furrow' on p. 224).

▸ **? 'ploughshare'**. Lt. *lẽmežis* 'ploughshare' ~ CS (Bes.) лемешь* (SJS II: 112) 'plough', R *лéмех*, dial. *лемéш*, SCr. *lèmeš* 'ploughshare' — In view of its *-s*-, perhaps Lv. *lemesis* 'ploughshare' is a loan from East Slavic. The *-s*- could be a hypercorrection after the oblique cases (e.g. *lemeša* GEN.SG.), cf. *vìksne*[2], GEN.PL.

92 For the deverbal suffix, cf. OCS гласъ 'voice, speech' ~ глаголати 'speak, proclaim'; CS кѫсъ* 'bit, crumb' ~ Lt. *kąsti* (*kánd*-) 'to bite'; OCS смѣхъ 'laughter' ~ смиꙗти сѧ 'to laugh'.

93 Snoj (2003: 836) considers the word for 'heather' (see below) to be related, and the word for 'fish trap' to originally have meant 'something woven (from heather)'. A fishing basket woven from heather does indeed appear to be found in the Highland Folk Museum, but I cannot verify whether such a tradition could have existed at an appropriate time in central Europe. See the doubts in РЭС (VI: 351–352), where all other etymological comparisons are also considered doubtful.

vīkšņu 'cherry tree' ← R *ви́шня* (cf. Būga 1922: 175–177 and also dial. *lemešs*, EH I: 733). Note that Lt. *lẽmežis* has itself been interpreted as a Slavic loan (Gołąb 1982: 130;[94] LKŽ s.v.; ERHJ I: 546); and while its limitation to a narrow group of Šiauliškiai dialects rouses suspicion, this is not sufficient to confirm or deny this proposal. In Slavic, the most common variant is **lemeše-* (trad. **lemešь*), continuants of which are found in every Slavic language. In addition, forms are found with a final *-ž*, but these look secondary, being largely limited to South Slavic: Sln. *lémež*, SCr. dial. *lèmež* (PCA XI: 327), Čak. *lemèž* (ERHJ I: 546), Bg. *лемѐж* (cf. the data in ЭССЯ XIV: 108–110). Perhaps one could assume the secondary influence of the deverbal noun suffix **-eže-* (trad. **-ežь*), which enjoyed a certain productivity in South Slavic (Berneker I: 700; cf. Vaillant 1974: 506).

Furthermore, some forms seem to lack the initial **l-*: CS емешь (Miklosich 1865: 1157), Bg. dial. *емѐш* (БЭР I: 495),[95] SCr. dial. (Montenegro) *jèmlješ* (RJA IV: 587), R dial. (N) *о́мех*, *оме́ш* (and variants, СРНГ XXIII: 198–199, 201–202; Мызников 2019: 556). Derksen (2015: 278) has considered the variant with **l-* the result of a secondary contamination with the root **lemH-* 'to break'. This is rather difficult to accept: forms with **l-* are much better represented in Slavic and the only forms found in Baltic. Provided the latter are not all Slavic loanwords, it would be highly improbable that the contamination could have occurred independently in both branches.[96] An interesting proposal is put forward by Bańkowski (2000 II: 19–20), who assumes contamination with a Proto-Slavic **lemęzi-* (**-že-*; trad. **lemęz/žь*) represented by Pl. dial. *lemiąże* PL. (Sł. Warsz. II: 714), OCz. *lemiez*, Sln. *lę́mez* 'rafter'. The assumption is that the latter would have been used in the sense 'plough shaft'. The weakness of this theory is that neither word is attested in this meaning, but such a confusion does indeed appear to have occurred in some forms meaning 'ploughshare': cf.

94 Cited according to the *Lithuanian Etymological Dictionary Database* (available at etimologija.baltnexus.lt, accessed 9 November 2023), s.v. *lẽmežis*.

95 БЕР claim that the development of /lʲ/ to /j/ is a typical dialectal phenomenon. It is true that around Vraca (where *емѐш* is recorded), we also find e.g. *пойѐ* for *полѐ* 'field' (БДА Ф 109); however, here we are dealing with a reflex of older */lj/, and not */l/, and the authors of БЕР do not quote any evidence for this supposed dialectal change.

96 It is notable that the given verb is attested (almost) exclusively in the *o*-grade in Slavic. Despite Schuster-Šewc (816; cf. ЭССЯ XIV: 113, 200), it seems unlikely that USrb. *lemić* 'to break', attested in some older sources beside *łomić* and corresponding to LSrb. *łomiś*, is a "Proto-Slavic archaism". It is most probably due to internal analogical processes. Similar considerations apply to the Serbo-Croatian iterative *lijèmati* 'beat, thrash' (RJA VI: 64).

OPl. ⟨lyemyąszem⟩ INST.SG. (SSP IV: 19), Kash. *lemiąż*, Slk. dial. (*apud* ЭССЯ XIV 109) *lemez*.

Kalima (1950) also considers the Slavic **l-* to be secondary, and interprets the whole family as an Iranian loanword, comparing Persian dial. *amāč*, *amāǰ* 'plough' (cf. also REW II: 267). This interpretation cannot be upheld, as the Persian word is itself a relatively recent loan from Turkic (cf. Turkish, Uighur *amač* 'plough'; Doerfer 1965: 124). For the same reason, Komi *amį́ś*, dial. (Upper Vyčegda) *ameʒ́*, Udmurt *ameź*, dial. *omeź* 'ploughshare' (Лыткин/Гуляев 1970: 32; Rédei 1986: 64) are likewise hardly of Iranian origin. The Turkic word is already attested since Kāşğarī (11th c. CE), but has a relatively limited distribution, being concentrated in Karluk Turkic and radiating from there into neighbouring sub-branches. Despite this, Starostin, Dybo & Mudrak (2003: 295–296) reconstruct the word for Proto-Turkic, offering the reconstruction **amač*, and further comparing Manchu *anja*, Mongolian *anjis* 'plough'. Regardless of whether one accepts the Altaic theory, the 6th millennium BCE dating for Proto-Altaic supported by Starostin et al. (idem: 237) clearly rules out the possibility of a shared inherited word for 'plough' (cf. Vovin 2005: 75). The overall picture is nevertheless of a cultural *Wanderwort* "with a complicated history" (to quote Helimski 1997b: 121).

At the least, it seems unattractive to separate Turkic (regional) **amač* 'plough' from Permic **amɛʒ́* 'ploughshare' and Slavic dial. **emeše-* (trad. **(j)emešь*) 'ploughshare'. It does not look likely, however, that Turkic could have been the source of either word, as the Permic voiced affricate cannot be explained on this basis, and the Slavic front-vocalism is aberrant. As it is doubtful that the **l-* in Slavic and Baltic can be considered folk-etymological, one may wonder whether this may also be attributed to non-IE borrowing. Rather than a phonetic motivation for an alternation between **l-* and **∅-*, a more reasonable account might be to assume the fossilization of a particle of some kind (such as in MDu. *lomre* 'shade' < Fr. *l'ombre*). However, no parallels of this alternation appear to be found within my corpus.

▸ **? 'heather'**. Lt. *vìržis* 'heather' ~ R dial. *вéрес* (СРНГ IV: 131; РЭС VI: 284), Cz. *vřes*, SCr. *vrȉjes* 'heather' (Machek 1950b: 158–159; Smoczyński 2018: 1680) — Derksen (2008: 516), reconstructs a variant **verska-* (trad. **verskъ*) on the basis of R *вéрескъ*, Uk. dial. (Makowiecki *apud* ЕСУМ I: 353) *верескъ*, although these are most easily viewed as secondary. In Czech dialects, one finds a whole host of obscure variants, including ones with a final *-k*: *vřesk*, *březek*, etc. (see ČJA II: 98; further on the initial *b-*, cf. ČJA V: 442–443). It is quite clear that these cannot all be old, and that we cannot explain the data without assuming convergence with unrelated plant names, cf. Cz. *břečťán* 'ivy', dial. 'heather', *břest* 'elm', dial. 'heather' (similarly R dial. *вéрестъ* 'heather' after *бéрестъ* '(field) elm'?), Cz. dial.

březa 'birch' beside *březek* 'heather'. The source of the final /k/ in R *вéрескъ* remains unclear (cf. *бересклéт* 'spindle tree'? see РЭС VI: 284), but it is unlikely to date back to Proto-Slavic.

The Balto-Slavic forms have long been compared with Gr. ἐρείκη 'heather' (assuming an earlier **wereikā*), on the one hand, and OIr. *fróich*, MW *gruc* 'heather' (< **uroik-o-*), on the other (e.g. Walde/Pokorny I: 273; REW I: 187). As this comparison is phonologically impossible in Indo-European terms, it has been suggested that these forms represent parallel loans from an unattested source (Machek 1950b: 158; Frisk I: 551; Matasović 2009: 431, 2013: 90; van Sluis forthc.). This would imply an underlying **ḱ* and suggest a loan predating satemization, which is chronologically difficult, as there is no agreement even between Baltic and Slavic. Furthermore, the initial **w-* is not ascertained for Greek, and the complete loss of the second-syllable diphthong in Balto-Slavic would be unparalleled. Thus, while the Celtic and Slavic forms potentially share three phonemes, the etymological equation of these forms is dubious.

Standard Latvian *vìrši* 'heather' shows **-s-*. Considering the variation within Slavic, one may argue that the choice of Lt. *vìržis* (and Lv. dial. *vîrži*[2], ME IV: 620) as a comparandum amounts to cherry-picking. Smoczyński (2018: 1680) suggests that *-ž-* may have arisen due to assimilation, or alternatively result from a folk-etymological connection with *veržti* 'tighten, tie up' (thus also T. Pronk *apud* Matasović 2013: 90). Neither of these explanations strike me as convincing, but at the same time, this cannot be classed as a certain example of a voicing alternation. On Žem. *birzdžiai* 'heather', see p. 223.

6.2.5.1 Conclusion

The certain and possible evidence for the alternation **ž* ∞ **š* is collected in Table 12, overleaf (see p. 185 for help reading the table). Forms which do not provide relevant data are presented in light grey. In Slavic, the cover symbol **S_1* may stand for quasi-IE **ḱ* or a cluster **(*T*)*s*. The cover symbol **S_2* may also reflect quasi-IE **s* directly.

At first sight, there appear to be a number of striking parallels for the irregular alternation between **ž* and **š* found in the word for oats (Pronk/Pronk-Tiethoff 2018: 295). However, after examining each case on its individual merits, the picture is somewhat less optimistic. Although we do indeed find a similar distribution between Baltic **ž* and Slavic **s*, the words for 'heather' and 'fishing basket' are ambiguous, and it remains uncertain that the word for 'ploughshare' is directly comparable as we seem to be dealing with a *Wanderwort* showing a broad Central Asian distribution.

TABLE 12 Possible examples of an alternation *ž̯ ∞ *š̯

	Baltic	Slavic	Elsewhere
'oats'	*au̯iź-	*au̯iS$_1$-	Lat. *au̯e(T)s-n-
? 'fishing trap'	*u̯arź-	*u̯rS$_1$-i̯-	
? 'ploughshare'	*lemeź-	*(l)emeS$_2$-i̯-	Tur. *amač
? 'heather'	*u̯rź-	*u̯erS$_1$-	
	*u̯rś-		

6.3 Sibilant Clusters

6.3.1 *Cs ∞ *sC

Some studies into non-Indo-European loanwords have drawn attention to doublets showing the metathesis of *s*-clusters (Oštir 1930: 5–6; Furnée 1972: 392–393; Šorgo 2020: 459). Of course, irregular metatheses do occur, and one might ask exactly what feature of a suggested substrate language could lie behind such an alternation (cf. Beekes 2014: 18). Here, it is worth remembering that our non-Indo-European source language was probably not a monolith, and that *regular* metatheses do occur. For instance, compare the regular developments *#*ks*- > *#*sk*- in Baltic (Stang 1966: 95), *-*ps*- > -*sp*- in Latin (Leumann 1977: 202; cf. Hamp 2003), and *-*sk*- > *-*ks*- in Ob-Ugric (Aikio 2015b: 2) and (often but sporadically) in late West Saxon (Hogg 1992: 298). Thus, one way in which such an irregularity could be explained would be to assume that one of the donor languages underwent a (regular) metathesis. Collecting examples of metathesis is therefore not necessarily irrelevant to the question of language contact.

▸ '**wax**'. Lt. *vãškas*, Lv. *vasks*; OCS воскъ 'wax' ~ ON *vax*, OHG *wahs* 'wax' (Machek 1968: 697; Polomé 1986: 661) — The Lithuanian *-šk-* is in itself problematic, as outside of a RUKI environment, it is difficult to derive it from any Indo-European cluster (Villanueva Svensson 2009: 15–16). The most frequent solution is to suggest a proto-form **uoḱs-ko-* (Lidén 1897: 28; Kiparsky 1934: 96; Kortlandt 1979a: 59; Derksen 2008: 529), but what does not seem to have been noted is that *-*ḱsk*- would hardly have yielded Germanic *-*hs*- in the first

place; compare OHG *misken* 'to mix' (< **mík-ske-*; LIV 428–429). The alternative reconstruction **-ḱk-* (ALEW 1386) equally fails to explain the Germanic evidence (cf. Arumaa 1976: 98).[97] Unless we assume an irregular, and apparently unmotivated, metathesis (thus e.g. Endzelīns 1911: 57; Smoczyński 2018: 1617;[98] РЭС VIII: 286), the Balto-Slavic and Germanic forms cannot be regarded as regularly cognate, and the disagreement between the two words might best be accounted for by assuming parallel borrowings from a non-IE source.

A number of examples of an alternation **-ks-* ∞ **-sk-* have been identified elsewhere in Europe. First of all, we can mention the comparison of OHG *dahs* (< **þahsa-*) 'badger' with the name of the Middle Irish legendary figure *Tadhg* (< **tazgo-*) *mac Céin*, who was associated with a taboo against eating badger meat (on which see Mac an Bhaird 1980) (Kroonen 2013: 531; van Sluis et al. 2023: 212). More reliable examples can be found between Greek and Latin: Gr. ἰξός ~ Lat. *viscum* 'mistletoe; birdlime', Gr. ἀξίνη ~ Lat. *ascia* 'axe' (Furnée 1972: 393; de Vaan 2008: 57).

▸ '**sturgeon**'. Lt. *eršķẽtas*; Pr. E *esketres* 'sturgeon'; Lat. *excetra* 'sea serpent' ~ R *осётр*, Cz. *jeseter*, SCr. *jèsetra* 'sturgeon' (Pisani 1968: 20–21; for further refs. and discussion of the Germanic comparanda, see pp. 236–237) — The correspondence between Baltic and Slavic is quite irregular. Although Prussian *esketres* · stoer[99] and Slavic **esetra-* (trad. **(j)esetrъ*) are hardly to be separated, the *-k-* in Baltic remains a problem. A change **ešetras* » **ešketras* due to the influence of Lt. *eršķẽtis*, *eršketỹs* 'wild rose' (Būga 1922: 195; Endzelīns 1943: 171; Топоров ПЯ II: 88–91) is hardly plausible; an association with this word can only come into question to explain the later metathesis **ešketras* » *eršķẽtas* but not the 'intermediate' form attested in Prussian and as *efchketras* 'walfisch' in Bretke.[100] In principle, if the Slavic *-s-* goes back to **-ḱs-*, the relationship between the Baltic and Slavic words could be understood as metathetic.[101]

---

97 Stang (1972: 61) does not see any need to comment on this irregularity; likewise Vasmer (REW I: 231). Fraenkel (LEW 1207) refers to Endzelīns (1911: 57), who operates with an unexplained sporadic alternation already in Proto-Indo-European (cf. Būga 1922: 176; Otrębski 1939: 133).

98 Smoczyński assumes an *ad hoc* metathesis only for Slavic, but overcomplicates the Baltic evidence through the assumption of an unattested reflex **vašas* (for a suggestion on Finnish *vaha*, see Chapter 3, fn. 163).

99 To be read /esketrĭs/? Compare *erßketris* · Wallfisch in *Lexicon Lithuanicum* (ALEW 303).

100 Žulys (1966: 152–153) plausibly interprets this word in Bretke as a Prussianism. Kortlandt (2000: 125), on the other hand, who expects **e- > a-* in Prussian, takes the initial *e-* as evidence that the word was loaned from Lithuanian (also ALEW 303).

101 The etymological connection with Pl. obs. (Sł Warsz. II: 171) *jesiory* PL. 'fishbones' and Lt. *ešerỹs* 'perch' (Brückner 1927: 206; REW II: 281–282, Derksen 2008: 144) is morphologically problematic (**es-et-r-* beside **es-er-*?).

The comparison with Lat. *excetra* 'sea serpent', however (e.g. Trautmann 1910: 331), suggests an original cluster **ḱsk*, similar to the one traditionally reconstructed for 'wax' (see above). In fact, the Latin word could even be analysed as a regular cognate of Lithuanian *ešketras*, assuming a reconstruction **eḱsketr-*. Nevertheless, to connect the Slavic word, we need to assume an *ad hoc* simplification **-ḱsk-* > **-ḱs-*, which is without parallel.

The main issue with the comparison is semantic. The oldest attested meaning of the word in Latin is a kind of mythological sea serpent (Pisani 1968: 21; TLL VI: 2165). As sturgeons are a particularly large fish, such a semantic shift is quite imaginable. Compare, for instance, Bretke's use of the word *ešketras* to render the biblical *Walfisch* (Žulys 1966: 153), or Finnish *sampi* 'sturgeon', dial. 'fish god' (Liukkonen 1999: 124). A loanword from Greek ἔχιδνα 'viper' through Etruscan mediation (Walde/Hofmann I: 425–426), as noted by Pisani, is phonologically problematic. As the similarity between the Balto-Slavic and Latin forms is so striking, and the semantic difference is easily bridgeable, it seems plausible that these words belong together.

▸ **? 'aspen'**. Lv. *apse*; Pr. E *abse*; R *осúна*, LSrb. *wósa*, Sln. *jesíka* 'aspen' ~ ON poet. *ǫsp* (cf. Ic. *ösp* 'aspen, poplar'), OHG *aspa* 'aspen' (Meillet 1909: 70; Machek 1954: 132; Skok II [1972]: 759; Boutkan/Siebinga 2005: 94; Kroonen 2013: 39; Matasović forthc.) — Arm. dial. *opʿi* 'poplar' most probably belongs here, too. The Armenian word can reflect **Hops-* (Friedrich 1970: 49–50; Witczak 1991; on the phonology, see also Clackson 1994: 99–100; Kümmel 2017a: 442), although a reconstruction **Hosp-*, matching Germanic, cannot be ruled out, either (Normier 1981: 24, fn. 23). It is usually assumed, however, that the metathesis was a Germanic-internal phenomenon (cf. IEW 55).[102] This metathesis would be irregular, but it could quite reasonably have been motivated by an association with **aska-* 'ash' (see Normier 1981: 25–26 with lit.; note also the discussion in Chapter 7, fn. 83). This example of metathesis is therefore uncertain. For a detailed discussion of the Lithuanian forms and Turkic comparanda, see pp. 278–279.

6.3.2 *Baltic* *sT ∞ *Slavic/Germanic* *(T)s

In a footnote, Endzelīns (1911: 43–44) has enumerated some examples of apparent alternations between **st* and **ts* in the Indo-European material. Although he does not make any claim as to the regularity of such a metathesis, Kroonen/Lubotsky (2009) have proposed that the development **ts-* > **st-* was indeed

102 *Contra* Kluge/Seebold (p. 189) and Kroonen (2013: 39), a Proto-Germanic variant **apsō-* cannot be posited on the basis of the OE variant *æpse**, which is the result of an internal development (Campbell 1959: 185). Contrast OHG *aspa* with *wefsa* 'wasp' (< **waps-jō-*).

regular in Germanic on the basis of the equation of Skt. *tsárati* 'sneak' and Go. *stilan* 'steal'. To this we may add another compelling example adduced by Endzelīns:

– Skt. *tsárati* 'sneak', Go. *stilan* 'steal', Lt. *seléti* 'lurk, sneak', Arm. *sołim* 'crawl, creep'
– Skt. *tsáru-* 'handle, hilt',[103] Gr. στελεά 'handle', ON *stjǫlr* 'butt, rump', OE *stela* 'stalk, stem' (cf. ME *stele* 'the handle of a tool or utensil'), ?Arm. *steł* 'stalk, branch'

The amount of data is quite limited, and Armenian shows conflicting reflexes of the initial cluster. Nevertheless, by assuming that *seléti* shows the regular Baltic reflex of **ts-*, we can also account for a few other unexpected cases of *s-* in Baltic:

– Lt. *sárgas*, Lv. *sar̂gs* 'guard' ~ OCS стражь, R *стóрож* 'guard', cf. Gr. στέργω 'feel affection' (cf. REW II: 20; Derksen 2008: 467)[104]
– Lt. *síena*, Lv. *siêna* 'wall' ~ OCS стѣна '(defensive) wall, barrier', metaphorically 'rock face' (Brückner 1927: 529; Kalima 1934: 552, who reject — perhaps unnecessarily — the old comparison with Go. *stains* 'stone')
– ? Lt. *súolas*, Lv. *suôls* 'bench' ~ Go. *stols* 'seat, throne'. The Germanic word has alternatively been derived from **sd-ōl-* to the root **sed-* 'sit' (Kerkhof *apud* Kroonen 2013: 481; cf. Martirosyan 2008: 610–611), but this can be viewed as a serious alternative.[105]

On the other hand, there are a couple of examples which show the opposite correlation, and which therefore cannot be accounted for with any Indo-European reconstruction. It is possible that these represent parallel loanwords from non-IE sources:

▸ **'bison'**. Lt. *stuñbras* 'bison', Lv. *stuñbrs* 'aurochs' ~ Pr. E ⟨wissambs'⟩ · ewer; R *зубр*, Pl. obs. (cf. Sł. Warsz. VIII: 374) *ząbr* 'bison'; OE *wesend*, OHG *wisunt* 'bison' (Schrader/Nehring II: 261; Machek 1968: 719; Kroonen 2012: 253; Šorgo

103 This *tsáru-* is hardly the same word as *tsáru-* RV 'ein schleichendes Tier' as maintained by EWA I: 687.

104 A change **ts- > *st-* in Slavic and Greek is perhaps unexpected, typologically speaking, as **ps-* and **ks-* are both preserved word-initially in Greek, and we have **ks- > *kṣ- > *x-* in Slavic (cf. Pl. dial. *chybać* 'rush; sway' ~ Skt. *ví kṣobhate* 'stagger'). But we should not *a priori* assume that **ts-* (in which the two phonemes have the same place of articulation) should have behaved similarly to other **Cs-*type clusters. Petri Kallio (p.c. March 2023) points out, for instance, Western Finnish *-tt-* < **-ts-* (e.g. *mettä* < *metsä* 'forest') beside preserved *-ps-*, *-ks-*.

105 In any case, the Baltic word, already in view of its acute intonation, is not, with Būga (1922: 280), to be compared with OCS село 'field, estate, settlement' (which might be ?< **sedla-*, trad. **sedlo*; Brückner 1927: 491–492; Stang 1972: 47) or Lat. *solium* 'seat, throne' (probably with **d > l*, de Vaan 2008: 571).

2020: 455–456) — The Prussian attestation is abbreviated in the original, and is normally restored to *wissamb*[*ri*]*s* (Trautmann 1910: 464; Endzelīns 1943: 276). If the *-b-* in Balto-Slavic is epenthetic in a cluster **-mr-* (Būga 1912: 45),[106] then the correlation between Germanic **-und-* ~ Baltic **-umr-* could reflect the same **d* ~ **r* alternation as in Lt. *sidãbras* ~ OCS сребро 'silver' and Pr. E *wobsdus* ~ Lt. *opšrùs* 'badger' (see pp. 225–227). However, note that it is in principle not possible to rule out a reading *wissamb*[*i*]*s* for Prussian.

Problematic are the Latvian variants in *s-*: *sũbrs* (ME III: 1129; EH II: 606) and *sumbrs* (ME III: 1120; LVPPV). The preserved *-m-* and accentuation of the Latvian forms seem to point towards borrowing. It is tempting to interpret *sũbrs* as a loan from East Slavic (Petersson 1921: 39; with secondary *s-*?), in which case *sumbrs* might be a Polonism. In any case, it seems obvious that the cited words for 'bison' cannot be separated from one another (cf. REW II: 107; Būga 1912: 44–46). In view of the numerous problems with reconstructing a common proto-form, it seems most probable that we are dealing with a word of non-IE provenance. On the element **wi-* in Germanic and Prussian, see 7.1.3. Note also the mismatch in vocalism between East and West Baltic (see 7.3.1).[107]

▸ **'roe'**. Lt. *stìrna*, Lv. *stiřna* ~ OR сьрна́ (Зализняк 2019: 205), Sln. *śŕna* 'roe deer' — Endzelīns (1909: 378; cf. EH II: 489) has pointed to a form ⟨firnos⟩ ACC.PL. 'roe', attested in Rehehusen's 16th century *Manuductio ad linguam Lettonicam*. If this is not merely an error (cf. Fennell 1982: 339), then it perhaps results from a contamination with the Slavic word. Despite Endzelīns and many who have followed him, I doubt it should be considered a unique archaism (but compare 'bison', above).

Most agree that *stìrna* is of IE origin and related to Pr. E *sirwis* 'roe deer', Lat. *cervus* 'deer' and further the root for 'horn' (Trautmann 1923: 260; Nussbaum 1986: 8, fn. 16; Derksen 2015: 429). The initial *st-* has been subject to numerous explanations. Early scholars suggested a loan from Slavic (e.g. J. Schmidt 1895: 37; Mikkola 1908: 14; also Mayer 1990: 102), assuming a pre-Slavic **ć* was adopted as Baltic **st*. There is no other evidence from early Slavic loans, however, that would support an affricate pronunciation at such a recent date. Alternatively, Andersen (2003: 53–54) has suggested a loan from an unknown IE dialect.[108]

---

106 Compare Lt. *dumblas* 'mud, sludge' ~ Lv. *dumûksnis* 'marsh' (see p. 210).

107 The involvement of the pan-Caucasian term for 'bison' (Oss. ID *dombaj*, Karachay *dommaj*, Bzyp Abkhaz *a-domp'éj*, Georgian *domba*; Иванов 1975; Абаев 1996: 206; Kroonen 2012: 253) in this equation is less certain, as the initial *d-* and the suffix both need to be accounted for.

108 Another issue with the traditional etymology is the accentual difference between Baltic and Slavic (cf. Meillet 1905: 446). Assuming *vr̥ddhi* per Petit 2004: 184; Villanueva Svensson 2011: 31 seems like an *ad hoc* solution, see Pronk 2012: 11–13.

In my opinion, it is worth asking whether the IE etymology might be wrong; after all the roe, compared to the red deer and the elk, has far less prominent horns.[109]

▸ **'thousand'**. Lt. *tűkstantis*, Lv. *tũkstuôtis* 'thousand' ~ OCS тысѫщи, тысащи 'thousand'; Go. *þusundi*, ON *þúsund* 'thousand' (Stang 1966: 282; 1972: 49) — Note that the **-s-* in Slavic cannot reflect a simple **s* (which should have become **x* by the RUKI law), but would be quite elegantly be explained from **ts*, a reconstruction which could also work for Germanic. See 3.5.4 for a detailed discussion of this word and the Uralic comparanda.

▸ **? 'fast'**. Lt. dial. *bruzgùs*, Žem. *bruzdùs* 'quick, agile' ~ OCS (Supr.) брьзо ADV. 'quickly', MR *борзы́и* 'fast (of horses)', OCz. *brzý* (Gebauer I: 111), SCr. *brȋz* 'fast' — Much has been made of the variant *бóрздьı*, attested in Middle Belarusian since the 15th century (ГСБМ II: 148–151). As support for the latter's antiquity, Ильинский (1910: 324) has adduced the SCr. dial. (Montenegro) *brzdica* 'rapids' from Vuk (RJA I: 695; PCA II: 157; Skok I: 222) and modern Polish *barzdo* (replacing OPl. *barzo* in the 16–17th centuries). Despite a general consensus, I consider the doubts voiced already by Потебня (1881: 1) still valid. SCr. *brzdica* is curiously paralleled by dial. *brzdar* (PCA II: 156) for *brȁzār* 'a kind of leather bag', in which Skok (I: 222) would see a contamination.[110] Perhaps Derksen (2008: 70) is correct in blaming the Belarusian variant on Baltic influence (but see below). Although these variants present some problems, I doubt that the evidence is sufficient to support a Proto-Slavic variant **burzda-* (trad. **bъrzdъ*).

The comparison of the Baltic and Slavic data implies multiple irregularities. First, there is the irregular correlation between Slavic **-ur-* (trad. **-ъr-*) and Baltic *-ru-*.[111] Secondly, there is a disagreement between Baltic *-zd-* ~ *-zg-* and Slavic **-z-*. If *-zd-* can be set up as original in Baltic (which should not be taken for granted; the Aukštaitian *-g-* would in any case be left unexplained), we might

109 Incidentally, I would also keep apart the words for 'cow', Lt. *kárvė*, R *коро́ва*, as neither the acute nor the initial velar are well accounted for. Pr. E *kurwis* 'ox', for what it is worth, would in my opinion suggest a labiovelar.

110 For Pl. *barzdo*, see Łoś (1922: 148), who also adduces Pl. *smardz* 'morel' « OPl. *smarsz* (SSP VIII: 318). Sln. *brzdit* 'stolz (von Pferden)' (Murko *apud* Pleteršnik I: 68) which Bezlaj (I: 50) included here, is derived from *bŕzda* 'bridle' (Furlan 2013: 119).

111 I am hesitant to put any weight on the variant *burzdùs*, which seems only to have been recorded by Kurschat (1883: 65) who himself marks it as an unfamiliar word with the note "in Südlitt.". Even more doubtful is the variant *burzgùs*. In the LKŽ, it is equated with *bruzgùs*, with a single illustrative sentence: "Mūs mergaitės tokios buřzgios". Yet a very similar example found in the *Papildymų kartoteka*, "Kõ tà mergáitė tokià burzgì?", is glossed as "niurzgùs" = 'grumpy'!

TABLE 13 Possible examples of sibilant metathesis

	Baltic	Slavic	Germanic	Elsewhere
'wax'	*u̯o(ḱ)sk-	*u̯o(ḱ)sk-	*u̯oks-	
'sturgeon'	*e(ḱ)sket-r-	*e(ḱ)set-r-	?*(ḱ)str-	Lat. *eksket-r-
? 'aspen'	*op(u)s-	*ops-	*osp-	Arm. *ops- (or *osp-)
'bison'	*stum(bʰ)r-	*(d)zam(bʰ)r-	*u̯i(t)snT-	
'roe'	*st(i)r̂n-	*(t)s(i)rn-	–	
'thousand'	*tûstant-	*tûts(a)nt-	*tū(t)snT-	PF *tušant-
				Md./Ma. *tüšäm
? 'fast'	?*BruzD-	*Bur(dʰ)z-	–	

be able to set up an irregular correspondence between Baltic **zd* and Slavic **(d)z*, parallel to the examples of **st* ∞ **(t)s*, above.

This correspondence could potentially find a parallel in the word for 'heather'. As against the standard *vìržis*, Mielcke (11: 270) cites *birzdzei* 'heydekraut'. The reality of this form seems to be confirmed by the form *brizdei* 'Calluna', attributed by Pabrėža (1834: 60) to Prussian Lithuanian (admittedly, this is perhaps simply miscopied from Mielcke). This would also show **b* ∞ **v* (see 6.4.2), but in view of the large amount of variants shown by the word for 'heather' in Slavic (see pp. 215–216), it would seem hasty to draw any dramatic conclusions on the basis of such scanty data.

6.3.2.1 Conclusion

The certain and possible evidence of 'sibilant metathesis' is collected in Table 13, above (see p. 185 for help reading the table). Forms which do not provide relevant data are presented in light grey. Shaded cells indicate sibilant-initial clusters.

The clearest pattern concerns the alternation **sT* ∞ **(T)s*: here we consistently find a sequence **st* in Baltic. In Germanic and Slavic, the surface realization is just a sibilant; however, in the word for 'thousand', there is indirect

support for the reconstruction **ts*, as only this reconstruction can unite the Slavic and Germanic data and explain the absence of the RUKI law in Slavic. Since a **t* could have been present in the other examples, and they show a comparable pattern, it seems reasonable to assume they result from the same substratal phenomenon.

6.3.3 *Other Alternations Involving Sibilants*

▸ (a) ? **'furrow'**. Lt. *bir̃žė* 'row, furrow; timber tract; border mark', Lv. *bìrze* 'furrow, row' ~ R *борозда́*, Cz. *brázda*, SCr. *brázda* 'furrow' — The Slavic word is traditionally compared with Skt. *bhṛṣṭí-* 'point, peak' (< **bʰrḱ-ti-*, cf. EWA II: 273). The implied suffix **-d-* in Slavic is difficult to set up (see Vaillant 1974: 490 for some doubtful examples), and the semantics are hardly compelling, in any case. The main disadvantage of the etymology is that we would need to abandon any connection with the Baltic synonym (cf. Holzer 1989: 53).[112] The inclusion of the Slavic evidence implies an additional alternation between Baltic **ž* and Slavic **zd*. Perhaps this can be compared with Lt. *triùšis* ~ OCS тръсть 'reed' (see p. 201). On the comparison with Lat. *porca* 'furrow' etc., see p. 208.

▸ (b) **'beard'**. Lt. *barzdà*, Lv. dial. *bȁrzda*; OE *beard*, OHG *bart* ~ OCS брада, R *борода́* (ACC.SG. *бо́роду*); Lat. *barba* 'beard' (Schrijver 1991: 448; Kuiper 1995: 66; Derksen 2015: 82; Pronk 2019a: 147) — Kroonen (2011b: 150–151) has presented a native etymology for this word. He assumes that the Germanic word for 'beard' is connected to ON *broddr*, OE *brord* 'point, tip; shoot' (< **bruzda-*; thus already Pedersen 1895: 73) and ON *borð*, OE *bord* 'board, plank; side of a ship' (< **bur(z)da-*). He opts for the reconstruction **barzda-* (cf. Kroonen 2013: 54), which, being the result of a reshuffling of ablaut within Germanic, would imply that the Latin and Balto-Slavic words are Germanic loanwords.

The reconstruction **barzda-* for Germanic (likewise e.g. Kluge/Seebold 93) would provide a natural explanation for Lt. *barzdà*. We may interpret the Baltic and Germanic words as cognate or, following Kroonen's model, view the Baltic word as a loan from Gothic. The Germanic loan etymology might be supported by the absence of the RUKI law in Lithuanian. It would incidentally be attractive to see Crimean Gothic *bars*, which has previously been considered a transmission error or a unique retention of NOM.SG. *-s* (Lehmann 1986: 62–63), as a direct reflection of this preform.[113]

---

112 Note that the Baltic word is left unmentioned by e.g. Berneker (I: 75), Vasmer (REW I: 109) and ЭССЯ (II: 220).

113 This would require a return to the more traditional view that the words for 'board' (cf. Go. *fotu-baurd* 'footstool') are unrelated, for which something can indeed be said; the partial semantic convergence in Norse may be secondary. Latvian *bȁrda* is in any case due

The main problems arise when considering the Slavic and Latin evidence. Despite the claim to the contrary in ALEW (102–103), the loss of *-z-* in Slavic would be irregular (cf. Pedersen 1895: 72–73). This could be remedied by assuming a Slavic loanword from West Germanic; however, mobile accentuation is generally thought to be atypical of Germanic loanwords (Meillet 1909: 69; Pronk-Tiethoff 2012: 242–244 with lit.). A Latin loan from Germanic faces chronological issues as the change **-rð-* > *-rb-* belongs to the preliterary period (e.g. Weiss 2020: 208). Moreover, it is likely that a preceding sibilant would have blocked the frication of inherited **dʰ*, whether inherited **sdʰ* merged with **st* (Meiser 1998: 119; Weiss 2020: 161) or with **sd* (Lubotsky 2004). As a result, the Latin form is only consistent with a pre-form without **s*.

▸ **'in calf'**. Lt. *beřgždžias* 'barren, fruitless' ~ R dial. *берёжая* 'in foal', SCr. dial. *brȅđa* 'in calf; pregnant' (< **berdjā*; trad. **berdja*) — Lat. *forda* 'in calf' is ambiguous, and could reflect either **bʰrd-* or **bʰrsd-* (cf. Leumann 1977: 210–211). Despite ALEW (116–117), the Baltic and Slavic forms are not formally identical, not only because the loss of *-z-* in Slavic would be irregular (see above), but also because the Slavic form exhibits an acute. The difference in intonation could be accounted for by reconstructing **bʰerd-* for Slavic and **bʰersd-* for Baltic (with **-s-* blocking Winter's law). The morphological function of this **s* would be unclear, however, and the parallelism of this example with the word for 'beard' makes it rather tempting to view both in the context of parallel loanwords.

6.4 Other Irregularities

6.4.1 *Alternations Involving Dentals*

▸ (a) **'silver'**. Lt. *sidãbras*, Lv. *sudrabs*, dial. *sidrabs* ~ Pr. III *sirablan* ACC.SG.; OCS сребро, Cz. *stříbro*, Sln. *srebrộ* 'silver'; Go. *silubr*, ON *silfr*, OHG *silabar** 'silver' (Ipsen 1924: 229–230; Stang 1972: 47; Huld 1990: 409–410; Boutkan/Kossmann 2001; Mallory/Adams 2006: 242; Kroonen 2013: 436; Šorgo 2020: 448; Thorsø et al. 2023: 108; van Sluis et al. 2023: 221) — This word has widely been considered an ancient *Wanderwort*. The original form must probably be reconstructed with **r–r*, with different dissimilations in Germanic and Prussian. Nevertheless, the East Baltic *-d-* is difficult to write off as dissimilatory,

---

to an internal development, as implied not only by the Lithuanian equivalent, but also by the Latvian dialect data (Kregždys 2004: 20–21; ALEW loc. cit.). Perhaps it is German-influenced: cf. Pr. E *bordus* 'beard' = */bārdus/ which is probably from MLG *bart*, NOM.PL. *bārde* (Smoczyński 2000: 178).

since such a dissimilation would be entirely without parallel. Further comparanda are found in Celtiberian ⟨śilaɒuŕ⟩ 'money, ?silver' (K.H. Schmidt 1977: 55) and Basque *zilhar* 'silver' (Boutkan/Kossmann 2001; Thorsø et al. 2023: 108). An additional issue within Balto-Slavic is the fact that the medial *-a-* in Baltic does not match Slavic **-e-* (see 7.2.2).

▸ **? 'badger'**. Pr. E *wobsdus* 'badger' ~ Lt. *opšrùs* 'badger' (*Apfchro* GEN.SG. in Bretke implies an *a*-stem) (cf. Bellquist 1993: 344) — The reality of the Prussian form is confirmed by the gloss *wobsdis* 'quod dicitur eyn luchs' (probably to be corrected to **eyn dachs*, Töppen 1867: 155; Gerullis 1922: 205) as well as perhaps Kashubian *jôpsc* (< **āpsti-*; trad. **apstъ*)[114] 'badger' (Лаучюте 1982: 78). In view of this, Smoczyński's dismissal of the Prussian form as having "no explanatory value" (2018: 885) is too hasty.[115] A similar form also seems to occur in the *Lexicon Lithuanicum*: ⟨opßcʒus⟩ 'fisch otter' (ALEW 721; the same form is also given beside ⟨ûbßrus⟩ in ClG 663). This comes particularly close to the Kashubian form, and one might suspect that both have been borrowed from Prussian. However, the difference in voicing remains to be explained.

The alternation between Lt. *-r-* and Prussian *-d-* is unlikely to be due to different suffixation, as *-d-* is not a productive suffix (LEW 517–518; Smoczyński SEJL^2 s.v.). In theory, one may compare the similar alternation in the word for 'silver', with the caveat that the distribution does not match. However, we must note that Lv. *âpsis*, dial. (Vidzeme) *âpša* 'badger' does not appear to contain either "suffix". ALEW suggests that Lv. *âpša* may derive from an earlier **âpsćā-* (< **âpštjā-*), thus coming close to the marginally attested Lt. *opščius* (ALEW 721; see above). It does indeed seem likely that the development **stj* > **š* was regular in Latvian (Endzelīns 1923: 125–126), but as the reconstruction of **t* remains hypothetical, it is uncertain whether the irregularities in this word can be used to support foreign origin. However, the relationship between the forms is also

114 The derivation seems acceptable so long as the voiceless auslaut can be attributed to word-final devoicing. Alternatively, Boryś (SEK II: 341) suggests the Kashubian word is cognate with Polish *jaźwiec* 'badger' through a development **jazvc* > **jasfc* > **japsc*. Indeed, this might better explain the variant *jôlsc* (?< **javzc*). Since both etymologies require an irregular development, it is difficult to decide between them. Perhaps the two options could be combined if we assume that the inherited word for 'badger' was influenced by the Prussian word.

115 I fail to comprehend Smoczyński's problematization of the initial *w-* in Prussian, especially since no such issue is taken with deriving Pr. E *wosux* 'he-goat' from **āž-uk-* (Smoczyński 2018: 886). A prothetic *w-* is regular before *o-* in the dialect of the Elbing Vocabulary, as has long been recognized (Trautmann 1910: 158); this is confirmed by the complete absence of words starting with *o-*.

difficult to explain in an IE context. On the possibility that the word for 'bison' also shows a *$d \infty *r$ alternation, see p. 221.

▸ (b) ? **'bee'**. Lt. *bìtė*, dial. (Žem.) *bitìs*, Lv. *bite*; Pr. E *bitte*, TC *bete* 'bee' ~ OCS (Ps. Sin.) бъчела, Cz. *včela*, Sln. *čəbə̑la* 'bee'; OIr. *bech* 'bee', MW *begegyr* 'drone' (Hamp 1971: 187; Vennemann 1998; Philippa et al. I [2003]: 308; Matasović 2009: 65; van Sluis 2022: 6–10, forthc.; note also Machek 1968: 679) — The relationship between the Baltic and Slavic words is difficult to account for. A reconstruction *_bit-kelā-_ (Knut-Olof Falk *apud* REW II: 471) could theoretically work, but the analysis of the second element remains unclear. It is therefore usually assumed that we are dealing with different suffixes, *b^hi-_t_- beside *b^hi-_k_- (Specht 1947: 46; IEW 116), added to the zero-grade of a root *b^hei-, which is indeed attested with ablaut in Germanic: OHG *bīa* (MoHG dial. *Beie*), Du. *bij* against OHG *bini* (with short vowel confirmed by Notker), MoHG *Biene*, MLG *bene* 'bee' (Kroonen 2011b: 228–231).

Vennemann (1998: 478–479; cf. Takács 2001: 109–110 with some older macrocomparatavist refs.) has drawn a further comparison with Egyptian  *bjt* 'bee' (Erman/Grapow I: 434; cf. the derivative in Coptic ⲉⲃⲓⲱ 'honey', Vycichl 1983: 38). The similarity is indeed striking, especially if the *-t* can be considered a feminine suffix (which is not certain; Takács 2001: 109). Such a suggestion is also historically plausible, since the first depictions of hive beekeeping derive from Egypt (Crane 1999: 162), although there is admittedly a great geographical distance between Egypt and the Northern Europe, to which our word is restricted (see van Sluis 2022: 7).

The main obstacle to uniting the European forms is the Celtic vocalism. Although *_biko-_ has sometimes been reconstructed (e.g. Berneker I: 116; Heiermeier *apud* LEW 1329; Matasović 2009: 65), the broader consensus among Celticists favours *_beko-_ (Stokes 1894: 166; Pedersen 1909: 88; LEIA B-24–25; van Sluis 2022: 8). This is supported by North Occitan *bèca* 'wasp', which is most probably a Gaulish loanword (Delamarre 2003: 70).[116] As noted by Pedersen (loc. cit.), the Slavic form could potentially reflect an earlier *_bečelā-_ (trad. *_bečela_) with the raising of unstressed *_e_ before a palatal; compare,

116 The Celtic etymology has been rejected by A. Thomas (editor's note, *Romania* 35, 139) and FEW (XIV: 344). They note that Creuse *bièco* would imply an earlier *_bęsca_, which they assume was metathesized from *_guespa_, deriving ultimately from Lat. *vespa* 'wasp'. However, the /s/ ought to have been preserved in Limousin, cf. *crespa* 'kind of pancake' (= French *crêpe*). In addition, the word was probably originally masculine (note the Limousin variant *bèc*), which would explain the preserved *-c* (Occitan *lac* 'lake' < *lacus*; cf. FEW V: 126; old *-ca* should have yielded *-cha* in Limousin, as in *pescha* 'fish' < *_pisca_; cf. Thomas loc. cit.). The diphthong in Creuse may be explained as due to contamination with *gyepo* (cited *apud* FEW XIV: 344) 'wasp', with which it is in competition in this area.

in a very similar environment, OCS вьчера 'yesterday' beside вечеръ 'evening' (cf. Kortlandt 1984–1985). As the comparison between Baltic **bʰit-* and Slavic/Celtic **bʰek-* would rest on the first phoneme alone, it is not entirely clear that there is enough material to draw a reliable comparison.[117]

6.4.2 *Alternations Involving Labials*

▸ (a) **'bean'**. Pr. E *babo*; R *боб*, Pl. *bób*, Sln. *bòb*; Lat. *faba* 'bean' ~ ON *baun*, OHG *bōna* 'bean' (Machek 1950b: 158; Kuryłowicz 1956a: 194; Schrijver 1991: 488; РЭС III [2009]: 283; Kroonen 2013: 55; Matasović 2013: 83; Pronk/Pronk-Tiethoff 2018: 282; Šorgo 2020: 435) — According to Walde/Pokorny (II: 131), the Germanic forms developed from **babnō-* by dissimilation, an *ad hoc* suggestion that has gained few serious proponents (ЭССЯ II: 149; Bańkowski 2000 I: 69 and with hesitation Kluge/Seebold 96). Instead, one has tended to keep the Germanic words apart (Kretschmer 1896: 146; Petersson 1909: 390; de Vries 1962: 29; implicitly Trautmann 1923: 23; REW I: 180). If the words are indeed related, the disagreement between Slavic and Latin **bʰ* and Germanic **w* would favour independent borrowings from a non-IE source.[118] Note also in this context the Latin *a*-vocalism (see 7.6). However, one must remain cautious due to the small amount of phonetic material compared.[119] Kretschmer (1896: 146) has assumed a connection with Lt. *pupà*, Lv. *pupa* 'bean', positing a loan from Slavic through a Finnic intermediary (Berneker I: 65; Walde/Pokorny loc. cit.). This can hardly come into question: Livonian *pubā* 'bean' is a Baltic loan (Thomsen 1890: 100; Petersson 1909: 390; Sabaliauskas 1959: 235), while Finnish *papu* is a loan from Slavic (Kalima 1956: 102). Pronk/Pronk-Tiethoff (2018: 282) have instead suggested that *pupà* is a loan from the same foreign source as the above forms. While the resemblance (a stem consisting of two labial stops) is indeed striking, neither of the implied alternations are precisely paralleled in my material

117 But note further van Sluis (forthc.), who presents some potential parallels for an alternation between **k*, **t* and **∅* in possible substrate words.

118 In this connection, note Berber **a-baw* 'faba bean', which Kossmann (1999: 113–114) states cannot be a direct loan from Latin. His current opinion (cf. Kossmann 2021: 16) is that we are dealing with a *Wanderwort* which has "spread over the Berber territory in post-proto-Berber times". If a Latin origin is ruled out, this begs the question as to whether it has been adopted from a related non-IE source.

119 An additional argument for foreign origin could be provided by the gloss *haba* 'faba', attributed to the "Falisci" by Terentius Scaurus. If this word really did belong to Faliscan proper, the absence of the change **-bʰ-* > **-f-* would imply a Proto-Italic **-b-*. This would not match the **-bʰ-* required by Balto-Slavic, and rule out the reconstruction of a common proto-form. However, since the development of initial *f-* > *h-* was probably not limited to Faliscan, and the reliability of glossators' attributions is often questionable, it is difficult to base much on this form (cf. Bakkum 2009: 83, 209).

(for *$p \infty *b^{(h)}$ see 6.2.2 and for *$u \infty *a$ see 7.3.1), which makes the inclusion of this form somewhat precarious.

It is also difficult to find reliable parallels for the alternation *$b^{(h)} \infty *w$. One relatively clear case without Balto-Slavic comparanda is the word for 'pea' (Gr. ἐρέβινθος 'chickpea' ~ OHG *arawīz* 'pea'; see e.g. Kroonen 2012: 242–244; Thorsø forthc. with lit.). In addition, Machek (1950b: 152–153; 1968: 132) has suggested we compare the Slavic word for 'oak' (OCS дѫбъ, etc.) with OHG *tanna* 'fir' (thus a virtual *d^honb^h- ~ *$d^honu̯$-). However, the semantic distance makes this comparison very uncertain.[120]

▸ (b) '**carrot**'. R *морко́вь*, SCr. *mr̂kva* 'carrot'; OE *moru* 'edible root', OHG *moraha* 'carrot' ~ Lv. *buřkãns* 'carrot'[121] (Machek 1950b: 158, 1954: 167; Kroonen 2013: 378; Matasović 2013: 88; ERHJ I [2016]: 639; Pronk/Pronk-Tiethoff 2018: 282; Šorgo 2020: 446) — The status of the Latvian word is somewhat problematic, as it may also have been loaned from Baltic German *Burkane* (see the discussion on pp. 31–32). However, it can be noted that while the word in Russian and German is clearly a late replacement for an older term for 'carrot', there is no other candidate for an old term within Baltic (Bentlin 2008: 247). A possible trace of this word in Lithuanian can be found in Szyrwid's *burkuntay* · pasternak, with an unclear *-t-*,[122] while the usual form is the very recent loanword *morkvà*, *morkà* 'carrot' (← Bel. *мо́рква*).

Moksha *puŕ'kä* 'carrot' has been derived from Russian (Mikkola 1894: 91; Helimski *apud* РЭС II: 223). However, the Russian form is not attested anywhere in the vicinity of Mordovia, being limited to the area adjacent to the Baltic-speaking territory (see pp. 31–32). There are also phonological obstacles

120 Within Baltic, one might also cite Lt. *kalavìjas* 'sword' as against Pr. E *kalabian* (= III *kalbīan* ACC.SG.). Yet it seems more attractive to explain this disagreement by assuming a Lithuanian loanword in Prussian. There are several cases of German /v/ being substituted as Prussian /b/, such as Pr. III *ebangelion* ACC.SG. 'gospel', *burwalkan* ACC.SG. 'yard' (← MHG *vorwërc* 'estate'), which implies that the Prussian still had a bilabial /w/ until recently, and therefore substituted a foreign /v/ with a labial stop. One wonders if a similar solution might be on the cards for the Lithuanian form *birzdžiai* 'heather', attested in Prussian Lithuania (see p. 223). Although the attested word for 'heather' in Prussian is E *sylo* (~ Žem. dial. *šilas*), this does not rule out the existence of dialectal synonyms.

121 North Žemaitian *buřkonas* is a loan from Latvian (cf. Sabaliauskas 1960b: 261). Perhaps the same can apply to the rather aberrant *burkúnas* given by Juška (I: 254), the geographical origin of which cannot be ascertained.

122 Another trace of this word could be found in Lv. *burkaņ̂ts*[2] (attested in Snēpele, EH I: 254) if this was borrowed from Lithuanian (differently Sabaliauskas 1960b: 261). The stem-final *-t-* has a curious parallel in Estonian *porgand -i* 'carrot' (cf. Būga 1925: 771), but this has been analysed as an excrescence within Estonian (Blokland 2005: 298–299). I will leave these forms out of consideration.

to a Russian source. Firstly, the loss of the final *-n* is unmotivated, the final sequence *-an* being known even in inherited words (cf. Md. M *kućkan* 'greater spotted eagle' < PU **kočka*), and *-n* having been preserved in other Russian loanwords (cf. Moksha dial. *karman* 'pouch, handbag' ← R *карма́н*; *praban* 'drum' ← R *бараба́н*, etc.). Second, the palatal *ŕ* (with subsequent fronting **-a* > **-ä*; Bartens 1999: 63) is not easily explained on the basis of the Russian data.[123] Both of these issues are equally prohibitive to the derivation of the Mordvin word directly from Baltic (thus Donner 1884: 266–267; Иллич-Свитыч 1960: 18).

According to Junttila (2019: 51), any word for 'carrot' must be recent, as carrot cultivation only became widespread in northern Europe in the Middle Ages.[124] Archaeologically, the evidence is "deplorably fragmentary" (Zohary/Hopf 2012: 160), so it seems difficult to draw firm conclusions. Part of the reason for this is the difficulty or impossibility of distinguishing between wild and cultivated specimens, and because cultivated varieties were usually harvested before going to seed (Karg/Robinson 2000: 137; Mueller-Bieniek 2010: 1725). The German word *Möhre* 'carrot', having regular cognates in Old English, must date at least to Proto-Germanic. However, it did not necessarily originally refer to the domesticated carrot. In several glosses, OHG *moraha* is given specifically as *pastinaca silvatica* 'wild carrot' (see AWb s.v. *mor(a)ha*); the exact referent of OE *moru* is not known except for the fact that it was distinct from the foreign *wealh-more* (glossed *pastinaca*, *daucus*; cf. *Dictionary of Old English Plant Names*,[125] s.v. *more* (1) with lit.); in Middle English, *more* referred to both inedible and edible roots.

In conclusion, there is a close resemblance between Slavic/Germanic **murk-* and Baltic **burk-*. If they go back to parallel borrowings from another source, then we might be dealing with an original term for 'edible root' which has become specialized in the sense 'carrot' in the individual languages. The word has spread into the Finnic languages (see p. 32) and Mordvin, but the route or

123 This argument is perhaps not as convincing, as Paasonen (MdWb) records dialectal variants of Erzya *morkov* 'carrot' (which is borrowed from Russian *морко́вь*) with a similar palatal — *moŕ'ko·v*, *ṁiŕkoṃ*. This does not appear to be a general phenomenon, however, so we might assume an exceptional solution, such as transfer of the palatal feature from the Russian final /v'/ to the previous syllable.

124 Junttila proposes a novel etymology (2014: 131; 2019: 51–52), deriving the Baltic German word for carrot from the place name *Burgundy* in the context of Hanseatic trade. As a parallel, he offers Hungarian *burgonya* 'potato', which is of the same origin. The obvious problem with this etymology is that Baltic German *Burkane* differs in consonantism, vocalism and place of stress from MoHG *Burgund*; there seem to be too many missing links in this etymology for it to be accepted.

125 Online database, accessed at http://oldenglish-plantnames.org/.

even trajectory of its spread is difficult to reconstruct. Since the word seems to be old in Germanic, having predated Grimm's law, we are probably dealing with an originally European word, which may have entered Germanic, Slavic, and Baltic, and ultimately Mordvin, independently from related non-IE sources.[126]

▸ [? **'goosefoot'**. Lt. *balánda*, Lv. dial. *baluôda* 'goosefoot, *Chenopodium*' ~ OS *maldia*, OHG *melta* 'orache, *Atriplex*' — See the discussion on pp. 177–178.]

▸ (c) **'oven'**. Pr. E *Vumpîs* 'oven' ~ Go. *auhns**, OHG *ovan* 'oven' (for refs. and more discussion, see p. 254) — In view of Pr. E *umnode* 'bakehouse', the word *Vumpîs* (for ?**umpins* < **umpns*) probably stands for an underlying */umnəs/, see also "[monticulus], qui dicitur Vmne prutenice, id est clibanus" 1331 (Gerullis 1922: 33; PKEŽ IV: 267; for the *-p-*, compare Pr. G *kampnit* ~ *kamnet* 'horse'). It is tempting to compare Prussian **umna-* with Germanic **ufna-* directly, which would imply an irregular correspondence **p* ∞ **m*. For further discussion, and on the possible connection with Gr. ἰπνός 'oven, furnace', see p. 254.

Fraenkel (1936c; see also LEW 1156–1157) attempts to derive both the Prussian word and Lt. dial. *ùblas* 'indoor oven for producing tar' from Germanic. He assumes the Lithuanian word was adopted "von der Weichselgermanen"; however, the attested Gothic *auhn* ACC.SG. 'oven', which shows a dissimilation **f* > **h* (see Kroonen 2013: 557), is hardly a suitable source, and Fraenkel's **ubnas* does not appear to be continued by any Germanic language.[127] For Prussian, he points to Sw. dial. (Rietz 486) *omn* and suggests a possible Scandinavian origin. However, there is no certain evidence of Scandinavian loans in Baltic (see Chapter 2). The Lithuanian word is phonologically rather difficult to compare with the other forms due to the need to assume a "suffix replacement", and its appurtenance remains uncertain.

A possible parallel for the alternation **P* ∞ **m* is found in the comparison of the Slavic word for 'oak' (OCS дѫбъ, etc.) with Finnic **tammi* (> F *tammi*, E *tamm*, Li. *täm*) 'oak', suggested by Machek (1968: 132). The Finnic word has regular cognates at least in Mordvin (E *tumo*, M *tuma* 'oak') and probably also in Mari (E *tumo*, W *tum* 'oak'; on the vocalism cf. Живлов 2014: 125; Metsäranta

126 Guus Kroonen (p.c. September 2021) points me towards some similar North-East Caucasian forms: Lak *marχ*, Dargwa *marqʷa* 'root'. I remain agnostic as to whether these could be somehow related.

127 In view of the substitution of Gothic lowered /ę̆/ in Lt. *pẽkus* ← Gothic *faihu* (see p. 41), we might expect Gothic */ǫ̆/ to turn up as Baltic */a/, although the existence of an East Germanic dialect which did not undergo **u* > *au* is conceivable. In any case, a later West Germanic origin is out of the question, as German **o* with open syllable lengthening is never adopted as Lt. *ŭ* (see Alminauskis 1934, *passim* and e.g. 144–145).

2020: 81).[128] The possibility that the Slavic and West Uralic words could go back to a shared substrate has been suggested again recently by Живлов (2015) and Aikio (*apud* Matasović forthc.). While the similarity between the words is curious, it is difficult to imagine a plausible way to bridge the geographical distance between Proto-Slavic and Volga Uralic.

▸ (d) ? **'aftermath'.** Lt. *atólas*, Lv. *atâls*; Pr. E *attolis* 'aftermath' ~ R dial. *омáва*, Sln. *otâva* 'aftermath' — Vasmer (REW II: 289) compares the Slavic words with Sln. *otáviti*, Cz. *otaviti* 'revive, strengthen', and considers them a derivative of the verbal root seen in OR тыти 'grow fat' (СРЯ 11–17 XXX: 257–258). It is equally possible that the verb in question is denominal, however: cf. Lv. dial. (ME I: 149, EH I: 133) *atãlêtiês* 'recover, get one's breath back' < *atãls* (cf. Thomsen 1890: 159; Gāters 1953: 113). Various root etymologies for Baltic are summarized in LEW (p. 22), but the similarity of the Baltic and Slavic words encourages a direct comparison (cf. Miklosich 1886: 228; Trautmann 1923: 16; Witczak 2001: 44–45). A segmentation of the Baltic word as **atâ-la-* and reference to the nominal prefix Lt. *ató-* is unlikely, as this prefix is unknown elsewhere in Baltic and is probably a Lithuanian innovation on the model of nominal *pó-*, *pró-* (etc.). As the Baltic and Slavic words are so similar, and the suffix **-âla-* would be unusual, one may consider an alternation **l* ∞ **w*, which is phonetically plausible, although not paralleled. On the question of the Finnic comparanda, see 3.5.3.

---

128 J. Häkkinen (2009: 37–38) considers the West Uralic term a probable substrate word, but without mention of the Slavic comparandum. The comparison between Slavic and West Uralic was already made by Топоров/Трубачев (1962: 246; see also Tomaschek 1883: 704), who saw the Slavic word as a loan from a dialect of Proto-Finnic. This can hardly be seriously considered due to the probable geographical distance between Proto-Slavic and Proto-Finnic and in the absence of convincing parallels. Напольских (2002: 143–145) rather sees the Uralic word as a borrowing from a lost Baltic dialect. Finally, Witczak (2020: 75–76) has interpreted the Slavic word as a loan from a West Uralic compound of **tammə* 'oak' + **puwə* 'tree'.

CHAPTER 7

Vocalism

7.1 Initial Vowels

In an important article, Schrijver (1997: 307–310) suggested the existence of a morpheme **a-*, which he supposed appeared in a number of non-Indo-European lexemes with a European distribution. He observed that the presence of the 'morpheme' in a couple of cases correlated with a 'reduced' stem. The most convincing example of this phenomenon is found in the word for 'blackbird':

**mesal-*	Lat. *merula*; MW *mwyalch*, Bret. *moualc'h* (< PCelt. **mesal-*(*s*)*kā-*)
**a-msl-*	OE *ōsle*, OHG *amsla* (< PGm. **amslōn-*) 'blackbird'

An excellent parallel is found in the comparison of Lat. *raudus* 'piece of copper or brass' and ODu. *arut* (attested in Latin context), OHG *aruz* 'ore' (Schrijver 1997: 308; Kroonen 2013: 37). As well as showing a similar correlation between the presence of **a-* and a 'reduced' stem (**raud-* ∞ **a-rud-*), the two variants also show an identical geographical distribution. This distribution is, however, disturbed by the addition with Sumerian *uruda* (< *aruda*; Jagersma 2010: 61) 'copper' (see Thorsø et al. 2023: 109). Although Schrijver's (2018: 363) suggestion that the language of Europe's first farmers could have been related to Hattic would somewhat resolve the geographical issue, there is also a huge time difference involved. It would be quite a stretch to assume that such morphological alternations as found in Hattic[1] would have been preserved in Europe intact for millennia after its colonization by farming populations.

Several more suggested examples of the morpheme **a-* have been collected by Iversen/Kroonen (2017: 518) and Schrijver (2018: 361–363; cf. also Matasović 2020: 338–342), although not all of them show the expected pattern of stem reduction. I have divided my evidence into those which do and those which do not follow this pattern.

1 In fact, it seems hardly possible to rule out that the relevant vowel reductions (associated with the definite article) constitute a young development unique to Hattic.

