



Universiteit
Leiden
The Netherlands

Unde venisti? The Prehistory of Italic through its Loanword Lexicon

Wigman, A.M.

Citation

Wigman, A. M. (2023, November 1). *Unde venisti?: The Prehistory of Italic through its Loanword Lexicon*. Retrieved from <https://hdl.handle.net/1887/3655644>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3655644>

Note: To cite this publication please use the final published version (if applicable).

UNDE VENISTI

The Prehistory of Italic through its Loanword Lexicon

Andrew Wigman

Unde vēnistī?

The Prehistory of Italic through its Loanword
Lexicon

Cover image: Photograph of Piedmont taken by Ettore Cauvin.

Copyright © 2023: Andrew Wigman. All rights reserved.

Unde vēnistī?

The Prehistory of Italic through its Loanword
Lexicon

PROEFSCHRIFT

ter verkrijging van

de graad van doctor aan de Universiteit Leiden,

op gezag van rector magnificus prof.dr.ir. H. Bijl,

volgens besluit van het college voor promoties

te verdedigen op woensdag 1 november 2023

klokke 10:00 uur

door

Andrew Michael Wigman

geboren te Voorhees (New Jersey, Verenigde Staten)

in 1991

Promotores: Prof. dr. G.J. Kroonen
Prof. dr. K. Kristiansen

Promotiecommissie: Dr. L.C. van Beek
Dr. B. Nielsen Whitehead (University of Copenhagen)
Prof. dr. M. Peyrot
Prof. dr. M.L. Weiss (Cornell University)

This research was carried out with funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 716732). I would also like to acknowledge the travel grant I received from the Leids Universiteits Fonds (LUF).

*cui dōnō lepidum novum libellum
āridā modo pūmice expolītum?*

To my parents

*namque vōs solēbātis
meās esse aliquid putāre nūgās.*

(Adapted from Catullus 1)

Table of Contents

Acknowledgements.....	xii
List of Figures.....	xiv
List of Tables.....	xiv
Abbreviations.....	xvi
1 Introduction.....	1
1.1 The Motivation.....	1
1.2 Prior Research.....	1
1.2.1 Indo-European Sources.....	4
1.2.1.1 The Balkans.....	4
1.2.1.1.1 Illyrian.....	4
1.2.1.1.2 Thracian and Macedonian.....	7
1.2.1.2 Indo-European Substrates.....	8
1.2.1.2.1 Pelasgian.....	9
1.2.1.2.2 Temematic.....	12
1.2.1.2.3 Indo-European Substrates directly involving Italy.....	13
1.2.1.2.3.1 Ribezzo's Ausonian.....	13
1.2.1.2.3.2 Haas's Frühitalische Element.....	14
1.2.1.2.3.3 Garnier and Sagot.....	15
1.2.1.3 Verdict on Indo-European Substrates.....	17
1.2.2 Non-Indo-European Sources.....	17
1.2.2.1 The Two Lineages.....	18
1.2.2.1.1 The Mediterranean Substrate.....	18
1.2.2.1.2 The European (Germanic) Substrate.....	20
1.2.2.2 Uniting the Lineages.....	21
1.2.2.3 An Excursion on Etruscan.....	25
1.3 The Consequences: Goals of this Dissertation.....	28
1.4 Methodology.....	29
1.4.1 Theoretical Perspective.....	29
1.4.2 How (not) to do it.....	30
1.4.2.1 Reservations about Semantic Category.....	30
1.4.2.2 Doubts about Limited Geographic Distribution.....	32
1.4.2.3 The Requirement of Positive Evidence.....	32
1.4.3 Terminology.....	33
1.4.3.1 Comparanda.....	33
1.4.3.2 Substrate.....	33
1.4.4 A Mild Disclaimer.....	34
1.5 Limitations.....	35
1.5.1 Scope.....	35
1.5.2 Methodological Blind Spots.....	35
2 The Linguistic Data.....	39
2.1 Introduction to the Data.....	39
2.1.1 Structure of the Data.....	39

2.1.2	Structure of the Entries	41
2.2	Non-inherited Origin in Latin Accepted.....	42
2.2.1	Phonotactic Reasons	42
2.2.1.1	Isolated to Latin but with Unrhotalized <i>S</i>	42
2.2.1.2	Isolated to Latin but with an Invalid Root Structure.....	46
2.2.2	Comparanda in Other Branches	49
2.2.2.1	Non-inherited Origin is Probable.....	49
2.2.2.2	Non-inherited Origin is Possible.....	130
2.2.3	Comparanda only in Latin and Romance.....	169
2.3	Origin Unclear.....	176
2.3.1	No Comparanda.....	176
2.3.2	Uncertain Comparanda	185
2.3.3	Conflicting Possibilities.....	213
2.3.3.1	Non-inherited vs. Inherited.....	213
2.3.3.2	Non-inherited vs. Loan from a Known Language	238
2.3.4	Core-Periphery Cases	247
2.3.5	Methodologically Difficult to Delimit Comparanda.....	252
2.4	Non-IE Origin in Latin Rejected.....	259
2.4.1	No Positive Evidence of Borrowing	259
2.4.2	Best Explained as Inherited	275
2.4.3	Loan from a Known Language	283
2.5	Latin Index for the Data Section	288
3	Feature Analysis.....	291
3.1	Introduction to the Feature Analysis	291
3.2	Phonological Alternations	292
3.2.1	Consonants.....	292
3.2.1.1	Alternations between PIE Rows	292
3.2.1.1.1	Non-Velars.....	293
3.2.1.1.1.1	Labials	293
3.2.1.1.1.1.1	Voicing.....	293
3.2.1.1.1.1.2	Aspiration.....	294
3.2.1.1.1.1.3	Voicing and Aspiration	294
3.2.1.1.1.2	Dentals.....	295
3.2.1.1.1.2.1	Voicing.....	295
3.2.1.1.1.2.2	Aspiration.....	295
3.2.1.1.1.2.3	Voicing and Aspiration	296
3.2.1.1.1.3	Interim Conclusion on Labials and Dentals.....	296
3.2.1.1.2	Velars.....	297
3.2.1.1.2.1	Voicing	297
3.2.1.1.2.2	Aspiration	297
3.2.1.1.2.3	Voicing and Aspiration.....	298
3.2.1.1.2.4	Palatalization	298
3.2.1.1.2.5	Labialization.....	299

3.2.1.1.2.6	Interim Conclusion on Velars	300
3.2.1.1.3	Conclusions on the QPIE Plosive Rows.....	300
3.2.1.2	Alternations Beyond the Plosive Rows.....	301
3.2.1.2.1	Labial Plosive ~ Labial Nasal Alternation	301
3.2.1.2.2	Labial Plosive ~ Labial Approximant Alternation	303
3.2.1.2.3	L ~ R Alternations.....	304
3.2.1.2.4	N ~ M Alternation.....	305
3.2.1.2.5	L ~ D Alternation.....	305
3.2.1.2.6	S ~ D Alternation	306
3.2.1.2.7	D ~ K ~ Ø Alternation.....	306
3.2.1.2.8	Further Irregular Involvement of a Sibilant.....	307
3.2.1.2.8.1	S mobile.....	307
3.2.1.2.8.2	S Insertion.....	308
3.2.1.2.8.3	SK Metathesis.....	309
3.2.1.2.9	Gemination.....	310
3.2.2	Vowels.....	312
3.2.2.1	Clearly non-IE Alternations.....	313
3.2.2.1.1	E ~ I	313
3.2.2.1.2	I ~ U.....	314
3.2.2.1.3	E ~ U.....	315
3.2.2.1.4	O ~ U.....	315
3.2.2.2	<i>a</i> -Vocalism	316
3.2.2.2.1	Reconstructed <i>a</i> -Vocalism	316
3.2.2.2.2	Alternations Involving A.....	317
3.2.2.2.2.1	A ~ \bar{A}	317
3.2.2.2.2.2	A ~ E	317
3.2.2.2.2.3	A ~ O.....	318
3.2.2.2.2.4	A ~ AU	319
3.2.2.2.2.5	A ~ U.....	319
3.2.2.2.2.6	A ~ AI.....	319
3.2.2.2.3	Wider Variation	320
3.2.2.2.4	Ablaut Phenomena.....	321
3.2.2.4.1	Ablaut Unparalleled in IE	321
3.2.2.4.2	Ablaut Difficult to Motivate from an IE Perspective	322
3.2.2.4.3	Vocalic Alternations That Can Occur in Ablaut Paradigms.....	322
3.2.3	Phonological Conclusions.....	323
3.3	Morphological Alternations	324
3.3.1	Pre-Greek Suffixes.....	324
3.3.1.1	Latin <i>-essus</i>	325
3.3.1.2	Latin <i>-undo</i>	325
3.3.1.3	Latin <i>*-ara</i>	326
3.3.1.4	Conclusion on Pre-Greek Suffixes in Latin.....	326

3.3.2	The <i>a</i> -Prefix	326
3.3.3	The Velar Suffix	328
3.3.4	The <i>n</i> -Suffix	330
3.3.5	Reduplication	334
3.3.6	Morphological Conclusions	336
4	Distribution Analysis	337
4.1	Introduction to Distribution Analysis	337
4.2	Partial Stratification Based on Phonology: The Most Recent Borrowings	340
4.2.1	Latin-Greek Isoglosses with Recurring Irregular Alternations	341
4.2.1.1	Voicing and Devoicing	341
4.2.1.2	Aspiration Alternations	341
4.2.2	Wider Implications: A Mediterranean Substrate	342
4.2.2.1	Other Words with a Mediterranean Distribution	342
4.2.2.2	Balkan Connections	343
4.2.2.3	Indo-Iranian Connections	345
4.2.2.4	Recurring Features	346
4.2.2.4.1	L ~ R Alternation	346
4.2.2.4.2	E ~ I Alternation	347
4.2.2.4.3	I ~ U Alternation	347
4.2.2.5	A Definitively Mediterranean Substrate	348
4.2.2.6	Mediterranean Languages with a Wider Distribution	350
4.3	Further Stratification Using Distribution	351
4.3.1	Potentially Recent Borrowings	351
4.3.1.1	Words Exclusive to Latin and Romance	351
4.3.1.2	Gemination	352
4.3.2	Earlier Strata	352
4.3.2.1	The Oldest Loans in Italy	352
4.3.2.2	Early Contact or Widespread Substrate?	354
4.3.2.3	Intermediate Contact Situations	357
4.3.2.3.1	Stratum Excluding Greek	357
4.3.2.3.2	Italo-Celtic Isoglosses and the Italo-Celtic Subnode	358
4.3.2.3.3	Italo-Germanic Group	360
4.3.2.3.4	Germanic Loans in General	361
4.3.2.3.5	Other Groups	362
4.4	A More Inclusive Visualization	363
4.5	Summary of Stratigraphy	365
5	Semantic Analysis	367
6	Population Genetics of Italy	371
6.1	Genetics Introduction	371
6.2	European Genetics from the Origins of Agriculture to the Homeland Debate 371	
6.2.1	Understanding the Origin and Spread of Agriculture	371
6.2.2	The Indo-European Homeland Debate	373

6.2.3	Refining and Overturning the Understandings.....	374
6.2.3.1	Building Genetic Databases.....	374
6.2.3.2	A Paradigm Shift Waiting to Happen: the Sequencing of Ancient DNA	376
6.2.3.3	The 2015 Paradigm Shift.....	379
6.2.4	Summary.....	383
6.3	The Italian Peninsula.....	383
6.3.1	Studies on Modern Populations	383
6.3.1.1	The Earliest Studies.....	383
6.3.1.2	More Modern Methods.....	384
6.3.2	Ancient DNA.....	386
6.3.3	Other Questions	388
6.3.3.1	Greek Colonization.....	388
6.3.3.2	Etruscans.....	389
6.4	Conclusions.....	390
7	Archaeological Theories on the Italicization of Italy.....	391
7.1	Single Origin Theories.....	392
7.1.1	The Copper Age Cultures	392
7.1.2	Side Note: Bell Beaker and Polada Cultures.....	392
7.1.3	Terramare Culture.....	393
7.1.4	Apennine Culture.....	395
7.1.5	Urnfield Horizon.....	395
7.2	Multiple Origin Theories and the Question of Proto-Italic Unity.....	397
8	Conclusion.....	403
8.1	Summary.....	403
8.2	Discussion: Triangulating Italic Prehistory.....	404
	Bibliography	413
	Nederlandse Samenvatting.....	479
	Curriculum Vitae	481

Acknowledgements

The trajectory that has resulted in this dissertation began long ago, with the language teachers who instilled in me the first sparks of linguistic curiosity (Tina Bechtel, Michael Korom, Cynthia Knisley, and Gregory Toates). It was then under the aegis of Philip Baldi at Penn State that I discovered Latin historical grammar and the field of historical comparative linguistics. My curiosity became insatiable as I learned that, from the balcony of Latin, one could look down upon the floorboards of Proto-Indo-European. My teachers at the Ludwig-Maximilians-Universität, Olav Hackstein, Peter-Arnold Mumm, Dieter Gunkel, and Filip De Decker, taught me the nuances of the architecture of this reconstruction. When I saw the glints of something older, just barely visible from between the floorboards, it was with Guus Kroonen and his project *The Linguistic Roots of Europe's Agricultural Transition* that I was able to begin excavating (along with the expertise of one who has literally excavated, my second supervisor Kristian Kristiansen) to see what sort of substrate the edifice was built upon. To all these teachers I am grateful.

I am grateful to several other scholars from whom I have learned much along the way. To Sasha Lubotsky, Alwin Kloekhorst, Tijmen Pronk, Lucien van Beek, Michaël Peyrot, Maarten Kossmann, David Stifter, and Peter Schrijver, I am thankful for the conversations and email exchanges. I will treasure the mentorship I received from Michael Weiss and Alan Nussbaum at Cornell University along with the friends I made there.

My research would never have proceeded without the insightful discussions on material, methodology, and motivation with my team- and office-mates Yvonne van Amerongen, Paulus van Sluis, Anthony Jakob, Rasmus Thorsø, Axel Palmer, and Cid Swanenvleugel.

Wonderful friends and colleagues have played a great role in making the hard work in Leiden worth it. To the Lunch Bunch (especially Niels Schoubben, Jesse Wichers Schreur, Ahmed Sosal, and Sophia Nauta), thank you for putting up with all the Germanic etymologies. To Stefan Norbruis, Annika Kramer, Rasmus Puggaard-Rode, Chams Bernard, Olga Nozdracheva, Ami Okabe, Ménghuī Shǐ, Niko Kontovas, Astrid van Alem, Laura Smorenburg, Jiāng Wú, Priscilla Lam, Aljoša Šorgo, Davide Procaccino, Eleonora Settembrino, Tíngtíng Zhèng, Meike de Boer, Aliza Glasbergen-Plas, Jiāqí Wáng, Shàoyǔ Wáng, and others, thank you for the dinners, picnics, and algemene gezelligheid. Abel Warries en Femke Montagne (translator extraordinaire), sommige Nederlandse vogels zijn inderdaad mooier dan Amerikaanse. En stor tak til Simon Poulsen og Tobias Søborg. Louise Skydsbjerg Friis (I, II, and III) *nī kārweñe*, thank you for always being there. Kate Bellamy, thank you for the tri-Georgian feast and the reyt good career advice. Petra Couvée, thank you for always believing in me. Yàn Wáng, thank you for the walks and talks. Sarah von Grebmer Zu Wolfsturn, your kindness and positivity has been an inspiration. Lis Kerr and Xander Vertegaal, you have been iridium quality friends.

I likewise could not have done without the support I have felt from friends in more far off places as well. Karla Gutberlet has been a constant source of strength these many years. Paweł Ślęczka, Daniel Méndez Aranda, and Ali Karity have helped keep me sane through the hardest parts of being away from home. Kittipong Vongagsorn, *bahu manye mitratvam avayoh*. Thank you Martin Riesenberg for the insightful discussions about steppe archaeology and the methodological difficulties of comparing different lines of evidence. Thank you Hannes Puchta, my Berlin-based book sleuth. Tori Parker, Hannah Hoganson, Neil Hong, and Amritha Mallikarjun, my oldest friends, come what may, thank you for keeping Downingtown free of werewolves. To Melissa DiJulio and Deirdre Graham, I loved our weekly lockdown check-ins, and I am still ready to buy that farm in Ireland with you.

I am also ever grateful to the support of my family. This book is dedicated to my parents, Dr. Larry and Mrs. Lisa Wigman. I thank my brothers, Jonathan and Benjamin, my cousins, aunts, uncles, and grandmother, who checked in with me along the way. At his doctoral defense, Dr. Benjamin Wigman said he always thought I was smarter than he, but even a chemist with a PhD can be wrong sometimes.

Grātia [est], in quā amīcitiārum et officiōrum alterius memoria et remūnerandī voluntās continētur (Cicero, *de Inventione* 2.53.161).

List of Figures

Figure 1.1 The linguistic diversity of the Italian peninsula represented by sites of inscriptions.	2
Figure 4.1 Latin words of loanword origin distributed by existence of comparanda in Celtic, Germanic, and Greek	338
Figure 4.2 The transmission of Lat. <i>cupressus</i> and its comparanda from a source form	339
Figure 4.3 Overlapping irregular alternations between words with a Mediterranean distribution	348
Figure 4.4 Range of <i>Rhamnus alaternus</i>	350
Figure 4.5 Range of <i>Ficus carica</i>	350
Figure 4.6 Range of <i>Buxus sempervirens</i>	350
Figure 4.7 Range of <i>Cupressus sempervirens</i>	350
Figure 4.8 Separate contact situations suggested by the distribution of non-inherited lexemes	357
Figure 4.9 One interpretation of contact incorporating Italo-Celtic considerations	359
Figure 4.10 An alternative scenario with an early Italo-Celtic split and later close contact	360
Figure 4.11 PCA of loose analysis	364
Figure 4.12 PCA of strict analysis	365

List of Tables

Table 3.1 Alternations between <i>*b</i> and <i>*p</i>	293
Table 3.2 Alternations between <i>*b^h</i> and <i>*b</i>	294
Table 3.3 Alternations between <i>*b^h</i> and <i>*p</i>	294
Table 3.4 Alternations between <i>*b^h</i> or <i>*b</i> and <i>*p</i>	294
Table 3.5 Alternations between <i>*b^h</i> , <i>*b</i> , and <i>*p</i>	295
Table 3.6 Alternations between <i>*d</i> and <i>*t</i>	295
Table 3.7 Alternations between <i>*d^h</i> and <i>*d</i>	295
Table 3.8 Alternations between <i>*d^h</i> and <i>*t</i>	296
Table 3.9 Alternations between <i>*g</i> and <i>*k</i>	297
Table 3.10 Alternations between <i>*g^h</i> and <i>*k</i>	298
Table 3.11 Alternations between <i>*g^h</i> or <i>*g</i> and <i>*k</i>	298
Table 3.12 Alternations between <i>*g^h</i> , <i>*g</i> , and <i>*k</i>	298
Table 3.13 Alternation between <i>*b^(h)</i> and <i>*m</i>	301
Table 3.14 Alternation between <i>*p</i> , <i>*b</i> (<i>/*b^h</i>), and <i>*m</i>	301
Table 3.15 Alternations between <i>*b^(h)</i> and <i>*mb^(h)</i>	302
Table 3.16 Alternation between <i>*b^h</i> , <i>*b</i> , and <i>*u</i>	303
Table 3.17 Alternation between <i>*b</i> and <i>*u</i>	303
Table 3.18 Group A alternations between <i>*l</i> and <i>*r</i>	304
Table 3.19 Group B alternations between <i>*l</i> and <i>*r</i>	305

Table 3.20 Group C alternation between <i>*l</i> and <i>*r</i>	305
Table 3.21 Alternations in the order of sibilant and velar in clusters.....	309
Table 3.22 Alternations in gemination.....	312
Table 3.23 Alternations between <i>*e</i> and <i>*i</i>	313
Table 3.24 Alternation between <i>*i</i> and <i>*u</i>	314
Table 3.25 Alternations between <i>*ī</i> and <i>*ū</i>	314
Table 3.26 Alternation between <i>*e</i> and <i>*u</i>	315
Table 3.27 Alternations between <i>*o</i> and <i>*u</i>	315
Table 3.28 Alternation between <i>ō</i> and <i>ū</i>	315
Table 3.29 Alternation between <i>ō</i> and <i>u</i>	315
Table 3.30 Alternations between <i>*e</i> and <i>*o</i>	322
Table 3.31 Alternations between <i>*e</i> and <i>*ϕ</i>	323
Table 3.32 Alternating presence of an n-suffix.....	332
Table 4.1 Consonantism correspondences between Germanic and other branches	361

Abbreviations

General

ca.	<i>circa</i> , around/approximately
cf.	<i>confer</i> , compare
dial.	dialectal
dim.	diminutive
e.g.	<i>exempli gratia</i> , for example
esp.	especially
et al.	<i>et alii</i> , and others
et alib.	<i>et alibi</i> , and elsewhere
etc.	<i>et cetera</i> , and so on
fn.	footnote
fthc.	forthcoming
id.	<i>idem</i> , the same
p.c.	personal communication
pg., pp.	page, pages
s.v.	<i>sub verbo</i> , under the entry
var(s).	variant(s)
with lit.	with literature

Dates

BCE	Before the Common Era (BC)
c.	century
CE	Common Era (AD)

Languages

Aeol.	Aeolic (Greek)
Akk.	Akkadian
Alb.	Albanian
Att-Ion.	Attic-Ionic (Greek)
Arab.	Arabic
Aram.	Aramaic
Arcad.	Arcadian (Greek)
Arm.	Armenian
Av.	Avestan
Boeot.	Boeotian (Aeolic Greek)
Bret.	Breton
Bulg.	Bulgarian
Cat.	Catalan
Cl. Arab.	Classical Arabic
CLuw.	Cuneiform Luwian
Copt.	Coptic
Dan.	Danish

Du.	Dutch
Egypt.	Egyptian
Engl.	English
Etr.	Etruscan
Fal.	Faliscan
Far.	Faroese
Fr.	French
Georg.	Georgian
Ger.	German
Gk.	Ancient Greek
Go.	Gothic
Hebr.	Hebrew
Hitt.	Hittite
HLuw.	Hieroglyphic Luwian
Hsch.	Hesychian (gloss in Greek)
Hurr.	Hurrian
Icel.	Icelandic
IE	Indo-European
It.	Italian
Kartv.	Kartvelian
Khot.	Khotanese
Lat.	Latin
Latv.	Latvian
LCo.	Late Cornish
Lith.	Lithuanian
Lyd.	Lydian
Mac.	Macedonian
MBret.	Middle Breton
MBulg.	Middle Bulgarian
MDu.	Middle Dutch
ME	Middle English
MHG	Middle High German
MÍr.	Middle Irish
MoDu.	Modern Dutch
MoGk.	Modern Greek
MoP	Modern Persian
MP	Middle Persian
MW	Middle Welsh
Myc.	Mycenaean (Greek)
Nw.	Norwegian
OBret.	Old Breton
ODan.	Old Danish
OE	Old English
OGeorg.	Old Georgian
OHG	Old High German

OLG	Old Low German	PSlav.	Proto-Slavic
ON	Old Norse	PU	Proto-Uralic
OP	Old Persian	PVasc.	Proto-Vasconic
OPol.	Old Polish	Pt.	Portuguese
OPr.	Old Prussian	Rom.	Romanian
OProv.	Old Provençal	Ru.	Russian
OPt.	Old Portuguese	RuCS	Russian Church Slavonic
ORu.	Old Russian	Sard.	Sardinian
Osc.	Oscan	SCr.	Serbo-Croatian
OSum.	Old Sumerian	Serb.	Serbian
OW	Old Welsh	SGael.	Scottish Gaelic
PAlb.	Proto-Albanian	Slk.	Slovak
PArm.	Proto-Armenian	Slov.	Slovene
PBalt.	Proto-Baltic	Sogd.	Sogdian
PBerb.	Proto-Berber	Sp.	Spanish
PBSl.	Proto-Balto-Slavic	SPic.	South Picene
PCelt.	Proto-Celtic	Sum.	Sumerian
PEBalt.	Proto-East Baltic	Sw.	Swedish
PGk.	Proto-Greek	Syr.	Syriac
PGm.	Proto-Germanic	Thess.	Thessalian (Aeolic Greek)
Phoen.	Phoenician	Toch. A	Tocharian A
PIE	Proto-Indo-European	Toch. B	Tocharian B
PIIr.	Proto-Indo-Iranian	Turk.	Turkish, Turkic
PIr.	Proto-Iranian	U	Umbrian
PItal.	Proto-Italic	Ugr.	Ugaritic
Pol.	Polish	Ved.	Vedic Sanskrit
PRom.	Proto-Romance	W	Welsh
Prov.	Provençal	YAv.	Young Avestan
PSem.	Proto-Semitic		

Grammar

abl.	ablative	masc.	masculine
acc.	accusative	neut.	neuter
adj.	adjective	nom.	nominative
dat.	dative	obl.	oblique
fem.	feminine	pass.	passive
fut.	future	pl.	plural
gen.	genitive	PPP	perfect passive participle
indecl.	indeclinable	pres.	present
inf.	infinitive	sg.	singular
loc.	locative	subj.	subjunctive

Archaeo-Genetics

ANE	Ancient North Eurasian (ancestry component)
calBCE	Calibrated radiocarbon date
CHG	Caucasus Hunter-Gatherer (ancestry component)
EEF	Early European Farmer (ancestry component)
EHG	Eastern Hunter-Gatherer (ancestry component)
LBK	Linearbandkeramik Culture
mtDNA	Mitochondrial DNA
WHG	Western Hunter-Gatherer (ancestry component)

Symbols

?	(appurtenance) uncertain
<	is from
>	develops to
>>	is borrowed as
~	alternates with
§	heading number (of this thesis)
<>	graphical element
//	phonemic transcription

Phonological Cover Symbols (in text; not in header reconstructions)

C	consonant
C _i	identical consonant
D	voiced stop
D ^h	voiced aspirated stop
H	laryngeal
R	resonant
T	unvoiced stop
V	vowel

1 Introduction

1.1 The Motivation

The Italian peninsula, reaching far into the Mediterranean Sea, served by virtue of its geography as an important crossroads of ancient cultures. Economic intercourse occurred across millennia between groups speaking many different languages that we know of (Figure 1.1), and likely countless others that have been lost to time. Many of these interactions left a mark on Latin. Here, the contact was intense and transformative. The colonization of Magna Graecia in the 8th century BCE resulted in ca. 5000 Greek words appearing in Latin amongst 44,000 core lexemes (Seidl 2003: 519) and Greek also affected written Latin syntax (Clackson & Horrocks 2011: 191-7, Weiss 2020: 509). Latin and Etruscan seem to have been in close enough contact that they both underwent the same areal shift to initial accent.¹ As will be discussed in detail in §5, the first traces of populations bearing steppe-derived ancestry appear in Northern Italy ca. 2000 BCE (Saupe et al. 2021) after the dispersal of Yamnaya populations ca. 3000 BCE (Haak et al. 2015), and they appeared on a peninsula that had been inhabited by Neolithic farmers since ca. 6000 BCE (Malone 2003: 242). With such intense contacts in the attested record, we must be able to see traces from *unattested* contacts as well.

Much research has been done on prehistoric substrates, but differences in methodology have produced a corpus of literature that on the one hand does not always agree on the exact nature and impact of these substrates and on the other hand deserves a fresh and more modern treatment.

1.2 Prior Research

The research on linguistic substrates is at its core a question about language contact. The search to understand potential contact between Latin and poorly attested or unattested languages has led in several directions. In general, the feasibility of recognizing different linguistic strata only came about at the end of the nineteenth century, after nuanced views of language contact began to develop. Thomason and Kaufman (1988: 1) capture the volatility of the early field: Max Müller in 1871 claimed “Es gibt keine Mischsprachen” against Hugo Schuchardt’s 1884 claim of the exact opposite, “Es gibt keine völlige ungemischte Sprache”.

¹ This is inferred for Etruscan and Italic on the basis of weakening and syncope in non-initial syllables (cf. Wallace 2008: 37-9, Weiss 2020: 118-19, 527). Etruscan spellings attest to what looks like vowel weakening by the end of the 7th century. Latin shows vowel weakening and syncope by ca. 500 BCE. Then Etruscan begins to show syncope in medial syllables beginning ca. 470 BCE. The Sabellic languages also developed initial accent. The order of the changes makes it difficult to determine if one group or the other initiated the change. Instead it looks areal.

2 *Unde vēnistī?* The Prehistory of Italic through its Loanword Lexicon

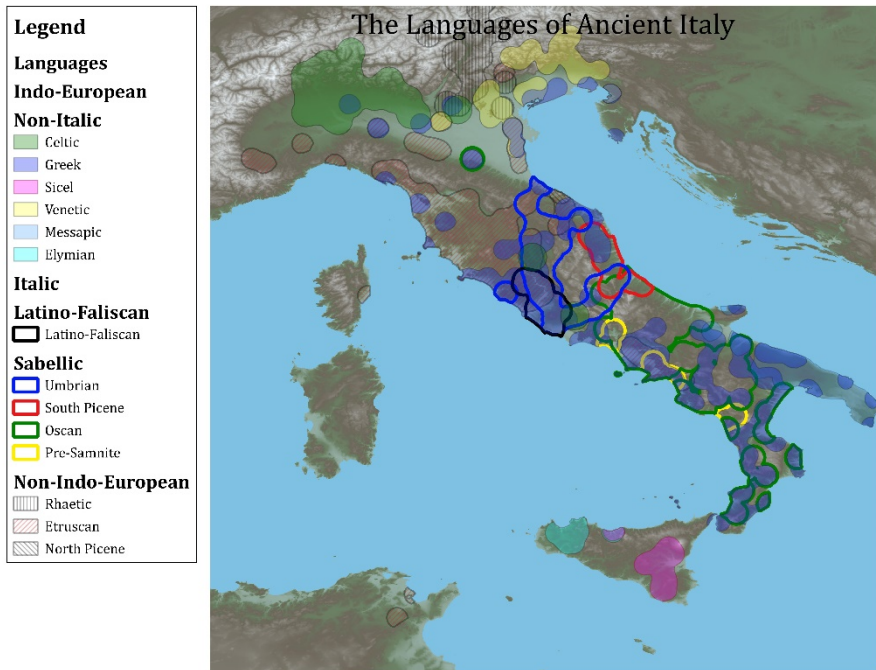


Figure 1.1 The linguistic diversity of the Italian peninsula represented by sites of inscriptions

Data taken mainly from a spreadsheet compiled by Katherine McDonald (<https://katherinemcdonald.net/research/maps/> with refs.), further supplemented as follows: Elymian from Marchesini (2012: 97); Faliscan from Bakkuum (2009); Latin from EAGLE (Electronic Archive of Greek and Latin Epigraphy, <http://www.edr-edr.it/>) with a search for inscriptions in Latin dating to before 400 BCE; Lepontic from the *Lexicon Leponticum* (https://lexlep.univie.ac.at/wiki/Main_Page); Rhaetic from the *Thesaurus Inscriptionum Raeticarum* (https://tir.univie.ac.at/wiki/Main_Page); Sicel from *Inscriptions of Sicily* (<http://sicily.classics.ox.ac.uk/>). Not shown are Punic/Phoenician and the Greek inscriptions of Sicily. The simplified grouping of the Sabellic languages into Oscan, Umbrian, South Picene, and Pre-Samnite follows Weiss (2020: 15-16). Note that Pre-Samnite pre-dates Oscan where it was attested.

The theory of linguistic substrates began to develop in the work of Romanists (Craddock 1969: 18-22), especially Graziadio Isaia Ascoli,² who were dealing with a wide variety of material attested in the numerous Romance languages and who seemed to understand

² Like his 1882 'Lettere glottologiche: prima lettera' in *Rivista di Filologica Classica* 10: 1-71, where he argued that Gallo-Romance sound shifts were caused by Celtic speakers learning Latin. Craddock (1969: 19-22) notes several scholars preceding Ascoli who wrote about similar ideas.

that the spread of Latin must not have involved simple language replacement. It began, at least with Ascoli, as a criticism of the Neogrammarian model. This discontentment with the stringency of the Neogrammarian model and the exclusive focus of Indo-Europeanists on discovering the regularities of the daughter languages continued to influence the work of Italian scholars into the 1940s. Bertoldi (1939a: 5-19) contrasted the approaches/attitudes of Indo-Europeanists like Walde with those of Romanists like his teacher Jules Gilliéron and later (Bertoldi 1942: 1-8) highly praised Schuchardt for taking into account the shifting linguistic contacts that must underlie the complexity of Romance development from Latin:

Schuchardt was therefore one of the first to rise up against the purely evolutionary conception of a language whose changes should be governed by rigid laws, who adheres instead to the principle that linguistic innovation in every system is the effect of contact of the individual with a more or less homogeneous collectivity of speakers in a varied game of expressive and receptive possibilities (Bertoldi 1942: 4-5).³

He saw that progress in the field relied on seeing linguistic change “no longer as the reflection of a pure and simple grammatical development, but as the result of historical events and cultural currents destined to accelerate or delay the rhythm of the consequent linguistic innovations” (Bertoldi 1942: 5).⁴ Alessio (1944a: 94) shared the frustration when describing research on the Mediterranean substrate (see §1.2.2.1.1), noting that the major etymological dictionaries “do not know how to completely free themselves from the shackles of traditional etymology”⁵ and calling for the languages of the Mediterranean to be thoroughly surveyed by experts in the field, not Indo-Europeanists “who have a very different sensitivity in dealing with linguistic problems.”⁶

As will be mentioned further in §1.4, words borrowed from the unknown pre-Latin languages of Europe show up in Latin as traditionally unetymologizable. They are either isolated without any cognates or have lookalikes in other languages that defy understood sound laws. Some words like this can be explained through inter-dialectal borrowing or internal processes like analogy. However, given the disgruntlement towards the field of Indo-European studies mentioned above, it is worth noting that the state of the field today takes both sides’ interests into account. It is only the adherence to the Neogrammarian model of the inviolability of sound law that allows for the identification of instances where regularity breaks down. It is the historical comparative method that

³ “Fra i primi ad insorgere contro la concezione puramente evolucionistica d’una lingua i cui mutamenti dovrebbero essere retti da leggi rigide fu, dunque, lo Schuchardt che s’attiene invece al principio che l’innovazione linguistica in ogni sistema è l’effetto del contatto dell’individuo con una collettività di parlanti più o meno omogenea in un gioco vario di possibilità espressive e ricettive.”

⁴ “Non più come il riflesso di un puro e semplice sviluppo grammaticale, ma come il risultato delle vicende storiche e delle correnti culturali destinate ad accelerare o a ritardare il ritmo delle conseguenti innovazioni linguistiche.”

⁵ “Non sanno liberarsi completamente dalle pastoie dell’etimologia tradizionale.”

⁶ “Che hanno una sensibilità molto differente nel trattare i problemi linguistici.”

4 *Unde vēnistī?* The Prehistory of Italic through its Loanword Lexicon

allows us to determine if a word is the expected reflex of an inherited root, if it is isolated, or if it shows irregular correspondences to would-be cognates.

In the history of the field, there have been many kinds of attempts to explain the origins of Latin etyma once it has been determined that they are not inherited. Numerous examples of borrowing from known, well-understood languages are still accepted today.⁷ But countless lexemes cannot be explained in this way. Given that the linguistic diversity of the Italian peninsula and the Mediterranean Basin has been documented since antiquity, various other contact scenarios have been proposed and used as explanations. What follows is a summary of several kinds of proposals: borrowings from poorly understood Indo-European languages, borrowings from lost Indo-European languages, and borrowings from non-Indo-European languages.

1.2.1 Indo-European Sources

1.2.1.1 The Balkans

As will be shown in detail, there are several cases of Latin-Greek correspondences with very slight aberration in consonantism. Two salient examples include Lat. *ballaena* ~ Gk. φάλλαίνα ‘whale’ and Lat. *fascinus* ‘evil spirit; charm’ ~ Gk. βάσκανος ‘bewitcher, slanderer’. In both cases, especially the latter where one can imagine descent from PIE **b^heh₂-* ‘to speak’, one reflex can be reconstructed as descending regularly from PIE **b^h* while the other represents **b*. This has led to the suspicion that the irregular one of the pair represents a loan from a known language in which PIE *mediae apsiratae* yielded *mediae*. There are three potential culprits, namely the poorly attested Balkan languages Illyrian, Thracian, and Macedonian (cf. Schwyzler I: 65-71, Biville I: 180).

1.2.1.1.1 *Illyrian*

Archaeological and onomastic material along with indications in the writings of ancient authors make it possible that Messapic was directly related to Illyrian (cf. Hamp 1957: 74), but there exists no inscriptional material to help confirm this (cf. Fortson 2010: 464-5, 467-8; Matzinger 2005: 29, de Simone 2018: 1842-3, Matzinger 2019: 20). In fact, Albanian is spoken in the geographic region where Illyrian is purported to have been spoken, but again, the absence of attested Illyrian material makes it very difficult to investigate any connection. On the other hand, some remarkable lexical correspondences between Albanian and Messapic, especially given the poor attestation of the latter and the nearly 1500 years separating their first attestations, make it quite likely that they are closely related (e.g. Matzinger 2005, followed lukewarmly by Hyllested & Joseph 2022: 240, 241). Thus, regardless of an existing relationship to the elusive Illyrian, Messapic seems to represent an originally Balkan language that came to be spoken on the Italian

⁷ From Greek (cf. Weise 1882, much more recently Biville I and II), Celtic (e.g. Schmidt 1966), Germanic (precious few: Green 1998: 182-200). On the possibility of inter-dialectal borrowing, i.e. Sabellicisms or regional variants appearing in attested urban Classic Latin see e.g. Rix (2005) and a careful treatment of the chronological and regional variation of Latin in Adams (2007).

peninsula.

Krahe (1955: 114-7) gave a summary of Latin and Greek words considered to be loans from Illyrian, often via Messapic, finding that they fell into two main categories—small sea-going vessels and horse-related words—along with some others. Even if we re-script Illyrian as “an ancestor of Albanian”, many of the proposed loans have little to support them. Lat. *gandeia* ‘an African vehicle’ is supposed to be from Messapic on comparison with Venetian *gondola* ‘type of boat’⁸ and the *-eia* suffix found in *sabaia* ‘beer’, given by glossators as Illyrian. Lat. *hōreia* ‘small fishing boat, pointed at the front’ would be Illyrian/Messapic on the evidence of the *-eia* suffix alone. Lat. *caballus* ~ Gk. καβάλλης ‘workhorse’ would have entered from Illyrian because of several personal names that all have the element *cabal(l)-*. But there is no reason to source this word specifically from Illyrian given its attestation in several other places.

The Messapic reflexes of PIE **b^h* and **d^h* really do seem to be *b* and *d* (cf. Messapic *berain* ‘may they bring’ < **b^her-o-ih₁-nt*, Matzinger 2019: 64, and further *hipa-des* ‘he/she dedicated’ < **supo-d^heh₁-s-t*, de Simone 2018: 1844). If this goes for Illyrian too, then a **D^h* ~ **D* discrepancy between languages might attest to one of them having borrowed an Illyrian reflex in place of its own inherited reflex. Several cases of this are given by Krahe (1955: 114-7) to explain irregular alternations between Latin and Greek, but there are methodological problems. According to him:

- Lat. *ballaena* ‘whale’ would be either from Gk. φάλλαινα ‘whale’ via Illyrian or both would be from Illyrian. But if the Illyrian reflex of **b^h* was *b*, then Gk. φάλλαινα cannot be from Illyrian. The Latin diphthong (unweakened in a non-initial syllable) attests to a late loan from Greek. In that case, we are not looking for a language whose reflex of PIE **b^h* was *b* but rather one that borrowed Gk. φ as *b*; there is no indication that Illyrian or Messapic did this.
- Lat. *dēda* ‘wet nurse’, if the vowel length is correct, would be from the Illyrian reflex of **d^heh₁d^hā-*: cf. Gk. τήθη ‘grandmother’ (cf. also Krahe 1937) and PBSI. **deʔd-* ‘grandfather/uncle’ (cf. Derksen 2007: 101, EDG 1477). But it is a *Lallwort* (cf. TLL s.v. *dida*) and thus its evidentiary value is dubious.
- Lat. *galaia* and Gk. γαλαία ‘racing vessel’ would be Illyrian along with Lat. *golaia* ‘sea turtle’, cognate with Gk. χέλως ‘tortoise’ and PSlav. **žely-* < **g^hel(H)-uH-*. Note here also the *-aia* suffix. But the connection of the ship and turtle words is not secure. Despite the turtle lexeme being reconstructible to a common proto-form in two languages of the Balkans, *golaia* appears in Latin recently, in glosses and the Latin translation of Dioscorides’ *de Materia Medica*. Thus a loan from Illyrian or even Messapic seems far from the only option.

⁸ This was used as evidence because Venetic had been thought to belong to the Messapic-Illyrian branch. It does not (already Beeler 1949: 48-57).

- Lat. *brīsa* ‘remains of pressed grapes’ would be from Illyrian **brīsa*, itself from Thracian **brūtja*, the source of Gk. βρῦτος ‘barley beer’ and βρῦτα ‘refuse of olives or grapes’ < PIE **b^hru-* (cf. Lat. *dē-frutum* ‘boiled down must’). The assumption of Thracian origin is not bulletproof, but otherwise this is the only example where a Balkan language seems to have been involved, albeit not Illyrian. The pre-form of Alb. *bërsí* ‘pomace, lees, dregs’, PALb. **brīšā*, is identical to Lat. *brīsa*. Since Albanian produces *s* < **tj*, a pre-form of Albanian may have been involved in the transmission of this lexeme into Latin. Krahe assumed that Illyrian was responsible for the change of *ū* to *ī*, but without any further evidence of this, it is *ad hoc*. All we can say is that βρῦτος reached Albanian (where a direct loan should have given ***brys-*, Demiraj 1997: 98) indirectly.

Kretschmer (1896: 248-9, fn. 4) had additionally suggested that a North Balkan treatment of **b^heh₂-* ‘to speak’ might be responsible for Gk. βάσκανος ‘who bewitches; sorcerer, slanderer’ beside Lat. *fascinus* ‘evil spirit, spell’ (cf. also Devoto 1943: 364). But given the semantic distance from the root in question and the lack of evidence of other IE attestations of this formal and semantic derivation, it does not seem fully warranted to achieve regularity by forcing **b^haskano-* through Illyrian on its way to Greek.

There are two cases of potential Illyrian loans given by Krahe where his reason seems to have been the aberrant *a*-vocalism in Latin. Lat. *mannus* ‘small horse’ (cf. also Brūch 1922: 246-7) would be from Illyrian **manda-* attested in a Messapic name of Jupiter *Menzana*, supposed to be from PIE **mez₂-* ‘to feed’ (cf. Alb. *mënd* ‘to suckle’ but also *mëz* ‘foal’). Orel (1998: 265) reconstructs for *mëz* PALb. **mandja-* and takes It. *manzo* ‘ox’ from its Messapic cognate. It. *manzo* requires a pre-form like **mandius* and could indeed theoretically be related to the Albanian form, but Lat. *mannus* would not regularly have developed from a form with **-nd-*.⁹ Lat. *parō*, Gk. παρόν ‘small boat’ would be from Illyrian, cognate to OHG *farm* ‘fast ship’, Ru. *poróm* ‘ferry’, etc. < **por-mo-*. But regardless of the source of the Greek form, the Latin is most easily explained as a loan from it. Illyrian need not have produced it.

Krahe additionally gives the case of Lat. *g* for Gk. κ in Lat. *grabātus*, Gk. κράβ(β)ατος ‘bed’. They would be from Illyrian/Macedonian **graba-* ‘oak’, also found behind Gk. γράβιον ‘torch, oakwood’ and the Umbrian epithet of Jupiter *Grabovius*. EDG (284) notes that the forms are also compared to e.g. Ru. *grab* ‘hornbeam’. The Slavic forms are the only indication of an originally Balkan source. Otherwise the forms could be from

⁹ The sequence *-nd-* is preserved in Latin, thus the assumption is in WH (II: 30) and EM (384) that it is a dialectal form (cf. *dispennite* for *dispandite* in Plautus’ *Miles Gloriosus*). Nor does Sabellic origin provide a good explanation. Weiss (2020: 188) notes that even there, the evidence for a development **nd > nn* is poor. Whether the forms with **-nd-* are Messapic/Illyrian to begin with is unclear, but such an origin does not straightforwardly explain how Latin ended up with *mannus* (which, WH [II: 30] mention, the grammarian Consentius ascribes Gallic origin).

anywhere. Perhaps Illyrian transmitted the χ of Gk. ἔγχελυς ‘eel’ to Latin as the gloss *enocilis* = *anguilla*, but this relies only on the Illyrian personal name *Enoclia* as evidence.

In the end, the evidence pointing to Illyrian (or Messapic/Proto-Albanian) being responsible for irregular sound correspondences involving Latin is too slim to confirm. When a solid IE root etymology can be proposed and several sound laws are involved, the case becomes stronger. But this is so far only the case for Lat. *brīsa*. Even there however, the vocalism suggests that its potential Proto-Albanian source did not receive the word from any attested source.

1.2.1.1.2 Thracian and Macedonian

Thracian is poorly attested and poorly understood. Our best information generally comes from Hesychius glosses (Fortson 2010: 463-4) beside a small number of inscriptions. Its name often appears in compounds like Thracio-Phrygian and earlier as Thracio-Illyrian, but the evidence is too fragmentary to connect Thracian to another subgroup with any amount of certainty. Thus, though Kretschmer (1923a: 229) gave as evidence of the Thracio-Phrygian development of PIE *mediae aspiratae* to *mediae* αββερετ < **b^her-* and αδδακετ < **d^heh₁-*, these are Phrygian forms and cannot be used to investigate Thracian. Our best evidence that Thracian had the same outcome of the voiced aspirates is Gk. βρῦτος and βρῦτεα (Hsch. βροῦτος, βρῦτιον), which seems to descend from **b^hru-* if related (as mentioned above) to Lat. *dēfrutum* and Engl. *brew*, *broth*. But as will be mentioned in §2.4.3, I am skeptical of the Thracian origin of this word. Nor does it explain the vocalism of Lat. *brīsa*. Its origins in Thracian may have nothing to do with the change and the mediating language could have been any language of the Mediterranean, known (like Illyrian) or unknown.

Because historical Thrace is not contiguous to Italic-speaking areas but does border on Greek-speaking areas, many of the hypotheses about words of Thracian origin are more relevant to Greek than to Latin. The Thracian pedigree of a Greek word has no bearing on how it is borrowed into Latin unless Latin has also borrowed the word from Thracian, which is geographically unlikely. For example, Boisacq (1911-12: 58-9) gives e.g. Gk. πίσος ‘pea’ and πύξος ‘box-tree’ as words of Thracian origin due to a lack of an etymology and the presence of the suffix *-aso* or *-so* with preserved intervocalic *s*. But Lat. *pisum* ‘pea’ is a direct borrowing from Greek neut. πίσον and a Thracian origin of Gk. πύξος (for which there is no morphophonological indication) does not explain the voiced consonant of Lat. *buxus*.

There are enough Macedonian personal names and glosses given as Macedonian that we can be relatively certain that it reflects PIE *mediae aspiratae* as *mediae*. Cf. for example Mac. Βίλιππος : Gk. Φίλιππος, Mac. Βερενίκη : Gk. Φερενίκη, and Mac. κεβαλή : Gk. κεφαλή ‘head’ (Fortson 2010: 464), though Méndez Dosuna (2012) argues that this is the result of a *Lautverschiebung* in what was originally a Greek dialect. Nevertheless, it seems that few Macedonian loans have been proposed for Latin, perhaps because it

suffers from the same problem as Thracian in that it is not contiguous with Latin-speaking regions.

The problem for all of these languages is that, while they could be used to try to explain a few mismatching forms between Latin and Greek by claiming that one language borrowed the form while the other language continues an inherited reflex, this is generally the extent of their usefulness. To account for larger numbers of words with irregular correspondences, we would need to assume either a substantial number of cases of this incongruous borrowing or perhaps that these languages were once more widespread—that they might have underlain Latin or Greek. It is unlikely that we will be able to claim this, at least for the languages in their historically attested states. Thus there have been a number of attempts to identify within Latin and/or Greek and/or other IE branches remnants of older Indo-European languages that may have reached the areas that Latin and Greek would come to be spoken in some time before they got there. The idea is certainly not absurd. Populations with steppe ancestry appear in Italy around 2000 BCE (Saupe et al. 2021) but Latin is not attested until much later. Part of this must be due to the lack of an alphabet on the Italian peninsula until the 8th c. BCE, but it is easy to image that other Indo-European languages got to the Italian peninsula before the Italic family did. Several lost PIE languages have been proposed to account for irregular reflexes, not all directly bearing on Italic.

1.2.1.2 Indo-European Substrates

Kretschmer (1896: 401-9), as will be mentioned in §1.2.2, in part summarizing what had been gathered up to that point, supported the idea of a non-IE speaking population having been present in Greece and Asia Minor before IE languages became settled in those areas. The convincing factor to him was the Gk. *-vθoς* suffix, not a gerundive in form or function, and which appeared on substantives of obscure etymology, several personal names, and placenames which corresponded to placenames in Asia Minor ending in *-ndos* and *-nda*. Gk. *-σ/-σσ/-ττ-* featured in his argument, but he noted that these can be the result of inherited morphology as well. He maintained this position for several decades (cf. still Kretschmer 1923a: 69). But by 1925, he began to change his mind, influenced by the paradigm-shifting discovery and decipherment of the Anatolian languages. He noticed a widespread, functionally diverse *nt*-suffix in words with a good IE etymology in Greek (e.g. *ἀνδριάς*, *ἀντος* ‘statue’), Slavic (diminutives like **ovъcъ* to **ovъcà* ‘sheep’), Italic (ethnic names like *Picentes*, *Fulcentes*, *Aventinus*), Germanic (**hri/unþiz-* ‘cow’ < **kr-ent-*), and Illyrian placenames in *-entum/-untum* (Kretschmer 1925a: 84-106). After listing several examples, he concluded that the non-IE *-vθoς/-anda* suffix was used in approximately the same ways as the *nt*-suffix he examined amongst the Indo-European languages. This could be due to chance, or it could suggest “daß schon vor der späteren Ausbreitung der Indogermanen sich eine indogermanische oder indogermanoide Welle nach dem Süden ergoß, und zwar über eine unindogermanische Urbevölkerung” (pg. 106).

In a second article in the same year, he used the early understanding of the recently understood Anatolian languages to make a logical postulation. If Lydian and Lycian are Indo-European, and if Etruscan is related to Lydian and to Tyrrheno-Pelasgian/Pre-Greek, then we should find Indo-European elements in Etruscan, Pre-Greek, and in the later Indo-European languages that settled over and around them, namely Italic and Greek (Kretschmer 1925b: 300-19). The relationships between the languages he mentioned may no longer be supported today, but the basic principle was clear: there could be traces of much older “protindogermanische” [*sic*] languages taken up by later Indo-European languages. And they should be visible: “In diesen Fällen erlaubte die unregelmäßige Vertretung der Verschlusslaute, die [protindogermanische] Herkunft von der [indogermanischen] zu unterscheiden” (pg. 310). He had set the stage for more systematic approaches to try to understand which irregular reflexes might be due to underlying but related languages—Indo-European substrates—with different sound laws.

1.2.1.2.1 *Pelasgian*

Taking archaeological investigations into account, Kretschmer updated his ideas in 1940 (231-78) and 1943 (84-218). He argued for two Pre-Greek layers: a non-IE Anatolian layer and an Indo-European stratum that he called Danubian. He identified the Anatolian layer with the ancient autochthonous *Leleges*, while the Danubian layer represented the *Pelasgoi*. The latter appeared as the archaeological Dimini Culture, bringing *Linearbandkeramik* elements from the North into the Balkans along with the *Protindogermansich* linguistic elements preserved in Greek. Katičić (1976: 57-87) in general provides a more detailed overview of the development of the Pelasgian theories than I could here ever hope to replicate. With the help of his summary, I will highlight a few of the main developments of the Pelasgian theory.

Milan Budimir, publishing from the 1920s to the end of the 1960s, preferred the name *Pelastik*, based on the *lectio difficilior* Πελαστικέ for Πελασγικέ in the scholiast to the *Iliad* 16.233. He saw Pelastik as preserving the three-way distinction of the original PIE velar series, and argued 1) that Albanian does this too and therefore continues this oldest PIE stratum of the Balkans and 2) that Slavic has many of the same features of this IE Pre-Greek language of the Balkans including being *satəm*, velarization of *-s-* after *u*, and preterit participles in *-lo*. Katičić’s criticism of his work is the same as for Kretschmer. The variation in the phonology of the proposed cognate sets is too great. It contradicts the main thesis that all the forms descend from one genetically homogenous language. An example of a word family claimed by Budimir to descend from PIE **d^heup/b-* ‘deep’ assumes *d ~ l* and *d ~ s* alternations, alternations between all nearly all possible vowels, and an *a-*prefix (Katičić 1976: 64-6):

Hsch. δάξα, var. δάψα ‘sea’	∴ Θέτις ‘Thetis’ < *Θέπις
	∴ Ὀδυσσεύς ‘Odysseus’ = prep. ο + δυσσ < *δυξ
	Lat. <i>Ulixes</i> = prep. u + lix
Gk. ζάψ ‘surf’ Alb. <i>det/dejet</i> ‘sea’	
Lat. <i>Tiberis</i> ‘Tiber river’	
Gk. σίμβλος ‘beehive’, σιπή ‘meal tub’,	
Gk. δέπας, δέπαστρον ‘beaker’	
Gk. ἀλάβαστρον/ἀλάβαστρος ‘vase for perfume’, λεπαστή ‘limpet-shaped cup’	
Gk. λαβρόνιον ‘large, wide cup’	
Lith. <i>dauburys</i> ‘valley’	

Vladimir Georgiev, especially by the time of his *Vorgriechische Sprachwissenschaft* (1941-5), had identified a pre-Greek *Pelasgian* underlying Greek that he considered an independent daughter branch of PIE. He was one of the first to propose a system of sound laws for Pelasgian. Based on his data, he found:

PIE *o > Pelasgian a		
PIE *R > Pelasgian uR or iR (*r, *l sometimes > ru, lu)		
Chain shift:	*b ^h > b	*b > p
	*d ^h > d	*d > t
	*g ^h > g	*g > k
		*p > ph
		*t > th
		*k > kh
PIE *k ^w , g ^w , g ^{wh} > Pelasgian k ^h , k, g		
PIE *k̥, g̥, g ^h > Pelasgian s (p); z (ð), i.e. it is a <i>satəm</i> language		
PIE *s preserved prevocally and intervocally		
Aspirate dissimilation occurred before any other changes		

Georgiev further claimed that this Pelasgian was the same language as the source of the Pre-Greek placenames. He does not convince Katičić on this account (pp. 79-80), who emphasizes that placenames are not reliably etymologizable. Several scholars built upon Georgiev’s premises. Weriand Merlingen followed Georgiev, but envisioned the IE Pelasgian language (he called it *Akhaean*) as a superstrate rather than a substrate. He additionally proposed a second IE language that influenced both Greek and Pelasgian. He called it *Psi-Greek* because of the first row of the consonant shift:

*p *t *k *k̥ > ps s ks ks
*b *d *g *g̥ > ph th kh kh
*b ^h *d ^h *g ^h *g ^h > b d g g
*k ^w *g ^w *g ^{wh} > ph bh b

Before the shift, aspirates were dissimilated in voicing and aspiration. As for vowels, *o > u, *e > i, *ā > ō, *ē > ā, but the syllabic nasals and resonants have irregular outcomes. This allowed etymological equivalences between Gk. θεός and Lat. *deus* ‘god’, Gk. ξανθός ‘blond’ and Lat. *candidus* ‘white’ as well as Gk. ἄνθρωπος and Gk. ἄνθρω, ἄνδρός ‘man’ (examples from Katičić 1976: 81). One sees how quickly this double system of superstrate languages can produce artificial results, especially when the matches are not perfect.

Otto Haas also followed Georgiev, and by 1960 had teased apart what he saw as two different layers: 1) That found by Georgiev (and Albrecht von Blumenthal in his 1930 *Hesychstudien*), which he called *hylleisch* after the Ἰλλυαῖοι of Istria. 2) An earlier layer that he called *vorgriechisch* to avoid historical interpretation, involving the change of **p*, **t*, **k* before **u* to *ps-*, *s-*, *ks-* (compare Merlingen's *Psi-Greek*) and **s > x* after **r* and **u* like in Balto-Slavic and Indo-Iranian. In this way, he explained e.g. Gk. ψύλλα 'flea' < Pre-Greek **phjulja* (cf. Lat. *pūlex* 'flea') and Gk. ὀξύς 'sharp' < Pre-Greek **akhjus* < **akús* 'sharp' (examples from Katičić 1976: 83-84).

Albert van Windekens (esp. van Windekens 1952) accepted the name Pelasgian, but only as a placeholder. He followed Georgiev's sound laws in general, amending **e > i* before a nasal, **u > o* in an initial syllable (but *u* in a second syllable), and noting that intervocalic **u* yielded *b* (Gk. ἐρέβινθος 'chickpea' vs. Lat. *ervium* 'bitter vetch'), which could be transformed into *m* through proximity to *n* (Gk. κυβερνάω 'to steer' vs. Cypriot *ku-me-re-na-i* 'they steer'). He also examined noun formation and explained the Pre-Greek suffixes through concatenations of Indo-European morphology (e.g. *-vθ-* < Pelasgian **-nth-* < IE **-n-t-*).

Albert Carnoy also followed Georgiev's sound laws and proposed numerous Pelasgian etymologies based on short roots with general meanings. (Cf. e.g. κόμωρος 'strawberry tree' < PIE **geu-* 'to bend, form a ball', μίνθη 'mint' < PIE **mei-* 'sweet, refreshing', νάρκισσος 'daffodil' < PIE **snerg-* 'to stiffen', σαλάμβη 'chimney' < PIE **swel-* 'absorb', Carnoy 1955a). This methodology led to criticism even from other Pelasgian scholars.

The problems with the methodology of the Pelasgian theories begin with the scholars themselves. Hester (1965: 347) writes that "the great ingenuity and erudition of the Pelasgianists," (the exact two words which Katičić (1976) also uses to preface his criticisms) "especially Georgiev and van Windekens, is generally praised, but...the greater the ingenuity, the greater the possibility of constructing a phantom Indo-European language from non-Indo-European material." Indeed, attempting to provide native etymologies for what had up to then begun to be suspected of being non-Indo-European was at the core of the Pelasgianists' methodology. In the end, they show that virtually anything can be provided with an IE etymology. Hester continues, "If we allow the Pelasgianists to postulate one new Indo-European language, we can hardly prevent them from postulating several (which is historically at least as plausible); but the postulation of them obviously increases the danger..." Of course there could have been numerous Indo-European languages in the area, now lost; the question is not one of plausibility but provability. Despite generally following the sound laws as proposed by Georgiev, there was disagreement amongst the Pelasgianists as to the exact phonology they reconstructed. Too many exceptions were allowed, defended on the basis of words entering Greek at different stages. This is theoretically valid—the way that language A borrows from language B changes as both languages' phonologies develop over time—but it is perilously difficult to prove. Another criticism shared by both Hester and Katičić is that most words given as Pelasgian show only one of the sound changes.¹⁰ Of those

¹⁰ Interestingly, Devoto (1943: 365) takes issue with exactly the cases that show more than one of the sound laws. However, this is because he sees several of the phenomena as having different sources. Thus

that show more than one, Hester (1965: 384) is only convinced by πύργος ‘tower, fortress’ (< **b^herǵ^h*), τάρρανον ‘vinegar’ (< **d^herǝ-g^h* or **ster-eg-*), τύμβος ‘grave mound’ (< **d^hmb^hos* cf. Gk. τάφος ‘grave’), and τρύγη ‘drought; harvest’ (< **d^hereg^h*- vel sim.), a list so small it could be due to coincidence. Katičić (1967: 76) places much weight on Hester’s (1965: 384) conclusion “It appears then that there is a small number of probable Indo-European loan-words in Greek,” but he does not finish the quotation. Hester continues “...borrowed most probably from neighbouring languages and not from a substrate or superstrate. These words are totally unconnected with ‘Aegean.’”

For these reasons, and particularly as evidence in favor of the Yamnaya theory of Indo-European dispersion has accumulated, Indo-European Pelasgian theories have generally been given up. Garnier and Sagot (2017) attempt a fresh approach with the criticisms in mind, as will be mentioned below. But the Pelasgianists were far from the only ones to attempt to find an Indo-European substrate amongst some of the daughter languages.

1.2.1.2.2 *Temematic*

Less relevant to Latin and the Mediterranean world but systematically similar to the Pelasgian hypothesis is the Temematic language described by Georg Holzer (1989). He proposes the existence of an Indo-European language from which Baltic and Slavic borrowed words with seemingly irregular outcomes. The sound laws are:

- 1) PIE *tenues* > *mediae* (The “*teme-*” of *temematic*)
**p*, **t*, **k*, **k̑*, **k^w* > *b*, *d*, *ǵ*, *g^w*
- 2) PIE *mediae aspiratae* > *tenues* (the “*-mat-*” of *temematic*)
**b^h*, **d^h*, **g^h*, **ǵ^h*, **g^{wh}* > *p*, *t*, *k*, *k̑*, *k^w*
- 3) PIE **r* > *ro*, **l* > *lo*
- 4) PIE **V* > *V* / _ *i*, *u*, *r*, *l*, *m*, *n*
- 5) PIE **er* > *ir* / _ *C*, #
- 6) PIE diphthongs immediately before a sequence **CV* received acute accentuation. This functioned after rules 4 and 5.

Holzer’s analysis benefits from a number of criteria he sets in order to ensure that his data are as objective as possible, including root length (shorter roots are more likely to be coincidence), root sharpness (the number of different root reconstructions allowed for by the comparanda due to the sound laws of the languages involved), probability of having

the connection of Gk. σοφός ‘clever, wise’ and Lat. *sapiō* ‘to know’ is impossible to him because it shows an Armenian-esque consonant mutation and an Illyrian-esque retention of **s*. What Georgiev saw as evidence of a pre-Greek IE Pelasgian language, according to Devoto, could be the result of “filoni diversi” (he mentions an orientalizing stratum and a central stratum) which fused to create Proto-Greek. He later argues for the existence of “peri-Indo-European”, languages on the margins of the IE linguistic area that are affected by non-IE features and lexemes. The ideas are not impossible, but it is interesting that he proposes such concepts as an alternative to Georgiev’s more systematic attempt to identify recurring sound laws.

an IE origin, semantic closeness, degree of abstractness of the semantics, and number of synonyms. Because of this, many of his comparisons are quite good. Kortlandt (2010: 73-80) accepts the existence of Temematic and takes it seriously enough to speculate on its placement in the IE tree.

The ability to establish that a root is of IE origin is per se difficult, especially when irregularities are involved. A large number of comparanda does not guarantee that a root is inherited (especially when the attestations are geographically close). For example, Holzer (1989: 121-2) connects PBSl. **tron-* ‘drone’ (PSlav. **trǫtъ*, Lith. *trānas*, Latv. *trans* ‘drone’) to Germanic and Greek drone words from **dʰrōn-*. But problems with reconstruction within Germanic (cf. Kroonen 2013: 101) and Greek (cf. EDG 105, 554) suggest instead that the whole group represents loans from a non-IE language. On the other hand, Holzer’s (1989: 51-4) connection of e.g. Lith. *būr̃žė* ‘furrow’ to Lat. *porca*, OHG *furh*, and MW *rhych* ‘furrow’ < **pŕk-* is more difficult to do away with. Its potential reflex in Ved. *pársāna-* ‘low sunken place?’ seems to lend credence to an IE origin. What makes it particularly compelling is its irregularity in Balto-Slavic compared to the common pre-form reconstructible for the other branches. Distributions like this seem quite significant; while an IE substrate is not the only explanation (perhaps the lexeme was mediated by a non-IE language, cf. Devoto’s 1943 *peri-indoeuropeo*), it should be kept in mind.¹¹

1.2.1.2.3 Indo-European Substrates directly involving Italy

1.2.1.2.3.1 Ribezzo’s Ausonian

Francesco Ribezzo postulated throughout several articles (1928, 1929, 1930, 1932, 1934a etc.) that pre-Italic Southern Italy from Rome to Sicily was home to an IE language that he named Ausonian after the inhabitants of Southern and Central Italy (called Ἀϋσωνες by the Greeks, *Aurunci* by the Romans). Its main feature was the reflection of *mediae aspiratae* as *tenues*, but almost all examples show **dʰ > t*:

Leuternii etc. < **leudʰer(i)no-* < **h₁leudʰ-* ‘people’
Rutulī < **rudʰ-* ‘red’ (1930: 92)
 Mount Aetna (Ἄιτνη) < **aidʰenā* < **h₂ejdʰ-* ‘to burn’ (1928: 192-4, 1934a: 91)
 Etr. *lautn* ‘free?’ = *līber* < **h₁leudʰ-* (1929: 64, 1932: 32)
 Sicel λίτρα ‘Sicilian coin’, Lat. *lībra* ‘Roman pound, scales’ < **tlēidʰrā-* < **telh₂-* ‘to carry’

Szemerényi (1991 II: 655-672) reviews the idea by means of updating the etymology of λίτρα/*lībra*.¹² He is convinced by Ribezzo’s arguments that “Sicel-Ausonian” had the

¹¹ Holzer (1989) lists only the cases that work with Temematic sound laws. It would be interesting to see a list of cases of alternation between Balto-Slavic and other IE branches that attest to different irregular correspondences.

¹² He asserts that Germanic forms (like OE *lēad*) are not borrowed from Celtic (cf. MlR. *lúaide*) but instead that the Germanic and Celtic lead words are from **loudʰo-* ‘lead’, whence Lat. *lībra* as “in the

change $*d^h > t$ and in fact further proposes to have found evidence for $*b^h > p$ as well. He notes some cases that Ribezzo adduces that do not show this development (like $*h_2elb^h$ - ‘white’ > ΑΑΛΙΒΑΝΟΝ = *Allifae*, $*reud^h$ - ‘red’ > *Rudiae*, pg. 663), but thinks they can be safely removed from consideration (pg. 671-2).

The $*b^h > p$ change is dubious: in Paulus *ex Festo* it is claimed that Latin has *album* ‘white’, Greek has ἀλφόν, but Sabine has *alpum*. Szemerényi (1991 II: 671) argues that it should be taken seriously and that it is a remnant of the dialect of their predecessors. Neither of these proposals is warranted in my mind. He further mentions *Alpēs* ‘Alps’ < $*h_2elb^h$ - ‘white’ and *Tiberis* ‘the Tiber’ < $*d^heub^h$ - ‘deep’, but as toponyms there is no guarantee that their original semantic value is recoverable. The etymological analysis of ethnic and geographic names is extremely problematic, and Etr. *lautn*, if borrowed from Italic, can simply have been borrowed before the change from $*h_1leud^h$ - > *lib-* (on this, cf. Rix 2005: 564, who suggests an Umbrian source).

The only case where it seems like Ribezzo and Szemerényi were on the right track is that of λίτρα ~ *libra*, but it does not carry the significance that they claim. Rather than evidence of an IE Ausonian language where $*d^h > t$, λίτρα is probably a loan from Sicel (Lejeune 1993: 2, Weiss 2021) where $*d^h > d$, and *d* was devoiced before sonorant consonants (Willi 2008: 22). Current scholarship leans towards Sicel being an Italic language (cf. Weiss 2020: 15 fn. 26), making it unlikely to have served as a substrate for the peninsula. Furthermore, the problems with a multi-wave origin of the Italic languages (as espoused here by Szemerényi 1991 II: 682-5) is discussed in §7.2.

1.2.1.2.3.2 Haas’s Frühitalische Element

Otto Haas, as mentioned above, had followed much of what Georgiev proposed for the Indo-European substrate language(s) in Greece, but would not call any of them Pelasgian. In 1960, he searched for traces of these in Latin. This too built off of work that Georgiev had done in his 1938 *Die Träger der kretisch-mykenischen Kultur, ihre Herkunft und ihre Sprache, II. Teil: Italiker und Urillyrier; Die Sprache der Etrusker*. But Haas set out to do it better by not assuming from the beginning that Etruscan was a dialect of Pelasgian.

His methodology included 1) that the only reason to compare an Italic word to another IE word despite a mismatch in form is its meaning. 2) The meaning must be well established; not hypothetical. He explicitly stated that this makes most of the Etruscan material unusable and that toponyms are generally not helpful. 3) The larger the portions of the words that correspond, the better (i.e. they do not share simply a root-like element). 4) The same strict standard must be used for counterarguments; alternative explanations do not take precedence simply by virtue of them being more conventional.

Bronze Age, the standard weight was of lead’ (Szemerényi 1991 II: 662). Lat. *plumbum* would thus be unrelated. In §2.2.2.1, I instead follow the etymology in Thorsø & Wigman et al. (2023: 117).

Despite his attempt at a strict methodology and his disclaimer that this is simply a preliminary attempt, his data leave an overall unconvincing impression. His comparisons led him to some of the following sound changes for *Frühitalisch*:

**b^h*, **d^h*, **g^h* > *p*, *t*, *k* (*parcō* ‘to spare’ < **b^horg^h*- cf. PGM. **bergan* ‘to keep (safe)’
 **b*, **d*, **g* > *p*, *t*, *k* (*columba* ‘dove’ vs. PSlav. *golqbb-* ‘dove’)
 **u* > *f* (*favus* ‘honeycomb’ < **uob^h-o-* cf. OHG *waba* ‘honeycomb’ but elsewhere
 ‘weave’)
 **o* > *a*
 **ou*, **eū* > *au*
 **au* > *iū* (*iubar* ‘starlight, sunshine’ < **ausr-* with *-br-* < **-sr-*, cf. Skt. *uṣar-* ‘dawn’)
 **āu* > *lu* (*luscinia* ‘nightingale’ < **āusi-kan-īā-* ‘Morgensängerin’)

Some of the developments have alternative outcomes or are otherwise suspicious. For example, after a nasal **b^h* > *b* and **g^h* > *g* (in *pīnguis* ‘fat’ < **b^hḡg^hu-*). But in *truncus* ‘mutilated; trunk’ < **d^hrong^hos*, **ng^h* instead produced *c*. While **āu* is meant to produce *lu*, *lutra* ‘otter’ (< **udrā-*) would suggest that **u* has the same result. Both **b^h* and **d^h* are meant to produce *b* after **u* in *Frühitalisch*, though these are simply regular Latin sound laws (the former is the expected development of non-initial **b^h* and the latter is the RUBL Rule [Weiss 2020: 84]).

There are furthermore issues with some of his etymologies. For example, he compares Lat. *filix* ‘fern’ to the Greek and Germanic willow words of the shape **welik-*. Lat. *culpa* ‘blame’ he compares to derivations of PGM. **glapp/bōn-* ‘to slip off’ such as ON *glapna* ‘to spoil, become useless’. He justifies the connection by comparison to Lat. *vinum culpatum* ‘spoiled wine’; this despite his warnings against hypothetical reconstructions of meaning. Lat. *tuba* ‘trumpet’ he compares to Ru. *dudá*, etc. ‘pipe (instrument)’, leading him to reconstruct **d^hud^h-* and **d^hou^hd^h-* respectively; but **d^he^hud^h-* is not a permissible root structure for PIE (unless perhaps an exception given its onomatopoeic function). He collected much material, but its usefulness is in the irregular correspondences he notes between forms rather than his IE etymological connections.

1.2.1.2.3.3 Garnier and Sagot

Garnier and Sagot (2017) set forth to re-analyze the situation, taking into account the criticisms of the Pelasgianists’ methodology (semantic implausibility, exceptions to the proposed sound changes, one sound change per word, non-conformance to PIE word-formation patterns), and propose to have found underlying Latin and Greek traces of an IE substrate language that was 1) *centum*, 2) underwent a change in which “voiced aspirated consonants (at least PIE **b^h*) would be reflected as non-aspirated voiceless consonants, at least in certain contexts”, 3) had syllabic sonorants, 4) had fortis consonants, and 5) underwent stress regression. The *centum* rather than *satəm* nature of the language and the change **D^h* > *T* rather than **D^h* > *D* puts their substrate in league with Temematic and Ausonian rather than Pelasgian.

Their analysis first proposes a frequent nominal morph **p*, for which I am not convinced there is sufficient PIE evidence.¹³ This allows proposals like (with their proposed phonetic laws in parentheses):

- (1) **leiH-* ‘to drip, pour’: **loi-p-eh₂-iéló-* > **loiβā̃-* (intervocalic lenition) > **loiβā-* (stress retraction). Gives Hsch. λoιβᾶται ‘libates’ and then deverbal Gk. λoιβή ‘libation’, secondary Gk. λείβω ‘to shed, pour’; Lat. *libāre* ‘to pour, spill’.
- (2) **dór-u-* ‘chopped piece of wood’: **dr-óu-p-o-* (action noun), **dr-ú-p-eh₂-* (collective), analogically levelled to **dr-ú-p-o-* > **drúPo-* (consonant fortition). Gives Gk. **drúp^hos*, denominal verb **drup^h-jo* > Gk. δρόπτω ‘to scratch’
- (3) **kél-* ‘to cover’:
- **kí-óu-p-o-* ‘covering’ > **kálúPo-* (pretonic fortition and anaptyxis). Gives Gk. κέλῦφος ‘shell’.¹⁴
 - **kí-u-p-éh₂* ‘set of huts’ > **klubá-* (posttonic lenition). Gives Gk. καλύβη ‘hut, cabin’.

They reconstruct several stages of derivations, which are difficult to prove but cannot be rejected outright. In fact, it is an important consideration: an Indo-European language, even one that served as the substrate to another, would have undergone the same complex derivations that we see in the attested languages. I take more issue with the semantic implausibility or sometimes vagueness of several of their suggestions. Gk. δρόπτω ‘to scratch’ above is difficult to envision from **dór-u-* ‘chopped piece of wood’ (actually ‘tree’). Lat. *porrum*, Gk. πράσον ‘leek’ they derive from **b^hers-*, traditionally ‘to point, burst’ but which they define as ‘to break into pieces; to break through the ground upwards (especially speaking of a plant)’. Gk. βροῦκος ‘edible locust’ they take from **(s)preug-* ‘to jump’. Gk. βλέπω ‘to look’, βλέφαρον ‘eyelid’ they derive from **b^hleg-* ‘to sparkle, shine’. To list these cases is of course to ignore the cases that do work well,¹⁵ but suggests that such cases may have other explanations (e.g. chance, onomatopoeia, origin in a non-IE substrate). In Garnier and Sagot (2020) they propose that some of the material they treated may be the result of an Anatolian (mainly Lydian, in part Lycian) substrate in Greek. Their reanalysis of several proposals still relies on a, to me, unconvincing link to PIE roots.

¹³ Without committing to a comparison, I would like to mention how this calls to mind Hubschmid (1963 I), most of which is dedicated to a study of a purportedly widespread substrate *p*-suffix. Though he does e.g. agree that κέλῦφος is of inherited origin (p. 34).

¹⁴ Merrit (2021) offers an alternative explanation of the labial element as part of a light verb construction with **b^huH-*.

¹⁵ Some of the sound laws they propose allow the comparison of Gk. βρέμω ‘to roar, grumble’ with Lat. *fremō* ‘to rumble, roar’, OHG *bremān* ‘buzz’, and MW *brefu* ‘roar’; Gk. πόνδαξ ‘bottom of a jar’ with Lat. *fundus* ‘bottom’; Gk. θρόνον ‘reed, rush’ with PBSI. **trus-* ‘reed, cane’; and a derivation of Gk. θρίναξ ‘trident’ from PIE **tri-* ‘three’.

1.2.1.3 Verdict on Indo-European Substrates

In sum, these attempts are all based on the perfectly plausible idea that the attested Indo-European languages were not the first Indo-European languages to be spoken in the area where they are attested. The irregularity of the data can *a priori* be due the intersecting effects of contact with 1) any number of different lost IE languages 2) over a long period of time in which both sides of the contact situation were undergoing sound changes.

On the other hand, it is equally plausible that such a complex pattern can have other explanations. It is certainly possible that Gk. τῦμβος is borrowed from a language in which it is the inherited reflex of **dʰmbʰ-o-*, like Gk. τᾶφος, but which deaspirated PIE *mediae aspiratae*. But if Greek has borrowed it, then it is equally plausible that the sound changes are due to a non-IE language's treatment of PIE **dʰmbʰ-o-*. It is also possible, in light of Lat. *tumulus*, that τῦμβος is not from the Indo-European formation **dʰmbʰ-o-* at all. The borrowing of a plosive into an IE language could only be mapped in terms of its voicedness, its aspiration, and (of velars) its palatal or labial quality. So a borrowing from a non-IE language that has its foreign sounds mapped onto native ones would look the same as a borrowing from an IE language with different sound laws. Given the amount of irregularity that such hypotheses do not explain, especially the cases where a link to an attested IE root is tenuous, perhaps other explanations are not simply plausible but even preferable. Alongside the field of explanation that proposed lost IE languages, a parallel field developed around the idea of the existence of non-IE substrates.

1.2.2 Non-Indo-European Sources

Whether or not traces of lost Indo-European languages exist in the attested daughter languages and whether these are visible as such, there are certainly a large number of cases for which an attempt at an Indo-European etymology is futile. Thus the focus of the rest of this chapter, and of this thesis in general is on the research of non-Indo-European sources of Latin vocabulary.

The history of research on non-Indo-European substrates began, as did that of the Pelasgianists et al. above, amongst the Romanists and especially Ascoli. Hirt (1894) had proposed that the differentiation of the Indo-European daughter languages was due in part to contact with different groups of languages. Kretschmer (1896: 401-9) importantly furnished evidence for what he understood as relics of an explicitly non-Indo-European language. From here, the field seems to have developed in several directions of which I will detail two for their direct impact on the motivation and methodology of this work. The first was a continuation of Kretschmer's work in the Mediterranean and the other, inspired by the phenomenon that Kretschmer identified, developed around Germanic.

1.2.2.1 The Two Lineages

1.2.2.1.1 *The Mediterranean Substrate*

Antoine Meillet (1908), inspired in part by Sommer's comparison of Lat. *plumbum* and highly diverse Gk. μόλυβδος, proposed a list of words that Latin and Greek had independently borrowed from "un troisième langue inconnue," explicitly ruling out Semitic as the source. These included *vaccinium*, *cupressus*, *menta*, *rosa*, *lilium*, *figus*, *libra*, and *vinum*. The article introduced the concept of a Mediterranean substrate to linguistics.¹⁶ Albert Cuny (1910) took from both Kretschmer and Meillet to further develop the list. He importantly noted not only the words with the already supposed non-IE suffixes, but went on to describe the phonological alternations that occurred within variants of these words as well as several other cases of irregular correspondence between e.g. Latin and Greek. His insights were remarkably prescient, and many of his suggestions withstood the test of time and are accepted in this work.

In Italy (cf. Craddock's 1969 overview of the field, which I follow here), research on the Mediterranean substrate was taken up early on by Francesco Ribezzo and Alfredo Trombetti. Ribezzo (1920a, esp. 1920b) proposed to have found evidence of a unitary substrate language in toponyms of the Mediterranean. Trombetti, known otherwise for his work on linguistic monogenesis, took Ribezzo's approach to the extreme and saw two strata underlying the Indo-European languages of the Mediterranean: a Basque-Caucasian one and an Etruscan-Asia Minor one (cf. one of his latest formulations in Trombetti 1927: 220).

Vittorio Bertoldi¹⁷ brought an amount of scientific rigor to the study of the Mediterranean substrate by investigating the lexicon in addition to the onomasticon. He laid out methodological considerations in a 1932 article and again in two books (Bertoldi 1939a, 1942). But Craddock (1969: 38-40), in reviewing Bertoldi's body of work, notes a general lack of transparency as to whether he considered the Mediterranean substrate to be a uniform language. While he grew to focus more and more on regional variation within the stratum, he also identified pieces of morphology that seemed to overlap broad areas. "In Bertoldi's writings it rarely, if ever, becomes crystal-clear exactly what languages belong to the Mediterranean substratum, what structural features words attributable to that substratum share, in short, what sort of linguistic reconstructions can be labeled uniquely Mediterranean" (Craddock 1969: 40).

Giacomo Devoto introduced the idea of Peri-Indo-European (Devoto 1943) as an

¹⁶ The development of this concept probably cannot be separated from the racial theory of "Mediterraneanism" proposed by Giuseppe Sergi in his 1895 *Origine e diffusione della stirpe mediterranea*. As an opposition to Nordicism, it proposed the existence of an autonomous, i.e. non-Nordic, non-African, Mediterranean race and culture with the further stipulation that they were the greatest race in the world (cf. Craddock 1969: 31 fn. 19).

¹⁷ As a student of Wilhelm Meyer-Lübke in Vienna, he met Carlo Battisti, who was at the time already a lecturer at the university.

alternative in response to the work of the Pelasgianists who were in many cases (especially Georgiev) all too willing to explain away all irregularities as traces of lost Indo-European languages. As an almost intermediate approach, Devoto's Peri-Indo-European languages/cultures/peoples were those of non-IE origin that became Indo-Europeanized, serving as vectors of non-IE phonology and morphology in the cases where they came to influence other Indo-European languages.

Giovanni Alessio, a student of Carlo Battisti, began to publish in the field of substrate studies with two articles (actually one, split into two) in 1935 and 1936. In 1955 he published a book overviewing his understanding of the situation. Similarly to Trombetti and in opposition to Bertoldi, he saw the Mediterranean substrate as a mostly homogenous unit. The patterns of phonology, morphology, and lexicon that he found reached from the Atlantic coast, through the Baltic, from Iberia to Asia Minor, on to the Caucasus and even to India (Alessio 1955: 220). He had earlier noted that "Mediterranean" was a term "which has gradually been emptied of its geographic significance"¹⁸ (Alessio 1946: 142).¹⁹ This inspired criticism from many, amongst them Alessio's rival Johannes Hubschmid.²⁰ Hubschmid was a student of Jakob Jud in Zürich, co-author of the 1920-40 *Sprach- und Sachatlas Italiens und der Südschweiz*. Thus the tradition of substrate research growing very visibly out of the study of the Romance languages was not restricted to the Italian scholars (many of whom had studied under Wilhelm Meyer-Lübke in Vienna or Jules Gilliéron in Paris). Beginning his publications in the field in 1943, his 1960 *Mediterrane Substrate* presents a less geographically extreme idea of the Mediterranean substrate, in fact two: an older *Eurafrikanische* (though he still calls it Mediterranean) substrate underlying most of Western Europe and a younger, Basque-related *Hispano-kaukasische* Mediterranean substrate (Hubschmid 1960a: 24). He identified further strata elsewhere as well, including several in Sardinia (Hubschmid 1953). His work began to culminate in the *Thesaurus Praeromanicus*, with volume 1 (1963) focusing on the evidence for a substrate *p*-suffix and volume 2 (1965) on Basque and Basque-Pre-Romance contact (including a discussion of Bertoldi's and Alessio's work). No further volumes were produced.

Thus from its inception, research on the Mediterranean substrate has been well informed by a large amount of material, especially from a detailed understanding of the Romance languages, using as important evidence the irregular phonological alternations between comparanda. But it has also suffered from a few problems. The reliance on toponyms as independent evidence has potentially over-extended the boundaries of the phenomenon. There has never been uniform agreement as to its geographic extent or the number of

¹⁸ "Che si è andata un po' alla volta svuotando del suo significato geografico."

¹⁹ Lewy (1934) had already used a definition of "Mediterranean" that extended to the Caucasus and in fact included parts of West Africa.

²⁰ Interestingly, Johannes Hubschmid's father, Johann Ulrich Hubschmied, was a critic of the (Italian) substrate linguists, writing: "In der Tat stehe ich all diesen Etymologien aus der voridg. *kala-pala-bara*-Sprache skeptisch gegenüber" (1942: 118).

strata it comprises.²¹ The spectrum ranged from Alessio's enormous, monolithic "Mediterranean" substrate, with regional variations, to e.g. de Simone's (1963: 196) conclusion that the evidence allows us neither to transgress well-delimited *Sprachräume* nor to postulate on the genetic relationship of these.

1.2.2.1.2 *The European (Germanic) Substrate*

In 1899, Bruno Liebich published *Die Wortfamilien der lebenden hochdeutschen Sprache als Grundlage für ein System der Bedeutungslehre*, in which he classified the German lexicon based on etymologizability.²² Hirt (1909: 59) followed Liebich's figures, wherein nearly one third of the German vocabulary, a significant part of it pertaining to maritime technology, was not attested outside of the Germanic family. Feist (1910: 350-1) then used these two works to conclude that 30% of the Germanic vocabulary came from a pre-Germanic non-IE substrate. Thus the Germanic substrate hypothesis was born²³ and the figure of 30% was repeated into the 2000s.²⁴ The idea that a pre-Indo-European language should have existed in North Europe was, especially after the discovery of Hittite, easily defensible: North Europe was almost certainly not the *Urheimat* of Proto-Indo-European.²⁵ Those defending the Germanic substrate hypothesis even drew on Kretschmer's work that had already been published on the non-IE strata in Latin and Greek (e.g. Feist 1932, Güntert 1934: 72). But Feist's 30% figure was the result of a misunderstanding of Liebich's categorization: the etymologies on which the categorization was made were very conservative (Prokosch 1939: 23) and the 30% included innovative combinations of nonetheless inherited morphology that simply did not occur outside of Germanic (Neumann 1971b: 78). Some scholars had sought other types of evidence, such as Ribezzo (1934a) and Güntert (1934) who suggested external influence triggered the first Germanic consonant shift. They pointed to what they understood as a similar change in Etruscan; thus the pre-Germanic substrate would be a relative, perhaps Alpine Rhaetic. But such a hypothesis was questioned already by Pokorny (1936:81-2) and is no longer considered today.

In contrast to research on the Mediterranean substrate, which relied in part on irregular phonological correspondences within and between languages, the earliest conceptions of the Germanic substrate hypothesis relied on the non-etymologizability (i.e. isolation) of lexemes. This was a critical weakness of the methodology. Streitberg and Michels (1927: 51) quite reasonably noted that, as long as there is nothing explicitly non-Indo-European about the word, its lack of etymology might be an ephemeral lacuna awaiting a solution

²¹ Karel Oštir (esp. 1921, 1930) is an outlier, in that his version of "Alarodian" included the Caucasian families, Sumerian, Urartian, Elamite, Egyptian, Berber, Basque, Etruscan, etc.

²² This summary follows Bichlmeier (2016).

²³ Though Neumann (1971b: 85) records even earlier ideas, for example Friederich Kluge's 1901 (second edition) *Urgermanisch. Vorgeschichte der altgermanische Dialekte*, in which he proposed *Apfel, Krug, Silber* and *Go. baira-* 'mulberry' were from an unknown *Urvolk*.

²⁴ Since Feist (1932). Cf. also recently Rifkin (2007: 54).