7.1.1 *a- ∞ *∅- *with 'Reduced' Stem*

▸ **'swan** (1)'. R *лéбедь*, Bg. *лѐбед* ~ Pl. *łabędź*, Sln. *labǫ́d*; ON *ǫlpt*, OHG *elbiz* (for refs. and further discussion, see pp. 176–177) — In East Slavic and Bulgarian, one finds the form **lebedi-* (trad. **lebedь*). This is the reconstruction given by early authors (e.g. Miklosich 1886: 162; Osthoff 1898: 65; cf. Andersen 1996a: 124),[2] although **elbedi-* (trad. **elbedь*) is nowadays more popular (Булаховский 1948: 118; ЭССЯ VI: 19; Sławski SP VI: 40). The older reconstruction is preferable, as the idea that **eRC-* should develop to *RěC-* throughout Slavic is doubtful: corroborating examples are lacking, and one would expect a parallel treatment to **aRC-* (trad. **oRC-*; cf. Vaillant 1950: 160–161; Derksen 2008: 143; Jakob forthc. a.).

Derksen (2000: 84) has suggested to account for the different Slavic reflexes by assuming a 'prefix' **a-* (thus **a-lb-* ∞ **leb-*). Although this idea was not taken over in his dictionary (2008: 143), it does seem a plausible way to account for the two forms. The irregular alternation between **-bǭd-* and **-bed-* in the second syllable already strongly suggests a non-IE origin. Furthermore, the geography, with the *a*-forms restricted to Germanic and the western part of the Slavic area, seems quite consistent with the examples adduced by Schrijver.

▸ **? 'elm** (2)'. OR *ильмъ*, OPl. *ilem** (hapax, attested ⟨Ylem⟩ 1472; SSP III: 15), dial. *ilmak* (Sł. Warsz. II: 78), Sln. dial. (Carinthia) *lìm* < **ilm* (Erjavec 1883: 293; Karničar 1990: 51); OE *elm*, OHG *elm*; Lat. *ulmus* 'elm' ~ MIr. *lem*; MW *llwyf* 'elm' (Machek 1954: 90; Polomé 1990: 334; Schrijver 1997: 311; van Sluis forthc.; Matasović forthc.) — Latin *ulmus* can probably reflect **elmo-* with **e-* > **o-* before velarized /l/, as in *olor* 'swan' ~ MW *alarch* (< **elar-* with Joseph's law; Schrijver 1995a: 76), followed by regular **olC* > *ulC* (cf. Weiss 2020: 150–151). Quite alone stands ON *almr* 'elm': perhaps ths initial *a-* has been carried over from other tree names (cf. ON *askr* 'ash', OSw. *asp* 'aspen', *al* 'alder'). Matasović (2009: 237), like Pedersen (1905: 313–314), has made an attempt to explain the words in terms of IE ablaut, but has later favoured a non-IE origin (Matasović forthc.).

The Slavic words have often been derived from Germanic (Miklosich 1886: 95; Berneker I: 424; Kiparsky 1934: 148), or more specifically, MHG *ilme* (attested since the 13th c.). In view of (1) the early attestation in Russian (already the Novgorod First Chronicle)[3] and its widespread appearance in Russian toponymy (Vasmer 1938: 452; В. Васильев 2012: 427–429), and (2) the non-trivial

2 Osthoff reconstructed Pl. *łabędź* etc. as **lōbʰ-*, comparing the Hesychian gloss ἀλωφούς · λευκούς, which occurs beside a parallel gloss ἀλφούς · λευκούς. The former is most probably a mere transmission error (Beekes 2010: 77; Gippert 2017: 184–185), meaning that Osthoff's reconstruction has no real basis.

3 See Folio 113b (under the year 6738) of the Synodal Codex.

development of Bel. dial. *лём*, LSrb. *lom*, showing the regular outcomes of a vocalized *yer* (cf. Schaarschmidt 1997: 62) and loss of **i-* (trad. **jь-*), a late Germanic loan is out of the question (thus also Friedrich 1970: 81–82; ЭССЯ VIII 222–223; Derksen 2008: 211).[4] Despite Machek's (1954: 90) claim to the contrary, all the forms can be united under a single proto-form, **ilima-*.[5]

The irregularity essentially depends on the Celtic evidence, namely MIr. *lem* (< **lemo-* / **limo-*) and MW *llwyf* (< **leimo-*) 'elm'. These words fail to correspond with each other even within Celtic. Schrijver (1997: 311) characterizes the relationship between the Celtic and other European words as **V-lm- ~ *lVm-*. Nevertheless, no precise parallels can be identified, and since the comparison only consists of two consonants, there is always a possibility that the similarity is due to chance.

Curiously, a very similar word is also found in several Turkic languages, cf. Chuv. *jĕlme*, Tatar *elmä*; Kumyk *elme* 'elm', Noghai *elmen* 'aspen'. On the basis of these forms, СИГТЯ (I: 126) offers the Proto-Turkic reconstruction **elmen*. If the final *-n* in Noghai is secondary after *emen* 'oak', the Caucasian Turkic forms could be combined under **elme*; however, the Volga Turkic forms imply a reconstruction **ilmä* (cf. А. Дыбо 2007: 129–130). The initial *j-* in Chuvash is irregular and would suggest a reconstruction **jilmä*, but it is perhaps secondary; Савельев (p.c. August 2021) has informed me of a form ⟨ѣльмя̀⟩ in an 18th century source, which would imply */ĕlmä/.

As the Volga Turkic vowel shift can be dated to the 15th or 16th centuries (Doerfer 1971: 329), even a Middle Russian origin could be considered, although in view of the lack of parallels, we may be inclined to date the loanword earlier. The reality of linguistic contact between Turks and early Slavs is

4 It is interesting to consider the possibility of a loan from *early* West Germanic, however. West Germanic **ŏ* may well in some cases have been adopted as **u* in Slavic, cf. OCS хлъмъ ~ OS (Heliand) *holm** 'hill', where attested Gothic uses a different word for 'hill', *hlain(s)**. There are no certain examples of the development WG **e* → Slavic **i* (trad. **ь*; the word for 'radish'? cf. Sabaliauskas 1960b: 258), but such a substitution might be anticipated. In this case, one could envisage an earlier Slavic loan from West Germanic **elma-* (cf. OHG *elm*). For the insertion of **i* (trad. **ь*) after **l* as a reflection of the Germanic 'clear l', compare OR Ольга < ON *Helga*.

5 *Pace* ESJS 448, there is no reason to reconstruct a Proto-Slavic variant **lima-* (trad. **lьmъ*). The loss of **i-* (trad. **jь-*) is semi-regular in West and South-East Slavic (Derksen 2003; the 'Russian' form *лём* cited in ESJS is in fact Belarusian, cf. СРНГ XVI: 346). The other forms, Pl. dial. *lim* (Sł. Warsz. II: 743), R dial. (Siberia) *лим* (СРНГ XVII: 47; ?cf. *и́лим*, Даль² II: 39), Sln. *lom* (Cigale 1860: 1306), do not show regular reflexes of **lima-* (trad. **lьmъ*) and must be explained otherwise. The CS form льмъ (Bes.), found twice on a single page (cf. SJS IV: 636), is evidently a scribal error for the Latin loan оульмъ 'elm', which is attested only in this text, and was apparently unfamiliar to the copyist.

proven by the existence of common Slavic borrowings from Turkic, cf. R arch. *толмáчь*, Cz. obs. *tlumač* 'interpreter'[6] ← Turkic **tïlmač* (> Old Turkic *tilmač*, Tatar *tïlmač* 'interpreter', Yakut *tïlbās* 'translation'; cf. REW III: 115–116; Doerfer 1965: 662–665; ЭСТЯ III: 233–235). Moreover, a couple more early Slavic loanwords have apparently found their way into Volga Turkic, most strikingly Tat. dial. *könǯälä*, Chuv. *kĕnčele* ~ dial. *kănčala*[7] 'flax prepared for spinning' ← Sl. **kǭželi-* (trad. **kǫželь*; > R dial. *кужель*, Bg. dial. *кѫ̀жел* 'flax prepared for spinning', Cz. *kužel* 'distaff'). From a phonological and geographical point of view, however, it seems difficult to derive the Caucasian Turkic forms directly from Slavic. Although the exact source of the Turkic words remains elusive, it is more probable that these are ultimately of Indo-European origin rather than representing independent witnesses of a non-IE *Wanderwort*.

▸ ? '**sturgeon**'. R *осётр*, Cz. *jeseter*, SCr. *jèsetra*; Lt. *eršķētas*; Pr. E *esketres* 'sturgeon' ~ Ic. *styrja*, OHG *sturio* 'sturgeon' (Oštir 1930: 6; Machek 1950b: 150; Bezlaj I: 228; Kroonen 2012: 240, 2013: 488; Iversen/Kroonen 2017: 513) — Both Baltic and Slavic point unambiguously to an initial **e-* (with regular development to *o-* in East Slavic, *pace* Andersen 1996a: 147). Despite the difficulties with interpreting initial vowels in Balto-Slavic (see 7.2), it might still be wise to keep this example apart from other examples of the **a*-prefix. Kroonen (2013: 488), who reconstructs **asetr-* for Balto-Slavic against **str-* in Germanic, would interpret the Germanic *u*-vocalism as resulting from a 'reduced stem' with a vocalic **r̥*.

As Kroonen notes, based on the other examples of *a*-prefixation, we should expect **a-str-* beside **setr-*. He argues that the original "ablaut" may have been "reshuffled", although since we do not have a clear understanding of the mechanisms underlying the alternations in the attested material, such an interpretation is risky; Šorgo (2020: 449–450) rejects the example altogether. An alternative non-IE analysis would be possible in the context of the **e* ∞ **u* alternation seen in **klen-* ~ **klun-* 'maple' (see 7.3.2), although an additional metathesis of **r* would have to be assumed.

Above (see pp. 218–219), I have argued in favour of a comparison of the Balto-Slavic words with Lat. *excetra* 'sea serpent'. If this is accepted, the comparison

6 Czech-Slovak shows a strange 'ablaut' between *tlumač* 'interpreter' : *tlumočiti* 'interpret'. This opens up the possibility that Slavic originally had a short vowel in the second syllable (in agreement with Turkic), preserved here in the derived verb, while in the noun it was secondarily assimilated to the agent noun suffix **-āče-* (trad. **-ačь*; on which see Vaillant 1974: 321–323).

7 This dialect form shows the expected Chuvash reflex with the development **küN-* > **kuN-* (as in Chuv. *kăn* = Turk. *gün* 'day'; *kămpa* = Tat. *gömbä* 'mushroom'). In this light, the more common front-vocalic form is perhaps loaned from Tatar.

with Germanic **stur-* becomes rather more impressionistic, even though the semantic correspondence with Germanic is perfect. If we follow Kroonen and reconstruct a 'reduced stem' without the initial vowel, we could set up a preform **ks(k)tr-*. While this could well develop into **stur-*, the fact that so much material has to be lost to achieve the Germanic forms makes the suggestion rather dubious.

▸ ? **'turnip'**. Lt. *rópė*, OHG *ruoba*, MDu. *rove*, Lat. *rāpum* 'turnip'; Gr. ῥάφανος 'cabbage, radish'; R *pŕna*, SCr. *rȅpa* (dial. *rȉpa*) 'turnip' ~ MW *erfin*, Bret. *irvin* 'turnip' (Oštir 1930: 64; Machek 1954: 57; Walde/Hofmann II [1954]: 418; Furnée 1972: 163; Čop 1973: 29; Chantraine DELG IV [1977]: 968; Bezlaj III [1995]: 171; Kroonen 2013: 415; Pronk/Pronk-Tiethoff 2018: 301) — Lt. *rópė* along with the Germanic and Latin forms support a common reconstruction **reh$_2$p-*. Nevertheless, the irregularities in the other cognates clearly point to a foreign origin, which has long been recognized. These irregularities include (1) the vocalic alternation between **-ă-* in Greek, **-ē-* or **-ai-* in Slavic, and **-ā-* elsewhere; (2) the mismatch of Greek *-φ-*, Celtic **-b-* as against **-p-* elsewhere; (3) the correspondence Gr. ῥ- ~ MDu. *r-*, in IE terms suggesting an initial **r-*, which is not typical of inherited vocabulary (e.g. Lehmann 1951).

If the Celtic forms are to be segmented **a-rb-*, then they may reflect a 'prefixed' variant of the pan-European word for 'turnip'. Despite the close formal parallel with the other cases of 'stem reduction', I still find it difficult to entirely rule out chance resemblance, given that very little material (i.e. **-rP-*) is being compared here.[8] Considering the broad geography, we must in any case be dealing with a cultural *Wanderwort*. In view of the narrow distribution of the *a*-prefixed variant, it seems more probable that it was formed locally on the basis of material loaned from elsewhere. In this case, perhaps this word could provide an indication of the productivity of *a*-prefixation among the pre-Indo-European languages of Northern Europe.

▸ † **'heron, stork'**. SCr. *róda* 'white stork' (uncommon in the dialects; Skok III: 153) ~ Gr. ἐρῳδιός (since Homer) 'heron, egret', Lat. *ardea* 'heron'; ?ON *arta* (attested in Þul *Fugla*; for the meaning, cf. Ic. *urt* 'teal', Sw. *årta* 'garganey') (Beekes 2000a: 27; Лигорио 2012; Kroonen 2013: 36; Iversen & Kroonen 2017: 518; Pronk 2019a: 154; Matasović 2020: 339) — The appurtenance of the Norse word seems doubtful to me for semantic reasons. The comparison between the others is obviously attractive. It should be noted, however, that Latin *ardea*

8 In addition, Celtic rarely shows a prefixed form in such alternations. In this respect, note Schrijver's comparison of Gaul. *alauda* (*apud* Pliny et al.; see TLL for attestations), Old French *aloe* and OE *lāurice*, *lāwerce* 'lark', but I must admit that I am not entirely convinced by this equation (cf. Matasović 2020: 340).

could be syncopated < **arVdejā-* (cf. Walde/Pokorny I: 146–147). In this case, the comparison with Greek need not be considered an unambiguous example of 'stem reduction', although the correlation between Latin *a-* and Greek ἐ- remains irregular. The supposition of an **a-* ∞ **∅-* alternation, on the other hand, depends entirely on the Serbo-Croatian form.[9]

Schrijver (1991: 65) has suggested that SCr. *róda* is a Romance loanword. Лигорио (2012: 23–25) rightfully criticizes this theory, since (1) Romance **-dja* would give SCr. *-ža* (cf. *lòpīž* 'earthen cooking pot' < (*)*lapideum*, FEW V: 160; M. Matasović 2011: 165); (2) the supposed Romance word **arda* is nowhere attested; (3) the expected reflex of **arC-* in South Slavic would be **raC-*. At the same time, the word's isolation certainly does give us cause for doubt (Matasović 2020: 339). In view of the widespread European and Slavic association of the stork with childbirth, one might, for instance, envisage a connection with SCr. *ròditi* 'give birth' (cf. Liewehr 1954: 90). Due to the doubtful Proto-Slavic status of this word, this example is too uncertain to be used here.

7.1.2 **a-* ∞ **∅-* *with No Stem Reduction*

▸ **'nut'.** Pr. E *buccareisis* 'beechnut'; Lt. *ríešutas*,[10] Lv. *riẽksts* 'nut' ~ R *opŕx*, Sln. *óręh* 'nut' (Fraenkel 1950b: 238; Polák 1955: 55; LEW [1965]: 731; Bezlaj II [1982]: 253; Matasović 2013: 93; Blažek 2018: 5–6) — Practically everyone who has accepted a non-IE origin has adduced additional forms such as Gr. κάρυον 'nut', ἄρυα (H.) 'walnuts', Alb. *arrë* 'walnut' in support of this claim. I am not optimistic about these comparisons; the only thing shared by ἄρυα and *ríešutas* is the phoneme **r*. Nevertheless, even without this evidence, the inexplicable initial vowel in Slavic allows us to make a fairly convincing case for foreign origin. It cannot contain the verbal prefix **ab-* (trad. **ob-*) 'around' because **-b-* would not have been lost before **-r-* (cf. OCS обрѣсти 'find, devise').[11] If we

9 The hapax ῥᾳδιός in Hipponax is probably due to aphaeresis, cf. μάσθλης (Hippocrates) 'leather' < ἱμάσθλη, and other examples in Strömberg (1944: 44–45), as well as θέλω < ἐθέλω 'want, wish'.

10 The oft-cited form *riešas* seems to derive from Miežinis (1894: 206), who has ⟨rieszas, rieszutas⟩. It seems doubtful that this is a genuine dialectal variant, and if genuine, it is probably a back formation. The suffix *-utas* is rare, but it is shared by *ãšutas* 'horse hair' and *degùtas* 'tar'. Note that both *ãšutas* and *ríešutas* decline as a consonant stem in East Vilniškiai dialects, as has been repeatedly pointed out (cf. Būga *apud* Trautmann 1923: 241; Specht 1947: 62; Fraenkel LEW 731; Ambrazas 1993: 57; Derksen 2015: 328), but this is of little value, as consonant stems became productive in this region (Zinkevičius 1966: 263).

11 This phonological issue is not even mentioned by ЭССЯ (XXIX: 71), who list the word under **obrěxъ*, and consider it (following Ильинский 1916: 153; Трубачев 1971: 65) a derivative of **rěšiti* 'to untie'. Note the criticism on this point by Крысько (2014: 104).

assume an ablauting **Hor-* : **Hr-* (Specht 1947: 62), this would imply that the element **-ois-* is a suffix, of which there is no indication.[12] A non-IE origin might also be favoured by the root structure (virtual **(H)roiHs-*; see 5.3.2).

▸ **? 'oats'**. Pr. E *wyse* 'oats' ~ Lt. *ãvižos*, Lv. *àuzas*; R *овёс*, Sln. *óvəs*; Lat. *avēna* 'oats' (Pronk/Pronk-Tiethoff 2018: 294–295; Kroonen et al. 2022: 19–20; for more refs. and discussion, see pp. 212–213) — Whether this example belongs here depends heavily on the analysis of the Prussian form, the only one in which initial **a-* is lacking. The interpretation of the Prussian data is unfortunately not straightforward. A similar word for 'oats' is attested in Simon Grunau's Prussian vocabulary: *wisge*. While we might be tempted to read /wizje/, the word must rather be identified with *wìzges* in Daukantas (Leskien 1891: 274; see LKŽ s.v. *vizgė* for additional data). On the other hand, *wyse* in the Elbing Vocabulary cannot be corrected to **wysge*, as its reality is confirmed by the grain tax *craysewisse* found in 15th century East Prussian documents.[13] The result is that we have little choice but to accept the existence of two dialectal synonyms for 'oats' in Prussian and Žemaitian. In light of this, it is plausible that the two forms influenced one another, and that pre-Prussian **awizē* lost its initial **a-* due to the influence of *wisge* (Pronk/Pronk-Tiethoff 2018: 294–295). The latter may be cognate with Latin *virga* 'shoot, twig, rod', OHG *wisc* · [*faeni*] *strues* 'bundle [of hay]' (LEW 1269).

A further potential piece of evidence for a form without an initial vowel can be found within Uralic. Although the comparison of Lat. **awe(T)sna* 'oats' with Karelian and East Finnish *vehnä* 'wheat' would seem almost obvious, I am not aware of any relationship having been proposed before.[14] The Karelian word cannot be separated from Md. EM *viš*, and Ma. E *wiste*, *wište*, NW *wištə*, meaning 'spelt'. Although the Mari sibilant is somewhat unexpected, the reflexes

12 Other Slavic examples are difficult to adduce here, as **ab-* (trad. **ob-*) can usually not be excluded on formal grounds. A case in point is OCS оскрѣдъ, Pl. *oskard* 'pickaxe' as against Pr. E *sturdis* ?*/skurdəs/ · bicke, where the Slavic prefix does not appear semantically motivated, but on formal grounds could represent **ab-*. Moreover, one is reminded of CS омлатъ beside OCS млатъ 'hammer' (cf. Mikkola 1898: 302). I leave such examples aside.

13 e.g. "schessel habir von czenden, phlughabir und craysewisse" 1431. For the data, see Töppen (1867: 151–152), who convincingly interprets the word as 'Heuhafer' (cf. Pr. E *crays* · hew).

14 I have now published the suggestion in Kroonen et al. (2022: 20). In addition, note that R. Matasović independently offered the same comparison during the workshop *Sub-Indo-European Europe* in August 2021. Koivulehto (2002: 592) has suggested a far less attractive loan etymology starting from IE **ḱueitnó-* 'white', a back-projection of Germanic **hwīta-* 'white', in turn the source of **hwaitja-* 'wheat'. Not only is the back-projection of this Germanic form to IE unwarranted, the suggested source also leaves the Uralic **š* unaccounted for.

seem largely consistent with a reconstruction **šn*.[15] However, the vocalism in Mari is irregular, and in Uralic terms would rather suggest **wäšnä* (cf. Aikio 2014a: 157 and *passim*). In this context, the very narrow distribution within Finnic seems to almost call for a loanword origin (cf. von Hertzen 1973: 90; Häkkinen/Lempiäinen 1996: 159). In combination, all of these facts solidify the impression of a word intrusive to Uralic.

Although there is a semantic discrepancy between 'oats' in Indo-European and 'wheat, spelt' in Uralic, it nevertheless seems quite possible that all of these forms derive from an earlier agricultural *Wanderwort*. If these words belong here, they would be a clear example of a form without initial *a-*. However, it is by no means clear that we are dealing with a morpheme **a-* or some other phenomenon, such as aphaeresis, which might be a symptom of the borrowing process resulting from the more restrictive phonotactics of Uralic.

▸ ? '**sedge**'. OE *secg*, Du. *zegge*, OHG *sahar* 'sedge'; OIr. *seisc*, MW *hesc* COLL. 'sedge'[16] ~ R *осóка*, Uk. *осокá* 'sedge' (Kroonen 2013: 421; Iversen/Kroonen 2017: 518) — The relationship between Germanic **sok-* and Celtic **se-sk-* exactly parallels the semantically similar Lt. *néndrė* 'reed' (< **ne-nd-*) beside Hitt. *nāta-* 'reed', MP *nʾy* 'pipe, flute' (< **nod-*; de Bernardo Stempel 1999: 68).[17] These words therefore permit an analysis in terms of Indo-European morphology. The Slavic word formally corresponds to Lt. *ãšaka* 'husk of grain; small fish bone', which is

15 Compare Võ. *pähn*, Md. M *pä(k)šä*, Ma. E *pište* 'lime tree' (< PU **pä(k)šnä*) and F dial. *hähnä*, Md. M dial. *šäkši* and Ma. E *šište* 'woodpecker' (?< **šä(k)ćnä*). It seems possible that the **k* in these Mordvin words is secondary (P. Kallio *apud* Holopainen 2019: 249, but note the different reflex in the word for 'spelt'), which would be supported by the Mari reflex of **pä(k)šnä*, where we fail to find the otherwise regular development **ä* > **ü* before **kš* (cf. Aikio 2014a: 155). Traditionally (e.g. UEW 716), the word for 'lime tree' is reconstructed as **päkšnä* with a three-consonant cluster. Note that UEW (p. 772) rejects the appurtenance of *šäkši* altogether. Perhaps as a result, the Mordvin word for 'woodpecker' is not even mentioned by Aikio (2015a: 44; in prep. 108–109), but it seems that it must belong here and that we should return to UEW's reconstruction **šäćnä* / **ćäšnä*? with various assimilations (as opposed to Aikio's preferred *ćäćnä*).

16 The Celtic forms have been adopted into Romance, cf. Occitan *sesca* 'bulrush', while Ibero-Romance (Spanish and Catalan *sisca*, *xisca* 'reed') suggests a divergent preform with **ī*. Coromines/Pascual (v: 264) attempt to solve this by assuming a borrowing through Mosarabian (see also FEW xi: 551).

17 Reduplication is also shown by OIr. *nenaid* as against Lt. *notrẽ* 'nettle' (see p. 203). Compare similarly Lt. *papařtis* (dial. *papártis*) 'fern' against MIr. *raith* 'ferns, bracken' < **prH-ti-* (Schrijver 1995a: 178; Zair 2012: 76), if this does not contain the prefix *pa-* (Gliwa 2009: 82). Typologically speaking, one may assume that the reduplication seen in these plant names had a collective function, although there is no actual evidence for this (awkwardly, the unreduplicated *raith* is in fact a collective, but this fact is not decisive due to the productivity of collectives in Celtic).

generally derived from the root of *aštrùs* 'sharp' (cf. ME I: 142 with lit.). Such a derivation makes sense for 'sedge', which has sharp leaves. On the other hand, the suffix *-*akā* has no close parallels within Slavic (Vaillant 1974: 543; though cf. Варбот 1984: 167), which remains an argument against an internal derivation (the suffix is also rare in Baltic, cf. Skardžius 1941: 125–126). As a result, although a non-IE origin seems possible, this example remains uncertain.

▸ **? 'grouse'.** Lt. *jerubẽ* 'hazel grouse', Lv. *iŕbe* 'partridge'; Pl. *jarząbek*, Sln. *jerębь* 'hazel grouse' ~ Lv. *rubenis* 'black grouse'; OHG *reba-huon*, MLG *rap-hōn* 'partridge' (Derksen 2000; see pp. 174–176 for more discussion) — The Baltic evidence implies an initial **je-* (or possibly **ę̄-*; see p. 175), but Lv. *rubenis* 'black grouse' lacks the initial syllable altogether. At first sight, the latter is a close match to ON *rjúpa*, Nw. *rype* 'ptarmigan' (Walde/Pokorny II: 360; Derksen 2000). Jóhannesson (1942: 223) has called the appurtenance of the Norse words into question, however, deriving them instead from the root of Ic. *ropa* 'belch; brag', synchronically used to describe the sound the ptarmigan makes during the breeding season (see also IEW 871; Kroonen 2013: 411).

This explanation chimes well with the alternative etymology deriving Lv. *rubenis* from *rubinât* 'kollern, falzen (von Birkhähnen)' (ME III: 552). On the other hand, it is reasonable to suspect that *rubinât* is denominal in origin (LEW 744 refers us to Fraenkel 1937: 362, where the parallel Cz. dial. *křepeliti* (Kott VI: 727) 'twitter (of a quail)' < *křepel* 'quail' is cited; cf. Derksen 2000: 81). This would be supported by the verb's isolation within Baltic.[18] In Germanic, the verb is unlikely to be denominal, as the primary meaning of Ic. *ropa* appears to be 'belch', cf. MDu. *ruppen* in the same sense, and the derived OHG *ropfezzen*, MDu. *op-ruspen*. The Germanic word is presumably of imitative origin (Kroonen 2013: 411).

Even if the position of ON *rjúpa* 'ptarmigan' remains unclear, it still seems attractive to compare the Balto-Slavic data with the West Germanic words for 'partridge' (Derksen 2000: 77, 79). While MLG *rap-hōn* can be explained as a folk etymological distortion after *rap* 'fast, agile' (DWb XIV: 334; Suolahti 1909: 256; Kluge/Seebold 2011: 750), OHG *reba-huon* 'partridge' does not have a convincing ulterior etymology (Suolahti mentions the call: 'zirrep'). The comparison with the Balto-Slavic forms presupposes (a) the irregular loss of the first syllable **i̯e-* and (b) an irregular vocalic relationship **e* ∞ **u*. On the latter, see 7.3.2.

---

18 Its supposed Lithuanian cognate, *rubėti*, is apparently attested just once in a list of Švėkšna dialect words sent into the newspaper *Viltis* by a K. Jazdauskis: "Rubėti, brazdėti, grumėti, bildėti — "sinonimai" " (see *Viltis* 1908, No. 114 [1 Oct.], p. 3). *Rubėti* (if not simply a printing error!) is evidently a secondary variant of the synonymous *rabėti* (LKŽ).

7.1.3 *wi-?

▸ **'bison'**. Pr. E ⟨wissambs'⟩ · ewer; OE *wesend, weosend*, OHG *wisunt* 'bison' ~ R *зубр*, Pl. obs. *zą̑br*, Sln. *zǫ̑bər*; Lt. *stum̃bras* 'bison' (for refs. and further discussion, see pp. 220–221) — The element *wi-* in Prussian and Germanic is unexplained and difficult to account for without recourse to *ad hoc* contaminations (cf. LEW 932 with lit.).[19] Compounded with the other irregularities (pp. 220–221), a non-inherited origin looks probable. As a potential, although speculative, parallel for the prefix **wi-*, Kroonen (2013: 457; cf. also Šorgo 2020: 449; van Sluis forthc.) has adduced the Gaulish *uisumarus* · trifolium (*apud* Marcellus of Bordeaux; Delamarre 2003: 322–333), as opposed to MoIr. *seamair*, Ic. *smœra* 'clover' (this suggestion in fact goes back to Oštir 1930: 26).

Kroonen (2012: 254, 2013: 571) has also suggested that the initial *i-* in R *изю́брь* 'Manchurian wapiti' might be identified with the **wi-* in Prussian and Germanic. While at first sight attractive, this suggestion is probably to be rejected. First, it is suspicious that the given word is limited to Russian, and that the species it refers to is only present to the east of Lake Baikal.[20] It can be noted that there are a couple of other words in Siberian dialects which show an epenthetic /i-/ before /z-/, cf. Siberian dial. *иза́боль* 'indeed' (СРНГ XII: 84) = dial. *за́быль* (Аникин 2003: 201) and *изу́фрь* = MR *зуфь*, *зуфрь* (СРЯ 11–17 VI: 70) 'a kind of woollen fabric', of Turkic origin (cf. Turk. *sof* 'woollen fabric'; see Аникин 2000: 215, 220). This is probably to be explained by assuming the interference of a substrate in which initial /z-/ is not permitted, cf. Khakas *izep* 'pocket' ← R dial. *зепь* (itself borrowed through a Turkic language, ultimately from Arabic, cf. Räsänen 1969: 124; Аникин 2003: 216), Yakut dial. *ïhīr* 'fat' ← R *жир* (Аникин 2003: 199).[21] Without the support of *изю́брь*, it is also difficult to assess whether this initial **wi-* can have anything to do with the initial /i/ in R *и́волга* 'oriole' (see p. 179 for a discussion of this element).

---

19 For example, Petersson (1921: 39–41) assumes the Prussian form has *wi-* after German *Wisent*, and connects the latter with Skt. *viṣā́ṇa-* 'horn' (thus also van der Meulen *apud* Derksen 2015: 433). Ильинский (1926: 56) assumes instead that the Prussian word itself is cognate with Skt. *viṣā́ṇa-*, with a second element **bʰr-* 'bearing'. Young (1998: 204–205) sees the element **wis-* in OHG *wisa* 'meadow'.

20 The word's earliest attestation is in the derivative *изубрина* (1495, СРЯ 11–17 VI: 209) in a report from a Moscow delegation about a mission to the Grand Dutchy of Lithuania. Logically, the meaning must be 'bison meat', but the specific context ("три бочки **и**зубрины") makes it tempting to assume a transmission error.

21 An epenthetic initial *i-* is also found in some Turkic loanwords already in Middle Russian: *изарбавъ* (17th c.; СРЯ 11–17 VI: 92) 'brocade' (~ Ottoman Turkish *zerbaf*), *изумрутъ* (15th c.; idem: 212) 'emerald' (~ Turk. *zümrüt*), which might suggest transmission through a similar substrate.

TABLE 14 Possible examples of 'prefixation'

	Baltic	Slavic	Germanic	Elsewhere
'swan (1)'	–	*a-lband^{h}-	*a-lb^{h}ed-	
		*leb^{h}ed^{h}		
'turnip'	*râp-	*rêp-/*raip-	*rāp-	Celt. *a-rb-
? 'sturgeon'	*e-(ḱ)sket-r-	*e-(ḱ)set-r-	*(k)st(u)r-	Lat. *e-ksket-r-
? 'elm (2)'	–	*i-lim-	*e-lm-	Tur. *e-lm-
			*a-lm-	Celt. *lĕm-
'nut'	*rais-	*a-rais-	–	
? 'oats'	*a-uiź-	*a-uiS-	–	Lat. *a-wesn-
	*u̯iź-			Ural. *wešnä
? 'sedge'	–	*a-sak-	*sak-i̯-	Celt. *se-sk-
? 'grouse'	*ie-rub^{h}-	*(i)e-remb^{h}-	*reb^{h}-	
bison	*u̯i-sam(b^{h})r-	*zam(b^{h})r-	*u̯i-(t)snT-	
	*stum(b^{h})r-			
? 'boar'	? *u̯-epr-i̯-	*u̯-epr-i̯-	*epr̥-	Lat. *aper-

▸ **? 'boar'**. OCS (Ps. Sin.) вепрь 'boar'; Lv. *vepris* 'castrated boar' ~ Lat. *aper*, OE *eofor*, OHG *ebur** (attested *epur*, *eber*) 'wild boar' (Machek 1968: 684; Polomé 1990: 335; Kroonen 2013: 114; Šorgo 2020: 438) — The comparison of these forms is obvious, although an explanation of the initial **w*- in Balto-Slavic is lacking (Walde/Hofmann 1: 56; Kluge/Seebold: 226; Derksen 2008: 515). Perhaps this **w*- can be identified with the element **wi*- found in the word for 'bison', discussed above. True, the distribution of this 'prefix' in the two

examples is almost diametrically opposite, and without further examples, it is difficult to draw any solid conclusions.

7.1.3.1 Conclusion

The result of this section is a rather heterogenous group of mostly uncertain examples, which are collected in Table 14, above (see p. 185 for help reading the table). Shaded cells indicate forms containing a 'prefix'.

In view of the rather different distributions and behaviours of these words, it is unlikely that they represent a single phenomenon. Most reminiscent of Schrijver's '*a*-prefix' is the word for 'swan', which meets three criteria: (a) an alternation **a*- ∞ *∅-, (b) an apparent 'stem reduction' and (c) a geographical distribution similar to that of the European words for 'blackbird' and 'ore'. The word for 'turnip' also appears to fit this pattern quite well, although it is more widespread, and must have partially spread as a cultural *Wanderwort*. A similar phenomenon has also been proposed to occur in the words for 'elm' and 'sturgeon', but both of these involve a number of issues and cannot be considered entirely certain.

The remaining cases do show an initial vowel in some of the continuants, but do not show the expected pattern of 'stem reduction'. While it cannot be ruled out that such alternations derive from a related substrate alternation, it is difficult to rule out that they result from an unrelated phenomenon, such as aphaeresis, either resulting from the borrowing process or taking place within the source language.

7.2 Alternations between Front and Back Vowels

Although there are some words in Balto-Slavic which appear to show an unclear alternation between **a*- and **e*- in initial position (see, for instance, pp. 276–277 on the word for 'alder'), the value of this alternation is unclear as a result of 'Rozwadowski's change' (cf. Rozwadowski 1915; Andersen 1996a: 102–104 and *passim*; Derksen 2002) — the observation that Balto-Slavic **e*- sometimes occurs in place of **a*- under as of yet unclear conditions. Since the development is also found in inherited words, such as in Lt. *erẽlis* 'eagle' < **h*$_3$*er*- (cf. Gr. ὄρνις 'bird', Hitt. *hāran*- 'eagle'), I agree with Andersen (1996a: 105) that little is gained by invoking non-IE substrates. This applies to examples such as Lv. dial. (Kurzeme) *ẽrcis* ~ Gr. ἄρκευθος 'juniper' (Beekes 2000a: 27; Derksen 2015: 533–534), even though the unclear Greek suffix quite possibly suggests at least the Greek word is of non-IE origin (cf. also κέλευθος 'way, journey', Chantraine 1933: 366–367).

In other positions, the evidence may also be ambiguous. Cases like Lt. *vãsara* 'summer' ~ OCS (Ps. Sin.) весна 'spring' (= Gr. ἔαρ 'spring') and Lt. *vãkaras*, Lv. *vakars* ~ OCS вечеръ 'evening' must result from a combination of assimilation and neo-ablaut (Otrębski 1955: 24–26; Hamp 1970b). One may in principle suggest a similar explanation for examples such as Lt. *sidãbras*, Pr. III *sirablan* ~ OCS сьребро, even though, in this case, there is plenty of other evidence for a non-IE origin (see pp. 225–226). In the following examples, the irregular vocalism is supported by examples beyond just Baltic.

7.2.1 *Baltic* *e ∞ *Elsewhere* *a

▸ **'hellebore'.** RCS чемерь 'hellebore, hemlock; poison'; Bel. *чэ́мер* 'white hellebore', Mac. dial. *чемер* '*Veratrum lobelianum*' (cf. ЭССЯ IX: 52–53);[22] OHG *hemera* 'hellebore' ~ Gr. κάμαρος, κάμμαρος 'a poisonous plant: ?aconite, ?larkspur' (Furnée 1972: 343; Huld 1990: 405–406; Beekes 2000a: 28; Kroonen 2013: 219; Derksen 2015: 236; ERHJ I [2016]: 125) — The Greek vocalism and geminate -μ- suggest a foreign origin. Šorgo (2020: 440) has doubted the appurtenance of the Greek word due to its different meaning. However, since it also refers to poisonous meadow plants, this doubt hardly seems unjustified.

The underived word in Slavic has undergone various semantic shifts connected to the plant's poison — Bg. dial. *чѐмер* 'distress; demon', SCr. *čȅmēr* 'bitterness; distress; venom', Slk. *čemer* 'a kind of disease associated with blood clots' — while the botanical sense has been recharacterized with suffixes: R *чемери́ца*, Pl. (dated) *ciemierzyca*; Sln. *čmeríka*, Bg. *чемерѝка*. The usual sense appears to be *Veratrum* ('white' or 'false' hellebore). Similarly, the sense *Veratrum album* is recorded in German for the Carinthian dialectal form *hammer* (Grimm DWb X: 316), matching the gloss of *hemern*, *hemer-wurz* in the 18th c. *Polyglotten-Lexicon der Naturgeschichte* (op. cit. 983).[23]

22 As I argue in detail elsewhere (Jakob forthc. b.), Lt. *kẽmeras* 'hemp-agrimony', although belonging to the modern standard language, originated as a ghost word. It was the result of Nesselmann's misinterpretation of the form *Kiemerai. Alpen* (ClG I: 73) as 'Alpkraut'. As Szyrwid translates the same word as *mára, incubus ephi*[*a*]*ltes* (see ALEW² s.v. *kiemerai*), *Alpen* is clearly to be understood as the plural of *Alp* 'daemon, incubus'. Nesselmann's *Kėmerai* was misinterpreted by Kurschat (1883: 177) as *kemerai*, whence it found its way into botanical reference literature and finally into the standard language.

23 The use of the word with reference to true Hellebores (*Helleborus sp.*; cf. Marzell IV: 1016) is perhaps due to the influence of classical nomenclature. For instance, the 13th c. Breslau Arzneibuch distinguishes the white and black *hemern* (MWb s.v. *hemere*), an obvious calque on Lat. *helleborus albus* and *niger*.

▸ **'ramsons'**. Lt. obs. *kermušis* 'wild garlic' (ClG I: 1088, see ALEW 553; Ruhig I: 59);[24] OCS чрѣмошъ ACC.SG. (Rosenschon 1993: 150), R *черемша́*, Pl. *trzemucha*, SCr. dial. (Lika) *crìjemuša* 'ramsons' ~ OE *hramsa*, MLG *ramese* (MoLG *Rāmsche*, Marzell I: 211); MW *craf* 'ramsons' (Machek 1950b: 158; Beekes 2000a: 29; Matasović 2009: 222, 2013: 89) — Further, with *o*-vocalism, Gr. κρόμμυον 'onion'. The Greek geminate is difficult to explain from an Indo-European perspective (Chantraine DELG II: 586; the Epic variant κρόμυον may be metrical; LSJ s.v.) and itself already points to foreign origin. As a result, the reconstruction of an ablauting *u*-stem **ḱremH-u-* : **ḱrmH-eu-* (Matasović 2005: 369; and already Hamp 1965: 232) is beside the point.

The original Slavic form is difficult to reconstruct: while R *черемша́* suggests an underlying **čermušā-* (trad. **čermъša*), the OCS hapax чрѣмошъ — which appears to match Sln. *črȇmoš* (Pleteršnik I: 109) and SCr. dial. *cremoš* (also widespread in toponymy, cf. Skok I: 273) — is not consistent with a medial **u* (trad. **ъ*; despite Sławski SP I: 154; ЭССЯ IV: 68). The word admittedly appears in a whole host of corrupt forms in South Slavic (cf. e.g. SCr. *srȉjemuž*,[25] Sln. *čémaž* (SSKJ²) 'ramsons'), but if the South Slavic form **čermaše-* (trad. **čermošь*) is old, then the suffix syllable does not match that of Lt. *kermušis*. The evidential value of this alternation is of course low, but there are other irregularities which make a hypothesis of non-IE origin probable.

Chantraine (DELG II: 586) explains the variant κρέμυον, attested in Hesychius, along with MoGr. κρεμμύδι 'onion' as the result of a labial dissimilation **o–u* > **e–u*. Some similar cases are indeed known from Modern Greek (cf. αλεπού 'fox' < MGr. ἀλωπού; cf. Holton et al. 2020: 68), but the development is by no means regular. MIr. *crem*, also *crim*, 'wild garlic' has also been interpreted as secondary for **kramu-* (> MW *craf*), like OIr. *tel* beside *taul* 'forehead; boss of a shield' (Thurneysen 1946: 52; Bernardo Stempel 1987: 101). In favour of this, one has cited the personal name *Craumthann*, which is a rare variant of *Crimthann* (see eDIL s.v.). The absence of spellings in *-au-* or *-u-* for the plant name itself is

24 The form **kermùšė* (Kurschat 1883: 178, marked as an unfamiliar word; also the citation form in Trautmann 1923: 128; LEW 243; Derksen 2015: 239–240; etc.) is dubious and seems to derive ultimately from Mielcke (1800 I: 116) who has *Kermußės* f. wilder Knoblauch, apparently miscopied from Ruhig (the German–Lithuanian part still has *Kermußis* in this sense, cf. idem II: 303). The dial. *kermušẽ*, *kérmušas* (Juška III: 85) 'tip of a drill', as shown by *kiáurmušis grąžtas*, literally 'through-beating drill' (Gegrėnai, LKŽ), is an unrelated compound consisting of *mùšti* 'beat' (cf. LEW 243) and *kiáurai* 'through'. For the phonological development, compare dial. *kelt̃vartis* Veliuona, Seredžius < *kiaũl-tvartis* 'pig sty' (Skardžius 1941: 427).

25 ЭССЯ (IV: 68) and ALEW (1175) seem certainly correct to dismiss this form as evidence for a Proto-Slavic variant with **s-*.

disturbing, but as the word is not attested particularly well or early, this might be put down to chance. According to Stifter (1998: 227, fn. 2), such fluctuation in vocalism is ultimately the result of the *i*-mutation of a second-syllable **u*. As such an explanation does not appear to be viable here, whether *crem* can be accounted for in this way remains uncertain.[26]

While the explanations in both cases are admittedly shaky, it must be acknowledged that the evidence for a stem **kremu-* outside of Balto-Slavic is of an uneven and marginal nature, and we may tentatively operate with an original distribution of **e* in Balto-Slavic, **a* in Celtic (and Germanic) and **o* in Greek.

Purely on the basis of the formal similarity, it hardly seems possible to separate Uk. *черéмха*, Cz. *střemcha*, Sln. *čŕ̨emsa*, obs. *čremha* (i.e. ⟨zhremha⟩ in Jarnik) 'bird cherry'.[27] Berneker's interpretation (I: 145) that the common factor is the strong smell (the bark of the bird cherry has a strong, acrid smell) has generally been followed by later authors (e.g. Walde/Pokorny I: 426; REW III: 321; ALEW 553–554). In Baltic, one finds Lt. *šermùkšnis*, Lv. *sę̀rmaûksis* (~ *sę̀rmûkslis*, etc.; ME I: 829–830) 'rowan', whose initial **š-* has been attributed to *Gutturalwechsel* (e.g. LEW 243; Derksen 2015: 240).[28] In view of the consistent meaning of the word outside of Balto-Slavic, the transference to a tree name must be considered secondary. It is therefore not of direct relevance to the word's ultimate origin.

▸ **'garlic'**. R *чеснóк*, Cz. *česnek*, Sln. *čésən* 'garlic' ~ OIr. *cainnenn* 'garlic; leek', OW *cennin* 'leeks (COLL.)' (< **kasnīnā-*) (Schrijver 1995b: 16–18; Derksen 2008: 86; Matasović 2009: 193, 2013: 89; van Sluis forthc.) — Schrijver challenged the old interpretation of the Slavic word as a derivative of the root for 'scratch,

26 It is interesting, however, that eDIL (s.v. *tul*) cites a NOM.SG. *taul* beside DAT.PL. *telaib* 'boss of a shield' from the Middle Irish *Lebor na hUidre*. The latter form would actually contain an *i*-mutating factor in its ending **-bi*, and one might wonder whether this could represent the original distribution of the variants.

27 There is no benefit in treating R dial. (Tver', Даль² IV: 610) *черéма* (= *черёма*, Даль³ IV: 1312) as the oldest variant (*pace* Sławski SP: 153; Matasović 2005: 369); this dialect variant is evidently back-formed from *черёмуха* or from dial. *черёмка* (e.g. СРГК V: 773). The latter is itself probably a corruption of *черёмха* after the diminutive suffix *-ка*; compare similarly Pl. dial. (Sławski op. cit.) *trzemka*, Cz. dial. (Machek 1968: 586) *střenka*.

28 ЭССЯ (IV: 68) would rather see the Baltic words as the result of an assimilation **k–š* > **š–š* and Matasović (2005: 369–370) assumes contamination of two originally distinct words. Strangely, Matasović assumes that it is the word for 'ramsons' that had **ḱ-*, which is precisely the opposite of what is found in Baltic. Note that the claim (ALEW 1175) that Lv. *cę̀rmaûksis* (with variants, ME I: 377–378, EH I: 268) 'rowan' agrees with Lt. *kermušis* is probably an illusion in view of the numerous other examples of secondary *c-* for **s-* cited by Endzelīns (1905: 183–185, 1923: 130–131).

comb' (cf. R *чесáть*; Miklosich 1886: 35; Meillet 1905: 453; REW III: 330; ЭССЯ IV: 89–90), and attractively compared the cited Celtic material. He pointed to the mismatch in vocalism as evidence that we are dealing with parallel borrowings from a substrate language.

Falileyev/Isaac (2003) have questioned Schrijver's appeal to substrate, and argue that Celtic could reflect a preform **ksno-* with *a*-epenthesis as in MW *adar* 'birds (PL.)' (< **ptarV-*; op. cit. 8, fn. 25; see also Zair 2012: 185, fn. 27 with references to earlier literature). On the other hand, it is questionable that epenthesis should have occurred in a cluster **ksn-*, especially in view of OIr. *sine* 'teat, nipple' (< **sfeṇio-*), which reflects a similar cluster **pst-* (cf. YAv. *fštāna-* 'breast', IEW 990). Moreover, the semantic association with 'scratch, comb' is tenuous. Berneker (I: 151) adduced OHG *kloba-louh** 'garlic' ~ *klioban* 'split' as a supposed parallel, but the Slavic forms do not mean 'split'; occasional senses like 'pluck (leaves, feathers)' (RJA I: 946) are unusual and clearly secondary. The basic meaning is 'comb' » 'scratch' (see also SCr. *kòsa* 'hair'; Hitt. *kis-^{zi}* 'to comb', IEW 585–586; Kloekhorst 2008: 481–482).[29]

7.2.2 *Balto-Slavic *ă ∞ *ĕ Elsewhere*

▸ [**'cottage'**. Lt. *trobà* 'peasant house; room' ~ Oscan tríibúm ACC.SG. 'house', OIr. *treb* 'residence, estate' (etc.) — See pp. 270–271.]

▸ [**'ground elder'**. Lt. *garšvà* ~ OHG *gires* — See p. 277.]

▸ **'honeycomb'**. Lt. *korỹs*, Lv. *kāre*[30] 'honeycomb' ~ Gr. κηρός 'wax; honeycomb', Lat. *cēra* 'wax' (Chantraine 1933: 371; Alessio 1946: 161–162; Ernout/Meillet [1951]: 114; Deroy 1956: 190; Pisani 1968: 19; Beekes 2010: 689–670; van Sluis 2022: 17–18; Kroonen forthc.)[31] — The Baltic word can only be connected

29 Falileyev/Isaac (op. cit. 5–6) also adduce some forms from Uralic languages: Komi (Permjak) *komiź*, (Jaźva) *ku·mić*, Udm. *kumiź* 'wild garlic' (< Proto-Permic **kǫmiź*, cf. Лыткин 1964: 47); Mansi (West) *kɔśśm*, (North) *χōsman* (< **kāšmā-*) 'onion', Hungarian *hagyma* 'onion'. These forms possibly reflect a common proto-form **kaćmə* (cf. UEW 164–165; Живлов 2014: 130), although the non-canonical phonotactics (**-CR-* cluster) make it improbable we are dealing with an inherited word in Uralic (Holopainen 2022: 106). One might be tempted to consider the Uralic words continuants of the same pre-European source word, but in view of the only approximate formal match and geographical distance, it is more probable that the similarity is coincidental.

30 LVPPV has *kàre*, which ME (II: 195) cite from Plāņi. This would be consistent with *kâre*[2] Dunika (EH I: 602) and High Latvian *kàre*[2]. The latter, however, could also correspond to dial. *kãres* (ME loc. cit.). Establishing the original intonation is difficult.

31 Adams (2013: 694) tentatively compares the Tocharian B hapax *śeriye*, but admits that his gloss '± wax, honeycomb' is based entirely on the comparison with Greek. The word is attested in a list of medical ingredients in a broken context, and no translation is attempted in the recent critical edition by Tatsushi Tamai (2020).

with the Greek if the latter was borrowed from Ionic-Attic into the other dialects, and Lat. *cēra* from Greek (Boisacq 1916: 450; Walde/Hofmann I: 202), but this possibility is generally viewed with scepticism (Ernout/Meillet 114; Chantraine DELG II: 527).[32] The suffixation in the derivative κήρινθος 'propolis' is usually mentioned as a key indicator of the word's non-Indo-European origin.

A possible *Wanderwort* is supported by the unclear Volgaic comparanda, on which see p. 133. In addition to these, we may also note Estonian *kärg* (dial. *käri*, Saaremaa *kärv*) < **kärjeh*[33] 'honeycomb' (cf. Vaba 1990b: 176–177), another form with front vocalism. In view of the difference in vocalism, a direct loanword from Baltic appears unlikely (compare 3.5.2), and for geographical reasons, a borrowing from a Turkic or Volga Uralic language is also extremely improbable. In conclusion, it must be admitted that the exact routes of movement of this word remain quite unclear, but the mismatch in vocalism between Greek and Baltic may nevertheless be used to support the analysis as a non-IE loanword.

▸ **? 'circle'.** OCS кржгъ 'circle' ~ ON *hringr*, MDu. *rinc* 'circle, curve' (Philippa et al. III [2007]: 132 s.v. *kring*) — The Germanic and Slavic words could in principle be combined by assuming apophonic variants **e* : **o*. However, this lacks a clear motivation, and the implied root **kreng^h-* would violate the Indo-European root constraints (Kroonen 2013: 247; see 5.3). Philippa et al. point to the Germanic variant **kringa-* (ON *kringr*, usu. *í/um kring* '(all) around', MHG *krinc*, *kranc* 'circle, vicinity', MDu. *crinc* 'circle, curve') and suggest the possibility that the word was borrowed from a substrate. However, it seems more attractive to interpret this variant as a *Reimbildung* based on the verbal root seen in MDu. *cringhen* 'turn (back)' (which is cognate with Lt. *grę̃žti* 'turn; bore, drill'; Stang 1972: 24). A non-IE origin might still be supported by the **e* ∞ **a* alternation, but this is naturally rather meagre evidence.

▸ **? 'people'.** Lt. *tautà*, Lv. *tàuta*; Pr. E *tauto* 'people' ~ Go. *þiuda*, OE *þēod*, OHG *diot*, *diota* 'nation, people' (Beekes 1998: 461–463; de Vaan 2008: 618; Derksen 2010: 38, 2015: 461) — The Germanic forms point to **teut-*, which could also account for Oscan *touto* 'civitas', OIr. *túath*, MW *tut* 'people, country'. This reconstruction would be supported by Venetic (Làgole) *teuta* 'civitas(?)'

32 Although denying the possibility of Proto-Greek **ā*, Frisk (I: 843–844) is still inclined to view the Latin word as a loan from Greek. There really seems to be no positive evidence for this (see already Osthoff 1901: 22) and the Lat. *-a* remains a potential obstacle to the loan etymology (Ernout/Meillet 114).

33 An original *käri*, GEN. *kärje* (which probably reflects **kärjeh*, like *puri* GEN. *purje* 'sail' < **purjeh*, cf. p. 90) has undergone various analogical reshufflings. The standard form *kärg* is built after cases such as *jälg*, GEN. *jälje* 'trail, track' = F *jälki*.

(Lejeune 1974: 110–111) and the Gaulish **teuta* reconstructed on the basis of onomastic evidence (Delamarre 2003: 295).[34]

The Baltic forms have often been interpreted as showing the regular reflex of **teut-* (e.g. Brugmann 1897: 202; Stang 1966: 73). According to the formulation of Zubatý (1898: 396) and Endzelīns (1911: 82–83; cf. also Vaillant 1950: 123) the development **eu* > **au* was regular before a consonant if there was a back vowel in the following syllable, but the only decent example of this is precisely Lt. *tautà* (cf. Endzelīns 1911: 83).[35] All other clear evidence points to the preservation of **euC* in Baltic (Berneker 1899; Pedersen 1935; Kortlandt 1979a: 57; Derksen 2010: 37).

If this rule is rejected, then one might assume apophonic variants (e.g. Endzelīns 1911: 82); however, nouns with *e*/*o* ablaut were rare in PIE (cf. recently Kloekhorst 2014: 151–161; van Beek 2018). Petit (2000: 143) has suggested that the unexpected Baltic vocalism might be due to neo-ablaut on the basis of other feminine *a*-stems, but then one wonders why other nouns of a similar structure (such as Lt. *žiáunos* 'gills' and *liaukà* 'gland', cited by Petit himself) were not subject to this analogical pressure. In sum, the Baltic vocalism still lacks a satisfactory explanation, so that we might consider Beekes' account as a non-Indo-European loanword a possible option. It should be admitted, however, that the words in all the other European languages can go back to a common proto-form.

▸ † **'bull'**. Lt. *tauras* 'buffalo, aurochs' (Bretke, Morkūnas; see ALEW 1248), dial. *taũris* 'calf, bull (vel sim.)' (cf. Arumaa 1930: 19, No. 2; LKŽ),[36] Pr. E *tauris* · wesant (for 'aurochs'?, PKEŽ IV: 186; Young 1998: 201–203); OCS тоуръ 'bull', Cz. *tur* 'bovid', OPl. *tur* · bubalus (SSP IX: 227–228) ~ ON *þjórr* '(young) bull' (Ipsen 1924: 227–228; Beekes 2000a: 30; Kroonen 2012: 250, 2013: 478, 540; Šorgo 2020:

34 Van der Staaij (1975: 197–198) has argued that Venetic *eu* is a secondary, dialectal phenomenon due to its geographical distribution, although the absence of early examples with *-eu-* may simply be due to the absence of early evidence from the relevant regions (cf. Lejeune 1974: 111). Matasović (2009: 386) has suggested that Gaul. *Teut-* is "just a spelling variant", but there seems to be other evidence for the preservation of **-eu-* in Gaulish. (cf. *Leucetius* epithet of Mars ~ OIr. *lóichet* 'lightning; gleam', Go. *liuhaþ* 'light', Delamarre 2003: 200; *Neuio-dunon*, placename in Pannonia, cf. *Nouio-dunum* 'Neu-Châtel', op. cit. 236).

35 As for Lt. *laũkas*, Lv. *làuks* 'blaze-faced', there is no reason to prefer a direct equation with Gr. λευκός 'bright, clear, white' over an *o*-grade adjective of the type Lt. *raũdas* 'reddish brown', Go. *rauþs** 'red' (cf. Berneker 1899: 164; Petit 2000: 120).

36 The accented form *taũras* of the standard language was apparently introduced by Būga (1912: 40–44). Interestingly, Būga had previously (RR II: 718) labelled **taũras* as erroneous; but has later defended it on the basis of toponymic evidence. Since the aurochs went extinct in the 17th century, this accented form must in any case be regarded a learned creation.

453; van Sluis forthc.) — With *a*-vocalism, cf. also Gr. ταῦρος, Lat. *taurus* 'bull' and Alb. *ter* 'steer' (with analogical umlaut from the plural; cf. Demiraj 1997: 384; Orel 2000: 224; Matzinger 2006: 56; Schumacher 2013: 228).[37]

OIr. *tarb*, MW *tarw* 'bull' reflect **tarwo-*. The metathesis is generally explained as the result of contamination (with **karwo-* 'deer', cf. MW *carw*, or less likely OIr. poet. *ferb* 'cow', Walde/Hofmann II: 651; LEIA T-31; de Bernardo Stempel 1999: 214–215).[38] By contrast, Latin lacks the expected metathesis to **tarvus*, which suggests a relatively recent loanword (cf. Трубачев 1960: 7; de Vaan 2008: 607; Weiss 2020: 170).[39] A possibility would be to assume a Sabellic origin, as with Lat. *bōs* 'cow', cf. Oscan ταυρομ ACC.SG. 'a sacrificial animal(?)' (Untermann 2000: 777–778). This would imply that the metathesis did not occur in Sabellic, although the argument is admittedly circular (there is no certain evidence either way).[40]

In Germanic, two similar words for 'bull' are found, both of which show *e*-vocalism. ON *þjórr* 'bull' (< **þeura-*) and Go. *stiur* 'ox, calf', OHG *stior* '(young) bull' (< **steura-*). Ipsen, and several others (see refs. above), have pointed to the fluctuation between **þ* and **st*, as well as the vocalic alternation, as evidence of a non-IE origin.

The older theory (Brugmann 1906: 353; Petersson 1921: 40–41; Walde/Pokorny I: 711; Mallory/Adams 1997: 135) separates the two Germanic words, taking **steura-* together with YAv. *staora-* 'pack animal; cattle', Parth. *ʿstwrʾn* /istōrān/ PL. 'cattle' (Durkin Meisterernst 2004: 91)[41] and assuming **þeura-* pro

37 Since a development **eu* > Alb. *e* has been widely assumed (e.g. Huld 1984: 155; Demiraj 1997: 46) one may be tempted, with Mallory/Adams 2006: 136, to equate Alb. *ter* directly with ON *þjórr*. However, Matzinger (2006: 57; cf. de Vaan 2018: 1739) has put this sound law into doubt. Furthermore, the expected vocalism appears to be preserved in Alb. *taroç*, *tarok* 'young bull'.

38 The regularity of the change **wr* > **rw* (Matasović 2009: 371) is uncertain. It is contradicted at least by OIr. *gúaire* ~ Lt. *gauraĩ* 'animal hair' and MIr. *glúair* (< **glauri-*) 'clear, bright' ~ ON *gløggr* (< **glawwa-*) 'clear, distinct; clever' (Zair 2012: 237).

39 De Vaan's treatment of this development as regular in Latin (cf. *alvus* 'belly', *nervus* 'sinew', *parvus* 'small' ~ Gr. αὐλός 'pipe, hollow tube'; νεῦρον, παῦρος) is certainly preferable to older notions of a 'sporadic' fluctuation (often assumed for PIE itself, thus Pedersen 1909: 176; Specht 1947: 35; Leumann 1977: 101). However, details need working out: cf. *caurus* 'northwest wind', *caulis* 'stalk, stem', *īnstaurō* 'repeat, restore'.

40 Sabellic shares **eu* > **ou* with Latin, and the development has usually been assumed for Proto-Italic (e.g. Brugmann 1897: 197; Leumann 1977: 61). Since *nervus* 'nerve, sinew' < **neuro-* must predate this change, this would force us to assume the metathesis was Proto-Italic, too. On the other hand, if the evidence for the preservation of **eu* into early Latin is taken seriously (see Weiss 2020: 112–113), this would imply the vowel development was independent in the two subfamilies (note also fn. 34, above, on Venetic).

41 Usually seen to be a derivative of the adjective in Skt. *sthūrá-* 'big, strong', Khot. *stura-* 'thick, large', MDu. *stuur* 'strong, fierce', OHG *stiuri* 'strong, proud' (e.g. IEW 1009–1010).

***þaura-* to be the result of a contamination with the former. It does indeed seem suspect to disregard this semantically and formally convincing match with Iranian in favour of an irregular comparison. If **steura-* and **þaura-* were in competition in Proto-Germanic, such a contamination seems quite imaginable.

A frequent argument in favour of a non-native origin is the existence of a similar word in Semitic, cf. Akkadian *šūru*, Aramaic *twr* */tōr/, Arabic *ṯawr*, Geʿez *sor* 'bull, ox' (< **ṯawr-*, Militarev/Kogan 2005: 307–309). While some have assumed a direct loan from Semitic into Indo-European (J. Schmidt 1891: 7; Vennemann 1995: 88–89) or in the other direction (e.g. Walde/Pokorny I: 711; Walde/Hofmann II: 651), several scholars have argued that the word is better derived from some other, third source (Feist 1913: 411; Ipsen 1924: 227–228; Schrader/Nehring II: 261; Гамкрелидзе/Иванов 1984: 519–520).[42]

Blažek (2003) has proposed a plethora of Afroasiatic comparanda, some of which were taken over by Militarev/Kogan (2005: 309–310). The latter additionally cite some Chadic words for 'elephant' and an isolated Kachama (Omotic) word for 'rhinoceros'. Leaving aside these semantically questionable comparisons, and since most of the Cushitic forms cited by Blažek reflect an unrelated **tsawadu* (Kießling/Mous 2003: 293), what we are left with is Ma'a/Mbugu *churú* 'bull' (cf. Militarev/Nikolaev 2020: 205–206). The evidence is therefore scarce, and if we add that Mous (1996: 202, 210) specifically warns against the reconstruction of proto-forms on the sole evidence of Ma'a, a mixed language with a complex history, the situation looks even less favourable. Thus, while the existence of Afroasiatic comparanda would more or less confirm a specifically Semitic → Indo-European loanword (Militarev/Nikolaev loc. cit.), the external evidence can hardly be relied on.[43]

To give better support to the Indo-European status of the word, Mallory/Adams (1997: 135; 2006: 140) adduce a Khotanese *ttura-* 'mountain goat', first included here by Bailey (1979: 132). True, as a late Khotanese word, *ttura-* could in theory reflect **taura-* (*pace* Simon 2008: 299) beside several other options (cf. Emmerick 1989: 212). However, the word is a hapax in a difficult passage

42 Delamarre (2003: 292), by contrast, considers the similarity coincidental.

43 Maarten Kossmann (p.c. March 2023) considers it possible to compare the Semitic word with Berber **a-zgăr* 'bull', providing (1) the **z* is assimilated from **s*, in which case the initial consonant correspondence would match Semitic **ṯn-* 'two' vs. Berber **sin* (vel sim.), and (2) the **g* is derived from **w*, which can be related to the known (but poorly understood) alternation between **w* and **g* in some Berber lemmas. For more detail, I refer to the original Twitter discussion between Kossmann (@ait_kisou) and @irzastan posted 8 February 2022. If this comparison is indeed correct, it would again tip the balance in favour of a Semitic → Indo-European loan.

whose meaning is far from certain (F. Dragoni p.c. November 2021),[44] and is therefore too unreliable to use. Абаев (1958: 390–391) has additionally mentioned Ossetic I *ʒæbīdyr*, D *ʒæbodur* 'West Caucasian ibex', but it is unclear whether *-dyr*, *-dur* can be segmented, as the first element remains obscure.[45]

Kroonen (2013: 540) associates the Germanic **eu* with the similar vowel attested in Etruscan *ϑevru-mineś* 'Minotaur'. The Etruscan *-ev-* is indeed problematic, and like the initial *ϑ-*, appears to rule out a direct loanword from Greek Μῑνώταυρος (cf. Fiesel 1928: 80–81; Kretschmer 1940: 266). On the other hand, due to the distance between the Germanic and Etruscan homelands (wherever the latter may be), it is unlikely the developments can be associated with each other. It is difficult to use the Etruscan evidence to support a non-IE origin, but it is not entirely clear where the word was adopted from.