²⁵ Published in English in the American journal *Language* because he had been so denounced for his *germanenfeindliche* ideas that no German journal would publish his work (Mees 2003: 20).

from improvements in the field. “Wer sagt uns, daß dies nicht morgen der Fall sein werde?” It took longer than a day, but Bichlmeier (2016: 323-4) notes two times (already Prokosch 1939: 23, Mees 2003: 26) at which it had already been claimed that most of the etymologies had been provided. Neumann (1971b), too, predicted that the number of pre-Germanic substrate words proposed on this basis would shrink as better etymologies came out. Pokorny, who was by no means opposed to the idea of a pre-IE substrate,²⁶ had taken this into account to conclude from other evidence (*Lautbestand*, *innere Sprachform*, and *Flexion und Wortbildung*), “that, from a purely linguistic perspective, there exist only few indications of a substrate. But these in all likelihood point to an ancient relationship with Finno-Ugric languages” (Pokorny 1936: 84).²⁷

The 1960s saw a development along the lines of Pelasgian. Łowmiański in 1963²⁸ criticized Pokorny for not taking into account the possibility of an *Indo-European* pre-Germanic substrate. Some Pelasgianists had already compared the sound changes they found to the Germanic *Lautverschiebung*. Hans Krahe (e.g. 1964) proposed to have found an inherited Old European stratum of linguistic unity behind Central and Western European hydronyms. But Mees (2003: 23) remarks that Krahe provided no actual proof that the hydronyms were from one monolithic stratum: “Having the model of Illyrian in mind, he assumed that together these elements represented the remnant of one archaic language, whereas in fact they may well represent the remains of a number of different Indo-European (palaeo-)dialects.” At about the same time as each other, Maurits Gysseling and Hans Kuhn (esp. Kuhn, Hachmann, and Kossack 1962) remarked on several place-names in the region of Belgium preserving a *p*, making it unlikely that they were native Germanic words (given the rarity of PIE **b*) and unable to be of Celtic origin (as **p* is lost). Other elements were built up that seemed to point to a third language originally native to the area. The former called it Belgian (cf. Gysseling 1975), the latter the *Nordwestblock*, but both agreed it was Indo-European. Meid (1986) defended the idea of the existence of a Germanic-related, non-Celtic language against criticism but remained unsure about its extent.

Credible progression on the research of the Germanic substrate would require updated methodology.

1.2.2.2 Uniting the Lineages

The topic of this thesis as well as its methodology probably has F.B.J. Kuiper to thank for, whether directly or not, uniting various lines of substrate research. He first wrote on the topic of substrates in his 1948 *Proto-Munda Words in Sanskrit*, then on the pre-Greek substrate in 1956. He would not write about the substrate of Germanic until 1995, but

²⁶ In fact, he wrote several articles on a non-IE substrate in Irish in the *Zeitschrift für Celtische Philologie* (1927 vol. 16: 95-144, 231-266, 363-394; 1928 vol. 17: 373-388; 1930 vol. 18: 233-248).

²⁷ “...daß rein sprachlich nur geringe Anhaltspunkte für ein Substrat vorhanden sind, die aber mit einer gewissen Wahrscheinlichkeit auf uralte Beziehungen zu den finno-ugrischen Sprachen hinweisen.”

²⁸ Łowmiański, Henryk. 1963 *Początki Polski: Z dziejów Słowian w I tysiącleciu n.e.* Warsaw: Państwowe Wydawnictwo Naukowe. Page 45. *Non vidi*, apud Witzczak (1996: 169-70).

two of his students, Edzard Johan Furnée and Robert S.P. Beekes, meanwhile made critical developments in the methodology of substrate research.

Kuiper's 1956 article formed the basis of Furnée's dissertation, published in 1972 as *Die wichtigsten konsonantischen Erscheinungen des Vorgriechischen*. The work undertook to describe all of the irregular consonant alternations (with an appendix on vowels) that occur within Greek and thus point to the influence of a non-IE substrate. Interestingly, he concludes that the variation was due to expressive alternation within the substrate language rather than to the borrowing process. Georgiev (1971: 164-67), who could think only of Pelasgian, reviewed Furnée harshly: "...These are apparently the sound laws(?) of Pre-Greek. The book presents a misguided attempt to rescue the already obsolete pan-Mediterranean idea. Its fundamental errors attest to an insufficient linguistic education,"²⁹ later saying that it is of no scientific value. Beekes on the other hand, who a year earlier had already used irregular alternations to argue for the European substrate origin of Gk. *λέπω* 'to peel' (Beekes 1971), valued it greatly, calling it "without a doubt a turning point in the study of the Greek substratum" (Beekes 1975). He would incorporate its insights into his 2010 Greek etymological dictionary.

Around this time, research on the topic began to appear in the United States as well. Eric P. Hamp wrote in 1975 and 1979 about substrate words, in the former noting a case of invalid root structure in words for 'to cut' and in the latter importantly noting the presence of **a* and an *a* ~ *o* alternation in reflexes of the apple word. Hamp and several others, especially Edgar Polomé, would continue to publish on substrate topics throughout the 1980s.

The late 80s and the 90s saw a boom of substrate research output, especially in Leiden. Beekes developed his ideas on the pre-Greek substrate, with their roots all the way back in Kretschmer's work, culminating in his 2010 etymological dictionary and a 2014 monograph. Beekesian Pre-Greek was supposed to have been restricted to the geographic area of Greece, though this assertion will be investigated in this thesis. But even more work was being done on the evidence for a non-IE substrate language present outside of and beyond Greece, including by Beekes (Beekes 1996, 1999, 2000). Kuiper's 1995 article³⁰ focused on the importance of non-IE **a* to detect three substrates within Germanic: Krahe's Old European (but this time interpreted as non-IE), a stratum specific to Germanic, and a wider "European substrate." Peter Schrijver, having studied in Leiden under Beekes, wrote a 1997 article on "Western European" substrate words, focusing on regular irregularities, that is "phonological and morphological alternations which are regular in the sense that they occur in more than one etymon according to a certain

²⁹ "...das sollen die Lautgesetze(?) des Vorgriechischen sein. Das Buch stellt einen verfehlten Versuch dar, die schon überholte panmediterransische These zu retten. Seine Grundfehler zeugen von ungenügender sprachwissenschaftlicher Schulung."

³⁰ He had also resumed the theme of his 1948 book in a 1991 article where he searched for non-IE words in the Rigveda. A decade later, Lubotsky (2001) would apply the updated methodology of the late 90s to search for substrate words that had entered Proto-Indo-Iranian.

pattern but irregular in the sense that they cannot be explained, for some reason or another, on the basis of Indo-European phonology and morphophonology” (pg. 296). Dirk Boutkan, a student of Beekes and Kortlandt, wrote in 1998 on “North European” substrate words that he found in Germanic, and many of his insights on the substrate were included in the 2003-2009 *Etymologisch Woordenboek van het Nederlands* by Philippa et al. Beekes in 2000 wrote on “European” substrate words in Greek, noting that they likely represent a different contact situation from that preserved by words with a more Mediterranean distribution.

Outside of the Netherlands, Eric Hamp and Edgar Polomé were joined by other Americans including Martin Huld and Joseph Salmons. Polomé (1986) provided a refreshed look at the Germanic lexicon lacking IE etymologies, indicating that, by this date, they were unlikely to receive them. Hamp (1990), drawing on his 1979 proposals about the apple word, attributed to a substrate in “Northern (Central) Europe” an *a ~ o* alternation arising from **ɔ*, **b*, and **DeD* shaped roots. Huld (1990) wrote on words shared in two distributions, an “Alpine” group and a “North Balkan” group, which nonetheless shared “similar phonological, morphological and syntactic patterns” such that they might be remnants of a Sprachbund.

The modern methodological considerations and assumptions that allowed substrate research to begin advancing again had been mentioned all throughout the works of these scholars, but especially in Polomé (1989)³¹ and Schrijver (1997).³² The former’s list was refined in a sense by Salmons (1992), who ranked the types of evidence. As most important he placed discrepant phonological or morphological features vis-à-vis IE, especially when they do not rely on isolated lexemes but rather show irregular correspondences. Weakest are items simply lacking a clear IE etymology. Schrijver’s list (used also by Lubotsky 2001) was even more refined. It presented the items with the understanding that none of them could stand in isolation; rather it was the accumulation of circumstantial evidence that made a substrate origin more likely:

1. Limited geographic distribution
2. Phonological or morphological irregularity
3. Remarkable word formation
4. Non-IE phonology
5. Semantic category

As Bichlmeier (2016: 327) notes, “Die ursprüngliche Liste der vermeintlichen Substratwörter spielt bei diesen Untersuchungen interessanterweise (so gut wie) keine Rolle mehr.” Thus the main early criticism of pre-Germanic substrate research had been

³¹ In part as a review of the work of Gysseling and why it failed to produce convincing results.

³² Schrijver’s 1997 paper crucially remarked on the phenomenon of the “*a*-prefix” with concomitant vocalic reduction (cf. Gallo-Lat. *alauda* vs. PGm. **laiyār/z-* ‘lark’). Interestingly, a “präfigiertes *a-*”, thought to be of Hattic origin, was first noticed as early as the 1930s (Kretschmer 1933a: 86) in placenames of Asia Minor.

addressed and solved. Research continuing the work on the Mediterranean substrate could benefit from this refined methodology as well, interesting in light of the fact that the consideration of irregular alternations always played an important role there. Two contradictory assumptions seem to have hindered its progression, revolving around uniqueness and exclusivity. On the one hand, there seems to have been an underlying assumption that it was the only substrate. Thus as more evidence of substrate words shared across Europe grew, the geographic range of the “Mediterranean” was expanded until the word became meaningless. On the other hand, the Mediterranean substrate was exclusively Mediterranean. This is interestingly continued in Beekesian Pre-Greek, where it is assumed the substrate language was not spoken outside Greece. In reality, several features supposed to be unique to it occur elsewhere.

Guus Kroonen continued the search for non-IE substrates in Germanic (Kroonen 2012a, 2013 [his etymological dictionary], Iversen & Kroonen 2017). An argument had been growing that we should expect one or more of the substrate languages of Europe to have been spoken by pre-Indo-European farming populations. Since at least Schrader (1883), the likelihood had been established that the speakers of Proto-Indo-European were pastoral nomads (a theory that would come into competition with and ultimately prevail over the idea that Proto-Indo-European spread *with* farming; see §6.2). V. Gordon Childe (1926) drew on Schrader’s work, finding it difficult to decide whether the PIE *Urheimat* was in Scandinavia or South Russia, but leaning towards the latter. In any case, he noted that the language of the original PIE-speakers would have been affected by contact. “Through migrations, intermingling with other races, commercial relations with alien civilizations and the autonomous local growth and specialization of arts and cults, many words may have been lost and replaced by others” (pg. 70). Marija Gimbutas built on Schrader as well, formulating the *kurgan hypothesis* (the direct predecessor of the *steppe hypothesis*) that saw Indo-European speakers expanding in waves from the Pontic-Caspian steppe (Gimbutas 1956), coming into conflict with and disrupting the previous inhabitants of agricultural “Old Europe”. These Old Europeans seem to have been culturally and perhaps linguistically relatively homogenous (based on the widespread shared symbolism and artistic representation, Gimbutas 1989; cf. also Shennan 2018: 105). Kallio (2003: 232-4) also explicitly noted that “the Northwest Indo-European speech area was already agricultural before the spread of Indo-European” and that the language of the farmers was unlikely to have been related to Indo-European (opposite to what e.g. Sherratt and Sherratt 1988 had proposed). Schrijver (2007: 21-2) noted that linguistic unity (at a family level) of the pre-Indo-European farmers could explain the pattern we see, in which words of non-IE origin and features like the *a*-prefix are found with a wide distribution amongst the daughter languages.

Kroonen (2012), encouraged by Kallio and Schrijver as well as the early ancient DNA results (see §6.2), argued in support of the *Agricultural Substrate Hypothesis*, providing evidence of its existence in Germanic. By the time of Iversen and Kroonen (2017), ancient genome analysis had confirmed that agriculture had indeed been brought to

Europe via migrations of people. Several studies had shown that farming populations have served as vectors for language spread (e.g. Bellwood & Renfrew 2002; Diamond & Bellwood 2003; Bellwood 2005). Thus, under the supervision of Guus Kroonen and Kristian Kristiansen, this thesis forms a part of the ERC EUROLITHIC Project, whose full title is “The Linguistic Roots of Europe’s Agricultural Transition”, aiming to identify the substrate lexicon of Latin (a large part of which likely were borrowed from the languages of the first European farmers).

1.2.2.3 An Excursion on Etruscan

Many have tried to show that Etruscan is Indo-European (cf. recently Steinbauer’s 1999 connection to the Anatolian branch), but the current *communis opinio* is that it is unrelated (cf. Wallace 2008: 1; refutations of Anatolian connections in Simon 2021: 227 fn. 2). As such, the close linguistic and cultural contact which we know from the historical record that Etruscan had with Latin represents an important source for non-IE material in the Latin lexicon. Additionally, if Etruscan is the remnant of the languages that were spoken on the Italian peninsula before the arrival of the Italic family, it could crucially be an (albeit quite poorly) attested example of one of the substrate languages. Unfortunately, the evidence is far too complex to allow an answer to any of these questions.

The question of Etruscan origins remains, despite the confidence asserted by Beekes (2003) and Kloekhorst (2022), undecided. Deciding one way or the other requires believing either Herodotus’ account that they came from Anatolia or the opposite report of an autochthonous Italian origin in Dionysius of Halicarnassus. The attestation of Lemnian on the island of Lemnos (two inscriptions, one the famous Lemnos Stele) in a slightly different alphabet than that used in Etruria shows that it was present there before the introduction of the alphabet. But whether it is autochthonous to Lemnos or represents an Etruscan colony cannot be determined. Two facts tip the balance in favor of an autochthonous origin, in my opinion. Firstly, the only other attested relative of Etruscan and Lemnian is Rhaetic. By the accounts of Oettinger (2010) and Marchesini (2013), Etruscan and Lemnian are more closely related to each other than either is to Rhaetic, suggesting Rhaetic split from the family first and that the split likely occurred in Italy. Secondly, a recent genetic study by Posth et al. (2021) found that by 800 BCE, populations in the area were genetically homogenous. That is, they found no difference between remains from Etruscan contexts and those from Latin contexts. If the Etruscans had come from elsewhere, then they did so long enough ago that they no longer retained a unique ancestry signature.

If the Etruscans were autochthonous, then their language is likely related to one or more of the European substrate languages. But if they are later intruders, then their linguistic influence on Latin is not as relevant to the research questions in this study. Unfortunately, it is also notoriously difficult to find solid proof of Etruscan lexical influence on Latin.

The main hindrance to understanding Etruscan lexical influence on Latin is its poor attestation. Despite thousands of texts, there are only around nine of any length (cf. Weiss 2020: 519). Our understanding of the grammatical structure is quite good, whereas our understanding of the lexicon is much more limited. Ernout (1930)³³ began to investigate the influence of Etruscan on Latin and his dictionary with Meillet (EM) along with that of Walde and Hofmann (WH) cautiously suspect many Latin lexemes to be of Etruscan origin. Often it is done even in the absence of an attested Etruscan source option. Given the lack of material, one can sympathize with this practice; this makes it no less reckless. More modern works base their proposals of Etruscan origin more concertedly on phonological and morphological considerations, yet the practice remains. Of the ca. 550 words that Breyer (1993) discusses (meaning that they have been proposed as loans from Etruscan), for only 123 is an Etruscan word attested (almost always of unknown meaning) that looks similar. Of the 198 words for which she accepts that Etruscan origin is either “sehr wahrscheinlich oder sicher” or “möglich, jedenfalls nicht auszuschließen”, a potential Etruscan attested form exists for only 101.

The phonological and morphological criteria used to identify Etruscan loans in Latin vary in quality (cf. a ranked list in Breyer 1993: 11-135). Some of the most frequently given include:

Phonological

o ~ *u* vacillation

e ~ *i* vacillation

media ~ *tenuis* vacillation

tenuis ~ *tenuis aspirate* ~ spirant alternation in the environment of liquids
or nasals

Morphological

-*rna* (-*rno*-)

-*enna* / -*ennus*

Etruscan had a four-vowel system in which the vowel they represented graphically with *u* is sometimes transcribed in Latin with *o* (Wallace 2008: 32-3, Weiss 2020: 523). Thus its pronunciation may well have been intermediate to Latin *o* and *u* and resulted in vacillation within Latin due to the borrowing process. On the other hand, Etruscan had both *e* and *i*. Their interchange (cf. Pfiffig 1969: 29-34) has been taken to be the culprit behind *e* ~ *i* alternations in Latin as well as between Latin and e.g. Greek (cf. Bertoldi 1939b: 89). While the “alternation” in Etruscan is likely explained by an umlaut of the first syllable and otherwise vowel weakening (Wallace 2008: 34-5), Latin could have borrowed from Etruscan before and after the changes occurred, creating a vacillation in Latin due to pre-existing formal diversity.

³³ The text of the journal to which I have access has amended the printed year 1929 to 1930. In the bibliography, I have listed it under its re-print as Ernout (1946).

The question of Etruscan consonant voicing is actually perhaps too complex to be useful. Outside of early abecedary inscriptions, the Etruscan alphabet did not use the graphemes for voiced stops. As late as Bonfante (1985: 203), it has been claimed that Etruscan had no voiced stops and that no Latin words with voiced stops can have originated in Etruscan. Indeed, Etr. *Catmite*, borrowed into Latin as *Catamītus*, is none other than Greek Γανυμήδης “Ganymede” with devoicing (though note it suggests a variant *Γαδυμήδης, cf. de Simone 1968-70 II: 189-90). On the other hand, Latin spellings of Etruscan names include *Tidi* for *Titi*, *Pergomnsa* for *Percumsna* and *Pabassa* for *Papaśa* (Rix 1985: 220). Varro writes that Lat. *sūbulō* ‘flute-player’ is of Etruscan origin, which Watmough (1997: 53-68) defends. Thus Latin speakers at least sometimes perceived Etruscan intervocalic stops as voiced. Whether this makes it in fact more likely that such alternations within Latin could be due to Etruscan or whether it is too poorly understood to be diagnostic, I leave open. The question of *tenuis* ~ *tenuis aspirata* ~ spirant alternation is likewise complex, but seems to have been exaggerated due to misunderstanding. There does in fact seem to have been an alternation between the *tenuis* and *tenuis aspiratae*, at least in contact with liquids and nasals (Pfiffig 1969: 38-41). But the alternation of Etr. *p* and Etr. *f* (cf. already Terracini 1929: 230-48) is attested for only some six lexemes. Pfiffig (1969: 42) asserts that it is late and regional, with Steinbauer (1999: 59) saying specifically North Etruscan.³⁴ Thus it seems unlikely that Etruscan caused this change in Latin (cf. Thorsø & Wigman et al. 2023 on *ferrum*).

Niedermann (1916: 152) noted that several Etruscan names attest to the sequence *-rn-* (e.g. *Macstrna*, *Perperna*, *Steprna*, *Θucerna*, and *Velθurna*), and that several Latin words with this sequence (e.g. *alaternus* ‘buckthorn’, *laburnum* ‘broom plant (Cystius)’, *santerna* ‘borax from gold smelting’, *vīburnum* ‘arrowwood’) are of unclear etymology. Despite this suffix appearing on roots of clear IE origin, he, followed by Ernout (1946: 29-35), suggested the suffix might be of Etruscan origin. Several Latin words have a native *-rn-* suffix such as *hibernus* (and other temporal adjectives) < **ǵʰejm-r-ino-* with zero-grade of **-er* as in Ved. *uṣar-būdḥ-* ‘waking at dawn’ and *hesternus* ‘of yesterday’ built on the stem of *heri* ‘yesterday’ with zero-grade of the oppositional suffix **-tero-* and the **-ino-* suffix (Weiss 2020: 311).³⁵ The sequence **-esino* would also yield *-ernus* (Michael Weiss, p.c.). Besides e.g. *taberna* ‘hut, inn’ < *trabs* ‘beam’ and *caverna* ‘a hollow’ < *cavus* ‘hollow’, Latin has *cisterna* ‘cistern’ alongside *cista* ‘box’ < Gk. κίστη ‘basket, chest’. Could the suffix have been added in Etruscan? Latin *lucerna* ‘oil lamp’ looks like it must be from the inherited root **leuk-*, but Latin has only reflexes of the full-grade in *lūx* ‘light’ and *lūceō* ‘to shine’. The short *ŭ* of *lucerna* is suspiciously unexplained (Alessio 1944a: 144 fn. 222, DV 355). Perhaps it went through Etruscan, which may not originally have vowel length (Wallace 2008:

³⁴ Additionally, it seems that the spellings with *f* are older; thus rather than *p > f*, perhaps through *ϕ*, we have in fact only evidence for the opposite direction *f > p* (Watmough 1997: 99, Steinbauer 1999: 59-60).

³⁵ Weiss (2020: 311) suggests that this suffix also occurs in PGm. **gestra-* ‘the previous or next day’; but Kroonen (2013: 176) notes that the **r* can simply be the result of the Germanic sound law **sr > *str* and thus PGm. **gestra-* would only attest to a suffix **-ro-*.

33-4, Weiss 2020: 523 on the vowel length). A crucial form is Lat. *santerna* ‘borax from gold smelting’ which seems likely to be related to Etruscan gold words *zamaθi* ‘gold’, *zamaθic* ‘golden’ (cf. Trombetti 1928: 128, further Breyer 1993: 274 fn. 307 on the Etruscan meanings). But then it is the Etruscan forms that specifically do not occur with the *-m-* suffix. Does this imply some other language was the source of the *-m-* in both languages?³⁶ In fact, one case proves that Etruscan has added an *-n-* suffix to a word that already ended in *-r-*: Etr. *Macstrna* is simply a borrowing of Latin *magister* with the (perhaps most³⁷) productive Etruscan suffix *-na*. Perhaps similarly, Lat. *lanterna* has the same meaning as Gk. λαμπτήρ ‘torch, lantern’ in which *-τήρ* is an inherited agentive suffix added to λάμπω ‘to shine’. Thus the status of *-m-* as an Etruscan suffix cannot be decided, but the *n-* suffix will importantly feature in later discussion.

Similarly based on its appearance in Etruscan names (transmitted in Latin) like *Perpenna*, *Sisenna*, *Spurinna*, *Vibenna*, and *Porsenna*, the suffixes *-enna* and *-inna* are considered to be of Etruscan (or Etruscicizing) origin (Ernout 1946: 27-9, Leumann 1977: 321, Weiss 2020: 514, etc.). They too occur on Latin bases like *sociennus* ‘partner’ < *socius* ‘ally’ and *Dossennus* ‘hunchback’ < *dossum* < *dorsum* ‘back’. Unlike *-erno-*, the geminate *-m-* is very unlikely to have a native source in Latin. However the proof that this suffix is specifically Etruscan feels just as underwhelming as for the *-m-* suffix above.

Thus the true nature of the extent of Etruscan influence on Latin and vice versa is plagued by difficulty. A new approach to Etruscan research might instead involve finding evidence to support or reject its status as one of the pre-IE substrate languages of Europe. This means, however, that Etruscan itself will play little role in the forthcoming analysis. In the end, my analysis here may be able to say more about Etruscan than the Etruscan language itself can benefit my analysis.

1.3 The Consequences: Goals of this Dissertation

The research on the pre-Indo-European substrate languages of Europe has come a long way since the idea of substrate influence was given practical attention in the 1880s. The methodology for identifying words as being of substrate origin has also been sharply refined. As concerns Latin, there is an immense corpus of work on which lexemes might be of substrate origin. Given the shortcomings of some of this prior work (see the concerns of uniqueness and exclusivity above) and its existence in several disparate places, a current survey of the substrate lexicon of Latin is greatly needed.

³⁶ Cf. Battisti (1929) who suggested that the suffix is too widespread in Mediterranean toponymy to have been originally just Etruscan. But furthermore, it is present in e.g. *palacurna* ‘gold ingot’ given by Pliny along with *palaga* of the same meaning and almost certainly related *balūx*, *bal(l)ūca* ‘gold sand’ as gold mining terminology from Iberia (WH I: 95, 851, II: 237; EM 65, 475). Ernout (1946: 32) wondered if this meant that the language of Iberia was related to Etruscan.

³⁷ Steinbauer 1999: 121

Therefore, the goals of the present work are to:

1. Use the most current methodology to collect the evidence for loanwords from unknown sources in Latin.
2. Determine the geographic extent of the attestations of these words and attempt to chronologize their appearance in Latin.
3. Use this corpus to understand what sorts of substrate languages were in Europe.
4. In combination with archaeology and genetics, use the linguistic data to postulate how and perhaps when Italic speakers (or their ancestors) entered the Italian peninsula.

1.4 Methodology

1.4.1 Theoretical Perspective

This work operates with the assumption that the steppe origin of the (core) Indo-European languages is correct. Thus, Latin has its origins with speakers of Proto-Indo-European who migrated over an as of yet unknown period of time, taking an as yet unknown route, to end up on the Italian peninsula. The *terminus post quem* is the ca. 3000 BCE initiation of the Yamnaya migrations from the Pontic Steppe while the literal *terminus ante quem* is the attestation of Very Old Latin in around the 7th century BCE. Ancient DNA analysis has dated the first populations with steppe-derived ancestry in Northern Italy to ca. 2000 BCE and in Central Italy to ca. 1600 BCE (Saupe et al. 2021). Given the proposed origins of Indo-European speakers, 2000 BCE therefore seems like a good *terminus post quem* for the arrival of Indo-European languages in Italy. Whether or not this was the Italic branch remains to be seen.

However, this means that there was a period of ca. 1000 years between the disintegration of Proto-Indo-European unity and the arrival of potentially Indo-European speaking populations in Italy during which contact would have occurred with the previous inhabitants of Europe (cf. already Hirt 1894: 38). Then there was a period of ca. 1400 years between the appearance of potentially Indo-European speaking populations in Italy and the written attestation of the Italic languages during which 1) the Italic languages must have arrived and 2) contact would have occurred with the previous inhabitants of the peninsula. Two major changes that speakers of Proto-Indo-European would have been exposed to on their way to Italy and which would likely have had a linguistic impact are 1) the introduction to and adoption of agricultural practices (and consequently the familiarity with vocabulary for new practices, tools, animals, and plants) and 2) the entry into the Mediterranean region (and consequently the familiarity with vocabulary for new plants, animals, etc.). These scenarios suggest several layers of lexical replacement would have occurred.

1.4.2 How (not) to do it

As mentioned above, Schrijver (1997: 294-6) discusses a list of five criteria that have been used to demonstrate that a word is of non-IE origin. He prefaces the discussion with the consideration that no one criterion alone is sufficient; only cumulatively can they be convincing. Indeed, objections can be raised to each, but I find the objections to some of them to be strong enough to warrant excluding them from the process. The criteria again are:

1. Limited geographic distribution
2. Phonological or morphological irregularity
3. Remarkable word formation
4. Non-IE phonology
5. Semantic category

Criteria 2-4 are valid according to the historical comparative method, whereas 1 and 5 are not actually based in valid *a priori* assumptions.

1.4.2.1 Reservations about Semantic Category

Schrijver (1997: 295) notes that words denoting economically and culturally unimportant animals and plants are possibly more likely to have a substrate origin. But he immediately expresses doubt about the robustness of this criterion due e.g. the ability to reconstruct **k̑rHs-* ‘hornet’ for PIE. I think the premise is faulty. Even despite the fact that, as mentioned above, many of the entities that the people participating in these migrations came across would have been unknown to them before, we must work with the caveat that the **absence** of an inherited *signifiant* does not necessarily mean the absence of its *signifié*. Nor, conversely, does the **presence** of an inherited *signifiant* always imply that it refers to its original *signifié*. In fact, there are generally four main input-output situations one can expect, which I would like to describe by use of an example from closer to (my) home. When English-speaking Europeans travelled to and began to live on the East Coast of North America, they came into contact with both new and familiar flora and fauna. They also had complex interactions with the indigenous peoples of the Americas and their languages.

The racoon, *Procyon lotor*, at the time did not exist outside of North America. It was a new animal to the Europeans, and they correspondingly adopted the Powhattan word transcribed at the time as *aroughcun* to refer to it.³⁸ The red fox, *Vulpes vulpes*, however,

³⁸ Other transcriptions of the Powhattan included *aroughcoune*, *rahaugcum*, *rarrowcun*, *raugroughcum*, *arathkone*, resulting in numerous early American English spelling variations (taken from the OED): with reduplication *rahaugcum*, *rahaugcum*, *rarrowcun*, *raugroughcum*; with an initial vowel *arocoun*, *aroughcoune*, *aroughcun*, *arroughcan*; and without an initial vowel *raccoone*, *rackcame*, *rackoone*, *racoone*, *racoune*, *racowne*, *rockoone*, *rokoone*, *rackoon*, *raccoon*, *raccoon*.

due to its native distribution across the Northern hemisphere, was already well known to them. Thus, rather than adopting a Powhattan word (such as that preserved in transcription as *onxe*), they kept the name ‘fox’. These are the two situations that would be the most informative to the question of the origins and earliest contacts of the Italic-speakers. The words for new things are borrowed from already-present languages and the words for familiar things stay the same.

This is not always the case, however. The exact opposite of each of these processes can also occur. The American robin, *Turdus migratorius*, was a previously unknown bird to the arriving Europeans as it is exclusively native to North America.³⁹ Rather than adopt an indigenous word for it,⁴⁰ its red-breasted similarity to the European robin (*Erithacus rubecula*) and perhaps its lesser economic significance than the racoon, led to the re-purposing of the English word ‘robin’ to refer to it. The robust cervid *Alces alces* was and is called ‘elk’ in Europe. Incoming English-speakers, however, adopted (an) Algonquian word(s) like Narragansett *moos* and Eastern Abenaki *mos* and it is now known in American English as the ‘moose’.⁴¹

These latter two examples did not occur without motivation (close similarity to a familiar animal in the first case and lack of familiarity with the “native” word in the latter case), but is the motivation strong enough that we would be able to predict its occurrence if we did not have the knowledge of hindsight? To be able to explain examples like this in prehistory furthermore already assumes the success of the primary task, that is being able to separate the racoons and foxes from the robins and moose. These examples suggest that semantics are not a robust way to do this. It is instead the morphophonotactics of the words that betray their non-native origin. A similar situation has been discovered for the names of cereals in Indo-European (Kroonen et al. 2022): many cereal terms are named with repurposed inherited material, rather than being foreign. Thus, semantic considerations are best kept for *after* non-inherited material has been identified, not used to identify it.⁴²

³⁹ No more than a few per year ever end up in the British Isles as a result of migration mishaps.

⁴⁰ Perhaps an Iroquoian word like Oneida *tsiskóko* or an Algonquian word like Ojibwe *alopichi*

⁴¹ The reason for this is likely that *Alces alces* was basically a new species for settlers coming from the British Isles. *Alces alces* was likely extirpated from the British Isles by 900 CE, and so English-speakers had actually come to use the word “elk” to describe very large deer. The *Alces alces* that they saw in America was nothing like the deer they were used to (red deer *Cervus elaphus*, roe deer *Capreolus capreolus*, and fallow deer *Dama dama*), unlike the *Cervus canadensis* that looked like a very large version of their familiar red deer. Thus it was the uniquely American *Cervus canadensis* that received the European name “elk” (although it is also known as the *wapiti* from the Algonquian language Shawnee, since even the European red deer *Cervus elaphus* to which it was closely related was already quite rare in the South of the British Isles) and it was the species which also existed in Europe—albeit basically unbeknownst to British settlers—whose name was replaced with an indigenous word.

⁴² Cf. Iversen & Kroonen (2017: 517) who use phonological irregularities to identify non-inherited words and can then conclude “One of the striking features of the lexical layer surviving from this landscape is that it contains a cluster of words belonging to the sphere of agriculture... On the basis of this clustering, it seems safe to claim that the language that donated the words for crops... was spoken by a culture whose subsistence strategy was more pervasively agricultural than that of the predominantly pastoral

1.4.2.2 Doubts about Limited Geographic Distribution

Schrijver (1997: 294) immediately acknowledges that this criterion by itself is not a valid piece of evidence, giving one example (Lat. *porrum*, Gk. *πράσον* ‘leek’ < **p̥r̥som*⁴³) where an IE-looking proto-form can be reconstructed and one (**mori* ‘sea’) that was thought to be limited to North and West Europe until an Iranian cognate was found in Ossetic. Even without cases like the second example, the first consideration is extremely important. When working within the historical comparative method, isolated items are understood to be able to represent either innovations or chance preservations. Without a compelling morphophonological reason, it cannot be ruled out that an isolated or geographically restricted lexeme is simply the sole survivor of an inherited construction. For this reason, in the data, items that have been suspected to be of substrate origin due only to their restriction to only a few daughter language are placed in the category §2.4.1 *No Positive Evidence of Borrowing*.

From the other side of things, a widespread distribution does not guarantee inherited status. As Beekes (1996: 215-16) notes, in the face of morphophonological irregularities, we cannot *a priori* exclude items for which there are matches in the Western languages as well as Greek and e.g. Sanskrit. These could represent early loans that were taken up just after the proto-language disintegrated and the daughter branches were still in close proximity.

1.4.2.3 The Requirement of Positive Evidence

Thus, the most secure indication that a word is of non-IE origin in Latin is the existence of positive morphophonological evidence. Very rarely can positive evidence exist in words isolated to Latin and without attested variants. But when (morpho)phonological sequences appear that are not reconstructible to PIE, it is good evidence of non-native origin. Two such instances appear in the data: cases that must reconstruct to an invalid PIE root structure⁴⁴ and cases of unrhotalized intervocalic *s* after a short vowel.

Otherwise, positive evidence can only be the result of comparison. Loanwords were taken up into the daughter languages after the dissolution of the parent language. They entered in slightly different forms due to pre-existing dialectal differentiation within the substrate languages and/or due to the borrowing process itself. Thus, while reconstruction to the proto-stage of the daughter languages should result in forms that are closer to each other, further reconstruction to PIE should make them look less similar

Indo-Europeans.”

⁴³ See the entry on this word for potential problems with reconstructing it this way.

⁴⁴ The root structure constraints relevant to the data set in §2 are those against **DeD* and **TeD^h* roots (e.g. Szemerényi 1990: 103, Beekes 2011: 171; on **TeD^h* roots see also Meillet 1912: 60). Even **TeRD^h* roots (excluding if the resonant is a nasal) were probably not permitted (de Vaan 1999). Nor were **C_ieC_i* roots (that is, roots beginning and ending with the same consonant, Szemerényi 1990: 103) or roots with initial **r* (Beekes 2011: 171).

again. It results in the reconstruction of irregular alternations, as if these words originated from different roots. Note that a Latin word can have remarkable phonotactics and still not be isolated to Latin. Cases like these are particularly strong. Likewise, however, other daughter languages can attest to forms that are not reconstructible in their branch despite being compared to a Latin word that is easily reconstructible. Even though the phonotactic problems are not present in Latin, they nonetheless occur in the family of the lexeme's attestation.

1.4.3 Terminology

1.4.3.1 Comparanda

The methodology for identifying words of non-IE origin as described above relies on locating irregular correspondences that appear upon comparison with forms in other daughter languages. I have decided to avoid using the word "cognate" for these groups and instead reserve that term for lexical items that have been inherited and developed regularly from the proto-language. Instead, the lexical items investigated in this thesis, since they must reconstruct to different proto-forms, will not look like cognates. But because of their morphophonological and semantic closeness, it is clear that they originate from the same non-native lexeme. They must still be compared. Thus I will employ the Latin gerundive *comparandum* to describe these quasi-cognates.

1.4.3.2 Substrate

Thomason and Kaufman (1988) in their seminal work *Language Contact, Creolization, and Genetic Linguistics* define **substratum interference** as:

a subtype of interference that results from imperfect group learning during a process of language shift. That is, in this kind of interference a group of speakers shifting to a target language fails to learn the target language (TL) perfectly. The errors made by members of the shifting group in speaking the TL then spread to the TL as a whole when they are imitated by original speakers of that language (pp. 38-9).

They further note that in cases like this, the target language often ends up borrowing few words from the shifting speakers' language. Because it requires a population that is shifting to a new language, its results are different from **borrowing**, "the incorporation of foreign features into a group's native language by speakers of that language: the native language is maintained but is changed by the addition of the incorporated features" (Thomason & Kaufmann 1988: 37).

In the case of Latin, both processes have likely occurred. Upon contact with the pre-Indo-European populations of Europe, speakers of what would eventually become Latin had both opportunity and cause to borrow words (migration to a new environment, shifting subsistence strategies upon contact with agricultural populations). But, at least upon reaching the Italian peninsula, the populations inhabiting the places where they

settled all eventually underwent a language shift to speak Italic languages. It is interesting to keep these inter-connected factors in mind.

Because the effects of both borrowing and language shift contributed to the non-IE lexicon of Latin, and because of its use in previous literature, I will focus on a more basic meaning of the word “substrate”; one that again connects to its Latin roots. *Substrātum* as the perfect passive participle of *substernō* means that which has been spread beneath. Used here, a “substrate” refers to an unknown language with which Latin speakers and their ancestors came into contact. This includes languages that might be classified as having served as adstrates or superstrates if we knew more about the intensity of contact and power relationships. I am generally of the opinion that the evidence points to the substrate languages being non-IE.

1.4.4 A Mild Disclaimer

It is important to note from the beginning of a study like this that there will always be a degree of subjectivity in the analysis and categorization of the data. Despite the shorthand description of ‘substrate’ that I assign the accepted cases and the confidence with which I include all data in the analyses, discussions, and conclusions, I recognize that other scholars with other theoretical perspectives may disagree with some of my analyses. I have discovered that, in order for an analytical approach to be falsifiable, the assumptions behind that approach must be made explicit.

The basic tenet of my approach is to identify the words in the Latin lexicon that can be shown to demonstrate problems in reconstruction, and that therefore are unlikely to have been directly inherited from the Proto-Indo-European parent language. Some of the cases are more widely accepted than others. For many, however, attempts can and have been made to assign the lexemes an Indo-European origin. My approach to these attempts is that, unless the irregularities can be demonstrated to participate in a systematic process, I remain suspicious of inherited origin as an explanation. Future developments in the field will certainly, as they have done in many cases in the past, uncover more systematic patterns. In the meantime, I am generally not persuaded by explanations that can be described as *ad hoc* in the face of more frequently occurring irregular alternations. Included in this category are unparalleled root extensions, unparalleled assimilation/dissimilation, unmotivated analogy, explanations proposing onomatopoeic origin, and explanations that involve several reconstructed (that is unattested) stages of derivation.

I am also opposed to assigning a word an Indo-European etymology if it requires that word to be derived from an inherited root with an only tenuous semantic link while at the same time removing it from comparison with an irregularly corresponding similar form that is a much better semantic match. For instance, Gk. βᾶσκανος ‘slanderer, sorcerer’ could theoretically be derived, like βᾶσκω ‘to come; go’ from *g^wem- ‘to go’ under the assumption that it underwent some sort of semantic change like ‘to go (towards)’ > ‘to

accost, attack’ > ‘to harm with magic or speech’, but it would require removing it from comparison with Lat. *fascinum* ‘witchcraft, phallus charm’ whose form can be reconstructed as identical but for the aspiration required (**b^haskano-*) and whose meaning is quite patently more similar.

It is my hope that, while certain lexemes will receive alternative interpretations based on differences in approach, the majority of my data holds up to scrutiny, and therefore that the results of my analysis remain viable.

1.5 Limitations

1.5.1 Scope

Much of the work of the early substratists was deeply informed by their training as Romance linguists. They located numerous substrate lexemes attested only in the Romance languages. This thesis restricts itself to words that have an attestation in Classical (urban) Latin⁴⁵ with the hope that 1) it limits the scope of the material but especially 2) it says the most about the earliest layers of Latin-substrate contact. The Romance languages are like a net that has been thrown over Europe and has caught substrate lexemes. But the net was thrown from Rome and therefore necessarily attests to later contact phenomena.

1.5.2 Methodological Blind Spots

As Alessio (1944a: 94) writes with suitably metaphoric language, “è anche vero...che la

⁴⁵ There are several words that Roman authors mention, but which are clearly foreign and which likely did not enter Latin parlance. While it seems counterintuitive to leave out what are basically sure-fire vestiges of lost languages, they do not comprise the substrate which entered Latin and therefore represent a different topic. For methodological reasons, I leave them out of the corpus. Pliny mentions several words in use in Iberian gold mines: *arrugia*, *corrugus* ‘shaft and pit in a gold mine’, *bāllux* ‘gold dust’, *palacuma*, *palaga* ‘gold ingot’, etc. But he gives these as foreign words, and it is clear they were never in use in Latin. *Camōx* ‘some quadruped’ occurs once in all of Latinity, in the calendar by the 5th century Polemius Silvius of Gaul. It has Alpine Romance descendants (cf. Bertoldi 1937a: 147, Hubschmid 1963-5 II: 94) but does not seem to have ever entered Latin. *Mufrō* and *mus(i)mō* ‘wild sheep, almost certainly the mouflon’, while fragmentarily mentioned in Cato, Lucilius, and Varro, is described as living in Corsica by Pliny (he also mentions Iberia) and Strabo. Amongst the Romance languages, most forms seem to have been borrowed from Corsican (REW no. 5715), suggesting that it may not have entered spoken Latin. *Capys* ‘bird of prey, probably falcon’ appears only in authors and glosses who attribute it to Etruscan and is not continued in Romance. (Isidore claims it is a native Italic word, but his spelling *capus* shows he has confused it with *capus* ‘capon’.) Despite its interesting implications in combination with PGM. **habuka-*, PSlav. **kobuzъ-*, and Arm. *k’owpič* ‘male peregrine falcon or hawk’ (cf. Thorsø fthc.), it does not seem to have actually entered Latin. The words *salmō* ‘salmon’ and *salar* ‘trout’ are considered by Diebold (1985: 59) and Huld (1990: 397) to be loans from a substrate. But salmon do not live in the Mediterranean Sea, and Pliny (*Nat.Hist.* 9.32) writes that the *salmō* is from Aquitania. Furthermore, *salar* appears in Ausonius’ (from Bordeaux) poem about the Moselle and in a letter by Sidonius Apollinaris (of Lyon) about his summer home. Thus they are likely loans from Celtic (cf. Holder 1896-1904 II: 1299, WH II: 467). The potential relationship to *salpa* ‘a kind of fish’ (cf. Alessio 1946a: 155) is interesting, but *salpa* itself is likely a loan from Gk. *σάλπη/σάρπη* ‘sea bream’. Thus there is no independent Latin evidence for this group.

buona riuscita della cernita dipende dalla bontà del vaglio e dall'abilità dell'operatore, e che attraverso una maglia rotta o addirittura balzando al disopra del cerchio,...alcuni chicchi di grano possono passare nel loglio o alcuni granelli di loglio sfuggire e restare nel vaglio."⁴⁶ Because the specification of positive evidence has above been given as the occurrence of irregular correspondences or remarkable phonotactics in the attested comparanda, some substrate words will be missed. If a particular phoneme of one of the substrate languages was similar to one existing in PIE, there is a chance that the daughter languages would borrow it in a way that would be too regular to be caught by this methodology.

Likewise there are bound to be false positives. For instance, there is always the chance that what has been identified as an irregularity will be understood as regular in the future. More insidious is possibility of chance resemblance. Schrijver (1997: 296) cautions:

When historical phonology no longer counts for much, the danger of producing false etymologies looms large. Any two words that lack a reliable IE etymology and resemble one another vaguely can now be lumped together under the label 'substratum', and we are deprived of our most important tool to check linguistic affiliation and perform linguistic reconstruction: sound laws.

Within the inherited corpus, it is the *Ausnahmslosigkeit der Lautgesetze* that allows us to rule out chance resemblances between lexemes. "A phonological 'near miss' is as good as useless" (Jakob fthc.). But substrate research operates outside of these laws. Schrijver (1997) solves this problem by exploring cases in which the same irregular (from an inherited perspective) correspondence recurs with the same distribution. This is similar to the methodology used in studies of loanwords between known languages. But as will be shown (§0), the correspondence patterns in the Latin words of loanword origin can be diverse. This is not unexpected given that the corpus presented by attested Latin represents all contact situations with 1) an unknown number of languages 2) throughout the prehistory of Latin from the dissolution of PIE until its attestation (some of the same arguments used to defend cases that did not match expected Pelasgian sound laws, cf. §1.2.1.3). But it means that quasi-*lautgesetzlich* approaches like that of Schrijver are unfeasible. I have only accepted for analysis cases where I believe the semantic match between forms is good. But semantic closeness can be subjective, and the risk of chance resemblance exists between even otherwise identical words. False positives and negatives are an unavoidable problem in any scientific investigation; we can only do our best to design a methodology that maximizes specificity and sensitivity.

The constraints of time and space have also meant that, amongst the list that would already include both false positives and false negatives, I have chosen to leave out some

⁴⁶ "It is also true that the success of the sorting depends on the goodness of the sieve and the ability of the operator, and that through a broken mesh or even a jumping out over the hoop, some grains of wheat can pass into the chaff or some bits of chaff escape and remain in the sieve."

suggested examples that I did not think would lead to productive discussion.⁴⁷ Despite the limitations mentioned here, I am confident that the material discussed is a thoroughly representative sample.