In conclusion, as far as Indo-European goes, the word is limited to Europe, but the arguments in favour of its foreign origin are somewhat circumstantial. The irregular form in Germanic could be explained as resulting from a fairly well-grounded contamination, and therefore the word cannot be classed as a certain case of an **e* ∞ **a* alternation. A reconstruction **th$_2$eu-ro-* could account for most of the other data. Perhaps the strongest evidence for a non-IE origin remains the irregular Celtic form, for which yet another contamination must be assumed. How exactly the Etruscan and Semitic words fit into the picture is unclear. Due to the complexity of this example, I will leave it out of consideration here.

▸ † '**poppy**'. R *мак*, Sln. *màk*; OSw. *val-mogha*; Gr. μήκων (Doric μάκων) 'poppy' ~ OHG *maho*, *mago* 'poppy' (Beekes 2000a: 29; Boutkan 2003a: 15; Matasović 2013: 89) — For a discussion of Lv. *maguône* (etc.) 'poppy', which is most likely of Germanic origin, see p. 41.[46]

On account of MoHG *Mohn*, the OHG variant *maho* has often been attributed a long vowel (e.g. Kluge/Götze 396). As Kluge/Seebold (484) point out, an OHG *ā* (< PGm. **ē*) would hardly be compatible with the Doric Greek *-ά-*.

---

44 Skjærvø (2002: 35) does not attempt a translation.

45 In an earlier publication (1949: 49), Абаев considered the word to be of native Caucasian origin, adducing Karachay *ǯuǧutur*, to which we can add Kabardian *šəqʷɬtər* in the same sense. He segments the Karachay word *ǯuǧu-tur* (with the second element assumed to be ← R *myp*?), comparing Georgian *ǯixvi* 'West Caucasian ibex'. Just how a form of this shape could be borrowed into Ossetic as *ʒæb-* is unclear to me. Bailey (1979: 132) sees in the first element the word Oss. ID *ʒæbæx* 'good'. This species of Ibex is referred to as *myp* in Russian, which is quite a surprising semantic shift: perhaps it was encouraged by the similar-sounding Karachay word?

46 It seems much less probable that these Baltic forms could show evidence of a **g* ∞ **k* alternation (Oštir 1929: 107).

However, as argued by Schaffner (2001: 358–361), the apparent MoHG evidence for a long vowel may be explained as the result of an early contraction over **h* (cf. MHG *mān* beside *mahen*).[47] The evidence is therefore consistent with an ablauting **meh$_2$k-* : **mh$_2$k-* (cf. Kroonen 2011b: 311–314, 2013: 371).

The word for 'poppy' has often been suspected to be a prehistoric loanword on the basis of cultural facts (Būga 1924a: 18; Kluge/Götze 396; Machek 1950b: 158; Sabaliauskas 1960a: 71, 1990: 261). On the other hand, some scholars have seen the distribution and evidence for ablaut as a clear indication of an inherited origin (Frisk II: 225; Kluge/Seebold 484). As we cannot prove a loanword on formal grounds, this word cannot be considered here.

7.2.3 *u ∞ *i

▸ **'oven'.** Pr. E *Vumpîs* 'oven'; Go. *auhns**, OHG *ovan* 'oven' ~ Gr. ἰπνός 'oven, furnace' (Kroonen 2013: 557) — For a discussion of the Prussian and Germanic forms, see further p. 231. The comparison between Gr. ἰπνός and Germanic **ufna-* seems nigh inescapable. The old, traditional, equation with Skt. *ukhắ-* 'cooking pot', on the other hand, is phonologically impossible and must be rejected (see Frisk I: 732–733; EWA I: 210; Kroonen 2013: 557). For the Greek word, Vine (1999: 19–23) has suggested an alternative etymology starting from **sp-no-* with *i*-epenthesis to the root of Gr. ἕψω 'boil', Arm. *ep'em* 'cook' < **seps-*, yet there is no other evidence that the final **s* in this root is suffixal.[48]

According to Vine (1999: 22), the inscriptional ηιπνε[ύεσθαι] is "exceedingly difficult to explain away", and while indeed Threatte (1980: 494) writes that non-etymological h- is "virtually unknown in fifth-century Attic texts", some of the examples accepted by him as etymologically justified are perhaps not, e.g. ηακόσια 'unwilling' (before 460 BCE) < **n̥-u̯ekontia* (cf. Beekes 2010: 400), and a handful of examples are still acknowledged as irregular (Threatte 1980: 495) so that a single attestation can be considered insufficient to prove an initial aspirate. It seems more likely that the equation with OHG *ovan* and Gr. ἰπνός should be maintained, and that we should assume a non-IE alternation **i* ∞ **u*. Unfortunately, this alternation does not appear to be supported by other certain examples.

47 This seems to be paralleled by Middle German *stol* 'steel' (Elbing Vocabulary; see Chapter 2, fn. 4) < OHG *stăhal* (Swiss *Stăchel*, cf. *Schw. Id.* X: 1197).

48 As for the widely acknowledged constraint against two like stops in an IE root, this evidently did not apply to fricatives, cf. Skt. *sásti*, Hitt. *ses-^{zi}* 'sleep' (< **ses-*, LIV 536–537; Kloekhorst 2008: 746), Hitt. *huhha-*, Lat. *avus* 'grandfather' < **h$_2$euh$_2$-* (op. cit. 352), possibly **h$_3$neh$_3$-mn-* 'name' (see Beekes 1987; van Beek 2011: 52–53).

▸ **? 'cod'.** R *треска́* 'cod' ~ ON *þorskr*, MLG *dorsch* 'cod' — The other Slavic forms (Uk. *тріска́*, Bel. *траска́*; Cz. *treska*, Sln. *tréska*, SCr. *trȅska*, etc.) are all regarded as recent loanwords from Russian (cf. ЕСУМ v: 645; Machek 1968: 650; Bezlaj iv: 220), so we must base our conclusions on the Russian evidence alone. In СРЯ 11–17 (xxx: 131–132), the forms are normalized under трѣска[2], consistent with the traditional etymological equation with RCS трѣска 'splinter' (REW iii: 137). However, none of the citations are actually spelled with ⟨ѣ⟩[49] and the MR variant *троска* (СРЯ 11–17 xxx: 180; cf. dial. (Karelia) *тро́ска*, СРГК v: 518) might instead imply an earlier *трьска (cf. Pedersen 1895: 72) with *yer* umlaut (as in *тресть* ~ *трость* 'reed', see p. 201). However, note similarly MR *тръска* ~ *троска* 'sharpened stick; stake', which I cannot explain.

Although the word must be old in Norse (cf. the early loanword in F *turska*, E *tursk*, Li. *tūrska* 'cod'; LÄGLOS iii: 322–323), MLG *dorsch* 'cod' and MDu. *dorsch** (attested *dorssch*) 'a kind of fish' need not be inherited, and have been interpreted as Norse loanwords (Philippa et al. i: 615; Kluge/Seebold 212). However, on formal grounds, cognancy cannot be ruled out. Likewise, R *треска́* has been interpreted as a loanword from Germanic (Tamm *apud* de Vries 1962: 618; Machek 1968: 650, allegedly from "an unattested northern variant"; Kluge/Seebold 212), but this is phonologically implausible. If we set up an original **tresk-*, this could be combined with Germanic **þurska-* by reconstructing an ablauting **trsk-* : **tresk-*. However, if the Russian form goes back to an earlier **trisk-*, the connection cannot be maintained in Indo-European terms. Without other Slavic cognates, it is difficult to make a convincing case for a non-Indo-European origin. Even though reconstructing old ablaut for a noun in the sense 'cod' is questionable, this word cannot serve as certain evidence.

7.2.3.1 Conclusion

The certain and possible evidence front ∞ back vocalic alternations is collected in Table 15, overleaf (see p. 185 for help reading the table). Forms which do not provide relevant data are presented in light grey. Shaded cells indicate forms with back vocalism.

49 Зализняк (2019: 185) keeps *треска́* 'cod' and *трьска́* 'splinter' apart. The word first appears in 16th century North Russian monasterial accounting books, but it is possible it had been in use earlier among illiterate fishing populations.

TABLE 15 Possible examples of front ∞ back alternations

	Baltic	Slavic	Germanic	Elsewhere
'hellebore'	–	*kemer-	*kemer-	Gr. *kam(m)ar-
'ramsons'	*kermus-	*kermus-	*kra/omus-	Gr. *krommus-
		? *kermas-		Celt. *kram-
'garlic'	–	*kesn-	–	Celt. *kasn-
'cottage'	*trâB-	–	*t(u)rb-	It.-Celt. *trĕb-
'honeycomb'	*kâr-	–	–	Gr. *kēr-
				Tur. *käräs
'ground elder'	*Gârsu̯-	–	*g^{h}erVs-	
'oven'	*umnV-		*upno-	Gr. *ipno-
? 'circle'	–	*krang^{h}-	*kreng^{h}-	
? 'people'	*taut-	–	*teut-	It.-Celt. *teut-
? 'cod'	–	*tri/esk-	*t(u)rsk-	

It seems that we can identify two main groups. In the former, Balto-Slavic shows **e* as against **a* (or **o*) elsewhere. It is remarkable that all three examples of this alternation show an initial **k-*. If this is not mere coincidence, we might assume a phonetic solution. It is reminiscent of the situation in Turkic, where due to the allophonic change **k* > */q/ in back-vocalic contexts, loanwords with /k/ are automatically adopted with front vocalism, as Chuv. *kămpa*, dial. *kŏmpa*, Tat. *gömbä* 'mushroom' (← Sl. **gǭbā-*, trad. **gǫba*), Chuv. *kĕrpe*, dial. *kŏ̆rpe* 'groats', Tat. *körpä* 'bran' (← R *крупá*).[50]

50 In this respect, one might note the unexplained East Baltic words for 'marten' — Lt. *kiáunė*,

The second group shows the opposite distribution, with the back vocalism being limited to Balto-Slavic. The coherence of this group is less certain, as in each case the comparanda show a distinct pattern of correspondences.[51] In this context, compare also the potential examples of front/back alternations among the Finnic-Baltic isoglosses discussed in 3.5.3.

7.3 Alternations between Low and High Vowels

7.3.1 *Baltic/Slavic High Vowel ∞ Low Vowel Elsewhere*

▸ [**'bison'**. Lt. *stum̃bras* 'bison', Lv. *stum̃brs*, OHG *wisunt* 'aurochs' ~ Pr. E ⟨wissambs'⟩ · ewer — See the discussion on pp. 220–221.]

▸ [**'oriole'**. Pl. *wilga*, Sln. *vólga* ~ Lt. *volungẽ* 3ª, Lv. *vãluôdze*, ME *wode-wale*, MHG *wite-wal* 'oriole' — See the discussion on pp. 178–180.]

▸ **? 'heel'**. Lt. *kulkšnìs* 'ankle; hind heel of an animal', Lv. dial. (ME II: 307) *kulksnis* 'ankle joint (in animals); leg' ~ Lat. *calx* 'heel (of a person or animal)' (Matasović 2013: 89) — Because of *kulfchnis* NOM.PL. 'ankles' in Bretke, it is attractive to view the *-k-* as intrusive (Berneker I: 660; Derksen 2015: 262; ALEW 618); compare Lt. *kùlšis*, *kùlšė*; Lv. dial. (ME II: 308) *kulša* 'hip, loins', Pr. E *culczi* 'hip'.[52] With this, one would like to compare SCr. *kȕk*, Bg. *кѫ̀лка* 'thigh, hip',[53] although this would require an unmotivated *Gutturalwechsel* (Berneker loc. cit.; Trautmann 1923: 145). The vowel in Latin *calx* is difficult to derive from any reasonable IE preform (Ernout/Meillet 89; against Schrijver's rule **ke-* >

Lv. *caûna* — which point to *e*-grade, as opposed to the other Balto-Slavic forms, which would imply an old **a* or **o*, cf. Pr. E *caune*, MR *куна́* (Зализняк 2019: 206), Cz. *kuna*, SCr. *kúna* 'marten'. The word is of unclear derivation (Derksen 2015: 242; ALEW 560). True, in the above examples, the *e*-vocalism was found throughout Balto-Slavic, so whether the phenomenon is the same is uncertain. Here one could also mention Lt. *pélkė*; Pr. E *pelky* 'marsh' ~ Gr. (H.) παλκός · πηλός 'earth, mud' (cf. Alessio 1946: 160; van Beek 2013: 548, fn. 21; Derksen 2015: 349–350), but it is difficult to base much on a mere gloss.

51 A potential additional example within Slavic would be the word for 'swan' (see pp. 176–177), cf. Sln. *labȍ́d* ~ R *лебедь*, although here the additional nasal in the former must be factored in.

52 It is tempting to further compare Lt. *kul̃nas*, *kulnìs*, Lv. dial. (Rucava) *kulna* 'heel', which could reflect earlier **kulkna-* (IEW 928) as in Lt. *bal̃nas* 'saddle' (= Pr. E *balgnan*) with the blocking of palatalization before **n* (cf. Pr. E *balsinis* 'cushion'). However, the difference between Pr. E *kulnis* 'ankle' and *balgnan* 'saddle' speaks against this.

53 Only South Slavic. Hardly here belong R (нарах?) *колкъ* ("коло́къ?" *sic*. Даль² II: 139) 'bony stump of a bovine horn' and Cz. *kelka* '(arch.) stump of a limb; (hunting term) tail of a deer'. The latter does not show a regular reflex of **kulk-* (trad. **kъlk-*). Similarly, USrb. *kulka* 'ankle' cannot be from **kulk-* and is probably merely a diminutive of *kula* 'bulge' (cf. Schuster-Šewc 723).

**ka*-, see Meiser 1998: 82–83), but due to the difficulties in analysing the Balto-Slavic data, attributing **a* ∞ **u* to borrowing from a non-IE source language may be premature.

A potential parallel can be found in the equation between RCS громѣждь, гремѣждь (СРЯ 11–17 IV: 129)[54] and Lat. *grāmae* (TLL VI: 2165) 'rheum in the eye' (cf. de Vaan 2008: 270; Matasović 2013: 84). The original form of the RCS word is uncertain, but would be consistent with **grim*- (trad. **grъm*-; СДРЯ 602; Berneker I: 360; Sławski SP VIII: 267; РЭС XII: 78) or **grum*- (trad. **grъm*-; ЭССЯ VII: 159; Derksen 2008: 194). In the latter case,[55] we would be dealing with a similar **a* ∞ **u* alternation. The long vowel in Latin[56] could also be accounted for if we assume an underlying **gra*(*k*)*sm*-. In Slavic, **x* (< **ks* or **s* + RUKI) would have disappeared without a trace before a resonant as in **lūnā*- (trad. **luna*) 'moon' < **louksneh*$_2$- (Pronk 2018: 300); thus, Slavic could theoretically reflect an earlier **gru*(*k*)*sm*-.[57]

▸ ? '**salmon**'. Lt. *lašišà*, Lv. *lasis* 'salmon' ~ Pr. E *lalasso* */lasasā/, R *лосо́сь*, Pl. *łososь* 'salmon'. Further ON *lax*, OHG *lahs* 'salmon' (cf. Laumane 1973: 116; Ariste 1975: 468)[58] — A Lithuanian variant *lãšis* was recorded by Nesselmann (1851: 350, "bei Memel"). Similar forms have been recorded all along the Western coastline,[59] as well as near the Latvian border (*lãšė* Ylakiai, Kivyliai; Vanagienė 2014), and these correspond formally to Lv. *lasis*. Generally, Lt. *lašišà* has been

54 Further Slavic forms have an unclear initial **k*-: Sln. *krmẹ́žəlj*, in addition to which Sławski (SP XIII: 267) adduces SCr. Čak. dial. *kȓmež*, Kajk. dial. *kȓmeželj* (cf. *kȓmežalj* in the dictionary of Popović *apud* PCA X: 216). Both languages also attest a shorter form: Sln. *krmẹ́lj* (lexicographically recorded), SCr. dial. *kȓmelj* (a kind of haplology?).

55 In favour of **grum*- (trad. **grъm*-), we can note that the form with -*o*- is attested some two centuries earlier (contrary to the 14th century date usually given (e.g. СРЯ 11–17 loc. cit.), Жолобов (2007: 35) has convincingly argued that the RCS Parenesis of Ephrem the Syrian should be dated to the 13th century), and that -*e*- would be more easily explained as a secondary assimilation.

56 The word is rare, but the long vowel is metrically secured in Plautus. The derived adjective *grammō*(*n*)*sus* would therefore show the littera rule.

57 However, if we are willing to permit an alternation **l* ∞ **r* (see p. 189), it would seem obvious to compare Greek γλαμυρός 'bleary-eyed' (cf. H. γλάμος 'mucus', *glamae* 'rheum' in Paulus ex Festo). This would speak against a reconstruction such as **gra*(*k*)*sm*-.

58 In view of its initial *l*-, Oss. D *læsæg* (recorded only lexicographically, cf. "Не документировано", Абаев 1973: 32–33) must be a borrowing (Абаев 1965: 37–38). Similarly, Arm. *losdi* 'salmon', attested only in the Armenian-Latin dictionary of Stefanus Roszka (V. Petrosyan on en.wiktionary.org, s.v. լոսդի [16 September 2019]; cf. Martirosyan 2008: 312) is most likely a local adoption of Romanian *lostrița*, *lostița* 'Danube salmon'. The latter ultimately derives from Slavic (cf. Diebold 1976: 368).

59 cf. *lašìs* Rusnė (on the Neman), *lãšė* Kintai, *lãšis* Kukuliškiai (*Papildymų kartoteka*), *lašìs* Palanga (LKŽ).

considered a derivative of *lãšis* (Skardžius 1941: 317; Specht 1947: 31), but it seems more probable that *laš-* (and Latvian *las-*) derives from *lašiš-* (> **lašš-*) by syncope. Note, with a similar distribution, the syncopated Žemaitian *vẽčas* (Daukantas, Juška) and Lv. *vęcs* as against Lt. obs. *vetušas* 'old'. In Latvian, the development is closely paralleled by *tacis* 'fishing weir' (< **tacsis* < **tacisīs*, cf. Lt. *takišỹs*; see Chapter 3, fn. 136). Compare also, with a different distribution, Lt. *lopšỹs* 'cradle' as against dial. Žem. *lopišỹs*. Due to the fluctuation in stem type, Baltic **lašiš-* may be an original root noun.

The difference between East Baltic **lašiš-* and Slavic/Prussian **lašas/š-* is difficult to account for in Indo-European terms. True, there are a couple of East Baltic forms which might show *-a-*, which could suggest that the stem **lašiš-* is a recent development. Szyrwid has *łafaßa* 'salmo, łosoś', and *Lexicon Lithuanicum* has *laßaßa* (ALEW 647). However, neither of these forms are entirely reliable: the form in Szyrwid might have been influenced by the Polish equivalent (the first *-s-* seems to imply this), and the other form might be a Prussianism (cf. PKEŽ III: 31). Nevertheless, the value of this alternation is not entirely certain.

The Balto-Germanic word for 'salmon' has almost universally been considered cognate to Tocharian B *laks* 'fish' (Walde/Pokorny II: 381; IEW 653; Derksen 2015: 274–275), and this has been seen as important in discussing the Indo-European homeland (e.g. Diebold 1976; Mallory/Adams 2006: 146). However, the comparison is phonologically irregular, as the Tocharian form is only consistent with an earlier **-u-* (Ringe 1992: 92). I must agree with Pinault (2009: 241, fn. 74), that the only correct solution is to reject the Tocharian form as a cognate. We are therefore dealing with a circum-Baltic term for local fauna which might plausibly be explained as a loanword from an unknown source.

▸ [? **'grouse'**. Lt. *jerubẽ* 'hazel grouse', Lv. *rubenis* 'black grouse' ~ OHG *rebahuon*, MLG *rap-hōn* 'partridge' — Note the reverse distribution. See the discussion on p. 241.]

▸ † **'shoe'**. Lt. *kùrpė* 'clog, shoe', Lv. *kuřpe*; Pr. E *kurpe* 'shoe'; Cz. dial. (Kott I: 827) *krpě* 'Schneereifen', Sln. *kŕplja* 'snow-shoe' ~ Gr. καρβάτιναι 'shoes of undressed leather' (Furnée 1972: 146; Beekes 2000a: 28; Derksen 2008: 263) — This example could possibly show the alternation **u* ∞ **a*,[60] but the Balto-Slavic

60 Another potential example is Pr. E *spurglis* 'sparrow' as opposed to Go. *sparwa*, OHG *sparo* 'sparrow' (Matasović 2013: 87), but it cannot be excluded that these represent an ablauting **sprg^{wh}-* : **sporg^{wh}-*. The reconstruction remains uncertain in view of numerous, but all doubtful, Greek comparanda: σπέργουλος, πέργουλον (H.) 'a wild bird', σπαράσιον (H.) 'a bird resembling a sparrow' (Frisk II: 1130; Schrijver 1997: 304) and ψάρ 'starling' (Walde/Pokorny II: 666; Kroonen 2013: 466).

stem might just as well be identical to the first syllable of Gr. κρηπίς, PL. κρηπῖδες 'kind of half boot' (?< *$k^{w}rh_{1}p$-, with Greek dissimilation *k^{w}–p > *k–p; on which recently see van Beek 2022: 466).[61]

7.3.2 *Slavic* *e ∞ *Germanic* *u

▸ **'maple'**. R *клён*, Pl. *klon* 'maple', SCr. *klȅn* 'field maple' ~ ON *hlynr* (attested in kennings; cf. Ic. *hlynur* 'sycamore maple'), MLG *lönenholt* 'maplewood' (MoLG *Löhn* 'maple', see Marzell II: 73; whence probably MoHG obs. *Lehne*, cf. DWb XII: 1137) (Oštir 1930: 22; Machek 1950b: 154; Matasović 2013: 85, forthc.)[62] — The lexicographically attested SCr. dial. *kûn* 'maple' is usually taken to represent an old *$klьnъ$ (trad. *$klьnъ$; Berneker I: 512; REW I: 567; ЭССЯ IX: 195).[63] Considering the isolation of this dialect form, it seems difficult to justify reconstructing it for Proto-Slavic. Skok (II: 95) remarks on the similarity to the "Macedonian" κλινότροχον quoted by Theophrastus,[64] referring to a kind of maple, and suspects a localized borrowing.

The position of Lt. *klẽvas*, Lv. *kļava*, *kļavs* 'maple' is unclear. The analysis *$klen$-$u̯o$- (e.g. Oštir 1930: 68; Bańkowski 2000 I: 706) would be in contradiction to Žem. *tę́vas*, Lv. *tiêvs* 'thin' < *$tenh_{2}$-$u̯o$- (~ Lat. *tenuis* 'thin, fine'; cf. also ЭССЯ IX: 194). In fact, the Latvian evidence appears to suggest an earlier *$kljawa$- (Endzelīns 1911: 94; Stang 1972: 28–29), which could be supported by the absence of *l*-hardening in Lithuanian dialects (Zinkevičius 1966: 160). It is difficult to account for the alternation *$kliou$- ∞ *$klen$- even within a non-IE context — as ALEW (p. 584) points out, all we are left with is a common element *kl-. Therefore, I am inclined to leave the Baltic terms out of the comparison.

---

61 OIr. *cairem*, MW *cryd* 'shoemaker' have been compared and derived from *$kerh_{1}pio$- (LEIA C-21; Matasović 2009: 189–190); however, this reconstruction is unlikely to yield the attested forms (Zair 2012: 83) and the Celtic words may more attractively be compared with Lat. *corium* 'leather' (de Bernardo Stempel 1987: 93; compare Gr. σκῡτεύς 'cobbler' < σκῦτος 'leather').

62 It almost seems a sheer accident that Schrader (1901: 33 and Schrader/Nehring I: 38) included here OCo. *kelin* 'holly'. He left the word without a gloss, and in the same unglossed form it was repeated by Trautmann (1923: 136), Berneker (I: 512, adding Welsh *celyn*) and Vasmer (REW I: 567); cf. Friedrich (1970: 64), where Welsh *celyn* is incorrectly glossed 'maple'. The etymology, with the correct gloss, is explicitly defended by Specht (1947: 60) and Fraenkel (LEW 270–271), while it is explicitly — and surely correctly — rejected by Stang (1972: 29; implicitly e.g. IEW 603).

63 Although *$kluna$- (trad. *$klъnъ$; Miklosich 1886: 118) or *$kulna$- (trad. *$kъlnъ$) would also be possible.

64 Alongside a variant γλεῖνος, of unspecified dialectal affiliation (cf. Meyer 1892: 325–326).

Germanic **hlun-* could theoretically be analysed as a secondary zero-grade to **hlen-*, thereby matching the Slavic form; however, since such a full-grade variant is unattested, this hypothesis is rather circular. It seems quite possible that the irregular correspondence could be explained as the result of borrowings from a non-Indo-European source. A similar explanation might also account for the "Macedonian" κλινό-, mentioned above, although too little is known about Ancient Macedonian for this form to be used here.

▸ [**'silver'**. OCS сребро, Cz. *stříbro*, Sln. *srebrọ̑* 'silver' ~ Go. *silubr*, ON *silfr*, OHG *silabar** 'silver' — See the discussion on pp. 225–226.]

▸ **? 'frogspawn'**. Lt. *kurkulaĩ*, Lv. *kuŕkuļi* (LVPPV: *kùrkulis*); Pl. *skrzek*, Cz. dial. (Kott V: 752) *žabo-křeky* 'frogspawn' ~ ON *hrogn*, OHG *rogo* '(fish) roe' (Polomé 1986: 661) — Germanic **kruk-* and Baltic **kurk-* seem to show an irregular metathetic relationship. The analysis of the Balto-Slavic data is difficult, however, due to repeated contaminations with words for 'croak' (cf. Machek 1924: 128–130). First and foremost, the Baltic forms look like derivatives of Lt. *kuřkti*, Lv. *kùrkt* 'to croak' (Būga 1923–1924: 139; cf. Nesselmann 1851: 212). In Slavic, compare R dial. *крякъ* 'frogspawn' (hardly with *-я-* < **-ę̄-*, *pace* REW I: 674) beside *кря́кать* 'croak' (СРНГ XV: 365–366), and further Pl. *skrzek* 'croaking; frogspawn'.

Since 'frogspawn' as 'the croaker' does not make much sense, I assume these contaminations are secondary. In this case, one is tempted to give preference to the Lithuanian **krekulai* (Miežinis 1894: 118),[65] dial. *krekùčiai* (LKŽ) 'frogspawn', which are closer to the Slavic forms. This could perhaps support the connection with Lt. *krèkti* 'coagulate, congeal' (Berneker I: 613–614; LEW 293) and suggest that the Germanic evidence is unrelated. On the other hand, the parallelism with **klen-* : **klun-* 'maple', discussed above, opens up the possibility of analysing **krek-* : **kruk-* as a non-IE borrowing.

7.3.3 ? *Lithuanian* *ā ∞ *Latvian* *ū

▸ **? 'millet'**. Lt. *sóros* ~ Lv. dial. (Kurzeme) *sûra*[2] 'millet' (Pronk/Pronk-Tiethoff 2018: 293; Kroonen et al. 2022: 22) — The inner-Baltic correspondence is irregular, but Elger's 1683 dictionary has Lv. ⟨sâre⟩, ⟨sare⟩ = */sāre/ 'milium' (see ME III: 806; Nieminen 1956: 164–165). This might suggest the Latvian variant with *-ū-* is secondary, although only *ad hoc* accounts can be given for it (cf. Nieminen 1956: 175–176). The best explanation is to assume a contamination with South Estonian *suurmaq*, Li. *sūrmõd* PL. 'groats', which is supported by

65 Attested as ⟨krakulai⟩, but its alphabetical position after ⟨krekinties⟩ implies a misprint.

the Salaca Livonian compound *kriev sūrmed* 'millet' (*kriev* 'Russian'; Winkler/Pajusalu 2009: 87, 182). The existing etymological explanations of the Baltic word are all unsatisfactory. A derivation from an older **psārā* and comparison with Slavic **prasa-* (trad. **proso*) 'millet' (Hirt 1927: 309; Otrębski 1939: 137), or Skt. *psāti* 'chew, consume' (Nieminen 1956: 170) remains highly hypothetical, especially in view of the contradictory development of initial **ps-* observed in Lt. *spenỹs* 'teat' < **psten-*.[66]

The Baltic word is obviously related to Md. E *suro*, M *sura* 'millet'.[67] Although some have derived the Mordvin word from Baltic (Thomsen 1890: 219; SSA III: 201; Kallio 2008a: 268), the opposite direction has often been preferred (Ojansuu 1921: 57–60; Kalima 1936: 210; Топоров/Трубачев 1962: 248). This was supported by the claim that the Mordvin lexeme is cognate with F *sora*, E dial. *sõra* 'gravel, coarse sand'. As an alternative, Toivonen (1928: 233) has adduced Komi *zẹ̑r*, (Jaźva) *zu̇·r* 'oats', Udmurt *zẹr* 'bromegrass', an equation which was taken up by UEW (766) and Лыткин/Гуляев (1970: 106, with hesitation). Should the Permic words belong here, the semantic shift could be explained as a result of a migration beyond the northern limits of millet cultivation (about 57° N in the Eastern Baltic, cf. Grikpėdis/Motuzaitė Matuzevičiūtė 2020: 163). If we admit a secondary voicing in Permic,[68] then we could suggest a Uralic reconstruction **sorə*. Such a reconstruction might also work for Mordvin, although admittedly, due to the PMd. final **-ə*, Mordvin is more consistent with a reconstruction **sora* rather than **sorə* (Pystynen 2020b).

Van Pareren (2008: 124; cf. Junttila 2015a: 23) is inclined to reject any relationship between the Baltic and Mordvin words because Baltic **-ā-* is difficult to square with Mordvin **-u-*. According to Nieminen (1956: 173), the comparison would only be possible should we assume an original **ō* in pre-Mordvin; however, the traditional reconstruction of Proto-Uralic **ō* has since

66 Witczak's (1997: 30–32) comparison with MP *xw'r* /xwār/ 'food', Oss. I *xor*, D *xwar* 'cereal, barley', allegedly from **sueh$_2$r-* is impossible. Not only is the loss of **u* in Baltic suspect, but the Iranian root is clearly **ĥwăr-* < **sueR-* (Cheung 2007: 147–148), i.e. not consistent with a laryngeal. See Kroonen et al. 2022: 22.

67 I cannot accept the suggestion of Nieminen (1956: 175) that the similarity is coincidental.

68 Sporadic initial voicing is attested in Permic, compare e.g. Komi *bẹ̑ž*, Udm. *bi̯ž* < **pončə* 'tail' (Sammallahti 1988: 547), but there remains a question as to whether this also applies to sibilants. The few possible equations, e.g. Komi *zi̯l'* 'diligent, hard-working' ~ Hungarian *ügyes* 'skilled, capable' (UEW 442–443) and Komi *za* 'stem, stalk, shaft' ~ Erzya *sad* 'stalk' (see Chapter 4, fn. 56), all involve some additional phonological obstacles. For **o–ə* > Komi/Udmurt *ẹ*, cf. Komi-Permjak *ẹs-*, Udmurt *ẹski̯-* 'vomit' < **oskə-* (cf. Metsäranta 2020: 103).

been shown to be faulty: most examples of **ō* should instead be reconstructed **a*(–*ə*) (see Aikio 2012b). If we set up an earlier **sarə* for Mordvin, this would bring it closer to the Baltic data, but force us to separate the Permic evidence; moreover, the Mordvin change **a*(–*ə*) > (**o* >) **u* has been dated very early and considered to be part of a chain of vowel shifts shared by Sámi (Живлов 2014: 116–117), which would make the reconstruction of a pre-Mordvin **sarə* anachronistic.

In conclusion, it is unclear whether such a hypothesis would shed any light on the Latvian *-ū-*, as (especially in view of the evidence from Elger's dictionary) the time depth is probably too shallow to assume influence of a substrate. However, a direct loan relationship between Mordvin and Baltic words for 'millet' remains doubtful, and it is quite possible that these represent parallel loanwords from a third source. Since the correlation between Baltic **ā* and Mordvin **u* could be characterized as one of height, this example may still belong here.

7.3.3.1 Conclusion

The certain and possible evidence high ∞ low vocalic alternations is collected in Table 16, overleaf (see p. 185 for help reading the table). Forms which do not provide relevant data are presented in light grey. Shaded cells indicate forms containing high vowels.

The evidence falls into two broad groups: those which show a high vowel in Balto-Slavic, and those which show a high vowel in Germanic. The former group consists mainly of more or less uncertain examples, and since there is little coherence in terms of distribution and co-occurring phenomena, it is quite unlikely that all of the examples can be directly compared.

Those which show a high vowel in Germanic form a much more promising group — all three show **e* (= **[æ]*) in Slavic and **u* in Germanic. Since the vowels **œ* and **u* could hardly be more different, representing opposite extremes of the vowel triangle, it may seem dubious to derive them from a common source. However, it is not too difficult to find such correlations between related languages, such as between e.g. Erzya Mordvin *kenže* (dial. *käнžä*) and Obdorsk Khanty *kuns* 'nail' (< **künčі*), while the regular development **e*(–*ä*) > Permic **ǫ* leads to examples such as Finnish *pesä* 'nest' ~ Udmurt *puz* 'egg'. Thus, the problem may be resolved by assuming distinct but related donor languages.

TABLE 16 Possible examples of high ∞ low alternations

	Baltic	Slavic	Germanic	Elsewhere
'bison'	*u̯i-sam(bʰ)r- *stum(bʰ)r-	*za/um(bʰ)r-	*u̯i-s(u)nT-	
'oriole'	*u̯ālanG-	*u̯(i)lg(ʷ)-	*u̯alk-	
? 'heel'	*kulḱ-ni-	? *kulk-	–	Lat. *kalk-
? 'salmon'	*laśiś-	*laSaS-	*laks-	
? 'grouse'	*i̯e-rubʰ-	*(i̯)e-re/imbʰ-	*rebʰ-	
'maple'	–	*klen-	*klun-	? Mac. *klin-
'silver'	*sidʰabʰr-	*sirebr-	*silubr-	Celt. ?*silapr-
? 'frogspawn'	? *krek-	*krēk-	*kruk-	
? 'millet'	*sār-	–	–	Md. *sur-

7.4 Alternations between Monophthongs and Diphthongs

7.4.1 ? *Baltic* *a ∞ *Slavic* *ai/ei

Schrijver (1997: 304–307) has adduced several examples in which Germanic **ai* appears to correspond to Celtic **a*, the clearest being MW *baed* (< **basio-*) against OE *bār* (< **baiza-*) 'boar'.[69] None of his examples involve Balto-Slavic, but a few cases can be identified in which Slavic potentially contains a diphthong. Although it is admittedly not possible to rule out an original long vowel based on the Slavic-internal evidence, in each case, a diphthong would be more easily reconcilable with the Baltic evidence.

69 Note, however, that van Sluis et al. (2023: 231) consider the Celtic word a possible West Germanic loan.

▸ **'leaf'.** OCS листъ, R *лист*, Cz. *list* 'leaf' ~ Lt. dial. *lãkštas* 'large flat leaf; leaf used as a baking sheet', dialectically and in older texts 'leaf' (see ALEW² s.v.), Lv. *laksti* PL. 'leaves of herbs; vegetable tops' — The Slavic word is usually compared with a different Baltic lemma, namely Lt. *láiškas* 'narrow leaf (e.g. of an onion), blade; green shoot', as a neologism '(postal) letter', Lv. *laiska* (hapax? cf. ME II: 410–411) 'leaf on a flax or cereal stalk'. According to this theory, the Slavic word would show a different ablaut grade and a derivation in *-*to*-. If we assume a non-inherited origin, the comparison with Lt. *lãkštas* appears more straightforward; moreover, this word is semantically a better fit, as it is well attested in the more general sense 'leaf'.

It seems attractive to compare both the Baltic and Slavic words with a group of West Uralic words for 'leaf': Sá. N *lasta*, Sk. *lõstt* (< **lęstę*); F *lehti*, E *leht*, Li. *lē'ḑ* (< **lehti*), Ma. E *ləštaš*, W *ləštäš* (< **lištäš*), all of which can regularly reflect PU **leštə* (UEW 689). Incidentally, the similarity to both the Baltic and Slavic words has already been noted: Sammallahti (1977: 123–124; cf. SSA II: 58–59) has assumed a direct loan from the Baltic **la*(*k*)*šta*-, while Viitso (1992: 189) and Напольских/Энговатова (2000: 229) have posited an early loan from Slavic. Finally, Blažek (2019: 216) has suggested a loan from an unattested Baltic **lišt*-, a zero-grade equivalent of Slavic **līsta*- (trad. **listъ*).

M. Живлов (p.c.) has noted that the similarity between these words might be better accounted for by assuming parallel loans from a substrate language. As the above solutions provide a convincing account of only part of the data, this is certainly worthy of consideration. A similar vocalic relationship is found between R *muc*, Cz. *tis*, Sln. (Pleteršnik II: 670) *tîs*, (SSKJ²) *tîsa* 'yew tree'[70] and Latin *taxus* 'yew tree', which has widely been regarded a non-IE borrowing (Schrader in Hehn 1911: 532; Schrader/Nehring I: 225; Oštir 1930: 22, 90; Machek 1950: 152; REW III: 107; Sławski SEJP I: 103),[71] a suggestion which, in principle, seems attractive: the Latin and Slavic words are semantically identical and

70 RCS тисъ translates Gr. κέδρος 'cedar' (cf. СДРЯ 960; СРЯ 11–17 XXIX: 350; cf. also OCS (Ps. Sin.) тіса glossing the Greek loanword кедрі, SJS IV: 457). Most likely, this is merely a localization of a Mediterranean dendronym and does not attest to a genuine currency of the word in this sense (*pace* Blažek/Janyšková 2015: 91). In any case, all of the modern languages are in agreement in meaning 'yew'.

71 Blažek/Janyšková (2015: 87) have suggested that the Slavic word may have been loaned from a Dalmatian **tis*, which would be the regular reflex of Lat. *taxeus* 'made of yew'. While an interesting suggestion, it is hampered, as the authors note, by the fact that this Dalmatian word is hypothetical (and the adjective *taxeus* does not appear to be continued in other Romance languages). More importantly, in a common Slavic loanword, one would undoubtedly anticipate a substitution **ĭ* → **ĭ* (trad. **ь*), as in e.g. Cz. *mše*, SCr. dial. *mȁša* 'mass' (← ML *missa*, M. Matasović 2011: 114–115).

formally similar (Slavic **s* may derive from **ḱs*, or perhaps from foreign **ks* in a borrowing postdating the RUKI law).

If Slavic **ī* (trad. **i*) in these cases derives from an earlier **ei*, then we are dealing with an underlying alternation **a* ∞ **ei*. At face value, this could be understood as a combination of a 'diphthongal' alternation (like the one described by Schrijver) and a front–back alternation (as in 7.2). Naturally, it is possible that both **a* and **ei* derive directly from some other source like **ai*; as usual, any explanation of this alternation will remain in the domain of speculation.

▸ † '**hazel** (1)'. Lt. dial. *lazdà* 'hazel', usually 'cane, stick', Lv. *lazda*; Pr E *laxde* 'hazel' ~ Uk. dial. *лíскa* (usu. *лищúна*), Cz. *líska*, SCr. *lijèska* 'hazel' — The comparison with Alb. *lajthi* 'hazel' (Meyer 1891: 234; Jokl 1923: 203–206; Huld 1990: 401) is suspect. As proven by the form *lakthi* in Dalmatian Albanian, an original cluster **-kϑ-* is to be reconstructed (see Demiraj 1997: 231–232). As **-kϑ-* is not the known reflex of any inherited cluster, *-th-* must be a suffix, while the stem **lak-* cannot easily be compared with the Balto-Slavic data.[72]

The Baltic and Slavic forms were already compared by Miklosich (1886: 167), but the comparison is generally viewed with scepticism (cf. REW II: 34; ALEW 652). Derksen (2008: 274), however, states that "there can hardly be any doubt" that the comparison is correct. I am not entirely convinced: even if we are willing to accept an alternation **a* ∞ *ai*, for which a partial parallel could be the word for 'turnip' (cf. Lt. *rópė* ~ R *рѣ́па*, p. 237), we are still left with the obscure relationship between Baltic **-zd-* and Slavic **-sk-*. In principle, given the lack of old attestations, the Slavic forms do not appear to exclude a reconstruction **lēs-ukā-* (trad. **lěsъka*), i.e. a derivative of the noun in OCS лѣсъ 'copse, thicket' (REW loc. cit.; cf. ЭССЯ XIV: 241 with lit.). Considering the vocalism and consonantism are both irregular, it is uncertain, despite Derksen, whether the similarity is sufficient to warrant a comparison at all.

7.4.2 ? *Baltic* *ā/ō ∞ *au *Elsewhere*

▸ '**seal**'. Lt. *rúonis*, Lv. *ruônis* 'seal' ~ OIr. *rón*, MW *moel-rawn*, Bret. *reunig* 'seal' (Ariste 1971: 10; Wagner 1981: 26; Sausverde 1996: 139; Stifter 2023, forthc.) — Considering that the similarity between the Baltic and the Celtic words seems obvious (Būga 1911: 37, 1922: 279), it is remarkable that the Celtic data is not even mentioned in most Baltic etymological dictionaries (LEW 746–747; Karulis II:

72 The direct equation of Lt. *lazdà* with OIr. *slat* 'rod, twig' (Kroonen 2011b: 217–218; ALEW 652) is not possible, as the Celtic form must be reconstructed **slattā* (Schrijver 1995a: 431; Matasović 2009: 345), cf. Modern Irish *slat* as against *nead* 'nest' (< **nisdo-*). The connection with OCS лоза 'vine' (Berneker I: 736; REW II: 43–44) is phonologically and semantically implausible.

129; Smoczyński 2018: 1115), which instead offer speculative root etymologies. The reason for this omission is that the Irish form was long ago explained as a Germanic loanword (Bezzenberger in Stokes 1894: 235; Pedersen 1909: 21; ME III: 581). However, the suggested source, Old English *hran* '(a small kind of) whale', is not phonologically suitable. The variant spelling *hron* shows the rounding of short /a/ before a nasal (Hogg 1992: 14), and demonstrates that we are dealing with a short vowel (see also the considerations of Stifter 2023: 183).

The Celtic reconstruction is problematic. Stifter (2023: 183) has stated that the British and Goidelic forms cannot be united under a common proto-form, but he has later (Stifter forthc.) suggested the reconstruction **rauno-*, comparing the homonym MoIr. *rón* 'horse-hair', MW *rawn* 'coarse animal hair' < **rauno-* (cf. R *руно́* 'fleece').[73] If this reconstruction is valid, then the only way to compare the Baltic and Celtic forms would be to reconstruct **reh$_2$u-no-* for Celtic and **roh$_2$u-n-* for Baltic. However, even then, the development of **-oh$_2$u-* > Baltic **-ô-* is highly suspect (Villanueva Svensson 2015).

Since it is *a priori* questionable that Proto-Indo-European could have had a word for 'seal', such manipulations feel superfluous. Instead, it is more probable that Celtic and Baltic loaned their respective words for 'seal' from related sources. Considering that seals are marine animals, this must have occurred relatively late in both branches.

▸ **? 'palate'.** Lt. *gomurỹs* 3ª 'palate', Lv. *gãmurs* 'larynx; windpipe' ~ OHG *goumo* · facia 'gullet; throat', MoHG *Gaumen* 'palate' (Derksen 2015: 184) — Almost all of the Germanic evidence points to **gōma(n)-*, cf. ON *gómr* 'palate; floor of the mouth', OE *gōma* 'palate; gullet', OHG *guomo* 'palate; throat' (cf. Kroonen 2011b: 302), which would harmonize nicely with the Baltic data, allowing for the reconstruction of a shared proto-form **g^{h}eh$_2$m-*.[74] However, this fails to account for the High German evidence, for which various solutions have been proposed.

In OHG, we find *guomo* beside *goumo*. The alternation *-ou- ~ -uo-* has been attributed to various reductions of a long diphthong **āu* (Winter 1982: 183; Kluge/Seebold 336). A possible condition for this variation was suggested by Kroonen (2013: 185): in his opinion, the pre-Proto-Germanic diphthong **ōu*

73 The *a*-vocalism raises problems for Slavic here, though, since the clear evidence for oxytone accentuation (cf. also Cz. *rouno*, SCr. (Vuk) *rúno* 'fleece'; Зализняк 1985: 135; Derksen 2008: 440) is hardly consistent with an internal laryngeal. See the detailed discussion in Stifter forthc.

74 The received connection with Gr. χάσκω 'yawn, gape' (IEW 449) appears to be contradicted by the initial *g-* in Baltic as opposed to the *ž-* in Lt. *žiótis* 'open one's mouth'. On the relationship of the Greek and Baltic forms, see Lubotsky 2011.

developed to *$\bar{o}$ in open syllables, while being shortened to **au* in closed syllables (cf. also idem: xv–xvi).[75] He therefore suggests that an earlier paradigm **gōumōn* (> **gōm-*), OBL. **gōumn-* (> **gaum-*) could account for both OHG variants. However, it should be noted that OHG spellings in *goum-*, *gaum-* are rare,[76] and a MHG continuation is uncertain (see MWb s.v. *guome*).

MoHG *Gaumen*, which only becomes common in the 16th century (see DWb IV: 1576–1578), is usually assumed to continue the OHG by-form *goumo*. However, a number of dialect forms appear to suggest a prototype **gūman-*, cf. Swiss (16th c.) *gūme* (*Schw. Id.* II: 308), Cimbrian *gaumo* (cf. Schmeller/Bergmann 1855: 39–40), Upper Saxon *gaumen* (DWb IV: 1577 under 3b), Prussian German *gūmə* (PrWb II: 261), as well as MLG *gume* (Schiller/Lübben I: 165) and Lower Saxon *gūmen* (NdsWb II: 135). The standard German form may in principle be derived from this preform, too. In that case, we might instead assume an old ablauting **gōman-* : **gūman-*, somewhat comparable to that observed in Go. *fon*, GEN.SG. *funins* 'fire', reflecting an earlier **peh$_2$ur*, OBL. **puh$_2$n-* (< **ph$_2$u-n-*) (for more potential examples, see Kroonen 2011b: 319–324).

Either interpretation of the Germanic evidence appears to require a root containing **u*, a reconstruction which is excluded by the Baltic evidence. This word may therefore possibly show evidence for an alternation **ā* ∞ **āu*. Note, however, that the interpretation of the Germanic ablaut alternation in laryngealistic terms may be anachronistic if we are actually dealing with a post-PIE loanword. Alternatively, we might interpret the continental Germanic evidence for **ū* as indicative of a non-IE alternation **ā* ∞ **ū*. Whatever the solution, the Baltic and Germanic forms are difficult to combine in an Indo-European context. Due to ambiguities in interpreting the Germanic evidence, this cannot be considered a certain example of a diphthong alternation.

7.5 Length Alternations

7.5.1 *Baltic Long ∞ Elsewhere Short*

▸ **'apple'.** Lt. *obuolỹs* 3^{a}, Lv. *âbuõls* 'apple' (beside Lt. *obelìs*, Lv. *âbele* 'apple tree', an old consonant stem); Pr. E *woble*; R *я́блоко*, Cz. *jablko*, Sln. *jábołkǫ* 'apple' ~ OE *æppel*, OHG *apful*; OIr. *ubull*, MW *aval* 'apple' (Kluge/Götze [1948]: 20–

75 Compare ON *nór* 'ship (as a kenning); tempering trough (= Ic. *nó-trog*)' (< **neh$_2$u-*; idem: 391) : *naust* 'boathouse' (< **neh$_2$u-st-*; idem: 384).

76 3× in AWb as against dozens in *guom-, guam-*; the interpretation of the hapax spelling *gaom-* is disputed, see EWAhd (IV: 562 with lit.), whose authors assume that *goumo* arose due to contamination with *goumen* 'eat, feast' (cf. also ALEW 400).

21; Hamp 1979: 163–166; Markey 1989: 599–600; Huld 1990: 398–400; Oettinger 2003; Kroonen 2013: 31) — The word for 'apple' occupies a curious position in the study of Indo-European. On the one hand, the word has often figured as a key example in the question of Indo-European *l*-stems (Fraenkel 1936b: 172–176; Adams 1985; Olsen 2010: 76; Beekes 2011: 195; Stifter 2019: 204–207); on the other, it has not infrequently been regarded as a probable loanword from a non-IE source.[77]

The evidence for an ablauting *l*-stem comes from the following: Lt. *obuolỹs* and OIr. *ubull* < **abūl-* (cf. in detail Stifter 2019) point to lengthened grade suffix **-ōl-*, while Germanic **apla-* and Slavic **ābl-uka-* (trad. **ablъko/ъ*) suggest a zero-grade **-l-*. Words for 'apple tree' tend to show full-grade: Pr. E *wobalne*; MR *яболонь* (Шахматов 1915: 151), Pl. *jabłoń*, Sln. *jáblana* 'apple tree' and OIr. *aball*, MW *avall* 'apple tree' suggest **-al-n-* (*a*-grade? cf. 7.6; or rather Slavic **-ol-n-*, Celtic **-l-n-*); ON *apaldr* also suggests **apal-*(*d*)*ra-* as against West Germanic **apla-*. Lt. *obelìs* points to **-el-* which could potentially be secondary for **-ol-*, as Lt. *sẽser-* 'sister OBL.' « **suesor-* (cf. Skt. *svásāram* ACC.SG.).

Aside from the limited distribution, the argument for a non-IE origin essentially comes down to the presence of the phoneme **b*. If one does not accept the existence of **b* in PIE, the word must be interpreted as a borrowing; by contrast, if one does accept such a phoneme, the word is unproblematic (cf. NIL 264). Hamp (1979: 163) initially speculated that the long vowel in Balto-Slavic might be a reflection of some non-IE feature, but retracted this view in an addendum, preferring to evoke the recently discovered Winter's law (cf. Winter 1978: 438). This has become the *communis opinio* (NIL 263), as the supposition of **b* in principle explains both the Germanic and Balto-Slavic data.

However, I do not consider there to be sufficient evidence for a phoneme **b* in PIE (see Pedersen 1951: 10–12; Гамкрелидзе/Иванов 1984: 6–7; Lubotsky 2013; see also Olander 2020). Aside from Lt. *trobà* 'peasant house', where the vocalism is unexpected (see the following entry), all of the other examples of Winter's law from **b* are ambiguous at best. Thus, Winter's only other example was Lt. *grébti* 'rake, gather up; snatch', where the secondary nature of the acute is shown by Lv. *grebt* 'carve, hollow out', R *грести́* 'rake up, gather together',[78]

77 I will not get into the attempts to connect the families of Lat. *mālum* and Pashto *maṇá* (both 'apple') to this word, except to say that the supposed irregular change **b* > **m* (Blažek 1995: 17) or, conversely, **m* > **b* (Гамкрелидзе/Иванов 1984: 639–640; Cheung/Aydemir 2015: 85–86) are both completely *ad hoc* (cf. Kroonen 2016: 88).

78 Kortlandt (1988: 393; followed by Derksen 1996: 321–322) has assumed the confusion of two roots, **gʰrebʰ-* 'to dig' and **gʰreb-* 'to grab', but the latter is based only on Lt. *grébti* (LIV 201; OCS грабити 'steal, snatch' belongs rather with Lt. *gróbti*; on the Germanic forms, see

and other cases do not inspire confidence, either.[79] If we reject the phoneme **b*, as I would recommend, then we should interpret the correspondences as irregular. The long, acute vowel in Balto-Slavic as opposed to the short vowel elsewhere can be viewed in this context (see in particular the following example).

On the other hand, it cannot entirely be excluded that the word was borrowed into Balto-Slavic prior to Winter's law. Note in this respect the fact that the word can be reconstructed to Proto-Balto-Slavic, and appears to show archaic ablaut, which would favour an early adoption. In this case, this alternation could not be characterized as one of length, and the only challenge to an Indo-European origin would be the necessity to reconstruct **b*.

▸ '**cottage**'. Lt. *trobà* 'peasant house; room', HLv. *tràba*^{2} 'old, worn-out building; improvised hut' (ME IV: 227; EH II: 692); Oscan trííbúm ACC.SG. 'house'[80] ~ OIr. *treb*, MW *tref* 'residence, estate'; further cf. Go. *þaurp* 'field', OE *þrop*, *þorp*, OHG *dorf* 'hamlet, estate' (Hamp 1978: 187; Huld 1990: 398; de Vaan 2008: 626) — There is no necessity in including Lat. *trabs* 'beam, tree' in this etymology (see Ernout/Meillet 698; Untermann 2000: 766 *contra* Walde/Hofmann II: 696–697; de Vaan 2008: 626, etc.). As with the word for 'apple', discussed immediately above, the analysis of this word is intrinsically linked to the status of the phoneme **b*. In this case, both Oscan and Germanic provide independent evidence in favour of this reconstruction, and the long vowel in Baltic could be attributed to Winter's law (Derksen 2015: 472).

However, the Baltic **ā*-vocalism also presents problems, as from an original **o* lengthened by Winter's law, we would anticipate **ô*.[81] ALEW (p. 1298) makes reference to "Neoablaut", but this is difficult, as the word does not show any evidence of ablaut within Baltic (leaving aside the doubtful *vien-trēb* 'alone' beside *-treĩb*, cf. ME IV: 667). In a similar context, Derksen has referred to "the well-known East Baltic reshuffling of the ablaut relations" (2002: 9); however, to justify this position, we would need more concrete argumentation. As it stands,

---

Kroonen 2013: 187). For the secondary acute, as well as the variant *grẽbti* (cf. 3PRES. *grẽbia* in Alytus), compare also Lt. *répti* (dial. *rẽpti*) 'take, rob' beside *rẽplės* 'tongs', Alb. *rjep*, *rrjep* 'skin, flay' (cf. LIV 507, where the secondary nature of the acute is taken for granted; likewise ALEW 991; for further discussion of this kind of metatony, see Pronk 2012: 29–32).

79 For Lt. *drẽbti* (3PRES. *drẽbia*), Derksen (2015: 138) is again content to assume contamination of two verbs, but the euphemistic sense 'strike' is hardly to be separated from senses such as 'pour (e.g. porridge); make from clay; slouch'. On OCS слабъ 'weak', which is reconstructed **sleb-* (LIV *addenda* s.v. **(s)leb-*), see Kroonen (2011a: 258–259).

80 The Umbrian hapax *trebeit* 3SG.PRES. 'lingers, dwells(?)' is normally ascribed a short **ĕ* (Buck 1904: 62; Untermann 2000: 759 with lit.). But as far as I can make out, the length is ambiguous. Buck (op. cit. 26) writes "oftenest there is no designation of the length" and on the spelling of **ē* in Umbrian (p. 34), "*i* occurs frequently [...] but *e* is far more common".

81 Lt. *núogas* ~ Skt. *nagná-* 'naked'; *púodas* 'pot' ~ OE *fæt* 'vessel, jar', etc. (Winter 1978: 345).

it would appear that **a* (see 7.6) and **b*, both of which are of doubtful status, would have to be assumed if we are to attribute the length and acute accent to Winter's law. On the other hand, Oscan also shows a long vowel in this word, suggesting that it may have some other origin.

Either way, the relationship between Baltic **-ă̄-*, Oscan **-ē-* and Celtic **-e-* is impossible to adequately explain in Indo-European terms (Beekes 1969: 191), independently of whether one assumes an originally long vowel in Baltic or one that was secondarily lengthened by Winter's law. For Baltic **a* against **e* elsewhere, see 7.2.2. On the other hand, the Germanic evidence is rather troubling, as at face value it seems to imply a syllabic **ŗ*. This might be more consistent with an inherited origin (compare 'furrow', p. 208). It might be possible to view the Germanic **u* vocalism in the context of the **e* ∞ **u* alternation discussed under 7.3.2, but this would require the additional assumption of the metathesis of **r* (compare similarly 'sturgeon', pp. 236–237).

▸ **'fresh'**. Lt. *préskas*; R *прѣсный*, Sln. *prę́sən* 'fresh, unleavened' ~ OE *fersc* 'fresh, unsalted', OHG *frisc* 'raw, fresh' — It was previously assumed that Balto-Slavic reflected **proisk-* and Germanic **prisk-* (still Torp 1919: 135; Walde/Pokorny II: 89). However, since Būga (1922: 277) demonstrated that the supposed Lithuanian variant **príeskas* does not exist, the etymology has largely been rejected, with the Germanic forms usually not even mentioned as possible comparanda (cf. Trautmann 1923: 231; LEW 652; REW II: 429–430; Smoczyński 2018: 1018).[82] It nevertheless seems difficult to imagine that the similarity is a mere coincidence in view of the precise agreement in meaning and correspondence of four consonants. The comparison can only be made by assuming parallel loanwords from another source. See also the discussion of Finnic **rēska* 'fresh', whose vocalism may also pose issues, in Chapter 3, fn. 159.

▸ [**'lynx'**. Lt. *lū́šis*, Lv. *lũsis*, Pr. E *luysis*; R *рысь*, Sln. *rîs* 'lynx' ~ OE *lox*, OHG *luhs* 'lynx' — See the detailed discussion on pp. 180–181.]

▸ **? 'ash'**. Lt. *úosis*, Lv. *uôsis*; Pr. E *woasis*; R *я́сень*, Slk. *jaseň*, SCr. *jȁsēn* 'ash tree' ~ Lat. *ornus*; OIr. *uinnius*, MW *onn* (COLL.) 'ash tree' (Machek 1954: 108, 1968: 217) — Bg. dial. *òсен* need not imply a variant with **ŏ-* in Balto-Slavic (*pace* БЕР IV: 936; Andersen 1996a: 142–143). Its distribution largely corresponds to that of *осѝка* 'aspen' (central Bulgaria, east of Sofia), which it was apparently influ-

82 The word is also omitted from Stang's treatment of the Balto-Slavic–Germanic isoglosses. He does (1972: 40), however, adduce a pair which is remarkably similar, both belonging to a similar semantic field and showing the same correlation in vocalism: Pl. *obrzazg*, *obrzask* 'tart flavour (of wine)', R *брѣзгать* 'be squeamish, fussy' ~ Nw. dial. *brisk* 'sharp or bitter taste'. Note, however, that an ablauting **bʰroisg-* : **bʰrisg-* would indeed be possible here.

enced by. The latter is a preserved archaism (cf. Cz. *osika* 'aspen'), while dial. *ясùка* 'aspen', on the contrary, shows the influence of *ѧ̀сен* 'ash' (thus already Zubatý 1892: 254 fn.).[83] I prefer to keep ON *askr*, Arm. *hacʻi* 'ash tree' Alb. *ah* 'beech' (< *h_2esk-o-*) apart.

Kortlandt (1988: 391) has suggested to start from a paradigm *Heh_3-s- : *Hh_3-es-*, a solution which has been followed by a number of Leiden-affiliated scholars (e.g. Schrijver 1991: 78; de Vaan 2008: 435; Kroonen 2013: 38). While such a paradigm seems possible on paper, it is difficult to imagine its survival into core PIE in a peripheral, non-basic vocabulary item.[84] Furthermore, this is not the only possibility. It would be equally possible, both phonologically and morphologically, to assume a reduplicated *h_3e-h_3s- in Balto-Slavic as against *h_3e/os- elsewhere. Compare Lt. *néndrė* against Hitt. *nāta-* 'reed' (see p. 240, where other possible parallels are adduced).

In this context, it is important to consider the similar correlation in the word for 'elbow'. A long acute vowel is found in Lt. *úolektis*, Lv. *uôlekts* 'ell (measure of length)', and possibly Pr. E *woaltis* 'ell, forearm' matching Gr. ὠλένη 'forearm', while Lt. *alkū́nė*, Pr. E *alkunis*, OCS лакъть, R *ло́коть* 'elbow' are not consistent with a laryngeal, and match Gr. ὀλέκρανον 'point of the elbow'.[85] Here, the reconstruction *Heh_3-l- : *Hh_3-el-* (Kortlandt loc. cit.; Lubotsky 1990: 131–132) is even more uncomfortable, as it would have to have survived into the respective prehistories of Greek and Balto-Slavic, while *Heh_3l- : *Hh_3l-* (Kroonen 2013: 22) is unlikely to work for the Balto-Slavic data.

Thus, if we were to explain the long vowel of Lt. *úosis* (etc.) as the result of borrowing from a non-IE source, consistency would demand we use the same explanation for 'elbow'. Yet considering that 'elbow' has plausible cognates at least in Indo-Iranian (Skt. *aratní-* 'elbow, ell') and more or less basic semantics,

---

83 Thus also SCr. *jàsika* beside *jȁsēn*, and Sln. *jasíka* and *jásen* reported in the same village by Erjavec (1883: 290). A variant with *e- perhaps underlies Sln. *jesíka*, whence *jésen* (and Kajkavian *jȅsēn*, cf. PCA VIII: 741). Or does this variation have something to do with the frequent occurrence of *ja-* for *je-* in South and West Slavic (Andersen 1996a: 74–76)? On the association of 'ash' and 'aspen', see Normier 1981: 25–26 with lit.