⁴⁷ This includes for example *bruscus* ‘type of frog’ (and *bruscum* ‘excrescence on a maple’, *brūscus* ‘butcher’s broom’, *rūscum* ‘butcher’s broom’, *ruscus* ‘toad’, Alessio 1944a: 119-22), *lāma* ‘bog’ (Alessio 1944a: 134-7), *lappa* ‘burdock’ (Alessio 1941b: 218-20), *lolium* ‘darnel’ (Alessio 1944a: 137), *parra* ‘unlucky bird’ (Matasović 2009: 334), *sturnus* ‘starling’ (cf. Matasović 2020: 340-1), *tālus* in the meaning ‘dice’ (Alessio 1944a: 147), *taxus* ‘yew’ (potentially IE, but the explanations feel like folk etymologies, cf. Alessio 1957), *tōfus* ‘volcanic rock, tufa’ (Fiesel & Groth 1932, Battisti 1943: 265, Alessio 1944a: 154-5).

2 The Linguistic Data

2.1 Introduction to the Data

This chapter presents discussions of the Latin lexical material that has been proposed to be of substrate origin in order to accept the cases that meet the criteria as discussed in §1.4 to use in further analysis, and to exclude cases that are inconclusive, methodologically unprovable, or inherited. The material is not exhaustive; the uncertain cases especially are limited to those which might prove valuable despite their inconclusive status. Some cases show that a frequent claim is based on precious little evidence. Others attest to the limits of the objective methodology, where too many subjective decisions are required to arrive at a conclusion. Together they form a representative sample.

2.1.1 Structure of the Data

The data have been categorized into three main groups, each with sub-divisions:

§2.2 *Non-inherited Origin in Latin Accepted*

These words show evidence of being borrowed by Latin from an unknown source.

§2.2.1 *Phonotactic Reasons*

§2.2.1.1 *Isolated to Latin but with Unrhotacized S*

An intervocalic *s* that remains in Classical Latin points to a loanword post-dating the fourth century BCE; the source of the loanword is unidentified.

§2.2.1.2 *Isolated to Latin but with an Invalid Root Structure*

Words in this category have no secure comparanda, but they cannot be reconstructed to a valid PIE root structure, pointing to non-IE origin.

§2.2.2 *Comparanda in Other Branches*

§2.2.2.1 *Non-inherited Origin is Probable*

The case for non-native origin in Latin based on irregular correspondences between comparanda is methodologically strongest in this group.

§2.2.2.2 *Non-inherited Origin is Possible*

A case for non-native origin can be made for this group based on irregular correspondences between comparanda, but some of the details are less certain (e.g. the security of the comparanda or whether irregularities can be explained by analogy).

§2.2.3 *Comparanda only in Latin and Romance*

A recent, non-native origin in Latin is likely for this group given the irregularities between attested Latin variants or between variants that can be reconstructed on the basis of Romance data.

§2.3 *Origin Unclear*

§2.3.1 *No Comparanda*

All proposed comparanda for these words can be ruled out.

§2.3.2 *Uncertain Comparanda*

Comparanda have been suggested for these words, but it is either difficult to determine how many are truly related or the evidentiary value of those that belong is compromised (e.g. they are onomatopoeic, semantically dubious, or perhaps themselves loans from/into Latin).

§2.3.3 *Conflicting Possibilities*

§2.3.3.1 *Non-inherited vs. Inherited*

For these words, choosing between existing interpretations (and therefore accepting or rejecting non-inherited origin) is too difficult.

§2.3.3.2 *Non-inherited vs. Loan from a Known Language*

For these words, it is too difficult to decide whether they represent independent evidence of a non-native lexeme or if they are borrowed from a known language (whether or not they are native to that language).

§2.3.4 *Core-Periphery Cases*

For the majority of cases of suspected non-native origin, there is irregularity between all of the comparanda. For a few cases however, a common pre-form can be reconstructed for several branches against a few branches that, if compared, require the reconstruction of irregular correspondences. It cannot be ruled out that some cases of non-IE words were of such phonology that they were coincidentally borrowed the same way into most languages. But it seems suspicious to treat these cases the same way as the more numerous others in which the irregularity is more ubiquitous. This is because it also cannot be ruled out that an inherited lexeme has been borrowed by an Indo-European language through some sort of indirect mediation.

§2.3.5 *Methodologically Difficult to Delimit Comparanda*

In several cases, the decision to exclude comparanda becomes particularly subjective. The semantics are a good fit and the irregular alternations between individual comparanda are paralleled in other more secure cases. But the end result is a very widespread distribution of very divergent forms that has only a very small chance of

representing a true substrate lexeme and a much larger chance of being the result of coincidental resemblance.

§2.4 *Non-IE Origin in Latin Rejected*

§2.4.1 *No Positive Evidence of Borrowing*

These words have been suggested to be of non-inherited origin but there are no positive formal criteria to make these claims beyond geographic distribution and semantic field. The comparanda can reconstruct to the same valid PIE root structures, despite the roots being otherwise unattested.

§2.4.2 *Best Explained as Inherited*

These words have been suggested to be of non-inherited origin but there are no positive formal criteria to make these claims and they can instead reconstruct to known PIE roots.

§2.4.3 *Loan from a Known Language*

It cannot be ruled out that these words were borrowed from a known language (or proto-form thereof), regardless of the deeper origin of the etymon in that language.

2.1.2 Structure of the Entries

The beginning of each entry lists the Latin word and its meaning(s). This is followed in the next line by a reconstructed pre-form, then by a section of comparanda and their reconstructions. At the right is the attested material or at least a representative sample of the attested material. Next to the left is the proto-form that can be reconstructed from the attested material on the branch level. Finally, to the far left is a quasi-Indo-European reconstruction that comes closest to being able to unite the intermediate proto-forms. In these reconstructions, there are a few features of notation that require explanation. Firstly, reconstructed **k*, **g*, and **g^h* for *centum* languages do not necessarily rule out the possibility that these sounds were borrowed as palatovelars (see §3.2.1.1.2.4), just as for *satəm* languages they do not rule out original labiovelars. I only explicitly reconstruct palatovelars for *centum* pre-forms in cases where the lexemes are clearly inherited (§2.4.2). Secondly, there are only a few cases where original quasi-PIE *a*-vocalism is the only reconstructible option for a form in a daughter branch (§3.2.2.2.1). Otherwise, the range of possibilities is reconstructed (e.g. **a/H* for Italic), including **a*, which can be interpreted as a shorthand for **h₂e* where possible. While **H* represents a laryngeal of unknown quality, other capital letters in the reconstructions do *not* represent cover symbols but rather elements that are not reconstructible (such as Latin intervocalic *s*). In the case that multiple phonemes are reconstructible, these are listed. Less certain comparanda are preceded with a question mark (?), in some circumstances two (??) indicating that they are not included in the strict version of the distribution analysis (§4.4).

Two checkboxes follow. “Irreg. correspondences” indicates when irregular phonological alternations between comparanda must be reconstructed. “Remarkable phonotactics” indicates the existence of sequences of phonology that are not reconstructible (e.g. for Latin: unrhotalized intervocalic *s* after a short vowel, lack of vowel weakening) or are not valid from a PIE perspective (*a*-vocalism, **b*, invalid root structures, gemination). The next line comprises a semantic categorization (analyzed in §5).

The last section before the text of the entry includes bibliographical information. The main sources utilized are the etymological dictionaries of Walde and Hofmann (abbreviated WH), Ernout and Meillet (EM) and de Vaan (DV). In the former two, the suspicion of substrate origin is often indicated via the designation “Mediterranean”. References to these three works as well as to Pokorny’s 1959 dictionary are given in the first line of citations in each entry, even if they do not feature in the text of the entry. Further literature appears in the next line. I have made frequent use of Schrijver’s 1991 *The Reflexes of the Proto-Indo-European Laryngeals in Latin* and the Leiden series of etymological dictionaries (esp. Derksen 2007, Beekes 2010 [EDG], Martirosyan 2009, Matasović 2009, Kroonen 2013, Derksen 2014). Other frequently recurring citations are those of Bertoldi, Battisti, Alessio, and Hubschmid.

2.2 Non-inherited Origin in Latin Accepted

2.2.1 Phonotactic Reasons

2.2.1.1 Isolated to Latin but with Unrhotalized *S*

asīlus ‘gadfly’

Pre-form: **h₂eS-* | PItal. **aSīlo-*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: animal, wild; insect

WH (I: 72), EM (51), DV (57)

Gil Fernandez (1959: 157), Breyer (1993: 335-6), EDG (1062), Mata Oroval (2017: 52-6), Weiss (2020: 301 fn. 88)

Latin *asīlus* is likely a loan due to its single intervocalic *s*, which does not occur in an environment where it could be the result of a simplified geminate (DV 57) unless by the *mamilla* rule (cf. *pusillus* ‘tiny’ < **pussillo-* < **putslo-lo-*, DV 502). However it has no comparanda to elucidate a potential source. EM (51) mention a connection with Gk. οἴστρος ‘gadfly’, but this is too formally dissimilar (cf. DV 57) and likely to be inherited in Greek, i.e. continuing **h₃eis-tro-* to a root **h₃eis-* ‘to irritate’, cf. Gk. οἴμα ‘rush, attack, rage’, Lat. *īra* ‘anger’ (Gil Fernandez 1959: 157, EDG 1062). Otherwise, WH (I: 72) and EM (51) accept the potential of Etruscan origin, as Servius claims the names

Asīlus and *Asīlās* are Etruscan. But the reliability of this evidence is difficult to determine, and Breyer (1993: 335-6) finds any connection with attested Etruscan material semantically and morphologically untenable. Inspired by the potential Etruscan connection and in light of the possible Anatolian origin of Etruscan, Mata Oroval (2017: 52-6) suggests that *asīlus* could be from a diminutive of *asinus* ‘donkey’ (since the donkey seems to have been introduced from the East). The *Benennungsmotif* would be similar to Engl. *horsefly*. But the attested diminutive of *asinus* is *asellus* (besides one Late Latin attestation of *asinulus*, cf. Du Cange’s *Glossarium mediae et infimae latinitatis*), making it difficult to prove that the required preform for *asīlus*, namely *asinulus*, existed in antiquity (s.v. *asinus* for more details). Additionally, while the proposed semantic change is possible, such changes, especially when the phonological details are complicated, are impossible to confirm. Lat. *asīlus* remains a recent loanword due to its lack of rhotacism.

asinus ‘donkey’

Pre-form: **h₂eS-in-* | PItal. **aSino-*

Comp.: ?**Ho(s?)-n-* | PGk. **ono-* | Gk. ὄνος ‘donkey’

?PSem. *‘*atān-* ‘female donkey’

?Sum. *anšu* ‘donkey’

??HLuw. *tarkasna-* ‘donkey’

□ Irreg. correspondences

■ Remarkable phonotactics

Semantics: animal, domestic

Pokorny (301-2), WH (I: 73), EM (51), DV (57)

de Lagarde (1877: 56-7), Solmsen (1888: 89-90, Meyer (1892: 319-20), Stolz (1902: 96-7), Niedermann (1903: 113), Pedersen (1906: 449), Brugmann (1908), Vasmer (1908), Cuny (1910: 160), Haupt (1915), Neumann (1964: 61), Leumann (1977: 143, 179, 306), EIEC (33-4), Schrijver (1991: 252), Janda (1999: 194-5), Melchert (2003: 195-6), Militarev & Kogan (2005: 29), Rix (2005: 568-9), EDG (274, 593, 1086), Kogan (2011: 206), Simon (2017: 328-9 fn. 58), Milevski & Horwitz (2019), Weiss (2020: 301 fn. 88), Todd et al. (2022)

Latin *asinus* is widely suspected of being a recent loan due to its single intervocalic *s*, which does not occur in an environment where it could be the result of a simplified geminate (Leumann 1977: 179, DV 57, Meiser 2010: 125, etc.). It must have entered Latin after around the fourth century BCE.

It has been proposed that there were several loci of donkey domestication occurring ca. 4500 BCE including North Africa and the Levant (Milevski & Horwitz 2019). A recent genetic study supports a single original domestication in North Africa around 5000 BCE, with a spread to Eurasia by 2500 BCE. Donkeys were present in Italy by the first

millennium BCE (Todd et al. 2022), and it seems that they were introduced to Italy and Greece from the East (EIEC 33-4). Thus, despite the donkey's agricultural importance as a beast of burden, it did not travel via the original spread of agriculture.

Gk. ὄνος 'donkey' is a good match for *asinus* semantically, but no common pre-form can be established. All who accept the comparison assume borrowing into Latin and Greek from a third source (e.g. Meyer 1892, Stolz 1902, Brugmann 1908, Cuny 1910: 160). But attempts to understand the relationship have all faltered. Meyer (1892: 319-20) and Stolz (1902: 96-7) assume independent borrowing in Greek and Latin from a pre-form **asnos*, which would have to have occurred after the date at which a form like PGk. **osnos* would have yielded Att-Ion. ***ὄβνος*. But **asnos* would have given Lat. ***ānos* and any later anaptyxis in this environment is unparalleled (Niedermann 1903: 113-14, Brugmann 1908: 200). Niedermann (1903: 114) assumes original **asenos* behind Lat. *asinus* which Brugmann (1908: 200-2) further reconstructs to **asonos*, allowing him propose that Greek underwent *Fernassimilation* to **osonos* > **ohonos* (in contrast, WH I: 73 start with initial **o-* and suggest the Latin *a* is the result of a Thraco-Illyrian treatment), which was metathesized to **hoonos* reanalyzed as ὄ ὄνος. But this requires *ad hoc* developments in Greek, and the Mycenaean attestation of *o-no* rules out any explanation involving an article. Nor can either Niedermann or Brugmann explain the lack of rhotacism in Latin.⁴⁸ While Schrijver (1991: 252) proposes rural dialectal origin and Rix (2005: 568-9) suggests it is a Sabellic loan, both are difficult to prove.

Given the difficulty of getting the Greek and Latin forms to match, several (cf. Vasmer 1908 with lit.) instead reject the connection and adduce Gk. ὄνος to Lat. *onus* 'burden' (< **h₃en-os-*, cf. Skt. *ānas-* 'heavy cart', etc.) on comparison with several other Balkan words that have the double meaning 'donkey' and 'burden'. Chronologically, since ὄνος appears already in Homer while *asinus* entered Latin after rhotacism, if the words are related after all, the solution may simply lie in the forms arriving from different sources.

The source of the lexemes, and further comparanda in general, are not entirely clear. Early on, de Lagarde (1877: 56-7) argued against comparison with PSem. **'atān-* 'female donkey' due to its semantic restriction to the female animal and its **t*. But despite Haupt's (1915) suggestion that the Semitic lexeme is deverbal from a root *'wt* 'to agree/consent' found a few times in the Old Testament, current scholarship does not seem to consider it to have any internal etymology (cf. Militarev & Kogan 2005: 29, Kogan 2011: 206). Nor is it present in South Semitic, making a loan into the Semitic languages conceivable.

Sum. *anšu* 'donkey' is identical in meaning and has a sibilant like Latin. The order of its

⁴⁸ The diminutive *asellus* is sometimes explained as the normal development of **asen-elo-* (Leumann 1977: 143, 306; Weiss 2020: 301 fn. 88), suggesting that the *i* of *asinus* is weakened from **e*. (An original **a*, closer to the Semitic forms, is also possible, but a pre-form **asano-* would have been liable to preserve its medial vocalism via the *alacer* rule.) It is difficult to confirm whether vowel weakening or rhotacism occurred first, but it is more likely that rhotacism is the later change (cf. Weiss 2020: 208-9). In this case, *asellus* would be an analogical diminutive on the model of e.g. *geminus* 'twin-born' : *gemellus*.

consonants might represent a metathesis from whatever the source form was (WH I: 73), given its geographic position farther to the East of the direction of the spread of the donkey. The original order of consonants may be preserved in Semitic and perhaps HLuw. *tarkasna-* ‘donkey’ if it is interpreted as *tark* + *asna* ‘draft donkey’ (Neumann 1964: 61, cf. EIEC: 34, DV 57). This interpretation is problematic however, since *-asna* is a relatively frequent suffix in Luwian. A more traditional etymological explanation takes *tarkasna-* as an internal derivation from **d^herǵ^h-* ‘to fasten’ (Janda 1999: 194-5, Melchert 2003: 195-6) as ‘*eine Last habend’. As this root is poorly attested, eDiAna⁴⁹ prefers a derivation from **d^hreg^h-* ‘to drag, haul’.⁵⁰

As early as Pedersen (1906: 449), Arm. *ēš* ‘donkey’ has been recognized as a reflex of **h₁ék_uos* ‘horse’ (though both WH [73] and EM [51] still disagree). Pedersen was willing to see *asinus* as derived via some intermediary from the Armenian collective formation *išan(k^c)* ‘donkeys’, the phonological details of which would all necessarily be *ad hoc*. WH rather see this form as borrowed along with Gk. ἴννος ‘hinny’ from a Pontic word ***išno-*, but further connect this to *asinus* despite correctly rejecting a preform like **asnos*. In fact, ἴννος ‘hinny’ cannot be separated from several other asinine terms in Greek (ἰννός [Hsch.], γίννος/γιννός/γίννος ‘offspring of a mare by a mule’, ὄννος, EDG [273, 593]). Gk. ὄνος ‘donkey’ looks admittedly more similar to these than it does to Lat. *asinus*.

Each of the potential comparanda to Lat. *asinus* requires extra assumptions. The Greek form(s) lack a sibilant; the Semitic forms have a dental instead of a sibilant; Sumerian requires the assumption of metathesis; Anatolian would be homophonous with a frequent suffix. An explanation may lie in the different time periods in which this lexeme was borrowed, but this is difficult to prove. In the end, the Latin form may be isolated (cf. the more or less exasperated *non liquet* of Solmsen 1888: 89-90). This entry would be placed in the uncertain category if it were not for the unrhotalized *s*, which at least guarantees its status as a recent loan in Latin regardless of the identification of its source.

casa ‘cottage, hut’

Pre-form: **ka/HS-* | PItal. **kaSā-*

Comp.: ?

□ Irreg. correspondences

■ Remarkable phonotactics

Semantics: architecture

Pokorny (534), WH (I: 175-6), EM (103), DV (96)

⁴⁹ <https://www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=319>, entry by Andreas Opfermann.

⁵⁰ Simon (2017: 328-9 fn. 58) however finds evidence that *tarkasna/-* actually means horse, and that it is the derivative *tarkasniya-* lit. ‘horse-like’ that means ‘donkey’. This would make an etymology calling the *tarkasna-* the ‘load-bearing one’ less likely. As an alternative, he proposes a root PIE **trg-* ‘goat/horse’ in HLuw. *tarkasna-* and Gk. τράγος ‘goat’.

Buck (1904: 66), Johansson (1906: 114), Berneker (1908-14 I: 589), Reichelt (1914: 340-1), Derksen (2007: 241, 244), Kroonen (2013: 313)

The non-geminate intervocalic *s* after a short vowel indicates a post-rhotacism borrowing (EM 103, DV 96). Etymologies that require a pre-form ***cassa* are thus difficult to defend but include a ‘dialectal’ development of **kat-ja* to a root **kat-* ‘to plait’ in i.e. Lat. *cassis* ‘hunting net’ (WH I: 175-6), though **tj > s* seems to be restricted to the Oscan of Bantia (Buck 1904: 66, though s.v. *rosa* for more details). Reichelt (1914: 340-1) compares descendants of PSlav. **kotъ* ‘booth, sty’ (on the form see Derksen 2007: 241) and **kotja* ‘hut’ (on the form see Derksen 2007: 244). DV (96) mentions further comparisons to OE *heador* ‘incarceration, jail’ (cf. also Johansson 1906: 114, WH I: 175) and Av. *kata-* ‘chamber’ (cf. also Berneker 1908-14 I: 589), though the vocalism of the latter has not palatalized the *k* or been lengthened by Brugmann’s Law, making the only possible reconstructions **kat-* or **knt-* (similar to PSlav. **kotja*). If Lat. *casa* is related to these isolated forms, it would establish an *s/t* alternation in the root.⁵¹ Otherwise, its source is simply unknown.

2.2.1.2 Isolated to Latin but with an Invalid Root Structure

faex ‘wine sediment, dregs’

Pre-form: **b^h/d^h/g^{wh}aik-* | PItal. **f/b/χ^vaik-*

Comp.: ?

□ Irreg. correspondences

■ Remarkable phonotactics

Semantics: viticulture

WH (I: 444-5), EM (213), DV (199, 229)

Bezzenberger (1911: 22), Alessio (1941a: 552)

Lat. *faex* is without certain comparanda.⁵² Explanations from an inherited perspective are not semantically convincing and are formally difficult (if it originally meant ‘dirt’, relationship to Lith. *bójus* ‘swamp’ or from **b^hoiH-* ‘to be afraid’ like *foedus* ‘foul, filthy’ [WH I: 444-5 with lit.]; if it originally meant ‘what is left behind’, related to Lith. *gaišti* ‘to dawdle’ [Bezzenberger 1911: 22]). EM (213) propose a Mediterranean loan because of its viticultural semantics alone. However, given its invalid **D^heT* root structure, there is a good chance that *faex* is not inherited (similarly s.v. *fracēs*) even without comparanda.

⁵¹ If it is a non-IE lexeme, then perhaps it can be connected to PGM. **kuta-* ‘shed’ (cf. ON *kot* ‘cottage, hut’, etc.) and PGM. **hudjan-* (cf. OHG *huttea*, MGH *hütte* ‘hut’), both classified as non-IE by Kroonen (2013: 313).

⁵² Alessio (1941a: 552) seems to compare Gk. τρύξις ‘unfermented wine, must; dregs’ but his reasoning is difficult to grasp. He gives the two as a Tyrrhenian-Aegean pair “in cui ad un elemento oscuro nel latino corrisponde un elemento oscuro nel greco, ma entrambi appartenenti a radicali diversi.” In essence, they are not worth comparing.

farciō, -īre ‘to fill completely, stuff’

Pre-form: **b^h/d^h/g^{wh}alHrk-* | PItal. **f/β/χ^warkje-*

Comp.: ?

□ Irreg. correspondences ■ Remarkable phonotactics

Semantics: action; culinary

Pokorny (110-11), WH (I: 456-7), EM (216-17), DV (202)

Schrijver (1991: 488-9), LIV2 (s.v. **b^hrek^u-*), EDG (1588), van Beek (2022: 402-9)

Lat. *farciō* reconstructs to an invalid **D^heT* root structure, and as such points to a loan. Within Latin it is potentially from the same root as *frequēns* ‘occurring at close intervals’ (WH I: 456-7) < **b^hrek^w-* (LIV2 s.v.). Schrijver (1991: 488-9) suggests that the *a*-vocalism is the result of a syllabic resonant in a complex cluster (here **b^hrk^wj-*). Its only potential match, Gk. φράσσω, Attic φράττω ‘to fence in, surround’ is semantically remote and cannot attest to the original voicing of its velar (Schrijver 1991: 489, LIV2 s.v. **b^hrek^u-*, DV 202, EDG 1588). Nor do any of the Greek forms attest to labiovelar, further weakening the comparison with *frēquens*. Van Beek’s (2022: 402-9) derivation of φράσσω from the root **b^herǵ^h-* ‘to rise’ additionally removes it from comparison with *farciō*, leaving the Latin verb likely isolated.

focus ‘hearth, fireplace’

Pre-form: **b^h/d^h/g^{wh}ok-* | PItal. **f/β/χ^woko-*

Comp.: ?

□ Irreg. correspondences ■ Remarkable phonotactics

Semantics: domestic life

Pokorny (495), WH (I: 521), EM (243), DV (228)

Schrijver (1991: 465-74), Hamp (1992), Martirosyan (2009: 191), Meiser (2010: 82), Weiss (2020: 150)

In order to find an explanation for Lat. *focus* that does not require an invalid **D^heT* root structure, Hamp (1992) proposed a backformation from **foculus* ‘brazier’: **d^hg^{wh}-e-tlo-* (cf. *foveō* ‘to warm’ < **d^heg^{wh}-*) > **g^{wh}-e-tlo-* > **χ^weklo-* > **f^weklo-* > **foklo-* > *foculum*. DV (228) notes chronological problems however. The change **e* > *o* / **w_C(C)V_[back]* (Schrijver 1991: 465-74, Meiser 2010: 82, though Weiss 2020: 150 requires the consonant to be a nasal) is not prehistoric. The pre-form of *bonus* is inscriptionally attested as DVENOS. It seems very unlikely that at this time, the reflex of **g^{wh}* was still **f^w*. Otherwise, Martirosyan (2009: 191) follows several before him (cf. WH I: 521 with lit.) in connecting Arm. *boc^c* ‘flame’ < **b^hok-so-*. There are few issues formally save that it too requires an invalid root structure. Martirosyan suggests it is a substrate lexeme with a distribution like that of Lat. *faber* ~ Arm. *darbin* ‘craftsman,

smith'. The invalid root structure indeed suggests a loan.

pampinus 'shoot or leaf of a vine'

Pre-form: *pa/Hmp- | PItal. *pampino-

Comp.: ?

□ Irreg. correspondences ■ Remarkable phonotactics

Semantics: viticulture

Pokorny (94-5), WH (II: 243-4), EM (478)

Niedermann (1909: 58-9), Lafon (1934: 42-3), Bertoldi (1942: 172-3), Alessio (1946b: 215), Furnée (1972: 272), EDG (91), Smoczyński (2018: 906), van Sluis (fthc.)

Indo-European did not have *C_ieC_i roots,⁵³ so the reconstruction *pa/Hmp- looks non-IE. Its exact relationships to proposed comparanda are unclear. Pokorny (94-5) compares Baltic words like Lith. *pam̃pti* 'to swell, bulge', but Smoczyński (2018: 906) explains this as onomatopoeic. Otherwise *pampinus* is frequently compared to Gk. ἄμπελος 'grape vine' as a loan from a common Mediterranean source (Niedermann 1909: 58-9, WH II: 244, EM 578, Alessio 1946b: 215). EDG (91) notes that there is no reason beyond a lack of IE explanation to suspect that ἄμπελος is a substrate word. Bertoldi (1942: 172-3) instead compares *pampinus* to several Romance words for 'raspberry' like Rhaeto-Romance *ámpua*, Tuscan *ámpola*, *lampone*, etc. Alessio (1946b: 215) rejects the comparison on semantic grounds, and van Sluis (fthc.) proposes a better match for the group: an *a*-prefix alternation relationship with PCelt. *mab- (cf. W *mafon* 'raspberries'). Note though that Sardinian *zampina* means 'wild grapevine' (Alessio 1946b: 215). Lafon's (1934: 42-3) comparison of Abkhaz *paṛəniž* 'black grape' and Georg. *babilo* 'tall vine stock (or vine)' is widely followed (WH II: 243-4, Furnée 1972: 272, EM 478), though he admits it is unclear if they are loans or not. Thus it is unclear if Lat. *pampinus* has any relatives after all.

tabānus 'gadfly'

Pre-form: *ta/Hb^h- | PItal. *tafāno-

Comp.: ?

□ Irreg. correspondences ■ Remarkable phonotactics

Semantics: animal, wild; insect

WH (II: 639), EM (672)

REW (no. 8507, 8601b), Ernout (1946: 41), Ernout (1954: 53), Alessio (1955: 654-5), Latte (1955: 196-7), Furnée (1972: 200, 231, 388), FEW (XIII[1]: 6), Breyer (1993: 388-90), EDG (303, 534), Weiss (2020: 504)

⁵³ Cf. fn. 44.

Lat. *tabānus* ‘gadfly’ has no comparanda. The Romance languages continue *tabānus* (e.g. Calabrian *tavanu*), *tabō*⁵⁴ (e.g. Rom. *tăun*, Fr. *taon*), and **tafānus* (e.g. It. *tafano*, Prov. *tavan*)(REW no. 8507). Sp. *tábano* seems to attest to **tábānus* (Weiss 2020: 504). Ernout (1946: 41) considers the Etruscan personal names *taφane* and *taφunias* to be the source of the Latin word and thus the explanation for the variant with *f*. But the forms with *f* for Lat. *b* are easily explained as continuing a Sabellic reflex of the word (cf. Weiss 2020: 504).⁵⁵ Breyer (1993: 389) also notes that Etr. *φ* does not equate to Lat. *f*. Later, Ernout (1954: 53) takes a more conservative approach, suggesting that both the Italic and the Etruscan could have been borrowed independently from a substrate or that Etruscan had borrowed from Italic. Later still, EM (672) merely mention that the form is found in the Etruscan names.

Alessio (1955: 655) links PRom. **tauna* ‘wasp, bee’ (cf. Lyonnais *tona*, South Fr. *tauna*). REW (no. 8601b) writes that the gender and accentuation mean that it cannot be linked with *tabānus*, which would make *tabānus* and PRom. **tauna* independent comparanda with a *b ~ w* alternation (cf. Furnée 1972: 231). However FEW (XIII[1]: 6) shows that this is not the case. The forms that reconstruct to **tauna* are actually developments from **tabōne* (the oblique of *tabō*) with a secondary accent shift.

The only potential external comparanda are Hesychian forms. They are problematic (cf. EDG 534). Furnée (1972: 200) links Lat. *tabānus* with Hsch. θάπτα· μυῖα, Κρητες ‘fly, Cretan’ and Gk. δάπτης ‘gnat’. The form δάπτης is however from δάπτω ‘to devour, consume’, and is better translated as ‘eater’ (EDG 303), perhaps referring to a carnivorous animal or person (Latte 1955: 196). The Hesychius gloss which caused the confusion, Latte (1955: 196) argues, is itself corrupted. Rather than θάπτα· μυῖα, Κρητες, it should read θάπτρα· μνήμα, Κρητες ‘monument, Cretan’, with θάπτρα related to θάπτω ‘to bury’. Furnée (1972: 388) further links to θάπτα the Hesychius gloss λάττα· μυῖα, Πολυρρήνιοι ‘fly, Polyrrhenian (in Crete)’, with λάττα probably from **λαπτα* (though this suggestion is based on comparison with θάπτα). Given the problems with the other forms, the comparison is too risky.

In the end, despite a lack of comparanda, PItal. **tafāno*- reconstructs to an invalid **TeD^h* root structure, making inherited origin unlikely.

2.2.2 Comparanda in Other Branches

2.2.2.1 Non-inherited Origin is Probable

alnus ‘alder’

Pre-form: **h₂el-s-no-* | PItal. **alsno-*

⁵⁴ Also attested in later Latin.

⁵⁵ WH (II: 639) however note that it is generally *not* in the formerly Sabellic areas that the forms with *f* occur. Alessio (1955: 655) adds that one of the forms with *f* is Tuscan *tafano*. He takes these two facts as evidence that the *f* forms are not Oscanisms but rather point to an Etruscan origin. But Umbrian is not unattested in Tuscany.

- Comparanda: **h₂el-is-* | PSlav. **olbxa-* | Ru. *ol'xá* 'alder', etc.
 **h₁el-is-* | PSlav. **elbxa-* | Slk. *jelcha* (dial.) 'alder', etc.
 **h₁el-(i)s-* | PBalt. *(*a*)*el(i)snio-* | Lith. *alksnis, elksnis* 'alder'
 **h₂el-us-* | PGm. **aluz-* | ON *qlr*, OE *alor* 'alder'
 **h₂el-is-* | PGm. **alis/zo-* | OHG *elira*, MoDu. *els* 'alder'
 ?**h₁el-(i)s-* | PGm. **elustrō-* | ON *jólstr* 'laurel willow', etc.
 ??Macedonian *ἄλιζα* (Hsch.) 'white poplar'
 ??**h₂el-s-no-* | PAlb. **alsno-* | Alb. *halë* 'black pine'

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, tree

Pokorny (302-4), WH (I: 31), EM (23), DV (34)

Pedersen (1895: 40), Specht (1947: 59), Szemerényi (1960: 227-9), Zinkevičius (1966: 131-5), Friedrich (1970: 70-3), Huld (1981: 304), Corominas & Pascual (1984-91 I: 175), Puhvel (I: 29-30), Magnusson (1989 s.v. *jölstur*), Schrijver (1991: 40), Andersen (1996: 130), Demiraj (1997: 193-4), Derksen (2002: 6), Derksen (2007: 370), Kroonen (2013: 22), Simon (fthc.)

The main difficulties in reconstructing the comparanda are the forms in *el-* beside *al-* and the alternation in suffix vocalism *-is-* ~ *-s-*. Unsatisfactory explanations have included reducing the root to **el-*, *ol-*, *l-* 'red, brown' referring to the color of the wood (WH I: 31, Pokorny 302-4), explaining the alternations as secondary (Szemerényi 1960: 228, Friedrich 1970: 70-3), and writing the variation off as a phenomenon common with tree names (EM 23).

An alternating reflex *a/e* of **e* within Baltic (Lith. *alksnis, elksnis*, dial. *aliksi*; Latv. *alksnis*, dial. *elksnis*; the *k* is secondary) is common due to Rozwadowski's change (Andersen 1996: 130, Derksen 2002: 6), but the same alternation cannot be explained away in Slavic. That the original vocalism of Slavic is **al-* (> **ol-*, preserved in Ru. *ol'xá*) and that the forms in **el-* represent contamination from *elka* 'spruce' < **edl-* (Schrijver 1991: 41) cannot explain Slk. dial. *jelcha* because the West Slavic reflex of **edl-* 'spruce' is *jedl-*, precluding the contamination (Derksen 2007: 370). The *a* ~ *e* alternation thus seems to be original in Slavic, suggesting that it may also be so in Baltic. The same alternation may also be present in Germanic. Most forms there go back to an initial **al-*. Gothic **alisa* might even survive in Spanish *aliso* (cf. Szemerényi 1960: 227, Schrijver 1991:40), although the latter has also been interpreted as independent evidence (Corominas & Pascual 1984-91 I: 175).⁵⁶ Even without a Gothic form, the West Germanic languages attest both Verner variants: MDu. *else* < **alisan-* vs. OHG *elira* (and *erila* with metathesis) < **alizō(n)* (Kroonen 2103: 22). It is within Old Norse that

⁵⁶ Go. **alisa* should yield Spanish ***álasa* or ***alésa*, not the attested *aliso*.

the forms *jōlstr* ‘laurel willow’ (< **elustrō-*, with unclear *-u-*) and *ilstri* ‘willow’ (< **elistrīo-*) show initial **el-* beside *ōlr* ‘alder’ < **aluz-*. But because both *jōlstr* and *ilstri* are types of willow (Magnusson 1989 s.v. *jōlstur*), not alder, their evidentiary value is reduced. If related, Schrijver (1991: 41) suggests that their initial vocalism might have arisen through analogy to the *elm* word. But in light of the Slavic situation, it may be original.

For the vocalism of the suffix, the two Verner variants in Germanic along with what seems to be a secondary vowel in OE *alor* < **aluz-* suggest that this lexeme was early remodeled into an *s*-stem,⁵⁷ meaning that Germanic does not actually offer evidence of original *-is-* vocalism of the suffix (Schrijver 1991: 41).⁵⁸ This vocalism *is* demonstrable for Balto-Slavic however, where the explanation given for Germanic cannot apply (cf. Derksen 2007: 370). The Baltic forms of the shape **a/elsnio-* can have arisen by late syncope (Szemerényi 1960: 228) which, despite Schrijver’s (1991: 42) dismissal and explanation that the Slavic forms with **-is-* have innovated an “ancient secondary ablaut,” *does* sometimes occur within Lithuanian (Zinkevičius 1966: 131-5). Weak further evidence of an **-is-* suffix is potentially to be found in Hsch. ἄλιζα ἢ λεύκη τὸ δένδρον. Μακεδόνες ‘white poplar, Macedonian’, but Schrijver correctly points out that we do not know enough about Macedonian to be able to make any claims.

Lat. *alnus* offers incontrovertible evidence of a zero-grade of the *s*-suffix. It can only go back to **alsno*-⁵⁹ (WH I: 31, DV 34, *pace* Pedersen 1895: 40, Szemerényi 1960: 228) because **alisino-* > ***alernus* and **alisno-* > ***alīnus*. Huld (1981: 304) reconstructs Alb. *halë* ‘black pine’ to the same preform (**Azēls-no-*), but its semantics are aberrant and Demiraj (1997: 193-4) notes that several other reconstructions are possible. In any case, Latin proves an alternation *-is-* ~ *-s-* in the suffix that is not reconcilable from a PIE perspective. This along with the likely *a* ~ *e* alternation within the vocalism of the first syllable shows we are dealing with a non-Indo-European lexeme. The *n*-suffix of Lat. *alnus* and the nasal element in the Baltic forms as well as the *-str-* suffix of Oic. *jōlstr* are potentially pieces of substrate morphology.

Puhvel (I: 29-30) considers the possibility that Hitt. ^{GiS}*alanza(n)-* (c.) ‘a kind of tree’ might rather be related to this group of words than to Gk. ἐλάτη ‘silver fir’ with which it is sometimes compared. Hittite would have to have metathesized **alsno-* > **alṣno-* > **alansa-* after which /ns/ > /nts/ is regular and would produce *alanza-*. Because we have no indications as to which tree ^{GiS}*alanza(n)-* refers, and because the metathesis is not regular, it cannot be compared with any certainty (cf. Simon fthc.).

ascia ‘axe, mason’s trowel’

⁵⁷ Kroonen (2013: 22) notes that it cannot be ruled out that the word originally inflected as a root noun, making it either very archaic or a foreign loan.

⁵⁸ Because the Germanic **-is-/us-* ablaut in *s*-stems is a reflex of an original **-es-/os-/s-* ablaut in PIE (Schrijver 1991: 41).

⁵⁹ Or an *n*-stem formation like **al-en-os* (Specht 1947: 59) but this seems highly unlikely in light of all other comparanda having a suffix containing *-s-*.

Pre-form: **h₂esk-jeh₂-* | PItal. **askia-*

Comp.: **h₂eg^(h)/ks-ih₂-n-* | PGk. **aksīn-* | Gk. ἄξιϋν ‘axe’

**h₂eg^wes-(ih₂-)* | PGm. **akwes(ī)-* | Go. *aqizi*, ON *øx*, OHG *acchus*
‘axe’, etc.

Akk. *ḥaššīnu*, Aram. *ḥšn*, etc. ‘axe’
>> Arm. *kac^cin* ‘axe’

Sum. *hazin* ‘axe’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: tool

Pokorny (9), WH (I: 71-2), EM (50), DV (57)

Cuny (1910: 160), Waldman (1972: 117), Ruijgh (1997: 540), Olsen (1999: 955), EDG (111), Kroonen (2013: 19), Rosół (2013: 21-3), Braune (2018: 242), Bernabé (2021: 115-16)

Lat. *ascia* would match the Greek and Germanic comparanda better if it were metathesized from ***aksia*, but the cluster *-ks-* does not regularly metathesize (cf. *axis*, *texō*, etc.). Lat. *viscum* ‘mistletoe’ against Gk. ἰξός ‘id.’ may represent an example where this has indeed happened (EM 50, DV 57), but the inherited status of the *viscum* lexeme cannot be confirmed (s.v. *viscum* and cf. Cuny 1910: 160). Thus, if this is the only other case of such a metathesis, it only adds suspicion.

If we accept the explanation of sporadic metathesis, and if this occurred after devoicing, putative PItal. ***aksīā* could be the result of **g^w*, **k^w*, **g*, or **k*. Gk. ἄξιϋν seems to rule out the possibility of a labiovelar⁶⁰ because something like **-k^ws-* should have given ψ.⁶¹ A labiovelar is however required by Go. *aqizi* < PGm. **akwesī-*. This creates a strange alternation **h₂eg^wesī ~ h₂eksī*,⁶² ruling out any connection with PIE **h₂ek-* ‘sharp’ (pace WH I: 72). It also makes the sibilant element look like a suffix, albeit one that is not explainable from an IE point of view (in fact, cf. that of *almus*). While the Greek has a suffix *-īn-* (Pre-Greek according to Ruijgh 1997: 540 fn. 11 and EDG 111), Latin a suffix **-jeh₂-*, and Gothic a suffix from PGm. **-ī-*, OHG *ackus* inflects as a root noun (Braune 2018: 242).

Given the peculiarities between the comparanda, it is attractive to consider the Semitic (and Sumerian) forms mentioned by Rosół (2013: 21-3 with lit.) and several before him. Akk. *ḥaššīnu* (other Semitic forms like Aram. *ḥšn*, Syr. *ḥaššīnā*, Arab. *ḥašīn*, and Ge’ez

⁶⁰ The Mycenaean hapax *a-qi-ja-i* is likely a misspelling for *i-qi-ja-i* ‘chariot [dat.pl.]’ (Bernabé 2021: 115-16).

⁶¹ Cf. Gk. πέψω ‘I will cook’ < **pek^w-s-*. That Gk. ξίφος ‘type of sword’ is represented as Myc. *qi-si-pe-e* with a labiovelar is explained by EDG (1036) as the result of a Pre-Greek consonant alternation.

⁶² This is Pokorny’s (9) *agu(e)sī*, *aksī*, the **u* of which is based on Zupitza’s (*apud* Pokorny) suggestion of **aguésī*: **agusjās*.

haššīn ‘axe’ are loans from Akkadian, Waldman 1972: 117) is similar enough to especially Greek ἄξινη with its *-n-* suffix to be compared. While Rosóġ takes the Greek to be a probable loan from Semitic, it is not without problems. In loans, Semitic *ħ* and *h* usually yield Gk. *χ* while Semitic *š* yields *σ* (cf. Gk. *χρυσός* < Phoen. *ħ[u]r[ō/ū]š*). Rosóġ proposes a metathesis from *haššīn-* > Gk. **αχσῖν-* > ἄξινη. Arm. *kac^cin* ‘axe’ looks also to have something to do with Semitic, in this case without the metathesis. But it is likely not a direct loan, as Semitic *ħ/h* should not yield Arm. *k* (Olsen 1999: 955). And to explain the Latin form, already proposed to be a metathesized form of the Greek word, would we have to propose two metatheses? Nor can a Semitic origin explain the lack of the nasal element in the Latin and Germanic forms. In fact, even within Akkadian, attestations of the word exhibit irregular *š ~ z* alternation (Waldman 1972: 117), so the word is probably a loan there as well (cf. Šorgo 2020: 432; EDG 111 suggests an origin in an Anatolian language). Notable also is the existence of Sum. *hazīn* ‘axe’ of nearly identical shape.

Thus we are dealing with a word spread through one or more intermediary languages. Its ultimate origin is uncertain, but it is not Indo-European.

avēna ‘oats; stalk, straw’

Pre-form: **h₂eu-e(k^(w)/g^(w))s-n-* | PItal. **awe(C)snā-*

Comp.: **h₂eu-ik/s-* | PSlav. **ovьsъ* | Ru. *ovēs*, Cz. *oves*, SCr. *òvas* ‘oats’, etc.
**h₂eu-ig^h/S-* | PEBalt. **(a)vižā?* | Lith. *avižà*, Latv. [nom.pl] *àuzas*
 ‘oats’

?West Uralic **we/äšnä* ‘wheat/spelt’

?**ka/o/Hb^h-a/e/os-on-* | PGm. **hab(a)zan-* | ON *hafri*, OHG *habaro*
 ‘oats’, etc.

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, domestic

Pokorny (88), WH (I: 81), EM (56), DV (64)

Strömberg (1940: 87, 137), Huld 1990, Schrijver (1991: 46-7), Cooper (2005: 228-9), Derksen (2007: 384), EDG (32) Kroonen (2013: 197), Aikio (2014: 157), Pronk & Pronk-Tiethoff (2018: 294-5), Kroonen et al. (2022: 19-20)

Despite clearly related forms in both Baltic and Slavic, no single Balto-Slavic pre-form can be reconstructed due to the differences in voicing. The Slavic suggests the reconstruction PBSl. **awiš-* with a voiceless sibilant whereas the Baltic forms require PBSl. **awiž-* (Derksen 2007: 384). Latin *avēna*, which is clearly related, has thus traditionally been reconstructed to Proto-Italic as **aweksna-*, since palato-velars, albeit still with a voicing discrepancy, can theoretically be reconstructed for the Balto-Slavic pre-forms.