84 Note in this respect that the very similar word for 'mouth' *h_1eh_3-s, OBL. *h_1h_3-s-* still preserved its archaic paradigm in Hitt. *ais* (for *$ās$, cf. CLuw. *āssa*), obl. *iss-* (thus Kloekhorst 2008: 166–167), but was levelled in the rest of PIE, where it was probably reinterpreted as a root noun: Lat. *ōs* (GEN.SG. *ōris*), Skt. *ā́s-*, OIr. (rare; cf. LEIA A-4) *á* 'mouth'. (For alternative views on the Hittite form, see Melchert 2010 and NIL 388–389 with lit.).

85 The short vowel is assured in all early attestations, while ὠλέκρανον is a corruption of later editors (Isépy/Primavesi 2014: 126–127). Lat. *ulna* 'ell, elbow', OIr. *uilen* 'elbow' and Go. *aleina* 'ell' reflect a short vowel, but are ambiguous as this may be the result of pretonic shortening (В. Дыбо 1961: 13, 25, 2008: 561; cf. Schrijver 1991: 352; Kroonen 2013: 22).

a non-IE etymology is not attractive. If the only thing separating *úosis* from *úolektis* is its semantics and geography, then its non-IE origin cannot be considered certain.

▸ [? **'nettle'**. Lt. *notrẽ*, Lv. *nâtre*; Pr. E *noatis* 'nettle' ~ OSw. *nätla*, *nätsla*, OE *netele*, OHG *nezzila* 'nettle' — See the detailed discussion on p. 203.]

▸ † **'harrow'**. Lt. *akéčios*, Lv. *ecêšas*, Pr. E *aketes* PL. 'harrow' ~ OE *egeþe*, OHG *egida*; MW *oget* 'harrow' (Oštir 1930: 15) — The alleged connection with Bel. dial. (cf. ДАБМ No. 233) *acéць* 'a kind of drying barn' (not a rack!) should be abandoned for semantic and phonological reasons. On the Slavic suffix, see now Pronk/Pronk-Tiethoff (2018: 285–286). Kroonen et al. (2022: 13) point out that Germanic can reflect **ageþjō-*. However, there is still a mismatch with regard to the medial syllable, which is long (and acute) in Baltic. In the opinion of Pisani (1968: 19–20), the foreign origin of the Baltic harrow is supported by the similarity of harrows used in the Baltic to those used in Rome. Lat. *occāre* 'till, harrow' (and the late *occa* 'harrow' — a back formation?) has also usually been adduced here.

It has been claimed that the Baltic vowel could be analogical after the verb seen in Lt. *akéti* 'to harrow' (Топоров ПЯ I: 67–68). On the other hand, the verb, like OHG *egen*, *eckan* 'to harrow' has itself been seen as a potential back-formation, which is supported by the *ja*-present (Lt. *akéja*, Lv. *ecêju*; cf. ALEW 13). One could alternatively assume secondary suffix replacement on the model of forms such as Lt. *vežéčios* 'one-horse cart' (cf. Smoczyński 2018: 10). Although I think this word may well be of non-IE origin, the potential for analogy means that there is no certain evidence. See also Kroonen et al. 2022: 13.

▸ † **'moss'**. Lt. *mūsaĩ* 'mould film (on beer, wine, etc.)'; Lat. *mūscus* 'moss'[86] ~ R *мох*, Cz. *mech*, SCr. dial. *mâh* 'moss'; OE *mos*, OHG *mos* 'moss; swamp' — Beside the Lithuanian evidence, Latin *mūscus* may also suggest an original long vowel, so we could suppose an alternation **ū* ∞ **ŭ* on this basis. On the other hand, the Latin form may equally reflect a full-grade **meus-*, corresponding to OE *mēos*, Du. obs. *mies* 'moss'. It is tempting to attribute the Baltic lengthened vowel to a secondary development (cf. Smoczyński 2018: 831), and indeed Būga (1914: 198–199; RR I: 585) has adduced Žemaitian evidence that would point to a short *-u-*,[87] supported by the forms *mọ́s*^{a} ACC.SG. and *mọsuôt*^{ę} 'mould over'

86 The Latin length is considered uncertain by Walde/Hofmann (II: 134) and de Vaan (2008: 397) and the vowel is given as short by TLL. While metrical evidence is lacking, the long vowel is clearly demonstrated by the Romance reflexes (cf. Ernout/Meillet: 424), cf. Italian *mùschio*, Spanish *musgo* 'moss', etc.

87 *Musomìs aptraukė* Salantai, *musojaĩ* Kvėdarna. As for ⟨Allus apmuſſójęs⟩ 'Kahmicht bier'

cited for Mosėdis by Vanagienė (2014: 455–456). Although the reason for this lengthened vowel is unclear,[88] it is very unlikely that the two variants should be explained as parallel loanwords from a non-Indo-European source. Note in addition that the vowel here is circumflex, while other potential examples of length alternations show an acute.

▸ † **'poplar'**. Lt. *túopa*; Lat. *pōpulus* 'poplar' ~ R *мóполь*, Slk. *topoľ*, SCr. dial. *topòla* 'poplar'[89] (Machek 1954: 132; Matasović forthc.) — The Lithuanian form, which is now part of the standard language, goes back to Būga (1908: 87; 1921: 433), where it is attributed to the East Aukštaitian dialect of Salakas. Gliwa (2008) is sceptical that this is an inherited word, and considers it more likely we are dealing with a clipping of the Slavic loanword *tópelis*. As for the *-úo-*, Gliwa's assumption of an original Žemaitian form jars with the reported East Lithuanian distribution. The vocalism could, however, be explained as a dialectal adaptation of literary short /ò/.[90] Therefore, despite disagreeing in the details, I would support Gliwa's suggestion, and suspect that *túopa* may indeed be a dialectal neologism based on a Slavic loanword.[91]

7.5.2 *Baltic Short ∞ Slavic Long*

▸ **'iron'**. Lt. *geležìs*, Lv. *dzèlzs* (dial. *dzelezs*), Pr. E *gelso* 'iron' ~ OCS желѣзо, Sln. *želẹ́zọ* 'iron' (Mikkola 1903: 41; Meillet 1909: 70; Machek 1968: 725; ALEW 351) —

(Mielcke II: 291), it is tempting to assume an error; cf. the immediately preceding "Kahmicht *mufőtas*". In the LKŽ, all of Būga's examples have been corrected to *mūs-* (cf. *Mūsõms aptraukė* cited for Salantai s.v. *mūsà*), but there is no reason to doubt their reliability.

88 It would seem most promising to start from the verb (*ap-*)*mūsóti -ja* 'become covered with a mould film', where for the lengthened grade we could compare iterative formations like *bylóti -ja* 'speak', although the verb in question does not have an iterative meaning.

89 The Slavic word is often considered a borrowing from Latin (e.g. Ernout/Meillet 924; Machek 1968: 647); however, finding a suitable source form presents difficulties (see REW III: 121).

90 Compare Salakas forms such as *puoľka* (= literary *pòlka*) 'a dance', *kaľiduõras* (= virtual **kolidòras*, literary *koridorius*) 'corridor' (Zinkevičius 1966: 69–70). In a similar area, we find *tòpalas* 'poplar' Kazitiškis (just 15 km from Salakas), *tòpolis* Kupiškis (LKŽ). Some of these forms from the LKŽ might even be normalizations of dialectal /túop-/, but note that Vosylytė (2013: 377) only cites forms with short /ò/, e.g. *tɔ́.pɔ.l'ɔ*. GEN.SG. Kupiškis.

91 On the other hand, Gliwa (2008: 241), is rather dismissive of the LBŽ's further citations from the South Aukštaitian dialects of Seiniai and Alytus, stating that these may ultimately trace back to Būga, but without evidence. Another of Gliwa's arguments is that the poplar (*Populus alba* and *Populus nigra*) is not found in Lithuania. This statement is consistent with the distribution maps on https://euforgen.org/, but not with those of the *European Atlas of Forest Tree Species*, where both species are marked as native to Lithuania. The only widespread dialectal term for 'poplar' quoted in the LKŽ which is not loaned from Slavic *mónoль* is *jõvaras* — another Slavic loanword (Skardžius 1931: 90; LEW 195).

In Lithuanian, there is rather a lot of evidence for an original root noun; particularly note the NOM.PL. *gẽležes* recorded in several dialects which otherwise only appear to have old consonant stems.[92] However, an ablauting **ǵʰ*-stem (Tremblay 2004) can hardly come into question here, first of all due to the lack of parallels, and secondly due to the acute intonation in Slavic (R *желя́зо*, cf. Зализняк 2019: 508; SCr. (Čak.) *želèzo*; cf. Derksen 2015: 555). Thirdly, there is the obvious chronological issue of reconstructing an archaic Indo-European nominal paradigm for a designation for 'iron'. The difference in vowel length would rather speak in favour of the word entering Baltic and Slavic independently. On the various unsuccessful external comparisons, particularly with Gr. χαλκός 'copper, bronze', see Thorsø et al. (2023: 113).

▸ '**ruffe**'. Lt. dial. (S Aukšt.) *ežgẽ*, also rarely *ežegỹs* (cf. eʒ́gis, Ruhig II: 220) 'ruffe'; Pr. E *assegis* · persk[93] ~ Kash. *jôżdż* (GEN.SG. *jażdża*), Pl. *jazgarz*, Cz. *ježdík* 'ruffe' — The Slavic forms require a reconstruction **ēzg-* or **āzg-*. Although the word has a limited distribution within Slavic, the discrepancy in vowel length rules out the possibility of a Baltic loanword. Derksen (2008: 155; 2015: 159) states that the *-g-* in Lithuanian "may be the well-known Baltic intrusive velar". This can clearly not be correct, first and foremost due to the trisyllabic form attested in Prussian and Lithuanian dialects.[94] ALEW (p. 309) explains the Slavic vocalism as due to the influence of Pl. *jaź*, Cz. *jesen* 'ide', yet this is a very different kind of fish (cf. Sławski SEJP I: 533). While the traditional etymological comparison with *ežỹs* 'hedgehog' might be semantically acceptable (Trautmann 1910: 305; Derksen 2015: 159), it cannot be substantiated without *ad hoc* morphological assumptions.

---

92 The form is widespread in Uteniškiai dialects: Dusetos, Užpaliai, Debeikiai (Zinkevičius 1966: 264), Leliūnai (*Papildymų kartoteka*). From these dialects, Zinkevičius otherwise cites only NOM.PL. *dùres* 'door', *díeveres* 'brothers-in-law', *óbeles* 'apples', *vóveres* 'squirrels' and (from Debeikiai) *aũses* 'ears'. All of these are probably or possibly old consonant stems. The form *gẽležes* is also cited from the South Aukštaitian dialect Seinai, and is the only form cited by Zinkevičius from this dialect (from the LKŽ we can also add *dùres* NOM.PL. 'door' and *žuvès* GEN.SG. 'fish'). I therefore do not think that the ALEW (351) is justified in calling the consonant stem inflection secondary here, despite the *i*-stem inflection in the earliest texts. See also Tremblay (2004: 239).

93 Interpreted by Trautmann (1910: 305) as *Perca fluviatilis*, i.e. 'perch' (thus also Endzelīns 1943: 145; Топоров ПЯ I: 133; PKEŽ I: 104), but as correctly noted by ALEW (309), it can hardly be excluded that the actual meaning of the Prussian word was 'ruffe', which is considered a kind of perch in German folk taxonomy ("Kaulbarsch").

94 But even without these forms, the idea that *ežgẽ* should somehow be a back formation(?) from *egžlỹs* (attested lexicographically, cf. egʒ́lys, Ruhig II: 220), which has "preserved the original constellation" is implausible.

TABLE 17 Possible examples of long ∞ short alternations

	Baltic	Slavic	Germanic	Elsewhere
'apple'	*âBōl-	*âBl-	*abl-	Celt. *aBōl-
'cottage'	*trâB-	–	*t(u)rb-	It. *trēb-
				Celt. *treB-
'fresh'	*prêsk-	*prêsk-n-	*prisk-	
'lynx'	*lûḱ-	*[l]ûḱ-	*luk-s-	Gr. *lunk-
? 'ash'	*âs-	*ôs-en-	–	It. *os-Vn-
				Celt. *os-n-
? 'nettle'	*nât-	*nât-	*nad-	Celt. *ninat-
'iron'	*Geleǵʰ-	*Gelêǵ⁽ʰ⁾-o-	–	
'ruffe'	*eǵʰegʰ-î̯-	*ēźgʰ-	–	

7.5.2.1 Conclusion

The certain and possible evidence long ∞ short vocalic alternations is collected in Table 17, above. As in previous tables, long vowels which turn up as acute in Balto-Slavic are written with the caret ⟨ ˆ ⟩ (see p. 185 for more help reading the table). Forms which do not provide relevant data are presented in light grey. Shaded cells indicate forms containing long vowels.

Quite a large number of examples have been identified which show an unexpected long vowel in Balto-Slavic by contrast to other European comparanda. In every case, the vowel is acute, and remarkably, a Proto-Balto-Slavic reconstruction is possible, suggesting that we are dealing with a relatively significant time depth. Where we find an acute vowel, it is possible that something other than length is responsible, such as glottalization.

7.5.3 i/u ∞ ∅

▸ **'alder'**. Lt. *aĨksnis*, dial. (NE) *alìksnis* (cf. p. 33), Lv. *àlksnis*; R *ольха́*, Pl. *olcha*; ON *ǫlr*, OHG *elira* 'alder' ~ Lat. *alnus* 'alder' (Machek 1954: 130; Polomé

1990: 334; Huld 1990: 401–402; Derksen 2002: 6, 2008: 307; de Vaan 2008: 34–35; Kroonen 2013: 22; Pronk 2019a: 154; Matasović forthc.) — The Slavic vocalism presents difficulties. South Slavic in general suggests **elixā-* (trad. **(j)elьxa*): SCr. obs. *jelha* (> *jóha*; Skok I: 772), Sln. *jéłša*, Bg. *елшà*, dial. *елхà* 'alder'. Polish *olcha* and Cz. *olše* demand initial **a-* (trad. **o-*), as does apparently East Slavic, where one usually anticipates the preservation of **e-* before **ĭ* (trad. **ь*) in the following syllable (Шахматов 1915: 140–141; REW I: 389, s.v. *ёвня*). It has been suggested that some forms may result from contamination with the word for 'spruce', cf. R dial. *ёлха* 'alder' beside *ёлка* 'spruce' (Kortlandt *apud* Schrijver 1991: 41), and Bg. *елхà* 'conifer, fir tree', dial. 'alder'. While it is unlikely that contamination with the word for 'spruce' can explain all cases of **e-* (Derksen 2008: 370),[95] the evidence is difficult to evaluate in view of the more general problems with initial vowels in Slavic (cf. p. 244 and Andersen 1996a: 128–130).

A more remarkable issue is posed by the Latin form. While all the remaining data points to **alis-*, Latin is only consistent with a reconstruction **als-no-* (Walde/Hofmann I: 31; Schrijver 1991: 42). This can be considered clear evidence for an irregular alternation **-i-* ∞ **-∅-* and therefore offers some empirical support for the hypothesis that the word for 'alder' is of non-IE origin.

▸ **'ground elder'**. Lt. *garšvà* (dial. *gáršva*), Lv. *gãrsa* 'ground elder' ~ OHG *gires* · macedonicum; cf. *gierisch* 1604 'aegopodium', *giersig* 1616 'wild angelica', modern *Giersch* 'ground elder' (DWb VII: 7388–7389); MLG *gers*, *gersele* · grot petercilie (MoLG *Heers* with unclear anlaut, cf. Marzell I: 125) — Although the Germanic and Baltic forms are usually compared without question (e.g. IEW 445; EWAhd IV: 370–372), the almost consistently disyllabic form in OHG (AWb IV: 285), which can hardly be explained as svarabhakti (cf. Reutercrona 1920: 137, 169), as well as perhaps Early MoHG *gierisch* (see above) and Swiss dial. *Gerrist* (*Schw. Id.* II: 404), seem to suggest a disyllabic preform, e.g. **gʰeru/is-* or **gʰirVs* (where **V* is not **a*), which cannot easily be compared with the Baltic words. Thus, if the comparison is correct, we are dealing both with an irregular loss of the second syllable vowel in Baltic, which can hardly be explained in Indo-European terms, as well as an alternation between a front vowel in Germanic and back vowel in Baltic (on which see 7.2.2).

95 The Russian dialectal evidence is in fact more complicated, as we also find forms like dial. *елóха* (РЭС XI: 325–326). However, even if we assume an original **eluxā-* (trad. **(j)elъxa*), this dialectal variant must in any case be analogical (after GEN.PL. *елóх*). Rather than multiplying entities, I suspect that this form is ultimately the result of a dialectal hardening of /l'/ in the sequence /l'x/, although more evidence would be desirable (Мирская *apud* Касаткин 1999: 177 mentions the dialectal forms *Óльга* 'Olga', *скóлько* 'how much', but the distribution of this phenomenon is not clear to me).

▸ **? 'aspen'.** Lt. *ẽpušė*, dial. *ãpušė* 'aspen' ~ Lv. *apse*; Pr. E *abse*; R *оси́на*; ON poet. *ǫsp* 'aspen', Arm. dial. *opʿi* 'poplar' (for refs. and further discussion, see p. 219) — In view of the Lithuanian data, the Latvian word has been derived by syncope from an earlier **apuse* (e.g. Schulze 1913: 288; Smoczyński 1989: 40). However, syncope does not generally occur in Slavic, so here **aps-* really is required, in line with Armenian. Būga (1922: 226) has suggested that the Lithuanian word arose through contamination with *pušìs* 'pine', an idea which has been taken over in other etymological works (Trautmann 1923: 11–12; LEW 14; Smoczyński 2018: 40; note also ALEW 45). However, a change **-ps-* > **-puš-* involves both a vowel epenthesis and a change in sibilant quality,[96] and seems hardly imaginable, especially since the trees in question are not very similar. Endzelīns (1943: 136; cf. Būga 1908: 118), starts instead from **aps-ušē-*, with loss of the first **s* by dissimilation, but such a dissimilation would be unparalleled (cf. Zinkevičius 1966: 181–182).

Already Hoops (1905: 123–124) drew attention to a group of similar forms in the Turkic languages. Chuvash *ăvăs* 'alder' seems to reflect an earlier **abus* (Räsänen 1969: 3; Мудрак 1993: 29) or **abïs* (ЭСТЯ I: 607–608). The **b* is also supported by e.g. Siberian Tatar *ausaq*, dial. (Tomsk) *apsaq* (Тумашева 1992: 25, 32) 'aspen', which reflects a derivative **abs-ak* with regular syncope (cf. А. Дыбо 2007: 130). The similarity of Turkic **abus* and the possible Proto-Baltic form **apuš-* is striking, but since the Turkic **-u-* may be due to anaptyxis in a final cluster **-bs* (cf. СИГТЯ VI: 65), it does not unambiguously support the reality of this Baltic reconstruction.

In fact, the reconstruction **abs* rather than **abus* might better account for Khakas *os* and Tatar *usaq* 'alder', which appear, at face value, to reflect Turkic **os* (the expected Khakas reflex of **abus* would be **ōs*, with a long vowel; А. Дыбо 2007: 19). Hoops (followed by СИГТЯ IV: 131) had considered these to be Slavic loanwords, but it would be rather remarkable if Tomsk *apsaq* were unrelated to Tatar *usaq*, with its identical suffixation; moreover, a suitable Slavic source is unattested.[97] A very similar correlation is found between Chuv. *avăt-* and Old Turkic *öt-* 'sing (of birds)'. which is reconstructed **ebt-* by Мудрак (СИГТЯ VI:

96 Būga had previously expressed the view that **s* became **š* after labials (1911: 3). This would help his case, but as it is clearly contradicted by Lt. *vapsvà* 'wasp', he had apparently already rejected the development by the time of this proposal, where he states explicitly that the expected Lithuanian form would be **apsė*.

97 Hoops suggests a Slavic donor form **osa*; however, such a form is only attested as a relic in West Slavic, and does not occur in East Slavic at all. Note that the Belarusian *acá* cited in ЭССЯ (XXXII: 93) does not exist. In the original source (Лемцюгова 1970: 7), it is only a reconstructed form (*ACA *sic.*) based on toponymic evidence.

166; in his notation **ĕvt-*), implying a reduction **eb-* > **ö-* before a consonant. Therefore, one way to connect the Khakas/Tatar forms would be to assume a parallel development **abs* > **ŏs*.

Perhaps more problematic, and something which does not seem to have been noted, is the fact that we find **-b-* in Turkic rather than **-p-*. This might in fact be better accounted for by assuming a disyllabic donor form in which **p* had become lenited intervocalically. We can recall here Hoops' (loc. cit.) suggestion of an Iranian source, but are faced again with the issue that no trace of the word has been found in Indo-Iranian.[98]

The main issue with deriving the Turkic words from a (para-)Baltic source is the word's broad distribution in Siberia. This could theoretically be accounted for by assuming a loan already into 'dialectal' Proto-Turkic (but from where?) or by assuming a later horizontal spread through the Turkic dialects, which could perhaps provide an alternative account for the irregularity within Turkic, but is difficult to substantiate in any detail. Compare also the discussions of the words for 'honeycomb' (p. 133), 'mink' (p. 143) and 'elm' (pp. 235–236).

All in all, the word for 'alder' raises a number of problems which preclude its reconstruction, and it is possible that this could be explained by assuming parallel loanwords from an unattested source language. However, this does not really help to resolve the word's problematic distribution within Siberian Turkic.

▸ **? 'beehive'.** Lt. *avilỹs*, (Žem.) *aulỹs* 'beehive', Lv. (Kurzeme) *aũlis*, also *avelis* 'wooden beehive' ~ R *у́лей*, Pl. *ul* 'beehive', Sln. *ûlj* 'hollow tree; beehive' — The word is generally connected to Lat. *alvus* 'belly', *alveus* 'hollow vessel', Gr. αὐλός 'pipe, hollow tube' (Trautmann 1923: 18; LEW 25–26; REW III: 181; ALEW 77–78). On semantic grounds, the comparison can hardly be faulted, as Latin *alvus* is also used in the sense 'beehive'. On the other hand, the Baltic forms are far easier to explain starting from **avil-*, with the *aul*-forms deriving by syncope, and indeed, the Lithuanian variants have led Zinkevičius (1966: 138) to doubt the IE etymology. ALEW speculates that *avilỹs* is due to reanalysis on the basis of an unspecified root **au-*, but the main issue is that the suffix *-ilỹs* is not ana-

98 Leaving aside the supposed connection with the Indo-Iranian word for 'shovel; shoulder-blade' (Friedrich 1970: 50–52; Гамкрелидзе/Иванов 1984: 627; Šorgo 2020: 434), which rests on pure speculation (cf. KEWA III: 547 with older lit.; Normier 1981: 24, fn. 21). The **b* is also problematic to Pedersen's suggestion of a "pre-Armenian" source (1906: 462) and A. Дыбо's suggestion of a Tocharian source (2007: 130). The discovery of Arm. *op'i* causes additional problems for both proposals with regard to vocalism.

lysable. Therefore, despite the attractiveness of the IE etymology, the unclear *-i-* might suggest an irregular correspondence with Slavic, which could point to a foreign origin.

▸ † **'hazel** (2)'. Lt. *kafsula* 'Schaft' (Bretke *apud* Bezzenberger 1877: 293), dial. (S Aukšt.) *kasulà* 'plough shaft' (LKŽ), *kaffúlas* 'Jägerspieß' (Ruhig I: 213) ~ Lat. *corylus* (< **kosVlo-*), OHG *hasal*, OIr. *coll* (< **koslo-*) 'hazel' (Huld 1990: 401; Matasović 2013: 84, forthc.) — As Latin can simply reflect **kose*/*olo-*, with suffixal ablaut, the irregularity depends entirely on the Baltic data, whose appurtenance is uncertain on semantic grounds. I therefore exclude this example.

7.6 IE **a*

It has been suggested that many of the words traditionally reconstructed with **a* are rather loans from non-IE sources (Kuryłowicz 1956a: 194–195; Kuiper 1995: 65–68; Pronk 2019a: 154). The argument is essentially that most words for which **a* has been reconstructed are geographically restricted, have a technical meaning, and often involve other irregularities. The following have already been regarded as probable loanwords on other grounds in the previous sections:

– **ǵʰuak-* 'torch' (p. 167): illegal root structure[99]
– **dʰraK-* 'dregs' (pp. 190–191): **gʰ* ∞ **k*, illegal root structure
– **dʰalK-* 'scythe' (p. 191): **gʰ* ∞ **k*, illegal root structure
– **kanaP-* 'hemp' (pp. 206–207): **b* ∞ **p*, **nn* ∞ **n*
– **bʰar(s)dʰ-* 'beard' (pp. 224–225): **sdʰ* ∞ **dʰ*
– **bʰaB-* 'bean' (pp. 228–229): **bʰ* ∞ **w*
– **trā*/*ĕb-* 'cottage' (pp. 270–271): **ā* ∞ *ē*, IE **b*

Still, given the continued disagreement as to whether **a* should be reconstructed, it might seem overly dismissive to label any word appearing to suggest the reconstruction **a* (for which the most decisive evidence comes from Italo-Celtic and Greek) as being of non-IE origin, especially where no other evidence supports this hypothesis (compare the similar considerations with regard to the word for 'apple' on pp. 268–270). Here, I will briefly treat a few words which fall into this category:

▸ ? **'post** (2)'. Lt. *stãbas* 'pillar; idol, statue'; ON *stafr* 'staff, cane; post, support' ~ OIr. *sab* 'pole, stake' (Beekes 2000b: 12) — The Irish word is unlikely to be bor-

99 Schrijver (1991: 465) has considered Lat. *fax* an example of the unrounding of **wo* in an open syllable (pre-Latin **fwak-* < **fwok-*), but the development is clearly contradicted by Lat. *forum* 'open space' (< **fworo-* < **dʰuoro-*; cf. idem: 472), and can hardly be accepted.

rowed from Old Norse due to the unique substitution *st-* → *s-* (cf. Marstrander 1915: 97, 125). A reconstruction **stHbʰ-* would be unproblematic for Germanic and Celtic, but whether it could account for the Baltic data is disputed. According to one widespread current view, in Balto-Slavic, "a laryngeal was never vocalized" (Beekes 1988: 23; cf. also Lubotsky 1981: 89; Smoczyński 2006: 187–188).

On the other hand, it seems the evidence is not exactly decisive. The "classic" view, at least, states that the Indo-European 'schwa' yielded Balto-Slavic **a* (cf. Brugmann 1897: 177; Arumaa 1964: 80–81; Stang 1966: 22; Matasović 2008: 89), for which the *Paradebeispiele* — Lt. *statùs* 'upright; steep' (= Gr. στατός 'standing') and OR споръ 'abundant', Cz. *sporý* 'stocky; abundant' (= Skt. *sphirá-* 'fat') — still maintain much of their initial appeal. The reluctance in accepting this sound law seems mainly to be based on the small number of examples, but as long as no counter-evidence exists, it cannot be rejected out of hand.[100] I therefore agree with Villanueva Svensson (2008: 12) that the issue is in need of "a full and unprejudiced study".[101]

▸ ? **'mast'**. OCS мостъ, SCr. *môst* 'bridge'; OE *mæst* 'mast', OHG *mast* 'mast, pole' ~ OIr. *maide* 'stick, staff; beam'; Lat. *mālus* 'mast, pole' (Kuryłowicz 1956a: 195; Pronk 2019a: 151) — Again, a reconstruction **mHsd-* is possible at least for the extra-Balto-Slavic evidence (Schrijver 1991: 167). The Slavic **t* might be more elegantly explained by positing a Germanic loanword (Stender-Petersen 1927: 281–283; Matasović 2008: 50; Kroonen 2013: 357), although this is uncertain for semantic and accentological reasons (REW II: 163; Pronk-Tiethoff 2012: 180).

---

100 Alternative explanations for both have been suggested, but the question is, even if these explanations are in principle possible, are they an improvement on the traditional etymologies? Since Kortlandt (1982: 26), in my view rightly, rejects the notion that OR стрыи 'father's brother' is derived from **ph₂tr-*, it is unclear on what basis he is opposed to the development. He takes *statùs* as a derivative of *statýti* 'put (upright)', in turn from a reduplicated **ste-sth₂-* (Kortlandt 1989b), but the opposite derivation seems more likely from a Baltic perspective (cf. Smoczyński 1999: 23); moreover, in Kortlandt's scenario, the *-a-* still has to be explained as secondary. For споръ, the reconstruction **su-para-* (i.e. **sъporъ*; Kortlandt 1980: 352) appears to be in contradiction to the attested Old Russian evidence (cf. СРЯ 11–17 XXVII: 72).

101 Other examples to consider here are Lt. *mãtas* 'measure' <? **mh₁-to-* (Darden 1990: 63; Smoczyński 1999: 23), and OCS столъ 'throne, bench' (beside Go. *stols* 'seat, throne'), cf. Smoczyński (1999: 20). For alternations such as CS носъ 'nose' : Lt. *nósis* and OCS соль 'salt' : Lv. *sàls*, Matasović (1997: 135) has reconstructed **nh₂s-*, **sh₂l-*. At first sight, this indeed seems preferable to assuming the preservation of an extremely archaic paradigm **naʔs-* : OBL. *nʔas-* into Proto-Balto-Slavic (Kortlandt 1985: 119), but Sanskrit *nás-* 'nose' must also be accounted for (cf. Lubotsky 1981: 90). For 'salt', an equally possible reconstruction is **sh₂-ēl* : **sh₂-el-*, provided Eichner's Law is rejected (see Pronk 2019b: 144–145).

Admittedly, assuming a suffixed formation **mazd-to-* for Slavic is not much of an improvement (Kiparsky 1934: 47; Derksen 2008: 326–327).

▸ **? 'corner'.** Lt. *kaṃpas* 'corner, angle' ~ Gr. κάμπτω 'bend (the knee), turn back (a chariot)' (Kuryłowicz 1956a: 195; Beekes 2000a: 28) — Here, a reconstruction **kh₂mp-* is improbable, if not impossible (Beekes 2010: 632). On the other hand, the possibility that Gr. κάμπ- is secondary for *κάπ- with the analogical restoration of -μ- from the full-grade (Pronk 2019a: 149) remains plausible, if somewhat convoluted (cf. χανδάνω 'hold' for *χαδάνω < **gʰnd-nH-* beside PERF. κεχόνδει). OIr. *camm*, MW *cam* 'crooked, bent', if they belong here, could possibly reflect a zero-grade **kmp-* (with a development **-mp-* > **-mb-*; Thurneysen 1946: 117).

∴

In conclusion, none of these potential examples of **a* are entirely watertight, and therefore the question as to whether the apparent presence of such a phoneme is sufficient to prove a non-Indo-European origin need not be answered here. However, the characterizations of authors such as Kuryłowicz do seem to be generally valid, and I consider this to be another potential criterion which could favour a non-Indo-European origin, at least where other evidence is available.

CHAPTER 8

Analysis

In the above pages (including two examples in 5.3.1), I have discussed 92 word families which might plausibly be explained as loanwords from unattested non-Indo-European sources. In 16 cases, it was found that the evidence is too ambiguous or uncompelling, and these cases will not be fed into the further analysis. Of the remaining examples, I have considered 46 to be probable loanwords, and a further 30 have been accepted as possible, but uncertain. In this section, I will analyse the data from an extra-linguistic perspective, as well as attempt to draw some broader conclusions about the dataset as a whole. In this context, the certain cases will be used as my core data set, with uncertain examples only being incorporated where this could provide additional useful information.

8.1 Semantics

The majority of the words treated here fall into the following broad semantic categories (uncertain cases are listed in square brackets):

a. **Wild animals** (12 + 7):
   Mammals: bison, lynx, roe (deer) [+ badger, boar]
   Birds: bird of prey (see 5.3.1), grouse, oriole, pigeon [× 2], swan (× 2)
   Aquatic animals: ruffe, seal, sturgeon [+ cod, frogspawn, salmon, fishing trap]

b. **Wild plants** (11 + 8):
   Trees: alder, hornbeam, maple [+ ash, aspen, elm (× 2)]
   Tree parts: leaf, nut [+ bast]
   Edible plants: (wild) carrot, goosefoot, ground elder, ramsons [+ nettle]
   Other: (false) hellebore, reed [+ heather, sedge]

c. **Cultivated plants and agriculture** (9 + 5):
   Crops: bean, hemp, oats, rye [+ lentil, millet]
   Fruits and vegetables: apple, garlic, pear, turnip
   Agriculture: scythe [+ furrow, ploughshare, aftermath]

These three categories already account for 70 % of the certain examples. Beyond this, three relatively clear semantic groups can be identified with at least two certain examples:[1]

d. **Apiculture** (3 + 2): drone, honeycomb, wax [+ bee, beehive]
e. **Structures** (2 + 2): cottage/estate, oven [+ mast, post]
f. **Metallurgy** (2): iron, silver

Each of these six categories will be discussed below in more detail, but first I would like to point out some absences. Most remarkable here is the absence of geographical terminology and terms for natural phenomena, especially considering that these semantic areas have received much attention from other researchers focusing on questions of substrate contact (cf. Kalima 1919: 257–258; Bertoldi 1932: 94; Ariste 1971: 9–10; Polomé 1986: 662; Saarikivi 2004; Aikio 2009: 41, 2012a: 85; Beekes 2014: 47–51). The only word fitting into this category is the uncertain case 'lightning', but even here the precise semantics might point towards borrowing in a religious or cult context (see p. 202). While it is true that some suggestions of non-IE origin have been made in connection to Baltic geographical terms (e.g. 'meadow', p. 198), I have found none of these to be compelling.[2]

Another semantic category which is under-represented, although perhaps less surprisingly, is that of animal husbandry. Outside of words connected with apiculture (on which see 8.1.4, below), the only term in my corpus which falls into this category is the adjective 'in calf'. As in many Indo-European languages, the Baltic lexicon for domestic livestock is conservative, with most important terms being directly inherited from the proto-language.[3] We have also observed that several words connected to livestock breeding were loaned into Proto-Finnic (see 3.6.1), and the main foreign source for words in this semantic field appears to have been Germanic (see Chapter 2). This points towards a continuity in animal husbandry practices among Baltic-speaking populations since Proto-Indo-European times, and relatively advanced stockbreeding practices compared to their non-Indo-European neighbours.

---

1 With one certain example, we can also note body parts: beard [+ heel, palate]. Note in this context Ariste's mention of "somatic words" as good candidates for substrate loans (1963: 17).

2 Lt. *mãrios* 'sea; (Curonian) lagoon' has often been considered to be of non-IE origin (Nehring 1959; Hamp 1979: 162–163; Sausverde 1996: 136), but since Latin *mare* 'sea' has been regarded as either a regular cognate (Schrijver 1991: 474–475) or analogical (Vine 2011), this word has not come into consideration here. See also Chapter 7, fn. 50 on Lt. *pélkė* 'marsh'.

3 For instance, Lt. *avìs* 'sheep', *ožỹs* 'goat', Lv. *gùovs* 'cow' (= Skt. *ávi-* 'sheep', *ajá-* 'goat', *gáv-* 'cow'), Lt. *paršas* 'piglet, castrated boar' (= Lat. *porcus* 'pig').

The remaining certain cases are difficult to group together in any meaningful way, largely because their meanings are too general to be categorized, or because they cover multiple possible semantic fields. For instance, the words for 'dregs' (also 'yeast') and 'fresh' (also 'unleavened'), as well as 'oven' (categorized here under Structures) could all be associated with breadmaking, but since the attested meanings for each term are not limited to this semantic domain, such a grouping is too optimistic.[4]

8.1.1 *Wild Animals*

Terms for animals have often been mentioned as especially strong candidates for borrowing in substrate contact situations (e.g. Bertoldi 1932: 94; Philippa et al. 1: 22; Matasović 2013: 76). In this context, it is notable that the words for wild animals show a more limited distribution in comparison to other semantic categories. Out of 12 probable cases, 9 are limited to Baltic, Slavic and Germanic. This suggests that we are dealing with localized terms rather than extensive horizontal transmission, which is consistent with a substrate mechanism.

Table 18, overleaf, illustrates the distribution of the most certain cases in this semantic category.[5] Forms for which a common proto-form can (theoretically) be reconstructed have been enclosed in dotted lines. Thus, the word for 'lynx' is potentially reconstructible for Balto-Slavic, and also for Graeco-Armenian (note that the Balto-Slavic and Graeco-Armenian reconstructions are not reconcilable).

As regards the motivation for the borrowing of animal names, it seems natural to assume that words for local species for which no term was previously available would be most prone to adoption (Haspelmath/Tadmor 2009: 51; Aikio 2012a: 85). However, this explanation can only apply in the minority of the cases in our corpus. One such case is 'lynx': the animal's current range does not extend beyond the forest steppe, and as the animal's preferred habitat is dense woodland (Nowak/Paradiso 1983: 1072),[6] it is unlikely it would have frequented the grasslands further south. The animal is absent in the steppe, but is recorded in the Neolithic from Trypillia (Mallory 1982: 208), a culture with

4 Both Baltic and Slavic borrowed the Germanic word for 'bread' (see p. 38), although Lithuanian has also preserved an older inherited term, *dúona* 'bread, loaf' (LEW 111). Latvian *màize* 'bread' is derived from *mìeži* 'barley'. Other uncategorized words are 'thousand', 'torch' (p. 167) and the uncertain cases 'circle', 'corner', 'fast' and 'people'.

5 The abbreviations used are as follows: B = Baltic, S = Slavic, G = Germanic, C = Celtic, It = Italic, Gr = Greek. Under "+", I have indicated all other comparanda (with the usual abbreviations).

6 To establish current distribution, in addition to the references cited, I have referred to the data on the IUCN Red List website (https://www.iucnredlist.org/); for birds, this data has been supplemented with the maps from Birds of the World (https://birdsoftheworld.org/).

TABLE 18 Distribution of borrowed animal names

	B	S	G	C	It	Gr	+
bison	✓	✓	✓				
lynx	✓	✓	✓			✓	Arm.
roe	✓	✓					
bird of prey		✓	✓				
grouse	✓	✓	✓				
oriole	✓	✓	✓				
pigeon (1)		✓	✓		✓		Eg.
swan (1)		✓	✓				
swan (2)	✓	✓					
ruffe	✓	✓					
seal	✓			✓			
sturgeon	✓	✓	?		✓		

which Indo-Europeans potentially came into contact during their early fragmentation (Kroonen et al. 2022: 33–34). Note that the wide distribution and possibility of reconstructing common proto-forms for multiple branches might suggest that this word was borrowed comparatively early.

Lexical gaps could also account for the borrowing of a word for 'seal', an animal which is not found inland, and perhaps also 'sturgeon'. The sturgeon is anadromous, meaning it migrates upriver to spawn. Migrations are usually relatively short, but as much as 1000 river kilometres may occasionally be travelled (Holčík et al. 1989: 376; Brevé et al. 2022: 1164–1165). In addition, the now endangered stellate sturgeon previously spawned in river basins across the Pontic-Caspian area (cf. Mallory 1983: 267, 275). Nevertheless, the adoption of a foreign term could have been motivated by differences in species, a geographical gap in the distribution of sturgeon species, or changing dietary habits among migrating populations, which might have caused the original term for the animal to have been lost.[7]

However, most of the animals discussed here must have been known to speakers of Indo-European. Among the *mammals*, the roe deer, as well as the badger and wild boar, are widely distributed throughout Europe and are

7 Similar considerations could apply to the word for 'salmon', which may have originally referred to the anadromous salmon trout (Diebold 1976). Note that Mallory stresses the paucity of salmonid remains in the Pontic-Caspian region (1983: 268).

also present in the Pontic-Caspian steppe (cf. Mallory 1982: 206–207, 211, 212; Bellquist 1993: 336–337; Anthony 2007: 175). The same can be said of the bison, which still occurred east of the Dnieper into the Middle Ages (Benecke 2005), although admittedly does not appear to have been very frequent in the steppe since the Neolithic (Mallory 1982: 213). Among the *birds*, the golden oriole breeds throughout all of Europe including the steppe, and the same is true of the wood pigeon (cf. Mallory 1991: 231). The rock dove also occurs natively in the steppe; the exact vectors of spread of the domesticated and feral pigeon are difficult to trace, but it is now of course ubiquitous (Gilbert/Shapiro 2013). The mute swan breeds in many parts of the steppe, and was therefore presumably known to Indo-European speakers. It is common in Northern Europe, although its distribution admittedly becomes more patchy east of the Dniester. Among the *fish*, the ruffe has a very extensive distribution throughout Eurasia. The above facts make it is improbable that these terms were borrowed to fill lexical gaps within the Indo-European languages.[8]

It has been remarked that substrate loanwords tend to involve animals of low economic significance (Schrijver 1997: 295; cf. also Matasović forthc.). Perhaps this idea derives from an expectation that terms for economically important animals should rather be transmitted horizontally, for instance through trade.[9] However, economic significance is rather a cultural and subjective phenomenon. Even migratory passerines, such as the golden oriole, may have economic value: in coastal Egypt, they are hunted for food and sold on as delicacies (Eason, Rabia & Attum 2016).

We therefore should seek a cultural motivation for borrowing. In the case of wild animals, the most obvious cultural context is hunting. In the Eastern Baltic context, it has been suggested that the transition from a hunter-gatherer to a stockbreeding economy passed through a transitional stage where the reliance on hunting and gathering remained significant (Zvelebil/Dolukhanov 1991: 268 with lit.; Piličiauskas et al. 2017: 541), and cultural exchange might have been amplified by a later in-flow of hunter-gatherer-derived populations during the Bronze Age (Mittnik et al. 2018; Saag et al. 2019). In a context of language shift,

8 A word for 'roe' can probably be reconstructed for Proto-Indo-European on the basis of Gr. (Herodotus) ζορκάς (~ δορκάς) and MW *iwrch* 'roe deer' (< **iork*-), cf. IEW 513. For the 'swan', only a common Italo-Celtic form can be given: Lat. *olor*, MW *alarch* (?< **h*$_1$*el-r*-) (Schrijver 1995a: 76).

9 One does indeed find, for instance, a clustering of terms for insects and reptiles among the Finnic substrate terms in the Russian dialects (Мызников 2004: 113–116). Above, on p. 117, I have suggested that the common denominator between these animal terms might rather be their negative perception. This brings us back to the idea of low-status vocabulary (cf. Chapter 5, fn. 20).

technical vocabulary associated with a particular economy may be transferred into the target language, and this is likely to be more robust where activities related to this economy continue to be practiced (Brenzinger 1992). The borrowing of the technical term 'fishing trap', if reliable, would also favour this interpretation.

8.1.2 *Wild Plants*

Like animal names, plant names have often been viewed as central candidates for substrate borrowings (Bertoldi 1932: 94; Philippa et al. I: 22; Matasović 2013: 76; Soosaar 2021). The names for wild plants were a key focus of Václav Machek, and the irregularities observed in these can be seen as having given rise to his version of the substrate theory (Machek 1944–1946; 1950b; 1954; see p. 155). Although a small number of wild plant names show a narrow distribution comparable with that of the wild animals, they on the whole tend to exhibit comparanda outside of the Baltic region (see Table 19, overleaf).

First of all, it is remarkable that several of the tree names which have come into question here have been assessed as uncertain, and have therefore been omitted from the table (thus 'aspen', 'ash' and two words for 'elm'). Indeed, words for trees very often seem to show minor phonological issues, to the extent that irregularities have been viewed as a mere quirk typical of tree names (cf. E. Itkonen 1946: 306; Friedrich 1970: 108), an opinion which is perhaps justified by the perception of such terms as belonging to a 'dialectal' phase of Indo-European (cf. Hirt 1905: 189; Schrader/Nehring II: 630; and e.g. Ernout/Meillet 23 s.v. *alnus*). A case could sometimes be made for such an interpretation, especially where identical forms are found in neighbouring branches (compare the example of 'alder', above).

Aikio (2015a: 45–46) has argued that a number of West Uralic terms in this semantic field should in fact be explained as substrate words, noting that they show irregular sound correspondences:

– F *haapa*, E *haab*; Sá. N *suhpi*, Ma. E *šopke* 'aspen'
– F dial. *vahtera*, E *vaher*; Md. E *ukštor*; Ma. EW *waštar* 'maple'
– F *pähkinä*, E *pähkel*; Md. M *päšťä* 'nut, hazelnut'; Ma. EW *pükš* 'nut'; Udm. *paš-pu* 'hazel' (*pu* 'tree')

At the *Sub-Indo-European* workshop in Leiden, September 2021, Aikio has additionally adduced the word for 'alder' (F *leppä*, Sá. N *leaibi*, Md. E *l'epe*), which shows a clear resemblance to the Balto-Slavic word for 'lime tree' (see p. 89). What is remarkable is that these words also show a rather broad geographical distribution, with cognates found from Sámi to Mari or from Finnic to Permic. This probably implies that the words were adopted at a time when these branches were closer together, and it might be possible to talk of 'dialectal

TABLE 19 Distribution of borrowed plant names

	B	S	G	C	It	Gr	+
alder	✓	✓	✓		✓		
hornbeam	✓	✓			✓		
maple		✓	✓			?	
leaf	✓	✓					F Sá Ma
nut	✓	✓					
carrot	✓	✓	✓				Md
goosefoot	✓	✓	✓				
ground elder	✓		✓				
ramsons	✓	✓	✓	✓		✓	
hellebore		✓	✓			✓	
reed	✓	✓				✓	

Proto-Uralic' here, too (the words for 'maple', for instance, can *almost* be treated as regular cognates). Of the uncertain cases, the words for 'aspen' and 'elm (2)' also have potential comparanda in Turkic.[10] In this case, some kind of borrowing must certainly have taken place, but we are still left with the question as to where were these words adopted from, and what motivated their borrowing.

As the hornbeam is not currently found in the steppe, and spread to south-eastern Europe only during the Atlantic Period, being earlier restricted to Italy (Sauer 1988: 152–154), it is unlikely that the Indo-Europeans would have known this tree, and its borrowing might have been motivated by a lexical gap. The alder, on the other hand, is very widespread in Europe and should have been present in the steppe (cf. Friedrich 1970: 72–73 with lit.). The same can be said of the aspen, ash and elm.

The motivation for borrowing must again have somehow been associated with differences in cultural practices. There are few reliably reconstructible words for specific trees, and it is quite possible that trees were of lesser importance to steppe pastoralists than to the European populations they replaced. This might be implied by the large-scale deforestation (or "steppification")

10 Here we can note that Kroonen (2013: 39) has compared the European words for 'aspen' with F *haapa*, etc. I am not convinced, however, that there is sufficient similarity to warrant a comparison.

of Northern Europe coinciding with the arrival of the Corded Ware Culture (Pelisiak 2016: 218–219; Haak et al. 2023: 71–72; Allentoft et al. forthc.). Although one can hardly conclude that the Indo-Europeans did not value wood, it may be suggested that distinguishing varieties of trees was not a top priority for steppe-derived pastoralists.

I have divided the remaining plants into 'edible' and 'other'.[11] This division is somewhat arbitrary, as it is difficult to know what was interpreted as food by prehistoric populations. Cultural groups may differ in plant preferences, despite there being no significant difference in plant availability (Welcome/Van Wyk 2019). Reeds are fully edible, and may have been eaten, but since reeds also have numerous other uses (e.g. weaving mats, producing ropes), I have categorized them under 'other'. On the other hand, nettles may also be twined into string and woven into textiles. My categorization as an 'edible' plant is partially influenced by the semantic shift to 'vegetable tops' observed for this word in Slavic (but this is not necessarily indicative of its earlier uses). Furthermore, the knowledge of which plants are poisonous (such as the hellebore) is obviously most vital to those gathering plants for consumption.

Goosefoot, *Chenopodium*, is a plant whose remains are found in abundance at Yamnaya sites, with indications that it was eaten (Anthony 2007: 326, 439), so it is probable that Indo-European speakers had a word for the plant. One might assume that a decline in the consumption of this plant could have been associated with a shift towards cultivated cereals, although there is plenty of evidence of *Chenopodium* consumption even in Iron Age agricultural contexts (Kroll 1990; Behre 2008: 68–69; Ślusarska 2021: 189). Evidence for both wild garlic and nettles have been recovered in the Bronze Age Srubnaya Culture in the same region (Anthony et al. 2005: 408) as well as pollen belonging to Apiaceae (the family to which the carrot and ground elder belong). Plants from this family might have been consumed as vegetables in Western Russia already during the Neolithic (Kittel et al. 2020: 196).

It appears that the borrowing of these plant names can in no case be confidently associated with a lexical gap; on the contrary, there is evidence that many of these species were actively consumed both in the steppe and in Europe. Notably, the wave of deforestation coinciding with the emergence of animal husbandry in Northern Europe actually coincides with an increase in evidence for both Chenopodiaceae and *Urtica* (nettles) in the palynological record (Pelisiak 2016: 218–219).

11 For information about plant uses, I have referred to the *Plants for a Future* database at https://pfaf.org/, where copious further references are provided.

The borrowing of these plant names into dispersing Indo-European dialects, if not associated with a change in dietary preferences which may be the result of shifting subsistence practices, could indeed be connected to the principle of "low economic significance" signalled by Schrijver (1997: 295). Wild plants form a small but integral part of both pastoralist and agriculturalist diets (cf. Zanina et al. 2021; van Amerongen 2016: 215–226), but gathering of plants outside of a hunter-gatherer economy is presumably perceived as of secondary importance. On the other hand, we may consider a sex bias in the transmission of these terms. Ethnological evidence shows that plant gathering and preparation is cross-culturally most often the sole domain of women (Murdoch/Provost 1973: 207, 210).[12] A male sex-bias in the migrations of steppe-derived populations, combined with female exogamy (Knipper et al. 2017; Saag et al. 2017; Mittnik et al. 2019), would provide a plausible context whereby indigenous terms, passed down from mother to daughter, could resist replacement during language shift.

8.1.3 *Cultivated Plants and Agriculture*

First, it should be noted that cultivated plants and their wild equivalents are not always linguistically differentiated. Following from the discussion on pp. 229–231, I have listed 'carrot' as a wild plant, even though the term normally refers to the cultivated variety in the modern languages. On the other hand, I have included 'garlic' as a cultivated plant in view of the fact that the word, wherever it occurs, is differentiated from the wild *Allium ursinum* (another probable loanword; see 'ramsons', pp. 246–247).

Many of the crop terms have comparanda beyond Balto-Slavic, Germanic and Italo-Celtic, and many show an extremely broad distribution, suggestive of largely horizontal rather than vertical borrowing. On this basis, it can be hypothesized that the majority of terms for cultivated plants spread as *Wanderwörter*. In the case of founder crops, the spread of these words is unlikely to be directly associated with the spread of the crops themselves. Instead, it is more likely to be related to the spread of peoples and the adoption of an agricultural lifestyle. The trajectory of spread is usually difficult to establish on linguistic grounds. The material covered in this category is as follows:[13]

12 "A few men, especially those who hunt and fish, also gather some edible plants from time to time. However, it was not customary, and their knowledge of these plants was quite limited in comparison to the women's" (Ertuğ 2000: 175 in a study of a community pursuing mixed hunter-gatherer/agricultural subsistence in Anatolia).

13 Additional abbreviations used in this table: Ro = Romance, Pm = Permic, T = Turkic.

TABLE 20 Distribution of borrowed agricultural terms

	B	S	G	C	It	Gr	+
bean	?	✓	✓		✓		? Berber
hemp	✓	✓	✓		Ro	✓	Arm. Iran.
oats	✓	✓			✓		F Md Ma
lentil	✓	✓	✓		✓	?	
rye	?	✓	✓	✓			Md Pm T
apple	✓	✓	✓	✓			
garlic		✓		✓			
pear	✓	✓					
turnip	✓	✓	✓	?	✓	✓	
scythe	✓				✓		

It is notable that a disproportionate number of borrowed terms for crops are shared with Italic (including a word for 'scythe'), which seems to bring the centre of gravity towards central and southern Europe. In addition, we often find historically identical forms in several branches. In certain cases, it is possible that a word spread within IE; as discussed above (p. 194), the word for 'rye' might well have entered Baltic, and possibly even Slavic, through Germanic mediation. Similarly, the Baltic word for 'hemp' may well have been borrowed from Slavic. Nevertheless, most of the forms cannot be explained as borrowings from any attested language.

The process of Neolithization in the Eastern Baltic is extremely interesting and differs markedly from that in Central Europe. While the arrival of Corded Ware can be dated to the early 3rd millennium BCE (Piličiauskas 2018), the first individuals do not show evidence of admixture with Anatolian Farmer populations, suggesting an independent, direct migration from the steppe (Mittnik et al. 2018: 8). Although later individuals do show evidence of this ancestry, there remains no solid evidence for agriculture until the middle of the second millennium, where a few barley grains have been recovered from western Lithuania (Grikpėdis/Motuzaitė Matuzevičiūtė 2017). Here still, we also find abundant wild plant remains, suggesting a mixed subsistence involving only small-scale cultivation; moreover, it cannot be decided with certainty whether the aforementioned grains were cultivated locally or imported (Grikpėdis/Motuzaitė Matuzevičiūtė 2020: 162).

A radical reassessment of early agriculture in the East Baltic has taken place in the past few years (cf. Piličiauskas et al. 2016; Girininkas 2019). Far from

earlier claims of cultivated grains already in the 3rd millennium (Rimantienė 1992: 109–110), it has now become apparent that there is no solid evidence of agriculture prior to the Late and Final Bronze Age, i.e. the 1st millennium BCE. This is, at least, not in contradiction with dietary data,[14] where a shift to a diet incorporating cereals can only be demonstrated from the Late Bronze Age (Piličiauskas et al. 2017).

If we examine the cereal terminology in the East Baltic languages, it becomes immediately apparent that the arrival of its speakers in the region cannot be equated with the first steppe migrations, as has sometimes been suggested (Mallory 1989: 108; Rimantienė 1992: 137–138; Parpola 2012: 133; Mittnik et al. 2018: 8). What we find is that the cereal terminology in Baltic is generally archaic, with some terms directly inherited from (core) Proto-Indo-European (see below). As a result, we must assume a continuity in agricultural practices among Balto-Slavic peoples during their migration from the Indo-European homeland. This points to a much later date for the arrival of Baltic-speaking populations in the Baltic Region, the most probable proxy being the Late Bronze Age hillfort phenomenon (Lang 2016: 18, 2018). At least the following Baltic cereal terms appear to be inherited:

– Lt. *javaĩ* M.PL. 'cereal' (= Skt. *yáva-* 'grain, crop')
– Lt. *sėmuõ*, PL. *sémenys* 'seed, linseed', Pr. E *semen* 'seed' (= Lat. *sēmen*)
– Lt. *dúona* 'bread' (= Manichaean Sogdian *δ'n* 'seed')
– Lt. *žìrnis*, Lv. *zir̃nis* 'pea', Pr. E *syrne* 'grain' (= Lat. *grānum* 'grain')
– Lt. *pẽlūs* M.PL, Lv. *pęlus* F.PL.; Pr. E *pelwo* 'chaff' (= Skt. *palā́va-* 'chaff')
– Lt. *árti*, Lv. *arît* 'to plough' (= Gr. ἀρόω, Lat. *arō*)
– Pr. E *wagnis* 'coulter' (= Gr. ὀφνίς 'ploughshare')[15]

While it cannot be excluded that some of these terms originally referred to wild grains, their consistent agricultural meaning favours an early association with agriculture (see the survey in Kroonen et al. 2022). Particularly relevant are terms connected to ploughing, as archaeological evidence for ploughs and other agricultural tools in the East Baltic appears to be unreliable before the Late Bronze Age, coinciding with the archaeobotanical evidence (Lang 2007: 107; Luik/Maldre 2007: 33; Piličiauskas et al. 2016: 190–191; Girininkas 2019: 68–72).

14 Inferred from the isotopic ratios of carbon and nitrogen in the bone collagen of ancient individuals.

15 This word has been replaced in East Baltic, however, so is only indirectly relevant to this question (see the discussion of another word for 'ploughshare' on pp. 213–215). The Greek word is only known from a Hesychian gloss, but the formal correspondence with Prussian is ideal, and the word is also known from Germanic.

Interestingly, most of the inherited terms are generic in character, suggesting that, while agriculture was certainly practiced, it remained rudimentary, with different crop types perhaps not being distinguished. One inherited term has become semantically specified in East Baltic — the term for 'pea', representing a semantic shift from a generic term for 'grain' (as preserved in Prussian *syrne* and OCS зрьно 'grain'). As noted above (see 3.6.2), the semantic shift from 'grain' to 'pea' is rather surprising, as the pea, while one of the earliest crops to appear in the East Baltic, is recorded in small quantities (Pollmann 2014: 409), making its status as a staple crop improbable.

Table 21, overleaf, shows the Balto-Slavic terms for various specific crops arranged in order of their appearance in the archaeobotanical record. The periodization is based on the useful survey of the archaeobotanical evidence by Grikpėdis and Motuzaitė Matuzevičiūtė (2020). Terms that are highlighted in bold have been suggested in this book to be borrowings from non-Indo-European sources. Shaded cells indicate that a common proto-form could theoretically be set up for multiple sub-branches.

If we start from the hypothesis that the arrival of East Baltic-speaking populations in the Baltic region was associated with the emergence of diversified agriculture, then the crops should fall into two groups: the earliest crops, which might have been brought by the Balts themselves, and for which terms might already have been present in the Baltic languages prior to their arrival (i.e. possible "pre-migration terms") and later crops, adopted already *in situ*, for which any terms must postdate such a migration.

Pre-migration crops: The only term reconstructible for Proto-Balto-Slavic refers to a variety of wheat. This term is continued by Lt. (Žem.) *pūraĩ*, Lv. dial. *pûŗi*[2] 'winter wheat', corresponding to RCS пъıро (rendering Gr. ὀλύρα, ζέα, cf. СДРЯ 1759), SCr. dial. *pȉr* 'spelt' (Skok II: 660), Sln. *píra* 'spelt; (dial.) millet' and further to Gr. πῡροί 'wheat'. Due to the meaning and limited distribution, a non-IE origin has been suggested (Frisk II: 631; Lubotsky 1988: 136); however, the comparison is impeccable on formal grounds, and we must reckon with the possibility of an inherited cereal term (Nieminen 1956: 170–172; Kroonen et al. 2022: 21). The semantic specialization in Baltic is explained by the word's marginalization in favour of the loaned *kviečiaĩ*, probably associated with a transition to free-threshing wheats (see below).[16]

Beyond this, a shared word for 'barley' can be reconstructed for Proto-Baltic. It is possible that this could be connected to the role of barley as a pion-

16 Note that according to ME (III: 449–450), *pūŗi* was used in some parts of Kurzeme as a general term for 'wheat'.

eer crop in more northern latitudes (Motuzaitė Matuzevičiūtė 2018), although according to our workinghypothesis, the earliest barley finds in the Eastern Baltic should predate the arrival of the Balts. Nevertheless, the existence of a shared Baltic term might suggest barley was one of the first crops to have been cultivated by Baltic speakers. The origin of the term is unknown, however (Smoczyński 2018: 798; Kroonen et al. 2022: 15–16), and a post-Proto-Baltic diffusion cannot be ruled out.

TABLE 21 Chronology of cereal terms in Baltic

	East Baltic	**Prussian (E)**	**Slavic**
LATE BRONZE AGE (1st millennium BCE) — "pre-migration"			
barley	Lt. *miẽžiai* Lv. *mìeži*	*moasis*	RCS ячьмъı Sln. *jéčmen*
hulled wheat (spelt)	? Lt. *pūraĩ* 'winter wheat'	_	RCS пъıро SCr. *pȉr*
broomcorn millet	**Lt. *sóros*** **Lv. obs. *sāre***	*prassan* [?← Pl.]	Pl. *proso* SCr. *prȍso*
pea	Lt. *žìrnis* Lv. *zir̃nis*	[*keckers* ← MLG]	R *горóх* SCr. *grȁh*
(broad) bean	**Lt. *pupà*** **Lv. *pupa***	*babo* [?← Pl.]	**R *боб*** **SCr. *bȍb***
false flax, *Camelina sativa*	Lt. *jùdros* Lv. *idra* [?← F]	_	? Pl. *rydz* ? Sln. *rîdžək*
ROMAN IRON AGE (1st–8th c. CE) — "post-migration"			
free-threshing (bread) wheat	[Lt. *kviečiaĩ* Lv. *kvìeši* ← Go.]	*gaydis*	OCS пьшеница
rye	[Lt. *rugiaĩ* Lv. *rudzi* ← Go.]	*rugis* [?← G]	**R *рожь*** **Sln. *rȓž*** [?← G]
oats	**Lt. *ãvižos*** **Lv. *àuzas***	***wyse***	**R *овёс*** **Sln. *óvəs***
flax	Lt. *linaĩ* Lv. *lini* [?← Sl.]	Pr. G *lino*, *lynno* [?← Sl.]	R *лён* Sln. *lân*

TABLE 21 Chronology of cereal terms in Baltic (*cont.*)

	East Baltic	Prussian (E)	Slavic
hemp	Lt. *kanãpės* [?← Sl.]	*knapios* [?← Sl.]	**R *конопля́*** **Pl. *konopie***
turnip	**Lt. *rópė***	–	**R *ре́па*** **SCr. *рѐпа***
MIDDLE AGES (13th–14th c. CE)			
lentil	**Lt. *lę̃šis*** [Lv. *lę̃ca* ← R]	[*lituckekers*]	**RCS лача** **SCr. *léća***
opium poppy	[Lt. *aguonà* Lv. *maguône* ← G]	*moke* [?← Sl.]	R *мак* Sln. *màk*

Although false flax (*Camelina sativa*) is normally interpreted as a weed in southern European Neolithic contexts (Zohary/Hopf 2012: 111), it appears that it was cultivated before flax in the Eastern Baltic, perhaps serving both as an oil plant and as animal fodder (Pollmann 2014: 412–413). No certainly old designation for false flax can be identified in Balto-Slavic.[17] It is conceivable that the modern word for 'flax', which could theoretically be reconstructed for Proto-Balto-Slavic, was applied to this plant, or served as a general designation of oil plants. Pollmann notes that the same area where abundant remains of *Camelina* were identified archaeologically was later known for flax cultivation (2014: 413). However, it cannot be entirely excluded that the East Baltic terms were adopted from North Russian as late as the Middle Ages (cf. 1.1).