Even still, the Latin requires *e*-vocalism against the Balto-Slavic *i*-vocalism of an otherwise unknown **-e/ik/ġ(h)-* suffix (cf. DV 64). A plosive need not even be reconstructed for Italic in the first place, as both **-VKsn-* and simply **-Vsn-* would yield *-ŷn-*. Huld (1990) mentions that the reconstruction of the shape **awig-* was a way to account for Gk. αἰγίλωψ ‘goat/oatgrass, kind of oak’, which is unrelated.⁶³ He says that instead, the Latin, Baltic, and Slavic can be reconciled under some non-IE spirant, which indeed seems to be the most obvious solution (cf. Kroonen et al. 2022: 19-20). These three words are of non-IE origin (EM 56, Schrijver 1991: 46-47, Derksen 2007: 384, DV 64).

Latin seems to have added an *-n-* suffix, which Baltic and Slavic did not. Of course, a suffix **-no/eh₂-* is a frequently occurring piece of PIE morphology, and perhaps Latin added it to nativize the foreign word. This possibility is brought into question if the appurtenance of West Uralic **wešnā* (Finn. *vehnä*, Mordvin *viš*), **wäšnā* (Mari *wiste*) ‘wheat/spelt’ is legitimate (Aikio 2014: 157). Aikio⁶⁴ identifies **wešnā* as a substrate word within Uralic, which, like ca. 45% of the substrate words he identifies, contains **š*. The reconstruction of a spirant is strikingly similar to that suggested for Latino-Balto-Slavic **awe/iS-(na-)*. The semantic match is not as exact, but is still within the realm of cultivated cereals. Most problematically, the Saami word is lacking the initial syllable. For this, OPrus. *wyse* ‘oats’ might be comparable, but it lacks the *n*-suffix and its form might not be thoroughly trustworthy due to the potential for contamination with synonymic OPr. *wisge*, *wysge* ‘oats’ (Pronk & Pronk-Tiethoff 2018: 294-5). If it is indeed related, it suggests that the *n*-suffix found sporadically attached to words of non-IE origin might not always be of IE pedigree.

Mention must be made of Huld’s (1990: 404) suggestion of adducing the Germanic words for oat. While Kroonen (2013: 197) reconstructs **habran-* as a secondary development from **hafra-* ‘billy-goat’ based on the double meaning of Faroese *havur* ‘goat; unthreshed grain’, he notes that the Cimbrian doublet *habaro/havaro* suggests an original **b*. Unless a Verner variant, this prevents a match with the goat word. Huld alternatively reconstructs PGm. **xavazan-*, which we can update to PGm. **haba/ezan-* as if from **ka/o/Hb^h-e/a/os-on-*. Huld explains the source of the PGm. **h* as some fricative that was preserved due to its late borrowing into Germanic as opposed to Latin, Baltic, and Slavic. Seeing as Baltic and Slavic seem to have borrowed the word separately, we can propose a relatively late spread anyways. Its appurtenance would create a *b ~ w* alternation akin to but, perhaps problematically, opposite to that between Lat. *faba* and PGm. **baunō-* (s.v. *faba*). Thus it is difficult to accept with any certainty.

⁶³ Cooper (2005: 228) summarizes earlier proposals where αἰγίλωψ is from ἀργ-. Given that αἰγίλωψ is also a kind of oak, it is possible that this meaning has resulted from a conflation with αἰγίλος ‘oat-grass’ (Strömberg 1940: 87, 137, followed by EDG 32). On the other hand, Strömberg’s derivation of αἰγίλος from αἶξ ‘goat’ may as well be folk etymological. The Greek word is simply too uncertain to compare.

⁶⁴ “The Layers of Substrate Vocabulary in Western Uralic”, talk at the workshop *Sub-Indo-European Europe: Problems, Methods and Evidence*. August 30-31, 2021. Leiden University.

baculum ‘stick, staff’

Pre-form: **ba/Hk-tlo-* / **ba/Hk-elo-* | PItal. **bake/olo-*

Comp.: **ba/HK-el-o-* | PRom. **bakillo-* | Prov. *bacèu* ‘washing staff’, etc.

**ba/h₂k-tro-* | PGk. **baktro-* | Gk. βάκτρον ‘stick, cudgel’

**ba/o/Hk-ió-* / **ba/o/Hg^h-io-* | PGm. **pagjō-* | Engl. *peg*, etc.

**ba/HK-o-* | PCelt. **bakko-* | OIr. *bacc* ‘crook, angle, bend’, etc.

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: tool

Pokorny (93), WH (I: 92), EM (64), DV (67)

Niedermann (1930: 5), FEW (I: 201), Thurneysen (1946: 92-3), Frisk (1960-72 I: 211), Lühr (1985: 283), Schrijver (1991: 100, 105), Matasović (2009: 52), EDG (194), Kroonen (2013: 395), Stifter (2023: 32)

The root in question here is immediately remarkable due to its comprising the two rarest phonemes **a* and **b*, for which reason, along with the root’s limited distribution, Schrijver (1991: 100, 105) is more convinced of a non-IE origin than a need to reconstruct a root shape like **bHK-* for Latin. There is, however, a more convincing argument for a non-IE origin in the inexplicable geminate of some forms.

Gk. βακτηρία beside βάκτρον suggests that the former is an abstract formation from *βακτήρ with the latter being a by-form of the same, as is the case with ἄροτήρ beside ἄροτρον (Frisk 1960-72 I: 211, EDG 194). Thus the Greek forms reconstruct to an original **bak-tro-*. Latin *baculum* superficially looks like it could be a diminutive **bak-elo-*, but in light of the semantics and the Greek form, it is more likely the reflex of the instrument noun suffix *-tlo-* (Niedermann 1930: 5, EM 64). In both cases, these productive formations need not be inherited as archaisms from PIE. In fact, since they are two different agentive suffixes, it is at best a common innovation. Schrijver (1991: 100) however uses the diminutive formation to reconstruct **bak-(k)elo-* for *baculum* because there is evidence of a geminate in some Latin and Romance forms. All extant Romance reflexes of the diminutive *bacillum* go back to **baccillum* with a geminate (FEW I: 201) and Lat. *imbēcillus* ‘weak, feeble’, if literally ‘without a staff’, could suggest the stem was **bāc-* and that the geminate forms within Romance originated by the littera rule. Given the rarity of a situation like this, it is impossible to know if a littera rule derived pre-form like **bakk-elo-* would yield Lat. *baculum*, but it seems like an extra assumption based on an already not fully secure etymology of *imbēcillus*. At face value, Latin and the Romance forms show an alternation **bak-* ~ **bakk-*. Celtic comparanda like OIr. *bacc* ‘bill-hook, angle, bend’ and OW *bach* ‘hook, peg’ securely

attest to a geminate in that they reconstruct to PCelt.. **bakko-* (Matasović 2009: 52).⁶⁵

The Germanic forms in **pag-* are traditionally considered Verner variants of **bak'* (cf. Kroonen 2013: 395), but could theoretically reconstruct to PIE **g^h* as well. In any case, they securely rule out a geminate. Thus, along with the Greek forms, Germanic proves that an alternation **bak ~ *bakk* must be reconstructed for this root, which, in addition to its remarkable phonetics, points to a non-IE root (Schrijver 1991: 100, DV 67, Kroonen 2013: 395, Stifter 2023: 32).

baiulus ‘porter, carrier’

Pre-form: **ba/Hg-* | PItal. **bagjelo-*

Comp.: **ba/o/Hg-nó-* / **ba/o/HK-* | PGm. **pakka-* | ME *packe*, etc. ‘bundle, pack’

**b^ha/o/HG^h-* | PGm. **bagg-* | ON *baggi* ‘pack, bundle’

**b^ha/h₂k-* | PGk. **p^hakel(l)o-* | Gk. *φάκελος, φάκελλος* ‘bundle’

?**b^ha/h₂sk-* | PGk. **p^haskōlo-* | Gk. *φάσκωλος* ‘leather pouch, satchel’, etc.

**b^(h)a/HK-* / **b^(h)a/Hsk-* | PCelt. **bakki-* / **baski-* | W *beich* ‘load, weight, burden’, etc.

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: tool

Pokorny (111), WH (I: 93-4, 459-60), EM (64, 218), DV (68, 203)

Osthoff (1893: 322), Solmsen (1904: 22-6), Bertoni (1910: 25-6), REW (no. 880), Hubschmid (1955: 91-7), Corominas & Pascual (1984-91 I: 453-5), Furnée (1972: 173, 295-301), Schrijver (1991: 100, 102-3), Demiraj (1997: 93-4), Matasović (2009: 58), Kroonen (2013: 396), EDG (1547), Šorgo (2020: 459), GPC (s.v. *baich*)

Osthoff (1893: 322) originally compared Lat. *baiulus* to Gk. *βαστάζω* ‘to lift up’, *βάσταγμα* ‘load’, with the Latin from **bad-jo-* and the Greek from **bad-to-*. But several Romance forms make a pre-form with **g* much more likely (Solmsen 1904: 22-6, Bertoni 1910: 25-6, Schrijver 1991: 100, DV 68, Kroonen 2013: 396). While Span. *baga* ‘flax seed capsule’ is often mentioned, it is likely from Lat. *bāca* ‘berry’ (Corominas & Pascual 1984-91 I: 453-4). However Aragonese *baga* in the sense ‘rope with which loads are tied’ belongs to a group of words including Prov. *baga*, Venetian, Lombardian, Emilian *baga*, and Friulian *bage* ‘bundle, bag, purse’ (REW no. 880, Hubschmid 1955: 91-7 lists many more), which can easily be connected to Lat. *baiulus*.⁶⁶ Its **b* and

⁶⁵ Lühr (1985: 283 with lit.) suggests that PCelt. **bakko-* could be from earlier **bak-no-*, basically a Celtic Kluge’s Law. But several cases must be reconstructed for Proto-Celtic that do not undergo Kluge’s law, such that it likely cannot explain Celtic geminates (Thurneysen 1946: 92-3).

⁶⁶ This makes the first syllable of *baiulus* heavy, such that the actual length of the *a* is indeterminate (cf.

a-vocalism make it unlikely to be inherited.

Two Germanic forms can be compared to *baiulus*. PGM. **pakka-* ‘bundle, pack’ (ME *packe*, MDu. *pac*, etc.) can be mechanically reconstructed to **bagg-* with a (non-IE) voiced geminate. Its geminate can however also be explained via Kluge’s Law < **bag-nó-*. The voiced geminate in ON *baggi* is not the classic outcome of Kluge’s Law, but sometimes occurs as the result of contamination (Kroonen 2011: 124). Otherwise, it points to a reconstruction with a geminate voiced aspirate. The initial consonant alternation between the two Germanic forms has no explanation, pointing to the reconstruction of QPIE **b^h ~ b* (and perhaps **g^h ~ g*) alternations.

Further comparanda are difficult to navigate. Kroonen (2013: 392) further compares Gk. φάκελος ‘bundle’ < **b^hak-*, whose variant φάκελλος has geminate λλ (EDG 1547), with an *l*-suffix reminiscent of *baiulus*. But these forms are difficult to separate from several other Greek words, also referring to bundles, but with an additional sibilant, including Gk. φάσκωλος ‘leather pouch, satchel’, Hsch. βάσκιον· δεσμαί φρύγανων ‘bundles of firewood’, and Hsch. βασκευταί· φασκίδες· ἀγκάλαι ‘bundles’. This leads Furnée (1972: 173) to compare this **b^(h)ak-/ *b^(h)ask-* group to Lat. *fascis* ‘bundle’. Several Celtic forms like W *beich* ‘load, weight, burden’ can be either from **bakki-* (similar to the Germanic) or **baski-* (more similar to *fascis*) (WH I: 94, Matasović 2009: 58, Kroonen 2013: 396, GPC s.v. *baich*).⁶⁷ Neither Schrijver (1991: 102-3) nor DV (203) is convinced that Lat. *fascis* is anything more than an Italo-Celticism,⁶⁸ nor does EDG (1547) mention *fascis* under his entry on the Greek forms.

The Romance forms have long been suspected of being loans from another IE language. For the latter, the REW (no. 880) has somehow suggested Dalmatian origin,⁶⁹ while DV (68) suggests they are either borrowings from Germanic (asserted also by Corominas and Pascual 1984-91 I: 454-5) or Celtic. In fact, he suggests that Lat. *baiulus* itself could be from Celtic. If *baiulus* is a more recent borrowing, then *fascis* may represent the independent Latin reflex of the substrate bundle word. The semantic match is quite good, and the vacillating presence of a sibilant has parallels in other potential substrate vocabulary (cf. Furnée 1972: 295-301 in Greek, Šorgo 2020: 459 on PGM. **aik-* vs. Lat. *aesculus* ‘oak’ and further the entries on *barba* and *turdus*). But even in taking the strictest approach—separating Lat. *fascis* and W *beich*, and having Lat. *baiulus* be a

similarly **mag-īōs > mai(i)or*, **h₂(e)ǵ-īoh₂ > ai(i)ō*, Weiss (2020: 172)). But there seems to be no reason to assume it was long (though it is given as *bāiulus* by LS and DV 68).

⁶⁷ Matasović (2009: 58) writes that OIr. *basc* ‘necklace’, often taken from the same pre-form, is scarcely attested. More importantly, its semantics seem too far removed to compare it with certainty.

⁶⁸ Though Demiraj (1997: 93-4) and Matasović (2009: 59) further adduce semantically distant Alb. *bāshkë* ‘fleece’. (Its homonym *bāshkë* ‘together, common’ is argued by Demiraj to be the same word, with a shift in meaning that has come about in the context of shepherding.)

⁶⁹ Hubschmid (1955: 91-7) takes this much further, claiming an Altaic/Turkic origin on comparison with several words including Old Turkish *bağ* ‘Warenbündel; Strick, Fessel’ and numerous modern Turkic languages with the same form and meaning.

more recent borrowing—does not account for the phonological alternations present between the other comparanda of *baiulus*. It remains a member of a group of words whose morphophonological alternations make them difficult to explain from an inherited perspective.

ballaena ‘whale’

Pre-form: **ba/Hl-d/n/s/ɥ*⁷⁰ | PItal. **ballAEnā*-

Comp.: **b^ha/Hl-ǰ*- | PGk. **p^hallaina*- | Gk. φάλλαινα ‘whale’

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: animal, wild; aquatic

Pokorny (120-2), WH (I: 94-5), EM (65), DV (68)

Brüch (1919a: 198-9), Kretschmer (1923b: 280-1), Frisk (1960-72 II: 987), Leumann (1977 I: 158-9), Biville (I: 178-81), EDG (1549)

The ultimate source of Lat. *ballaena* is Gk. φάλλαινα ‘whale’, generally recognized to be an -αινα derivative of φαλλός ‘penis’ (Kretschmer 1923b: 280-1, Brüch 1919: 199, Frisk 1960-72 II: 987, EDG 1549), with the latter analyzable as either native (< **b^hel-* ‘to swell’ cf. Frisk 1960-72 II: 987) or Pre-Greek (because of evidence of a λ ~ λλ alternation cf. EDG 1549). In any case, it was borrowed into Latin late enough for non-initial -ae- < -ai- to have missed the vowel weakening and monophthongization to -ī- (DV 68). As a reflex of loaned Greek φ, we expect Lat. *p* (especially in early loans) or *ph* (in later and learned loans).⁷¹ Before the imperial period, we have two examples of Lat. *b* for Gk. φ: *ballaena* < φάλλαινα and *Bruges* < Φρύγες (Biville I: 178-82). Given the geographic positioning of Italy and Greece, the suspicion has fallen, not without reason, on the languages of southern Italy.⁷² Brüch (1919a: 199) saw Proto-Greek **bhallaina* entering Illyrian where **bh* > *b*⁷³ as the cause for the Latin consonantism while Kretschmer (1923b: 280-1) preferred Messapic directly. What can be said with certainty is that a third language has mediated the transmission of the Greek word into Latin.

burrus ‘red, reddish-brown’

Pre-form: **bur-so-* | PItal. **burso-*

⁷⁰ Several other clusters can produce *ll* in Latin, but as Greek also has a geminate and there are further irregularities between the forms, it would be artificial to provide all possible pre-forms as though the word were inherited.

⁷¹ Leumann (1977 I: 159) is incorrect in proposing that *ballaena* could be directly from φάλλαινα; his reasoning is that Gk. π sometimes yields Lat. *b* in loans. Despite the ancient grammarians using this exact argument, Lat. *b* for Gk. π is also irregular.

⁷² The explanation for *Bruges*, as it is a para-Balkan ethnonym, is more complex. It could be due to other languages that likely had the *b* < **bh* change that include Macedonian, Thracian, and Phrygian itself.

⁷³ cf. Messapic *berad* = Lat. *ferat* (Biville I: 180).

Comp.: **ph₂ur-s-uo-?* | PGk. **purswo-* | Gk. *πυρρός* ‘blazing red, tawny’

■ Irreg. correspondences □ Remarkable phonotactics

Semantics: color

WH (I: 124), EM (78-9)

Cuny (1910: 160), Kretschmer (1928: 166), Schulze (1933: 115-16), REW (nos. 1117, 1410, 1416), Kahane & Kahane (1960: 138-142), Furnée (1972: 157), Biville (I: 237-8), EDG (1264)

The same correspondence is found between the proper names Lat. *Burrus* and Gk. Πύρρος. But here the discussion is of the adjectives for ‘red’. The adjective is attested only in glosses, lexicographers, and in Paulus *ex Festo*, who tells us it is an old word for *rufus* ‘red’. But as Biville (I: 237-8) notes, it must have been in widespread use as it has many Romance descendants (from *birrus*, REW no. 1117; *burrus*, REW no. 1416; and *būrius*, REW no. 1410). The correspondence between Lat. *b* and Gk. *p*, like in *buxus* (s.v.) is irregular, leading Biville (I: 238) to conclude that it attests to a borrowing from Greek through an unknown language⁷⁴ (cf. also EM 79) or that it occurred within the Greek of Southern Italy. There does not seem to be any evidence of the latter.

Biville along with e.g. WH (II: 124) and EM (78-9) are certain that the Latin word, directly or not, was borrowed from Greek. The geminate in *πυρρός* is likely though a pre-form **purswo-* (Schulze 1933: 115-16, EDG 1264), which seems to contain the color adjective suffix *-*yo-* (esp. if the Corinthian horse name Πυρρός is related). It has been suggested to derive from Gk. *πῦρ* ‘fire’ (already called a *pis-aller* by Cuny 1910: 160) or to be related to Lith. *pu̯vas* ‘dirt, mud’ (Schulze 1933: 115-16). Furnée (1972: 157) proposes that *πυρρός* ‘red’, *πυρρός* ‘torch, firebrand’, and *burrus* descend from a pre-IE color term, perhaps for horses (cf. Lat. *būricus* ‘small horse’, though EM 78 give this another etymology), coincidentally similar to *πῦρ* ‘fire’ and thus later folk etymologically connected to it. In the end, *πυρρός* seems to have a pedigree in Greek. Its presence in Latin as *burrus* shows it arrived there via an intermediary language.

buxus ‘box-tree’

Pre-form: **buk^(w)/g^(w)(^h)-so-* | PItal. *bukso-*

Comp.: **puk-so-* | PGk. **pukso-* | Gk. *πόξος* ‘box-tree’

■ Irreg. correspondences □ Remarkable phonotactics

Semantics: plant, tree

⁷⁴ Kretschmer (1928: 166) proposed Etruscan, especially based on the Hesychius gloss *βυρρός*: *κάνθαρος ὑπὸ Τυρρηγῶν*. But 1) the semantics do not match, 2) the use of words meaning ‘Etruscan’ by Hesychius and Pseudo-Dioscorides often simply refers to Italy and not the Etruscans themselves (cf. Breyer 1993: 133), and 3) it is generally only intervocalically that the consonants of Etruscan names were perceived by Latin-speakers to be voiced (cf. s.v. *taeda*). Thus we can rule out Etruscan intermediation. Kahane and Kahane (1960: 138) simply explain it as ‘rustic’, but this does not feel like an explanation.

WH (I: 125), EM (79)

Cuny (1910: 160), Boisacq (1911-2: 58-9), Fohalle (1925: 171), Furnée (1972: 157), Biville (I: 240), Breyer (1993: 180), EDG (1259)

The mismatch in voicing between Lat. *buxus* and Gk. *πόζος* ‘box-tree’ is not normal in loans between the two languages (Cuny 1910: 160). Boisacq (1911-2: 58-9) and EM (79) take them to be from a language of Asia Minor (Boisacq specifically names Thracio-Phrygian). EDG (1259) is unconvinced by all attempts to feed an Indo-European root through another IE language to arrive at the attested forms (like **puk-s-o-* to **peuk-s-* ‘fir tree’). Nor is it likely that Etruscan intermediation could produce voicing in this position as a loan from Greek (Breyer 1993: 180). WH (I: 125), Furnée (1972: 157) and EDG (1259) assert that the box-tree is native to Italy, not Greece or Asia Minor, so the Greek form would have to be a borrowing from Latin (perhaps with devoicing through Etruscan). Biville (I: 240) instead follows the argumentation that this word represents a loan into both Latin and Greek from a third language in the Mediterranean (cf. also Fohalle 1925: 171).⁷⁵

caballus ‘horse, esp. a nag or workhorse’

⁷⁵Kretschmer (1928: 166-7) preferred to take cases like these as loans from Greek, but it seems that the only alternative that he considered was the one presented by Fohalle (1925), namely that both Latin and Greek had borrowed from a third language. “Gegen [diese] Möglichkeit,” he wrote, “besteht das Bedenken, daß die voridg. Urbevölkerung der Appenninhalbinsel für uns viel weniger greifbar ist als die der Balkaninsel, daß die Berührungen der Römer und anderer idg. Stämme Italiens mit ihr in sehr alte Zeiten zurückgehen müssen und wir nicht wissen, ob Wörter wie *gōbius cobius*, *conger* so alt im Lateinischen sind.” His apprehension and tendency toward rejection seems thus to have been based on uncertainty about the past. But another option exists, like for *ballaena* and *burrus* above—namely that a third language was responsible for transmitting many of these words from Greek to Latin. Biville (I: 245) for example mentions several nautical terms that have been argued to have entered Latin indirectly from Greek (*gubernāre* [s.v.], *ancora* ‘anchor’ < Gk. *ἄγκυρα* ‘id.’, *anquīna* ‘halyard’ < Gk. *ἀγκοίνη* ‘bent arm’?, *aplustre* ‘ornamented ship stern’ < Gk. *ἄπλαστον* ‘ship’s poop’, *struppus* ‘band, strap’ < Gk. *στρόφος* ‘twisted band or chord’, *supparus* [s.v.]). Even so, there are several examples mentioned by Fohalle that I do not include because their analysis is difficult. Lat. *gōbiō* and *gōbius* ‘goby, gudgeon’ occur beside forms in *-c-*, from Gk. *κοβιός* ‘gudgeon’. Biville (I: 244) considers the voiceless variants learned Hellenisms, but they appear in some Italian dialects. The REW (no. 3815) says they are Southern Italian and show the *κ-* of Greek, which would rule out a Greek dialect of Southern Italy being responsible for the voicing. This is suspicious, but even EDG (812) considers the Latin loaned from Greek. The idea seems to be that the voicing is somehow secondary. It sounds *ad hoc*, but Gk. *κόμμη* ‘gum’ appears first in Latin as *cummi* and later as *gummi*. EDG (744) does not believe that independent borrowings from Egyptian (the source of the Greek) would yield such similar forms, while Biville (I: 257) simply stresses that the form with *g-* is later (suggesting that it developed in Italy post-borrowing). From Gk. *καμπή* ‘bend’, *gamba* ‘horse ankle (pastern)’ seems to have been borrowed in a veterinary context, but it appears late (4th c.) and appears besides *camba*. Thus when both variants are in circulation, which mysteriously seems only to be the case for velars, it is difficult to reach a conclusion. Lat. *amurca* ‘the watery part that flows out of pressing olives’ seems to be a borrowing of Gk. *ἀμόργη* ‘watery part which runs out when olives are pressed’ (itself a derivation of *ἀμέργω* ‘to pluck, squeeze out’) with the devoicing and vowel change suspected to be due to Etruscan (cf. Biville I: 233-4), but no such Etruscan word is attested and the variant *amurga* is also attested and preserved in Romance. It is difficult to decide in favor of 1) a situation involving multiple borrowings but only affecting words with a variation in a velar or 2) a Latin-internal situation that only seems to have affected these few words.

- Pre-form: *ka/Hb^(h)- | PItal. *kab/fallo-
 ??*ka/Hb^(h)-ōn- | PItal. *kab/fōn- | Lat. *cabō* ‘nag’
- Comp.: *ka/Hb- | PGk. *kaballā- | Gk. καβάλλης ‘workhorse, nag’
 *ka/ob^h- | PSlav. *ka/ob- | OCS, Ru. *kobyła*, Pol. *kobyła*, etc. ‘mare’
 *ka/HP- | PCelt. *kappe/ilos | OIr. *capall* ‘horse’, W *ceffyl*, etc. ‘horse’
 *kab^(h)/p- | PIr. *kab/parda- | MoP *kawal* ‘second class horse of mixed blood’

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: animal, domestic; equestrian

WH (I: 125), EM (80), DV (77)

Boisacq (1916: 388-9), Maass (1925: 469), Tafrahi (1925: 259), Kretschmer (1928: 191-2), Kretschmer (1932a: 247-8), van Windekens (1959), Schmidt (1966: 161), Bailey (1979: 52), Emmerick (1981: 185), Huld (1990: 403), Schrijver (1991: 425-35), Breyer (1993: 509), Gamkrelidze & Ivanov (1995 II: 474), Watmough (1997: 54-5), Polomé (1998: 673), Delamarre (2003: 69), Simon (2005: 405-16), Derksen (2007: 231), Schwarz (2008), EDG (611), Weiss (2020: 130)

The shape of Lat. *caballus* is strange enough that it is unanimously considered a loan, but its age is difficult to securely determine. While it looks to have entered Latin after the 4th c. BCE vowel weakening that would have produced **cabellus* (Simon 2005: 411), the absence of weakening could be the result of the *alacer* rule (on the rule, cf. Weiss 2020: 128). Thus it could have entered Latin much earlier, but in either case, the second vowel was *a* when it did so. Both its proximal and ultimate sources and therefore its relationship to Lat. *cabō* are a matter of debate. *Caballus* cannot be regularly derived from *cabō* (pace Polomé 1998: 673 as anything like **kab-on-elo-* should have resulted in *-ullus*, cf. Weiss 2020 301, fn. 88). Nor do there seem to be good parallels for the *n*-stem *cabō* to be a shortening of *caballus* (Nehring 1949: 166, Simon 2005: 407 pace e.g. Maass 1925: 469).

Gk. καβάλλης [masc.] is nearly identical, and the 3rd c. BCE attestation of καβάλλειον (Tafrahi 1925: 259) shows that the lexeme is old in Greek. With *caballus* first attested in Lucilius (2nd c. BCE), it is not the age of the forms that rules out a borrowing from Greek (cf. DV 77) as much as the unexplained difference in endings.⁷⁶ Several (Maass 1925: 469, heartily followed by Kretschmer 1928: 191 and Kretschmer 1932a: 247-8, WH I: 125-6, EM 80, EDG 611) have suggested a Wanderwort with its origins in the name of some horse-breeding ethnic group (on comparison with names for types of horse such as

⁷⁶ Van Windekens (1959: 80) suggests that a borrowing from Greek to Latin cannot be ruled out, but nevertheless prefers an ultimate origin of the forms from PIE **g^hab^(o)l-* ‘forked’ via Pelasgian. We can rule out the latter suggestion on semantic grounds.

Ger. *Gaul*, *Wallach*, and Fr. *hongre*). Maass and Kretschmer preferred a source in Asia Minor due to e.g. Herodotus' mention of the Καβηλλέες in Anatolia, but Nehring (1949: 165) and Simon (2005: 407) reject this as too speculative. Nehring still argues for a proximal Anatolian origin, but on the basis of ethnic names there that vacillate between *-αλος* and *-ων*. He suggests that *caballus* and *cabō* are of Etruscan origin, with the latter representing the tendency of Latin to borrow Etruscan nouns in *-u* with Lat. *-ōn* (cf. further Breyer 1993: 509, Watmough 1997: 54-5). There is no further evidence that could make this any more than a guess.

Nehring (1949: 168-70) noted Persian (MoP *kawal*) and Turkic (11th c. *kevel* in *kevel at* 'well-bred, swift horse') words to suggest an ultimate origin in Central Asia. Others have placed more emphasis on the Slavic words.⁷⁷ Boisacq (1916: 388-9) interpreted them as pointing to a Northeastern European origin of the words. Huld (1990: 403) suggested a North Balkan substrate origin, linking Lat. *cabō*, OCS *κοηβ* 'stallion', and Lith. *kumėlė* 'mare' via *m ~ b* and vocalic alternations. Simon (2005: 408) rejects the connection on the grounds of having only initial *k* in common, and I am likewise weary of putting so much variation under the same roof.⁷⁸ Simon (2005) proposes an Iranian origin of *caballus* and an ultimately PIE origin of the root, but in the end, MoP *kawal* is probably a recent loan.⁷⁹ Another possibility sometimes mentioned is that *caballus* entered Latin from Gaulish (Schmidt 1966: 161, Gamkrelidze and Ivanov 1995 II: 474, EM 80) where it occurs in personal names and toponyms (Delamarre 2003: 96). But Insular Celtic forms (OIr. *capall*, W *ceffyl*, etc.) reconstruct to **kappe/ilo-* (not **cappallus*, *pace* Delamarre 2003: 96, as the OIr. dat.pl. *caiplib* and acc.pl. *caipliu* show; nor does the geminate *ll* of the Old Irish form need to be original, *p.c.* David Stifter). The Insular

⁷⁷ Kretschmer (1928: 192) used Hsch. κάβηλος, κάληβος: ἀπεσκολυμμένος τὸ αἰδοῖον. οἱ δὲ ὄνος, suggesting its use in referring to castrated donkeys, to propose that OCS *kobyła* is a labialized borrowing of a later Greek pronunciation /*kabilos*/. But EDG (611) correctly warns that we cannot be certain of the Hesychian forms' appurtenance.

⁷⁸ Derksen (2007: 231) suggests an "either/or" scenario in which OCS *κοηβ* 'stallion' could be related to *cabō* if it goes back to **kab-n-io-* with a non-glottalic (i.e. foreign) **b* (though technically **b^h* would work for both as well) or it derives from **kom-nb* and instead belongs with ORu. *komonb*, Cz. *komoň* 'horse', and perhaps Lith. *kumėlė* 'mare'.

⁷⁹ Simon (2005) adduces Khot. *kabä* 'horse' as a dialectal development from expected ***kava-* < PIr. **kaba-*, but PIr. **kapa-* would work just as well (Emmerick 1981: 185 for Khotanese, for MoP cf. *nawa* 'grandson' < **napat-*) and does not require the reconstruction of an invalid root structure **keb^h* or **keb-* with rare **b*. Nor is the meaning of *kabä* even certain (Bailey 1979: 52). In any case, the vocalism of the Iranian forms has to have been **a* (**e* would have palatalized and **o* would have undergone Brugmannian lengthening), which is suspicious. Without *kabä*, the only forms that do not have an *l*-suffix are 1) Lat. *cabō* (whose development from **keb-ōn-* relies on Schrijver's [1991: 425-35] proposal of **e > a* after a plain velar), attested only in Late Latin glosses and without continuants in Romance, and 2) Finnish *hepo* 'steed, stallion' if it represents a loan from an otherwise unattested PGm. **hepa-*. As to the forms with the *l*-suffix, Simon explains the geminate of Latin and Greek as independent suffixations of a PIr. **kabala-*, but it is already highly unlikely (though not completely impossible, cf. Schwartz 2008) that the *l* of MoP *kawal* is inherited from Proto-Iranian; thus the PIr. reconstruction here with **rd*. Simon further relies on two other independent suffixations with *l*, one in which Turk. *kevel* is the addition of the Turkic denominative suffix *-(V)l* to borrowed Khot. *kabä*, and another one to get OCS *kobyła* < **kob-ōn-* + *-la*.

Celtic forms therefore look quite old, and are on phonological grounds *not* loans from Latin. The lexeme's presence in the Celtic languages could possibly explain its wide distribution, but in fact the Gaulish forms look more likely to be loans from Latin than vice versa.

In the end, all secure comparanda attest to a Wanderwort of the general shape **kabal-* of unknown origin, though it looks older in the West (Proto-Insular-Celtic) and younger in the East (MoP *kawal*). It is more likely that the *l* was originally present than that the individual borrowing languages all added their own *l*-suffix. Thus any direct connection with *cabō* is difficult to confirm.⁸⁰ If related, its late attestation in glosses and absence from the Romance languages means that it does not need to have been borrowed at the same time or from the same source as *caballus*.

calix 'vessel for food or drink'

Pre-form: **k_lH-ik-*, **ka/Hl-ik-* | PItal. **kalik-*

Comp.: *?*sk_lH-ik-* | PItal. **skalik-* | U **skalčeta** [abl.sg. + postpos. **-ta**], etc.
'sacrificial vessel'

**kul-ik-* | PGk. **kulik-* | Gk. κύλιξ '(drinking) cup'

*?*k^(w)al(H)-ek-* | PIIr. **kalaś-* | Skt. *kalása-* 'pot, jug, bowl'

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: vessel

Pokorny (440-1), WH (I: 138-9), EM (87), DV (83)

KEWA (I: 179, 181), Frisk (1960-72 II: 47), EWAia (I: 321), Furnée (1972: 132), Schrijver (1991: 207), Giacomelli (1994: 36-7, 40), Untermann (2000: 683-4), EDG (628, 800), Beekes (2014: 67), van Beek (2022: 23)

WH (I: 138-9) and EM (87) connect Lat. *calix* and Gk. κύλιξ via a zero-grade formation of a root **(s)kel-* 'to split'. Schrijver (1991: 207), followed by DV (83), proposes that **k^wH-ik-* can regularly yield κύλιξ via *u*-coloring of the vowel that arises to the left of **_l* (cf. γυνή), and the form with *s*-mobile preserved in Umbrian caused the change **sk^w-* > **sk-* whereupon the delabialized velar was spread to the *s*-less forms in Latin. But given Latin words like *squāma* '(fish)scale', it is clear that **s* does *not* delabialize **k^w* in Italic. Furthermore, **k^wH-V-* in Greek ought to give **παλ-V-* (cf. βαρύς 'heavy' < **g^wH-u-*, van Beek 2022: 23). Instead, the *a ~ u* alternation between κύλιξ and *calix* seems to be

⁸⁰ An ultimate connection between the two via something like an original **kabVn-* ~ **kabanlo-*, despite seeming attractive since **nl > ll* in Latin and Greek (and Old Irish if the *ll* of *capall* is original), is made unlikely by the fact that in Slavic, the addition of the *l*-suffix would have to be secondary (after the change **-ōn# > *-ī*). At best, the change **nl > ll* might be typologically frequent and the assimilation could have occurred already in the donor language. At worst, *cabō*, which occurs in glosses also as *cabōnus*, and whose meanings include *caballus grandis/magnus*, really is a late clipping of *caballus* with the *-on(e)* augmentative suffix.

original, indicating a loan (cf. EDG [628, 800], though his position on the appurtenance is unclear; further Frisk 1960-72 II: 47, Schrijver 1991: 207, DV 83).⁸¹ Furnée (1972: 132) took Gk. κολίχνη⁸² ‘small cup’ to attest to a **k* ~ **k^h* alternation, but it instead probably attests to the *n*-suffix that seems to trigger aspiration (cf. Beekes 2014: 67, s.v. *laurus*).⁸³ Also compared is Skt. *kaláśa-* ‘pot, jug, bowl’ (WH I: 138-0, KEWA I: 179, EWAia I: 321, EM 87⁸⁴, Untermann 2000: 684), though DV (83) follows Schrijver (1991: 207) in being more skeptical. However, given the absence of palatalization or Brugmann’s Law, the Sanskrit as well as the Italic words reconstruct to original *a*-vocalism.

Gk. κάλυξ ‘seed-vessel, husk’ has the same vocalism of Latin, but is semantically more distant. Hsch. σκάλλιον· κολίκον μικρόν ‘small cup’ is formally and semantically similar to the Umbrian form, but this may be coincidental. Finally, Furnée (1972: 132) compares semantically similar Hsch. κίλλιξ· στάμνος ‘jar’, but this introduces several more irregularities.

In the end, Lat. *calix* and Gk. κύλιξ attest to an irregular *a* ~ *u* alternation. This points to a loanword. If U *skalçeta* represents the same lexeme, its initial *s* is not the IE *s* mobile. The appurtenance of Skt. *kaláśa-* is semantically and formally possible, but any archeo-linguistic reality for such a borrowing scenario has yet to be identified (similarly, s.v. *carbasus*).

calpar, -āris ‘vessel, cask, pitcher’

Pre-form: **ka/Hlp-eh₂r-* | PItal. **kalpār-*

Comp.: **ka/h₂lp-id-* | PGk. **kalpid-* | Gk. κάλις, -ιδος ‘jug, urn’

**ka/h₂lp-eh₂-* | PGk. **kalpā-* | Gk. κάλιη ‘pitcher’

**kelp-ur-n-* | PCelt. **kelqurno-* | OIr. *cilorn* ‘pitcher, vessel’, W *celwrn* ‘pail, pitcher, vessel’, Bret. *kelorn* ‘tub’

?Assyrian (Akk.) *karpu, karpatu* ‘vase, pot’

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: vessel

⁸¹ Giacomelli (1994: 36-7, 40) proposes that the vocalic alternation is the result of lower register variation in a population with widespread Greek-Latin diglossia, but this idea has been criticized by e.g. Ruijgh (1986).

⁸² Latin borrows this as *culigna* ‘drinking vessel’ (Biville I: 183). All the changes are expected in an early loan (with Latin borrowing *χ* as *c* and then voicing it to *g* before *n*), such that there is no need to propose an Etruscan intermediary (cf. Breyer 1993: 156-7, who does not rule it out).

⁸³ The variant κολίσκη has a sigmatic element that is difficult to explain.

⁸⁴ EM (87) also compare Skt. *kalikā-* ‘bud’, but KEWA (I: 181) takes it as either a by-form of *kaláśa-* with a *-ka-* suffix or, since it is first found in classical literature, a derivation from *kalā-* ‘small part, a sixteenth part’ and thus not at all related to the family.

Pokorny (555), WH (I: 142), EM (88)

Scheftelowitz (1904: 149), Ernout (1946: 49), Holmes (1947), CAD (K: 219, 221), Untermann (2000: 374), EDG (627), Weiss (2020: 168), van Sluis (fthc.)

Lat. *calpar* does not occur outside of glosses and the grammarians (EM 88), which makes it difficult to ascertain whether it was actually in currency in Latin. WH (I: 142) suggest that it represents **calp-āli-*, an *-ālis* derivative of a form loaned from Gk. κάλπη (EDG 627). The suffix *-ālis* does indeed normally dissimilate to *-āris* when attached to roots with *l* (Weiss 2020: 168). EM (88) alternatively suggest that it came through Oscan, which has *-āri-* derivatives of this shape (cf. *casnar* ‘old man’, Untermann 2000: 374). Following Ernout (1946: 49), EM (88) consider the possibility that it has been transmitted via Etruscan with the plural ending *-ar*, though WH (I: 142) consider it unlikely.

The Etruscan connection is potentially bolstered by the Celtic evidence, with the formation **kelqurno-*. The suffix *-erna* appears in many Etruscan and Etrusco-Latin personal names (Niedermann 1916: 152), and though there are many native sources of a suffix **-erno-* in Latin (Holmes 1947), several Latin words in *-erna/us* are without secure etymology. Similar cases include Lat. *cisterna* ‘cistern’ with an *-erna* element attached from Gk. κύστη ‘vessel’ and Lat. *lanterna* ‘lantern’ with a *-na* element attached from Gk. λαμπτήρ ‘lantern’; in both cases the element is attached for no discernable native Latin reason. Here we see Gk. κάλπη which ends up in Proto-Celtic with what looks like an *-urno-* suffix, pointing to Etruscan mediation into Celtic (cf. van Sluis fthc.). The Latin form exists somewhere along the Wanderwort chain.

Nor are the Greek forms, themselves lacking secure etymology (EDG 627), necessarily the originals. Also compared with varying degrees of security (beginning with Scheftelowitz 1904: 149) are Assyrian *karpu* and *karpatu* ‘pot, earthenware container’ (CAD K: 219, 221). The parallel to the two Greek forms, one with and one without a dental element, is striking. This would make for a Wanderwort with its origins far to the East. Even if only Latin, Greek, and Celtic are compared, the *e ~ a* alternation attested cannot be accounted for in an inherited way.

cant(h)ērius ‘castrated horse; donkey’

Pre-form: **ka/Hnt-eHr-* | PItal. **kantērio-*

Comp.: **g^h/k(a)nd^h-eHl-* | PGk. **kant^hēl-* | Gk. κανθήλιος ‘pack ass’

■ Irreg. correspondences □ Remarkable phonotactics

Semantics: animal, domestic

WH (I: 155), EM (94)

Boisacq (1916: 406 fn. 2), Nehring (1949: 166), Deroy (1956a: 190-1), Furnée (1972: 130, 290), EDG (635)

Lat. *canthērius* would very obviously be a direct borrowing from Gk. καθήλιος, reflecting the theta in spelling as well as the vowel length, but it has an *r*-suffix rather than an *l*-suffix. No Greek forms have an *r*-suffix, and while both suffixes are common in both languages, there is no reason for Latin to replace this suffix in a borrowing from Greek. That καθήλιος originated in Greek (where it is probably a derivation of καθήλια ‘panniers on both sides of the pack-saddle; rafters’ [EDG 635] on comparison with forms like καθία ‘baskets’⁸⁵) is evidenced by the existence of unsuffixed forms like Gk. κάθων ‘pack ass’, καθίς ‘donkey dung’ (though EDG 635 suggests it may be a shortening of καθήλιος akin to *cabō* from *caballus*, hinted at by Boisacq 1918: 406 fn. 2; Nehring 1949: 166 disagrees). Given the Greek pedigree of the forms, regardless of their ultimate origin (likely Mediterranean), *cant(h)ērius* looks very much like Gk. καθήλιος was brought to Latin by a third language in which an *l* ~ *r* alternation occurred (rather than being an independent borrowing in both Latin and Greek as suggested by Boisacq 1918: 406 fn. 2, WH I: 155, EM 94). We would then have to assume that the spelling with *th*, for which there is no native source either way, is a learned spelling.

carbasus ‘fine linen; sail’, var. *carbasa* (pl. of *carbasum*)

Pre-form: *ka/Hrb^(h)-⁸⁶ | PItal. *karb/faSo-

Comp.: *ka/hzrp- | PGk. *karpaso- | Gk. κάρπασος ‘fine flax’

*k^(w)a/orp- | PIIr. *karpaso- | Skt. *karpāsa*- ‘cotton’
 > MoP *karpās* ‘fine linen’
 > Hebr. *karpas* ‘fine garment’

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: textiles

WH (I: 165), EM (99)

Lewy (1895: 126), Cuny (1910: 161), Fohalle (1925: 172-5), Porzig (1927: 272-4), Pisani (1938), Mackenzie (1971: 49), Klein (1987: 287), EWAia (I: 317), Guggenheimer

⁸⁵ Deroz (1956a: 190-1) suggests a derivation from κάνα etc. ‘reed’ (thus reed > basket > pannier). Furnée (1972: 130, 290) further adduces ἀνθήλιον ‘pack-saddle’ and κάθος ‘basket’, assigning it Pre-Greek origin.

⁸⁶ Schrijver (1991: 111-13) suggests that a root shape *CHR(C) was rare. But he provides evidence that for some lexemes it can be reconstructed on the evidence that in some cases, under circumstances that are obscured by its rarity, it metathesized to *CRHC. It is not immediately clear what the sequence *CHR(C) ought to yield in Latin. If vocalization proceeded from the right, one might expect the resonant to vocalize producing *CHorC- > *CorC-. Thus a Latin sequence *CarC-* might require the reconstruction of *a*-vocalism. On the other hand, Schrijver (1991: 114-15) gives two cases where Latin seems to require the reflex of a zero-grade root with internal laryngeal before a resonant for which the result is still *a*-vocalism. These are *cantus* < *khn-to- and *callum* < *khl-no-. The vocalization of the laryngeal would regularly yield *a*, whereas the vocalization of the resonant of *khl-no- ought to yield ***collum*. But Schrijver (1991: 72-3) notes that in word-initial position (*HRC-), the result is almost always *aRC-* (*ursus* must be regarded as irregular). Thus, if the development is the same word-internally, perhaps *CaRC* is the normal reflex of *CHR(C).

(1998), Biville (I: 201-2, 240-2), Shorto (2006: 490), EDG (648), Alves (2022: 32)

Reconstructing any sort of PIE pre-form for the Latin or Greek words is only to demonstrate that they are irregular loanwords, already indicated by the preserved intervocalic *s*. While Shorto (2006: 490) takes Skt. *karpāsa-* ‘cotton’ as a loan from Proto-Austroasiatic **kpaas* ‘cotton’, Alves (2022: 32) notes that cotton production likely post-dates the split of Austroasiatic by a millennium. Thus the Austroasiatic words are borrowings from an Indic source (Sanskrit or Pali). Nevertheless, Skt. *karpāsa-* ‘cotton’⁸⁷ is itself likely a loanword (Cuny 1910: 161, EWAia I: 317). Gk. *κάρπασος* originally referred to high quality linen, and only much later means ‘cotton’ (EDG 648). It is often considered to have come from the Indic word (EWAia I: 317, EDG 648). Lat. *carbasus* ‘fine linen; sail’ must be from this same source, though to consider it directly borrowed from Greek (EDG 648) is not straightforward due to the difference in consonant voicing (Biville I: 240-2). It fits better into the series of words that Latin seems to have borrowed from Greek through an intermediary (cf. *burrus*, *buxus*, see also Fohalle 1925: 172-5).

An additional Greek word, homophonous with *κάρπασος* except in gender and likely also of Mediterranean origin, has entered Latin as well. Columella gives to a poisonous plant the name *carpasum*, which is clearly the same as Gk. *κάρπασον* ‘white hellebore’. An inscription at Pompeii calls the juice of this plant *opocarpasum*, which is Gk. *ὀποκάρπασον*. But Pliny writes *carpathum*, which is not attested in Greek. Only from two Mycenaean women’s names *Ka-pa-si-ja* vs. *Ka-pa-ti-ja* and toponym pairs like *Καρπασία* (Cyprus) vs. *Κάρπαθος* (Aegean island) is it suggested that a form like this must also have existed in Greek (Biville I: 201, EDG 648), though the meanings of these onomastics cannot be substantiated. This would make Greek a likely source for the Latin, but at the same time would require a word of non-IE origin in Greek (EDG 648).⁸⁸ Post-biblical (Mishnaic) Hebrew has *karpas* ‘celery’⁸⁹ (Lewy 1895: 126). Aramaic has *krafsā*, *krefśā* ‘celery’ and Arabic *karafs* ‘celery’. The latter is borrowed from Persian *karafs* (MP *klps* [Mackenzie 1971: 49]), though it is unclear whether the Hebrew and Aramaic forms can be as well (Klein 1987: 287). The (likely) alternation within Greek and the distribution make this look like a Mediterranean substrate word for an herbaceous plant.

EDG (648) keeps the two word families (linen and poisonous plant) strictly separate (cf. also Boisacq 1911-2: 58), though Porzig (1927: 272-4) and Pisani (1938) support the

⁸⁷ The Sanskrit must also be the source of Persian *karpās* ‘fine linen’. The Hebrew hapax *karpas* in Esther 1:6 most likely means ‘fine garment’, and is borrowed from Persian (Klein 1987: 287). In fact, this may explain the shape of *karpas* ‘celery’. Guggenheimer (1998) suggests that the scribe who provided the vocalism for *krps* in the Qaddēsh u-Rechatz was only familiar with the vocalism of *krps* ‘fine linen’ in Esther, and mistakenly wrote them the same way, whereas *krps* ‘celery’ might originally have more closely reflected Persian *karafs*.

⁸⁸ It is unclear if this could have anything to do with the Laconian change from $\theta > \sigma$ that occurred by the 4th century BCE (cf. Allen 1987: 26). This seems too late, and Beekes (2014: 18) considers $\theta \sim \sigma$ alternations as part of the wider Pre-Greek $\tau (\delta, \theta), \tau\tau (\theta\theta) \sim \sigma (\zeta), \sigma\sigma$ alternations.

⁸⁹ Klein (1987: 287) translates ‘parsley’. This might perhaps be influenced by Pesach traditions.

idea that they were originally from the same substrate word. Porzig saw a Mediterranean substrate word (perhaps for plant that was both poisonous and used for its fiber) entering India via an Anatolian source while Pisani theorized about a linguistic substrate that could have existed across the whole area in the Bronze Age. Any archeo-linguistic reality for such a scenario has yet to be identified (see §4.2.2.3).

carpinus ‘hornbeam’

Pre-form: **ka/Hrp-* | PItal. **karpino-*

Comp.: **g^(h)ra/ob^h-r-* | PSlav. **grabrь* | Ru. *grab*, Cz. *habr*, Sln. *gâber*, *grâber*, etc. ‘hornbeam’

?**grab-* | PGk. **grab-* | Gk. γράβιον ‘torch, oakwood’

?**sk(e)rp-* | PBalt. **ske/irp-* | OPr. *skerptus* ‘field elm’, Lith. *skir̃pstas* ‘elm, alder buckthorn, hornbeam, honeysuckle, beech’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, tree

Pokorny (938-47), WH (I: 171-2), EM (101), DV (94)

Alessio (1936: 185), Alessio (1944a: 123-4), Hubschmid (1958: 212), Wagner (1960-4 I: 311), ESSJa (VII: 99-100), Corominas & Pascual (1984-91 I: 856), Puhvel (IV: 99-100), EDG (284), Matasović (2013: 84)

Lat. *carpinus* ‘hornbeam’ is identical in meaning and close in form to descendants of PSlav. **grabrь* (on the form, ESSJa VII: 99-100). In this interpretation, the Slavic descendants dissimilated one or the other of two resonants (with both preserved in Sln. *grâber*). Perhaps *carpinus* could have originated via dissimilation from **crarpinus*,⁹⁰ or the second resonant in Slavic could be a suffix. But both explanations still require the assumption of metathesis to arrive at matching pre-forms. OPr. has *wosi-grabis* ‘European spindle(tree)’, which is semantically removed and thus of unclear appurtenance.

I am not convinced by Pokorny (938-47), WH (I: 171-2), and DV (94) who would rather derive *carpinus* from *(s)*ker-* ‘to cut’ because of the serrated leaves of the hornbeam even though the Slavic forms are a perfect semantic match. They, along with EM (101) instead compare OPr. *skerptus* ‘field elm’ and Lith. *skir̃pstas* ‘elm, alder buckthorn, hornbeam, honeysuckle, beech’. Smoczyński (2018: 1200) does not commit, but notes the *-st-* of the Lithuanian forms could be the result of reanalysis of a frequentative formation of a verb **skirp-*. Nor is he certain about Lith. *skrōblas* ‘hornbeam’, another perfect semantic match, as the cognates are formally diverse.

⁹⁰ Lat. *pōrtiō* ‘degree, portion’ seems to be derived via dissimilation from **prō* *ratione* (DV 448) and Lat. *prōcērus* ‘tall, lofty’ is perhaps from **prō* + **krēros* < **kreh₁-ro-* (DV 491).

WH (I: 171-2) go further to suggest that Gk. γράβιον ‘torch, oakwood’ is related via the “Reimwurzel” **greb^h*- in γράφω ‘to write’, a suggestion which can be ignored along with the connection to U *Graboui*, an epithet of Jupiter. It is interesting that, though they reject the connection with Slavic, they consider γράβιον a Macedonian/Illyrian word. More interesting is that EDG (284) agrees, though he excludes *carpinus* and connects the Greek and Slavic words. He follows Furnée’s (1972: 169) connection with Hsch. γοβρία: φανοί, λαμπτήρες ‘bright, torches’ to conclude that the family is of non-IE Balkan substrate origin, with the Modern Greek forms γράβος (Epirus), γάβρος (Arcadia) ‘hornbeam’ continuing this word. But the Modern Greek words are probably loans from Slavic.

Matasović (2013: 84) accepts only the words that mean ‘hornbeam’, i.e. the Latin, Baltic, and Slavic. This approach makes the most sense (though the Baltic forms are formally the most distant and semantically broad; thus their appurtenance is uncertain). Alessio (1936: 185, 1944a: 123-4) had done the same, though he went too far in deriving them from a substrate root **karra* ‘rock’ and further comparing Gk. κάρφος ‘small dry body’. It is this semantic distance which removes Gk. γράβιον ‘torch, oakwood’ from consideration. Interestingly, there are several Romance (and Basque) forms for ‘branch’ and ‘firewood’ and even ‘oak’ that are similar to one or both of these Greek words: (Nuorese *kárva* ‘branch’, Asturian *garbu* ‘small firewood’, Basque *karbasta* ‘stick with branches’, Wagner 1960-4 I: 311; Sp. *carba* ‘scrubland full of coarse oak trees’, Pt. *carvalho* ‘oak’, Corominas and Pascual 1984-91 I: 856; etc.). But none is likely related to *carpinus* (cf. Hubschmid 1958: 212). We are left with irregular voicing (and perhaps aspiration) alternations in a root with likely original *a*-vocalism present in Latin and Slavic.

Hitt. *karpina*- ‘a tree/shrub’ has also been compared (Puhvel IV: 99-100, not necessarily in an inherited way), but its meaning is too vague to ascertain its appurtenance.

caulae ‘railing or lattice barrier, sheepfold; pores of the skin’

Pre-form: **ka/Hg^h-e/o/ul-* | PItal. *kaxe/o/ulā*

Comp.: **kog^h*- | PItal. **koχom* | Lat. *cohūm* ‘part of a yoke’, *incohāre* ‘to begin’

**ka/o/Hg^h*- | PGm. **haga(n)*- | OHG *hag* ‘hedge, fence’, OE *haga* ‘fence’

**ka/Hg^(h)*- | PCelt. **kagyo-* | Gaul. *caii* ‘fence’, W *cae* ‘fence, collar’

□ Irreg. correspondences

■ Remarkable phonotactics

Semantics: tool

Pokorny (518), WH (I: 187-8, 243-4), EM (107, 131), DV (99, 123, 126)

Thurneysen (1887: 155-6), Schrijver (1991: 141, 462), EWAia (I: 288), Untermann (2000: 362, 380-1), Matasović (2009: 184), Kroonen (2013: 198), van Sluis, Jørgensen &

Kroonen (2023: 216)

Latin *caulae* is reconstructed to an original **kaχulā-* (Pokorny 518, WH I: 187, EM 107, DV 99) because of its close semantic match with Germanic **haga(n)-*, etc. (Kroonen 2013: 198) and Celtic **kagyo-* (Matasović 2009: 184) of the same root shape. All go back to what could be reconstructed as **kHg^h-*, but the invalid **TeD^h* root structure points to *a*-vocalism.

Lat. *cohūm* ‘part of a yoke’ is poorly attested outside of glosses, and its meaning is not completely understood, interpreted from accounts by Varro and Paulus *ex Festo* (Schrijver 1991: 141). Varro, followed at least in the beginning by Thurneysen (1887: 155-6) and by EM (131), favored a relation to *cavus* through a form like **coūm*. WH (I: 243-4) reject this in favor of deriving it from the same root as *caulae*. DV (123) as well supports the connection with *caulae*, and champions the connection with *incohāre* ‘to begin’ as developing from ‘to start work’ < **‘to yoke a plow to a team of oxen’*. This would be from **kog^h-*, pointing to an *a ~ o* alternation (since **kHg^h- ~ *kHog^h-* is still of an invalid structure). Several Sabellic words including Osc. **kahad** [3sg.pres.subj.] ‘to take?’ (Untermann 2000: 362 with lit.), U *cehefi* [pass.inf.], and U **kukehes** [*com* + 2/3sg.fut.] ‘to take/get?’ (Untermann 2000: 380-1 with lit.) of relatively uncertain semantics also reconstruct to **kalog^h-*. Widely considered related to *cohūm/incohāre*, van Sluis, Jørgensen & Kroonen (2023: 216) support a connection with the Celtic and Germanic words.

Interestingly, EWAia (I: 288) argues that Skt. *kākṣa-* ‘bush, scrub’ is to be separated from *kākṣa-* ‘armpit’, leaving it to be potentially connected with this family. The semantics are not extremely far off if we consider either narrowing from ‘bush’ > ‘hedge’ > ‘fence’ or a broadening in the opposite direction. But **g^h-s-* is by far not the only source of Skt. *kṣ*. Additionally, the potential appurtenance of Alb. *thanë* ‘cornel; winter stall for sheep’ < **kald(C)-neh₂* (van Sluis, Jørgensen & Kroonen 2023: 216) to the Italic, Celtic, and Germanic family would rule out the connection with Sanskrit, as it cannot reflect **k̄*. The inclusion of one excludes the other, but in fact neither is semantically close enough to compare to the Italo-Celto-Germanic group with certainty.

caupō ‘trader, huckster; innkeeper’

Pre-form: **kh₂eup-/ *keh₂up-* | PItal. **kaupōn-*

Comp.: **kh₂(e)p-* | PGk. **kap-* | Gk. *κάπηλος* ‘huckster, innkeeper’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: economic

WH (I: 189), EM (107), DV (100)

Ernout (1946: 42-3), Nehring (1949: 165), Furnée (1972: 115-6, 257-8), Puhvel (III: 125-7), Bonfante (1985: 207), Breyer (1993: 507-13), Kloekhorst (2007: 295), EDG (638), Oettinger (2021: 120-2)

Despite the circulation of the forms *cōpō*, *cōpa*, and *cūpō*, the borrowing of this word into Proto-Germanic resulting in OHG *kaufōn* ‘to buy’ etc. establishes the form with the diphthong as original rather than a hyper-urbanism of *ō* (WH 189).

It is often considered a Mediterranean substrate word based on its isolation beyond Gk. *κάπηλος*, likewise without etymology and of precisely the same semantics but aberrant in phonology and morphology (WH 189, EM 107, DV 100). A derivation of the Greek from *κάπη* ‘crib’ with the assumption that this originally meant ‘chest’ and therefore ‘one who sells from a chest’ is unlikely (WH 189, EDG 638). Furnée (1972: 115-6) includes *κάπηλος* as one of several Greek words with the Pre-Greek suffix *-ηλος* while the Latin form is a *praedō*-type *n*-stem, which is otherwise easily Indo-European. There is no strong evidence that the word entered Latin via Etruscan: neither in its *n*-stem (Bonfante 1985: 207, *pace* Ernout 1946: 42-3, Breyer 1993: 507-13) nor the *a* ~ *au* alternation (Breyer 1993: 251, *pace* WH 189).⁹¹ Etruscan lookalikes (*caupis*, *caupnal*, *caupne*) are onomastic and of unknown meaning.

Furnée (1972: 257-8) hesitantly compares Hitt. *happar-* ‘business, payment, price’ (inherited from **h₃ép-r-* < **h₃ep-* ‘work’, Kloekhorst 2007: 295) on the uncertain grounds that cuneiform *h* may correspond to Gr. *κ* in some potential loans. Puhvel (III: 127) accepts the connection with Hittite, which makes the Latin and Greek words “Mediterranean” (his scare quotes) loans of Anatolian origin. Oettinger (2021: 120-2) argues that Gk. *κάπηλος* was directly borrowed from an unattested Lydian **kapala-* ‘merchant’, from the same root as Hitt. *happar-*. Interestingly, this would match Nehring’s (1949: 165) argument that Lat. *cabō* ~ *caballus* is of ultimately Anatolian origin due to the vacillation there in some names between *-αλος* and *-ων* (cf. thus also Gk. *κάνθων* ~ *κανθήλιος* ‘pack ass’). I am not fully convinced by Oettinger’s evidence of the ability to reconstruct a Lydian form from the Greek.⁹² Nor is he able to explain the *au* of Lat. *caupō*. If the root is indeed ultimately Anatolian, the word reached at least Latin through (an) intermediary language(s), resulting in the difference in vocalism. PItal. **kaupōn-* against PGk. **kapēl-* at first blush seem to attest to an *n* ~ *l* alternation, but they may instead be two different suffixes; the alternation between **-ōn-* and **-āl(l)-* seems to have a few parallels. In any case, *caupō* remains a foreign word in Latin whose direct donor(s) is/are unclear.

citrus ‘citron (*Citrus medica*); arar, sandarac gum/Sictus tree (*Tetraclinis articulata*)⁹³’

Pre-form: **kit-ro-* | PItal. **kitro-*

⁹¹ Etruscan attests the monophthongization of *au* > *a*, but would not produce the reverse (Breyer 1993: 251).

⁹² Greek has *Γύγης* from Lyd. *Kuka-*. The first *γ* for *k* is explained as being due to the neutralization between voiced and unvoiced stops in word-initial position leading to a variance in pronunciation. The second *γ* for *k* is explained via assimilation. But with so few secure examples of Greek loans from Lydian, it feels perilous to explain the correspondences away so quickly.

⁹³ At different times formerly assigned to the genera *Thuja* (thus sometimes called the ‘thuja’) and *Callitris*.

Comp.: **ked-ro-* | PGk. **kedro-* | Gk. κέδρος ‘cedar, juniper’

■ Irreg. correspondences □ Remarkable phonotactics

Semantics: plant, tree; fruit

WH (I: 223-4), EM (123-4), DV (116)

Fohalle (1925: 166-70), Battisti (1960: 375), Leumann (1977: 198), Hamp (1978: 185-95), Biville (I: 223-4), Breyer (1993: 188), EDG (663), Weiss (2020: 192)

Lat. *citrus* means both the fruit-bearing ‘citron’ and the aromatic-wooded ‘arar’. While clearly related to Gk. κέδρος ‘cedar, juniper’, a direct loan is unlikely due to the differences in phonology and semantics. In contrast, Lat. *cedrus* ‘cedar, juniper’ is a direct loan in form and meaning from Gk. κέδρος, while Gk. κίτρος ‘citron’ is a direct loan from Latin *citrus* (Biville I: 223-4, EDG 663).

It is often accepted that Lat. *citrus* was taken from Gk. κέδρος through Etruscan mediation (WH I: 223-4, EM 123-4). This is not beyond questioning however, given the lack of an attested Etruscan form. The devoicing of **-dr-* > *-tr-* in Latin may have an internal explanation, though beyond the near total lack of the sequence *dr* in Latin, there are few secure examples of the devoicing.⁹⁴ While Etruscan sometimes attests *-l-* from Gk. ε (cf. Etr. *Elina* < Ἐλίνα), this is not attested in an initial syllable (Biville I: 223-4, Breyer 1993: 188). Given that there is nothing about the words that requires the direction of borrowing to have gone from Greek to Latin (via an intermediary), both Latin and Greek may well have borrowed the word from another language of the Mediterranean (Fohalle 1925: 166-70, Battisti 1960: 375, Biville I: 224, DV 116).

columba ‘pigeon, dove’

Pre-form: **ke/ol-o/umb*^(h)- | PItal. **ke/olo/umb/fā-*

Comp.: **g*^(h)*ol-omb*^(h)- | PSlav. **gǫlǫbъ-* | OCS *golǫbъ* ‘pigeon, dove’, etc.

**gul-ub*^h- | PGm. **kulubrōn-* | OE *culfre*, *culfure* ‘dove’

Copt. ⲉⲣⲟⲟⲙⲡⲉ /*kjroompe*/ ‘dove’ < Egypt. *gr-n-p.t* ‘dove’

?**kol-* | PArm. **salámba-* | Arm. *salamb* ‘partridge, francolin’

??**kol-umb-* | PGk. **kolumbo-* | Gk. κόλυμβος ‘little grebe’⁹⁵

⁹⁴ Hamp (1978: 185-95) discusses this at length, proposing examples in word-initial position. Word-internally, the best example is *uter* ‘wineskin, water bottle’ < **udris* (cf. Gk. ὕδρια ‘water jar’) the water word, attested otherwise only in *unda* ‘wave’. *Lutra* ‘otter’ on comparison with e.g. Skt. *udrá-* ‘otter’ is also probably an example from the water root, but the initial *l-* is of unknown origin. Leumann (1977: 198) gives *taeter* ‘foul, disgusting’ < **taidro-*, cf. *taedet* ‘s/he is tired/disgusted’, but the latter is without etymology. The element *quadru-* seems to show the opposite phenomenon, but Weiss (2020: 192) suggests it is the reflex of the cluster **h₁r*.

⁹⁵ Often given as *Podiceps minor*, but this is not a recognized species name and should be amended to *Tachybaptus ruficollis* (cf. also Battisti 2021: 211).

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: animal, bird

Pokorny (429-34, 547-8), WH (I: 249), EM (134), DV (126)

Skeat (1888: 146), Brugmann (1906: 386-7), Oštir (1921: 49), Worrell (1934: 67), Frisk (1960-72 I: 906), Erman and Grapow (1971 V: 181), Furnée (1972: 170), Lockwood (1990: 262), Vycichl (1990: 249), Schrijver (1991: 375), Biville (II: 265), Johnston & Janiga (1995: 6), Peust (1999: 280), Derksen (2007: 175), EDG (741), Martirosyan (2010: 565), Neri (2016: 14), Allen (2020: 115), Batisti (2021), Jakob (fthc.)

Lat. fem. *columba* is the primary form to which masc. *columbus* is a secondary derivation (WH I: 249, Schrijver 1991: 375, EM 134). Though remarkably similar to Gk. κόλυμβος ‘little grebe’, the latter’s semantic mismatch in light of all other comparanda makes it difficult to compare with any certainty. Frisk (1960-72 I: 906) notes that the *v* of Greek makes it difficult to reconstruct a proto-form that is not a “lautliches Unding” from an inherited perspective,⁹⁶ and Biville (II: 265) suggests that the large semantic distance shows that neither is borrowed from the other. Even if the verb κολυμβάω ‘to dive, submerge, jump into the water, swim’, whose variants like κολυμφάω attest to features that Furnée (1972: 170) and EDG (741) take as evidence of a Pre-Greek origin (a $\varphi \sim \beta$ alternation and -υμβ- suffix), is denominal from κόλυμβος (and therefore indicates that κόλυμβος is also Pre-Greek), it only further points to the original meaning of the Greek word being something like ‘diver’ and to the similarity with *columba* being coincidental.

The Slavic evidence provides the first phonological peculiarity indicative of a non-IE origin in that it requires the reconstruction of voiced initial *g. There are Baltic relatives as well (WH I: 249, EM 134, DV 126, Derksen 2007: 175) but there the avian meanings are lacking: Lith. *gelumbė* ‘cloth’ and OPr. *golimban* ‘blue’. Within Slavic, Derksen (2007: 175) reconstructs **golobъ* ‘blue’ (cf. Ru. *golubój* ‘pale blue’, SCR. *golùbjiĭ* ‘dove-colored, blue-gray’, Slov. *golôbji* ‘dove-’) and widely-attested **gòlobъ* ‘pigeon, dove’ (cf. OCS *golobъ*, Cz. *holub*, Slov. *golôb* ‘pigeon, dove’). Since it is the color meaning that is found in both branches, it is sometimes assumed (cf. WH I 249, EM 134, DV 126) that the dove meaning has been derived from the meaning ‘blue/gray’.⁹⁷ But Lockwood (1990: 262) and Derksen (2007: 175) argue that it is instead the avian meaning that is primary in Balto-Slavic, with the color meanings being derived from it. The attestation of avian meanings outside of Balto-Slavic strengthens this idea wherein

⁹⁶ Neri (2016: 14) proposes the effect of Cowgill’s Law, but Batisti (2021: 212) doubts that it occurred in the given phonetic environment.

⁹⁷ This is similar to an alternative explanation for Lat. *columba*. Gk. κελαινός ‘dark, black’ could reconstruct to **kel-Vn-ios*, attesting to an *n*-stem shared with Lat. *columba* < **kol-on-bʰā* (EDG 667) and with further links via an *aniĭ*-root **kel-* (Schrijver 1991: 427), cf. Lat. *calidus* ‘with a white spot on the forehead’, Lith. *kalyb/vas* ‘dog with a white spot on the neck’, Swiss German *helm* ‘spot on the forehead of cattle’, potentially Skt. *karkī-* ‘white cow’, etc. But this explanation is doubted by Frisk (1960-72 I: 906) and Furnée (1972: 170) and would still require the Balto-Slavic color words to be unrelated.

Balto-Slavic innovated the adjective ‘dove-colored’ and the meaning ‘pigeon’ was lost in Baltic after the split. To explain the *g ~ *k alternation, Lockwood (1990: 262) proposes that Latin has undergone taboo deformation to *kol-. This seems unlikely. Batisti (2021: 206-7) proposes parallel constructions with the color suffix *-b^ho- on two different roots: Latin from the root behind Lat. *color*⁹⁸ and Slavic from *g^hleh_r- ‘glow, be bright’. But this too seems unlikely given at least one further dove word of similar yet irreconcilable shape (in Germanic). We would have to assume several independent and coincidentally very similar formations.

Jakob (fthc.) adduces OE *cul(u)fre* ‘dove’ (Engl. dial. *culver* ‘wood-pigeon’) as an irregular comparandum. The traditional explanation (cf. Skeat 1888: 146) is a loan from Lat. *columba*, but the nasal loss in a process like this would be unexplained. Taken at face value, the lack of the nasal suggests this is not an example of one of Brugmann’s (1906: 386-7) inherited *-ŋ-b^ho- morphemes (cf. Batisti’s 2021 explanation) but rather an irregular alternation. The consonantal alternations within this family of comparanda is enough to show that it is not of IE origin.

Whatever source it originated from seems to have given it to New Egyptian as well. Oštir (1921: 49) noted the similarity to Coptic ⲉⲣⲟⲟⲙⲡⲉ /kʲroompe/ ‘dove’ < Egypt. *gr-n-p.t* ‘dove’. The Egyptian word is attested from the New Kingdom (1300-1075 BCE), and looks transparently like a compound *gr* ‘bird’ *n* ‘of’ *p.t* ‘the sky’. Peust (1999: 280) takes this at face value and suggests that Egyptian is the source of the European forms. Worrell (1934: 67) had already hinted that the European and Egyptian forms were independent borrowings, and Vycichl (1990: 249) reasonably suggests that this spelling ‘bird of the sky’ is a play on words, something akin to a folk etymology. The variation in spellings listed by Erman and Grapow (1971 V: 181) indeed seems to point to this. It appears not only with the genitival *n* but also with *m* (suggesting that it was not originally the genitival element at all) and both with the expected initial element *gr* ‘bird’ but also the homophonous *gr* ‘to become still/silent’⁹⁹. This together with its appearance in the New Kingdom suggests a loan in Egyptian. At the time of the New Kingdom, the word written *gr-n-p.t* would have been pronounced something like /kʲVránipV/ (Allen 2020: 115). Since the *n* might be folk etymological and is also spelled with *m*, amending that to /kʲVrámpV/ (cf. Jakob fthc.) means that the form looked strikingly similar to the European comparanda even before it developed into Coptic /kʲroompe/.¹⁰⁰ The Coptic evidence proves that the *r* is real and not a spelling for *l* (cf. Loprieno 1995: 31), which

⁹⁸ He follows an analysis by Höfler (2015) that proposes Lat. *color* is from a root *k^huel- ‘dark, black’, but I do not find that the etymologies that Höfler proposes (Arm. *šatax* ‘clay, mud, mortar’, Gk. πηλός ‘clay, earth, mud’ [with the length unexplained] < *k^huel-; Gk. κόλα ‘the parts under the eyes’, Lat. *culex* ‘gnat’ [the plural *culices* that seems to refer to floaters in the vitreous humor of the eye should be compared to their name e.g. in Fr. *mouches volantes*] < *k^hul-o-) to be more convincing than the traditional etymology of *color* < *kel- ‘to cover’ (cf. the parallel Ved. *várṇa-* ‘color’ alongside ‘covering’).

⁹⁹ Spelled in full <g - r - MAN WITH HAND TO MOUTH> (Gardiner’s W11-D21-A2), so we know it is this verb.

¹⁰⁰ ⲉⲣⲟⲟⲙⲡⲉ is not the only Coptic spelling; different dialects have *o*, *a*, and *aa*. This points to an original **a* in the Egyptian parent form (Allen 2020).

along with its appearance already before 1000 BCE proves it is not borrowed from any of the forms we have attested. Instead both the European and the Egyptian form were borrowed from a third source. Notably, the rock dove (*Columba livia*) was domesticated in the eastern Mediterranean between 5,000 and 10,000 years ago (Johnston & Janiga 1995: 6).

Martirosyan (2009: 565) adduces Arm. *salamb* ‘partridge or francolin’, via PArm. **salámba-* as if < **kól(o)mb^h-(e)h₂-*, arguing that this is a Mediterranean word. The initial palatal needed for Armenian further removes the possibility of a link with explicitly unpalatalized (required by the Lithuanian and Sanskrit reflexes) **kel-* (see fn. 97) and also discredits **gol-* as the original root. For a similar suffix in a columbid bird, s.v. *palumbēs* ‘wood-pigeon’. Though like Gk. κόλυμβος, the semantic difference makes this a less certain comparandum.

cotōneum ‘quince’

Pre-form: **kot-ōn-ejo-* | PItal. **kotōnejo-*

Comp.: **kud-ōn-ih₂* | PGk. **kudōnia-* | Gk. κυδώνια (μᾶλα) ‘quinces’
**kod-u-* | PG. **kodu-* | Gk. κοδύμαλον ‘quince’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, tree; fruit

WH (I: 281), EM (146)

Solmsen (1911: 241-5), Nehring (1923), Fohalle (1925: 170-1), Berger (1956: 8-13), Battisti (1960: 380-1), Biville (I: 225-8), Breyer (1993: 189), EDG (797), Beekes (2014: 60)

Lat. *cotōneum* ‘quince’, often called *mālum cotōneum*, is often proposed to have been borrowed/calqued from Greek μήλον κυδώνιον, as if ‘Cydonian apple’, with Etruscan mediation potentially explaining the change from *v* > *o* and the devoicing of *δ* (Solmsen 1911: 243, WH I: 281, EM 146, Breyer 1993: 189). However, the Greek word seems to have been connected with the Cretan city of Κυδωνία by folk etymology only (Solmsen 1911: 242, Fohalle 1925: 170-1, EDG 797, Beekes 2014: 60). The preservation in Alcman of κοδύ-μαλον¹⁰¹ seems to attest the older, original form. Biville (I: 227-8) and EDG (797) consider it to be specifically Anatolian, but this seems to be based on toponymic evidence (cf. also Nehring 1923).¹⁰² In any case, this is crucial evidence that

¹⁰¹ Hesychius also has κοδώνεα. His definition of the term as σῦκα χειμερινά. καὶ καρῶν εἶδος Περσικῶν ‘winter figs; a kind of Persian nut’ is argued by EDG (797) to have been based on confusion with κόττανον ‘small kind of fig’. This is not related, at least not in any close way, and is of Semitic origin (Solmsen 1911: 242).

¹⁰² Solmsen (1911) and Biville (I: 227-8) also propose the word might be Lydian, since the Etruscans are purported to have come from Lydia; this after Biville just discussed the problems with the theory of Etruscan mediation for this word. There are many problems with this analysis. Nehring (1923) argues that neither the *o* nor the *t* is proof of Etruscan mediation because they could also be from a language of Asia

strongly suggests that Latin and Greek have borrowed the quince word independently of one another, perhaps from an Anatolian language,¹⁰³ but otherwise from an unknown source (WH I: 281, Biville I: 228, EM 146, EDG 797, Beekes 2014: 60). Berger (1956: 8-13) followed by Battisti (1960: 380-1) compares Burushaski *jaṭúr/jaṭór*, purportedly reconstructible to **koḍú-ur*, where the suffix *-ur* is common in other plant names. But the changes from **k > j* and **ḍ > t* are not well understood and may be without parallel.

cucumis ‘Armenian cucumber/snake melon’

Pre-form: **ku-kum-es-* | PItal. **kukumes-*

Comp.: **ku-ku-* | PGk. **kuku-* | Hsch. κύκκον· τὸν σικκόν ‘cucumber’, κυκῦζα· γλυκεῖα κολόκυντα ‘sweet round gourd’

**tik-u-* / **t/kjik-u-* | PGk. **tiku-* / **t/kyiku-* | Gk. σικύα ‘bottle gourd’

**t/kjek-u-* | PGk. **t/kyeku-* | Hsch. σεκούα· σικύα ‘bottle gourd’

**kék(ʷ)-* | PArm. **sekʰ-* | Arm. *sex* ‘muskmelon’

**tūkū-* | PSlav. **tyky-* | Ru. *tykva* ‘gourd’¹⁰⁴

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: plant, domestic

WH (I: 299-300), EM (154), DV (148)

Alessio (1944a: 109-10), Alessio (1946b: 36), Neumann (1971a: 265), Furnée (1972: 243, 251), André (1978: 49), Puhvel (IV: 250-1), RLA (X: 20), Jannick, Paris, and Parrish (2007), Martirosyan (2009: 574), EDG (1330), Sebastian, Schaefer, Telford & Renner (2010), Kogan (2011: 203), PSD (s.v. *tikil*, *ukuš*, *ukuštikil*)

Lat. *cucumis* and its comparanda originally referred to various cultivars of the muskmelon *Cucumis melo*, most likely the non-sweet snake melons (Jannick, Paris, and Parrish 2007). The Latin and some of the Greek forms look like they might be reduplicated (André 1978: 49),¹⁰⁵ but this does not explain the other variants.

In fact, there are several peculiarities in this family of comparanda that cannot be explained from a native IE perspective. Within Greek, there exists *i ~ e ~ u* alternation in σικύα, σεκούα, and σικύα (EDG 1330). Arm. *sex* might preserve an initial **s*, which would be irregular, and could reflect a final unvoiced aspirate **kʰ* (Martirosyan 2009: 574). However given Hsch. κύκκον ‘cucumber’, both the Greek and Armenian forms with irregular initial *s* could have been borrowed from a source starting in **kj*

Minor. While this is not provable, it shows that Etruscan is not the only explanation for the changes.

¹⁰³ It is unclear if this is then assumed to be a non-IE language of Anatolian, a non-IE word in an IE Anatolian language, or an unattested inherited word.

¹⁰⁴ Some, like Alessio (1946b: 33-43) have suggested that this specifically, and in fact the root more generally, is the same as in *ficus* ‘fig’.

¹⁰⁵ André suggested it was perhaps due to the shape and volume of the vegetable.

(paralleling PIE **k̑*- > Arm. *s*-). Less likely is Furnée's (1972: 251) suggestion (cf. also Alessio 1946b: 36), based on the shape of Slavic **tyky*-, that the first *k* of Latin and some of the Greek forms might be due to assimilation of an original **t* to the second *k*. If Gk. σικύα etc. are from *σικύφα, it might be evidence of an *m* ~ *w* alternation with *cucumis*. Alessio (1944a: 109-10) proposes that the *-mo-* suffix is Mediterranean, more specifically Tyrrhenian, but I am skeptical of this. It might otherwise be related to the suffix of Lat. *racēmus* (s.v.).

Semitic forms like Ge'ez *k'äsya*, Akk. *qiššû*, Hebr. *qiššû* 'cucumber' (cf. EDG 1330) can only be related if we assume metathesis of the sibilant and velar elements.¹⁰⁶ Neumann (1971a: 265) suggested that Hitt. *kunkumati-* is a reflex of this culture word, with which Puhvel (IV: 250-1) agrees. Its meaning cannot be determined beyond the name of a plant, perhaps a vegetable, so that it cannot be adduced with certainty.

cupressus 'cypress'

Pre-form: **kup-Vr-et-to-* | PItal. **kup(V)resso-*

Comp.: **kup-ar-it-jo-* | PGk. **kuparisso-* | Gk. κηάρισσος 'cypress'
**kub^h-ar-it-ino-* | PGk. **kup^harissino-* | κηαρίσσινος 'made of
cypress'

Hebr. *gofer* 'gopher wood'

■ Irreg. correspondences □ Remarkable phonotactics

Semantics: plant, tree

WH (I: 313), EM (159)

Strong (1890 no. 1614), Fraenkel (1886: 153), Brown-Driver-Briggs (172), Cuny (1910: 162), Zimmern (1915: 53), Pisani (1938b), Ernout (1946: 36), CAD (K: 178, 333, 553), Furnée (1972: 159-60), Klein (1986: 284), Breyer (1993: 198), Biville (II: 146), EDG (803), Weiss (2020: 507)

Lat. *cupressus* and Gk. κηάρισσος are clearly related, but the relationship is not one of direct borrowing. The syncope of the stressed Greek *α* is expected if the word entered Latin before the shift to initial accentuation and syncope,¹⁰⁷ however there is no way to

¹⁰⁶ RLA (X: 20) equates Akk. *qiššû* with Sum. *ukuš* 'member of Cucurbitaceae' (cf. PSD s.v. *ukuš* 'cucumber') implying but not explicitly stating a loan from Sumerian. The word occurs in a compound *ukuštikil* 'colocynth, the bitter cucumber' (PSD s.v. *ukuštikil*), the second element of which is *tikil* 'pointed' (PSD s.v. *tikil*). This looks close to the potential reconstruction **TVkV-* for the IE comparanda. But it is almost certainly coincidence, especially given the more likely reconstruction **k(j)Vku-* for the IE forms. Furthermore, Kogan (2011: 203) reconstructs PSem. **kV_l(l)V-* for the Semitic forms. This would mean at best a loan into both Proto-Semitic and Sumerian independently. Given that the wild progenitor of *Cucumis melo* occurs in India and that both this melon and the cucumber (*C. sativus*) are likely of Asian origin (Sebastian, Schaefer, Telford & Renner 2010), a Wanderwort that left a trace in Sumerian would not be unexpected. However, the large phonological distance between the forms makes them difficult to connect, even if we assume they must have traveled over large geographical distances.

¹⁰⁷ The shift to initial accent in this word can have occurred within Latin, or it could have entered Latin

get Lat. *e* from Gr. ι in this position (Biville II: 146). EM (159) and WH (313) consider both independent borrowings from a Mediterranean substrate. The *-issos* ending of the Greek form is considered an example of a Pre-Greek suffix (cf. EDG 803), and Lat. *cupressus* seems to be the only example of this Pre-Greek suffix in a Latin word not borrowed directly from Greek.¹⁰⁸ The adjective κυπαρίσσινος ‘made of cypress’ demonstrates the existence of a variant with an aspirate (Furnée 1972: 159-60), further suggesting that the word is truly at home in Greece. Thus it is particularly tempting to see the Latin form as a borrowing (possibly through some intermediary) from Pre-Greek. That the intermediary could have been Etruscan however, as Ernout (1946: 36) believes and WH follow, is unlikely. There are no attested Etruscan words of this shape (Breyer 1993: 198; Biville II: 146) and Breyer further notes that Etruscan mediation does not solve the problem of Latin *e* for Gk. ι.

Latin has either received this word from Pre-Greek, from Pre-Greek through an intermediary, or from Greek through an intermediary. But in any case, its most proximal source is unknown.

A form of the word without the Pre-Greek suffix seems to have existed in the Mediterranean region, where it was borrowed into Hebrew as the hapax¹⁰⁹ in Genesis 6:14 *gofer*, the wood used to build the ark, thus often simply translated as ‘gopher wood’ (WH I 313, Furnée 1972: 160, as recently as Weiss 2020: 507).¹¹⁰

ervum ‘bitter vetch’

Pre-form: **h₁er(H/-V-)u_o*- | PItal. **er(V)wo*-

Comp.: **h₁orh₃-bo*- | PGk. **orobo*- | Gk. ὄροβος ‘bitter vetch’
 **h₁erh₁-bind^ho*- | PGk. **erebint^ho*- | Gk. ἐρέβινθος ‘chickpea’
 >> OGeorg. *erbindi*, Georg. *erevindi* ‘pea’

**h₁or-u-īd*- | PGm. **arwīt*- | ON *ertr* ‘peas’, OHG *arawīz* ‘pea’, etc.

**h₁orVb*- | PArm. **ari/uw*- | Arm. *arowoyt* ‘alfalfa’

with the shift having already taken place. In any case, we have a *terminus ante quem* of the 3rd century BCE for its borrowing. Alternatively, the syncope could have occurred in the donor language.

¹⁰⁸ Lat. *platessa* ‘flatfish’ does not seem to occur in Greek, but has the *-essa* suffix built on Gk. πλατύς ‘broad, flat’, which has no Latin cognate.

¹⁰⁹ Despite *gofer* being a hapax, Strong (1890 no. 1614) and Brown-Driver-Briggs (172) take the word *gofriyt* ‘brimstone’ (Mod.Hebr. *gofriyt* ‘sulfur’) as derived from it (the latter follow the assumption that *gofer* is a misreading of *kōfer* ‘pitch’, therefore ‘pitched wood’). However in light of Arab. *kibrīt* ‘sulfur’, a loan from Aramaic *kībrīā* ‘sulfur’ (Fraenkel 1886: 153) with a cognate in Akkadian *ki/ubrītu* ‘sulfur’ (CAD K: 333), this must be an unrelated family of words.

¹¹⁰ Cuny (1910: 162) additionally compares Hebr. *kōfer* ‘pitch, tar; henna’, but this is going too far. It has Semitic cognates in e.g. Akk. *kupru* ‘bitumen’ and *kupāru* ‘to smear on; to wipe off’ (CAD K: 178, 553, Klein 1987: 284). Interestingly, Hebr. *kōfer* seems to have formed a verb *kfr*, which occurs as a hapax—where else but Genesis 6:14 (Klein 1986: 284). Pisani (1938b) compared MoP *sarw* ‘cypress’, but I see no need to reject the alternative etymology as a loan from Semitic (cf. Zimmern 1915: 53).

?*Hreb^(h)-e/ont/d^(h)- | PIIr. *Hrab^(h)ant/d^(h)- | Rushani *ravand* ‘wild chickpea (*Cicer songaricum*),’ etc.

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: plant, domestic

Pokorny (429-34, 547-8), WH (I: 419-20), EM (202), DV (195)

Alessio (1944b: 410), Thurneysen (1946: 175), Hubschmid (1955: 238-45), Mayrhofer (KEWA I: 48), Furnée (1972: 198), Corominas & Pascual (1984-91 III: 85), Puhvel (V: 134), Schrijver (1991: 36, 423), Mayrhofer (EWAia III: 13), Matasović (2009: 40), EDG (451), Kroonen (2013: 37), EDIL (s.v. 1 *arbor*), Kroonen (fthc.), Thorsø (fthc.)

Gk. ἐρέβινθος ‘chickpea’ shows the textbook Pre-Greek suffix -ινθος (EM 202, EDG 451).¹¹¹ Gk. ὄροβος ‘bitter vetch’ is clearly a related form, but the vocalism of the second syllable cannot be reconstructed to the same pre-form and it lacks the -ινθος suffix. Lat. *ervum* ‘bitter vetch’ with the same meaning as ὄροβος cannot be separated as a comparandum, but contributes to the impossibility of constructing a common pre-form. The Greek and Latin forms technically allow the reconstruction of *g^w, but PGM. *arwīt- ‘pea’,¹¹² which must also be the same etymon, allows only *w. Thus a labial element in Greek and Latin are most likely. In this case, Latin requires *w and Greek requires *b. This b ~ w alternation is still irregular, but not unattested (cf. Kroonen 2013: 37). Thorsø (fthc.) shows that we must adduce Arm. *arowoyt* ‘alfalfa’ as a comparandum, and suggests that it is a hypercorrection of earlier *arowowt. In this case it would end with a suffix *-oūd, perhaps an un-nasalized form of the vθ-suffix, which might also be present in the *-ī- of Germanic (see Kroonen fthc. with lit.).