The East Baltic designations for 'millet' and 'bean' are both possible borrowings from non-IE sources, although for 'bean', I have considered the inclusion of the Baltic data uncertain (see pp. 228–229). For millet, the main evidence is the existence of comparanda in Mordvin, which cannot be explained as direct borrowings. In principle, it is possible that the Balts picked up millet cultivation from Central Europe, where millet was well established from the 2nd millennium (Filipović et al. 2020). However recent investigations demonstrate

17 Perhaps the best candidate for a Proto-Slavic term is R *ры́жик*, Pl. *rydz* (see atlas.roslin.pl/plant/6517, accessed 9 November 2023), Sln. *rîdžək* (Pleteršnik II: 426) 'false flax', which all derive from an adjective continued by R *ры́жий* 'red-haired', Pl. dial. *rydzy* 'copper-red', Sln. (Pleteršnik) *rîdž* 'fuchsgelb'.

another centre of spread in Central Asia (Motuzaitė Matezuvičiūtė et al. 2022). Widespread evidence of millet can be identified in the Pontic steppe region as well as in northwest Kazakhstan from the 1st millennium BCE. It is possible that an eastern centre of spread could account for the linguistic facts more effectively, although more evidence is required to establish the archaeological plausibility of this scenario. If true, the word for 'millet' can be identified as a *Wanderwort* with its roots in an unidentified Central Asian language.

Post-migration crops: In Chapter 2, I have argued that the East Baltic term for 'wheat' is a loan from East Germanic. Since the possible timeframe for contacts with Germanic coincides more or less with the first reliable evidence for free-threshing wheat, in particular bread wheat, *Triticum aestivum* (Grikpėdis/Motuzaitė Matezuvičiūtė 2020: 164), there is a plausible archaeological context for the adoption of this foreign term (note also that a term for 'bread' was borrowed from Germanic). Considering the similar chronology of rye cultivation in the region, it is probable that the word for 'rye' was taken from the same source. The Baltic word for 'hemp' is possibly a Slavic loanword, as is the word for 'flax' (see above).

Interestingly, at least two "post-migration" crop names — 'oats' and 'turnip' — are clear borrowings from unknown sources.[18] The comparanda for both of these point towards central or southern Europe. Both terms are shared with Italic, and are actually attested in literary sources in Latin several centuries before they emerge in the Baltic archaeological record,[19] which strongly implies a trajectory from south to north. However, a proximate source of borrowing cannot be identified in any known language. Both words are also present in Slavic, but the reconstruction of a Balto-Slavic prototype is impossible, implying the Balts and Slavs must have been in contact with distinct Central European agricultural groups carrying related words.

Evidence of the linguistic landscape in north-eastern Europe is practically non-existent until the Late Middle Ages, so that the existence of unrecorded languages during the first millennium CE which later went extinct need not surprise us. However, since we are clearly dealing with *Wanderwörter*, even if the terms are originally of non-Indo-European origin, it cannot be ruled out that they were transmitted into Balto-Slavic through unattested Indo-European languages. This is imaginable in cases such as 'turnip', where the Baltic term is historically identical to the equivalent in Germanic and Latin. On the other

18 For the Baltic word for 'lentil', see the discussion on pp. 201–202.

19 A derivative of the word for 'turnip' is also found early in Greek, but in a secondary meaning.

hand, little can be said with certainty; neither can it be established that the languages with which Baltic and Slavic were in contact were related with each other, despite possessing similar words for crops.

I have also included the fruit trees apple and pear in this subsection, although they might be better described as wild. Both were first domesticated after the dispersal of the Indo-Europeans, as the cultivation of these plants must be done through grafting rather than from seed (Mallory/Adams 1997: 26; Zohary/Hopf 2012: 138, 140). The distribution of the crab apple and wild European pear is similar, encompassing most of Europe, and the western half of the Pontic-Caspian steppe (see Zohary/Hopf 2012: 137, 139). The pear is not found north of Latvia (cf. Schrader/Nehring 1: 147), and as a consequence, there is no old word for 'pear' in Finnic. Both plants can be and are consumed in their wild form.

A possible candidate for an inherited word for 'apple' is Gr. μῆλον, which has convincingly been argued to be cognate to Hittite *samlu-* 'apple' (Kroonen 2016). If this originally referred to the wild apple, then the spread of the Greek word into Lat. *mālum* and Alb. *mollë* (Schrader/Nehring 1: 53) might be associated with the emergence of domesticated varieties in the early historical period. However, it is difficult to rule out a post-PIE loanword.[20] For pear, we have no comparisons which go beyond two neighbouring branches, and no inherited term can be reconstructed with confidence, although it is theoretically possible that Gr. ἄπιον and Lat. *pirum* 'pear' could reflect an inherited **h*$_2$*pis-o-*. In this case, the term would originally refer to a wild variety and only secondarily to the cultivated pear.

8.1.4 *Apiculture*

As words for 'honey' and 'mead' can be reconstructed for Proto-Indo-European, it has been assumed that PIE speakers must have been involved in apiculture (Гамкрелидзе/Иванов 1984: 603); however, since wild honey hunting has been practiced since the Mesolithic, there is no necessity to believe the Indo-Europeans were familiar with domesticated honey bees (Schrader/Nehring 1: 139–140; van Sluis 2022: 4, 26; cf. Crane 1999: 162). The complete absence of beeswax residues on pottery in the Neolithic Eurasian Steppe, despite good conditions for its preservation, probably speaks against any active apiculture (Roffet-Salque et al. 2015: 229). Three words have been classed as probable loanwords in this semantic field (see Table 22, overleaf).

20 If Kroonen's comparison (2016: 88–89) with Georgian *msxali* 'pear' is valid, then the loanword would have to be very early, predating the loss of the laryngeals.

TABLE 22 Distribution of borrowed apicultural terms

	B	S	G	C	It	Gr	+
drone	✓	✓	✓			✓	
honeycomb	✓					✓	Md Ma Tur
wax	✓	✓	✓				

The borrowing of terms for bees along with the technology for their domestication would be unsurprising, as bees may have been a mere pest to honey hunters, and therefore of less importance (Vennemann 1998: 477–478). Interestingly, however, in an actual case of language shift studied by Brenzinger (1992), we find the opposite situation: after shifting to speak Maasai, originally Yaaku beekeeping communities continued to use a substrate word for 'honey', while words for various kinds of 'bee' had recently fallen out of use (idem: 234–235). This of course need not worry us too much, as we cannot expect all cases of language shift to be identical.

Among the apicultural terms, only the words for 'wax' and 'drone' can potentially be reconstructed for Proto-Balto-Slavic, both of which show good evidence of foreign origin. The terms for 'bee', and also perhaps 'beehive' (see pp. 279–280), which have been considered uncertain pre-European loanwords, cannot be reliably reconstructed for Balto-Slavic. As a result, it is uncertain whether speakers of Proto-Balto-Slavic were engaged in apiculture. Tree beekeeping is already attested in Latvia in the Middle Ages, and was only completely superseded in the East Baltic region by (log) hive beekeeping in the 18th century (Crane 1999: 132–133, 233–234). A potentially Proto-Balto-Slavic term related to the use of tree hives is Lt. dial. *genỹs*, *geinỹs*, Lv. *dzeĩnis*, dial. *dzenis* 'climbing rope (for accessing tree hives)' (apparently → the Võro hapax *kõno* in the same sense; Vaba 1990b: 173) which corresponds regularly to R dial. *жень* (Nižnij Novgorod, Kostroma; СРНГ IX: 129), Bel./Uk. (Polesia) *жэнь*, *жêнь*, *жинь* (ДАБМ No. 313; ЕСУМ II: 193; Никончук *apud* ЭСБМ III: 270) 'climbing rope' (Būga 1916: 156).[21]

21 ЕСУМ (II: 193) suspect that the Slavic word is loaned from Baltic. The distribution would appear to favour this, even though the Russian word is attested rather far from the Baltic territories. It is uncertain whether a Baltic loanword can be expected to have undergone the first palatalization. Note, however, OR ижера *pro* *игера 'Ingrians' ← Ingr. *Inkeroin* cited on p. 18 and the hydronym *Селижáровка* (beside OR Серегѣрь), also of presumed Finnic origin (REW II: 605; Крысько 1994: 83).

8.1.5 *Structures*

Despite a perception of the Indo-Europeans as primarily nomadic (Kuhn 1862: 371; Anthony 2007: 321–322; Anthony/Ringe 2015: 211; see also the literature review in Häusler 2002: 3–48), there is evidence that at least the later stages of Indo-European unity were associated with a level of sedentarization (Kroonen et al. 2022: 32–36), and some clear inherited terminology exists relating to the erection of fixed or semi-fixed structures (Mallory/Adams 2006: 219–229), most notably the verbal root in HLuw. *tama-*, Gr. δέμω 'to build', which is the basis of the nominal derivatives in most branches, e.g. Lycian *tãma* 'building', Arm. *tun*, Lat. *domus* 'house', and probably Lt. *nãmas* 'house' (cf. IEW 198–199).

Insight into the technology of house-building among Indo-European speakers may be provided by Skt. *dehí-*, Osc. feíhúss ACC.PL., Gr. τεῖχος, Sln. *zîd* '(surrounding) wall', apparently derived from the verbal root for 'to mould (clay)'.[22] Mallory/Adams (2006: 223) emphasize that this word does not generally mean 'wall of a house', although Pr. E *seydis* · want 'wall' and Gr. τοῖχος, derived from the same root, do appear to be generic terms.[23] In any case, it is tempting to speculate that this might be a reflection of the construction of temporary wattle and daub huts, as known from ethnographical parallels of nomadic pastoralists (e.g. Evans/Pritchard 1940: 65).

On the other hand, a word for some kind of fortification must be reconstructed on the basis of Lt. *pilìs*, Lv. *pils*, Skt. (RV) *púr-* 'fortress, stronghold', Gr. πόλις 'city, citadel'. Considering the possible association of the arrival of the Balts in the region with the appearance of fortified settlements (Lang 2016: 18, 2018a) and the implication of continuity provided by the linguistic data, it seems attractive to assume the construction of hillforts already started in the Indo-European homeland. Note, for instance, the Early Bronze Age hillfort at Mykhailivka on the Lower Dnieper (Anthony 2007: 324). Whatever the details of Indo-European house construction, it is likely to have greatly differed from that of Neolithic Europe. According to Della Volpe (1996: 152), timber-framed longhouses, generally being devoid of any defensive structures, predominate in the pre-Indo-European context.

---

22 Compare Go. (*ga-*)*digan** (rendering Gr. πλάσσω 'mould, form'), Lat. *fingō* 'mould, fashion (clay, wax, etc.); sculpt', ToB *tsaikaṃ* 'mould (pottery); build', as well as (with apparent metathesis) Lt. *žíẽsti* 'mould (pottery)', OCS съзьдати 'build, create' (IEW 245).

23 Likewise, Mac. *sид* is a generic term for 'wall'. The usual word for 'wall (of a house)' in Balto-Slavic is Lt. *síena*, R *стьна́* (while in Mac., *стена* means 'rock face'). As this word is possibly related to Go. *stains* 'stone', it might reflect a shift towards stone architecture (for a discussion of the relationship between these words, see 6.3.2).

TABLE 23 Distribution of borrowed terms for structures

	B	S	G	C	It	Gr	+
cottage	✓		✓	✓	✓		
oven	✓		✓			✓	

In the centre of an Indo-European home, there was presumably a hearth (Hitt. *hāssā-* 'hearth, fireplace', Lat. *āra* 'altar'; IEW 68–69). On the other hand, ovens are considered to have spread into Europe as part of the Neolithic package, emerging during the eighth millennium BCE in Anatolia (Barbaro et al. 2021: 1161). Domed clay ovens are known from households in Neolithic sites immediately adjacent to Yamnaya (Anthony 2007: 143, 166), and it seems quite probable that a word for 'oven' would have been taken over from such farming populations.

8.1.6 *Metallurgy*

The only metal term in Balto-Slavic with direct Indo-European cognates is the word for 'gold', Lt. *áuksas*, Pr. E *ausis* (III *ausin* ACC.SG.) = Lat. *aurum* 'gold'. While the narrow distribution has led to speculations of a direct or indirect loan relationship (Kretschmer 1896: 150; Pisani 1968: 11), on formal grounds, a common inheritance cannot be excluded (see Driessen 2003).[24] The following terms can theoretically be dated to Proto-Balto-Slavic, two implying ablaut (for a discussion of the further etymologies of the metal names, which remain uncertain, see Thorsø et al. 2023: 117):

– Lt. *švìnas*, Lv. *svins* 'lead' (< **ḱuin-*) ~ OR свиньць, Sln. *svínəc* 'lead' (< **ḱuein-*)
– Lt. obs. *álvas* (*álwu* INST.SG. in Daukša), Lv. *alva*, dial. *alvs* 'tin' (cf. Endzelīns 1923: 157) (< **HolH-u̯-*) ~ OCS олово, SCr. *ȍlovo* 'lead', R *óлово* 'tin' (< **HolH-eu̯-*)
– ? Pr. E *wutris* 'smith' ~ CS вътрь 'smith' (see Miklosich 1865: 113; SJS I: 352)

On the other hand, the terms for two other metals do not permit the reconstruction of a common Balto-Slavic preform, and these may be interpreted as loanwords from unknown sources (see Table 24, overleaf).

24 I am rather convinced by the interpretation of To. B *yasa*, A *wäs* 'gold' as a loanword from Samoyed **wäsa* (> Ngan. *basa* 'metal, iron', Taz Selkup *kȩ̄sį* 'iron'; cf. Kallio 2004: 132–133). In any case, connecting the Tocharian with the European terms raises serious morphological issues (see Thorsø et al. 2023: 105–106).

TABLE 24 Distribution of borrowed metallurgical terms

	B	S	G	C	It	Gr	+
iron	✓	✓					
silver	✓	✓	✓	?			Basque

The absence of a common Balto-Slavic term for 'iron' is hardly surprising, as the split of this branch undoubtedly predated the Iron Age. While some iron artefacts may have been imported into the East Baltic region from elsewhere already in the Final Bronze Age (Lang 2007: 121), local iron production probably began during the first centuries CE, where it was produced in smelting furnaces from bog ores (Stankus 2001; Rundberget et al. 2020: 96).[25] The Slavic word was evidently adopted from a related source, pointing to the spread of a localized smelting practice.

An Indo-European word for 'silver' can be reconstructed on the basis of YAv. *ərəzata-*, Lat. *argentum*, OIr. *argat*, and probably Arm. *arcat'* 'silver', but this word appears to have been replaced in the northern European branches. The word for 'silver' in Balto-Slavic and Germanic is a widespread *Wanderwort*, whose centre of spread might be located in Iberia (Thorsø et al. 2023: 118), an idea that would be supported by the comparanda in Basque and Celtiberian (the latter probably being adopted locally after the southward migration of Celtic speakers). Although the word seems to be reconstructible to Proto-Germanic, it cannot be reconstructed for Proto-Balto-Slavic, or even Proto-Baltic, again suggesting that the word was absorbed into already diffuse linguistic groups.

8.2 Stratification

The main methodological novelty in this section has been an attempt to identify alternations which do not merely reoccur, but which show a particular geographical patterning. I reasoned that a geographical distribution would both support the validity of an alternation, and potentially provide us with some information on the dialectal makeup of the underlying substratum. In

25 Although A. Merkevičius *apud* Lang 2018b dates the appearance of iron metallurgy in Lithuania to 300 BCE.

TABLE 25 Alternations showing a geographical patterning

	B	S	G	C	It	Gr	Examples
*-*VNT*- ∞ *-*VT*-							5
*$g^{(h)}$ ∞ *k							7
*$b^{(h)}$ ∞ *p							2 [+ 2]
*T ∞ *T^h							3 [+ 1]
*$\check{z}$ ∞ *$\check{s}$					?		1 [+ 3]
*st ∞ *$(t)s$							3 [+ 1]
*sd ∞ *d							2
*a-CC ∞ *CVC							1 + 2
*ke- ∞ *ka/o-							3
*$\check{a}$ ∞ *$\check{e}$							3 [+ 2]
*$\bar{V}$ ∞ *V							4 [+3]
*V ∞ *$\bar{V}$							2

total, I have identified seven consonantal and five vocalic alternations which can be said to show a geographical distribution on the basis of at least three (or two certain) examples. This is presented in Table 25, above.

Dark shaded cells consistently show the rightmost variant, while light shaded cells indicate a hesitation between the two. The shading is based on both certain and uncertain examples (the number of the latter is indicated in brackets). In the case of *a-CC ∞ *CVC, I have also included Schrijver's examples of 'blackbird' and 'ore' (see 7.1.1), which seem to show a related phenomenon, even though they have not fallen under the scope of this work.

Previous studies have often tended to treat the palaeo-European contact languages as a monolithic layer, whereby the irregularities present in the Indo-European reflexes are reflections of synchronic features of a single substrate language (Kuiper 1968; Schrijver 1997; Beekes 2014; see p. 166). The presence of geographical patterns contradicts this assumption, as such distributions are more easily explained as the result of dialectal or diachronic differences in the source language. In any case, it seems highly improbable that the linguistic landscape was homogenous among sedentary Neolithic farming populations prior to the expansion of Indo-European (see the discussion in Anthony 2007: 80–81).

A deeper analysis of the stratification based on distribution alone is very difficult, as none of the alternations obviously correlate with each other. An

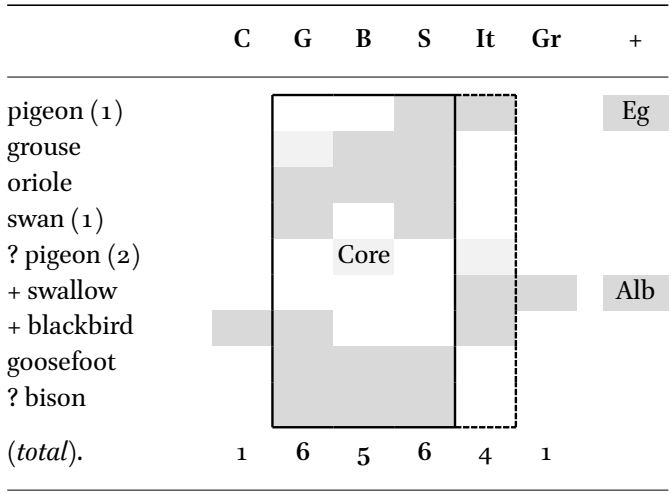
exception is *$g^{(h)}$ ∞ *k and *$b^{(h)}$ ∞ *p, but this actually results from the fact that the two alternations co-occur in two of the relevant word families, and we thereore cannot speak of the coherence of two independent sets. The fact that no clear patterns emerge on this higher level need not dishearten us. On the one hand, the number of examples of each alternation is small, and there is perhaps simply insufficient evidence for meaningful patterns to emerge. On the other hand, these alternations represent manifestations of complex contact situations which may have taken place in different locations and at different times, and therefore a complex picture is exactly what we should expect.

It is perhaps more instructive to examine which kinds of alternations co-occur (cf. Šorgo 2020: 461–462). The word for 'pigeon (1)', for instance, shows both *$g^{(h)}$ ∞ *k and *-*VNT*- ∞ *-*VT*-. This might well suggest that the two alternations are somehow related. Indeed, on p. 185 (and in Jakob forthc. a.), I have noted that the word for 'pigeon' shows a similar structure to several other bird names, including another word for 'pigeon (2)' which potentially shows the alternation *$b^{(h)}$ ∞ *p. The full set of words (including one plant name) is as follows:

'pigeon (1)'	OCS голѫбь, OE *culufre*, Lat. *columba*
'grouse'	Lt. *jerubẽ* 'hazel grouse', Sln. *jerẽb*, ?OHG *reba-huon* 'partridge'
'oriole'	Lt. *volungẽ*, Pl. *wilga*, MHG *wite-wal*
'swan (1)'	Pl. *łabędź*, R *лебедь*, ON *ǫlpt*
'goosefoot'	Lt. *balánda*, R *лобода́* 'goosefoot', OHG *melta* 'orache'
+ 'swallow'	Lat. *hirundō*, Gr. χελῑδών, Alb. *dallëndyshe*
? 'pigeon (2)'	Lt. *balañdis*, Lat. *palumbēs*

This collection of terms is perhaps the strongest evidence for a particular stratum: as well as clustering in a particular semantic field, they show similar kinds of alternations, in particular, a semi-regular correlation between voiced stops in the north (always Baltic, usually Slavic) and voiceless in the south (i.e. Italic), and a second syllable of the shape **VND*, whereby the nasal is sometimes absent (although always present in Italic). In addition, I have noted 'swan (1)' as a plausible example of the alternation **a-CC* ∞ **CVC*. This might encourage us to view this alternation as yet another feature of this stratum. Indeed, the classic example of this alternation is another bird name (cf. Lat. *merula*, OHG *amsla* 'blackbird'). Aside from this, potential 'prefixal' elements have been identified in the words for 'grouse' and 'oriole'. Finally, it is tempting to adduce the word for 'bison' here, as OHG *wisunt* shows a similar disyllabic root structure with a second syllable in **VND*, although here the initial syllable appears itself to be a 'prefixal' element.

TABLE 26 The **VND* substrate



Viewing the stratum as a whole (see Table 26, above), we can see a Central European 'core', consisting of Balto-Slavic, Germanic and (slightly less so) Italic, and a periphery. Notably, all of the words attested in the peripheral languages are also attested in Italic, and indeed Italy can be seen as a sort of interface between Central Europe and the Mediterranean on the one hand, and with Celtic on the other. It is certainly not the case, however, that the words were borrowed into the 'peripheral' branches directly from Latin or an Italic language. If words belonging to this stratum are not originally Mediterranean, they must have been carried into the region by speakers of unattested, presumably non-Indo-European languages. This implies a significant antiquity, which is already suggested by the attestation of 'pigeon (1)' in Egyptian in the 12th c. BCE (see p. 189).

On the other hand, the words 'swan (1)' and 'oriole' show irregular variation even within Slavic, suggesting that at least one variant was adopted after the dialectal fragmentation of this branch. This places us in a very broad timeframe stretching some two millennia, and raises serious doubts as to the internal coherence of the stratum. One suggestion, borrowing the analytical tools of botany, would be to interpret Slavic as a "centre of diversity", and suggest that Slavic was geographically closest to the 'core'. This is potentially supported by the fact that Slavic takes an intermediate position in the voicing alternations, implying contacts with multiple source languages or dialects.

Considering the limitation to animal and plant names, and in particular bird names, we are most probably dealing with a largely vertical borrowing context; in other words, a linguistic substrate. By contrast, a number of tech-

nological borrowings, particularly those showing a broad distribution, were probably borrowed horizontally through trade, and can be characterized as *Wanderwörter*. Here we may include most of the terms for crops and cultivated plants (see 8.1.3), as well as the word for 'silver' and probably 'thousand' (see 3.5.4).

It is probable that other (sub)strata existed. If we consider the alternations which do not occur in any of the words in the 'bird name' stratum, it is curious that **ke-* ∞ **ka-* typically involves a "non-core" distribution: of the three examples, two involve Celtic, and two involve Greek. The word for 'ramsons' shows a particularly broad distribution. Somewhat comparable is the alternation **T* ∞ **T^h*, which always (by necessity) involves Greek. The words in these categories seem to cluster semantically in the domain of wild and cultivated plants. At least 'drone' can be reconstructed for Proto-Balto-Slavic, suggesting a certain antiquity.[26]

The nasal alternation in the word for 'lynx' is also unlike that attested in the 'bird names'. Above, I have briefly mentioned that this word could be a particularly old loanword. There are multiple possible indications of this:

– The word shows an unusually large distribution, being present in five Indo-European branches.
– It can possibly be reconstructed for Proto-Balto-Slavic,[27] as well as for Proto-Graeco-Armenian.
– Importantly, the sibilant in Balto-Slavic implies that these borrowings predated satemization.

Aside from the word for 'lynx' only half a dozen words can be securely reconstructed for Proto-Balto-Slavic. The following may be mentioned:

– In three branches: 'apple', 'lynx', 'nettle'
– East Baltic + Slavic: 'alder', 'drone', 'fresh'[28]
– Prussian + Slavic: ?'bean' (if not a Slavic loan in Prussian)

While the possibility of reconstructing a word for Proto-Balto-Slavic may be seen as implying its relative antiquity, it is not a watertight indication: 'rye' and 'hemp', which would theoretically be reconstructible for Proto-Balto-Slavic, must have entered Baltic recently in view of the cultivation history of these

26 A shared reconstruction might also be attempted for 'ramsons', although this requires that Baltic **š* results from the RUKI law, which I consider dubious (Chapter 1, fn. 14). A similar obstacle exists to the reconstruction of the words for 'wax' and 'reed' for Proto-Balto-Slavic (on the latter, see the note under 'furrow' on p. 224).

27 The only obstacle to this is the Slavic **r-*, on which see Chapter 6, fn. 18.

28 More dubiously, we may be able to reconstruct words for 'aspen' (provided Baltic **u* is not old; cf. pp. 277–278), 'hornbeam' (irregular Baltic *s-*) and 'lightning' (the dental in Slavic and East Baltic is ambiguous).

plants. Nevertheless, the six certain examples cited above appear to be good candidates for Proto-Balto-Slavic loanwords. Support for the antiquity of the word for 'apple' may be seen in its adoption into a fairly unproductive noun class (the *l*-stems). Depending on one's analysis (see pp. 268–270 for a discussion), this word — like 'cottage' — may additionally be interpreted as predating Winter's law, which would certainly imply a Proto-Balto-Slavic antiquity.[29]

If the East Baltic comparanda for 'bean' are accepted (see p. 228), then the word can no longer be reliably reconstructed for Proto-Balto-Slavic. In fact, it would point to the opposite extreme: a word which is even irregular between East and West Baltic. A few other such words can be cited which exhibit similar behaviour: among the certain cases, we can mention 'bison' and 'thousand', both of which exhibit the alternation **st* ∞ *(*t*)*s*, and also 'silver', a widespread *Wanderwort*. Less certain examples are 'badger', 'salmon', and also 'oats', depending on the analysis of the Prussian data (pp. 239–240).

To summarize, it is clear, at least, that we are not dealing here with a chronologically or geographically localized borrowing event; however, due to the number of variables and small number of examples, it is difficult to comprehensively stratify the material. Nevertheless, there are indications of at least three chronological layers — one early layer, exemplified by the word for 'lynx', which may represent a borrowing event close to the steppe chronologically aligned with the disintegration of the proto-language, a late layer, apparently postdating the split of East and West Baltic (providing a Proto-Baltic stage ought to be reconstructed at all), and an intermediate layer. In addition, one group of words, primarily comprising bird names, seems to form a robust cluster and perhaps represents a set of loanwords from related source languages.

29 Other candidates for loanwords predating satemization, and therefore potentially contemporaneous with the word for 'lynx', are the uncertain cases 'elm (1)' and 'furrow'. It must be noted, however, that the reconstruction of the word for 'elm' to Proto-Balto-Slavic is not entirely straightforward, as much of the Slavic evidence speaks in favour of accentual mobility (cf. В. Дыбо 2002: 469), and the word for 'furrow' is irregular between Baltic and Slavic (see p. 224).

Conclusion

One of the first goals set out in this monograph was to scrutinize the stereotype of the Baltic languages as 'pure' Indo-European languages which have experienced minimal external influences. Through a detailed study of all possible contact events, both with attested and unattested languages, I have attempted to demonstrate that a significant part of the Baltic lexicon cannot be explained on internal grounds. However, most interactions which have contributed to the Baltic lexicon appear to have been with unknown languages, by contrast, for instance, to Finnic, which is known to have been in contact with several Indo-European languages throughout its history.

No Indo-European loanword layers can be identified with certainty in Baltic prior to those with Gothic at the start of the Common Era. Contacts with Slavic, as far as we can make out, only started after the northern migration of (pre-)North Russian speakers. In addition, a notable layer of Baltic loanwords can be identified in Finnic, suggesting a significant contact event. However, even though the source of these loans seems to be more closely affiliated to East Baltic than West Baltic, and there is evidence that the source language has undergone some specifically East Baltic semantic and formal developments, it remains improbable that this was the direct ancestor of the attested East Baltic languages. There does not appear to be any old Finnic contribution to the Baltic lexicon, and the evidence seems to support the notion of an East Baltic substrate, most probably spoken to the east of the current Baltic territories, which was absorbed by Finnic some time before the Common Era. In addition, we see small layers of loanwords in both Sámi and Mordvin, suggesting some peripheral contact with this or a closely related Baltic language.

We have approached the question of non-Indo-European components in the Baltic lexicon from multiple angles. First, we have attempted to find words common to both Baltic and Finnic which are unlikely on phonological grounds to have been adopted directly from one to the other. Although there are few relatively clear cases, there are a number of convincing examples which allow for a hypothesis that Baltic and Finnic were independently in contact with similar, and probably also distinct, 'autochthonous' populations upon their arrival in the Baltic Sea region.

As we are able to operate in the context of regular sound correspondences (or more specifically, their absence), it is somewhat easier to make a case for a non-Indo-European element in Baltic lexical items with Indo-European comparanda. Almost fifty relatively clear cases were identified. Some initial efforts can be made to stratify this material, and at least one relatively coherent and

distinct stratum can be identified, consisting primarily of bird names with a second syllable of the shape **V(N)D*. In addition to this, we can distinguish a number of widespread *Wanderwörter*, most significantly in the domain of cultivated plants, whose proximate source in the Baltic languages cannot be identified in any known language, and which may be reasoned at least partly to have originated among Neolithic farming populations.

Contact with unattested languages is an area of study which has long been marginalized, partly for the reason that it is considered impossible to study, *a priori* unscientific, or inevitably circular. As a result of these prejudices, this subfield remains in its infancy. One of the goals of this work has been to demonstrate the feasibility of constraining the study of such contacts within scientific principles. It is important that a 'substrate' hypothesis is treated as strictly and objectively as a hypothesis of cognancy, and built on the basis of positive evidence. In this context, a suggestion of non-Indo-European origin can be viewed as a reasoned scientific solution to the problem of irregular sound correspondence, and not as a last resort or throwaway suggestion.

It is certainly not true that the Baltic languages have developed in a vacuum, void of contact with other languages. Instead, most of the language contact has taken place in a preliterary context, with languages which never came to be written down, or of which no written trace has yet been uncovered. This is actually precisely what we should expect, since the area where the Balts have come to reside has been populated since the end of the last Ice Age by numerous genetically distinct populations, undoubtedly bringing with them different languages, while writing has only reached the region in the Middle Ages. In this context, traces of foreign languages preserved in the modern Baltic languages can be seen as a valuable resource and a potential key to unlocking the population history of the region.

Bibliography

Note on alphabetization: titles and names written in the Cyrillic script are presented in the original spelling, but alphabetized according to the usual academic transcription into the Latin alphabet:

а б в г д е ж з и й к л м н о п р с т у ф х ц ч ш щ ъ ы ь э ю я ѣ
a b v g d e ž z i j k l m n o p r s t u f x c č š šč ” y ’ e ju ja ě

For the purposes of alphabetization, diacritics are ignored, so that *Çabej* precedes *Цаболов* (= *Cabolov*). The Cyrillic letter *х* is alphabetized under *x* (rather than *h*, as sometimes found). For ease of navigation, Ukrainian/Belarusian *ґ* have been alphabetized under *g*.

Web links are valid as of 9 November 2023.

Abbreviated Titles

ALEW — Wolfgang Hock (2019, ed.), *Altlitauisches etymologisches Wörterbuch* [Version 1.1]. Berlin: Humboldt-Universität.

ALEW² — Wolfgang Hock (2020, ed.), *Altlitauisches etymologisches Wörterbuch* [Version 2.0]. Accessed online at https://alew.hu-berlin.de/.

АУМ — Тетяна В. Назарова et al. (1984–2001, eds.), *Атлас української мови*. 3 vols. Київ: Наукова думка.

AWb — Elisabeth Karg-Gasterstädt & Theodor Frings (1952– , eds.) *Althochdeutsches Wörterbuch*. 7 vols. [A–R]. Leipzig: Sächsische Akademie der Wissenschaften.

БДА — Иван Ю. Кочев (2001, ed.), *Български диалектен атлас. Обобщаващ том*. София: Труд.

БЭР — Владимир И. Георгиев, Иван Дуриданов et al. (1971– , eds.), *Български етимологичен речник*. 8 vols. [*а–фѣкалка*]. София: Българската академия на науките/М. Дринов.

CDIAL — Turner, Ralph L. (1973 [1966]). *A comparative dictionary of Indo-Aryan languages*. London: Oxford University Press.

ČJA — Jan Balhar & Pavel Jančák (1992–2011, eds.), *Český jazykový atlas* [2nd edition]. 5 vols. Praha: Academia. Accessed online at https://cja.ujc.cas.cz/.

ClG I — Adolfas Ivaškevičius et al. (1995–1997, eds.), *Clavis Germanico-Lithvana*. Vols. 1–2. Vilnius: Mokslo ir enciklopedijų leidykla.

ДАБМ — Рубэн І. Аванесаў, Кандрат Крапіва & Юзэфа Ф. Мацкевіч (1963,

eds.), *Дыялекталагічны атлас беларускай мовы*. Мінск: Акадэмія навук БССР.

ДАРЯ I Рубен И. Аванесов & Софья В. Бромлей (1986, eds.), *Диалектологический атлас русского языка* I. *Фонетика*. Москва: Наука.

DOE Angus Cameron, Ashley Crandell Amos, Antonette diPaolo Healey et al. (2018, eds.), *Dictionary of Old English*: *A to I*. Accessed online at tapor.library.utoronto.ca/doe (subscription resource).

DWb Moriz Heyne et al. (1854–1971, comp.), *Deutsches Wörterbuch von Jacob Grimm und Wilhelm Grimm*. 33 vols. Leipzig: Hirzel.

DWb² *Deutsches Wörterbuch von Jacob Grimm und Wilhelm Grimm*. *Neubearbeitung* (*A–F*). Accessed online at https://www.woerterbuchnetz.de/?sigle=DWB2&lemid=A00001.

eDIL Gregory Toner, Maire Ní Nhaonaigh et al. (2019, eds.), *Electronic Dictionary of the Irish Language*. Accessed online at dil.ie.

EES Iris Metsmägi, Meeli Sedrik & Sven-Erik Soosaar (2002, eds.), *Eesti etümoloogiasõnaraamat*. Tallinn: Eesti Keele Sihtasutus.

EH Janis Endzelīns & Edita Hausenberg (1956). *Ergänzungen und Berichtigungen zu K. Mühlenbachs Lettisch-deutschem Wörterbuch*. 2 vols. Chicago.

EMS Anu Haak, Evi Juhkam, Mari Kendla et al. (1994– , eds.), *Eesti murrete sõnaraamat*. 7 vols. [*a–päevätuss*]. Tallinn: Eesti Teaduste Akadeemia.

ERHJ Ranko Matasović, Dubravka Ivšić Majić & Tijmen Pronk (2016–2021, eds.), *Etimološki rječnik hrvatskoga jezika*. 2 vols. Zagreb: Institut za hrvatski jezik i jezikoslovlje.

ЭСБМ Віктар У. Мартынаў & Генадзь А. Цыхун (1978– , eds.), *Этымалагічны слоўнік беларускай мовы*. 14 vols. [*а–тэчка*]. Мінск: Навука і тэхніка.

ESJS Eva Havlová et al. (1989–2018, eds.), *Etymologický slovník jazyka staroslověnského*. 19 vols. Praha: Academia; Brno: Tribun EU.

ЭССЯ Олег С. Трубачев & Андрей А. Журавлев (eds., 1971–), *Этимологический словарь славянских языков*. 42 vols. [**a–*perzъ*]. Москва: Наука.

ЭСТЯ Эрванд В. Севортян, Лия С. Левитская et al. (1974–2003, eds.), *Этимологический словарь тюркских языков*. 7 vols. Москва: Наука/Восточная литература РАН.

ЕСУМ Олександр С. Мельничук et al. (1982–2012, eds.), *Етимологічний словник української мови*. 6 vols. Київ: Наукова думка.

EWA Mayrhofer, Manfred (1992–2001). *Etymologisches Wörterbuch des Altindoarischen*. 3 vols. Heidelberg: Winter.

EWAhd Albert L. Lloyd, Otto Springer & Rosemarie Lühr (1988– , eds.), *Etymologisches Wörterbuch des Althochdeutschen*. 7 vols. [*a–skazzôn*]. Göttingen–Zürich: Vandenhoeck & Ruprecht.

GPC Gareth A. Bevan & Patrick J. Donovan (1967–2002, eds.), *Geiriadur Prifysgol Cymru*. 4 vols. Cardiff: Gwasg Prifysgol Cymru. Also 2nd online edition, accessed at geiriadur.ac.uk/gpc/gpc.html.

ГСБМ Аляксандр М. Булыка (1982–2007, ed.), *Гістарычны слоўнік беларускай мовы*. 37 vols. Мінск: Навука і тэхніка/Беларуская навука.

IEW Pokorny, Julius (1959). *Indogermanisches etymologisches Wörterbuch*. Bern–München: Francke.

KEWA Mayrhofer, Manfred (1956–1980). *Kurzgefaßtes etymologisches Wörterbuch des Altindischen*. 4 vols. Heidelberg: Winter.

KKS Pertti Virtaranta & Raija Koponen (1968–2005, eds.), *Karjalan kielen sanakirja*. 6 vols. Helsinki: Suomalais-Ugrilainen Seura. Accessed online at kaino.kotus.fi/kks.

LÄGLOS Andries D. Kylstra, Sirkka-Liisa Hahmo, Tette Hofstra & Osmo Nikkilä (1991–2012, eds.), *Lexikon der Alteren Germanischen Lehnworter in den Ostseefinnischen Sprachen*. 3 vols. Amsterdam: Rodopi.

LEIA Vendryes, Joseph, Édouard Bachallery & Pierre-Yves Lambert (1959–1996). *Lexique étymologique de l'irlandais ancien*. 7 vols. Dublin: Institute for Advanced Studies.

LEW Fraenkel, Ernst (1955–1966). *Litauisches etymologisches Worterbuch*. 2 vols. Heidelberg: Winter/Göttingen: Vandenhoeck & Ruprecht.

LIV Helmut Rix et al. (2001, eds.), *Lexicon der indogermanischen Verben* [2nd edition]. Wiesbaden: Reichert.

LIV addenda Kümmel, Martin (2015). *Addenda und Corrigenda zu LIV²*. Accessed online at martinkuemmel.de/liv2add.pdf.

LKA Elena Grinaveckienė, Kazys Morkūnas et al. (1977–1991, eds.), *Lietuvių kalbos atlasas*. 3 vols. Vilinus: Mokslas.

LKŽ Juozas Balčikonis & Vytautas Vitkauskas (1941–2002, eds.), *Lietuvių kalbos žodynas*. 20 vols. Vilnius: Mokslas, Lietuvių kalbos instituto leidykla.

LSJ Henry J. Liddel, Robert Scott, Henry S. Jones & Roderick McKenzie (1940, eds.), *A Greek–English Lexicon* [9th edition]. Oxford: Clarendon Press.

LVPPV Laimdots Ceplītis, Aina Miķelsone, Tamāra Porīte & Silvija Raģe (1995, eds.), *Latviešu valodas pareizrakstības un pareizrunas vārdnīca*. Riga: Avots.

MdWb Kaino Heikkilä (1990–1996, comp.), *H. Paasonens Mordvinisches Wörterbuch*. 6 vols. Helsinki: Suomalais-Ugrilainen Seura.

ME Janis Endzelīns (1923–1925, ed.), *K. Mühlenbachs Lettisch-deutsches Wörterbuch*. 4 vols. Rīga: Latvju grāmata.

MWb *Mittelhochdeutsches Wörterbuch* (online edition) [*a–kochlêhen*]. Accessed online at http://mhdwb-online.de/.

NdsWb Hans Janßen et al. (1965–). *Niedersächsiches Wörterbuch*. 10 vols. [*a–skrofulös*]. Neumünster–Kiel–Hamburg: Wachholtz.

NIL Dagmar S. Wodtko, Britta Irslinger & Carolin Schneider (2008). *Nomina im Indogermanischen Lexikon*. Heidelberg: Winter.

ODS *Ordbok over det danske Sprog* (online edition). Accessed online at https://ordnet.dk/ods.

OED *Oxford English Dictionary* (online edition). Oxford University Press. Accessed online at https://www.oed.com/ (subscription resource). [N.B. entries updated for the OED Third Edition are marked OED[3]].

OstWb Steinitz, Wolfgang, Liselotte Böhnke, Gert Sauer & Brigitte Schulze (1966–1993). *Dialektologisches und etymologisches Wörterbuch der ostjakischen Sprache*. 15 issues. Berlin: Akademie.

PKEŽ Mažiulis, Vytautas (1988–1997). *Prūsų kalbos etimologijos žodynas*. 4 vols. Vilnius: Mokslas.

PrWb Erhard Riemann & Ulrich Tolksdorfs (1974–2005, eds.), *Preußisches Wörterbuch. Deutsche Mundarten Ost- und Westpreußens*. 6 vols. Neumünster: Wachholtz.

РЭС Аникин, Александр Е. (2007–). *Русский этимологический словарь*. 16 issues [*а–житьё*]. Москва: Рукописные памятники Древней Руси.

REW Vasmer, Max (1953–1958). *Russisches etymologisches Wörterbuch*. 3 vols. Heidelberg: Winter.

RJA Đuro Daničić, Pero Budmani et al. (1880–1976, eds.), *Rječnik hrvatskoga ili srpskoga jezika*. 23 vols. Zagreb: Jugoslavenska akademija znanosti i umjetnosti.

РСА *Речник српскохрватског књижевног и народног језика* (1959–). 21 vols. [*а–покупити*]. Београд: Српска академија наука.

Schw. Id. Friedrich Staub et al. (1881– , eds.), *Schweizerisches Idiotikon*. 17 vols. [*a–zih*]. Frauenfeld: Huber/Basel: Schwabe.

SD Constantinus Szyrwid [= Konstantinas Sirvydas] (1677). *Dictionarium trium lingvarum*. 4th edition. Vilnæ: Societas Iesu. [SD[1] = first edition, digitized online at seniejirastai.lki.lt/db.php?source=20].

СДРЯ Измаил И. Срезневский (1893–1912). *Матеріалы для словаря древне-русскаго языка по письменнымъ памятникамъ*. Санктпетербургъ.

СДРЯ 11–14 Рубен И. Аванесов (vols. 1–5), Игорь С. Улуханов (vol. 6), Вадим Б. Крысько (7–) (1988– , eds.), *Словарь древнерусского языка (XI–XIV вв.)*. 12 vols. [*а–съотъходьнъ*]. Москва: Русский язык.

SEK Boryś, Wiesław & Popowska-Taborska, Hanna (1994–2006). *Słownik etymologiczny Kaszubszczyny*. 6 vols. Warszawa: Slawistyczny Ośrodek Wydawniczy.

СИГТЯ IV Эдхям Р. Тенишев (2001, ed.), *Сравнительно-историческая грамматика тюркских языков*: *Лексика*. Москва: Наука.

СИГТЯ V Эдхям Р. Тенишев (2002, ed.), *Сравнительно-историческая грамматика тюркских языков*: *Региональные реконструкции*. Москва: Наука.

СИГТЯ VI Эдхям Р. Тенишев & Анна В. Дыбо (2006, ed.), *Сравнительно-историческая грамматика тюркских языков*: *Пратюркский язык-основа*. Москва: Наука.

SJS Joseph Kurz (1966–1997, ed.), *Slovník jazyka staroslověnského*. 4 vols. Praha: Academia, Euroslavica. Also corrected online edition at http://gorazd.org/gulliver/.

SKES Yrjö H. Toivonen, Erkki Itkonen & Aulis J. Joki (1955–1981, eds.), *Suomen kielen etymologinen sanakirja*. 7 vols. Helsinki: Suomalais-Ugrilainen Seura.

Сл. Акад. *Словарь Академіи Россійской* (1789–1794). 6 vols. Санктпетербургъ: Императорская Академія Наукъ.

Sł. Warsz. Jan Karłowicz et al. (1898–1927, eds.), *Słownik języka polskiego*. Warszawa.

SMS *Suomen murteiden sanakirja* (online edition). Accessed at kaino.kotus.fi/sms.

СРГК Александр С. Герд. (1994–2002, ed.), *Словарь русских говоров Карелии и сопредельных областей*. 6 vols. Санкт-Петербург: Санкт-Петербургский государственный университет.

СРГС Александр И. Федоров (1999–2006, ed.), *Словарь русских говоров Сибири*. Новосибирск: Наука.

СРЯ 11–17 Степан Г. Бархударов, Галина А. Богатова et al. (1975– , eds.), *Словарь русского языка XI–XVII* вв. 31 vols. [*а–улокъ*]. Москва: Наука.

СРНГ Федот П. Филин (vols. 1–23), Федор П. Сороколетов (vols. 24–45), Сергей А. Мызников (46–) (1966– , eds.), *Словарь русских народных говоров*. 52 vols. [*а–цванки*]. Ленинград–Санкт-Петербург: Наука.

SSA Erkki Itkonen, Ulla-Maija Kulonen et al. (1992–2001, eds.), *Suomen sanojen alkuperä*. 3 vols. Helsinki: Suomalaisen Kirjallisuuden Seura.

SSKJ^2 *Slovar slovenskega knjižnega jezika* [2nd edition] (2014). 2 vols. Ljubljana: Inštitut za slovenski jezik.

SSN Ivor Ripka & Adriana Ferenčíková (1994– , eds.), *Slovník slovenských nárečí*. 3 vols [*a–r*]. Bratislava: Veda.

SSP Stanisław Urbańczyk et al. (1953–1995, eds.), *Słownik staropolski*. 11 vols. Warszawa–Wrocław: Ossolineum.

TLL *Thesaurus linguae latinae* (1900–). Leipzig: Teubner/Berlin: De Gruyter.

TPC *Татарско-русский словарь* (2007). 2 vols. Казань: Магариф.
UEW Rédei, Károly (1986–1992). *Uralisches etymologisches Wörterbuch*. 3 vols. Wiesbaden: Harrassowitz.
VKS Elna Adler, Silja Grünberg & Merle Leppik (2012, eds.), *Vadja keele sõnaraamat*. Tallinn: Eesti Keele Sihtasutus.
VMS Valdek Pall (1989, ed.), *Väike murdesõnastik* (online edition). Accessed online at https://eki.ee/dict/vms/.

List by Author

Абаев, Василий И. (1949). *Осетинский язык и фольклор* I. Москва: Наука.
Абаев, Василий И. (1958). *Историко-этимологический словарь осетинского языка* I. *A-K'*. Москва–Ленинград: Наука.
Абаев, Василий И. (1965). *Скифо-европейские изоглоссы: на стыке Востока и Запада*. Москва: Главная редакция восточной литературы.
Абаев, Василий И. (1973). *Историко-этимологический словарь осетинского языка* II. *L-R*. Ленинград: Наука.
Абаев, Василий И. [= V.I. Abaev] (1996). Corrections and additions to the Ossetic etymological dictionary. In: Hans H. Hock (ed.), *Historical, Indo-European, and Lexicographical Studies* [FS Zgusta], Berlin–New York: De Gruyter, 197–220.
Adams, Douglas Q. (1985). The Indo-European Word for 'apple' Again. *Indogermanische Forschungen* 90, 79–82.
Adams, Douglas Q. (2013). *A Dictionary of Tocharian B: Revised and Greatly Enlarged*. Amsterdam–New York: Rodopi.
Ahlqvist, August (1871). *De vestfinska språkens kulturord*. Helsinki: Finska litteratursällskapet.
Aikio, Ante (2002). New and Old Samoyed Etymologies. *Finnisch-Ugrische Forschungen* 57, 9–57.
Aikio, Ante (2004). An essay on substrate studies and the origin of Saami. In: Irma Hyvärinen, Petri Kallio & Jarmo Korhonen (eds.), *Etymologie, Entlehnungen und Entwicklungen* [FS Koivulehto]. Helsinki: Société Néophilologique, 5–34.
Aikio, Ante (2006a). On Germanic-Saami contacts and Saami prehistory. *Suomalais-Ugrilaisen Seuran Aikakauskirja* 91, 9–55.
Aikio, Ante (2006b). Etymological Nativization of Loanwords: a Case Study of Saami and Finnish. In: Ida Toivonen & Diane Nelson (eds.), *Saami Linguistics*. Amsterdam–Philadelphia: John Benjamins, 17–52.
Aikio, Ante (2009). *The Saami Loanwords in Finnish and Karelian*. Diss. University of Oulu.
Aikio, Ante (2012a). An essay on Saami ethnolinguistic prehistory. In: Riho Grünthal

& Petri Kallio (eds.), *A Linguistic Map of Prehistoric Northern Europe*. Helsinki: Suomalais-Ugrilainen Seura, 63–117.

Aikio, Ante (2012b). On Finnic long vowels, Samoyed vowel sequences, and Proto-Uralic **x*. In: Tiina Hyytiäinen, Lotta Jalava, Janne Saarikivi & Erika Sandman (eds.), *Per urales ad Orientem: iter polyphonicum multilingue* [FS Janhunen]. Helsinki: Suomalais-Ugrilainen Seura, 227–250.

Aikio, Ante (2014a) On the reconstruction of Proto-Mari vocalism. *Journal of Language Relationship* 11, 125–157.

Aikio, Ante (2014b). Studies in Uralic Etymology II: Finnic Etymologies. *Linguistica Uralica* 50, 1–19.

Aikio, Ante (2014c). Studies in Uralic Etymology III: Mari Etymologies. *Linguistica Uralica* 50, 81–93.

Aikio, Ante (2015a). The Finnic 'secondary e-stems' and Proto-Uralic vocalism. *Suomalais-Ugrilaisen Seuran Aikakauskirja* 95, 25–66.

Aikio, Ante (2015b). Studies in Uralic Etymology IV: Ob-Ugric Etymologies. *Linguistica Uralica* 51, 1–20.

Aikio, Ante (2021). Studies in Uralic Etymology V: Permic Etymologies. *Linguistica Uralica* 57, 161–179.

Aikio, Ante (2022). *Proto-Uralic*. In: Marianne Bakró-Nagy, Johanna Laakso & Elena Skribnik (eds.), *The Oxford Guide to the Uralic Languages*. Oxford: University Press, 3–27.

Aikio, Ante (in prep.). *Uralic Etymological Dictionary* (*draft version of entries A–Ć*). Draft paper published on academia.edu (17 January 2020).

Aitzetmüller, Rudolf (1958). *Das Hexaemeron des Exarchen Johannes* I. Graz: Akademie.

Alatalo, Jarmo [= Ярмо Алатало] (1998). *Сӱссыӷӱй ӛӂипсан. Ӄӛтӄый ӄӱланни*. Максимкин Яр–Helsinki: Национальный клуб максимоярочка.

Alatalo, Jarmo (2004, ed.). *Sölkupisches Wörterbuch*. Helsinki: Finnisch-Ugrische Gesellschaft.

Alessio, Giovanni (1946). Un' oasi linguistica preindoeuropea nella regione baltica? *Studi Etruschi* 19, 141–176.

Allen, James (2020). *Ancient Egyptian Phonology*. Cambridge: University Press.

Allentoft, Morten E. et al. (forthc.). *Population genomics of Stone Age Eurasia*. Preprint version posted on bioRxiv, 5 May 2022. doi.org/10.1101/2022.05.04.490594.

Аллмяэ, Райли, Маре Аун, & Лиина Малдре (2018). Предварительные результаты изучения остеологического материала курганных могильников Рысна-Сааре I и II в Северной Сетумаа (Юго-Восточная Эстония). *Археология и история Пскова и Псковской земли. Материалы LII заседания*. Псков, 298–310.

Alminauskis, Kazimieras (1934). *Die Germanismen des Litauischen* I. *Die deutschen Lehnwörter im Litauischen*. Diss. Universität Leipzig.

Ambrazas, Saulius (1993). On the development of diminutives in the Baltic languages. *Linguistica Baltica* 2, 47–67.

Ambrazas, Saulius (2000). *Daiktavardžių darybos raida* II. *Lietuvių kalbos vardažodiniai vediniai*. Vilnius: Mokslo ir enciklopedijų leidybos institutas.

Ambrosiani, Björn (2005). Birka and Scandinavia's trade with the East. *Russian History* 32, 287–296.

van Amerongen, Yvonne (2016). *Wild West Frisia*. Diss. Universiteit Leiden.

Andersen, Henning (1996a). *Reconstructing Prehistorical Dialects: Initial Vowels in Slavic and Baltic*. Berlin: Mouton de Gruyter.

Andersen, Henning (1996b). Взгляд на славянскую прародину: доисторические изменения в экологии и культуре. *Вопросы языкознания* 1996/5, 65–106.

Andersen, Henning (2003). Slavic and the Indo-European migrations. In: Henning Andersen (ed.), *Language Contacts in Prehistory: Studies in Stratigraphy*. Amsterdam–Philadelphia: Benjamins, 45–76.

Anderson, Walter (1938). Zum Lehnwort im Baltendeutsch. *Zeitschrift für Mundartforschung* 14, 146–149.

Аникин, Александр (1992). [Review of] G. Holzer. Entlehnungen aus einer bisher unbekannten indogermanischen Sprache im Urslavischen und Urbaltischen. *Вопросы языкознания* 1992/5, 169–173.

Аникин, Александр (1998). *Этимология и балто-славянское лексическое сравнение в праславянской лексикографии. Материалы для балто-славянского словаря* I. [**a–*go-*]. Новосибирск: Сибирский хронограф.

Аникин, Александр (2000). *Этимологический словарь русских диалектов Сибири* [2nd edition]. Москва–Новосибирск: Наука.

Аникин, Александр (2003). *Этимологический словарь русских заимствований в языках Сибири*. Новосибирск: Наука.

Аникин, Александр (2004). [Review of] Słownik prasłowiański. Tom XIII. Goda — Gyža. *Балто-славянские исследования* 16, 369–386.

Аникин, Александр (2005). *Опыт словаря лексических балтизмов в русском языке*. Новосибирск: Наука.

Аникин, Александр (2014). Проблемы изучения балтизмов в русском языке. *Siberian Journal of Philology* 2014/4, 189–195.

[Аникин РЭС — see РЭС under Abbreviated Titles.]

Anthony, David W. (2007). *The Horse, the Wheel and Language*. Princeton–Oxford: Princeton University Press.

Anthony, David W. & Ringe, Don (2015). The Indo-European Homeland from Linguistic and Archaeological Perspectives. *Annual Review of Linguistics* 1, 199–219.

Anthony, David W. et al. (2005). The Samara Valley Project: Late Bronze Age Economy and Ritual in the Russian Steppes. *Eurasia Antiqua* 11, 395–417.

Ariste, Paul (1962). Mõnda substraadist. *Keel ja Kirjandus* 5, 13–17.

Ariste, Paul (1971). The Earliest Substrata in the Baltic-Finnic Languages. *Néprajz és Nyelvtudomány* 15, 5–11.

Ariste, Paul (1975). Substraatseid kalanimetusi. *Keel ja Kirjandus* 18, 467–472.

Arumaa, Peeter (1930). *Litauische mundartliche Texte aus der Wilnaer Gegend*. Dorpat: Universitas Tartuensis.

Arumaa, Peeter (1964). *Urslavische Grammatik* I. *Einleitung*. *Lautlehre*. Heidelberg: Winter.

Arumaa, Peeter (1976). *Urslavische Grammatik* II. *Konsonantismus*. Heidelberg: Winter.

Ахметьянов, Рифкат Г. (1981). *Общая лексика духовной культуры народов Среднего Поволжья*. Москва: Наука.

Ахметьянов, Рифкат Г. (2015). *Этимологический словарь татарского языка*. 2 vols. Казань: Магариф–Вакыт.

Babel, Anna M. (2016). Affective motivations for borrowing: Performing local identity through loan phonology. *Language & Communication* 49, 70–83.

Bacevičiūtė et al. (2004) = Rima Bacevičiūtė, Audra Ivanauskienė, Asta Leskauskaitė & Edmundas Trumpa (eds.), *Lietuvių kalbos tarmių chrestomatija*. Vilnius: Lietuvių kalbos institutas.

Bailey, Harold W. (1979). *Dictionary of Khotan Saka*. Cambridge–New York–Melbourne: Cambridge University Press.

Bakkum, Gabriël C.L. M. (2009). *The Latin Dialect of the Ager Faliscus*. Amsterdam: University Press.

Balaišis, Vytautas (1994). Das Problem der gotischen Diphthonge *ai*, *au* und die litauischen Lehnwörter *kvietỹs* „Weizen“, *kliẽpas* „Laibbrot“. *Baltistica* 4 *Priedas*, 5–12.

Bańkowski, Andrzej (2000). *Etymologiczny słownik języka polskiego*. 2 vols. Warszawa: Wydawnictwo naukowe PWN.

Barbaro, Cecilia Conati et al. (2021). A Multidisciplinary Approach to the Study of Early Neolithic Pyrotechnological Structures. The Case Study of Portonovo (Marche, Italy). *Open Archaeology* 7, 1160–1175.

Barber, Elizabeth J.W. (1991). *Prehistoric Textiles*. Princeton: University Press.

Bartens, Raija (1999). *Mordvalaiskielten rakenne ja kehitys*. Helsinki: Suomalais-Ugrilainen Seura.

Batisti, Roberto (2021). Birds of a feather? Latin *columba* ‘pigeon, dove’ and Greek κόλυμβος ‘grebe’. *Indogermanische Forschungen* 126, 205–228.

Becker, Johann P. (1904). Kurische sprache in Perwelk. *Beiträge zur Kunde der indogermanischen Sprachen* 28, 257–281.

Bednarczuk, Leszek (1976). Zapożyczenia ugrofińskie w językach bałtosłowiańskich. *Acta Baltico-Slavica* 9, 39–64.

van Beek, Lucien (2011). The “Saussure effect” in Greek: a reinterpretation of the evidence. *Journal of Indo-European Studies* 39, 129–175.

van Beek, Lucien (2013). 'Struck (with Blood)'. The Meaning and Etymology of παλάσσω and Its Middle Perfect πεπάλακτο. *Mnemosyne* 66, 541–565.

van Beek, Lucien (2018). Greek πέδιλον 'sandal' and the Origin of the e-Grade in PIE 'foot'. In: Lucien van Beek et al. (eds.), *Farnah* [FS Lubotsky]. Ann Arbor–New York: Beech Stave Press.

van Beek, Lucien (2022). *The Reflexes of Syllabic Liquids in Ancient Greek*. Leiden–Boston: Brill.

Beekes, Robert S.P. (1969). *The Development of the Proto-Indo-European Laryngeals in Greek*. Den Haag–Paris: Mouton.

Beekes, Robert S.P. (1971). A European substratum word. *Orbis* 20, 132–137.

Beekes, Robert S.P. (1975). [Review of] E.J. Furnée, Die wichtigsten konsonantischen Erscheinungen des Vorgriechischen. *Lingua* 36, 69–82.

Beekes, Robert S.P. (1982). GAv. *må* the PIE word for 'moon, month,' and the perfect participle. *Journal of Indo-European Studies* 10, 53–64.

Beekes, Robert S.P. (1987). The PIE Words for 'Name' and 'Me'. *Die Sprache* 33, 1–12.

Beekes, Robert S.P. (1988). PIE. *RHC-* in Greek and Other Languages. *Indogermanische Forschungen* 93, 22–45.

Beekes, Robert S.P. (1996). Ancient European Loanwords. *Historische Sprachforschung* 109, 215–236.

Beekes, Robert S.P. (1998). The origin of Lat. *aqua*, and of **teutā* 'people'. *Journal of Indo-European Studies* 26, 459–466.

Beekes, Robert S.P. (2000a). *European substratum words in Greek*. In: Michaela Ofitsch & Christian Zinko (eds.), *125 Jahre Indogermanistik in Graz*. Graz: Leykam, 21–31.

Beekes, Robert S.P. (2000b). Roots with nasal infix in Pokorny. *Amsterdamer Beiträge zur älteren Germanistik* 54, 3–23.

Beekes, Robert S.P. (2010). *Etymological Dictionary of Greek*. Leiden–Boston: Brill.

Beekes, Robert S.P. (2011) = Michiel de Vaan (ed.), *Comparative Indo-European Linguistics* [2nd edition]. Amsterdam–Philadelphia: Benjamins.

Beekes, Robert S.P. (2014) = Stefan Norbruis (ed.), *Pre-Greek*: *Phonology*, *Morphology*, *Lexicon*. Leiden–Boston: Brill.

Behre, Karl-Ernst (1992). The history of rye cultivation in Europe. *Vegetation History and Archaeobotany* 1, 141–156.

Behre, Karl-Ernst (2008). Collected seeds and fruits from herbs as prehistoric food. *Vegetation History and Archaeobotany* 17, 65–73.

Беляева, Ольга П. (1973). *Словарь говоров Соликамского района Пермской области*. Пермь: Пермский государственный педагогический институт.

Bellquist, Julie B. (1993). "Badger" in Indo-European. *Journal of Indo-European Studies* 21, 331–346.

Benecke, Norbert (2005). The Holocene distribution of European bison — the archaeozoological record. *Antropologia-Arkeologia* 57, 421–428.

Bennett, William H. (1949). The Monophthongization of Gothic *ái áu*. *Language* 25, 15–21.

Bentlin, Mikko (2008). *Niederdeutsch-finnische Sprachkontakte*. Helsinki: Suomalais-Ugrilainen Seura.

Bereczki, Gábor (1994). *Grundzüge der tscheremissischen Sprachgeschichte* I. Szeged.

Bereczki, Gábor (2013). *Etymologisches Wörterbuch des Tscheremissischen*. Wiesbaden: Harrassowitz.

de Bernardo Stempel, Patrizia (1987), *Die Vertretung der indogermanischen Liquiden und nasalen Sonanten im Keltischen*. Innsbruck: Institut für Sprachwissenschaft.

de Bernardo Stempel, Patrizia (1999). *Nominale Wortbildung des älteren Irischen*. Tübingen: Niemeyer.

Berneker, Erich (1899). Von der Vertretung des idg. *ĕu* im baltisch-slavischen Sprachzweig. *Indogermanische Forschungen* 10, 145–167.

Berneker = Berneker, Erich (1908–1914). *Slavisches etymologisches Wörterbuch*. 2 vols. Heidelberg: Winter.

Bertoldi, Vittorio (1932). Problèmes de substrat. Essai de méthodologie dans le domaine préhistorique de la toponymie et du vocabulaire. *Bulletin de la Société de linguistique de Paris* 32, 93–184.

Bezlaj = Bezlaj, France (1977–2007). *Etimološki slovar slovenskega jezika*. 5 vols. Ljubljana: Mladinka knjiga; Založba ZRC.

Bezzenberger, Adalbert (1877). *Beiträge zur Geschichte der litauischen Sprache*. Göttingen: Peppmüller.

Bezzenberger, Adalbert (1882). Ein lettisches lautgesetz. *Beiträge zur Kunde der indogermanischen Sprachen* 7, 273–277.

Bezzenberger, Adalbert (1885). *Lettische Dialekt-Studien*. Göttingen: Vandenhoeck & Ruprecht.

Bezzenberger, Adalbert (1887). Mundartliche Texte. *Mitteilungen der Litauischen litterarischen Gesellschaft* 2, 29–48.

Bezzenberger, Adalbert (1911). [Review of] Reinhold Trautmann, Die altpreußischen Sprachdenkmäler. *Zeitschrift für vergleichende Sprachforschung* 44, 285–331.

Bichlmeier, Harald (2016). Zur Frage des nichtindogermanischen Substrats im Germanischen: Ein kurzer Überblick. *Slovo a slovesnost* 77, 316–336.

Bjørnflaten, Jan I. (2006). Chronologies of the Slavicization of Northern Russia Mirrored by Slavic Loanwords in Finnic and Baltic. In: Juhani Nuorluoto (ed.), *The Slavicization of the Russian North: Mechanisms and Chronology*. Helsinki: Yliopisto, 50–77.

Blažek, Václav (1995). Indo-European 'Apple(s)'. *Sborník prací Filozofické fakulty brněnské univerzity* A 43, 15–20.

Blažek, Václav (2003). Semitic **t̠áwar-* "bull" and its relatives. In: Lionel Bender, David Appleyard & Gábor Takács (eds.), *Afrasian* [FS Diakonoff]. München: Lincom, 7–13.

Blažek, Václav (2018). Indo-European Dendronyms in the Perspective of External Comparison. *Journal of Indo-European Studies* 46, 1–45.

Blažek, Václav (2019). Tocharian B *lasto* in Balto-Slavic Perspective. *Journal of Indo-European Studies* 47, 213–222.

Blažek, Václav & Janyšková, Ilona (2015). Slavic **tisъ* — an evergreen problem of Slavic etymology? In: Ilona Janyšková & Helena Karlíková (eds.), *Etymological Research into Old Church Slavonic*, 69–98.

Bliujienė, Audronė & Curta, Florin (2011). Exotic Lands, Quixotic Friends: Eastern Lithuania and the Carpathian Basin in Late Antiquity and the Early Middle Ages (AD c. 380 to c. 620). *Medieval Archaeology* 55, 29–64.

Blokland, Rogier (2005). *The Russian Loanwords in Literary Estonian*. Diss. Universiteit Groningen.

Boček/Malčík = Vít Boček & Petr Malčík (2011, eds.), *Václav Machek*: *Korespondence*. 2 vols. Praha: Lidové noviny.

Boisacq, Émile (1916). *Dictionnaire Étymologique de la Langue Grecque*. Heidelberg: Winter.

Bopp, Franz (1833). *Vergleichende Grammatik des Sanskrit, Zend, Griechischen, Lateinischen, Litthauischen, Gothischen und Deutschen*. Berlin: Akademie der Wissenschaften.

Boryś, Wiesław (1977). Czakawskie *brńuse*. *Rocznik Slawistyczny* 38, 93–95.

Boryś, Wiesław (2008). *Słownik etymologiczny języka polskiego*. Kraków: Wydawnictwo Literackie.

Boutkan, Dirk (1998). On the form of North European substratum words in Germanic. *Historische Sprachforschung* 111, 102–133.

Boutkan, Dirk (2003a). On Gothic *magaþ* ~ Old Frisian *megith* and the form of some North European substratum words in Germanic. *Amsterdamer Beiträge zur älteren Germanistik* 58, 11–28.

Boutkan, Dirk (2003b). Lithuanian *šlãkas*, Norse *slag*. In: Alfred Bammesberger & Theo Venneman (eds.), *Languages in Prehistoric Europe*. Heidelberg: Winter, 245–252.

Boutkan, Dirk & Kossmann, Maarten (2001). On the etymology of 'silver'. *North-Western European Language Evolution* 38, 3–15.

Boutkan, Dirk & Siebinga, Sjoerd M. (2005). *Old Frisian Etymological Dictionary*. Leiden–Boston: Brill.

Braune, Wilhelm (1876). Ueber den altpreussischen diphthong ai. *Beiträge zur vergleichenden Sprachforschung* 8, 91–101.

Breidaks, Antons B. [= А.Б. Брейдак] (1983). Некоторые данные балтизмов финно-угорских языков. *Baltistica* 19, 47–51.

Brenzinger, Matthias (1992). Lexical retention in language shift: Yaaku/Mukogodo-Maasai and Elmolo/Elmolo-Samburu. In: Matthias Brenzinger (ed.), *Language*

Death: *Factual and Theoretical Explorations with Special Reference to East Africa*. Berlin: Mouton de Gruyter, 213–254.

Brevé, Niels W.P. et al. (2022). Historical reconstruction of sturgeon (*Acipenser* spp.) spatiotemporal distribution and causes for their decline in North-Western Europe. *Biodiversity and Conservation* 31, 1149–1173.

Brodowski = Jacob Brodowskỹ. *Lexicon Germanico-Lithuanicum et Lithuanico-Germanicum* [Manuscript]. Accessed online at epaveldas.lt/preview?id=MAB04-0000708 16.

Brückner, Aleksander (1877). *Litu-slavische Studien* I. *Die slavischen Fremdwörter im Litauischen*. Weimar: Böhlau.

Brückner, Aleksander (1927). *Słownik etymologiczny języka polskiego*. Kraków: Spółka Wydawnicza.

Brückner, Aleksander (1929). Die germanischen Elemente im Gemeinslavischen. *Archiv für Slavische Philologie* 42, 125–146.

Brugmann, Karl (1897). *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen. Zweite Bearbeitung* I. *Einleitung und Lautlehre*. Strassburg: Trübner.

Brugmann, Karl (1906). *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen. Zweite Bearbeitung* II. *Vergleichende Laut-, Stammbildungs- und Flexionslehre*. Strassburg: Trübner.

Buck, Carl Darling (1904). *A Grammar of Oscan and Umbrian*. Boston: Ginn.

Būga, Kazimieras (1908). *Aistiški studijai*: *Tyrinėjimai lygintijo prūsų, latvjų ir liėtuvjų kalbômoksljo srityję* I. Peterburgas: Imperatoriškoji Mokslų Akademija.

Būga, Kazimieras (1911). *Apie lietuvių asmens vardus*. Vilnius: Kukta [= RR I: 201–269].

Būga, Kazimieras [= К.К. Буга] (1912). Lituanica. *Извѣстія отдѣленія русскаго языка и словесности императорской академіи наукъ* 17/1, 1–49 [= RR I: 339–383].

Būga, Kazimieras (1913). Kann man Keltenspuren auf baltischem Gebiet nachweisen? *Rocznik sławistyczny* 6, 1–38 [= RR I: 496–530].

Būga, Kazimieras [= Казимиръ Буга] (1914). Славяно-балтійскія этимологіи 135–160. *Русскій филологическій вѣстникъ* 72, 187–202 [= RR I: 460–472].

Būga, Kazimieras [= Казимиръ Буга] (1915). Славяно-балтійскія этимологіи 161–185. *Русскій филологическій вѣстникъ* 73, 335–343 [= RR I: 473–479].

Būga, Kazimieras [= Казимиръ Буга] (1916). Славяно-балтійскія этимологіи 186–261. *Русскій филологическій вѣстникъ* 75, 141–156 [= RR I: 479–495].

Būga, Kazimieras (1922). *Kalba ir senovė*. Kaunas: Svietimo Ministerija [= RR II: 7–330].

Būga, Kazimieras (1923a). Upių vardų studijos ir aisčių bei slavėnų senovė. *Tauta ir žodis* 1, 1–44 [= RR III: 493–550].

Būga, Kazimieras (1923b). [Review of] *H. H. Bender*. A Lithuanian etymological index. *Tauta ir žodis* 1, 397–405 [= RR III: 648–657].

Būga, Kazimieras [= Kasimir Būga] (1923–1924). Die Metatonie im Litauischen und Lettischen. *Zeitschrift für vergleichende Sprachforschung* 51, 109–142; 52, 91–98, 250–302 [= RR II: 386–483].

Būga, Kazimieras (1924a). *Lietuvių kalbos žodynas*. Kaunas: Svietimo Ministerija [= RR III: 9–483].

Būga, Kazimieras (1924b). Šis-tas iš lietuvių ir indoeuropiečių senovės. *Tauta ir žodis* 2, 98–110 [= RR III: 584–599].

Būga, Kazimieras (1925). Die litauisch-weißrussischen Beziehungen und ihr Alter. *Zeitschrift für slavische Philologie* 1, 26–55 [= RR III: 749–778].

Būga RR = Zigmas Zinkevičius (1958–1962, comp.), *K. Būga. Rinktiniai Raštai*. 4 vols. Vilnius: Valstybinė politinės ir mokslinės literatūros leidykla.

Bugge, Sophus (1888). Etymologische Studien ueber germanische Lautverschiebung III. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 13, 311–339.

Булаховский, Леонид А. (1948). Общеславянские названия птиц. *Известия академии наук СССР. Отделение литературы и языка* 7/2, 97–124.

Булаховский, Леонид А. (1958). *Исторический комментарий к русскому литературному языку* [5th edition]. Киев: Радянська школа.

Булаховский, Леонид А. (1968). Морфологическая проблематика русских наименований птиц. *Вопросы языкознания* 1968/4, 100–106.

Byrd, Andrew M. (2010). *Reconstructing Indo-European Syllabification*. Diss. University of California.

Çabej, Eqrem (1976). *Studime gjuhësore*: *Studime etimologjike në fushë të shqipes*. 2 vols. Prishtinë: Rilindja.

Цаболов, Руслан Л. (2001). *Этимологический словарь курдского языка* I. *A–M*. Москва: Восточная литература.

Цаболов, Руслан Л. (2010). *Этимологический словарь курдского языка* II. *N–Ž*. Москва: Восточная литература.

Campbell, Alistair (1959). *Old English Grammar*. London: Oxford University Press.

Черепанова, Евгения А. (1973). К географии балтизмов в говорах полесья. *Baltistica* 9, 71–74.

Chantraine, Pierre (1933). *La formation des noms en grec ancien*. Paris: Klincksieck.

Chantraine DELG = Chantraine, Pierre (1968–1980). *Dictionnaire étymologique de la langue grecque*. 4 vols. Paris: Klincksieck.

Charpentier, Jarl (1930). Der Name des Roggens. *Arkiv för nordisk filologi* 46, 63–73.

Cheung, Johnny (2002). *Studies in the Historical Development of the Ossetic Vocalism*. Wiesbaden: Reichert.

Cheung, Johnny (2007). *Etymological Dictionary of the Iranian Verb*. Leiden–Boston: Brill.

Cheung, Johnny & Aydemir, Hakan (2002). Turco-Afghanica: On East Iranian **amarnā* and Turkic *alma*, *alïmla*, *almïla* 'apple'. In: Михаил С. Пелевин (ed.), «*На пастбище Мысли Благой*» [FS Стеблин-Каменский]. Санкт-Петербург: Контраст, 73–94.

Cigale, Matej (1860). *Deutsch-slovenisches Wörterbuch*. 2 vols. Ljubljana: Blasnik.

Clackson, James (1994). *The Linguistic Relationship between Armenian and Greek*. Oxford: Wiley–Blackwell.

Clauson, Gerard (1972). *An Etymological Dictionary of Pre-Thirteenth-Century Turkish*. Oxford: University Press.

Collinder, Björn (1955). *Fenno-Ugric Vocabulary*: *An Etymological Dictionary of the Uralic Languages*. Stockholm: Almqvist & Wiksell.

Collinder, Björn (1960). *Comparative Grammar of the Uralic Languages*. Stockholm: Almqvist & Wiksell.

Čop, Bojan (1973). Oštirs sprachwissenschaftliche Ideenwelt. *Linguistica* 13, 13–96.

Cordes, Gerhard (1973). *Altniederdeutsches Elementarbuch*. Heidelberg: Winter.

Craddock, Jerry R. (1969). *Latin Legacy Versus Substratum Residue*. Berkeley–Los Angeles: University of California Press.

Crane, E. Eva (1999). *The World History of Beekeeping and Honey Hunting*. New York–London: Routledge.

Даль^2 = Даль, Владиміръ И. (1880–1882). *Толковый словарь живаго великорусскаго языка* [2nd edition]. 4 vols. С.-Петербургъ–Москва: Вольфъ.

Даль^3 = Иван А. Бодуэн-де-Куртенэ (1903–1909, ed.), *Толковый словарь живого великорусскаго языка Владиміра Даля* [3rd edition]. 4 vols. С.-Петербургъ–Москва: Вольфъ.

Darden, Bill J. (1990). Laryngeals and Syllabicity in Balto-Slavic and Indo-European. *Papers from the 26th Annual Meeting of the Chicago Linguistics Society* 11. *The Parasession on the Syllable in Phonetics & Phonology*, 61–70.

Daukantas, Simonas (1846). *Dajnes Źiamajtiû* (*pagał Zôdiû Dajnînikû îszraszytas*). Petropìlie: Kray.

Decaux, Étienne (1966). La neutralisation n'est pas un changement structurel. *Bulletin de la Société de linguistique de Paris* 61, 57–81.

Delamarre, Xavier (2003). *Dictionnaire de la langue gauloise* [2nd edition]. Paris: Editions Errance.

Della Volpe, Angela (1996). Indo-European Architectural Terms and the Pre-Indo-Europeans. In: Karlene Jones-Bley & Martin E. Huld (eds.), *The Indo-Europeanization of Northern Europe*. Washington: Institute for the Study of Man, 148–165.

Demiraj, Bardhyl (1997). *Albanische Etymologien. Untersuchungen zum albanischen Erbwortschatz*. Amsterdam–Atlanta: Rodopi.

Derksen, Rick (1996). *Metatony in Baltic*. Amsterdam–Atlanta: Rodopi.

Derksen, Rick (1999). North European **KulP-* 'beak, snout'. *Baltistica* 34, 69–76.

Derksen, Rick (2000). Old Icelandic *jarpi* 'hazel-grouse', *rjúpa* 'ptarmigan' and their Germanic and Balto-Slavic cognates. In: Dirk Boutkan & Arend Quak (eds.), *Language Contact. Substratum, superstratum, adstratum in Germanic Languages*. Amsterdam–Atlanta: Rodopi, 75–87.

Derksen, Rick (2002). "Rozwadowski's change" in Baltic. *Baltu filoloģija* 9, 5–12.

Derksen, Rick (2003). Slavic **jь-*. In: Peter Houtzagers et al. (eds.), *Dutch Contributions to the Thirteenth International Congress of Slavists, Ljubljana: Linguistics*. Amsterdam–Atlanta: Rodopi, 97–105.

Derksen, Rick (2007). Balto-Slavic Etymological Studies and Winter's Law: A Concise Review of Dybo 2002. In: Mate Kapović & Ranko Matasović (eds.), *Tones and Theories: Proceedings of the International Workshop on Balto-Slavic Accentology*. Zagreb: Institut za hrvatski jezik i jezikoslovlje, 39–46.

Derksen, Rick (2008). *Etymological Dictionary of the Slavic Inherited Lexicon*. Leiden–Boston: Brill.

Derksen, Rick (2010). The development of PIE **eu* in Baltic and Slavic. In: Elena Stadnik-Holzer & Georg Holzer (eds.), *Sprache und Leben der frühmittelalterlichen Slaven* [FS Katičić]. Frankfurt am Main: Peter Lang, 37–41.

Derksen, Rick (2015). *Etymological Dictionary of the Baltic Inherited Lexicon*. Leiden–Boston: Brill.

Derksen, Rick (2019). 'Henbane' in Slavic. In: Uwe Bläsing, Jasmine Dum-Tragut & Theo M. van Lint (eds.), *Armenian, Hittite, and Indo-European Studies* [Gedenkschrift Weitenberg]. Leuven: Peeters, 113–119.

Derksen, Rick (2020). On the earliest Slavic borrowings in East Baltic. *Baltu filologija* 29, 33–53.

Deroy, Louis (1956). La valeur du suffixe préhellénique *-nth-* d'après quelques noms grees en *-νθος*. *Glotta* 35, 171–195.

Diebold, A. Richard (1976). Contributions to the Indo-European Salmon Problem. In: William M. Christie, Jr. (ed.), *Current Progress in Historical Linguistics*. Amsterdam–New York: North-Holland, 341–387.

Dini, Pietro U. (2014) = Milda B. Richardson & Robert E. Richardson (2014, trans.), *Foundations of Baltic Languages*. Vilnius.

Doerfer, Gerhard (1965). *Türkische und Mongolische Elemente im Neupersischen* II. Wiesbaden: Steiner.

Doerfer, Gerhard (1971). Bemerkungen zur Methodik der türkischen Lautlehre. *Orientalistische Literaturzeitung* 66, 325–344.

Donner, Otto (1884). Über den Einfluss des Litauischen auf die finnischen Sprachen. *Internationale Zeitschrift für allgemeine Sprachwissenschaft* I, 257–271.

Donner/Joki = Aulis J. Joki (1944, ed.), *Kai Donners kamassisches Wörterbuch*. Helsinki: Suomalais-Ugrilainen Seura.

Dressler, Wolfgang (1974). [Review of] Die wichtigsten konsonantischen Erscheinungen des Vorgriechischen [...] by Edzard J. Furnée. *Language* 50, 736–738.

Driessen, Michiel (2003). **h_2é-$h_2$us-o-*, the Proto-Indo-European term for 'gold'. *Journal of Indo-European Studies* 31, 347–362.

Dunkel, George E. (2014). *Lexikon der indogermanischen Partikeln und Pronominalstämme* II. *Lexikon*. Heidelberg: Winter.

Durkin-Meisterernst, Desmond (2004). *Dictionary of Manichaean Texts* III/1: *Dictionary of Manichaean Middle Persian*. Turnhout: Brepols.

Дыбо, Анна В. (2007). *Лингвистические контакты ранних тюрков. Лексический фонд. Пратюркский период*. Москва: Восточная литература.

Дыбо, Анна В. (2008). Материальный быт ранних тюрок. Жилище. In: Анна В. Дыбо (ed.), *Природное окружение и материальная культура пратюркских народов*. Москва: Восточная литература, 219–272.

Дыбо, Владимир А. (1961). Сокращение долгот в кельто-италийских языках и его значение для балто-славянской и индоевропейской акцентологии. *Вопросы славянского языкознания* 5, 9–34.

Дыбо, Владимир А. (2000). *Морфонологизованные парадигматические акцентные системы* I. Москва: Языки русской культуры.

Дыбо, Владимир А. (2002). Balto-Slavic Accentology and Winter's Law. *Studia Linguarum 3*, 295–515.

Дыбо, Владимир А. (2008). Германское сокращение индоевропейских долгот, германский «Verschärfung» (закон Хольцмана) и балто-славянская акцентология. *Аспекты компаративистики* 3, 537–608.

Eason, Perri, Basem Rabia & Omar Attum (2016). Hunting of migratory birds in North Sinai, Egypt. *Bird Conservation International* 26, 39–51.

Eckert, Rainer (1983). *Die Nominalstämme auf -i im Baltischen*. Berlin: Akademie der Wissenschaften.

Ehrlich, Freydis et al. (2021). Application of morphometric and stable isotope analyses for distinguishing domestic and wild geese. *International Journal of Osteoarchaeology* 32, 457–466.

Elger, Georg (1683). *Dictionarium Polono-Latino-Lottauicum*. Vilnius: Societas Iesu.

Emmerick, Ronald E. (1968). *Saka Grammatical Studies*. London: Oxford University Press.

Emmerick, Ronald E. (1989). Khotanese and Tumshuqese. In: Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum*. Wiesbaden: Reichert, 204–229.

Endzelīns, Jānis [= I. Эндзелинъ] (1899). Латышскія заимствованія изъ славянскихъ языковъ. *Живая старина* 9/3, 285–312 [= DI I: 80–112].

Endzelīns, Jānis [= J. Endzelin] (1905). Zur erweichung der gutturale im Lettischen. *Beiträge zur Kunde der indogermanischen Sprachen* 29, 178–195 [= DI I: 290–306].

Endzelīns, Jānis [= И. Эндзелинъ] (1907). О происхожденіи литовско-латышскаго ё. *Извѣстія Отдѣленія русскаго языка и словесности Императорской Академіи наукъ* 12/1, 40–66 [= DI II: 9–30].

Endzelīns, Jānis [= J. Endzelin] (1909). Lettische Miscellen. *Zeitschrift für vergleichende Sprachforschung* 42, 375–380 [= DI II: 84–89].

Endzelīns, Jānis [= И. Эндзелинъ] (1911). *Славяно-балтійскіе этюды*. Харьковъ: Зильбербергъ [= DI II: 167–354].

Endzelīns, Jānis [= J. Endzelin] (1914a). Miszellen. *Indogermanische Forschungen* 33, 119–127 [= DI II: 528–537].

Endzelīns, Jānis [= J. Endzelin] (1914b). Zu den kurischen Bestandteilen des Lettischen. *Indogermanische Forschungen* 33, 96–104 [= DI II: 504–512].

Endzelīns, Jānis [= J. Endzelin] (1923). *Lettische Grammatik*. Heidelberg: Winter.

Endzelīns, Jānis [= J. Endzelin] (1924). Germanisch-baltische Miszellen. *Zeitschrift für vergleichende Sprachforschung* 52, 110–128 [= DI III/1: 395–412].

Endzelīns, Jānis [= J. Endzelin] (1932). [Review of] E. Blese, Latviešu personu vārdu studijas I. *Zeitschrift für Slavische Philologie* 9, 245–262 [= DI III/2: 498–517].

Endzelīns, Jānis [= J. Endzelin] (1933). Zur Herkunft der baltischen Lehnwörter des Finnischen. *Mémoires de la Société finno-ougrienne* 67, 76–83 [= DI III/2: 128–134].

Endzelīns, Jānis (1937). Sīkumi. *Filologu biedrības raksti* 17, 162–166 [= DI III/2: 286–292].

Endzelīns, Jānis (1943). *Senprūšu valoda*. Riga: University Press [= DI IV/2: 9–351].

Endzelīns DI = Daina Zemzare, Hermanis Bendiks et al. (eds., 1974–1982). *Jānis Endzelīns. Darbu izlase*. 4 vols. Rīga: Zinātne

Endzelīns/Schmalstieg 1971 = William R. Schmalstieg & Benjamiņš Jēgers (1971, trans.), *Jānis Endzelīns' Comparative Phonology and Morphology of the Baltic Languages*. Den Haag–Paris: Mouton.

[Endzelīns ME — see ME under Abbreviated Titles.]

Erhart, Adolf (1995). Archaisch oder konservativ? Das Anatolische und das Baltische. *Linguistica Baltica* 4, 1–12.

Erjavec, Fran (1883). Is pótne torbe. *Letopis Matice Slovenske* 1882/1883, 195–351.

Erman/Grapow = Erman, Adolf & Grapow, Hermann (1926–1961). *Wörterbuch der aegyptische Sprache*. 5 vols. Leipzig: Hinrichs.

Ernout/Meillet = Ernout, Alfred & Meillet, Antoine (1951). *Dictionnaire étymologique de la langue latine* [3rd edition]. Paris: Klincksieck.

Ertuğ, Füsun (2000). An Ethnobotanical Study in Central Anatolia (Turkey). *Economic Botany* 54/2, 155–182.

Falileyev, Alexander & Isaac, Graham R. (2003). Leeks and garlic: the Germanic ethnonym *Cannenefates*, Celtic **kasn-* and Slavic **kesn-*. *North-Western European Language Evolution* 42, 3–12.

Falk/Torp = Falk, Hjalmar S. & Torp, Alf (1910–1911). *Norwegisch-dänisches etymologisches wörterbuch*. 2 vols. Heidelberg: Winter.

Федотов = Федотов, Михаил Р. (1996). *Этимологический словарь чувашского языка*. 2 vols. Чебоксары: Чувашский государственный университет гуманитарных наук.

Feist, Sigmund (1910). Die germanische und die hochdeutsche lautverschiebung sprachlich und ethnographisch betrachtet. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 37, 307–354.

Feist, Sigmund (1913). *Kultur, Ausbreitung und Herkunft der Indogermanen*. Berlin: Weidmann.

Feist, Sigmund (1932). The Origin of the Germanic Languages and the Indo-Europeanising of North Europe. *Language* 8, 245–254.

Fennel, Trevor G. (1982). *The First Latvian Grammar*: *J.G. Rehehusen's "Manductio ad linguam lettonicam ..."*. Melbourne: Latvian Tertiary Committee.

Fiesel, Eva (1928). *Namen des griechischen Mythos im Etruskischen*. Göttingen: Vandenhoeck & Ruprecht.

Filipović, Dragana et al. (2020). New AMS ^{14}C dates track the arrival and spread of broomcorn millet cultivation and agricultural change in prehistoric Europe. *Scientific Reports* 10:13698.

Foley, James (1977). *Foundations of theoretical phonology*. Cambridge: University Press.

Forssman, Berthold (2001). *Lettische Grammatik*. Dettelbach: Röll.

Fortescue, Michael & Vajda, Edward (2022). *Mid-Holocene Language Connections between Asia and North America*. Leiden–Boston: Brill.

Fraenkel, Ernst (1936a). [Review of] E. Dickenmann, Untersuchungen über die Nominalkomposition im Russischen. *Zeitschrift für Slavische Philologie* 13, 205–236.

Fraenkel, Ernst (1936b). Die indogermanischen *-l*-Stämme. *Zeitschrift für vergleichende Sprachforschung* 63, 168–201.

Fraenkel, Ernst (1936c). Zur Suffixsubstitution im Litauischen. *Zeitschrift für vergleichende Sprachforschung* 63, 201–206.

Fraenkel, Ernst (1937). Zur *eu*-Erweiterung indogermanischer Wurzeln unter besonderer Berüksichtigung [*sic*.] von Baltisch, Slavisch und Griechisch. *Mélanges Émile Boisacq* 1, 355–381.

Fraenkel, Ernst (1950a). *Die baltischen Sprachen*. Heidelberg: Winter.

Fraenkel, Ernst (1950b). [Review of] J.B. Hofmann: Etymologisches Wörterbuch des Griechischen. *Gnomon* 22, 235–239.

Fraenkel, Ernst (1951). Zur indoeuropäischen Wortbildung und Etymologie. *Lingua Posnaniensis* 3, 113–134.

[Fraenkel LEW — See LEW under Abbreviated Titles.]

Freudenthal, Axel O. & Vendell, Hermann (1886). *Ordbok öfver estländsk-svenska dialekterna*. Helsingfors: Tidnings- & Tryckeri-Aktibolagets tryckeri.

Friedrich, Paul (1970). *Proto-Indo-European Trees*. Chicago–London: University of Chicago.

Frisk = Frisk, Hjalmar (1954–1972). *Griechisches Etymologisches Wörterbuch*. 3 vols. Heidelberg: Winter.

Frischbier = Frischbier, Hermann (1882–1883). *Preussisches Wörterbuch: Ost- und westpreussische Provinzialismen in alphabetischer Folge*. 2 vols. Berlin: Enslin.

Furlan, Metka (2013). *Novi etimološki slovar slovenskega jezika. Poskusni zvezek*. Ljubljana: Znanstvenoraziskovalni center.

Furnée, Edzard J. (1972). *Die wichtigsten konsonantischen Erscheinungen des Vorgriechischen*. Den Haag–Paris: Mouton.

Галинская, Елена А. (1993). О хронологии некоторых изменений в системе вокализма праславянского языка. In: Борис А. Успенский & Мария Н. Шевелёва (eds.), *Исследования по славянскому историческому языкознанию* [Gedenkschrift Хабургаев]. Москва: Московский государственный университет, 35–46.

Гальковскій, Николай М. (1913). *Борьба христіанства съ остатками язычества въ древней Руси* II. *Древне-русскія слова и поученія, направленныя противъ остатковъ язычества въ народѣ*. Москва: А. Снегирева.

Гамкрелидзе, Тамаз В. & Иванов, Вячеслав В. (1984). *Индоевропейский язык и индоевропейцы*. Тбилиси: Университет.

Garde, Paul [= П. Гард] (1974). К истории восточнославянских гласных среднего подъема. *Вопросы языкознания* 1974/3, 106–115.

Garnier, Romain (2014). Nouvelles considérations sur l'effet Kortlandt. *Glotta* 90, 139–159.

Garnier, Romain (2015). [Review of] Robert Stephen Paul Beekes, Pre-Greek: Phonology Morphology, Lexicon. *Journal asiatique* 303, 329–332.

Gāters, Alfred (1953). Der baltische Name für das Grummet. *Zeitschrift für vergleichende Sprachforschung* 71, 113–117.

Gebauer = Gebauer, Jan (1903–1916). *Slovník staročeský*. 2 vols. [*A–N*]. Praha: Česká akademie věd a umění/Unie.

Gelumbeckaitė, Jolanta (2017). The evolution of Baltic. In: Jared Klein, Brian Joseph & Matthias Fritz (eds.), *Handbook of Comparative and Historical Indo-European Linguistics*. Berlin: De Gruyter, 1712–1715.

Георгиев, Владимир И. (1957). *Тракийският език*. София: Българска академия на науките.

Георгиев, Владимир И. (1973). [Review of] Furnée, Edzard J.: Die wichtigsten konsonantischen Erscheinungen des Vorgriechischen. *Kratylos* 17, 164–167.

Герд, Александр С. (1970). Из истории прибалтийско-финских названий рыб (sintti, tintti, tint). *Советское финно-угроведение* 6/2, 87–91.

Герд, Александр С. (1980). Лексические заимствования или слова-автохтоны. *IV Всесоюзная конференция балтистов. Тезисы докладов*. Рига, 125.

Герд, Александр С. (1981). Из истории прибалтийско-финских названий рыб (siika, siiga; muhju). In: Георгий М. Керт & Мария И. Зайцева (eds.), *Прибалтийско-финское языкознание*, 50–54.

Геров = Геровъ, Найденъ (1895–1904). *Рѣчник на Блъгарскый языкъ*. 5 vols. Пловдивъ: Съгласие.

Gerullis, Georg (1922). *Die altpreußischen Ortsnamen*. Berlin–Leipzig: De Gruyter.

Gharib, Badrozzaman [= Qarīb, Badr al-Zamān] (1995). *Sogdian Dictionary*. Tehran: Farhangan.

Giannelli, Ciro & Vaillant, André (1958). *Un lexique macédonien du XVI^e siècle*. Paris: Institut d'études slaves.

Gilbert, M. Thomas P. & Shapiro, Michael D. (2013). Pigeons, Domestication of. In: Claire Smith (ed.), *Encyclopedia of Global Archaeology*. New York: Springer.

Gippert, Jost (2017). Armeno-Albanica II: Exchanging doves. In: Bjarne S.S. Hansen, Adam Hyllested, Anders R. Jørgensen & Guus Kroonen (eds.), *Usque ad radices* [FS Olsen]. København: Museum Tusculanum, 179–192.

Girininkas, Algirdas (2019). Hoes and the First Farmers on Lithuanian Territory: From Hoe to Ard. *Archaeologia Baltica* 26, 66–79.

Gliwa, Bernd (2008). Ist lit. *túopa* 'Pappel' wirklich ein Erbwort? *Baltistica* 43, 239–244.

Gliwa, Bernd (2009). Pflanzennamen mit Präfix *pa-* und angrenzendes im Litauischen und Lettischen. *Baltistica* 44, 77–90.

Gołąb, Zbigniew (1982). Kiedy nastąpilo rozszczepienie językowe Baltów i Słowian. *Acta Baltico-Slavica* 14, 121–134.

Gołąb, Zbigniew (1992). *The Origins of the Slavs*. Columbus: Slavica.

Gorbachov, Slava [Yaroslav] (2006). "North Indo-European" *$tuh_x s\underset{\circ}{n}tih_2$ '1000': a new proposal [Conference handout]. *AATSEEL Annual Meeting, Philadelphia* (29 December 2006).

Grabowski, Radosław (2011). Changes in cereal cultivation during the Iron Age in southern Sweden. *Vegetation History and Archaeobotany* 20, 479–494.

Griffin, David A. (1959). Los mozarabismos del «Vocabulista» atribuído a Ramón Martí (Continuación). *al-Andalus* 24, 333–380.

Grikpėdis, Mindaugas & Motuzaitė Matuzevičiūtė, Giedrė (2016). The beginnings of rye (*Secale cereale*) cultivation in the East Baltics. *Vegetation History and Archaeobotany* 25, 601–610.

Grikpėdis, Mindaugas & Motuzaitė Matuzevičiūtė, Giedrė (2017). A review of the earliest evidence of agriculture in Lithuania and the earliest direct AMS date on cereal. *European Journal of Archaeology* 21/2, 1–16.

Grikpėdis, Mindaugas & Motuzaitė Matuzevičiūtė, Giedrė (2020). From barley to buckwheat: Plants cultivated in the Eastern Baltic region until the 13th–14th century AD. In: Santeri Vanhanen & Per Lagerås (eds.), *Archaeobotanical studies of past plant cultivation in northern Europe*. Groningen: Barkhuis.

Grinaveckienė, Elena & Mackevič, Juzefa F. (1989). Lie. (*i*)*au* vokalizmo atliepai baltarusių lituanizmuose. *Baltistica* 3 *Priedas*, 70–80.

Grinaveckis, Vladas (1972). Žem. *jeĩnis*, *jeĩnė* ir *muñsiai*. *Baltistica* 8, 73–77.

Grünthal, Riho (2012). Baltic loanwords in Mordvin. In: Riho Grünthal & Petri Kallio (eds.), *A Linguistic Map of Prehistoric Northern Europe*. Helsinki: Suomalais-Ugrilainen Seura, 297–343.

Haak, Wolfgang et al. (2023). The Corded Ware Complex in Europe in Light of Current Archeogenetic and Environmental Evidence. In: Kristian Kristiansen, Guus Kroonen

& Eske Willerslev (eds.), *The Indo-European Puzzle Revisited*. Cambridge: University Press, 63–80.

Haas, Otto (1960). *Das frühitalische Element: Versuch über die Sprache der ersten Indogermanen Italiens*. Wien: Notring.

Hackstein, Olav (1992). Eine weitere griechisch-tocharische Gleichung: Griechisch *πτῆξαι* und tocharisch B *pyāktsi*. *Glotta* 70, 136–165.

Häkkinen, Jaakko (2009). Kantauralin ajoitus ja paikannus: perustelut puntarissa. *Suomalais-Ugrilaisen Seuran Aikakauskirja* 92, 9–56.

Häkkinen, Kaisa (2004). *Nykysuomen etymologinen sanakirja*. Juva: WSOY.

Häkkinen, Kaisa & Lempiäinen, Terttu (1996). Die ältesten Getreidepflanzen der Finnen und ihre Namen. *Finnisch-ugrische Forschungen* 53, 115–182.

Hamann, Silke R. (2003). *The Phonetics and Phonology of Retroflexes*. Diss. Universiteit Utrecht.

Hamp, Eric P. (1965). Evidence in Keltic. In: Werner Winter (ed.), *Evidence for Laryngeals*. The Hague: Mouton, 224–235.

Hamp, Eric P. (1970a). Lithuanian *ugnìs*, Slavic *ognь*. In: Thomas F. Magner & William R. Schmalstieg (eds.), *Baltic Linguistics*. University Park: Pennsylvania State University, 75–79.

Hamp, Eric P. (1970b). Productive suffix ablaut in Baltic. *Baltistica* 6, 27–32.

Hamp, Eric P. (1971). Varia III. *Ériu* 22, 181–187.

Hamp, Eric P. (1978). Sound Change and the Etymological Lexicon. In: Donka Farkas, Wesley M. Jacobsen & Karol W. Todrys (eds.), *Papers from the Parasession on the Lexicon*. Chicago: Linguistic Society, 184–195.

Hamp, Eric P. (1979). The North European word for 'apple'. *Zeitschrift für Celtische Philologie* 37, 158–166.

Hamp, Eric P. (1981). Refining Indo-European Lexical Entries. *Zeitschrift für vergleichende Sprachforschung* 95, 81–83.

Hamp, Eric P. (1990). The Pre-Indo-European language of Northern (Central) Europe. In: Thomas L. Markey & John A.C. Greppin (eds.), *When Worlds Collide*. Ann Arbor: Karoma, 291–305.

Hamp, Eric P. (2003). On **ri* in Latin and Albanian *krip*. *Glotta* 59, 230–231.

Hao, Chen (2019). *Bark*: A Study on the Spiritual World of the Early Türks. *Studia Uralo-Altaica* 52, 15–20.

Hasiuk, Michał (1970). Die Ferndissimilation des *k*, *g* in den litauischen Dialekten. In: Velta Rūķe-Draviņa (ed.), *Donum Balticum* [FS Stang]. Stockholm: Almquist & Wiksell, 352–360.

Haspelmath, Martin & Tadmor, Uri (2009, eds.), *Loanwords in the World's Languages: A Comparative Handbook*. Berlin: De Gruyter Mouton.

Häusler, Alexander (2002). *Nomaden, Indogermanen, Invasion. Zur Entstehung eines Mythos*. Halle: Orientwissenschaftliches Zentrum.

Havlová, Eva (1966). Slav. *člověkъ*. *Scando-Slavica* 12, 80–86.

Havlová, Eva (1994). Etymologické metody Václava Machka. *Slavia* 63, 389–394.

[Havlová ESJS — see ESJS under Abbreviated Titles.]

Hehn, Victor (1870). *Kulturpflanzen und Hausthiere in ihrem Übergang aus Asien nach Griechenland und Italien sowie in das übrige Europa*. Berlin: Gebrüder Borntraeger.

Hehn, Victor (1911) = [8th edition of Hehn 1870, ed. Otto Schrader].

Heidermanns, Frank (1993). *Etymologisches Wörterbuch der germanischen Primäradjektive*. Berlin–New York: De Gruyter.

Heikkilä, Mikko (2013). Itämerensuomen ehdollisesta äänteenmuutoksesta /ai/ > /ei/. *Virittäjä* 117, 583–592.

Heikkilä, Mikko (2014). *Bidrag till Fennoskandiens språkliga förhistoria i tid och rum*. Helsinki: Unigrafia.

Heikkinen, Marja E. et al. (2015). Relationship between wild greylag and European domestic geese based on mitochondrial DNA. *Animal Genetics* 46, 485–497.

Helimski, Eugen (1997a). *Die matorische Sprache*. Szeged: JATE.

Helimski, Eugen (1997b). The southern neighbours of Finno-Ugrians: Iranians or an extinct branch of Aryans ("Andronovo Aryans")? In: Tette Hofstra et al. (eds.), *Finnisch-ugrische Sprachen in Kontakt*, 117–125.

Helimski, Eugen [= Хелимский, Евгений] (2001). Уральцы и их предшественники: Белые пятна на этноисторической карте Северной Евразии и уральские языки. In: *Congressus Nonus Internationalis Fenno-Ugristarum* IV. *Dissertationes sectionum: Linguistica* I. Tartu: Paar, 332–336.

Hendriks, Pepijn (2014). *Innovation in Tradition: Tönnies Fonne's Russian-German Phrasebook* (*Pskov, 1607*). Amsterdam–New York: Rodopi.

Henning, Walter B. (1946). The Sogdian Texts of Paris. *Bulletin of the School of Oriental and African Studies* 11, 713–740.

Henning, Walter B. (1963a). The Kurdish Elm. *Asia Major* 10, 68–72.

Henning, Walter B. (1963b). Coriander. *Asia Major* 10, 195–199.

Henning, Walter B. (1965). A Grain of Mustard. 29–47. *AIΩN-Linguistica* 6, 29–47.

Herne, Gunnar (1954). *Die slavischen Farbenbennenungen: Eine semasiologisch-etymologische Untersuchung*. Uppsala: Almqvist & Wiksell.

Hertzen, Erik von (1973). Itämerensuomen lainakerrostumien ikäämisestä. *Suomalais-Ugrilaisen Seuran Aikakauskirja* 72, 77–105.

Hill, Eugen (2016). Phonological evidence for a Proto-Baltic stage in the evolution of East and West Baltic. *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 13, 205–232.

Hillman, Gordon (1978). On the Origins of Domestic rye. *Anatolian Studies* 28, 157–174.

Hintze, Almut (1994). *Der Zamyād-Yašt: Edition, Übersetzung, Kommentar*. Wiesbaden: Reichert.

Hirt, Hermann (1892). Vom schleifenden und gestossenen Ton in den indogermanischen Sprachen I. *Indogermanische Forschungen* 1, 1–42.

Hirt, Hermann (1896). Akzentstudien I: Germ. got. *þūsundi*. *Indogermanische Forschungen* 6, 344–349.

Hirt, Hermann (1898). Grammatisches und etymologisches. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 23, 288–357.

Hirt, Hermann (1905). *Die Indogermanen: Ihre Verbreitung, ihre Urheimat und ihre Kultur* I. Strassburg: Trübner.

Hirt, Hermann (1907). *Die Indogermanen: Ihre Verbreitung, ihre Urheimat und ihre Kultur* II. Strassburg: Trübner.

Hirt, Hermann (1909). *Etymologie der neuhochdeutschen Sprache*. München: Beck.

Hirt, Hermann (1927). *Indogermanische Grammatik* I. Heidelberg: Winter.

Hogg, Richard M. (1992). *A Grammar of Old English* I. *Phonology*. Oxford: Blackwell.

Holčík, Juraj et al. (1989). *Acipenser sturio* Linnaeus, 1758. In: Juraj Holčík (ed.), *The Freshwater Fishes of Europe*. Vol. 1/2. Wiesbaden: AULA, 367–393.

Holopainen, Sampsa (2019). *Indo-Iranian borrowings in Uralic*. Diss. University of Helsinki.

Holopainen, Sampsa (2021). On the question of substitution of palatovelars in Indo-European loanwords into Uralic. *Suomalais-Ugrilaisen Seuran Aikakauskirja* 98, 197–233.

Holopainen, Sampsa (2022). Uralilaisen lingvistisen paleontologian ongelmia — mitä sanasto voi kertoa kulttuurista? In: Kaisla Kaheinen et al. (eds.). *Hämeenmaalta Jamalille* [FS Salminen]. Helsinki: Helda Open Books, 101–112.

Holopainen, Sampsa & Junttila, Santeri (2022). *Die alten arischen und baltischen Lehnverben der uralischen Sprachen*. Dettelbach: Röll.

Holopainen, Sampsa, Juha Kuokkala & Santeri Junttila (2017). Indoeurópai jövevényszavak és a második szótagi labiális magánhangzók fejlődése az uráli nyelvekben. *A nyelvtörténeti kutatások újabb eredményei* 9, 109–136.

Holthausen, Ferdinand (1899). Engl. *culver* — russ. *gólubĭ* 'Taube'. *Indogermanische Forschungen* 10, 112.

Holthausen, Ferdinand (1934). *Altenglisches etymologisches Wörterbuch*. Heidelberg: Winter.

Holton, David et al. (2020). *The Cambridge Grammar of Medieval and Early Modern Greek*. 4 vols. Cambridge: University Press.

Holub, Josef & Kopečný, František (1952). *Etymologický slovník jazyka českého*. Praha: Státní nakladatelství učebnic.

Holzer, Georg (1989). *Entlehnungen aus einer bisher unbekannten indogermanischen Sprache im Urslavischen und Urbaltischen*. Wien: Akademie der Wissenschaften.

Holzer, Georg (2001). Zur Lautgeschichte des baltisch-slavischen Areals. *Wiener Slavistisches Jahrbuch* 47, 33–50.

Honti, László (1982). *Geschichte des obugrischen Vokalismus der ersten Silbe*. Budapest: Akadémiai Kiadó.

Hoops, Johannes (1905). *Waldbäume und Kulturpflanzen im germanischen Altertum*. Straßburg: Trübner.

Hoops, Johannes (1915–1916, ed.). *Reallexikon der Germanischen Altertumskunde* III. Straßburg: Trübner.

Huld, Martin E. (1981). Albanian corrigenda to Friedrich's Proto-Indo-European trees. *Zeitschrift für vergleichende Sprachforschung* 95, 302–308.

Huld, Martin E. (1984). *Basic Albanian Etymologies*. Columbus: Slavica.

Huld, Martin E. (1990). The linguistic typology of the Old European substrata in North Central Europe. In: Karlene Jones-Bley & Martin E. Huld (eds.), *The Indo-Europeanization of Northern Europe*. Washington, DC: Institute for the Study of Man, 389–423.

Hupel, August W. (1818). *Ehstnische Sprachlehre für die beyden Hauptdialekte, den revalschen und dörptschen, nebst einem vollständigen ehstnischen Wörterbuche*. 2 vols. Mitau: Steffenhagen und Sohn.

Hyllested, Adam (2009). PIE *-*bʰ*- in Nouns and Verbs: Distribution, Function, Origin. In: Rosemarie Lühr & Sabine Ziegler (eds.), *Protolanguage and Prehistory*. Wiesbaden: Reichert, 202–214.

Ильинскій, Григорій А. (1910). Славянскія этимологіи XXI–XXV. *Русскій филологическій вѣстникъ* 63, 322–341.

Ильинскій, Григорій А. (1916). Звукъ ch въ славянскихъ языкахъ II. *Извѣстія Отдѣленія русскаго языка и словесности* 20/4, 135–186.

Ильинский, Григорий А. (1926). К этимологии балтийско-славянских названий „зубра". *Tauta ir žodis* 4, 51–56.

Иллич-Свитыч, Владислав М. (1960). К этимологии слов морковь и тыква. *Этимологические исследования по русскому языку* 1, 16–26.

Иллич-Свитыч, Владислав М. (1963). *Именная акцентуация в балтийском и славянском*. Москва: Академия наук СССР.

Ильминский, Николай И. (1860–1861). Матеріалы къ изученію киргизскаго нарѣчія. *Ученыя записки, изданныя императорскимъ казанскимъ университетомъ* 1860/3–4, 1861/1.

Imberciadori, Giulio (2022). The Etymology of Toch. B *tapre*, A *tpär* 'high' [Conference handout]. *International Etymological Symposium* (Greifswald, 16–18 May 2022).

Ipsen, Gunther (1924). Der Alte Orient und die Indogermanen. In: Johannes Friedrich et al. (eds.), *Stand und Aufgaben der Sprachwissenschaft* [FS Streitberg]. Heidelberg: Winter, 200–237.

Isépy, Peter & Primavesi, Oliver (2014). Helladios und Hesychios. Neues zum Text der *Bibliotheke* des Photios (*Cod.* 279). *Zeitschrift für Papyrologie und Epigraphik* 192, 121–142.

Ишкильдина, Линара К. (2018). Этимология слов с анлаутным «ш» в башкирском языке. *Томский журнал ЛИНГ и АНТР* 22, 31–38.

Itkonen, Erkki (1946). Zur Frage nach der Entwicklung des Vokalismus der ersten Silbe in den finnisch-ugrischen Sprachen, insbesondere im Mordwinischen. *Finnisch-Ugrische Forschungen* 29, 232–337.

Itkonen, Erkki (1948). Vokaalikombinaatiot ja vartalotyypit. *Virittäjä* 52, 124–144.

Itkonen, Erkki (1953). Zur Geschichte des Vokalismus der ersten Silbe im Tscheremissischen und in den permischen Sprachen. *Finnisch-Ugrische Forschungen* 31/3, 149–345.

Itkonen, Erkki (1977). Die Umwandlung einiger *a*- und *ä*-Stämme zu *e*-Stämmen im Urfinnischen. *Suomalais-Ugrilaisen Seuran Aikakauskirja* 75, 5–12.

Itkonen, Terho (1982). *Laaja*, *lavea*, *lakea* ja *laakea*: lisiä suomen *kj*-sanojen vaiheisiin. *Virittäjä* 86, 121–139.

Itkonen, Terho (1987). Erään vokaalivyyhden selvittelyä. Pitkän vokaalin lyhenemistä vai lyhyen pitenemistä *h*-loppuisessa ensi tavussa? *Virittäjä* 91, 164–208.

Itkonen, Terho (1997). Reflections on Pre-Uralic and the "Saami-Finnic protolanguage". *Finnisch-Ugrische Forschungen* 54, 229–266.

Itkonen, Toivo I. (1958). *Koltan- ja kuolanlapin sanakirja* I. Helsinki: Suomalais-Ugrilainen Seura.

Иванов, Вячеслав В. (2002). Ранние коптские заимствования в славянском. In: Татьяна М. Николаева (ed.), *Славянская языковая и этноязыковая системы в контакте с неславянским окружением*. Москва: Языки славянской культуры, 55–60.

Иванов, Вячеслав В. & Топоров, Владимир Н. (1958). *К постановке вопроса о древнейших отношениях балтийских и славянских языков* (IV Международный съезд славистов: Доклады). Москва: Академия наук СССР.

Iversen, Rune & Kroonen, Guus (2017). Talking Neolithic. Linguistic and Archaeological Perspectives on How Indo-European Was Implemented in Southern Scandinavia. *American Journal of Archaeology* 121, 511–525.

Jagersma, Abraham H. (2010). *A descriptive grammar of Sumerian*. Diss. Universiteit Leiden.

Jakob, Anthony (forthc. a.). Three pre-Balto-Slavic bird names, or: A more austere take on Oštir. To appear in: Guus Kroonen (ed.), *Sub-Indo-European Europe*. De Gruyter Mouton.

Jakob, Anthony (forthc. b.). On some potential ghost words in Baltic. To appear in *Baltistica* 58(1).

Jakob, Anthony (forthc. c.). The palaeo-Baltic substrate: a methodological exploration. To appear in the proceedings of *Hansakansan kanssa. Internationales etymologisches Symposium, Greifswald*.

Jakob, Anthony (forthc. d.). Symmetry, sound change and the Baltic loanwords in Finnic. To appear in: Sampsa Holopainen & Juho Pystynen (eds.), *Areal effects*

in prehistoric contacts between Uralic and Indo-European [working title]. Helsinki: Suomalais-Ugrilaisen Seura.

Jamison *Commentary* = Jamison, Stephanie. *Rigveda Translation: Commentary*. Accessed online at rigvedacommentary.alc.ucla.edu.

Janhunen, Juha (1977). *Samojedischer Wortschatz*: *Gemeinsamojedische Etymologien*. Helsinki.

Janko, Josef (1912). *O pravěku slovanském*. Praha: Otto.

Jasanoff, Jay. H. (2002). The nom. sg. of Germanic *n*-stems. In: Alfred Wedel & Hans-Jörg Busch (eds.), *Verba et Litterae: Explorations in Germanic Languages and German Literature*. Newark: Linguatext, 31–46.

Jēgers, Benjamiņš (1969). Einige baltische und slavische Verwandte der Sippe von lit. *dirti*. *Studi Baltici* 10, 63–112.

Jóhannesson, Alexander (1942). Isländische Beiträge zum indogermanischen Wörterbuch. *Zeitschrift für vergleichende Sprachforschung* 67, 220–223.

Joki, Aulis J. (1973). *Uralier und Indogermanen*. Helsinki: Suomalais-Ugrilainen Seura.

Jokl, Norbert (1923). *Linguistisch-kulturhistorische Untersuchungen aus dem Bereiche des Albanischen*. Berlin–Leipzig: De Gruyter.

Jokl, Norbert (1924). Albaner. In: Max Ebert (ed.), *Reallexikon der Vorgeschichte* I. Berlin: De Gruyter, 84–94.

Junttila, Santeri (2011). A collection of forgotten etymologies: revisiting the most improbable Baltic loanwords in Finnic. In: Cornelius Hasselblatt et al. (eds.), *Language Contact in Times of Globalization*. Amsterdam–New York: Rodopi, 77–132.

Junttila, Santeri (2012). The prehistoric context of the oldest contacts between Baltic and Finnic languages. In: Riho Grünthal & Petri Kallio (eds.), *A Linguistic Map of Prehistoric Northern Europe*. Helsinki: Suomalais-Ugrilainen Seura, 261–296.

Junttila, Santeri (2014). Die alten baltischen Lehnwörter im neuen estnischen etymologischen Wörterbuch. *Baltu filoloģija* 23, 126–135.

Junttila, Santeri (2015a). Proto-Finnic loanwords in the Baltic languages? An old hypothesis revisited. In: Santeri Junttila (ed.), *Contacts between the Baltic and Finnic languages*. Helsinki: Suomalais-Ugrilainen Seura, 12–38.

Junttila, Santeri (2015b). Revisiting the 'separate Proto-Balto-Slavic borrowings' of Saami. In: Harry Mantila et al. (eds.), *Congressus duodecimus internationalis Fenno-Ugristarum, Oulu. Book of abstracts*, 476–478.

Junttila, Santeri (2016a). *Tiedon kumuloituminen ja trendit lainasanatutkimuksessa*. Helsinki: Yliopisto.

Junttila, Santeri (2016b). Die baltisch-slawische Frage im Lichte der alten Baltischen Lehnwörter des ostseefinnischen. *Baltistica* 51/2, 217–238.

Junttila, Santeri (2017). Lähtökielen sanansisäisten soinnittomien klusiilien edustus kantasuomen balttilaislainoissa I. *Suomalais-Ugrilaisen Seuran Aikakauskirja* 96, 127–148.

Junttila, Santeri (2018). Altlettgallische Lehnwörter in den mordwinischen Sprachen? *Finnisch-Ugrische Forschungen* 64, 72–91.

Junttila, Santeri (2019). Lähtökielen sanansisäisten soinnittomien klusiilien edustus kantasuomen balttilaislainoissa II. *Suomalais-Ugrilaisen Seuran Aikakauskirja* 97, 35–73.

Junttila, Santeri (in prep.). *Kantasuomen balttilaislainat. Sanakirjan alustava versio keskustelun pohjaksi*. [2nd draft, dated 17 April 2020]. Accessed online at: blogs.helsinki.fi/santerijunttila/luonnokset/

Juška = Juška, Antanas [= А. Юшкевичъ] (1898–1904). *Литовскій словарь съ толкованіемъ словъ на русскомъ и польскомъ языкахъ*. 3 issues [A–K]. Санктпетербургъ: Императорская Академія Наукъ.

Kalima, Jalo (1919). *Die ostseefinnischen Lehnwörter im Russischen*. Hesinki: Société Finno-Ougrienne.

Kalima, Jalo (1924). Zur karelisch-olonetzischen lautgeschichte. *Finnisch-Ugrische Forschungen* 24, 163–176.

Kalima, Jalo (1928). Eräiden sanojen alkuperästä. *Virittäjä* 32, 102–111.

Kalima, Jalo (1929). Zur Herkunft der slavischen Lehnwörter im Ostseefinnischen. *Zeitschrift für Slavische Philologie* 6, 154–172.

Kalima, Jalo (1934). Russ. *стержень* und slav. *stěna*, d. *Stein*. In: *Germanisch-romanische Studien* [FS Suolahti]. Helsinki: Suomalainen tiedeakatemia.

Kalima, Jalo (1936). *Itämerensuomalaisten kielten balttilaiset lainasanat*. Helsinki: Suomalaisen kirjallisuuden seura.

Kalima, Jalo (1939–1940). Fi. *ativo*. *Finnisch-Ugrische Forschungen* 26, 211–214.

Kalima, Jalo (1950). Zur slavischen Etymologie. *Zeitschrift für Slavische Philologie* 20, 406–417.

Kalima, Jalo (1956). *Die slavischen Lehnwörter im Ostseefinnischen*. Berlin–Wiesbaden: Harrassowitz.

Kallio, Petri (1998). Vanhojen balttilaisten lainasanojen ajoittamisesta. In: Riho Grünthal & Johanna Laakso (eds.), *Oekeeta asijoo* [FS Suhonen]. Helsinki: Suomalais-Ugrilainen Seura, 209–217.

Kallio, Petri (2006). On the Earliest Slavic Loanwords in Finnic. In: Juhani Nuorluoto (ed.), *The Slavicization of the Russian North: Mechanisms and Chronology*. Helsinki: Yliopisto, 154–166.

Kallio, Petri (2007). Kantasuomen konsonanttihistoriaa. In: Jussi Ylikoski & Ante Aikio (eds.), *Sámit, sánit, sátnehámit* [FS Sammallahti]. Helsinki: Suomalais-Ugrilainen Seura, 229–249.

Kallio, Petri (2008a). On the "early Baltic" loanwords in common Finnic. In: Alexander Lubotsky et al. (eds.), *Evidence and counter-evidence: Essays in honour of Frederik Kortlandt* I. *Balto-Slavic and Indo-European linguistics*. Amsterdam–New York: Brill, 265–277.

Kallio, Petri (2008b). The Etymology of Finnish *sauna* 'Sauna'. In: Kees Dekker, Alasdair MacDonald & Hermann Niebaum (eds.), *Northern Voices* [FS Hofstra]. Leuven–Paris–Dudley: Peeters, 314–319.

Kallio, Petri (2009). Stratigraphy of Indo-European loanwords in Saami. In: Tiina Äikäs (ed.), *Máttut–máddagat: the Roots of Saami ethnicities, societies and spaces / places*. Oulu: Giellagas Institute, 30–45.

Kallio, Petri (2012). Jälkitavujen diftongit kantasuomessa. *Fenno-Ugrica Suecana Nova Series* 14, 31–40.

Kallio, Petri (2014). The Diversification of Proto-Finnic. In: Joonas Ahola & Frog (eds.), *Fibula, Fabula, Fact: The Viking Age in Finland*. Helsinki: Suomen Kirjallisuuden Seura, 155–168.

Kallio, Petri (2015). The language contact situation in prehistoric Northeastern Europe. In: Robert Mailhammer, Theo Vennemann & Birgit A. Olsen (eds.), *The Linguistic Roots of Europe: Origin and Development of European Languages*. København: Museum Tusculanum, 77–102.

Kallio, Petri (2016). Historical phonology from Proto-Finnic to Proto-Livonian. *Journal of Estonian and Finno-Ugric Linguistics* 7/1, 39–65.

Kallio, Petri (2018). Ensitavun diftongit kantasuomessa. In: Sampsa Holopainen & Janne Saarikivi (eds.), Περὶ ὀρθότητος ἐτύμων: *Uusiutuva uralilainen etymologia*. Helsinki: Suomalais-Ugrilainen Seura, 251–268.

Kallio, Petri (2021). The Position of Leivu. In: Uldis Balodis & Karl Pajusalu (eds.), *Studies on the South Estonian language islands* (*Leivu, Lutsi, Kraasna*). Tartu: Ölikool, 123–143.

Kallio, Petri (forthc.). Substrates in Finnic. To appear in: Ante Aikio & Santeri Palviainen (eds.), *Substrate Languages in Northern Europe*. De Gruyter Mouton.

Kapović, Mate (2006). *Reconstruction of Balto-Slavic personal pronouns*. Diss. University of Zagreb.

Karaliūnas, Simas (1993). Some remarks on the names for pigeon in Lithuanian. *Linguistica Baltica* 2, 109–113.

Karaliūnas, Simas (2004). *Baltų praeitis istoriniuose šaltiniuose* I. Vilnius: Lietuvių kalbos institutas.

Karg, Sabine & Robinson, David (2000). Secondary food plants from medieval sites in Denmark: fruits, nuts, vegetables, herbs and spices. In: Karin Viklund (ed.), *Nordic Archaeobotany–NAG 2000 in Umeå* (= *Archaeology and Environment* 15). University of Umeå, 133–142.

Karničar, Ludwig (1990). *Der Obir-Dialekt in Kärnten: die Mundart von Ebriach/Obirsko im Vergleich mit den Nachbarmundarten von Zell/Sele und Trögern/Korte*. Wien: Akademie der Wissenschaften.

Karsten, Torsten E. (1915). *Germanisch-finnische Lehnwortstudien*. Helsinki: Finnische Literaturgesellschaft.

Karulis, Konstantīns (1992). *Latviešu etimoloģijas vārdnīca*. 2 vols. Rīga: Avots.

Касаткин, Леонид Л. (1999). *Современная русская диалектная и литературная фонетика как источник для истории русского языка*. Москва: Наука.

Касьян, Алексей С. (2012). Отражение праславянской фонемы **y* в старопсковском диалекте разговорника Т. Фенне как архаизм. *Вопросы языкознания* 2012/4, 73–100.

Kaukienė, Audronė & Jakulytė, Dalia (2015). *Stonais po leipen zaidiantẽ*: kitokia XV a. baltiško teksto interpretacija. *Res humanitariae* 17, 40–52.

Kazlauskas, Jonas (1968). *Lietuvių kalbos istorinė gramatika*. Vilnius: Mintis.

Kendla, Mari & Viikberg, Jüri (2015). Protoeurooplaste keelepärandist. *Emakeele Seltsi aastaraamat* 61, 135–154.

Keresztes, László (1986). *Geschichte des modwinischen Konsonantismus* II. *Etymologisches Belegmaterial*. Szeged.

Keresztes, László (1987). *Geschichte des mordwinischen Konsonantismus* I. Szeged.

Kettunen, Lauri (1930). *Vatjan kielen äännehistoria*. Helsinki: Suomalaisen Kirjallisuuden Seura.

Kettunen, Lauri (1938). *Livisches Wörterbuch mit grammatischer Einleitung*. Helsinki: Suomalais-Ugrilainen Seura.

Kießling, Roland & Mous, Maarten (2003). *The Lexical Reconstruction of West-Rift Southern Cushitic*. Köln: Rüdiger Köppe.

Kilian, Lothar (1986). Zur Frage eines westfinnischen Substrats in Litauen aus der Sicht von Archäologie und Hydronymie. *Zeitschrift für Ostforschung* 35, 481–502.

Kim, Ronald I. (2018). The phonology of Balto-Slavic. In: Jared Klein, Brian Joseph & Matthias Fritz (eds.), *Handbook of Comparative and Historical Indo-European Linguistics*. Berlin: De Gruyter, 1974–1985.

Kiparsky, Valentin (1934). *Die gemeinslavischen Lehnwörter aus dem Germanischen*. Helsinki: Suomalainen Tiedeakatemia.

Kiparsky, Valentin (1936). *Fremdes im Baltendeutsch*. Helsinki: Société néophilologique.

Kiparsky, Valentin (1948). Chronologie des relations slavobaltiques et slavofinnoises. *Revue des Études Slaves* 24, 29–47.

Kiparsky, Valentin (1949). *Tuhkuri* 'vesikko'. *Virittäjä* 53, 58–68.

Kiparsky, Valentin (1952). The Earliest Contacts of the Russians with the Finns and Balts. *Oxford Slavonic Papers* 3, 67–79

Kiparsky, Valentin (1959a). Kielitieteen peruskysymyksiä. *Virittäjä* 63, 421–424.

Kiparsky, Valentin (1959b). Über etymologische Wörterbücher. *Neuphilologische Mitteilungen* 60, 209–230.