A connection with Skt. *aravinda-* ‘lotus’ mentioned by WH (420) and several earlier sources is very uncertain according to Mayrhofer (KEWA I: 48; EWAia III: 13). The lotus indeed has round, edible, high-protein seeds, but the word does not appear until the period of the epics, which is problematically late for a word that might have been picked up in Europe. At the same time, potential Dravidian sources like Kannada *are-viri* and Telugu *ara-viri* ‘to be half-opened (as a flower)’ do not seem any more convincing. Instead, more reliable might be several Iranian forms, albeit isolated to the Pamir languages, which Kroonen (fthc.) reconstructs to PIIr. *Hrab^(h)ant-: Shughni *rivand*, Rushani *ravand* ‘wild chickpea (*Cicer songaricum*),’ and Yazgulyam *raván* ‘pea’.

While sources as recently as DV (195) and EDG (451) adduce Mİr. *orbaind* ‘grains’ as a

¹¹¹ The Georgian forms are almost certainly borrowed from Greek (pace Furnée 1972: 198). Lafon (1934: 34) had placed great weight on the *a* of a form *erevandi*, but noted that, in an updated version of his dictionary, Soulxan-Saba had replaced it with *erevindi*. The form with *a* was indeed likely a mistake (cf. Kroonen fthc.).

¹¹² Forms like Old Spanish *arvanço*, *ervanço* could be a borrowing from otherwise unattested Go. *arwāits romanized as *arwātius, with forms like Sp. *garbanzo* etc. having been contaminated by *garroba* ‘carob’. But they have also been argued to represent independent comparanda < PRom. *ervantios (Alessio 1944b: 410, Hubschmid 1955: 238-45, Corominas & Pascual 1984-91 III: 85).

related form, this is likely a coincidental look-alike that actually belongs to OIr. *arbor* ‘grain’ and W *erwain* ‘meadowsweet’.¹¹³ Puhvel (V: 134) suggests that hapax Hitt. *arwana-* might mean ‘pea’ and be related, but its context (“pours wine...into the pit and throws *arwana-*”) is too vague to confirm this interpretation.

The irregular but certain correspondences between Latin, Greek, Germanic, and Armenian along with the widespread non-IE suffix make a strong case for a substrate borrowing (WH I: 419-20; EM 202; Schrijver 1991: 36, 423; DV 195).

faba ‘bean’

Pre-form: **b^ha/Hb-* | PItal. *fabā*

Comp.: **b^(h)a/ob^h-* | PSlav. **bòbъ* | Ru. *bob* ‘bean’, etc.

**b^(h)a/ob^h-* | PBalt. **babō-* | OPr. *babo* ‘bean’

**b^h2eu-n-* | PGM. **baunō-* | ON *baun*, OE *bēan*, OHG *bōna* ‘bean’, etc.

Proto-Berber **ā-βāw* ~ **ā-bāw* ‘bean’

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: plant, domestic

Pokorny (106), WH (I: 436), EM (208), DV (197)

Alessio (1955: 368), Furnée (1972: 175), André (1978: 50), Schrijver (1991: 488), Kuiper (1995: 65), Biville (I: 187), Demiraj (1997: 94), Orel (1998: 94), Boutkan & Kossmann (1999: 88), EDG (1547, 1556), Kroonen (2013: 55)

Faliscan *haba* beside Lat. *faba*, if it is genuine,¹¹⁴ requires the reconstruction of initial **b^h* but medial **b*.¹¹⁵ Given the Balto-Slavic accentuation,¹¹⁶ Latino-Faliscan must reconstruct to original *a*-vocalism (cf. Kuiper 1995: 65). Already requiring the reconstruction of two very rare phonemes, this word does not look inherited. The Germanic bean words reconstruct to **b^hau-n-*, whose nasal element does not appear in any of the other comparanda and might be an example of the non-IE *n*-suffix. It cannot

¹¹³ This pair, as well as the Old Irish paradigm itself, demonstrate that from the PIE root **h₂erh₃-* ‘to plow’, Proto-Celtic continued a heteroclitic noun **ar(a)war* ~ **ar(a)wen*, an archaic and thus likely inherited formation (Matasović 2009: 40). While the *b* of OIr. *arbor* should have disappeared intervocally, this is a case of paradigmatic leveling due to the heteroclitic stem. **h₂erh₃-ur* > **arur*, **h₂erh₃-ūen* > **arawen* with leveling to **arwur* ~ **arwen*. In Old Irish, this was continued as an irregular paradigm neut. nom. sg. *arbor*, nom. pl. *arbanna* (EDIL s.v. 1 *arbor*). The nom. pl. *arbanna* (treated as an *n*-stem in the oblique cases) seems to have been reinterpreted as a more regular neut. *o*-stem plural (cf. Thurneysen 1946: 175) implying a nom. sg. **arbann*. After the neuter gender was lost in Middle Irish, the form would have become a masc. *o*-stem, the nom. pl. of which is the form we have attested: *orbaind* (the *nd* for *nn* and *o* for *a* in these positions do not make a phonemic difference).

¹¹⁴ It occurs in two glosses (Velius Longus, *CGILat.* VII 69.6-10; Terentius Scaurus, *CGILat.* VII 13.8-9), the latter of which explicitly ascribes it to Faliscan. But Bakkum (2009 I: 82-3, 210) is cautious.

¹¹⁵ Thus André’s (1978: 50) assessment of a reduplicative origin expressing shape/form is untenable.

¹¹⁶ Any sort of full-grade ~ zero-grade alternation involving a laryngeal such as **b^hh₃b-* ~ **b^hh₃eb-* is not allowed by the Baltic accentuation.

be reconciled with the Italic and Balto-Slavic forms in any regular way (cf. EM 208), suggesting that all are independent loans from a non-IE language (cf. Schrijver 1991: 488, DV 197). In fact, a further irregularity is required by the Slavic evidence, which must descend from **b^(h)ab^h-*, as **b^(h)ab-* would trigger Winter's Law and give PSlav. ***bábъ*. Thus between the Italic, Balto-Slavic, and Germanic forms, the second consonant shows a non-IE *b ~ b^h ~ w* alternation, pointing to loans from an unknown language.

Proto-Berber **b* is quite rare (Boutkan & Kossmann 1999: 88). Thus Proto-Berber **ā-βāw ~ *ā-bāw* 'bean' is likely to have been borrowed at a post-Proto-Berber date and Maarten Kossmann (*p.c.*) suspects from something like Italic. The final **w* does not easily correspond to the **b* of PItal. **fabā* however, which to me suggests it could still be an independent loan from a third source.

Gk. φακός 'lentil' (EDG 1547) and Alb. *báthë* 'broad bean' (Orel 1998: 94) are compared, but attest to **k* where Italic, Balto-Slavic, and Germanic have a labial.¹¹⁷ Their appurtenance is thus very uncertain.¹¹⁸ A more likely, albeit indirect, Latin comparandum for these forms is *phaselus* 'bean' (Demiraj 1997: 94), but this is a borrowing from Gk. φάσηλος 'edible bean; small boat' (cf. Biville I: 187).¹¹⁹

far, -rris 'husked wheat, emmer, grain, flour'

Pre-form: **b^ha/Hrs-* | PItal. **fars-*

Comp.: **b^ha/o/Hr(V)s-* | PGM. **bariz-* | ON *barr* 'grain, barley', Go. *barizeins* 'of barley'

**b^(h)a/HrV-* | PCelt. **baragi(-nā)* | OIr. *bairen* 'bread, food'

**b^(h)a/ors-ino-* | PSlav. **bǫrš-ьno* | OCS *brašьno* 'food', SCR. *brášno* 'flour, food'

¹¹⁷ To Gk. φακός 'lentil', EDG (1547) wonders if ἀφακη, ἄφακος 'vetch' should be connected, which could make it Pre-Greek.

¹¹⁸ WH (I: 436) saw in Lat. *faba* a reduplicated *Lallwort* for something swollen. This would help adduce the Greek and Albanian comparanda, whereby a pre-form **b^hak-* (Demiraj 1997: 94, Orel 1998: 94) could represent **b^ha-* with a non-IE suffix **-ko-* and PGM. **baunō-* would attest to the *a ~ au* vocalic alternation found in *caupō-κάπηλος*. But as Italic reconstructs to **b^ha-ba-*, it is not truly reduplicated after all.

¹¹⁹ Some interpretations in the 1930s had it go the other way. Pisani (1930: 184) took Roman *faceōlo* as evidence of a **faceolus* besides Tuscan *fagiōlo* < **faseolus* suggesting to him Umbrian origin, with the word brought to Greece through Magna Graecia. Kretschmer (1933b: 181-2) rather saw the word as a *satəm*-treatment of **b^hak-* like the Albanian word, and theorizes that the word entered Latin through Illyrian before being brought to Greece. WH (I: 436) reject the connection entirely, perhaps too hastily, but (pg. 299) assert correctly that the direction is from Greek to Latin. Its further origin in Greek (loan from a *satəm*-treatment of **b^ha-k-*?) can only be speculated on. Any consideration of Lat. *basēlus* as representing a Lat. *b* for Gk. φ alternation (cf. Alessio 1955: 368 [who admits that the comparison is uncertain], Furnée 1972: 175, EDG 1556) cannot be upheld. The form *basēlus* is only found in Isidore, and clearly represents a Late Latin development (Biville I: 187-8).

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: plant, domestic

Pokorny (111), WH (I: 455-6), EM (216), DV (201)

Meiser (1986: 172, 174), Schrijver (1991: 113-4), Untermann (2000: 265-6), Derksen (2007: 57), Matasović (2009: 56), Kroonen (2013: 52), Kroonen et al. (2022: 5)

Lat. *far* ‘emmer, flour’ can have arisen via syncope from an *s*-stem like **b^hrH-os* > **faros* (cf. *vir* < **wiros*), but U *farsio* (= Lat. *farreum*) cannot be the result of syncope; intervocally **s* > *z* and then, post-syncope, the resulting cluster **rz* would have given U **farfio* (Meiser 1986: 172, 174; Schrijver 1991: 113). Nor is **b^hrH-s-* possible as it would give Lat. **frās-* (Schrijver 1991: 113, Untermann 2000: 265-6 with lit., Kroonen et al. 2022: 5). Slavic also points to a root **b^hars-* (or **b^hors-*, which is equally unlikely in an inherited *s*-stem) in PSlav. **böršbno-* (Derksen 2007: 57), with a nasal suffix. With the traditional explanation of an inherited *s*-stem (cf. Pokorny 111) effectively ruled out, PGM. **bariz-* has either reanalyzed **b^hars* as an *s*-stem to which it introduced ablaut (DV 201) or it borrowed the lexeme as **b^hare/is-* (Kroonen et al. 2022: 5). PCelt. **baragi(-nā)* reconstructs to **b^harV-*, conspicuously lacking the **s* (if the segmentation is correct). This lexeme is quite likely a loan from an unknown source (Schrijver 1991: 113-114, DV 201, Matasović 2009: 56, Kroonen et al. 2022: 5; less explicitly EM 216, Untermann 2000: 265).

fascinus, *-um* ‘evil spirit; charm, spell; apotropaic phallus’

Pre-form: **b^ha/Hsk-Vno-* | PItal. **faskVno-*

Comp. **ba/hzsk-ano-* | PGk. **baskano-* | Gk. βάσκανος ‘who bewitches; sorcerer, slanderer’

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: magico-religious

Pokorny (91-2, 105-6), WH (I: 459), EM (218), DV (203)

Wharton (1890: 34), Kretschmer (1896: 248-9 fn. 4), Frisk (1960-72 I: 223-4), Leumann (1977: 167), Schrijver (1991: 102), EDG (191, 203), Magni (2017), Weiss (2020: 308 fn. 121)

The only comparandum of Lat. *fascinum* is Gk. βάσκανος (Schrijver 1991: 102) and despite their close similarity,¹²⁰ the irregular correspondence of Lat. *f* to Gk. β shows that neither is derived from the other.¹²¹

¹²⁰ The suffix *-ino-* is generally borrowed from Greek to make material adjectives, with the native Latin suffix being *-īno-* (Magni 2017, Weiss 2020: 308, fn. 121). Here it is clearly something else, and is the product of vowel weakening, perhaps from an *a* like in the Greek form.

¹²¹ Theoretically, the alternation could also be between **g^{wh}* and **g^w*. Because the difference is still one of aspiration, it is not of great typological consequence. However, in §4.3.2.1 it will be suggested that the *f*~

Wharton (1890: 34) suggested that this was due to the Thracian reflex of **b^h*, and Kretschmer (1896: 248-9, fn. 4) agrees that it originated in the North in his section on Illyrian. Several have followed (e.g. WH I: 459, Pokorny 105-6, Leumann 1977: 167, EM 218) because it is attractive to see this as cognate with Gk. φημί 'to speak' and φάσκω 'to declare, think'. While Latin attests denominal *fascināre* and Greek βασκαίνω 'to bewitch', the latter has further related forms. While βάζω 'to speak, say (often of nonsense) and βάζις 'word, rumor' are sometimes considered onomatopoeic (cf. WH I: 459), βάσκειν· λέγειν, κακολογεῖν (Hsch.) cannot be done away with. Frisk (1960-72 I: 223-4) followed by EDG (203) suggests that in the sense of κακολογεῖν, βάσκειν might have been influenced by βάσκανος. I wonder if they are simply of the same origin and βάσκειν has been influenced by φάσκω in the senses of (κακο)λογεῖν.

Thraco-Illyrian origin might explain the Greek forms, but this assumption is based solely on the purported etymological link with φάσκω. Moreover, it does not explain the Latin. Unlike in *ballaena* ~ φάλλανα (s.v.), Latin shows the expected reflex of **b^h*. If, as EM (218) assert, βάσκανος is derived from βάσκειν, and βάσκειν is from the Thraco-Illyrian version of φάσκω, then Latin would have to have produced *fascinum* independently. While it is possible that the verb for 'to speak' could be used to mean 'to cast a spell', the Latin and Greek forms are probably too similar in shape and derived semantics to be coincidence. This leads Schrijver (1991: 102), DV (203), and EDG (203) to propose that they are common borrowings from a substrate. However this pair came about, it was not due to internal developments within Greek or Latin, nor was it borrowed from any attested source.

ferrum 'iron, steel'

Pre-form: **b^hers-* | PItal. **fersom*

Comp.: **b^hros-* | PGM. **brasa-* 'brass' | OE *bræs* 'bronze, brass', OFri. *bress* 'copper'

Luw. **parza-* 'iron'

>>Akk. *parzillu-* 'iron'

>> Ugr. *brđl*, Hebr. *barzel*, Phoen. *brzl*, Aram. *przl*, etc.

Svan *berež* 'iron'

?Ingush/Chechen *borza* 'bronze'

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: metallurgy

b alternation is not as useful for stratification because both words could be from a source form with /v/ or /β/. If however the reconstruction with **g^{wh}* for *fascin* is correct, then this word too belongs to the group of the oldest lexemes borrowed on the Italian peninsula.

WH (I: 485-6), EM (229), DV (214)

Hommel (1881: 3386), Vaniček (1881: 109), Muller (1918: 148), Krogman (1937: 268-9), Alessio (1941a: 552), Pokorny (109-10), de Simone (1968-70 II: 179), Furnée (1972: 232 fn. 13), Breyer (1993: 444), Pleiner (1996: 287), Watmough (1997: 99), Valério & Yakubovitch (2010), Garnier (2017a: 252), Thorsø & Wigman et al. (2023: 111-12)

Attempts to derive Lat. *ferrum* from PIE have treated it as isolated. Early on, Vaniček (1881: 109) derived it from a root **b^hers-* ‘to fixate, solidify’, but the root (Pokorny 109-10) rather means ‘point, stubble, bristle’ (cf. Lat. *fastigium* ‘sharp point, tip’, OIr. *barr* ‘tip, top’, OHG *burst*, *borst* ‘bristle’, etc.). Recently, Garnier (2017a: 252) derived it from **d^her-* ‘to hold, support’ through a backformation of **con-ferer-atus* > **conferrātus* ‘resoldered’ from an *s*-stem **d^her-elos-*.

It is not isolated, however, and the external comparanda make it clear that it is a Wanderwort. Within Indo-European, *ferrum* cannot be separated from PGM. **brasa-* ‘brass’. Krogman (1937: 268-9) linked the two under an ablauting *s*-stem **b^her-s-*, **b^her-os-* to a root **b^her-* ‘to shine; bright, brown’ but these are now seen as different roots; nor is it clear what pattern of ablaut this would reflect. Adducing Svan *berež* ‘iron’ (Furnée 1972: 232 fn. 13) and Ingush/Chechen *borza* ‘bronze’ (Thorsø & Wigman et al. 2023: 111-12) suggests that the sigmatic element is a part of the root. The sigmatic element is further present in a group of related Semitic words including Ugr. *brdl*, Hebr. *barzel*, Phoen. *brzl*, Aram. *przl*, Cl. Arab. *firzil*, etc. (Muller 1918:148, Alessio 1941: 552, WH I: 485-6, DV 214, hesitantly EM 229). The Semitic forms are all borrowed from Akk. *parzillu-* ‘iron’ (known since Hommel 1881: 3386), which Valério and Yakubovich (2010) have suggested is from a Luwian word meaning ‘iron ore’. The lexeme **parza-* occurs in *parzassa-* ‘made of *parza-*’ and *parzagulliya-* ‘having loops made of *parza-*’. Thorsø & Wigman et al. (2023: 111-12) argue that **parza-* meant ‘iron’ rather than ‘iron ore’ and that the *l*-suffix of the Semitic forms could have been added via a Hurrian intermediary.

Despite identifying its ultimate source, the immediate source of Lat. *ferrum* remains unknown. Thorsø & Wigman et al. (2023: 111-12) show that there is no understood mechanism to explain how initial Phoenician *b* might be borrowed as Latin *f*. Latino-Punic underwent fricativization of *p* to *f*, but Plautus’ *Poenulus* uses to spell *b*. Several have instead suggested Etruscan mediation for the word (Alessio 1941a: 552, WH I: 485-6, Furnée 1972: 232, Breyer 1993: 444), which is archaeologically attractive seeing as the earliest iron production on the Italian peninsula is from Etruria (Pleiner 1996: 287). However, Thorsø & Wigman et al. (2023 fn. 34) show that, although Etruscan would likely have de-voiced initial *b* > *p*, a change within Etruscan of *p* > *f* is not frequent or regular enough to count on. Sporadic *p* > *f* changes within Etruscan are likely late and regional. Direct contact with *r*, *l*, *n*, *m*, or *s* has been interpreted as leading to a change *p* > *f* (de Simone 1968-70 II: 179), but Watmough (1997: 99) shows via a chronological ordering of attestations that the change was actually from a fricative to a

stop. Thus the mediating language from which Latin borrowed *ferrum* remains unknown.

ficus ‘fig’

Pre-form: **dʰik-o-* | Pltal. **ḫiko-*

Comp.: **dʰiljūūk-o-* | PGk. **ʰ(j)wūko-* | Gk. Boeot. τῦκον, Att-Ion. σῦκον ‘fig’

**tu/ūḡʰ-* | PArm. **tuz-* | Arm. *tʰuz* ‘fig’

Hebr. *šiqmā* ‘the sycamore fig’, Aram. *šiqmīn* [pl.] ‘mulberry trees’

>> Gk. συκάμῖνον ‘mulberry tree’

>> Gk. συκόμορον ‘mulberry’ (influenced by μόρον ‘mulberry’)

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: plant, tree; fruit

WH (I: 492), EM (232), DV (218)

Lewy (1895: 23), Berger (1956: 21-2), Battisti (1960: 359, 381), Turner (1966-9 I: 509), Hoffner (1967: 43, fn. 58), Friedrich (1970: 150), Furnée (1972: 262), Puhvel (III: 232), Martirosyan (2009: 295), EDG (1421), Simon (fthc.)

Latin *ficus* ‘fig’ can be reconstructed with initial **bʰ*, **dʰ*, or **gʷʰ*, but in light of the comparanda, **dʰ* is the obvious choice.¹²² In any case, Latin *f* is the reflex of a voiced aspirate, which forms an invalid **DʰeT* root structure. Arm. *tʰuz* ‘fig’ reconstructs to a different invalid root **tu/ūḡʰ-*, though Martirosyan (2009: 295) suggests it might represent an underlying **tu/ūk-* influenced by the suffix *j/z* found in plant and animal names. The palatalization is automatic after *u*, and need not be original. This pre-form is very similar to that behind the Greek forms Boeot. τῦκον and Att.-Ion. σῦκον, with a glide initiating the change to *σ*. In fact, both Armenian and Greek pre-forms can be reconstructed with **tʰ*. In Greek this is the reflex of a PIE **dʰ*, but in Armenian it is only the result of **tʰ*.

This group of words is widely accepted to be independent loans from a non-IE language, perhaps from a word with the shape of **tʰik-*, **tʰūk-*, or **tʰwīk-* (Furnée 1972: 262, WH I: 492, EM 232, DV 218, etc.). An initial **tʰ* can be reconstructed for Italic, Greek, and Armenian, though it is neutralized in the latter two. A non-IE origin is additionally supported by the existence of Hebr. *šiqmā* ‘the sycamore fig’ and Aram. *šiqmīn* [pl.] ‘mulberry trees’, the isolation of which suggests they are not native to Semitic (Battisti 1960: 359). The latter was borrowed back into Greek as συκάμῖνον ‘mulberry tree’ and less directly as συκόμορον ‘mulberry’, which seems to have been affected through folk etymology by both σῦκον and μόρον ‘mulberry’ (Lewy 1895: 23 with lit.).

Sometimes compared (Berger 1956: 21-2, Battisti 1960: 381, Friedrich 1970: 150) is

¹²² This rules out a connection with Semitic forms like Phoen. *pg* ‘ripe fig’ and Hebr. *paggā* ‘unripe fig’ that are clearly only superficially similar to the most recent Latin form.

Burushaski *phaák* ‘fig’, reconstructed by Berger (not without reason) as **twoq*. Nevertheless, it seems much more likely that an Iranian word from **phālgu* ‘fig’ (Turner 1966-9 I: 509 gives e.g. Shina *phāg*) is the source of the Burushaski word. Hoffner (1967: 43 fn. 58, see further Simon fthc.) had suggested that the *-sik(k)a* element of Hittite plant names like *ḫas(s)ik(k)a-* ‘a tree and its fruit’ and *marsikka-* ‘id.’ might be comparable to Gk. *σῦκον*.¹²³ As Puhvel (III: 232) notes however, this relies on *ḫas(s)ik(k)a-* being translated as ‘fig tree’, whereas in several lists, it is mentioned alongside ^{GIS}MA ‘fig’, suggesting it means something else. Finally worth mentioning is Udi *ṭaxan* ‘fig’, whose shape is potentially quite close to the other forms and whose source in the Caucasus is not far removed from Armenia.

filix, felix ‘fern, bracken’

Pre-form: **b^hel-ik-* | PItal. **fe/ilik-*

Comp.: **blēg^h-n/r-* | PGk. **blēk^hn/r-* | Gk. *βλήχρον, βλήχρον* ‘male fern’

**b^hreg-n-* | PGM. **brekna(n)-* | Dan. *bregne*, Sw. *bräken*, etc. ‘fern, bracken’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, wild

WH (I: 497), EM (234), DV (220)

Petr (1896: 209), Pokorny (120), Falk & Torp (1960: 100), Furnée (1972: 132, fn. 64, 65), Leumann (1977: 101), Schrijver (1999: 37-8), EDG (221), Beekes (2014: 37)

Lat. *felix* probably shows the original vocalism, with *filix* being the result of assimilation from the *i* of the next syllable (cf. Leumann 1977: 101). Lat. *filix/felix* has been compared since Petr (1896: 209) to reflexes of PIE **b^hel-* ‘henbane’ (Pokorny 120) in Germanic (Ger. *Bilsenkraut*, etc.) and Slavic (Ru. *belená*, etc.), but this has been given up for semantic reasons (cf. WH I: 497 and EM 234). DV (220) however, following Schrijver (1999: 37-8), revives the link: “The stems of henbane show a superficial resemblance to the feathered leaves of fern, and both plants have well-known medicinal properties. This often suffices to create formal similarities in languages.” Battisti (1960: 352) saw it as either a reflex of **b^hel-* having undergone Mediterranean changes or a complete Mediterraneanism, though his only evidence of this is that it is isolated. Instead, it seems most likely that *filix/felix* is neither inherited nor isolated. It is best compared to words with the same meaning (‘fern/bracken’) in Greek and Germanic.

Latin *f* goes back to a voiced aspirate, and cannot regularly correspond to Gk. *βλήχρον* ‘fern’ (Grassmann’s Law should result in *π*). EDG (221) also shows that the *r ~ n* suffix alternation in Greek is not the result of an inherited heteroclitic stem, but must be something peculiar. PGM. **brekna(n)-* (cf. Falk & Torp 1960: 100) ‘bracken’ shows an *r*

¹²³ He more specifically mentions Myc. *su-za*, which EDG (1421) reads as */σῦκία/*.

~ *l* alternation with the two other comparanda. Its velar reflects PIE **g* as opposed to **k* for the Latin. It seems unlikely that this is the result of leveling after devoicing in the nominative, since this change is not normally leveled (cf. *rēx*, *rēgis* ‘king’). The Greek velar reconstructs to **g^h*, which could be the result of the additional suffix (cf. Furnée 1972: 132, fn. 64, 65; Beekes 2014: 37) shared by Germanic (which may have resulted in the alternation in vocalism in the first syllable). This indicates that the *-ix/-ex* suffix of the Latin forms is not of IE origin. Alternatively, if the whole lexeme was **BleG-n-*, then Latin, which attests to a form without the suffix, has interpreted the foreign root-final velar (particularly after the addition of a nominative *-s*) as a native suffix.

fracēs ‘lees, oil dregs’

Pre-form: **d^hrak-* | PItal. **prak-*

Comp.: **d^hra/og^h-* | PGm. **dragjo-* | ON *dregg* ‘dregs, lees, yeast’, etc.

**d^(h)ra/og^h-* | PBalt. **dragia[?]-* | OPr. *dragios* ‘yeast’, Lith. *drāgēs* ‘dregs’, etc.

**d^(h)ra/osg^h-* | PSlav. **drozgija-* | OCS *droždbe* ‘dregs’, etc.

**d^(h)ra/o/Hs-* | PALb. **dras-* | Alb. *dra* ‘dregs, sediment’

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: viticulture / oil

WH (I: 538-9), EM (251), DV (238)

Fraenkel (1962-5: 103), Kortlandt (1987), Orel (1987: 140), Schrijver (1991: 486), Demiraj (1997: 141), Orel (1998: 141), Derksen (2007: 121), EDG (553), Kroonen (2013: 99), Schumacher & Matzinger (2013: 262-3), Derksen (2014 s.v. *derkti*), de Vaan (2018: 1746), Weiss (2018: 444)

Germanic **dragjo-* (Kroonen 2013: 99) and PBalt. **dragia[?]-* (cf. Derksen 2007: 121) can be reconstructed to the same root shape. Lat. *fracēs* is a good semantic match, but the quality of its velar does not match theirs.¹²⁴ Taken at face value, *fracēs* reconstructs to an invalid **D^heT* root structure. WH (I: 528-9) and DV (238) suggest that voiceless *k* could have arisen in the nominative singular, devoiced before the ending *-s*. The singular *frac* is attested rarely, once in the Philoxenus Glossary.¹²⁵ This does not seem enough to affect a word used mainly in the plural, and would be regular in any case (**g* is preserved in e.g. *rēx*, *rēgis*).¹²⁶ However, an additional indication of non-IE origin is the sigmatic

¹²⁴ Kroonen (2013: 99) separates it by deriving it from *frangō* ‘to break’, but all other derivatives of this verb maintain the *g*.

¹²⁵ And once in another gloss according to the TLL.

¹²⁶ If the *k* were originally *g* from a root **d^hrag^h-*, Weiss (2018: 444) writes that this is exactly the root shape in which Limited Latin Grassmann’s Law should operate. The potential blocking of the phenomenon (the expected result would be **drag^h-* > **drages* > ***tragēs*) is a parallel to the same blocking in the nominative of Gk. θρίξ, τριγός.

element that appears in PSlav. *drāzġija-*.

Alb. *dra* ‘sediment, dregs; smudged butter; sweepings, dirt’ is derived from **drag-* < **d^hra/og^h-* by Demiraj (1997: 141) and Orel (1998: 141). The former supports this with a proposal that it is the source of the verb *ndrag* ‘to make/get dirty’. But it is not fully clear if **g^h* should disappear,¹²⁷ and de Vaan (238) considers a pre-form **drab-*. Less problematic might be a reconstruction **d^(h)ras-*, with the sibilant of the Slavic forms yet lacking the velar.¹²⁸ Perhaps this makes it possible to further compare (cf. DV 238, Schrijver 1991: 486, WH I: 538-9) forms with no velar but a long vowel + *sn* (PGm. **drōsna-*: OE *drōsne*, OHG *truosana*, MoDu. *droesem* ‘dreg’, perhaps also OE *drōme* [Schrijver 1991: 486]).^{129, 130}

Latin requires *a* vocalism, as a laryngeal would produce a **CRHC* sequence yielding **frācēs*. The Baltic accentuation (EDG 553) and vocalism also prohibit a laryngeal in the root. Thus Derksen (2007: 121) favors reconstructing *a*-vocalism for the Balto-Slavic formations despite *o* being a possibility. This, in addition to the irregular velar correspondence, the vacillating appearance of the sigmatic element, and the invalid root structure make this family look very much non-IE.

frīgō ‘to roast’

Pre-form: **b^hreig-* / **b^hriHg-* | PItal. **frīg-*

Comp.: **b^hruHg-* | PGk. **p^hrūg-* | Gk. φρῦγω ‘to roast, dry, fry’

?**b^hreǵ-* / **b^herǵ-* | PIIr. **b^hra(i)ǵ-* | Skt. *bhrajj-* ‘to fry, roast’, MP *bryz-* ‘to roast, bake’, etc.

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: action; culinary

Pokorny (137), WH (I: 548), EM (259), DV (254)

Thurneysen (1890: 353), Biville (I: 194, II: 290-3), EDG (1593), Giacomelli (1994: 36), Cheung (2007 s.v. *bra(i)ǵ*)

Lat. *frīgō* ‘to roast’ is attested since Plautus, but its perfect *frīxī* is not attested in Classical Latin. Giacomelli (1994: 36) takes it as a loan from or influenced by Greek φρῦγω ‘to roast, dry, fry’. But this is difficult to defend. There are examples of Lat. *i* in

¹²⁷ Orel (1987: 140) suggests it had to do with original accentuation (e.g. *shteg* ‘path’ < **staiga* < **stóig^hos* vs. *ve* ‘widow’ < **widewā* < **uid^héueh₂*).

¹²⁸ According to de Vaan (2018: 1746) building on Kortlandt (1987), intervocalic **s* does not disappear but rather yields Alb. *sh*. De Vaan notes that it is still debated, and Schumacher & Matzinger (2013: 262-3) propose that intervocalic **s* only yields Alb. *sh* before a front vowel, whereas it disappears before a back vowel.

¹²⁹ WH (I: 538-9) further mention similar forms with an *st* suffix: OE *dærst(e)*, *dræst* ‘yeast, dregs’, OHG pl. *trestir* ‘marc (remains of crushed grapes)’.

¹³⁰ EM (251) compare *marcēre* ‘to be withered, wrinkled, weak’ along with e.g. OIr. *mraich* ‘malt’, but Schrijver (1991: 458) notes that **mr* yields Lat. *br*, not *f*.

loans from Gk. *v* since the archaic period, but these can often be explained phonologically or by alternations already circulating in Greek (Biville II: 290-3). Nor does Lat. *f* borrowed from Gk. *φ* occur before the 1st c. BCE (Biville I: 194). EDG (1593) and DV (243) support both forms being borrowed from a third language or a Latin borrowing from Greek via an intermediary language (cf. the same vocalic alternation in *fīcus* ~ σῦκον ‘fig’). Cheung (2007 s.v. *bra(i)j*) compares them to a widespread root amongst the Iranian languages (with *i* introduced into the full-grade from the zero-grade), also occurring in Skt. *bhraj-* ‘to fry, roast’ < **b^hreǵ-* / **b^herǵ-* ‘to roast’ (cf. also WH I: 548; EM 254 as evidence of an expressive word), but neither the Latin nor the Greek can be a regular reflex of this root.¹³¹

fulica, fulix ‘water bird, probably coot’

Pre-form: **b^hul-Vk-* | PItal. **fulVkā-*

Comp.: **b^ha/ol-ig-* | PGm. **balikōn-* | OHG *belihha* ‘coot’

**b^(h)o/ul-a/oK-* | PCelt. **bo/ul-a/okkagno-* | SGael. *bolachdan* ‘coot’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: animal, bird; aquatic

Pokorny (118-20), WH (I: 559-60), EM (259), DV (248)

Niedermann (1905-6: 78), Persson (1909: 60-2), Furnée (1972: 192), Sommer & Pfister (1977: 60), Fruyt (1986: 229-30), EWA (I: 530), EDG (1550), van Sluis (fthc.)

Lat. *fulica* is traditionally linked to **b^hel-* ‘shining, white’ (cf. Pokorny 118-20), allowing a connection with Gk. φαληρίς ‘coot’, Hsch. φάλος· λευκός (EDG 1550).¹³² This is not without issue, however. One must assume dialectal¹³³ *u* for *o* < ***b^hol-* (Sommer & Pfister 1977: 60), but this is *ad hoc* (DV 248). It is also obvious to compare it to OHG *belihha* (WH I: 558-9, EM 259 with lit.), but in fact, this creates a **k* ~ **g* alternation in the suffix (interpreted as different inherited velar suffixes by EWA I: 530).¹³⁴ Van Sluis (fthc.) identified SGael. *bolachdan* ‘coot’ as a comparandum. In the region where the word is attested, the reflexes of OIr. *-cht* and *-cc* merge into /xg/, thus behind *bolachdan* could be the pre-form **bo/ul-a/oxtagno-* < **bo/ul-a/okt-* with a further unexplained dental element or **bo/ul-a/okk-agno-* < **bo/ul-a/okk-* + the diminutive suffix. The latter looks more similar to the Italic and Germanic comparanda, with gemination of the velar.

¹³¹ Thurneysen (1890: 353) earlier tried to derive them via a “vocalic *z*” from **b^hrzǵō-*.

¹³² Further also Skt. *balākā-* ‘white heron, egret’, but this requires the assumption that it has been contaminated by *baka-* ‘heron’ (Niedermann 1905-6: 78, followed in e.g. KEWA II: 418, Fruyt 1986: 230, EWA I: 530).

¹³³ The regular development of **o* > *u* / *_(l)i* proposed by Persson (1909: 60-2), followed by Fruyt (1986: 229-30) is unlikely given e.g. *folium*.

¹³⁴ The form *fulica* might help show that the unvoiced velar is original, rather than the result of devoicing in the nominative singular. But if it is secondary to *fulix*, then it cannot be ruled out that the change was leveled to the oblique forms of *fulix* before the formation of *fulica*.

Furnée (1972: 192) and EDG (1550) note that, amongst the Greek attestations of words related to φαλός, there is consonantal alternation that makes it not look particularly native. Thus, even if the Greek form does share a root with the Italic, Germanic, and Celtic words, it does not prove that it is inherited. Furthermore, DV (248) notes that it is uncertain whether *fulica* even refers to the coot.¹³⁵ Its etymologization under a lexeme meaning ‘white’ may in fact be a learned folk etymology. The irregular correspondences we must reconstruct for this Latin, Germanic, and Celtic bird lexeme point to non-IE origin.

funda ‘leather strap, sling’

Pre-form: **b^h/g^{wh}und^(h)-* | PItal. **f/χ^wund/pā*

Comp.: **sb^h/g^{wh}end-* | PGk. **sp^hendonā* | Gk. σφενδόνη ‘sling’

■ Irreg. correspondences □ Remarkable phonotactics

Semantics: tool

WH (I: 562), EM (260), DV (249)

Cuny (1910: 158), Meillet (1922a: 73), Leumann (1977: 162), Biville (I: 197-8), EDG (1430)

Lat. *funda* could semantically be from **b^hend^h-* ‘to bind’, but morphologically its *u* precludes derivation from this root. A connection with *fundere* ‘to pour’ in the sense that slinging is like pouring (cf. Walde 1921: 83) is gratuitous. Instead, the best comparandum for *funda* is Gk. σφενδόνη ‘sling’ of similar form and meaning. A direct borrowing from Greek should have resulted in Lat. ***spend-* (WH I: 562, Biville I: 197). Thus like the pair ~ σφόγγος, the Latin and Greek forms are independent relatives. Biville (I: 197-8) suggests that through *s* mobile and IE ablaut, both can go back to an inherited formation: **b^hond-* (Latin) ~ **sp^hend-* (Greek).¹³⁶ But this again does not explain the *u* of *funda*. The pair thus most likely represents loans from third source (Cuny 1910: 158, Meillet 1922a: 73, EM 260, DV 249, EDG 1430).

fungus ‘fungus, mushroom, sponge’

Pre-form: **b^h/g^{wh}ong-* | PItal. **fongo-*

Comp.: **sp/b^hong-* / **sk^w/g^{wh}ong-* | PGk. **spongo-*, **sp^hongo-* | Gk. σπ/φόγγος
‘sponge, spongy object; gland’

**sʷomb^h-* / **sʷong^{wh}-?* | PGm. **swamb/ppan-* | Go. *swammis* ‘sponge’,

¹³⁵ Vergil *Georgics* 1.363: ‘when the marine *fulicae* play on dry land’; Ovid *Metamorphoses* 8.625: ‘now the waves are frequented by diving birds and swampy *fulicis*’. Perhaps the closest, (but yet why does he not mention their color?) is Pliny *Naturalis historia* 11.44(37).122: ‘(nature) has given to the *fulicarum* kind (a crest) residing from the beak through the middle of the head’. Additionally, the Romance descendants do indeed mean ‘coot’.

¹³⁶ The root **b^hend-* is attested in Skt. *bhandate* ‘feels happy’ (LIV2 s.v. *?*b^hend-*), but semantically this cannot be the same root.

ON *svöppr* ‘mushroom’, etc.

Avar *sa:k* ‘tinder’, Tsez *zik’u* ‘mushroom’, Udi *ša’mk:al* ‘mushrooms’
 >>? Georg., Megrel., Laz *sok’o*, Svan *sok’(w)*¹³⁷ ‘mushroom’

?**psong-* / **kong*^(w)- | Arm. *sunkn, sungn, sunk, sung* ‘tree-mushroom’

?**g*^(w)(*h*)*u/omb*^(h)- | PSlav. **gōba* ‘(tree-)fungus’ | OCS *gōba* ‘sponge’,
 Sln. *gōba* ‘mushroom, tree-fungus’, etc.

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: fungus / tool

WH (I: 566-7), EM (262), DV (250)

Pedersen (1904), Cuny (1910: 158), Otrębski (1939: 184), Bartholomae (1961: col. 925), Machek (1971: 179), Furnée (1972: 164, 232, 360 etc.), ESSJa (VII: 78-80), Rédei (1986: 75), Rédei (1988: 355), EWAia (II: 240-1), Biville (I: 198-9), Derksen (2007: 182), EDG (1385), Martirosyan (2009: 586), Kroonen (2013: 495), Kurdadze et al. (2015: 193), Holopainen (2019: 186-8), eDIL (s.v. *spongē, sponc*)

WH (I: 566-7) take Lat. *fungus* ‘mushroom; sponge’ as a borrowing from Gk. *σπ/φόγγος* ‘sponge’, assuming that the meaning ‘mushroom’ must have existed but is unattested in Greek literature. They propose that the borrowing of *f* from *σπ/φ* happened under the influence of *fungor* ‘to perform, administer’. This seems untenable, as it ignores the initial *σ* and appeals to an analogy that is semantically indefensible. Biville (I: 198-9) suggests that the *σφ* cluster in *σφόγγος* would have to have been pronounced *sf*, and then heard and rendered in Latin as *fungus*. But when Latin certainly has borrowed Gk. *σπογγία*, it is as *spongia* (and in Isidore as *spungia/sfungia*). The shape of Lat. *fungus* instead suggests an independent relative of the Greek forms rather than a borrowing from them (cf. *funda* ~ *σφενδόνη*). Given the *π* ~ *φ* alternation within Greek (Furnée 1972: 164, 232, 360 etc., EDG 1385) and irregular comparanda elsewhere, it is widely suspected that the words are loaned from an unknown third source (Cuny 1910: 158, EM 262, DV 250, EDG 1385, Martirosyan 2009: 586 with lit.), and could well represent a Wanderwort.

The lexeme occurs as Arm. *sunkn, sungn, sunk, sung* ‘tree-mushroom, cork tree’, where Martirosyan (2009: 586) reconstructs **psongo-*. Such a reconstruction makes it look intermediate to Greek **sp*^(h)- and words that are similarly lacking the plosive in Kartvelian and Nakh-Dagestanian. The Kartvelian forms are taken by Martirosyan as independent comparanda due to their widespread distribution, but they cannot be ruled out as borrowings from Nakh-Dagestanian (Peter Schrijver p.c.). The Nakh-Dagestanian forms are complex and difficult to reconstruct, but Udi *ša’mk:al* crucially attests to a nasal otherwise lacking at the surface of the Dido and Avar-Andic forms. Furthermore,

¹³⁷ Martirosyan (2009: 586) gives this Svan form, though Kurdadze et al. (2015: 193) give only *t’q’ubul* (which seems to exist in Georgian as well, e.g. *t’q’ubla-sok’o* ‘*Armillaria tabescens*’).

the West Dido forms (Tsez *zik'u*, Hinuq *zek'u* ‘tree fungus, mushroom, tinder’) hint at an original paradigm in which the oblique was **sink'(w)ú-* (Peter Schrijver, p.c.).

The Slavic material¹³⁸ looks like it rules out a reconstruction with **b^h* and instead favors something like **g^{wh}* as the original first plosive. But as Derksen (2007: 182) notes, a connection between PSlav. **gōba* ‘(tree-)fungus’ and Gk. *σπ/φόγγος* is difficult formally because of the final **b^(h)* that must be reconciled with the velar of all other proposed comparanda so far. Interpretations have thus varied. Pedersen (1904) suggested that the Slavic and potential Germanic comparanda represented **sg^{wh}omb^ho-* or **sg^uomb^ho-*, metathesized variants of the root behind the Greek forms (cf. further Otrębski 1939: 184, Machek 1971: 179, without mention of Germanic). Smoczyński (2018: 404-5) instead takes the Slavic forms and several Baltic words for swellings on plants and persons (cf. Lith. *gūmbas* ‘bump, gall, ulcer, etc.’) from PBSl. **gumb-*, a neo-root reanalyzed from the nasal infix present of the root **g^(w)ub^h-* ‘to bend, curve.’

On the Germanic forms, Kroonen (2013: 495) reconstructs PGm. **swamb/ppan-*, an *n*-stem as if from **suomb^h-*. It indeed shares the problem of a final labial in place of a velar with Slavic. Given the likelihood of this being a non-IE lexeme, perhaps the Germanic and Slavic forms were borrowed without the velar element and the *b* is secondary from the nasal. Within Germanic, Kroonen (2013: 598) has, on the basis of **wulfa-* for expected **wulhwa-* < **ulk^w-*, suggested that **k^w* > **p* after resonants in words with an initial labial (cf. also **fīmfe* < **penk^we*). If an anlauting sibilant would not block this, perhaps **swamp-* is from something like **swank^w-*. Though if this works more generally on labiovelars, then **swang^{wh}-* may yield **swamb^h-*. And if this occurred before Kluge’s Law, then **-mb^h-* > **-mpp-* > would explain the **b ~ pp* alternation in Germanic. The apurtenance of the Slavic forms is still difficult, since this explanation does not apply there.