Kiparsky, Valentin (1963). *Russische historische Grammatik* I. *Entwicklung des Lautsystems*. Heidelberg: Winter.

Kiparsky, Valentin (1968). Slavische und baltische b/p-Fälle. *Scando-Slavica* 14, 73–97.

Kiparsky, Valentin (1972). Zur Etymologie des lit. *šeškas*, lett. *sesks* 'Iltis'. *Baltistica* 8, 157–160.

Kiparsky, Valentin (1973). О балтизмах русского литературного языка. *Baltistica* 9, 67–70.

Kiparsky, Valentin (1975). *Russische historische Grammatik* III. *Entwicklung des Wortschatzes*. Heidelberg: Winter.

Kitson, Peter R. (1997). Old English bird-names. *English Studies* 78, 481–505.

Kittel, Piotr et al. (2020). On the border between land and water: The environmental conditions of the Neolithic occupation from 4.3 until 1.6 ka BC at Serteya, Western Russia. *Geoarchaeology* 36, 173–202.

Kleyner, Svetlana D. (2005). From Yellow to Blue — or Not? In: Dafydd Johnston, Elena Parina & Maxim Fomin (eds.), *'Yn llawen iawn, yn llawn* iaith'. Aberystwyth: Centre for Advanced Welsh and Celtic Studies, 47–52.

Klimas, Antanas (1957). Lithuanian and Indo-European. *Lituanus* 1957/4, 14–16.

Klimas, Antanas (2002). Lithuanian in the 21st Century: Linguistic Snippets and Tidbits. *Lituanus* 48/3, 52–80.

Klingenschmitt, Gert (1982). *Das altarmenische Verbum*. Wiesbaden: Reichert.

Kloekhorst, Alwin (2008). *Etymological Dictionary of the Hittite Inherited Lexicon*. Leiden–Boston: Brill.

Kloekhorst, Alwin (2014). The Proto-Indo-European Acrostatic Inflection Reconsidered. In: Norbert Oettinger & Thomas Steer (eds.), *Das Nomen im Indogermanischen*. Wiesbaden: Reichert, 140–163.

Kluge, Friedrich (1907). Zum altpreußischen Wortschatz. *Indogermanische Forschungen* 21, 358–361.

Kluge/Götze = Alfred Götze (1948, ed.), *Friedrich Kluge. Etymologisches Wörterbuch der deutschen Sprache* [14th edition]. Berlin: De Gruyter.

Kluge/Seebold = Elmar Seebold (2011, ed.). *Kluge: Etymologisches Wörterbuch der deutschen Sprache* [25th edition]. Berlin–New York: De Gruyter.

Knipper, Corina et al. (2017). Female exogamy and gene pool diversification at the transition from the Final Neolithic to the Early Bronze Age in central Europe. *Proceedings of the National Academy of Sciences* 114/38.

Koivulehto, Jorma (1970). Suomen *laiva*-sanasta. *Virittäjä* 74, 178–183.

Koivulehto, Jorma (1971). Germanisch-finnische Lehnbeziehungen I. Finn. Vordervokal für germ. Hintervokal. *Neuphilologische Mitteilungen* 72, 577–607.

Koivulehto, Jorma (1972). Germanisch-finnische Lehnbeziehungen II. Wirtschaftliches und Landwirtschaftliches. *Neuphilologische Mitteilungen* 73, 575–628.

Koivulehto, Jorma (1979a). Lainoja ja lainakerrostumia. *Virittäjä* 83, 267–301.

Koivulehto, Jorma (1979b). Baltisches und Germanisches im Finnischen: die finn. Stämme auf *-rte* und die finn. Sequenz *VrtV*. In: Erhard F. Schiefer (ed.), *Explanationes et tractationes Fenno-Ugricae in honorem Hans Fromm*. München: Fink, 129–164.

Koivulehto, Jorma (1981). Reflexe des germ. /ē¹/ im Finnischen und die Datierung der

germanisch-finnischen Lehnbeziehungen. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 103, 167–203.

Koivulehto, Jorma (1983). Suomalaisten maahanmuutto indoeurooppalaisten lainasanojen valossa. *Suomalais-Ugrilaisen Seuran Aikakauskirja* 78, 107–132.

Koivulehto, Jorma (1984). Itämerensuomalais-germaaniset kosketukset. In: Sven-Erik Åström (ed.), *Suomen väestön esihistorialliset juuret*. Helsinki: Societas Scientiarum Fennica, 191–204.

Koivulehto, Jorma (1986). Die Sieverssche Regel im Lichte der germanisch-finnischen Lehnbeziehungen. In: Bela Brogyanyi & Thomas Krömmelbein (eds.), *Germanic Dialects: Linguistic and Philological Investigations*. Amsterdam–Philadelphia: John Benjamins, 249–294.

Koivulehto, Jorma (1990). Zu den ältesten Kontakten zwischen Ostseefinnisch und Balto-Slavisch. In: Sari Vaula (ed.), *Itämerensuomalaiset kielikontaktit*. Helsinki: Valtion painatuskeskus, 148–153.

Koivulehto, Jorma (1992a). Direkte Kontakte des Lappischen mit dem Baltischen. In: Pál Deréky et al. (eds.), *Festschrift für Károly Rédei zum 60. Geburtstag*. Wien–Budapest: MTA Nyelvtudományi Intézet, 299–304.

Koivulehto, Jorma (1992b). Der Typus *palje* '(Blase)balg', *turve* 'Torf' unter den Lehnwörtern des Ostseefinnischen. *Suomalais-Ugrilaisen Seuran Aikakauskirja* 84, 163–190.

Koivulehto, Jorma (1993a). Zur Etymologie von fi. *susi*, *tosi* — und *kesi*: eine Entgegnung. *Linguistica Uralica* 29/1, 21–36.

Koivulehto, Jorma (1993b). Suomi. *Virittäjä* 97, 400–406.

Koivulehto, Jorma (1998). *Kiire 'päälaki' ja muuta etymologista rajankäyntiä*. In: Riho Grünthal & Johanna Laakso (eds.), *Oekeeta asijoo* [FS Suhonen]. Helsinki: Suomalais-Ugrilainen Seura, 235–246.

Koivulehto, Jorma (1999). *Verba mutuata*. Helsinki: Suomalais-Ugrilainen Seura.

Koivulehto, Jorma (2000). Reflexe des urbaltischen **ā* in baltischen Lehnwörtern des Ostseefinnischen. *Linguistica Baltica* 8, 103–124.

Koivulehto, Jorma (2001). The earliest contacts between Indo-European and Uralic speakers in the light of lexical loans. In: Christian Carpelan, Asko Parpola & Petteri Koskikallio (eds.), *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations*. Helsinki: Suomalais-Ugrilainen Seura, 235–263.

Koivulehto, Jorma (2002). Contact with non-Germanic languages II: Relations to the East. In: Oskar Bandle et al. (eds.), *The Nordic Languages* I. Berlin: De Gruyter Mouton, 583–594.

Колесов, Владимир В. (1980). *Историческая фонетика русского языка*. Москва: Высшая школа.

Kortlandt, Frederik (1977). Historical laws of Baltic accentuation. *Baltistica* 13, 319–330.

Kortlandt, Frederik (1978). I.-E. palatovelars before resonants in Balto-Slavic. In: Jacek Fisiak (ed.), *Recent developments in historical phonology*. Den Haag: Mouton, 237–243.

Kortlandt, Frederik (1979a). Three problems of Balto-Slavic phonology. *Zbornik za Filologiju i Lingvistiku* 22/2, 57–63.

Kortlandt, Frederik (1979b). On the history of the Slavic nasal vowels. *Indogermanische Forschungen* 84, 259–272.

Kortlandt, Frederik (1980). [Review of] V.M. Illich-Svitych, Nominal accentuation in Baltic and Slavic. *Lingua* 51, 346–354.

Kortlandt, Frederik (1982). IE **pt* in Slavic. *Folia Linguistica Historica* 3, 25–28.

Kortlandt, Frederik (1983). On final syllables in Slavic. *Journal of Indo-European Studies* 11, 167–185.

Kortlandt, Frederik (1984–1985). On reduced vowels in Slavic. *Zbornik Matice srpske za filologiju i lingvistiku* 27–28, 367–368.

Kortlandt, Frederik (1985). Long vowels in Balto-Slavic. *Baltistica* 21, 112–124.

Kortlandt, Frederik (1988). Remarks on Winter's law. In: Adriaan A. Barentsen, Ben M. Groen & Rob Sprenger (eds.), *Dutch Contributions to the Tenth International Congress of Slavists*: *Linguistics*. Amsterdam: Rodopi, 387–396.

Kortlandt, Frederik (1989a). Od praindoevropskog jezika do slovenskog (fonološki razvoj). *Zbornik za filologiju i lingvistiku* 32/2, 41–58.

Kortlandt, Frederik (1989b). Lithuanian *statýti* and related formations. *Baltistica* 25/2, 104–112.

Kortlandt, Frederik (1993). Tokie šalti rytai. *Baltistica* 28, 45–48.

Kortlandt, Frederik (1995). Lithuanian verbs in *-auti* and *-uoti*. *Linguistica Baltica* 4, 141–143.

Kortlandt, Frederik (1998). The Language of the Old Prussian Catechisms. *Res Balticae*, 117–129.

Kortlandt, Frederik (2000). Initial *a-* and *e-* in Old Prussian. *Linguistica Baltica* 8, 125–127.

Kortlandt, Frederik (2007). The development of the Indo-European syllabic resonants in Balto-Slavic. *Baltistica* 42, 7–12.

Kortlandt, Frederik (2018). Proto-Baltic? *Baltistica* 53, 175–185.

Kossmann, Maarten (1999). *Essai sur la phonologie du proto-berbère*. Köln: Köppe.

Kossmann, Maarten (2020). Proto-Berber phonological reconstruction: An update. *Linguistique et Langues Africaines* 6, 11–42.

Kott = Kott, František Št. (1878–1893). *Česko-německý slovník zvláště gramaticko-frazeologický*. 7 vols. Praha: Kolář/Simáček.

Krahe, Hans (1962). Die Struktur der alteuropäischen Hydronomie. *Abhandlungen der geistes- und sozialwissenschaftlichen Klasse* 1962/5.

Krajčovič, Rudolf (1975). *A Historical Phonology of the Slovak Language*. Heidelberg: Winter.

Kregždys, Rolandas (2004). Liet. *barzdà* 'barba': samplaikos *-zd-* kilmės klausimu. *Acta Linguistica Lithuanica* 51, 15–27.

Kregždys, Rolandas (2012). *Baltų mitologemų etimologijos žodynas* I. *Kristburgo sutartis*. Vilnius: Lietuvos kultūros tyrimų institutas.

Кренке, Николай А. (2019). *Древности бассейна москвы-реки от неолита до средневековья*. Москва–Смоленск: Свиток.

Kretschmer, Paul (1896). *Einleitung in die Geschichte der griechischen Sprache*. Göttingen: Vandenhoeck & Ruprecht.

Kretschmer, Paul (1921). Der Götterbeiname Grabovius auf den Tafeln von Iguvium. In: *Festschrift Adalbert Bezzenberger*. Göttingen: Vandenhoeck & Ruprecht, 89–96.

Kretschmer, Paul (1940). Die vorgriechischen Sprach- und Volksschichten. *Glotta* 28, 231–278.

Kroll, Helmut (1990). Melde von Feudvar, Vojvodina: Ein Massenfund bestätigt *Chenopodium* als Nutzpflanze in der Vorgeschichte. *Prähistorische Zeitschrift* 65, 46–48.

Kroonen, Guus J. (2011a). False exceptions to Winter's law. In: Rick Derksen & Tijmen Pronk (eds.), *Accent Matters: Papers on Balto-Slavic Accentology*. Amsterdam–New York: Rodopi–Brill, 251–261.

Kroonen, Guus J. (2011b). *The Proto-Germanic n-stems: a study in diachronic morphology*. Amsterdam: Rodopi.

Kroonen, Guus J. (2012). Non-Indo-European root nouns in Germanic: evidence in support of the Agricultural Substrate Hypothesis. In: Riho Grünthal & Petri Kallio (eds.), *A Linguistic Map of Prehistoric Northern Europe*. Helsinki: Suomalais-Ugrilainen Seura, 239–260.

Kroonen, Guus J. (2013). *Etymological Dictionary of Proto-Germanic*. Leiden–Boston: Brill.

Kroonen, Guus J. (2016). On the Origin of Greek μῆλον, Latin *mālum*, Albanian *mollë* and Hittite *šam(a)lu-* 'apple'. *Journal of Indo-European Studies* 44, 85–91.

Kroonen, Guus J. (2017). The development of the Proto-Indo-European instrumental suffix in Germanic. *Indogermanische Forschungen* 122, 105–110.

Kroonen, Guus J. (forthc.). For the nth time: The Pre-Greek νθ-suffix revisited. To appear in: Guus Kroonen (ed.), *Sub-Indo-European Europe*. De Gruyter Mouton.

Kroonen, Guus J. & Lubotsky, Alexander (2009). Proto-Indo-European **tsel-* 'to sneak' and Germanic **stelan-* 'to steal, approach stealthily'. *Studia Etymologica Cracoviensia* 14, 237–241.

Kroonen, Guus J., Anthony Jakob et al. (2022). Indo-European cereal terminology suggests a Northwest Pontic homeland for the core Indo-European languages. *PLoS ONE* 17/10.

Крысько, Вадим Б. (1994). Древний новгородско-псковский диалект на общеславянском фоне. *Вопросы языкознания* 1994/6, 16–93.

Крысько, Вадим Б. (2014). Маргиналии к «Этимологическому словарю славянских языков» (вып. 34–38). *Вопросы языкознания*. 2014/1, 100–119.

Kuhn, Adalbert (1862). [Review of] Adolphe Pictet. Les origines Indo-Europennes ou les Aryas primitifs I. *Beiträge zur vergleichenden Sprachforschung* 2, 369–382.

Kuiper, Frans B.J. (1956). The Etymology of ἄνθρωπος. In: Heinz Kronasser (ed.), ΜΝΗΜΗΣ ΧΑΡΙΝ: *Gedenkschrift Paul Kretschmer* I. Wien: Sprachgesellschaft, 211–226.

Kuiper, Frans B.J. (1968). Pre-Hellenic Labio-Velars? *Lingua* 21, 269–277.

Kuiper, Frans B.J. (1995). Gothic *bagms* and Old Icelandic *ylgr*. *North-Western European Language Evolution* 25, 63–88.

Кулешов, Вячеслав С. (2010). Вокализм ранних славянских заимствований в прибалтийско-финских языках и реконструкция вокалической системы северно-словенского диалекта. *Finnisch-Ugrische Mitteilungen* 32/33, 345–354.

Куликовскій, Германъ И. (1898). *Словарь областного олонецкаго нарѣчія*. Санктъ-Петербургъ: Императорская академія наукъ.

Kulonen, Ulla-Maija (1988). Ovatko *reuna* ja *leuka* sittenkin balttilaisia lainasanoja? *Virittäjä* 92, 84–87.

Kümmel, Martin (2000). *Das Perfekt im Indoiranischen*. Wiesbaden: Reichert.

Kümmel, Martin (2017a). Even more traces of the accent in Armenian? In: Bjarne S.S. Hansen, Adam Hyllested, Anders R. Jørgensen & Guus Kroonen (eds.), *Usque ad radices* [FS Olsen]. København: Museum Tusculanum.

Kümmel, Martin (2017b). Agricultural terms in Indo-Iranian. In: Martine Robbeets & Alexander Savelyev (eds.), *Language Dispersal Beyond Farming*. Amsterdam–Philadelphia: Benjamins.

Kuokkala, Juha (2012). *Jälkitavun labiaalivokaalit saamessa ja itämerensuomessa*. MA Thesis Helsingin yliopisto.

Kurila, Laurynas (2021). Lithuania in the Migration Period — Quiet Refuge or Battlefield? *Lithuanian Heritage* January/February 2021, 17–21.

Kurschat, Friedrich (1883). *Wörterbuch der littauischen Sprache* II. *Littauisch-deutsches Wörterbuch*. Halle: Buchandlung des Waisenhauses.

Kuryłowicz, Jerzy (1956a). *L'apophonie en indo-européen*. Wrocław: Ossolineum.

Kuryłowicz, Jerzy (1956b). Die Doppelvertretung von idg. *ei*, *oi* im Lettolitauischen. In: Heinz Kronasser (ed.), ΜΝΗΜΗΣ ΧΑΡΙΝ: *Gedenkschrift Paul Kretschmer* I. Wien: Sprachgesellschaft, 227–236.

Кузьменко, Юрий К. (2013). К вопросу о неиндоевропейских субстратах в германском. *Индоевропейское языкознание и классическая филология* 17, 511–535.

Lacey, Mohamed Eric R. (2013). *Birds and Bird-lore in the Literature of Anglo-Saxon England*. Diss. University College London.

Lamprecht, Arnošt, Dušan Šlosar & Jaroslav Bauer (1977). *Historický vývoj češtiny*. Praha: Státní pedagogické nakladatelství.

Lang, Valter (2007). *The Bronze and Early Iron Ages in Estonia*. Tartu: Ölikool.

Lang, Valter (2015). Formation of Proto-Finnic — an archaeological scenario from the Bronze Age / Early Iron Age. In: Harri Mantila et al. (eds.), *Congressus Duodecimus Internationalis Fenno-Ugristarum. Plenary Papers*. Oulu: Yliopisto, 63–84.

Lang, Valter (2016). Early Finnic-Baltic contacts as evidenced by archaeological and linguistic data. *Journal of Estonian and Finno-Ugric Linguistics* 7, 11–38.

Lang, Valter (2018a). Fortified Settlements in the Eastern Baltic: From Earlier Research to New Interpretations. *Achaeologia Lituana* 19, 13–33.

Lang, Valter (2018b). The Eastern Baltic. In: Colin Haselgrove et al. (eds.), *The Oxford Handbook of the European Iron Age*. Oxford: University Press, 173–192.

Lange, Jacob (1772–1773). *Vollständiges deutschlettisches und lettischdeutsches Lexicon*. 2 vols. Mitau: Steffenhagen.

Lanszweert, René (1984). *Die Rekonstruktion des baltischen Grundwortschatzes*. Frankfurt am Main–Bern–New York: Peter Lang.

Ларин, Борис А. (1959). *Русско-английский словарь-дневник Ричарда Джемса (1618–1619 гг.)*. Ленинград: Университет

Larsson, Jenny H. (2018). The lexicon of Baltic. In: Jared Klein, Brian Joseph & Matthias Fritz (eds.), *Handbook of Comparative and Historical Indo-European Linguistics*. Berlin: De Gruyter, 1681–1698.

Лаучюте, Юрате А. (1982). *Словарь балтизмов в славянских языках*. Ленинград: Наука.

Laumane, Benita (1973). *Zivju nosaukumi latviešu valodā*. Rīga: Zinātne.

Лаумане, Бенита [= Laumane, Benita] & Непокупный, Анатолий П. (1968). К постановке вопроса об атласе морской и рыболовецкой лексики балтийского моря. *Baltistica* 4, 69–78.

Legzdiņa, Dardega et al. (2020). Re-evaluating the Bronze and Earliest Iron Age in Latvia: changes in burial traditions in the light of ^{14}C dates. *Radiocarbon* 62, 1845–1868.

Lehmann, Winfred P. (1951). The Distribution of Proto-Indo-European /r/. *Language* 27, 13–17.

Lehmann, Winfred P. (1986). *A Gothic Etymological Dictionary*. Leiden: Brill.

Lehtinen, Meri (1967). On the Origin of the Balto-Finnic Long Vowels. *Ural-altaische Jahrbücher* 39, 147–152.

Lehtiranta, Juhani (2001). *Yhteissaamelainen sanasto*. Helsinki: Suomalais-Ugrilainen Seura.

Lejeune, Michel (1974). *Manuel de la langue vénète*. Heidelberg: Winter.

Лемцюгова, Валянціна П. (1970). *Беларуская айканімія: лінгвістычны аналіз назваў населеных пунктаў Мінскай вобласці*. Мінск: Навука і тэхніка.

Lemeškin, Ilja (2014). Petrus Wickerau vs. Petrus Turnau. Kretos pėdsako autorystės klausimu. *Baltistica* 49, 139–161.

Leskien, August (1891). Die Bildung der Nomina im Litauischen. *Abhandlungen der philologisch-historischen Classe der königlich sächsischen Gesellschaft der Wissenschaften* 12, 151–618.

Leskien, August (1922). *Handbuch der altbulgarischen Sprache* [6th edition]. Heidelberg: Winter.

Leumann, Manu (1942). Idg. *sk̂* im Altindischen und im Litauischen (Schluß.). *Indogermanische Forschungen* 58, 113–130.

Leumann, Manu (1977). *Lateinische Grammatik* I. *Lateinische Laut- und Formenlehre*. München: Beck.

Левичкин, Александр Н. & Мызников, Сергей А. (2014). *Словарь пинежских говоров*: *Проект*. *Пробные словарные статьи*. Санкт-Петерург: Нестор-История.

Levin, Jules F. (1973). *-jā* Stems and *-ē* Stems in the Elbing Vocabulary. In: Arvids Ziedonis, Jr. (ed.), *Baltic Literature and Linguistics*. Toronto: University, 189–196.

Levin, Jules F. (1974). *The Slavic Element in the Old Prussian Elbing Vocabulary*. Berkeley: University of California Press.

Levin, Jules F. (1992). Pigeons, cows and April in Lithuania. *Linguistica Baltica* 1, 85–91.

Levin, Jules F. (2003). The North Slavic-Lithuanian contact area: mutual influence and resistance. In: Robert A. Maguire & Alan Timberlake (eds.), *American Contributions to the 13th International Congress of Slavists* I. *Linguistics*. Bloomington: Slavica, 139–148.

Lewis/Short = Lewis, Charlton T. & Short, Charles (1879, eds.), *A Latin Dictionary*. Oxford: Clarendon Press.

Librado, Pablo et al. (2021). The origins and spread of domestic horses from the Western Eurasian steppes. *Nature* 598, 634–640.

Lidén, Evald (1897). *Studien zur altindischen und vergleichenden Sprachgeschichte*. Upsala: Almqvist & Wiksell.

Lidén, Evald (1906). Ein gotisches Lehnwort im Altpreussischen. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 31, 600–602.

Liewehr, Ferdinand (1954). Zur Ausdrucksverstärkung im Slawischen. *Zeitschrift für slavische Philologie* 23, 89–115.

Liewehr, Ferdinand (1956). Über expressive Sprachmittel im Slawischen. *Zeitschrift für slavische Philologie* 25, 11–27.

Лигорио, Орсат (2012). С.-х. *рода*, лат. *ardea*. *Зборник Матице српске за филологију и лингвистику* 55, 21–28.

van Linde, Paul (2001). Borrowed or inherited? *Congressus Nonus Internationalis Fenno-Ugristarum* V. *Dissertationes sectionum*: *Linguistica* II. Tartu, 288–293.

van Linde, Paul (2007). *The Finnic vocabulary against the background of interference*. Diss. Universiteit Groningen.

Liukkonen, Kari (1973). Lisiä balttilais-suomalaisten lainasanasuhteiden tutkimukseen. *Virittäjä* 77, 17–32.

Liukkonen, Kari [= Лиукконен, Кари] (1987). *Восточнославянские отглагольные существительные на* -m-. Helsinki.

Liukkonen, Kari (1999). *Baltisches im Finnischen*. Helsinki: Finnisch-Ugrische Gesellschaft.

Lockwood, William B. (1990). Latin *columba*, *palumbēs*, Greek *κόλυμβος*. *Historische Sprachforschung* 103, 261–263.

Loewenthal, Wilhelm (1901). *Die slavischen Farbenbezeichnungen*. Leipzig: Preis.

Loma, Aleksandar (2012). Vom Indogermanischen zum Slavischen und zurück. Probleme und Perspektiven einer vertiefenden Wortforschung im Bereich der slavischen Sprachen. In: Bohumil Vykypěl & Vít Boček (eds.), *Methods of Etymological Practice*. Praha: Lidové noviny.

Lõo, Jaan (1911). Meie koduloomade nimetused. *Eesti kirjandus* 6, 49–61, 81–89.

Łoś, Jan (1922). *Gramatyka polska* I. *Głosownia historyczna*. Lwów–Warszawa–Kraków: Ossolineum.

Lõugas, Lembi, Aivar Kriiska & Lina Maldre (2007). New dates for the Late Neolithic Corded Ware Culture burials and early animal husbandry in the East Baltic region. *Archaeofauna* 16, 21–31.

Lubotsky, Alexander (1981). On the reduced grade *-a-* in Sanskrit. *Lingua* 55, 75–95.

Lubotsky, Alexander (1988). *The system of nominal accentuation in Sanskrit and Proto-Indo-European*. Leiden–New York–København–Köln: Brill.

Lubotsky, Alexander (1990). La loi de Brugmann et **H₃e-*. In: Jean Kellens (ed.), *La reconstruction des laryngeals*. Liège–Paris: Université de Liège, 129–136.

Lubotsky, Alexander (1994). Avestan *ϑβōrəštar-* and the Indo-European root *√turḱ-*. *Die Sprache* 36, 94–102.

Lubotsky, Alexander (1995). Sanskrit *h* < **dh, bh*. In: Ярослав В. Васильков & Никита В. Гуров (eds.), *Стхапакашраддха. Сборник статей памяти Г. А. Зографа*. Санкт-Петербург: Петербургское востоковедение, 124–144.

Lubotsky, Alexander (1997). The Indo-Iranian reflexes of PIE **CRHUV*. In: Alexander Lubotsky (ed.), *Sound Law and Analogy* [FS Beekes]. Amsterdam–Atlanta: Rodopi.

Lubotsky, Alexander (2001). The Indo-Iranian substratum. In: Christian Carpelan, Asko Parpola & Petteri Koskikallio (eds.), *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations*. Helsinki: Suomalais-Ugrilainen Seura, 301–317.

Lubotsky, Alexander (2004). Avestan *siiazd-*, Sanskrit *sedh-*, Latin *cēdere*. In: Adam Hyllested et al. (eds.), *Per aspera ad asteriscos* [FS Rasmussen]. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft, 322–332.

Lubotsky, Alexander (2011). The origin of Sanskrit roots of the type *sīv-* 'to sew', *dīv-* 'to play dice', with an Appendix on Vedic *i*-Perfects. In: Stephanie W. Jamison, H. Craig Melchert & Brent Vine (eds.), *Proceedings of the 22nd UCLA Annual Indo-European Conference*. Bremen: Hempen, 105–126.

Lubotsky, Alexander (2013). False Labels in Indo-European Reconstruction: Laryngeal Loss in Compounds and Marginal Phonemes [Conference handout]. *Comparative-Historical Linguistics of the 21st century: Issues and Perspectives* (Moscow, March 20–22, 2013).

Lühr, Rosemarie (1993). Zur Semantifizierung von Zahlwörtern: das Wort 'Tausend' — eine germanisch-baltoslavische Isoglosse? *Linguistica* 33, 117–136.

Luik, Heidi & Maldre, Lina (2007). Bronze Age bone artefacts from Narkūnai, Nevieriškė and Kereliai fortified settlements. Raw materials and manufacturing technology. *Archaeologia Lituana* 8, 5–39.

Lunt, Horace G. (2001). *Old Church Slavonic Grammar* [7th edition]. Berlin–New York: Mouton de Gruyter.

Лыткин, Василий И. (1952). *Древнепермский язык*. Москва: Академия наук СССР.

Лыткин, Василий И. (1964). *Исторический вокализм пермских языков*. Москва: Наука.

Лыткин, Василий И. & Гуляев, Евгений С. (1970). *Краткий этимологический словарь коми языка*. Москва: Наука.

Mac an Bhaird, Alan (1980). Tadhg mac Cein and the Badgers. *Ériu* 31, 150–155.

Machek, Václav (1924). Slov. *klъkъ*. *Listy Filologické* 51, 125–131.

Machek, Václav (1930). Slovanské jméno piva *olъ*. *Slavia* 8, 209–217.

Machek, Václav (1934). *Recherches dans le domaine du lexique Balto-Slave*. Brno: Filosofická fakulta.

Machek, Václav (1944–1946). Drobné výklady o jménech rostlin. *Naše řeč* 28, 106–109, 175–180; 29, 107–114; 30, 61–68.

Machek, Václav (1947). Einige slavische Fischnamen. *Zeitschrift für Slavische Philologie* 19, 53–67.

Machek, Václav (1950a). Einige slavische Vogelnamen. *Zeitschrift für Slavische Philologie* 20, 29–51.

Machek, Václav (1950b). Quelques noms slaves de plantes. *Lingua Posnaniensis* 2, 145–161.

Machek, Václav (1951a). Quelques mots slavo-germaniques. *Slavia* 20, 200–218.

Machek, Václav (1951b). Trois noms slaves de couleurs. *Lingua Posnaniensis* 3, 96–111.

Machek, Václav (1954). *Česká a slovenská jména rostlin*. Praha: Akademie věd.

Machek, Václav (1961). Zum Wortschatz des Litauischen: Teil 2. *Zeitschrift für Slavische Philologie* 29, 345–356.

Machek, Václav (1968). *Etymologický slovník jazyka českého*. Praha: Academia.

Mägiste, Julius (1923). Etümologiseerimiskatseid. *Eesti Keel* 2, 33–39.

Mägiste, Julius (1939). Läänemere-sm. ja volga-sm. keelte võrdluste alalt. *Eesti Keel* 18, 60–81.

Mägiste, Julius (1959). Gibt es im Tscheremissischen baltische Lehnwörter? *Ural-altaische Jahrbücher* 31, 169–176.

Mägiste, Julius (1962). Sm. *lunka*, *lunki* 'tuohi' ja mord. *l'eńǵe* 'niini'. *Virittäjä* 66, 244–247.

Mägiste, Julius (1970). Estn. (*h*)*õis* 'Blüte' und seine Sippe — ein balt. Lehnwort? In: Velta Rūķe-Draviņa (ed.), *Donum Balticum* [FS Stang]. Stockholm: Almquist & Wiksell, 352–360.

Mailhammer, Robert (2007). *The Germanic Strong Verbs*. Berlin–New York: Mouton de Gruyter.

Mallory, James P. (1982). Indo-European and Kurgan Fauna I. Wild Mammals. *Journal of Indo-European Studies* 10, 193–222.

Mallory, James P. (1983). Proto-Indo-European and Kurgan Fauna II. Fish. *Journal of Indo-European Studies* 11, 263–279.

Mallory, James P. (1989). *In Search of the Indo-Europeans*. London: Thames & Hudson.

Mallory, James P. (1991). Kurgan and Indo-European Fauna III. Birds. *Journal of Indo-European Studies* 19, 223–234.

Mallory, James P. & Adams, Douglas Q. (1997). *Encyclopedia of Indo-European Culture*. London–Chicago: Fitzroy Dearborn.

Mallory, James P. & Adams, Douglas Q. (2006). *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*. Oxford: University Press.

Marchand, James W. (1973). *The Sounds and Phonemes of Wulfila's Gothic*. Den Haag–Paris: Mouton.

Markey, Thomas L. (1989). The spread of agriculture in western Europe: Indo-European and (non-) pre-Indo-European linguistic evidence. In: David R. Harris & Gordon C. Hillman (eds.), *Foraging and Farming*. London–New York: Routledge, 585–619.

Marstrander, Carl J.S. (1915). *Bidrag til det norske sprogs historie i Irland*. Kristiania: Dybwad.

Martirosyan, Hrach K. (2008). *Etymological dictionary of the Armenian inherited lexicon*. Leiden–Boston: Brill.

Мартынов, Виктор В. (1985). Прусско-славянские эксклюзивные изолексы. *Этимология* 1982, 3–13.

Marzell = Marzell, Heinrich (1943–1979). *Wörterbuch der deutschen Pflanzennamen*. 4 vols. Leipzig: Hirzel/Stuttgart–Wiesbaden: Steiner.

Masing, Oskar (1926). *Niederdeutsche Elemente in der Umgangssprache der baltischen Deutschen*. Riga: Löffler.

Matasović, Maja (2011). *Analiza najstarijih latinskih posuđenica u hrvatskom*. Diss. Sveučilište u Zagrebu.

Matasović, Ranko (1997). Odrazi indoeuropskih laringala u slavenskim jezicima. *Croatica* 45/46, 129–146.

Matasović, Ranko (2004). The Proto-Indo-European syllabic resonants in Balto-Slavic. *Indogermanische Forschungen* 109, 337–354.

Matasović, Ranko (2005). The Centum Elements in Balto-Slavic. In: Gerhard Meiser & Olav Hackstein (eds.), *Sprachkontakt und Sprachwandel*. Wiesbaden: Reichert, 363–374.

Matasović, Ranko (2008). *Poredbeno-povijesna gramatika hrvatskoga jezika*. Zagreb: Matica Hrvatska.

Matasović, Ranko (2009). *Etymological Dictionary of Proto-Celtic*. Leiden–Boston: Brill.

Matasović, Ranko (2013). Substratum words in Balto-Slavic. *Filologija* 60, 75–102.

Matasović, Ranko (2020). Language of the bird names and the pre-Indo-European substratum. In: Romain Garnier (ed.), *Loanwords and Substrata*, 331–344.

Matasović, Ranko (forthc.). Proto-Slavic Forest Tree Names: Substratum or PIE origin? To appear in: Guus Kroonen (ed.), *Sub-Indo-European Europe*. De Gruyter Mouton.

Mathiassen, Terje (1995). Nochmals zum *ie*-Komplex im Ostbaltischen. *Linguistica Baltica* 4, 41–53.

Maticsák, Sándor (2014). The Erzya-Mordvin continuation of nominal derivational suffixes in the proto-language composed of a single sibilant or affricate. *Finnisch-Ugrische Mitteilungen* 38, 115–141.

Matzinger, Joachim (2006). *Der Altalbanische Text Mbsuame e Krështerë* (*Dottrina cristiana*) *des Lekë Matrënga von 1592*. Dettelbach: Röll.

Mayer, Harvey E. (1978). Die Divergenz des Baltischen und des Slavischen. *Zeitschrift für Slavische Philologie* 40/1, 52–62.

Mayer, Harvey E. (1990). Baltic via Glottalic Indo-European. *Baltistica 26*, 100–104.

Mayrhofer, Manfred (1986). *Indogermanische Grammatik* I. Heidelberg: Winter.

[Mayrhofer EWA, Mayrhofer KEWA — See EWA, KEWA under Abbreviated Titles.]

Mažiulis, Vytautas (1979). [Review of] В. Н. Топоров. Прусский язык. Словарь (A–D). *Baltistica* 15, 146–149.

[Mažiulis PKEŽ — See PKEŽ under Abbreviated Titles.]

McKenzie, Roderick (1918). Notes sur l'histoire des diphtongues *ie* et *uo* dans les langues baltiques. *Bulletin de la Société de linguistique de Paris* 21/2, 156–174.

Meillet, Antoine (1894). De quelques difficultés de la théorie des gutturales indo-européennes. *Mémoires de la Société de linguistique de Paris* 8, 277–304.

Meillet, Antoine (1905). *Études sur l'étymologie et le vocabulaire du vieux slave* II. Paris: Champion.

Meillet, Antoine (1907). Les alternances vocalique en vieux slave. *Mémoires de la Société de linguistique de Paris* 14, 332–390.

Meillet, Antoine (1908–1909). De quelques emprunts probables en grec et en latin. *Mémoires de la Société de linguistique de Paris* 15, 161–164.

Meillet, Antoine (1909). [Review of] Slavisches etymologisches Wörterbuch von Dr Erich Berneker. *Rocznik Slawistyczny* 2, 57–71.

Meillet, Antoine (1912). À propos de Avestique *zrazdā-*. *Mémoires de la Société de linguistique de Paris* 18, 60–64.

Meillet, Antoine (1913). La langue Lituanienne et la langue Lette. *Annales des nationalités* 2/5–6, 204–205.

Meillet, Antoine (1920). Les noms du "feu" et de l'"eau" et la question du genre. *Mémoires de la Société de linguistique de Paris* 21, 249–256.

Meillet, Antoine & Vaillant, André (1933). Le slave commun **jarębĭ* et le suffixe en *-b-*. *Revue des études slaves* 13, 101–102.

Meineke, Eckhard (1998). Glas. §1. Sprachliches. In: Rosemarie Müller et al. (ed.), *Reallexikon der germanischen Altertumskunde* [2nd edition] XII. *Getränke–Greiftierstil*. Berlin–New York: De Gruyter, 139–142.

Meiser, Gerhard (1998). *Historische Laut- und Formenlehre der lateinischen Sprache* [2nd edition]. Darmstadt: Wissenschaftliche Buchgesellschaft.

Meissner, Torsten (2013). Griechische Etymologie [Review of Beekes 2010]. *Kratylos* 58, 1–31.

Melchert, H. Craig (2010). The Word for 'mouth' in Hittite and Proto-Indo-European. *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 7, 55–63.

Meringer, Rudolf (1909). Sprachlich-sachliche Probleme. *Wörter und Sachen* I, 164–210.

Merkevičius, Algimantas & Muradian, Lijana (2016). Ankstyviausi žirgų palaikai laidojimo objektuose Lietuvoje. *Archaeologia Lituana* 16, 28–39.

Metsäranta, Niklas (2020). *Periytyminen ja lainautuminen: Marin ja permiläisten kielten sanastontutkimusta*. Diss. Helsingin yliopisto.

Meyer, Gustav (1891). *Etymologisches Wörterbuch der albanesischen Sprache*. Strassburg: Trübner.

Meyer, Gustav (1892). Etymologisches. *Indogermanische Forschungen* 1, 319–329.

Mielcke = Mielcke, Christian (1800). *Littauisch-Deutsches und Deutsch-Littauisches Wörter-Buch*. 2 vols. Königsberg.

Miežinis, Mykolas (1894). *Lietuviszkai-latviszkai-lenkiszkai-rusiszkas žodynas*. Tilžė: Noveskis.

Mikkola, Jooseppi J. (1894). *Berührungen zwischen den westfinnischen und slavischen Sprachen*. Helsinki: Suomalais-Ugrilainen Seura.

Mikkola, Jooseppi J. (1896a). Litauische lehnwörter im Slavischen. *Beiträge zur Kunde der indogermanischen Sprachen* 21, 118–121.

Mikkola, Jooseppi J. (1896b). Etymologische beiträge. *Beiträge zur Kunde der indogermanischen Sprachen* 21, 218–225.

Mikkola, Jooseppi J. (1897). Baltische etymologien. *Beiträge zur Kunde der indogermanischen Sprachen* 22, 239–255.

Mikkola, Jooseppi J. (1898). Slavica II. *Indogermanische Forschungen* 8, 302–304.

Mikkola, Jooseppi J. (1899). Baltische etymologien II. *Beiträge zur Kunde der indogermanischen Sprachen* 25, 73–76.

Mikkola, Jooseppi J. (1903). Baltisches und Slavisches. *Öfversigt af Finska Vetenskaps-Societens Förhandlingar* 45/4, 2–47.

Mikkola, Jooseppi J. (1906). Eräs baltilainen lainasana. *Virittäjä* 10, 78.

Mikkola, Jooseppi J. (1908). [Review of] Vergleichende slavische Grammatik von Dr Wencel Vondrák. *Rocznik Slawistyczny* 1, 3–19.

Mikkola, Jooseppi J. (1930). Vecākie sakari somu un baltu valodu starpā. *Izglītības Ministrijas Mēnesraksts* 1930, 436–446.

Mikkola, Jooseppi J. (1938). *Die älteren Berührungen zwischen Ostseefinnisch und Russisch*. Helsinki: Suomalais-Ugrilainen Seura.

Miklosich, Franz von (1865). *Lexicon palaeoslovenico–graeco–latinum*. Wien: Braumüller.

Miklosich, Franz von (1878). *Altslovenische Lautlehre* [3rd edition]. Wien: Braumüller.

Miklosich, Franz von (1886). *Etymologisches Wörterbuch der slavischen Sprachen*. Wien: Braumüller.

Militarev, Alexander & Kogan, Leonid (2005). *Semitic Etymological Dictionary* II. *Animal Names*. Münster: Ugarit-Verlag.

Militarev, Alexander & Nikolaev, Sergei (2020). Proto-Afrasian names of ungulates in light of the Proto-Afrasian homeland issue. *Journal of Language Relationship* 18/3, 199–226.

Minkevičius, Karolis et al. (2020). New evidence on the southeast Baltic Late Bronze Age agrarian intensification and the earliest AMS dates of *Lens culinaris* and *Vicia faba*. *Vegetation History and Archaeobotany* 29, 327–338.

Mittnik, Alissa et al. (2018). The genetic prehistory of the Baltic Sea region. *Nature Communications* 9/442.

Mittnik, Alissa et al. (2019). Kinship-based social inequality in Bronze Age Europe. *Science* 366, 731–734.

Младеновъ, Стефанъ (1909). Старитѣ германски елементи въ славянскитѣ езици. *Сборникъ за народни умотворения, наука и книжнина* 25.

Motuzaitė Matuzevičiūtė, Giedrė (2018). The possible geographic margin effect on the delay of agriculture introduction into the East Baltic. *Estonian Journal of Archaeology* 22, 149–162.

Motuzaitė Matuzevičiūtė, Giedrė et al. (2022). The integration of millet into the diet of Central Asian populations in the third millennium BC. *Antiquity* 96 (387), 560–574.

Mous, Maarten (1996). Was There Ever A Southern Cushitic Language (Pre-) Ma'a? In: Catherine Griefenow-Mewis & Rainer Voigt (eds.), *Cushitic and Omotic Languages*. Köln: Köppe, 201–211.

Мудрак, Олег А. (1993). *Исторические соответствия чувашских и тюркских гласных: опыт реконструкции и интерпретации*. Москва: Институт востоковедения.

Мудрак, Олег А. (2007). Аварская надпись на сосуде из клада Надь-Сент-Миклош. In: Михаил Е. Алексеев (ed.), *Аспекты алтайского языкознания: материалы Тенишевских чтений*. Москва: Советский писатель, 81–103.

Mueller-Bieniek, Aldona (2010). Carrot (*Daucus carota* L.) in Medieval Kraków (S. Poland): a cultivated form? *Journal of Archaeological Science* 37, 1725–1730.

Munkácsi, Bernat & Kálman, Béla (1986). *Wogulisches Wörterbuch*. Budapest: Akadémiai kiadó.

Murdock, George P. & Provost, Caterina (1973). Factors in the Division of Labor by Sex: A Cross-Cultural Analysis. *Ethnology* 12/2, 203–225.

Мызников, Сергей А. (2004). *Лексика финно-угорского происхождения в русских говорах Северо-Запада*. Санкт-Петербург: Наука.

Мызников, Сергей А. (2019). *Русский диалектный этимологический словарь*. Москва–Санкт-Петербург: Нестор-История.

Надькин, Дмитрий Т. (1988). Проблема редуцированного гласного и особенности развития ударения в мордовских языках. *Актуальные вопросы мордовского языкознания* 94, 3–11.

Naktinienė, Gertrūda, Aldona Paulauskienė & Vytautas Vitkauskas (1988). *Druskininkų tarmės žodynas*. Vilnius: Mokslas.

Напольских, Владимир В. (1990). Палеоевропейский субстрат в составе западных финно-угров. *Uralo-Indogermanica: Балто-славянские языки и проблема урало-индоевропейских связей* II. Москва: Академия наук СССР, 128–134.

Напольских, Владимир В. (1997). Происхождение субстратных палеоевропейских компонентов в составе западных финно-угров. *Балто-славянские исследования* 1988–1996, 198–208.

Напольских, Владимир В. [= Napoľskich, Wladimir] (2002). Zu den ältesten Beziehungen zwischen Finno-Ugriern und zentraleuropäischen Indogermanen. In: Rogier Blokland & Cornelius Hasselblatt (eds.), *Finno-Ugrians and Indo-Europeans: linguistic and literary contacts*. Maastricht, 265–271.

Напольских, Владимир В. (2006). Балто-славянский языковой компонент в Нижнем Прикамье в середине I тыс. н. э. *Славоведение* 2006/2, 3–19.

Напольских, Владимир В. (2010). Названия спорыни в удмуртском языке. *Acta Linguistica Petropicani* 6/1, 51–60.

Напольских, Владимир В. (2015). К происхождению названий соли в финно-пермских языках. *Linguistica Uralica* 51, 161–176.

Напольских, Владимир В. & Энговатова, Ася В. (2000). [Review of] Симпозиум "Контакты между носителями нидоевропейских и уральских языков в неолите, энеолите и бронзовом веке в свете лингвистических и археологических данных". *Российская археология* 2000/4, 224–232.

Nehring, Alfons (1959). Idg. **mari, *mori*. In: Wolfdietrich Rasch (ed.), *Festschrift für Franz Rolf Schröder*. Heidelberg: Winter.

Nesselmann, Georg H.F. (1851). *Wörterbuch der Littauischen Sprache*. Königsberg: Gebrüder Borntraeger.

Nichols, Johanna (1998). The Eurasian spread zone and the Indo-European dispersal. In: Roger Blench & Matthew Spriggs (eds.), *Archaeology and Language* II. *Archaeological Data and Linguistic Hypotheses*. London: Routledge, 220–266.

Niedermann, Max (1918). *Essais d'étymologie et de critique verbale latines*. Neuchâtel: Attinger Frères.

Nieminen, Eino (1934). Der stammauslaut der ins urfinnische entlehnten baltischen *ā*-feminina und die herkunftsfrage. *Finnisch-Ugrische Forschungen* 22, 5–66.

Nieminen, Eino (1940). Itämerensuomalaisten ja balttilaisten kielten välisten kosketusten tutkimuksesta [Review of Kalima 1936]. *Virittäjä* 44, 376–384.

Nieminen, Eino (1944). Eräistä balttilaisista lainasanoistamme ja niiden vastineista. *Virittäjä* 48, 238–250.

Nieminen, Eino (1945). *Reuna* ja *leuka*. *Virittäjä* 49, 40–55.

Nieminen, Eino (1953). Ein Beitrag zu der ostslavischen und ostseefinnischen Badeterminologie. *Lingua Posnaniensis* 4, 211–228.

Nieminen, Eino (1955). Die baltischen und ostseefinnischen Ausdrücke für Segel. *Zeitschrift für vergleichende Sprachforschung* 72, 129–160.

Nieminen, Eino (1956). Die Benennungen der Hirse in den baltischen Sprachen. *Zeitschrift für vergleichende Sprachforschung* 74, 162–176.

Nieminen, Eino (1957). Über einige Eigenschaften der baltischen Sprache, die sich in den ältesten baltischen Lehnwörtern der ostsefinnischen Sprachen abspiegelt. *Sitzungsberichte der Finnischen Akademie der Wissenschaften* 1956, 185–206.

Nieminen, Eino (1963). Ausdrücke für Bau-, Arbeits- und Nutzhölzer fremden Ursprungs im Finnischen. *Zeitschrift für Slavische Philologie* 31, 233–241.

Nikkilä, Osmo (1982). Wörter für 'Körper' germanischen Ursprungs im Ostseefinnischen. *Советское финно-угроведение* 18, 251–260.

Nikkilä, Osmo (1999). Sporadischer Konsonantenwechsel im Ostseefinnischen. *Finnisch-Ugrische Mitteilungen* 21/22, 129–160.

Nikkilä, Osmo (2002). Suomen sanojen alkuperät päätökseen [Review of SSA III]. *Virittäjä* 106, 124–136.

Nikolaev, Alexander (2018). Greek Etymology in the 21st century [Conference handout]. *Society for Classical Studies 149th Annual Meeting*, Boston, Massachusetts.

Николаев, Сергей Л. (1988). Следы особеностей восточнославянских племенных диалектов в современных великорусских говорах I. Кривичи. *Балто-славянские исследования* 1986, 115–154.

Николаев, Сергей Л. (1989). Следы особеностей восточнославянских племенных диалектов в современных великорусских говорах I. Кривичи (окончание). *Балто-славянские исследования* 1987, 187–225.

Николаев, Сергей Л. (1990). К истории племенного диалекта кривичей. *Советское славяноведение* 4, 54–63.

Николаев, Сергей Л. (1995). Вокализм карпато-украинских говоров I. Покутско-буковинско-гуцульский ареал. *Славяноведение* 1995/3, 105–122.

Николаев, Сергей Л. (2001). Из исторической фонетики и просодии северо-западных говоров. *Вопросы русского языкознания* 9, 86–121.

Николаев, Сергей Л. (2017). К этимологии и сравнительно-исторической фонетике имен северогерманского (скандинавского) происхождения в «Повести временных лет». *Вопросы ономастики* 14/2, 7–54.

Николаев, Сергей Л. (2020). «*Слово о полку Игореве*»: *Реконструкция стихотворного текста*. Москва–Санкт-Петербург: Нестор-История.

Николаев, Сергей Л. & Хелимский, Евгений А. (1990). Славянские (новгородо-псковские) заимствования в прибалтийско-финских языках: *-a* и *-i* в рефлексах имен мужского рода. *Uralo-Indogermanica: Балто-славянские языки и проблема урало-индоевропейских связей* II. Москва: Академия наук СССР, 41–43.

Nikulin, Andrey (2016). *Six Uralic Etymologies* [Unpublished article]. Accessed online at academia.edu/28497819.

Nilsson, Torbjörn K. (1996). Is Balto-Finnic *kypärä* 'helm, casque' really a loanword from Baltic, or is it indigenous? *Word* 47, 185–191.

Nilsson, Torbjörn K. (2001). A new treatise on Baltic loanwords in Finnish [Review of Liukkonen 1999]. *Linguistica Baltica* 9, 177–194.

Nirvi, Ruben E. (1964). *Sanoja ja käyttäytymistä*: *Sanahistoriallisia tutkimuksia* II. Helsinki: Suomalaisen Kirjallisuuden Seura.

Noreen, Adolf (1894). Etymologisches. *Indogermanische Forschungen* 4, 320–326.

Noreen, Adolf (1923). *Altnordische Grammatik* I. *Altisländische und altnorwegische Grammatik* [4th edition]. Halle: Niemeyer.

Normier, Rudolf (1981). Zu Esche und Espe. *Die Sprache* 27, 22–29.

Nowak, Ronald M. & Paradiso, John L. (1983). *Walker's mammals of the world* [4th edition]. Baltimore–London: Johns Hopkins University Press.

Nussbaum, Alan (1986). *Head and Horn in Indo-European*. Berlin: De Gruyter.

Nuutinen, Olli (1987a). *Malo* — balttilainen lainasanarelikti. *Virittäjä* 91, 514–530.

Nuutinen, Olli (1987b). Kalannimi seipi. In: Mauno Koski, Eeva Lähdemäki & Kaisa Häkkinen (eds.), *Fennistica festiva in honorem Göran Karlsson septuagernarii*. Åbo Akademis förlag, 107–115.

Nuutinen, Olli (1987c). Balttilaisten lainojen substituutiotapauksia. *Virittäjä* 91, 52–63.

Nuutinen, Olli (1989). *Järvi* — balttilainen laina. *Virittäjä* 93, 497–503.

Nuutinen, Olli (1992). *Pälvi* — balttilainen laina. *Virittäjä* 96, 403–406.

Oettinger, Norbert (2003). Neuerungen in Lexikon und Wortbildung des Nordwest-Indogermanischen. In: Alfred Bammesberger & Theo Vennemann (eds.), *Languages in Prehistoric Europe*. Heidelberg: Winter, 183–193.

Ohala, John J. (1978). Southern Bantu vs. the World: The Case of Palatalization of Labials. *Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society*, 370–386.

Ojansuu, Heikki (1908). Hehko, hehvo, hieho. *Virittäjä* 12, 33–35.
Ojansuu, Heikki (1916). *Suomen kielen tutkimuksen työmaalta. Sarja esitelmiä* I. Jyväskylä: Gummerus.
Ojansuu, Heikki (1921). *Lisiä suomalais-balttilaisiin kosketuksiin*. Helsinki: Suomalaisen Kirjallisuuden Seura.
Ojansuu, Heikki (1922). Eesti Etümologiad. *Eesti Keel* 1, 137–139.
Olander, Thomas (2015). *Proto-Slavic Inflectional Morphology*. Leiden–Boston: Brill.
Olander, Thomas (2020). To **b* or not to **b*: Proto-Indo-European **b* in a phylogenetic perspective. *Historische Sprachforschung* 133, 182–208.
Olsen, Birgit (1999). *The Noun in Biblical Armenian*. Berlin–New York: De Gruyter.
Olsen, Birgit (2010). *Derivation and Composition*. Innsbruck: Institut für Sprachen und Literaturen.
Olsen, Birgit & Thorsø, Rasmus (2022). Armenian. In: Thomas Olander (ed.), *The Indo-European Language Family: A Phylogenetic Perspective*. Cambridge: University Press.
Orel, Vladimir (2000). *A Concise Historical Grammar of the Albanian Language*. Leiden–Boston–Köln: Brill.
Orel, Vladimir (2003). *A Handbook of Germanic Etymology*. Leiden–Boston: Brill.
Osthoff, Hermann (1898). Griechische und lateinische Wortdeutungen III. *Indogermanische Forschungen* 8, 1–68.
Osthoff, Hermann (1901). *Etymologische Parerga* I. Leipzig: Hirzel.
Oštir, Karel (1921). *Beiträge zur alarodischen Sprachwissenschaft* I. Wien–Leipzig: Beyers.
Oštir, Karel (1929). Japodi. *Etnolog* 3, 87–113.
Oštir, Karel (1930). *Drei vorslavisch-etruskische Vogelnamen*. Ljubljana: Znanstveno Drustvo.
Otrębski, Jan (1939). *Indogermanische Forschungen*. Vilnius: Towarzystwo przyjaciół nauk.
Otrębski, Jan (1955). Aus der Geschichte der litauischen Sprache. *Lingua Posnaniensis* 5, 23–40.
Otrębski, Jan (1966). Die ältesten germanischen Lehnwörter im Baltischen und Slavischen. *Die Sprache* 12, 50–64.
Ozoliņš, Kaspars (2015). *Revisiting Proto-Indo-European Schwebeablaut*. Diss. University of California.
Paasonen, Heikki (1896). Kielellisiä lisiä suomalaisten sivistyshistoriaan. *Suomi* 3/13.
Paasonen, Heikki (1897). *Die türkischen Lehnwörter im Mordwinischen*. Helsingfors: Finnische Litteratur-Gesellschaft.
Paasonen, Heikki (1903). *Mordvinische Lautlehre*. Helsinki: Société Finno-Ugrienne.
Paasonen, Heikki (1906). Über die benennung des roggens im syrjänisch-wotjakischen und im mordwinischen. *Journal de la Société Finno-Ougrienne* 23, 1–7.
Paasonen, Heikki (1909a). *Mordwinische Chrestomathie* [2nd edition]. Helsinki: Suomalais-Ugrilainen Seura.

Paasonen, Heikki (1909b). Eräs liettualaisperäinen sivistyssana. *Suomalainen Tiedeakatemia. Esitelmät ja pöytäkirjat* 1, 16–17.

[Paasonen MdWb — See MdWb under Abbreviated Titles.]

Pabrėža, Jurgis A. (1834). *Waardaa tayslyynee Augimiu atsyrądąntiu Augmyniićźioo źemaytyyniee* [Manuscript]. Accessed online at https://www.epaveldas.lt/preview?id=C10000689960.

Pajusalu, Karl, Arvo Krikmann & Eberhard Winkler (2009). Lexical relations between Salaca Livonian and Estonian dialects. *Linguistica Uralica* 45, 283–298.

Pakerys, Antanas (2003). *Lietuvių bendrinės kalbos fonetika*. Vilnius: Enciklopedija.

Palionis, Jonas (1995). *Lietuvių literatūrinės kalbos istorija* [2nd edition]. Vilnius: Mokslas.

Palmér, Axel I., Anthony Jakob et al. (2021). Proto-Indo-European 'fox' and the reconstruction of an athematic *ḱ*-stem. *Indo-European Linguistics* 9, 234–263.

van Pareren, Remco (2006). *Baltische leenwoorden in het Mordwiens*. MA Thesis Universiteit Groningen.

van Pareren, Remco (2008). Die direkten baltischen Lehnwörter im Mordwinischen. *Finnisch-Ugrische Mitteilungen* 30/31, 69–147.

Parpola, Asko (2012). Formation of the Indo-European and Uralic language families in the light of archaeology: Revised and integrated 'total' correlations. In: Riho Grünthal & Petri Kallio (eds.), *A Linguistic Map of Prehistoric Northern Europe*. Helsinki: Suomalais-Ugrilainen Seura, 119–184.

Parry-Williams, Thomas H. (1923). *The English element in Welsh*. London: Cymmrodorion.

Павловъ, Алексѣй С. (1890). *Неизданный памятникъ русскаго церковнаго права XII вѣка*. С.-Петербургъ.

Pedersen, Holger (1895). Das indogermanische *s* im Slavischen. *Indogermanische Forschungen* 5, 33–87.

Pedersen, Holger (1905). Die nasalpräsentia und der Slavische akzent. *Zeitschrift für vergleichende Sprachforschung* 38, 297–421.

Pedersen, Holger (1906). Armenisch und die nachbarsprachen. *Zeitschrift für vergleichende Sprachforschung* 39, 334–485.

Pedersen, Holger (1909). *Vergleichende Grammatik der keltischen Sprachen* I. *Einleitung und Lautlehre*. Göttingen: Vandenhoeck & Ruprecht.

Pedersen, Holger (1913). *Vergleichende Grammatik der keltischen Sprachen* II. *Bedeutungslehre* (*Wortlehre*). Göttingen: Vandenhoeck & Ruprecht.

Pedersen, Holger (1935). Lit. *iau*. *Studi Baltici* 4, 150–154.

Pedersen, Holger (1951). *Die gemeinindoeuropäischen und die vorindoeuropäischen Verschlusslaute*. København: Munksgaard.

Пекарский = Пекарский, Эдуард К. (1907–1930). *Словарь якутского языка*. 13 issues. Санкт-Петербург–Ленинград: Академия наук.

Pelisiak, Andrzej (2016). The beginnings of mobile husbandry in the mountain periphery of Southeastern Poland. In: Martin Furholt, Ralf Grossmann & Marzena Szmyt (eds.), *Transitional landscapes? The 3rd millennium BC in Europe*. Bonn: Habelt, 210–227.

Перевощиков, Петр Н. (1962, ed.). *Грамматика современного удмуртского языка*. Ижевск: Удмуртское книжное издательство.

Persson, Per (1912). *Beiträge zur indogermanischen Wortforschung*. Uppsala: Akademiska bokhandeln.

Petersson, Herbert (1909). Etymologien. *Indogermanische Forschungen* 23, 384–404.

Petersson, Herbert (1921). *Studien über die indogermanische Heteroklisie*. Lund: Gleerup.

Petit, Daniel (2000). Lituanien *taũsti, čiùtnas*, et le nom du «peuple» en indo-européen. *Bulletin de la Société de linguistique de Paris* 45, 119–146.

Petit, Daniel (2003). Note sur le lituanien *kraĩtis*. In: Norbert Ostrowski & Ona Vaičiulytė-Romančuk (eds.), *Prace bałtystyczne: Język, literatura, kultura*. Warszawa: Uniwersytet Warszawski, 92–100.

Petit, Daniel (2004). *Apophonie et catégories grammaticales dans les langues baltiques*. Leuven–Paris: Peeters.

Petit, Daniel (2005). Das litauische Adverb *ankstì* 'früh'. *Studia Etymologica Cracoviensia* 10, 141–156.

Petit, Daniel (2010). *Untersuchungen zu den baltschen Sprachen*. Leiden–Boston: Brill.

Peust, Carsten (1999). *Egyptian Phonology: An Introduction to the Phonology of a Dead Language*. Göttingen: Peust & Gutschmidt.

Peyrot, Michaël (2018). Tocharian Agricultural Terminology: Between Inheritance and Language Contact. In: Guus Kroonen, James P. Mallory & Bernard Comrie (eds.), *Talking Neolithic*. Washington, DC: Institute for the Study of Man, 242–277.

Philippa et al. = Marlies Philippa, Frans Debrabandere & Arend Quak (2003–2009, eds.). *Etymologisch woordenboek van het Nederlands*. Amsterdam: University Press.

Пичхадзе, Анна А. (2002). О происхождении славянского перевода Хроники Георгия Амартола. *Лингвистическое источниковедение и история русского языка* 2001, 232–249.

Piggot, Stuart (1969). The Earliest Wheeled Vehicles and the Caucasian Evidence. *Proceedings of the Prehistoric Society* 34, 266–318.

Pijnenburg, Wilhelmus J.J. (1989). Eine germanisch-baltoslawische Isoglosse. *Historische Sprachforschung* 102, 99–106.

Piličiauskas, Gytis (2018). *Virvelinės keramikos kultūra Lietuvoje 2800–2400 cal BC*. Vilnius: Lietuvos istorijos institutas.

Piličiauskas, Gytis, Dalia Kisielienė & Giedrė Piličiauskienė (2016). Deconstructing the concept of Subneolithic farming in the southeastern Baltic. *Vegetation History and Archaeobotany* 26, 183–193.

Piličiauskas, Gytis et al. (2017). The transition from foraging to farming (7000–500 cal BC) in the SE Baltic. *Journal of Archaeological Science: Reports* 14, 530–542.

Pinault, Georges-Jean (2008). *Chrestomathie tokharienne*. Leuven–Paris: Peeters.

Pinault, Georges-Jean (2009). On the formation of the Tocharian demonstratives. In: Elisabeth Rieken & Paul Widmer (eds.), *Pragmatische Kategorien*: *Form*, *Funktion und Diachronie* [Proceedings]. Wiesbaden: Reichert, 221–245.

Пирейко, Лия А. (1976). *Талышско-русский словарь*. Москва: Русский язык.

Pisani, Vittore (1968). Rom und die Balten. *Baltistica* 4, 7–21.

Pleteršnik = Pleteršnik, Maks (1894–1895). *Slovensko-nemški slovar*. 2 vols. Ljubljana: Knezoškofijstvo.

Plöger, Angela (1973). *Die russischen Lehnwörter der finnischen Schriftsprache*. Wiesbaden: Harrassowitz.

Plöger, Angela (1982). Über die Entstehung des finnischen Stammtyps *CV̄C*(*C*)*a*/*ä*. *Finnisch-Ugrische Forschungen* 44, 66–98.

Подвысоцкій, Александръ (1885). *Словарь областного архангельскаго нарѣчія*. С.-Петербургъ: Академія наукъ.

Pogatscher, Alois (1898). Altenglisch *ƀr* aus *mr*. In: Jakob Schipper (ed.), *Festschrift zum VIII. allgemeinen deutschen Neuphilologentage in Wien Pfingsten 1898*. Wien–Leipzig: Braumüller, 97–107.

Погодинъ, Александръ Л. (1903). *Слѣды корней-основъ въ славянскихъ языкахъ*. Варшава: Варшавскій учебный округъ.

Polák, Václav (1955). Trois notules étymologiques concernant les mots de substrat. *Časopis pro moderní filologii* 37, 54–56.

Pollmann, Britta (2014). Environment and agriculture of the transitional period from the Late Bronze to early Iron Age in the eastern Baltic: an archaeobotanical case study of the lakeshore settlement Luokesa 1, Lithuania. *Vegetation History and Archaeobotany* 23, 403–418.

Polomé, Edgar C. (1986). The Non-Indo-European Component of the Germanic Lexicon. In: Annemarie Etter (ed.), *o-o-pe-ro-si* [FS Ernst Risch]. Berlin–New York: De Gruyter, 661–672.

Polomé, Edgar C. (1989). Substrate Lexicon in Germanic. *North-Western European Language Evolution* 14, 53–73.

Polomé, Edgar C. (1990). The Indo-Europeanization of Northern Europe: the Linguistic Evidence. In: Karlene Jones-Bley & Martin E. Huld (eds.), *The Indo-Europeanization of Northern Europe*. Washington, DC: Institute for the Study of Man, 331–338.

Polomé, Edgar C. (1992). Indo-European and Substrate Languages in the West. *Archivio Glottologico Italiano* 67, 66–85.

Popowska-Taborska, Hanna & Boryś, Wiesław (1996). *Leksyka kaszubska na tle słowiańskim*. Warszawa: Akademia Nauk.

Porzig, Walter (1954). *Die Gliederung des indogermanischen Sprachgebiets*. Heidelberg: Winter.

Posti, Lauri (1946). Etymologisia huomioita. *Virittäjä* 50, 385–387.

Posti, Lauri (1953). From Pre-Finnic to Late Proto-Finnic. *Finnisch-Ugrische Forschungen* 31/1, 1–91.

Posti, Lauri (1962). *Silakka* sanan alkuperästä. *Virittäjä* 66, 282–287.

Posti, Lauri (1977). Some new contributions to the stock of Baltic loanwords in Finnic languages. *Baltistica* 13, 263–270.

Потебня, Александръ А. (1881). *К исторіи звуков русскаго языка* III. *Этимологическія и другія замѣтки*. Варшава: Земкевичъ и Нолковский.

Prellwitz, Walther (1891). *Die deutschen Lehnwörter im Preussischen und Lautlehre der deutschen Lehnwörter im Litauischen*. Göttingen: Vandenhoeck & Ruprecht.

Prellwitz, Walther (1897). Studien zur indogermanischen etymologie und wortbildung. *Beiträge zur Kunde der indogermanischen Sprachen* 22, 76–114.

Prellwitz, Walther (1909). Etymologische Miszellen. *Zeitschrift für vergleichende Sprachforschung* 42, 385–387.

Преображенский = Преображенскій, Александръ Г. (1910–1916). *Этимологическій словарь русскаго языка*. 2 vols. Москва: Лисснеръ & Совко.

Press, J. Ian (1986). *Aspects of the Phonology of the Slavonic Languages*: *The Vowel* y *and the Consonantal Correlation of Palatalization*. Amsterdam: Rodopi.

Priestly, Tom M.S. (1997). Vandals, Veneti, Windischer: The Pitfalls of Amateur Historical Linguistics. *Slovene Studies* 19, 3–41.

Pronk, Tijmen (2012). Proto-Indo-European long vowels and Balto-Slavic accentuation. *Baltistica* 47, 205–247.

Pronk, Tijmen (2013). On the Development of **in*, **im*, **un* and **um* in Slavic. In: Marko Jesenšek (ed.), *Miklošičeva monografija. Ob dvestoletnici rojstva Franca Miklošiča*. Ljutomer: Gimnazija Franca Miklošiča, 117–140.

Pronk, Tijmen (2015). Singulative *n*-stems in Indo-European. *Transactions of the Philological Society* 113/3, 327–348.

Pronk, Tijmen (2018). Old Church Slavonic (*j*)*utro*, Vedic *uṣár*- 'daybreak, morning'. In: Lucien van Beek et al. (eds.), *Farnah* [FS Lubotsky]. Ann Arbor: Beech Stave Press, 298–306.

Pronk, Tijmen (2019a). Proto-Indo-European **a*. *Indo-European Linguistics* 7, 122–163.

Pronk, Tijmen (2019b). Eichner's law: a critical survey of the evidence. *Münchener Studien zur Sprachwissenschaft* 73, 121–155.

Pronk, Tijmen (2022). Balto-Slavic. In: Thomas Olander (ed.), *The Indo-European Language Family: A Phylogenetic Perspective*. Cambridge: University Press.

Pronk, Tijmen & Pronk-Tiethoff, Saskia (2018). Balto-Slavic Agricultural Terminology. In: Guus Kroonen, James P. Mallory & Bernard Comrie (eds.), *Talking Neolithic*. Washington, DC: Institute for the Study of Man, 278–314.

Pronk-Tiethoff, Saskia (2012). *The Germanic loanwords in Proto-Slavic*. Amsterdam–New York: Rodopi.

Puhvel, Jaan (1990). On Terms for 'nit' in Baltic and Balto-Finnic. In: Baiba Metuzāle-

Kangere & Helge Rinholm (eds.), *Symposium Balticum* [FS Rūķe-Draviņa]. Hamburg: Buske.

Puhvel, Jaan (1997). *Hittite Etymological Dictionary* IV. *Words beginning with K*. Berlin–New York: Mouton De Gruyter.

Pystynen, Juho (2014a). Palatal Unpacking in Finnic [Conference talk]. *19th Conference of the Finno-Ugric Studies Association of Canada*.

Pystynen, Juho (2014b). Interplay of minor soundlaws: Samic glide clusters. [Blog post, dated 19 July 2014, hosted at protouralic.wordpress.com].

Pystynen, Juho (2016). Another Phonological Relict in South Estonian. [Blog post, dated 23 July 2016, hosted at protouralic.wordpress.com].

Pystynen, Juho (2017). **ü* > **i*, **ü* in Samoyedic. [Blog post, dated 7 October 2017, hosted at protouralic.wordpress.com].

Pystynen, Juho (2018). *Itämerensuomen pitkien vokaalien alkuperä*. Diss. Helsingin yliopisto.

Pystynen, Juho (2019). *Seistä*. In: Santeri Junttila & Juha Kuokkala (eds.), *Petri Kallio Rocks*: *Liber semisaecularis*. Helsinki, 153–160.

Pystynen, Juho (2020a). On the development of **i* in Permic. *Finnisch-ugrische Forschungen* 65, 62–97.

Pystynen, Juho (2020b). Secondary apocope in Mordvin. [Blog post, dated 14 February 2020, hosted at protouralic.wordpress.com].

Pystynen, Juho (2020c). Notes on Proto-Mansi word-final vocalism. In: Sampsa Holopainen et al. (eds.), *Ёмас сымыӈ нӕкве вӧртур ӕтпост самын патум* [FS Ulla-Maija Forsberg]. Helsinki: Suomalais-Ugrilaisen Seuran Toimituksia, 249–261.

Pystynen, Juho (2023). *Sorsa* ja *sotka*. Äännesuhteita sukelluksissa. In: Sampsa Holopainen, Jeongdo Kim & Niklas Metsäranta (eds.), *Elämä ja etymologia* [FS Saarikivi]. Helsinki, 336–364.

Qvigstad, Just (1893). *Nordische Lehnwörter im Lappischen*. Christiania: Videnskabs-Selskabs forhandlinger.

Rannamäe, Eve (2016). *Development of sheep populations in Estonia as indicated by archaeofaunal evidence and ancient mitochondrial DNA lineages from the Bronze Age to the Modern Period on sheep*. Diss. University of Tartu.

Räsänen, Martti (1920). *Die tschuwassischen Lehnwörter im Tscheremissischen*. Helsinki: Suomalais-Ugrilainen Seura.

Räsänen, Martti (1969). *Versuch eines etymologischen Wörterbuchs der Türksprachen*. Helsinki: Suomalais-Ugrilainen Seura.

Rasmussen, Jens E. (1999). *Selected Papers on Indo-European Linguistics* I. København: Museum Tusculanum.

Ravila, Paavo (1935). Die stellung des lappischen innerhalb der finnisch-ugrischen sprachfamilie. *Finnisch-Ugrische Forschungen* 23, 20–65.

Rédei, Károly (1968). Szófejtések. *Nyelvtudományi közlemények* 70, 155–161.

Rédei, Károly (1986). *Zu den indogermanisch-uralische Sprachkontakten*. Wien: Akademie der Wissenschaften.

Rédei, Károly (2000). [Review of] Kari Liukkonen, Baltisches im Finnischen. *Linguistica Uralica 36*, 226–229.

[Rédei UEW — see UEW under Abbreviated Titles.]

Reutercrona, Hans G. (1920). *Svarabhakti und erleichterungsvokal im altdeutschen bis ca. 1250*. Heidelberg: Winter.

Rhea. Samuel A. (1872). Brief Grammar and Vocabulary of the Kurdish Language of the Hakari District. *Journal of the American Oriental Society* 10, 118–155.

Rießler, Michael (2022). Kildin Saami. In: Marianne Bakró-Nagy, Johanna Laakso & Elena Skribnik (eds.), *The Oxford Guide to the Uralic Languages*. Oxford: University Press, 219–239.

Rietz = Rietz, Johann E. (1862–1867). *Svenskt dialektlexikon: ordbok öfver svenska allmogespråket* [1962 reprint]. Lund: Gleerup.

Rimantienė, Rimutė (1992). The Neolithic of the Eastern Baltic. *Journal of World Prehistory* 6, 97–143.

Ritter, Ralph-Peter (1977). Zur Frage der finnischen Evidenz für die Sievers'sche Regel im Germanischen. *Die Sprache* 23, 171–179.

Ritter, Ralph-Peter (1993). *Studien zu den ältesten germanischen Entlehnungen im Ostseefinnischen*. Frankfurt am Main: Lang.

Ritter, Ralph-Peter (1998). Zu den Reflexen der baltischen Liquiddiphthonge im Ostseefinnischen. *Res Balticae* 1998, 113–116.

Rix, Helmut (1976). *Historische Grammatik des Griechischen*: *Laut- und Formenlehre*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Rix, Helmut (1999). Schwach charakterisierte lateinische Präsensstämme zu Seṭ-Wurzeln mit Vollstufe I. In: Heiner Eichner & Hans Christian Luschützky (eds.), *Compositiones indo-germanicae: in memoriam Jochem Schindler*. Praha–Wien: Enigma, 515–535.

Robbeets, Martine (2004). Does Doerfer's *Zufall* mean 'cognate'? The case of the initial velar correspondence in Altaic. *Turkic Languages* 8, 147–179.

Roffet-Salque, Mélanie et al. (2015). Widespread exploitation of the honeybee by early Neolithic farmers. *Nature* 527, 226–231.

Rosenschon, Ursula (1993). Sechs Seiten medizinischer Rezepte im glagolitischen Psalter 3/N des Sinaiklosters. *Sudhoff's Archiv* 77, 129–159.

Rozwadowski, Jan M. (1915). Przyczynki do historycznej fonetyki języków słowiańskich. *Rocznik slawistyczny* 7, 9–21.

Ruhig = Ruhig, Philipp (1747). *Littauisch-Deutsches und Deutsch-Littauisches Lexicon*. 2 vols. Königsberg: Hartung.

Rūķe-Draviņa, Velta (1955). Die Benennungen des Wacholders im Baltischen. *ORBIS*: *Bulletin International de Documentation Linguistique* 4/1, 390–409.

Rūķe-Draviņa, Velta (1964). Die Expansion der russischen Buchweizenbenennungen nach Westen. *Scando-Slavica* 10, 107–131.

Rundberget, Bernt et al. (2020). Bloomery ironmaking in Latvia — a comparative study of Iron Age and medieval technologies. *Historical Metallurgy* 52/2, 96–109.

Saag, Lehti et al. (2017). Extensive Farming in Estonia Started through a Sex-Biased Migration from the Steppe. *Current Biology* 27, 2185–2193.

Saag, Lehti et al. (2019). The Arrival of Siberian Ancestry Connecting the Eastern Baltic to Uralic Speakers further East. *Current Biology* 29, 1701–1711.

Saarikivi, Janne (2004). Is there Palaeo-European substratum interference in western branches of Uralic? *Suomalais-Ugrilaisen Seuran Aikakauskirja* 90, 187–214.

Saarikivi, Janne (2011). Saamelaiskielet — nykypäivää ja historiaa. In: Irja Seurujärvi-Kari, Petri Halinen & Risto Pulkkinen (eds.), *Saamentutkimus tänään*. Helsinki: Suomalaisen Kirjallisuuden Seura, 77–119.

Saarikivi, Janne (2018). Finnic and other western Uralic borrowings in Permian. In: Sampsa Holopainen & Janne Saarikivi (eds.), Περὶ ὀρθότητος ἐτύμων: *Uusiutuva uralilainen etymologia*. Helsinki, 269–355.

Saarikivi, Janne (2022). The divergence of Proto-Uralic and its offspring: A descendent reconstruction. In: Marianne Bakró-Nagy, Johanna Laakso & Elena Skribnik (eds.), *The Oxford Guide to the Uralic Languages*. Oxford: University Press, 28–58.

Sabaliauskas, Algirdas (1959). Относительно происхождения названия растений в балтийских языках. In: Ēvalds Sokols (ed.), *Rakstu krājums* [Gedenkschrift Endzelīns]. Rīga: Zinātņu akadēmija, 219–242.

Sabaliauskas, Algirdas (1960a). О происхождении названия мака в балтийских языках. *Краткие сообщения Института славяноведения* 28, 70–73.

Sabaliauskas, Algirdas (1960b). Dėl kai kurių baltų kalbų žemės ūkio augalų pavadinimų kilmės. *Lietuvių kalbotyros klausimai* 3, 257–267.

Sabaliauskas, Algirdas (1968). Baltų kalbų naminių gyvulių pavadinimai. *Lietuvių kalbotyros klausimai* 10, 101–190.

Sabaliauskas, Algirdas (1974). Zur Herkunft von litauisch *paĩvė*. *Zeitschrift für Slawistik* 19, 200–211.

Sabaliauskas, Algirdas (1990). *Lietuvių kalbos leksika*. Vilnius: Mokslas.

Sadziński, Witold & Witczak, Krzysztof T. (2016). Rozważania o nazwach bałtyckich fok. *Pomorania Antiqua* 25, 49–66.

Salminen, Tapani (1999). Euroopan kielet muinoin ja nykyisin. In: Paul Fogelberg (ed.), *Pohjan poluilla*. Helsinki: Societas Scientiarum Fennica, 13–26.

Salminen, Tapani (2002). Problems in the taxonomy of the Uralic languages in the light of modern comparative studies. In: Александр Е. Кибрик (ed.), *Лингвистический беспредел*. Москва: Московский государственный университет, 44–55.

Salmons, Joseph [Joe] (1992). Northwest Indo-European Vocabulary and Substrate Phonology In: *Perspectives on Indo-European Language, Culture and Religion* [FS Polomé]. McLean, Virginia: Institute for the Study of Man, 265–279.

Salmons, Joseph (2004). How (non-)Indo-European is the Germanic lexicon? … And what does that mean? In: Irma Hyvärinen, Petri Kallio & Jarmo Korhonen (eds.), *Etymologie, Entlehnungen und Entwicklungen* [FS Koivulehto]. Helsinki: Société Néophilologique, 311–321.

Sammallahti, Pekka (1977). Suomalaisten esihistorian kysymyksiä. *Virittäjä* 81/2, 119–136.

Sammallahti, Pekka (1988). Historical phonology of the Uralic languages with special reference to Samoyed, Ugric and Permic. In: Denis Sinor (ed.), *The Uralic languages: description, history and foreign influences*. Leiden: Brill, 478–554.

Sammallahti, Pekka (1998). *The Saami Languages: an Introduction*. Kárášjohka: Davvi Girji.

Sammallahti, Pekka (1999). The Indo-European loanwords in Saami. In: Christian Carpelan, Asko Parpola & Petteri Koskikallio (eds.), *Early contacts between Uralic and Indo-European: linguistic and archaeological considerations*. Helsinki: Suomalais-Ugrilainen Seura, 397–415.

Sauer, Jonathan D. (1988). *Plant Migration: The Dynamics of Geographic Patterning in Seed Plant Species*. Berkeley: University of California Press.

Saukkonen, Pauli (1962). Eräs etymologiointiperiaate. *Virittäjä* 66, 342–344.

Sausverde, Ērika (1996). *Seewörter* and Substratum in Germanic, Baltic and Baltic Finno-Ugric Languages. In: Karlene Jones-Bley & Martin E. Huld (eds.), *The Indo-Europeanization of Northern Europe*. Washington, DC: Institute for the Study of Man, 133–147.

Шахматовъ, Алексѣй А. (1893). *Изслѣдованія в области русской фонетики*. Варшава: Варшавскій учебный округъ.

Шахматовъ, Алексѣй А. (1915). *Очеркъ древнѣйшаго періода исторіи русскаго языка*. Петроградъ: Императорская академия наукъ.

Шахматовъ, Алексѣй А. (1916a). *Повѣсть временныхъ лѣтъ* I. *Вводная часть. Текстъ. Примѣчанія*. Петроградъ: Археографическая комиссия.

Шахматовъ, Алексѣй А. (1916b). *Введеніе въ курсъ исторіи русскаго языка* I. *Историческій процессъ образованія русскихъ племенъ и нарѣчій*. Петроградъ: Научное дѣло.

Schaarschmidt, Gunter (1997). *A Historical Phonology of the Upper and Lower Sorbian Languages*. Heidelberg: Winter.

Schaffner, Stefan (2001). *Das Vernersche Gesetz und der innerparadigmatische grammatische Wechsel des Urgermanischen im Nominalbereich*. Innsbruck: Institut für Sprachen und Literaturen.

Schiller, Karl & Lübben, August (1875–1881). *Mittelniederdeutsches Wörterbuch*. 6 vols. Bremen: Kühtmann/Fischer.

Schindler, Jochem (1970). [Review of] Anttila, Raimo: Proto-Indo-European Schwebeablaut. *Kratylos* 15, 146–152.

Schmalstieg, William R. (1983). Slavic *kamy* and *mati*: a structural view. *Journal of Indo-European Studies* 11, 151–165.

Schmeller/Bergmann = Joseph Bergmann (1855, ed.), *Johann Andreas Schmeller's sogenanntes Cimbrisches Wörterbuch*. Wien: Kaiserlich-königliche Hof- und Staatsdruckerei.

Schmid, Wolfgang P. (1968). Alteuropäisch und Indogermanisch. *Abhandlungen der geistes- und sozialwissenschaftlichen Klasse* 1968/6.

Schmid, Wolfgang P. (1977). Sprachwissenschaftliche Bemerkungen zu den Wörtern für Bauer und Dorf im Slavischen. In: Herbert Jankuhn, Rudolf Schützeichel & Fred Schwind (eds.), *Das Dorf der Eisenzeit und des frühen Mittelalters*. Göttingen: Vandenhoeck & Ruprecht, 41–61.

Schmidt, Johannes (1875). *Zur Geschichte des indogermanischen Vocalismus. Zweite Abteilung*. Weimar: Böhlau.

Schmidt, Johannes (1891). Die urheimath der Indogermanen und das europäische zahlsystem. *Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin aus dem Jahre 1890*.

Schmidt, Johannes (1895). *Kritik der Sonantentheorie*. Weimar: Böhlau.

Schmidt, Karl H. (1977). Der Beitrag der keltiberischen Inschrift von Botorrita zur Rekonstruktion der protokeltischen Syntax. *Word* 28, 51–62.

Schmidt, Peter [= Šmits, Pēteris] (1930). Sprachliche Zeugnisse über die Urheimat der Balten und ihre Kultur. In: Fritzis Adamovičs (ed.), *Die Letten*. Riga: Walters & Rapa, 71–88.

Schmidt-Brandt, Robert (1967). *Die Entwicklung des indogermanischen Vokalsystems*. Heidelberg: Groos.

Schrader, Otto (1883). *Sprachvergleichung und Urgeschichte*. Jena: Costenoble.

Schrader, Otto (1901). *Reallexikon der indo-germanischen Altertumskunde*. Strassburg: Trübner.

Schrader/Nehring = Schrader, Otto & Nehring, Alfons (1917–1929). *Reallexikon der indogermanischen Altertumskunde* [22nd edition]. 2 vols. Berlin–Leipzig: De Gruyter.

Schrijver, Peter (1991). *The reflexes of the Proto-Indo-European laryngeals in Latin*. Amsterdam: Rodopi.

Schrijver, Peter (1995a). *Studies in British Celtic Historical Phonology*. Amsterdam–Atlanta: Rodopi.

Schrijver, Peter (1995b). De etymologie van de naam van de Cannenefaten. *Amsterdamer Beiträge zur älteren Germanistik* 41, 13–22.

Schrijver, Peter (1997). Animal, vegetable and mineral: some Western European substratum words. In: Alexander Lubotsky (ed.), *Sound Law and Analogy* [FS Beekes]. Amsterdam–Atlanta: Rodopi, 293–316.

Schrijver, Peter (1999). On Henbane and Early European Narcotics. *Zeitschrift für celtische Philologie* 51, 17–45.

Schrijver, Peter (2001). Lost Languages in Northern Europe. In: Christian Carpelan, Asko Parpola & Petteri Koskikallio (eds.), *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations*. Helsinki: Suomalais-Ugrilainen Seura, 417–425.

Schrijver, Peter (2018). Talking Neolithic: The Case for Hatto-Minoan and its Relationship to Sumerian. In: Guus Kroonen, James P. Mallory & Bernard Comrie (eds.), *Talking Neolithic*. Washington, DC: Institute for the Study of Man, 336–374.

Schuchardt, Hugo (1922). Heimisches und fremdes Sprachgut. *Revista Internacional de los Estudios Vascos* 13, 70–82.

Schuhmann, Roland (2016). Where is the Substrate in the Germanic Lexicon? In: Bjarne S.S. Hansen et al. (eds.), *Etymology and the European Lexicon*. Wiesbaden: Reichert, 377–384.

Schulze, Wilhelm (1910). Etymologisches. *Sitzungsberichte der königliche preussische Akademie der Wissenschaften* 1910, 787–808.

Schulze, Wilhelm (1913). Der Fuchs. *Zeitschrift für vergleichende Sprachforschung* 45, 287–288.

Schumacher, Stefan (2004). *Die keltischen Primärverben*. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.

Schumacher, Stefan (2013). Historische Phonologie. In: Schumacher, Stefan & Matzinger, Joachim. *Die Verben des Altalbanischen*. Wiesbaden: Harrassowitz, 205–276.

Schuster-Šewc = Schuster-Šewc, Heinz (1978–1996). *Historisch-etymologisches Wörterbuch der ober- und niedersorbischen Sprache*. 24 issues. Bautzen: Domowina.

Schützeichel, Rudolf (2004). *Althochdeutscher und altsächsischer Glossenwortschatz*. 12 vols. Niemeyer: Tübingen.

Seewald, P. (1865). Einige lettische Ausdrucksformen. *Magazin, herausgegeben von der Lettisch-Literärischen Gesellschaft* 13/2, 38–74.

Sehwers, Johann (1936). *Sprachlich-kulturhistorische Untersuchungen: vornehmlich über den deutschen Einfluss im Lettischen*. Wiesbaden: Harrassowitz.

Senn, Alfred (1925). *Germanische Lehnwortstudien*. Heidelberg: Winter.

Senn, Alfred (1933). Nachträge zu Kluge's Wörterbuch. *The Journal of English and Germanic Philology* 32, 504–529.

Senn, Alfred (1943). The Historical Development of the Lithuanian Vocabulary. *Bulletin of the Polish Institute of Arts and Sciences in America* 1, 946–969.

Seržant, Ilja A. (2006). Die Vermittlungsrolle des Hochlettischen bei den altrussischen und litauischen Entlehnungen im Lettischen. *Acta Linguistica Lituanica* 55, 89–106.

Seržant, Ilja A. [= Сержант, И.А.] (2008). Относительная хронология процессов палатализации прабалтийских **k* и **g* в латышском. *Вопросы языкознания* 2008/1, 121–129.

Setälä, Eemil, N. (1902). Zur etymologie von Sampo. *Finnisch-Ugrische Forschungen* 2, 141–164.

Setälä, Eemil, N. (1909). Beiträge zur finnisch-ugrischen wortkunde 9. Fi. *kataja* 'wacholder'. *Finnisch-Ugrische Forschungen* 9, 126–128.

Setälä, Eemil, N. (1929). Itämerensuomalaisten kielten kosketukset slaavilaisten kanssa. *Suomalais-Ugrilaisen Seuran Aikakauskirja* 43, 30–39.

Shevelov, George Y. (1964). *A Prehistory of Slavic*: *The Historical Phonology of Common Slavic*. Heidelberg: Winter.

Shevelov, George Y. (1979). *A Historical Phonology of the Ukrainian Language*. Heidelberg: Winter.

Sillasoo, Ülle & Hiie, Sirje (2007). An archaeobotanical approach to investigating food of the Hanseatic period in Estonia. In: Sabine Karg (ed.), *Medieval Food Traditions in Northern Europe*. Copenhagen: National Museum of Denmark, 73–96.

Simon, Zsolt (2008). How to find the Proto-Indo-European homeland? A methodological essay. *Acta Antiqua Academiae Scientiarum Hungaricae* 48, 289–303.

Simon, Zsolt (2022). The Hurro-Urartian loan contacts of Armenian: A revision. *Hungarian Assyriological Review* 3, 63–89.

Simon, Zsolt (forthc.). Indo-European Substrates: The Problem of the Anatolian Evidence. To appear in: Guus Kroonen (ed.), *Sub-Indo-European Europe*. De Gruyter Mouton.

Skach, Elisabeth (2010). Die slavische Monophthongierung **ej* > *ī* im Lichte der Lehnbeziehungen zum Baltischen und Ostseefinnischen. *Wiener slavistisches Jahrbuch* 56, 137–140.

Skardžius, Pranas (1931). *Die slavischen Lehnwörter im Altlitauischen*. Kaunas: Spindulys.

Skardžius, Pranas (1941). *Lietuvių kalbos žodžių daryba*. Vilnius: Lietuvos mokslų akademija.

Skeat, Walter W. (1882). *An Etymological Dictionary of the English Language*. Oxford: Clarendon Press.

Skjærvø, P. Oktor (2002). *Khotanese manuscripts from Chinese Turkestan in the British Library*: *a complete catalogue with texts and translations*. London: British Library.

Skok = Skok, Petar (1971–1974). *Etimologijski rječnik hrvatskoga ili srpskoga jezika*. 3 vols. Zagreb: Akademija znanosti i umjetnosti.

Скворцов = Михаил И. Скворцов (1985, ed.), *Чувашско-русский словарь*. Москва: Русский язык.

Sławski, Franciszek (1960). Dubrownicka nazwa łabędzia. *Rocznik sławistyczny* 21, 37–40.

Sławski SEJP = Franciszek Sławski (1952–1982). *Słownik etymologiczny języka polskiego*. 5 vols. [**a*–**łżywy*]. Kraków: Towarzystwo miłośników języka polskiego.

Sławski SP = Franciszek Sławski (1974–2001, ed.), *Słownik prasłowiański*. 8 vols. [**a–*gyža*]. Wrocław: Ossolineum.

van Sluis, Paulus (2022). Beekeeping in Celtic and Indo-European. *Studia Celtica* 56, 1–28.

van Sluis, Paulus (forthc.). Substrate alternations in Celtic. To appear in: Guus Kroonen (ed.), *Sub-Indo-European Europe*. De Gruyter Mouton.

van Sluis, Paulus, Anders R. Jørgensen, & Guus Kroonen (2023). European Prehistory between Celtic and Germanic: The Celto-Germanic isoglosses revisited. In: Kristian Kristiansen, Guus Kroonen & Eske Willerslev (eds.), *The Indo-European Puzzle Revisited*. Cambridge: University Press, 193–243.

Ślusarska, Katarzyna (2021). Wild Resources in the Economy of Bronze and Early Iron Ages Between Oder and Bug Rivers — Source Overview. *Open Archaeology* 7, 177–210.

Smalinskienė, Gražina (1994). Lie. *salà*, la. *sala* „kaimas“. *Lietuvių kalbotyros klausimai* 31, 177–180.

Сметанина, Зоя В. & Иванова, Галина А. (2018). Фонематические варианты в «Областном словаре вятских говоров». In: Людмила В. Калинина (ed.), *Семантика. Функционирование. Текст* [FS Черновая]. Киров: Радуга, 203–210.

Śmieszek, Antoni (1909). Etymologica. *Materyały i prace Komisyi Językowej Akademii Umiejętności w Krakowie* 4, 391–408.

Smith, Emil (1910). Om opindelsen til fiskenavene *sild* og *sil*. In: Magnus Olsen (ed.), *Maal og minne*. Kristiania: Bymaals-Laget, 139–142.

Smoczyński, Wojciech (1983). Staropruskie lekcje i etymologie I. *Biuletyn Polskiego Towarzystwa Językoznawczego* 40, 167–183.

Smoczyński, Wojciech (1989). *Studia bałto-słowiańskie* I. Wrocław: Ossolineum.

Smoczyński, Wojciech (1998). Etimologijos pastabos V. Nomina postverbalia *indas*, *pergas*, *išeiga* ir kt. *Baltistica* 33, 251–260.

Smoczyński, Wojciech (1999). Заметки о балто-славянских рефлексах и.-е. ларингальных. *Kalbotyra* 48/2, 7–26.

Smoczyński, Wojciech (2000). *Untersuchungen zum deutschen Lehngut im Altpreussischen*. Kraków: Uniwersytet Jagielloński.

Smoczyński, Wojciech (2006). *Laringalų teorija ir lietuvių kalba*. Vilnius: Lietuvių kalbos institutas.

Smoczyński, Wojciech (2018). *Lithuanian Etymological Dictionary*. Berlin: Peter Lang.

Smoczyński SEJL² = Smoczyński, Wojciech. *Słownik etymologiczny języka litewskiego* (online version). Accessed at https://rromanes.org/pub/alii/.

Snoj, Marko (2003). *Slovenski etimološki slovar*. Ljubljana: Modrijan.

Соболевскій, Алексѣй И. (1910). Мелкія замѣтки по славянской и русской фонетикѣ. *Русскій филологическій вѣстникъ* 64, 178–179.

Соболевскій, Алексѣй И. (1914). Нѣсколько замѣтокъ по славянскимъ вокализму и лексикѣ. *Русскій филологическій вѣстникъ* 71, 431–452.

Solmsen, Felix (1904). Slavische etymologien. *Zeitschrift für vergleichende Sprachforschung* 37, 575–601.

Sommer, Ferdinand (1905). *Griechische Lautstudien*. Strassburg: Trübner.

Sommer, Ferdinand (1914). *Die indogermanischen* iā- *und* io-*Stämme im Baltischen*. Leipzig: Teubner.

Soosaar, Sven-Erik (2021). Võimalikke protoeuroopa substraadi jälgi eesti keele dendronüümides. *Emakeele Seltsi aastaraamat* 66, 295–309.

Šorgo, Aljoša (2020). Characteristics of lexemes of a substratum origin in Proto-Germanic. In: Romain Garnier (ed.), *Loanwords and substrata*. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft, 427–472.

Сорвачева, Валентина А. & Безносикова, Люция М. (1990). *Удорский диалект коми языка*. Москва: Наука.

Specht, Franz (1932). Die Flexion der *n*-Stämme im Baltisch-Slavischen und Verwandtes. *Zeitschrift für vergleichende Sprachforschung* 59, 213–298.

Specht, Franz (1935). Zur baltisch-slavischen Spracheinheit. *Zeitschrift für vergleichende Sprachforschung* 62, 248–258.

Specht, Franz (1947). *Der Ursprung der Indogermanischen Deklination*. Göttingen: Vandenhoeck & Ruprecht.

van der Staaij, Robert J. (1975). *A Reconstruction of Proto-Italic*. Leiden: University Press.

Stang, Christian S. (1957). *Slavonic Accentuation*. Oslo: Universitetsforlaget.

Stang, Christian S. (1966). *Vergleichende Grammatik der Baltischen Sprachen*. Oslo: Universitetsforlaget.

Stang, Christian S. (1972). *Lexikalische Sonderübereinstimmungen zwischen dem Slavischen, Baltischen und Germanischen*. Oslo: Universitetsforlaget.

Stankus, Jonas (2001). Geležies gamybos lietuvoje apžvalga. *Lietuvos archeologija* 21, 171–182.

Starostin, Sergei, Anna Dybo & Oleg Mudrak (2003). *An Etymological Dictionary of Altaic Languages*. Leiden: Brill.

Steingass, Francis J. (1892). *A Comprehensive Persian-English Dictionary*. London: Routledge & K. Paul.

Steinitz, Wolfgang (1964). Zur Periodisierung der alten baltischen und germanischen Lehnwörter im Finnischen. *Wissenschaftliche Zeitschrift der Humbolt-Universität zu Berlin. Gesellschafts- und sprachwissenschaftliche Reihe* 13, 335–338.

Stender-Petersen, Adolf (1927). *Slavisch-Germanische Lehnwortkunde*. Göteborg: Elander.

Stifter, David (1998). Old Irish ²*fén* 'bog'. *Die Sprache* 40, 226–228.

Stifter, David (2019). An apple a day … *Indogermanische Forschungen* 124, 171–218.

Stifter, David (2023). With the Back to the Ocean. The Celtic Maritime Vocabulary. In: Kristian Kristiansen, Guus Kroonen & Eske Willerslev (eds.), *The Indo-European Puzzle Revisited*. Cambridge: University Press, 172–192.

Stifter, David (forthc.). Prehistoric layers of loanwords in Old Irish. To appear in: Guus Kroonen (ed.), *Sub-Indo-European Europe*. De Gruyter Mouton.

Stika, Hans-Peter & Heiss, Andreas G. (2013). Bronzezeitliche Landwirtschaft in Europa — Der Versuch einer Gesamtdarstellung des Forschungsstandes. In: Karl-Heinz Willroth (ed.), *Siedlungen der älteren Bronzezeit*. Neumünster: Wachholtz, 183–216.

Stockfleth, Nils Vibe (1852). *Norsk-lappisk Ordbog*. Christiania: Cappelen.

Stokes, Whitley (1894). *Urkeltischer Sprachschatz* (= *Vergleichendes Wörterbuch der indogermanischen Sprachen* [4th edition], Vol. 2). Göttingen: Vandenhoeck & Ruprecht.

Strömberg, Reinhold (1944). *Griechische Wortstudien*. Göteborg: Elander.

Suhonen, Seppo (1973). *Die jungen lettischen Lehnwörter im Livischen*. Helsinki: Société Finno-Ougrienne.

Suhonen, Seppo (1988). Die baltischen Lehnwörter der finnisch-ugrischen Sprachen. In: Denis Sinor (ed.), *The Uralic Languages*: *Description, History and Foreign Influences*. Leiden–New York–Copenhagen: Brill, 596–615.

Suolahti, Hugo (1899). *Die althochdeutschen Tiernamen*. Darmstadt: Otto.

Suolahti, Hugo (1909). *Die deutschen Vogelnamen*. Straßburg: Trübner.

Szemerényi, Oswald (1967). Славянская этимология на индоевропейском фоне. *Вопросы языкознания* 1967/4, 3–24.

Takács, Gábor (2001). *Etymological Dictionary of Etyptian* II. *b-, p-, f-*. Leiden–Boston–Köln: Brill.

Tamai, Tatsushi (2020). *Tocharian Mss. of Weber and Macartney* [Unpublished]. Accessed online at https://academia.edu/43390019.

Tamm, Fredrik (1881). Svenska ord belysta genom slaviska och baltiska språken. *Upsala Universitets Årsskrift* 1881.

Терентьев, Владимир А. (1990). Corrections to the "Suomen kielen etymologinen sanakirja" concerning Germanic, Baltic and Slavic loanwords. In: *Uralo-Indogermanica: Балто-славянские языки и проблема урало-индоевропейских связей* II. Москва: Академия наук СССР, 30–32.

Thomsen, Vilhelm (1870). *Über den Einfluss der germanischen Sprachen auf der Finnisch-Lappischen*. Halle: Buchhandlung des Waisenhauses.

Thomsen, Vilhelm (1877). *The Relations between Ancient Russia and Scandinavia and the Origin of the Russian State*. Oxford: Parker.

Thomsen, Vilhelm (1890). *Berøringer mellem de finske og de baltiske* (*litauisk-lettiske*) *Sprog*. København: Blanco Lunos.

Thomson, Alexander I. (1927). Beiträge zur Geschichte des Diphthongs ы in den slavischen Sprachen. *Zeitschrift für Slavische Philologie* 4, 342–356.

Thörnqvist, Clara (1948). *Studien über die nordischen Lehnwörter im Russischen*. Uppsala: Almqvist & Wiksell.

Thorsø, Rasmus (forthc.). Alternation of diphthong and monophthong in Armenian

words of substrate origin. To appear in: Guus Kroonen (ed.), *Sub-Indo-European Europe*. De Gruyter Mouton.

Thorsø, Rasmus, Andrew Wigman, Anthony Jakob et al. (2023). Word mining: metal names and the Indo-European dispersal. In: Kristian Kristiansen, Guus Kroonen & Eske Willerslev (eds.), *The Indo-European Puzzle Revisited*. Cambridge: University Press, 105–126.

Threatte, Leslie (1980). *The Grammar of Attic Inscriptions* I. *Phonology*. Berlin–New York: De Gruyter.

Thurneysen, Rudolf (1946). *A Grammar of Old Irish* [Revised and Enlarged Edition]. Dublin: Institute for Advanced Studies.

Toivonen, Yrjö H. (1917). Huomioita lainasanatutkimuksemme alalta. *Suomalais-Ugrilaisen Seuran Aikakauskirja* 34, 1–47.

Toivonen, Yrjö H. (1928). Über alter und entwicklung des ackerbaus bei den finnisch-ugrischen völkern. *Mémoires de la Société Finno-ougrienne* 58, 229–240.

Толстой, Никита И. (1969). *Славянская географическая терминология*. Москва: Наука.

Tomaschek, Wilhelm (1883). Ethnologisch-linguistische Forschungen über den Osten Europa's. *Das Ausland* 36, 701–706.

Tomaschek, Wilhelm (1889). Kritik der ältesten Nachrichten über den skythischen Norden. *Sitzungsberichte der philosophisch-historischen Classe der kaiserlichen Akademie der Wissenschaften* 107, 1–70.

Топоров, Владимир Н. (1972). О балтийском элементе в подмосковье. *Baltistica* 1 *priedas*, 185–224.

Топоров, Владимир Н. (1988–1997). Балтийский элемент в гидронимии поочья. *Балто-славянские исследования* 1986, 154–177; 1987, 47–69; 1988–1996, 276–310.

Топоров ПЯ = Топоров, Владимир Н. (1975–1990). *Прусский язык*. 5 vols. Москва: Наука.

Топоров, Владимир Н. & Трубачев, Олег Н. (1962). *Лингвистический анализ гидронимов верхнего поднепровья*. Москва: Академия наук.

Töppen, Max (1867). Einige Reste der altpreussischen Sprache nebst antiquarischen Bemerkungen. *Altpreußische Monatsschrift* 4, 136–156.

Torp, Alf (1919). *Nynorsk etymologisk ordbok*. Kristiania: Aschehoug.

Trautmann, Reinhold (1910). *Die altpreussischen Sprachdenkmäler*. Göttingen: Vandenhoeck & Ruprecht.

Trautmann, Reinhold (1923). *Baltisch-Slavisches Wörterbuch*. Göttingen: Vandenhoeck & Ruprecht.

Treimer, Karl (1954). *Ethnogenese der Slawen*. Wien: Gerold.

Tremblay, Xavier (2004). Sur χαλκός, lit. *geležìs* et turc *qoruyžin*. *Historische Sprachforschung* 117, 238–248.

Трубачев, Олег Н. (1960). *Происхождение названий домашних животных в славянских языках*. Москва: Академия наук СССР.

Трубачев, Олег Н. (1971). Заметки по этимологии и сравнительной грамматике. *Этимология* 1968, 24–67.

Трубачев, Олег Н. (1994). Праславянское лексическое наследие и древнерусская лексика дописьменного периода. *Этимология* 1991–1993, 3–23.

[Трубачев ЭССЯ — see ЭССЯ under Abbreviated Titles.]

Тумашева, Диляра Г. (1992). *Словарь диалектов сибирских татар*. Казань: Издательство Казанского университета.

Uhlenbeck, Christianus C. (1900). *Kurzgefasstes etymologisches Wörterbuch der gotischen Sprache* [2nd edition]. Amsterdam: Müller.

Ulmann, Carl Christian (1872). *Lettisches Wörterbuch. Erster Theil: Lettisch-deutsches Wörterbuch*. Riga: Brutzer.

Untermann, Jürgen (2000). *Wörterbuch des Oskisch-Umbrischen*. Heidelberg: Winter.

Uotila, Eeva (1983). Baltofennica III: Finnish *viekas*, *veikeä*, *vaikku* ~ Baltic **veika-*. *Euroasiatica* 4/6. Napoli: Istituto Universitario Orientale, 5–12.

Uotila, Eeva (1986). Baltic impetus on the Baltic Finnic diphthongs. *Finnisch-Ugrische Forschungen* 47, 206–222.

Uotila, Eeva (1990). More Baltic loans from ancient everyday contexts: Fi. *suhta* and *hinta*. In: *Uralo-Indogermanica: Балто-славянские языки и проблема урало-индоевропейских связей* II. Москва: Академия наук СССР, 134–140.

de Vaan, Michiel (1999). The PIE root structure **Te(R)Dʰ-*. *Historische Sprachforschung* 112, 1–25.

de Vaan, Michiel (2003). *The Avestan Vowels*. Amsterdam–New York: Rodopi.

de Vaan, Michiel (2008). *Etymological Dictionary of Latin and the other Italic Languages*. Leiden–Boston: Brill.

de Vaan, Michiel (2018). The phonology of Albanian. In: Jared Klein, Brian Joseph & Matthias Fritz (eds.), *Handbook of Comparative and Historical Indo-European Linguistics*. Berlin: De Gruyter, 1732–1749.

Vaba, Lembit (1990a). Die baltischen Sonderentlehnungen in den ostseefinnischen Sprachen. In: Sari Vaula (ed.), *Itämerensuomalaiset kielikontaktit*. Helsinki: Valtion painatuskeskus, 125–139.

Vaba, Lembit (1990b). Gibt es ein baltisches Fragment im estnischen Wortschatz der Waldimkerei? *Linguistica Uralica* 26/3, 173–179.

Vaba, Lembit (1992). [Review of] A.D. Kylstra, Sirkka-Liisa Hahmo, Tette Hostra, Osmo Nikkilä, Lexikon der älteren germanischen Lehnwörter in den ostseefinnischen Sprachen, Bd. I. A–J. *Linguistica Uralica* 28, 219–223.

Vaba, Lembit (1997). Ergänzungen zur Substitution des baltischen wortanlautenden **st*. *Linguistica Uralica* 33/3, 177–180.

Vaba, Lembit (1998). Die Rolle des altpreussischen Sprachmaterials in etymologischen Untersuchungen ostseefinnischer Baltismen. In: Wojciech Smoczyński (ed.), *Colloquium Pruthenicum Secundum*. Kraków: Universitas, 177–175,

Vaba, Lembit (2011). Balti laenude uurimine avab meie kauget minevikku. *Keel ja Kirjandus* 54, 734–763.

Vaillant, André (1929). Deux emprunts slaves au roman de Dalmatie. *Revue des études slaves* 9, 270–272.

Vaillant, André (1950). *Grammaire comparée des langues slaves* I. *Phonétique*. Lyon: IAC.

Vaillant, André (1958). *Grammaire comparée des langues slaves* II. *Morphologie. Première Partie*: *Flexion nominale*. Lyon: IAC.

Vaillant, André (1974). *Grammaire comparée des langues slaves* IV. *La formation des noms*. Paris: Klincksieck.

Vanagienė, Birutė (2014). *Šiaurės vakarų žemaičių žodynas* I. *A–O*. Lietuvių kalbos institutas.

Варбот, Жанна Ж. (1984). *Праславянская морфонология словообразование и этимология*. Москва: Наука.

Васильев, Василий Л. (2012). *Славянские топонимические древности Новгородской земли*. Москва: Языки славянской культуры.

Васильев, Василий Л. (2015). Проблематика изучения гидронимии балтийского происхождения на территории России. *Linguistica* 55, 173–186.

Васильев, Леонид Л. (1929). *О значении каморы в некоторых древнерусских памятниках XVI–XVII веков*. Ленинград: Академия наук.

Vasmer, Max [= М. Р. Фасмеръ] (1909). *Греко-славянскіе этюды* III. *Греческіе заимствованія въ русскомъ языкѣ*. Санктпетербургъ: Академія наукъ.

Vasmer, Max (1913). Kritisches und Antikritisches zur neueren slavischen Etymologie V. *Rocznik slawistyczny* 6, 172–214.

Vasmer, Max (1922). Prie germanybių baltų kalbose. *Švietimo darbas* 1922/3–6, 272–274.

Vasmer, Max (1932). Beiträge zur historischen Völkerkunde Osteuropas I. Die Ostgrenze der baltischen Stämme. *Sitzungsberichte der Preussischen Akademie der Wissenschaften. Philosophisch-historische Klasse* 1932, 637–666.

Vasmer, Max (1938). [Review of] J. Mikkola, Die älteren Beziehungen zwischen Ostseefinnischen und Russisch. *Zeitschrift für Slavische Philologie* 15, 448–455.

[Vasmer REW — see REW under Abbreviated Titles.]

Вахрос, Игорь С. (1963). К истории бани в свете ее наименований в славянских языках. *Scando-Slavica* 9, 153–168.

Vennemann, Theo (1995). Etymologische Beziehungen im Alten Europa. *Der Ginkgo-Baum: Germanistisches Jahrbuch für Nordeuropa* 13, 39–115.

Vennemann, Theo (1998). Germania Semitica. Biene und Imme. *Sprachwissenschaft* 23, 471–487.

Vennemann, Theo (2003). Zur Frage der vorindogermanischen Substrate in Mittel- und Westeuropa. In: Theo Vennemann. *Europa Vasconica — Europa Semitica*. Berlin–New York: Mouton de Gruyter.

Vennemann, Theo (2011). English as a Contact Language: Typology and Comparison. *Anglia* 129, 217–257.

Вершинин = Вершинин, Валерий И. (2004–2011). *Этимологический словарь мордовских (эрзянского и мокшанского) языков*. 5 vols. Йошкар-Ола.

Viitso, Tiit-Rein (1978). The history of Finnic *õ* in the first syllable. *Советское финно-угроведение* 14, 86–106.

Viitso, Tiit-Rein (1981). Finnic Gradation: types and genesis. *Советское финно-угроведение* 17/3, 176–185.

Viitso, Tiit-Rein (1990). On the earliest Finnic and Balto-Slavic contacts. In: Sari Vaula (ed.), *Itämerensuomalaiset kielikontaktit*. Helsinki: Valtion Painatuskeskus, 140–147.

Viitso, Tiit-Rein (1992). Finnic and its prehistoric Indo-European neighbors. In: László Honti et al. (eds.), *Finnisch-ugrische Sprachen zwischen dem germanischen und dem slavischen Sprachraum*. Amsterdam–Atlanta: Rodopi, 185–196.

Viitso, Tiit-Rein (1998). Possible prehistoric contacts of Livonian. *Linguistica Uralica* 34, 8–22.

Viitso, Tiit-Rein & Ernštreits, Valts (2012). *Līvõkīel-ēstikīel-leţkīel sõnārõntõz = Livonian-Estonian-Latvian Dictionary*. Tartu–Rīga: Tartu Ülikool.

Villanueva Svensson, Miguel (2008). [Reivew of] Wojciech Smoczyński, Laringalų teorija ir lietuvių kalba. *Baltistica* 43, 115–123.

Villanueva Svensson, Miguel (2009). Indo-European **sḱ* in Balto-Slavic. *Baltistica* 44, 5–24.

Villanueva Svensson, Miguel (2010). Baltic *sta*-presents and the Indo-European desiderative. *Indogermanische Forschungen* 115, 204–233.

Villanueva Svensson, Miguel (2011). Indo-European long vowels in Balto-Slavic. *Baltistica* 46, 5–38.

Villanueva Svensson, Miguel (2014). On the relationship between West Baltic and East Baltic. In: Tatjana Civjan, Marija Zavjalova & Artūras Judžentis (eds.), *Baltai ir slavai: Dvasinių kultūrų sankirtos*. Vilnius: Versmė, 162–176.

Villanueva Svensson, Miguel (2015). The development of **ōu* in Baltic. *Indogermanische Forschungen* 120/1, 313–328.

Vine, Brent (1999). Greek ῥίζα 'root' and "Schwa secundum". In: Vyacheslav V. Ivanov & Brent Vine (eds.), *UCLA Indo-European Studies* I, 5–30.

Vine, Brent (2011). Initial **mo-* in Latin and Italic. *Münchener Studien zur Sprachwissenschaft* 65, 261–286.

Viredaz, Rémy (2020). Notes d'étymologie slave. In: Harald Bichlmeier, Ondřej Šefčík & Roman Sukač (eds.), *Etymologus. Festschrift for Václav Blažek*. Hamburg: Baar, 403–422.

Vosylytė, Klementina (2013). *Kupiškėnų žodynas* IV. *Š–Ž*. Vilnius: Lietuvių kalbos institutas.

Vovin, Alexander (2005). The End of the Altaic Controversy. *Central Asiatic Journal* 49, 71–132.

de Vries, Jan W. (1962). *Altnordisches etymologisches Wörterbuch*. Leiden: Brill.

de Vries, Jan W. (1971). *Nederlands Etymologisch Woordenboek*. Leiden: Brill.

Vycichl, Werner (1983). *Dictionnaire étymologique de la langue copte*. Leuven: Peeters.

Vycichl, Werner (1990). *La vocalisation de la langue égyptienne* I. *La phonétique*. Cairo: Insitut Français d'archéologie orientale.

Wagner, Heinrich (1981). Origins of Pagan Irish Religion. *Zeitschrift für celtische Philologie* 38, 1–28.

Wälchli, Bernhard (1997). Mordwinisch und ältere baltische Wortsemantik. *Finnisch-Ugrische Forschungen* 54, 305–324.

Walde, Alois (1910). *Lateinisches etymologisches Wörterbuch*. Heidelberg: Winter.

Walde/Hofmann = Johann B. Hofmann (1938–1954, ed.), *Lateinisches etymologisches Wörterbuch von A. Walde* [3rd edition]. 2 vols. Heidelberg: Winter.

Walde/Pokorny = Julius Pokorny (1927–1934, ed.), *Alois Walde. Vergleichendes Wörterbuch der Indogermanischen Sprachen*. 3 vols. Berlin–Leipzig: De Gruyter.

Weber, Dieter (1997). Osset. *bælon*, D *bælæu* 'Taube' und lit. *balañdis*. *Res Balticae* 1997, 119–128.

Weiss, Michael (1994). Life Everlasting: Latin *iūgis* "everflowing", Greek ὑγιής "healthy", Gothic *ajukdūþs* "eternity" and Avestan *yauuaēǰī-* "living forever". *Münchener Studien zur Sprachwissenschaft* 55, 131–156.

Weiss, Michael (2020). *Outline of the Historical and Comparative Grammar of Latin* [2nd edition]. Ann Arbor: Beech Stave Press.

Welcome, Ashton K. & van Wyk, Ben-Erik (2019). Spatial patterns, availability and cultural preferences for edible plants in southern Africa. *Journal of Biogeography* 47, 584–599.

Wexler, Paul (1977). *A Historical Phonology of the Belorussian Language*. Heidelberg: Winter.

Wichmann, Yrjö (1902). Samojedisches lehngut im syrjänischen. *Finnisch-Ugrische Forschungen* 2, 165–183.

Wichmann, Yrjö (1906). Zur geschichte der finnisch-ugrischen anlautenden *s-* und *č-*laute im tscheremissischen. *Finnisch-Ugrische Forschungen* 6, 17–39.

Wichmann, Yrjö (1911). Zur geschichte der finnisch-ugrischen anlautenden affrikaten nebst einem exkurs über die finnisch-ugrishen anlautenden klusile. *Finnisch-Ugrische Forschungen* 11, 173–290.

Wiik, Kalevi (1992). Suomen kielen ja murteiden syntykysymyksiä. In: Hannu Remes (ed.), *Lähivertailuja* 6. Joensuu: Joensuun yliopisto, 85–110.

Wiklund, Karl B. (1896). *Entwurf einer urlappischen Lautlehre* I. *Einleitung, Quantitätsgesetze, Accent, Geschichte der hauptbetonten Vokale*. Helsinki: Suomalais-Ugrilainen Seura.

Wilkin, Shevan et al. (2021). Dairying enabled Early Bronze Age Yamnaya steppe expansions. *Nature* 598, 629–633.

Winkler, Eberhard & Pajusalu, Karl (2009). *Salis-livisches Wörterbuch*. Tallinn: Teaduste Akadeemia.

Winter, Werner (1950). *Studien zum „Prothetischen Vokal" im Griechischen*. Hamburg: Hansischer Gildenverlag.

Winter, Werner (1962). Die Vertretung indogermanischer Dentale im Tocharischen. *Indogermanische Forschungen* 67, 16–35.

Winter, Werner (1978). The distribution of short and long vowels in stems of the type Lith. *ésti* : *vèsti* : *mèsti* and OCS *jasti* : *vesti* : *mesti* in Baltic and Slavic languages. In: Jacek Fisiak (ed.), *Recent developments in historical phonology*. The Hague: Mouton, 431–446.

Winter, Werner (1982). Indo-European words for 'tongue' and 'fish': A reappraisal. *Journal of Indo-European Studies* 10, 167–186.

Witczak, Krzysztof T. (1991). Armenian *op'i* 'white poplar, *Populus alba* L.' and the development of PIE **ps* in Armenian. *Annual of Armenian Linguistics* 12, 65–75.

Witczak, Krzysztof T. (1996). The Pre-Germanic Substrata and Germanic Maritime Vocabulary. In: Karlene Jones-Bley & Martin E. Huld (eds.), *The Indo-Europeanization of Northern Europe*. Washington, DC: Institute for the Study of Man, 166–180.

Witczak, Krzysztof T. (1997). Millet in Lithuanian and other Indo-European languages. *Baltistica* 32, 25–39.

Witczak, Krzysztof T. (2001). Lith. *atúodogiai*, *atúoriečiai* 'summer rye', *atólas* (m.) 'aftermath, after-grass' and the Indo-European name for 'grass'. *Baltistica* 36, 43–47.

Witczak, Krzysztof T. (2003). *Indoeuropejskie nazwy zbóż*. Łódź: Uniwersytet.

Witczak, Krzysztof T. (2020). Are There Traces of a Finno-Ugric Substratum in Proto-Slavic? *Slavistična revija* 68/1, 73–89.

Witsen, Nicolaas (1785). *Noord en Oost Tartarye* [3rd edition]. Amsterdam.

Worrel, William H. (1934). *Coptic Sounds*. Ann Arbor: University of Michigan.

Wrede, Ferdinand (1891). *Über die Sprache der Ostgoten in Italien*. Strassburg: Trübner.

[Хелимский — see Helimski]

Yakubovich, Ilya (2009). Two Armenian Etymologies. *Caucasian and Near Eastern Studies* 13, 266–272.

Young, Steven (1998). Old Prussian bovids revisited. In: Wojciech Smoczyński (ed.), *Colloquium Pruthenicum secundum*. Kraków: Universitas, 201–216.

Young, Steven (2009). Tone in Latvian borrowings from Old Russian. In: Tijmen Pronk & Rick Derksen (eds.), *Stressing the past: Papers on Baltic and Slavic accentology*. Amsterdam–New York: Rodopi, 177–192.

Zair, Nicholas (2012). *The reflexes of the Proto-Indo-European laryngeals in Celtic*. Leiden: Brill.

Зайцева, Мария И. & Муллонен, Мария И. (1972). *Словарь вепсского языка*. Ленинград: Наука.

Зайцева, Нина Г. (2008). Отражение вепсско-северно-русских контактов в вепсском лексическом тезаурусе. *Ученые записки петрозаводского государственного университета* 2008/4, 79–83.

Зайцева, Нина Г. (2010). *Новый вепсско-русский словарь*. Петрозаводск: Периодика.

Зализняк, Андрей А. (1985). *От праславянской акцентуации к русской*. Москва: Наука.

Зализняк, Андрей А. (2004). *Древненовгородский диалект* [2nd edition]. Москва: Языки славянской культуры.

Зализняк, Андрей А. (2019). *Древнерусское ударение: Общие сведения и словарь* [2nd edition]. Москва: Языки славянских культур.

Zanina, Oksana G. et al. (2021). Plant food in the diet of the Early Iron Age pastoralists of Altai. *Journal of Archaeological Science: Reports* 35.

Желеховский = Желеховский, Євгений (1886). *Малоруско-нѣмецкий словар*. 2 vols. Львів: Шевченко.

Zeps, Valdis J. (1962). *Latvian and Finnic Linguistic Convergences*. Bloomington: Indiana University.

Zinkevičius, Zigmas (1966). *Lietuvių dialektologija*. Vilnius: Mintis.

Zinkevičius, Zigmas (1980–1981). *Lietuvių kalbos istorinė gramatika* I–II. Vilnius: Mokslas.

Zinkevičius, Zigmas (1984). *Lietuvių kalbos istorija* I. *Lietuvių kalbos kilmė*. Vilnius: Mokslas.

Zinkevičius, Zigmas (1985). Lenkų–jotvingių žodynėlis? *Baltistica* 21/1, 61–82.

Zinkevičius, Zigmas (1987). *Lietuvių kalbos istorija* II. *Iki pirmųjų raštų*. Vilnius: Mokslas.

Живлов, Михаил А. (2008). Этимология прапермского **våž* ‘старый’. In: *Языковые контакты в аспекте истории*. Москва.

Живлов, Михаил А. [= Zhivlov, Mikhail] (2014). Studies in Uralic vocalism III. *Journal of Language Relationship* 12, 113–148.

Живлов, Михаил А. (2015). Неиндоевропейский субстрат в финно-волжских языках [Conference handout]. *X Чтения памяти С.А. Старостина* (27 March 2015).

Живлов, Михаил А. [= Zhivlov, Mikhail] (2023). Reconstruction of Proto-Uralic. In: Daniel Abondolo & Riitta-Liisa Valijärvi (eds.), *The Uralic Languages* [2nd edition]. London–New York: Routledge, 117–175.

Zohary, Daniel, Maria Hopf & Ehud Weiss (2012). *Domestication of Plants in the Old World* [4th edition]. Oxford: University Press.

Жолобов, Олег Ф. (2007). Корпус древнерусских списков Паренесиса Ефрема Сирина. I: РГАДА, Син. 38. *Russian Linguistics* 31, 31–59.

Zubatý, Josef (1892). Zum baltischen ů. *Beiträge zur Kunde der indogermanischen Sprachen* 18, 241–266.

Zubatý, Josef (1894). Slavische Etymologien. *Archiv für slavische Philologie* 16, 385–425.

Zubatý, Josef (1898). [Review of] Recherches sur l’emploi du génitif-accusatif en Vieux-Slave, par A. Meillet. *Archiv für slavische Philologie* 20, 392–397.

Žulys, Vladas (1966). Keleto retų žodžių istorija. *Baltistica* 1, 151–161.

Zupitza, Ernst (1896). *Die germanischen Gutturale*. Berlin: Weidmann.

Zupitza, Ernst (1900). Über doppelkonsonanz im Irischen. *Zeitschrift für vergleichende Sprachforschung* 36, 202–245.

Zupitza, Ernst (1904). Miscellen. *Zeitschrift für vergleichende Sprachforschung* 37, 387–406.

Журавлев, Анатолий Ф. (1980). К этимологии слав. **vorb-* 'птица Passer, воробей'. *Этимология* 1978, 52–58.

Zvelebil, Marek & Dolukhanov, Paul (1991). The Transition to Farming in Eastern and Northern Europe. *Journal of World Prehistory* 5/3, 233–278.

Word Index

Baltic Languages

Lithuanian

Latvian

Note: ⟨uo⟩ is alphabetized under *o*

Prussian

Unmarked = Elbing Vocabulary. Note: ⟨c⟩, ⟨y⟩ and ⟨z⟩ are alphabetized under *k*, *i* and *s*, respectively.

Slavic Languages

Russian

Note: ⟨ѣ⟩ is alphabetized under *е*.

Ukrainian

Belarusian

Lechitic

(unmarked = Polish)

Czech-Slovak

Unmarked = Czech

Church Slavic

Unmarked = Old Church Slavic

English

Unmarked = Old English

Dutch / Frisian

Unmarked = Middle Dutch

Low German

Unmarked = Middle Low German

High German

Unmarked = Old High German

Other Indo-European Languages

Goidelic

Unmarked = Old Irish

Fragmentary Italo-Celtic

Romance

(It. = Italian, Occ. = Occitan, Sp. = Spanish)

Albanian

Greek

Iranian

Uralic Languages

***Finnish* (*Karelian*)**

Veps

Votic

Estonian (*Võro*)

Livonian

Western Sámi

Unmarked = North Sámi

Eastern Sámi

Unmarked = Skolt Sámi

Other Uralic

(Hu. = Hungarian, Kh. Khanty, Ms. = Mansi)

Other Languages

***Turkic* (*Mongolian, Tungusic*)**

The East Baltic languages are well known for their conservative phonology as compared to other Indo-European languages, which has led to a stereotype that the Balts developed in isolation without much contact with other speech communities. This book challenges that view, taking a deep dive into the East Baltic lexicon and peeling away the layers of prehistoric borrowings in the process. As well as significant contact events with known languages, the lexicon also reveals evidence of contact with unattested languages from which previous populations must have shifted.

Anthony Jakob, Ph.D. (2023), Leiden University, is a comparative linguist who specializes in palaeolinguistics and etymology with particular reference to the Baltic languages. He is also an emerging scholar in the field of Uralic linguistics and has presented on topics in this field at international fora.

ISBN 978-90-04-68646-5
ISSN 0926-5856
brill.com/lsie