It is difficult to decide how to reconstruct the first plosive. The only reconstruction with a velar allowed by the Armenian forms is **kong(w)-* (p.c. Rasmus Thorsø). It looks suspicious because we must assume that the *s* of the Armenian form, present in most of the other comparanda, is not from a pre-form with **s* but rather happens to have developed coincidentally due to a palatovelar that is not required by any of the other reconstructions. On the other hand, the Armenian forms also look like they could plausibly post-satəmization loans from a form with a sibilant, like those in Kartvelian and Nakh-Dagestanian (now that a nasal can be reconstructed for the latter). I am not convinced that we can fully rely on the Armenian forms as independent evidence.

Without the secure (independent) apurtenance of the Slavic or Armenian forms, either a labial or a velar could have been original. The **w* of the Germanic forms could attest to a *b ~ w* alternation like that of PGm. *baunō-* against Lat. *faba* (s.v.). Alternatively, a non-IE labialized velar could have been borrowed into Latin and Greek as an aspirated

¹³⁸ The Slavic forms also mean ‘lip’, which is likely a secondary semantic development (ESSJa VII: 79, Derksen 2007: 182).

labiovelar while in Germanic the velar element was overtaken by the labial nature of the foreign phoneme. A similar situation might underlie Gk. δάφνη vs. Gk. δαύχνα ‘laurel’ (s.v. *laurus*). Potential evidence in favor of the labial comes in the form of further comparisons with Uralic, but these are not at all straightforward.¹³⁹

What we have here in any case is a widespread substrate word or Wanderwort. If one wonders what would give a word for mushroom such a broad distribution, it should be noted that certain fungi are indispensable fire-starting tools. The Ice Man of the Ötztal Alps was found with a pouch containing a fire-starting kit, composed of iron pyrites, flints, and shelf fungus (cf. Dickson, Oeggl & Handley 2003: 76). Thus this family of forms might be an ancient cultural word.¹⁴⁰

gubernō, -āre ‘to plot/steer a ship; to govern, manage’

Pre-form: *gub^(h)- | PItal. *gub/f-

Comp.: *kub- | PGk. *kub- | Gk. κυβερνάω ‘to steer, head for; to govern’
*kum- | PGk. *kum- | Gk. (Cypriot) *ku-me-re-na-i* ‘they steer’

??Lith. *kūmbryti* ‘to steer a ship’

??PSlav. *kr̥miti | OCS *kr̥miti* ‘to steer’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: action; maritime

WH (I: 625), EM (284)

Cuny (1910: 156), Boisacq (1916: 527-8), Fohalle (1925: 164-5), Otrębski (1939: 153), Ernout (1954: 24 fn. 4), Machek (1955: 61-4), Fraenkel (1962-5: 308-9), Lejeune (1972: 152), Neumann (1987: 64-9), Biville (I: 242-3), EWAia (I: 385-6), Beekes (1992: 188), Neumann (1992: 188), EDG (793), Egetmeyer (2010: 110, 202, 159) Smoczyński (2018: 629)

¹³⁹ For the comparisons, cf. Rédei (1986: 75, 1988: 355) and Martirosyan (2009: 586). Holopainen (2019: 186-8) reconstructs for the group *p̥ī̯ka ‘psychedelic mushroom’. He concludes that *p̥ī̯ka could be a borrowing from an Indo-Iranian form of the shape *b^hanga- in the meaning ‘narcotic’ vel sim., if it existed. This itself is a complicated question. Traditionally, YAv. *baṅha/bangha* (cf. Bartholomae 1961: col. 925) is interpreted as the name of a narcotic plant, but EWAia (II: 340-1) disagrees, instead following Henning (1951: 33-4). There is Skt. *bhaṅgá-* ‘hemp’, which seems to have given MoP *bang* ‘hemp’. MP *bang, mang* ‘henbane’ is unrelated. But the idea that YAv. *baṅha/bangha* refers to a narcotic plant seems to have resulted interpolation between these meanings. The word occurs in one context describing god as *abāṅha*, ‘without *baṅha*’, and thus ‘without narcotics’ makes little sense. Instead, Henning relates *baṅha/bangha* to Skt. *dvamsa-* ‘perishing, destruction’. Holopainen’s interpretation relies heavily on the psychedelic semantics of PU *p̥ī̯ka being original rather than the fungal semantics. This is indeed perhaps supported by Ngasasan *h^haṅkud’a* ‘to be drunk’. If, however, this is a secondary development (cf. Mansi (East) *p̥ē̯jk*, (West) *p̥ē̯jk*, (North) *p̥ā̯jx* ‘fly agaric; intoxication’, East Khanty *pa̯jkał-* ‘to sing after having consumed fly agaric’), the Uralic forms stand a chance of being loans from the same source as the Indo-European forms.

¹⁴⁰ Note OIr. *spong*, *sponc*, which, though borrowed through Lat. *spongia* from Gk. σπογγία, means both ‘sponge’ but also ‘touchwood, tinder’ (EDIL s.v. *spong*, *sponc*), suggesting a long-maintained but unattested meaning of the Latin word.

Lat. *gubernāre* ‘to steer a ship’ is borrowed from Gk. κυβερνάω of the same meaning, but its *g* for Gk. κ is irregular.

Greek attests to two variants: one with β (κυβερνάω) and one with μ (*ku-me-re-na-i*, probably /kumernāhi/ < **kumernansi*, cf. Egetmeyer 2010: 110, 202). Lejeune (1972: 152) prefers the explanation that κυβερνάω is a dissimilation from *κυμερνάω (cf. Homeric μαρνάμενος vs. inscriptional βαρναμενος) to do away with what otherwise looks like a *b ~ m* alternation. Neumann (1987: 64-9) explains it as metathesized from *κυρβ- and therefore related to κύρβις ‘triangular tablets forming a three-sided pyramid, turning on a pivot, upon which the early laws were inscribed at Athens’ (later ‘pillars or tablets with inscriptions’). He reconstructs it to the root **k^werb-* ‘to turn’. But Beekes (1992: 188) notes the root is otherwise always **k^werp-* (including, supposedly, Gk. καρπός ‘wrist’) and takes issue with the number of assumptions required to get from **kurb-nā-* to *κυβερνᾶ (upon which the verb was built). Neumann (1992: 188) responded in defense, saying in fact it only requires the assumption of metathesis and anaptyxis. These are indeed two extra assumptions that are used to reconstruct the word back to an otherwise unattested root. Thus I follow EDG (793) who still disagrees with Neumann and takes the irregular *b ~ m* alternation at face value (cf. also Egetmeyer 2010: 159).

Fohalle (1925) discussed the possibility that the “faiblesse articulatoire” of voiceless Greek plosives was perceived by Latin speakers as voicedness, but found no evidence of this. Thus he concluded, especially in cases where the word in question does not have an IE etymology, a voicing discrepancy between Latin and Greek points to a pre-Greek origin. Ernout (1954: 24 fn. 4) and EM (284) follow, and write that it is therefore not necessary to suppose that the word came to Latin from Greek via an intermediary; both forms could have been borrowed from the same Aegean substrate source. But Biville (I: 242-3) rightly notes that they are so close that one cannot help but suspect borrowing. Rather than independent loans, this seems like another case of a Greek word loaned into Latin via an intermediary (cf. already Cuny 1910: 156), like *ballaena* (s.v.). That it does not have an etymology within Greek does not change this.

Boisacq (1916: 527-8) accepted a connection with Skt. *kūbara-* ‘transom of a wagon’ and Lith. *kum̃bryti* ‘to steer a ship’, but wrote that their connection was unclear as they required the form with **kub-* in Greek to be older; difficult, as he too accepted the explanation that κυβερνάω is a dissimilation from *κυμερνάω. Machek (1955: 61-4, following Otrębski 1939: 153) adds OCS *kr̃miti* ‘to steer’, still supporting an IE etymology.

The connection with the Sanskrit material can be rejected on semantic grounds (EWAia I: 385-6). EDG (793) follows Fraenkel (1962-5: 308-9) in rejecting the comparisons to Baltic. The latter writes that in order to be related, we would need to assume that Cypriot *ku-me-re-na-i* is earlier than any of the other Greek dialectal forms. (Note that this is the opposite of the problem as formulated by Boisacq.) However, if there is a *b ~ m*

alternation within Greek, then neither form has to be earlier than the other, and *kum̃bryti* might instead represent an additional alternation, namely *mb*. The stronger argument is semantic. Lith. *kumbrỹs* (vars. *kum̃buras*, *kumburỹs*, *kumbras*) refers to the bent wooden portion of a yoke or rudder, as well as a hill or peak. Smoczyński (2018: 629) does not even mention the nautical semantics, taking *kumbrỹs* as a voiced variant of **kumprỹs*, a derivation from *kum̃pti* ‘to bend, stoop’. This suggests, as does the limitation of *kum̃bryti* to the area of the Curonian Lagoon (Fraenkel 1962-5: 308), that the meaning ‘rudder’ and the derived verb ‘to steer’ was a secondary, dialectal development. To connect OCS *kr̃m̃ia* ‘back end of a ship’, *kr̃miti* ‘to steer’ requires, as Fraenkel points out, the assumption of *r* metathesis. Thus the Baltic and Slavic words stand a good chance of being only coincidentally similar.

hasta ‘spear, staff’

Pre-form: **g^ha/Hst-* / **g^ha/Hzd^h-?* | PItal. **χastā*

Comp.: **g^ha/Hzd^(h)-* | PCelt. **gazdo-* ‘withe’ | Mlr. *gat* ‘osier, withe’
**g^ha/Hst-* | PCelt. **gasto-* | OIr. *gass* ‘twig, branch’

**g^ho/Hzd^h-* | PGm. **gazda-* | Go. *gazds* ‘sting’, OHG *gart*, ON *gaddr* ‘goad’

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: tool

Pokorny (412-3), WH (I: 636), EM (290), DV (278)

Schrijver (1991: 134-5), Untermann (2000: 336-7), Lubotsky (2004: 329-30), Matasović (2009: 155), Meiser (2010: 270-1), Kroonen (2013: 172)

Latin *hasta* reconstructs to an invalid **D^heT* root structure. If the Latin form is not isolated, which despite its more specific semantics, does not need to be the case, we are likely looking at a non-IE word. The other comparanda show that the dental element of *hasta* is part of the root and not, for example, a feminine *-to* suffix. Lubotsky (2004: 330) takes the differing vocalism in Oscan **hostatu** as an additional peculiarity pointing to non-IE origin. However most others are more cautious (Untermann 2000: 336-7, DV 278, etc.): since the meaning of **hostatu** is unknown, we can only speculate on its connection with *hasta* and the significance of its aberrant vocalism.¹⁴¹ There are other features within the more securely related forms that hint at non-IE origin.

Szemerényi (1952) and Meiser (2010: 119) assume that **-zd^h-* yields *-st-* in Latin, such that *hasta* could be from a root **g^hazd^h-* of permissible structure (thus Schrijver 1991: 134-5). However Lubotsky (2004: 329-30) argues that **zd^h-* yields Latin *d* with

¹⁴¹ As it appears in the phrase **hostatu anhostatu**, the *o* for expected *a* could for instance be the result of weakening in its non-initial position in **anhostatu** whereupon it was leveled to its initial position in **hostatu** (Untermann 2000: 337, Meiser 1986: 270-1).

compensatory lengthening rather than *-st-*.¹⁴² Thus we can reconstruct **-st-* for the Latin form, which creates an irregular correspondence with the Germanic forms. Even without the Latin material, this alternation also occurs within the Celtic comparanda (WH I: 636, EM 290, Matasović 2009: 155), indicating that we are dealing with a non-IE loanword.

hedera ‘ivy’

Pre-form: **g^hed^(h)-a/es/r-* | PItal. **χeda/es/rā*

Comp.: **k/g^hid^h-ar-* | PGk. **kit^hara-* | Gk. κιθάρα ‘ivy’
**k/g^hitl^h-iō-* | PGk. **kil^(h)’o-* | Gk. κισσός ‘ivy’

■ Irreg. correspondences □ Remarkable phonotactics

Semantics: plant, wild

Pokorny (437-8), WH (I: 638), EM (291), DV (281)

Furnée (1972: 256-7), Hamp (1974), Leumann (1977: 315), Meiser (2010: 83)

The traditional explanation of Lat. *hedera* follows Festus in connecting it to *-hendō* ‘to grab’ < **g^hed-*, *quod edera vincit ad quodcumque se applicat* ‘because ivy overcomes anything it attaches itself to’ (Pokorny 437-8, Leumann 1977: 315, Meiser 2010: 83). The explanation smacks of a folk etymology, and is not fully accepted by WH (I: 638), EM (291), or DV (281). Instead, a connection to Gk. κισσός ‘ivy’ is attractive. Appearing as κίσσαρος in a gloss, and with the alternate form κιθάρα,¹⁴³ the Greek words are not inherited (Furnée 1972: 256-7, EDG 704). Given Grassmann’s Law in Greek and vowel weakening in Latin, all forms can reconstruct to an original **g^hed^h-ar-* ~ **g^hid^h-ar-* with irregular *e* ~ *i* vocalic alternation pointing to a loan. The *θ* ~ *σσ* alternation in the Greek forms (cf. other pairs like *carpasum* ~ *carpathum* ‘poisonous plant’ and *ἄν(ν)ησ(σ)ον* ‘anise’ ~ *ἄν(ν)ηθον* ‘dill’) points to vacillating palatalization (cf. EDG 704).

Hamp (1974), on the assumption that the initial *h* in Latin is not etymological (given the occasional spelling *edera*) proposes an explanation in which Lat. *hedera* < **h₁ed-is-a* with comparative morphology and OIr. *edenn* ‘ivy’ < **h₁ed-ies-no-*, W *eiddew* and Bret. *ilyau* ‘ivy’ < **h₁ed-ies-uo-* with comparative morphology would be extremely archaic active intensive agentive formations with the meaning ‘voracious’ to the root **h₁ed-* ‘to eat’. This is unlikely. The Celtic forms may still be related if they attest to an alternation **g^hed^h-* ~ **ed^h-*, but this remains very uncertain.¹⁴⁴

hirundō, -inis ‘swallow, martin, and similar birds’

¹⁴² There are only four potential examples of Lat. *-st-* < **-zd^h-*, one of which is this very word.

¹⁴³ Alb. *qisár* ‘ivy’ is probably a loan from unattested fem. **κισσάρα*.

¹⁴⁴ Furnée (1972) gives examples of this sort of alternation occurring within Greek: *κάρυον* ~ *ἄρυα* ‘nut’, *γίνυος* ~ *ίνυός* ‘hinny’ (pg. 391); *κίχλη* ‘thrush, wrasse (fish)’ ~ *ίχλα* ‘a sea fish’, *κάδδυχος* ‘urn, pitcher’ ~ *ἄδδιξ* ‘a measure of volume’, *καλιנדέομαι* ‘to roll around’ ~ *άλινδέω* ‘id.’, *κανθήλιον* ‘packsaddle’ ~ *ἀνθήλιον* ‘id.’, etc. (pg. 300 fn. 59). Possibly *καβάτινα* ‘rawhide shoes’ ~ *ἄρπις* ‘high boot’?

Pre-form: **g^hir-o/und^(h)-ōn-* | PItal. **χiro/undōn-*

Comp.: **g^hel-iHd-ōn-* | PGk. **k^helidōn-* | Gk. *χελιδών* ‘swallow’

**g^ho(u)l(H)-(o)nt/d^(h)-* | PAlb. **da(u)lant/d(h)-* | Alb. *dallëndyshe*
‘swallow’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: animal, bird

WH (I: 652), EM (296), DV (286)

André (1967: 92-4), Furnée (1972: 272), Çabej (1976 I: 105-6), CAD (S: 295), Orel (1998: 55), Lockwood (2001: 217-18), Newmark (2005 s.v. *dallëndyshe*), EDG (706, 1622), Weiss (2020: 153), Kroonen (fthc.)

Lat. *hirundō* ‘swallow, martin’ is traditionally taken as a derivation of *hirriō* ‘to snarl’ (WH I: 652), a verb attested late and reserved for describing dogs (André 1967: 93). Italian dialectal forms like *ríndina* and Sicilian *rinnina* suggest that a byform **hirindo* was also in circulation. It is of course possible that the call of such birds was thought to sound like barking or snarling, but a more robust explanation is at hand in light of the Gk. *χελιδών* ‘swallow’.

The Corinthian female name *Χελιδῶν* leads André (1967: 93) to conclude that it is the original form, and that this lexeme is not simply **g^hel-* ‘to call’ plus the small animal suffix *-δών*. EDG (1622), due to the rarity of the suffix *-ῶν* in post-consonantal position, instead proposes that this is a false archaism. They nonetheless find the derivation from **g^hel-* unconvincing,¹⁴⁵ and recognize a Pre-Greek suffix *-id-* in the word.

The similarity of the Latin and Greek forms, both in form and meaning, is remarkable. They both begin with **g^h-* and end with **-dōn*. There is a mismatch in vocalism in the primary syllable, and in fact one might expect the *i* of *hirundō* to be lowered to *e* in the open syllable as in *serō* < **sisō* (cf. Weiss 2020: 153, though s.v. *pirum* for reasons to doubt this).¹⁴⁶ There is an *l ~ r* alternation at the end of the first syllable and a nasal in the Latin form that is not present in the Greek. These are alternations that are not uncommon amongst other substrate lexemes. André (1967: 94) proposes that, if we assume some initial vocalic variation, both forms could be the result of different dissimilation. **k^henin^hdwon* > **k^helindwon* > **kelidwon* > Gk. *χελιδών*; **hinundo* > **hirundo*. This of course does not explain the source of the initial variation. Furnée (1972: 272) cites Akk. *hinundu* ‘swallow’, but this must be a misreading for *sinuntu* (CAD S: 295). André

¹⁴⁵ Though *κίχλη* ‘thrush’ is usually explained as a reduplicated formation to **g^hel-*, EDG (706) is skeptical but separate it from *χελιδών* for other reasons.

¹⁴⁶ Cf. the opposite expectation in Lockwood (2001: 217-18) where the swallow is named after its forked tail, and thus derives from *harundō* ‘reed’, which somehow becomes **herundō*, upon which we would get “popular *i*” for *e*. It should be noted however that at a later date, *hirundō* and *harundō* were being confused with each other. The Appendix Probi has *hirundo non harundo* and Fr. has *aronde* ‘swallow’ < *harundō* (André 1967: 92).

(1967: 94) wants to see *sinuntu* as related, perhaps the source, but finds the difficulties insurmountable. Thus we are left with an irregular match between Latin and Greek.

Anthony Jakob (p.c.) has noticed that Alb. *dallëndyshe* 'swallow' can be reconstructed to a similar but likewise aberrant pre-form. Kroonen (fthc.) reconstructs **ǵ^ho(u)l(H)-(o)nt/d^(h)*-. Çabej (1976 I: 106) had previously interpreted the word as containing the diminutive suffix *-ushe*.¹⁴⁷

Thus the range of this non-IE lexeme includes Latin, Greek, and Albanian. The original quality of the final dental in Alb. *dallënd-* is neutralized by its position (Kroonen fthc.) but on comparison with the *-und-* of Lat. *hirundō* it recalls the Gk. *νθ*-suffixes. Interestingly here however, the Greek form does not have the suffix with a nasal.

lacerna 'a cloak fastened at the neck'

Pre-form: **la/Hk-* | PItal. **lakernā*

Comp.: **la/h₂K-* | PGk. **lakko-* | Gk. λάκκος 'a garment'
**lo/h₃K-* | PGk. **lokka-* | Hsch. λόκκη 'a type of cloak'

■ Irreg. correspondences □ Remarkable phonotactics

Semantics: textiles

Pokorny (674), WH (I: 743), EM (336)

Furnée (1972: 344), EDG (826, 871)

Pokorny (674) derives *lacerna* 'cloak' from *lacer* 'torn, mutilated' as originally a torn piece of cloth used as an overcoat. *Lacinia* 'edge of the fabric' might connect them. WH (I: 743) follows, noting rightly that the suffix *-erna*, whether of Etruscan origin or not, is found attached to clearly Latin bases (s.v. *trabs* for more examples). EM (336) say this is nothing more than a folk etymology, though they indeed take *lacinia* from *lacer*. Greek has λακίς, -ίδος 'rag, tatters of clothes' from the same root as Lat. *lacer* (EDG 826). Its meaning seems to strengthen the connection between *lacer* 'torn' and *lacinia* 'edge of the fabric'.

Crucially, there is another Greek word with semantics more similar to *lacerna*. Furnée (1972: 344) followed by EDG (871) compares Gk. λάκκος 'a garment', which seems to be the same lexeme given by Hesychius as λόκκη γλαμύς, ἔραπτίς 'a type of cloak'. The *a ~ o* alternation is indicative of non-IE origin, and the geminate κκ means it cannot be related to λακίς (at least not in an inherited way). The etymology of *lacerna* seems thus far to have been contaminated by coincidence. PIE **lh₂(n)k-* 'to tear' produced

¹⁴⁷ Orel (1998: 55) had analyzed it as a compound of **dalluan dysh* 'appearing to be double' in reference to the bird's forked tail. This recalls Lockwood's (2001: 217-18) comparison, for the same reason, of *hirundō* to *harundō* '(forked) reed' and draws upon further meanings of *dallëndyshe* in Albanian: 'forked part of a loom framework, frog of a horse's hoof, etc. (Newmark 2005 s.v. *dallëndyshe*). The Latin comparison creates more problems than it solves however, and the extended meanings in Albanian can have originated from the avian meaning.

derivatives referring to an often torn material: cloth. Another lexeme **la/ok(k)-* with non-IE alternations looks nearly identical but refers to untorn cloth. *Lacinia* could derive from either, as the edge resulting from a tear or the finished selvedge.

laena ‘a garment of long-haired wool worn over the *pallium* or *toga*’

Pre-form: **(g^h)leh₂i-neh₂-* | PItal. **(χ)lainā*

Comp.: **g^hla/h₂m/n-ih₂-* | PGk. **k^hlaina-* | Gk. *χλαῖνα* ‘upper garment, mantle’
**g^hla/h₂n-id-* | PGk. **k^hlanid-* | Gk. *χλανίς* ‘a light upper garment’
**g^hla/h₂m-ud-* | PGk. **k^hlamud-* | Gk. *χλαμύς* ‘cloak, robe, mantle’

Hebr. *glōm* ‘wrap, mantle, cloak’, Late Babylonian *gulēnu* ‘upper garment’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: textiles

WH (I: 749-50), EM (337)

Fraenkel (1910-12 II: 178 fn. 2), de Simone (1968-70 II: 283), Furnée (1972: 388), Szemerényi (1974: 148), Breyer (1993: 169), EDG (1635), Rosoł (2013: 107-9), Gernier (2017), Garnier & Sagot (2020: 187), Weiss (2020: 177 fn. 26)

Lat. *laena* is a fleece garment originally used in the religious sphere, as reported by Servius (WH I: 749). It is clearly identical to Gk. *χλαῖνα* ‘upper garment, mantle’, and is generally assumed to have been borrowed from it. Because Latin should have borrowed *χλαῖνα* as ***claena*, Etruscan intermediation is often proposed but always admitted to be problematic. Festus hints at Etruscan origin: *laena vestimenti genus habitu duplicis. quidam appellatam existimant tusce, quidam graece; quam χλανίδα dicunt.*¹⁴⁸ But EM (337) note it is difficult to determine if this means it came from Etruscan. A form like Gk. *χλαῖνα* ought to give Etr. ***χlaina*, then Lat. ***c/glaena*, and a change like **khl > *hl > l* is not attested anywhere else within Etruscan (de Simone 1968-70 II: 283, Breyer 1993: 169). An alternative is that Latin simply represents the reflex of the same pre-form with initial **g^hl-* as Greek. There are no solid examples of this, but it might well have worked the same way as **g^hr- > *hr- > *r-* (cf. Weiss 2020: 177 fn. 26).

This would not be the full explanation, however. Greek also attests to *χλανίς* ‘a light upper garment’. The discrepancy in vocalism can be explained if *χλαῖνα* represents earlier **g^hlh₂n-ih₂-* (Fraenkel 1910-12 II: 178 fn. 2). To these forms can be added Gk. *χλαμύς* ‘cloak, robe, mantle’. Fraenkel (1910-12 II: 178 fn. 2) tried to take them all from a root *χλαμ-*, as did Szemerényi (1974: 148). The latter started with **klam-ja- > *k^hlan-ja- > *k^hlaina-*, and then secondary **k^hlain-id-* being dissimilated to **k^hlan-id- = χλανίς*. Furnée (1972: 388) followed by EDG (1635) instead sees this as a non-IE *m ~ n*

¹⁴⁸ “The *laena* is a type of garment, bipartite in appearance. Some think it named in the Tuscan language, some in Greek; they call it *χλανίδα*.”

alternation. Given all the variation in Greek for this lexeme, and that fact that *χλαῖνα* can be explained as a *-ia derivation of the root **χλαν/μ-*, it is attractive to see Greek as the source form indirectly mediated into Latin *laena*. The suffixes of the Greek words, esp. -υδ- are considered by EDG (1635) to be Pre-Greek.

The source of the Greek words (and thus ultimately Lat. *laena*) seems to be Semitic, cf. Hebr. *glōm* ‘wrap, mantle, cloak’, Aram. *glīmā*, etc. Late Babylonian has borrowed this lexeme as *gulēnu* ‘upper garment’, already producing an *m ~ n* alternation in Semitic (Szemerényi 1974: 148). Thus Rosoł (2013: 107-9) sees *χλαῖνα ~ χλανίς* entering Greek from a Semitic source with *n* and *χλαμύς* entering separately from a source with *m*.¹⁴⁹

lapis, -idis ‘stone, pebble’

Pre-form: **la/Hp-id-* | PItal. **lapVd-*

Comp.: **le/h₁p-ad-* | PGk. **lepad-* | Gk. *λέπας* ‘bare rock, mountain’¹⁵⁰

**la/Hpp-* | PRom. **lappa* | Pt., Sp., *lapa* ‘stone slab’

?**leh₁u/p-*, **lē/tū/p-* | PCelt. **l̥ē/l̥ō/wank-* | OIr. *líe, lía* ‘stone, pillar’

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: geography

Pokorny (678), WH (I: 761-2), EM (340-1), DV (326, 344)

Wood (1910: 82), Hubschmid (1943), Hubschmid (1953: 62-3), Battisti (1959: 147, 332), Hubschmid (1960a: 49), Furnée (1972: 239, 346), Orel (1998: 219), Corominas & Pascual (1984-91 III: 580-581), Untermann (2000: 823-4, 838), Newmark (2005 s.v. *lerē*), EDG (848), Weiss (2010a: 172-5), Kroonen (2013: 328), FEW (V: 173-5), eDiAna (s.v. Lycian B *lacra-*), OED (s.v. *lap*, v.2), van Sluis, Jørgensen & Kroonen (2023: 239)

Lat. *lapis* ‘stone’ has cognates in Italic. U *vapeře* [loc.sg.] (and several other case forms) likely meaning ‘stone seat’ is probably from **laped-* and U *vapeřia* ‘of stone’ from **laped-iā* (Untermann 2000: 823-4).¹⁵¹ The most promising comparandum is Gk. *λέπας* ‘bare rock, mountain’. In previous scholarship, Gk. *λέπας* has been compared to a larger family including Gk. *λέπω* ‘to peel (off)’.¹⁵² Despite the formal difficulties, the link

¹⁴⁹ By rejecting the relationship between *χλαῖνα/χλανίς* and *χλαμύς*, Garnier (2017c) proposes that the former are loans from a Lycian reflex of the inherited Anatolian wool word (Garnier & Sagot 2020: 187 propose Lydian). But the semantic difference between the Greek forms does not seem large enough to separate them.

¹⁵⁰ Cf. also Hsch. *λεπάς*: τὸ τῆ πέτρα προσσχομένον κογχύλιον ‘limpet’ and *λεπάδες*: τὰ πρὸς ταῖς πέτραις κεκολλημένα κογχύλια ὀστρέων ἐλάττω ‘mollusks which stick to rocks’, which are likely derived from the basal meaning ‘rock’.

¹⁵¹ South Picene *vepeten* [loc.sg.] (and related forms) seems to mean ‘monument’ and might also be related, though its vocalism is divergent (Untermann 2000: 838).

¹⁵² WH (I: 761-2) make a chain of analogies to this effect: *lapis* ‘stone’ : *λέπω* ‘to peel (off)’ :: *saxum* ‘stone’ : *secāre* ‘to cut’ :: *rūpēs* ‘rock, cliff’ : *rumpere* ‘to break, rupture’ etc. DV (541) disagrees with the link between *saxum* and *secāre* due to the unexplained *a*-vocalism. He maintains the link between *rūpēs*

between *λέπας* and *lapis* is semantically much more defensible.

This geographically isolated irregular match has led Battisti (1959: 147, 332) and Hubschmid (1960a: 49) to call it a Mediterranean substrate word. To Lat. *lapis* FEW (V: 171) adduces pre-Romance **lappa*,¹⁵³ **l̥bba*,¹⁵⁴ and **lawara*.¹⁵⁵ Hubschmid (1943, 1953: 63) adds Swiss German *lore* ‘heap of collected stones’, Alb. *lerë* ‘heap of stones, pebble bank’, and Gk. *λαύρα* ‘narrow passage, alley’.¹⁵⁶ Furnée (1972: 239) adds Gk. *λάας* ‘stone’ to this list, noting that the form *λαιάι* ‘loomweights’ seems to show a non-IE *a ~ ai* alternation.

Corominas and Pascual (1984-91 III: 580-581) are not so sure that **lappa* is of pre-Romance origin, suggesting other possibilities including cognancy with Engl. (*over*)*lap* (through Gothic). But the English verb is not attested before Middle English (OED s.v. *lap*, v.2) and thus seems to be an inner-English development. There seems to be no other reason to separate **lappa* from the *lapis* beyond its aberrant formation. PRom. **l̥bba* is more aberrant, and **law(a)ra* is both more aberrant and quite isolated. Their appurtenance is therefore more uncertain.

EDG (817) notes in relation to *λάας* that neither Mycenaean *ra-e-ja* /lāhejā/ nor Cypriot *la-o-se* show any trace of a digamma, so there was never a labial element in *λάας*. They therefore reconstruct **lāh-*, which would have to be an unrelated root. Alb. *lerë* ‘heap of

and *rumpere* however (DV 529). If words for ‘rock’ are indeed often derived from verbs for ‘to separate’, then the meaning ‘to peel’ probably counts. An alternative for the analogy *lapis* : *λέπω* would be *lapis* : *lapit* ‘cuts, injures’ (cf. Weiss 2010a: 172-5), and in fact the latter has itself been compared to *λέπω* (Wood 1910: 82). EDG (848) instead considers *λέπω* to be non-Indo-European. DV (335) disagrees with the non-IE interpretation of *λέπω*, connecting it to *lepōs* ‘charm, grace’ and *lepidus* ‘charming’ as well as *λεπίς*, *λοπίς* ‘rind, peel’, *λοπός* ‘scale, rind’, *λεπρός* ‘scaly, coarse’, Alb. *lapë* ‘rag, leaf’, Lith. *lāpas* ‘leaf’, Latv. *lapa* ‘leaf’, Ru. *lépest* ‘petal’, and Lith. *lepūs* ‘weak, soft’. DV admits that the connection of the Latin word is tenuous, but not impossible. I do not find it very appealing, but I wonder if the other words have anything to do with the seemingly non-IE group established in Kroonen (2013: 328) under PGm. **lauba-* ‘leaf, foliage’. In any case, WH (I: 761) suggests connecting Lat. *lapis* to *lepōs* and *lepidus*, but this is far from certain.

¹⁵³ FEW (V: 173-5): In France: Landese *lapa* ‘type of ferruginous rock’ Aurillac *soulapo*, Ytrac *sulápo* ‘cavern on the edge of a river’ (with prefix *sub*). Western Spain: Santander *lapes* ‘stone slabs for covering the roof’, Salamanca *lapa* ‘overhanging cliff that forms a cave’, Montañese *treslape* ‘part of the upper stones that on roofs covers the lower ones’. Pt. *lapa* ‘stone slab’, *solapa* ‘hidden cave’ (with prefix *sub*).

¹⁵⁴ FEW (V: 294): Middle Fr., MoFr. *libe* ‘stone block, stone used in small rubble masonry’, *libbe*, saintongeais *libe* ‘slab’, ‘large flat stone raised in the quarry’, Minot *līpe* ‘slab of stone cut flat to cover a wall’, Beaunotte *līp* ‘beautiful and large stone of masonry that contributes to the solidity of a wall’, dauphinois *lepō* ‘large pebble, cobblestone’, Aveyronnais *libo* ‘slice of turf removed for écobuage’

¹⁵⁵ Hubschmid (1953: 62-3): Campanian *lāver^a* ‘slab of rock’, Friulian *lāvare* ‘large stone slab’

¹⁵⁶ EDG (819) mentions that *λαβύρινθος* ‘labyrinth’ and *λάβρυς* ‘double-headed axe’ might be related to *λαύρα*, and the resulting *b ~ w* alternation is present between other comparanda. Their relationship to each other is doubtful however, nor are their semantics close enough or well understood enough to adduce here. Güntert (1933: 7) too hopefully added what he understood as Lycian *laβra* ‘stone slab?’ and Lydian *laprisa* ‘wall?’. These are misreadings for Lycian *lacra* of unknown meaning and Lyd. *laqrīša* ‘woodwork’ (eDiAna s.v. Lycian B *lacra*-). Hubschmid (1953: 63, fn. 2) rejects the connection to the forms with *b* because he finds no other examples of a *b ~ w* alternation in the Mediterranean.

stones, pebble’ is reconstructed to **laurā* by Orel (1998: 219) but Demiraj (1997: 237-8) proposes a heteroclitic **leh_r-ur* or **leu_r-*. On the other hand, Newmark (2005 s.v. *lerē*) gives the definitions ‘thin mud; mudhole’, ‘grime, dirt’, ‘quicksand’, and ‘scree; stretch of sand with an accumulation of rocks; creek bed full of rocks from the mountains’ with the adverb meaning ‘completely filthy’. This makes an alternative reconstruction of **h₂loi-ro-* (cf. Lat. *linō*, Hsch. ἀλίειν· ἀλείφειν ‘to smear’, W. *Ilynu* ‘to besmear’, Hitt. *ḫalīna* ‘clay?’, cf. DV 344 for Latin and the cognates) possible. Gk. λαύρα ‘narrow passage, valley’ is semantically quite different. Thus the links with forms containing **w* are too uncertain to propose a **p/b ~ *w* alternation.

On a related note, OIr. *líe*, *líā* ‘stone, pillar’ is reconstructed to **līwank-* by Matasović (2009: 242) but the intervocalic consonant could also be **φ* and *e*-vocalism would result in OIr. *í* in hiatus position (van Sluis, Jørgensen & Kroonen 2023: 239). A reconstruction **leφank-* from a root **lep-* matches very well with **lep-* in Gk. λέπας. Thus, this lexeme likely has attestation in Celtic as well. The *-ank* suffix remains obscure.

Thus Lat. *lapis* attests to a non-IE *a ~ e* alternation and gemination in Romance. This must represent a family of non-IE words with the meaning ‘stone’ of the shape **IVP-*.

laurus ‘laurel/bay tree’

Pre-form: **lH(e)u-r-* / **lH(e)ug^{wh}-r-* | PItal. **lauro-*

Comp.: **da/h₂b^h/g^{wh}-n-* | PGk. **dap^hnā-* | Gk. δάφνη ‘laurel/bay tree’
**da/h₂ug^(w)h-n-* | PGk. **dauk^hna-* | Gk. Thess. δαύχνα ‘laurel/bay tree’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, tree

WH (I: 775-6), EM (346)

Cuny (1910: 159 fn. 1), Niedermann (1909: 43-44), Güntert (1932: 21 fn. 1), Lafon (1934: 32-3), Furnée (1972: 132), Furnée (1979: 22), Beekes (2014: 67), Weiss (2020: 176), Kroonen (ftch.)

The connection between *laurus* and the δάφνη family is widely accepted and widely ascribed a substrate origin (Cuny 1910: 159 fn. 1, WH I: 775-6, Furnée 1979: 22, EM 346). The variation within Greek alone is remarkable. Niedermann (1909: 43-44) attributed it to contamination between different substrate forms as they were borrowed into the Greek dialects. Güntert (1932: 21 fn. 1) attributed it to an *l ~ d* alternation in Asia Minor.¹⁵⁷ Perhaps this is supported by the Hesychius gloss λάφνη· δάφνη. Περγαῖοι, which in any case shows that the alternation existed within the Greek-speaking world. This must be different from the *l ~ d* phenomenon termed the “Sabine *l*”, which occurs in

¹⁵⁷ Often compared are also λαβύρινθος ‘labyrinth’ vs. the Carian deity Δαβρανδός, but Güntert here mentions the Greek name *Lygdamis* vs. Akkadian *Dugdammē*, the name of a Cimmerian king who settled in Cilicia.

inherited words (cf. *lacruma* ‘tear’ vs. Gk. δάκρῦμα ‘tear’).

Furnée (1972: 132) links Gk. δάφνη through its variants with a velar like Thess. δαύχνα to δαῦκος ‘several types of umbelliferous plant’, noting the *-n-* suffix in δάφνη and considering the φ ~ χ/κ a substrate alternation. EDG (306) agrees with the connection, reconstructing a **dak^w-(n)-* for the Greek forms. Semantically the pairing of ‘laurel tree’ with ‘umbelliferous herb’ seems difficult to defend. The analysis as descendants of a non-IE phoneme similar to **k^w* still holds however. Even excluding δαῦκος, the alternation within Greek is between *-αφ-* and *-αυχ-* (i.e. never ***αυφ-*). This suggests one of two things. 1) The velar aspect of the non-IE labio-velar was interpreted variously behind (**k^w > p*) or in front (**k^w > uk*) of the velar, and Lat. *laurus* shows that the labial aspect overtook the velar aspect of the articulation. 2) Alternatively, *a ~ au* vocalic alternation before a **g^{wh}* would in the latter case have triggered the *boukolos* rule. The result of **laug^h-ro-* (with *boukolos* **ug^{wh} > ug^h*) within Latin would be **lauhro-* > **lauro-* (Weiss 2020: 176 gives examples of this development of **g^{hr}* word initially¹⁵⁸). If the first situation occurred, Beekes (2014: 67) suggests that the *n*-suffix of the Greek forms might have something to do with the aspiration. If the second occurred, the velar must have been borrowed with aspiration.

Latin has an *r*-suffix in *laurus* while the Greek forms have an *n*-suffix. The same *r ~ n* suffixal alternation, this time within Greek, occurs after an aspirated velar in Gk. βλήχρον vs. βλήχρον ‘fern’ (s.v. *filix*). Perhaps this points to an origin in the same substrate, but while the fern word also occurs in Germanic **brekna(n)-*, the laurel word is limited to Latin and Greek and refers to a Mediterranean plant.¹⁵⁹ It potentially also appears in Georg. *rapindi* ‘laurel tree’ (Lafon 1934: 32-3, Furnée 1979: 22), which attests to something like the Gk. *vθ*-suffix otherwise unattested in the Greek laurel words. However, if the lexeme originally contained something akin to **g^{wh}* as the Latin and Greek comparanda seem to suggest, then the Georgian word must be a loan from Greek; it is due to Greek sound laws that the *p* arose (cf. Kroonen fthc. See s.v. *ervum* for a similar suggestion that Georg. *erevandi* is likewise a Greek loan). The Georgian form further shows that its Greek source form started with *r*, which would have to be factored into the *d ~ l* alternation between the attested Latin and Greek forms.

lēns, -tis ‘lentil’

Pre-form: **l(e)nt-* | PItal. **lenti-*

Comp.: **lq^d-ur-* | PGk. **lat^huro-* | λάθυρος ‘pulse, chickling’

?PBerb. **līntī-* ‘lentil’ | Sous Berber *tilintit ~ tinilitit* ‘lentil’

■ Irreg. correspondences

□ Remarkable phonotactics

¹⁵⁸ Word-internally, e.g. Meillet and Vendreyes (1979: 73) argue that **(^g)^h > g /_u* because of cases like *figūra* ‘form’, *figulus* ‘potter’, and *ligurriō* ‘to lick’. But Weiss (2020: 87 fn. 61) suggests these forms may be analogical to *fingō* and *lingō* (where **(^g)^h > g /n_* is regular).

¹⁵⁹ Perhaps this does not exclude an origin in the same substrate, see §4.2.2.5.

Semantics: plant, domestic

WH (I: 783), EM (351), DV (334)

Fraenkel (1962-5: 359), Puhvel (III: 19-20), Boutkan & Kossmann (1999: 95), EDG (882), Beekes (2014: 41), Weiss (2020: 255), Simon (fthc.)

The similarity yet irregularity between *lēns* and its comparanda has been widely noted and suggested to be indicative of a non-IE origin (WH I: 783, EM 351, DV 334). At the most straightforward level, PItal. **lenti-* and PGk. **lat^huro-* are the most certain independent forms. EDG (882) remarks that the Greek word only barely resembles the rest, and it is indeed the most semantically remote. However, the vocalism between it and the Italic form can be explained if both were borrowed with a syllabic nasal, with the only remaining alternation being the final dental. Given that these loans are assumed to have occurred after the disintegration of PIE, there is no reason to believe the Greek form actually descends from PIE **d^h* and we can rather take the unvoiced aspirate at face value: Pre-PItal. **lnt-i-*, Pre-PGk. **lnt^h-ur-*. While Latin seems to have borrowed the lexeme as (or produced) an *i*-stem, there is always a chance that it was originally a root noun: consonant stems of Latin are a result of the merger of PIE consonant and *i*-stem classes (cf. Weiss 2020: 255). This means that, between the two forms, perhaps only Greek has added a suffix, *-ur-*, which Beekes (2014: 41) argues is Pre-Greek.

The independence of other comparanda is difficult to determine. If the *i*-suffix of Latin *lēns* is not original, then Baltic and Slavic are difficult to explain unless loans from Latin. PSlav. **lętja-* (cf. OCS *lęšta*, Ru. *ljač* ‘lentil’) might attest to a syllabic nasal like Latin and Greek, but PBalt. **lęši-* cannot, suggesting either vocalic alternation resembling IE ablaut or a borrowing from a language whose reflex of the syllabic nasal is *e* (Italic or Slavic). Such borrowing scenarios are not straightforward however. The *š* of Lithuanian *lęšis* ‘lentil’, despite the other indicators of a loan from another form mentioned above, does not seem to be obviously sourced from Latin or a Slavic language¹⁶⁰ (cf. Fraenkel 1962-5: 359), though perhaps Germanic is the source. If taken at face value, the *š* would reconstruct to PIE **k*, but it is clear that, if not borrowed from another IE language, the Lithuanian attests to a sibilant in its source form. OHG *linsa* has been suspected to be a loan from Latin (WH I: 783, EM 351), but Kluge and Seebold (1989: 444) note that in such a case we should expect the oblique stem **lent-* to be borrowed. We find attested the reflex of a sibilant, but the nasal could perhaps be neutralizing the dental element of an affricate here (cf. MHG *banse* ‘lean-to shelter’ < **b^hond^h*, Kroonen 2013: 52). Its isolation within German and the fact that all comparanda attesting to *e* vocalism rather than a syllabic nasal are insecure make it difficult to accept as an independent comparandum. Thus Latin and Greek attest to a **t ~ *t^h* (as if PIE **t ~ *d^h*) alternation, with the possibility that this further alternates with some sort of sibilant.

Sous Berber *tilintit ~ tiniltit* ‘lentil’ (with feminine *t-...-t* circumfix) can be reconstructed

¹⁶⁰ Latv. *lēca* with its *c* is a loan from Slavic.

to Proto-Berber **l̄int̄-*. The other Berber languages have Arabic loans for the lentil word, which means that the Sous form could be the original Berber lexeme, but it could also be a later loan from Romance (Boutkan & Kossmann 1999: 95). Pisani (1967: 403)¹⁶¹ compared Hitt. *ḫalenzu-*, which at the time was glossed as *Wasserlinse*. Puhvel (III: 19-20) however shows that *ḫalenzu-* actually means ‘overgrowth’, and that the association between *l̄ens* and *ḫalenzu-* would never have come about if the gloss were English ‘duckweed’ instead of German *Wasserlinse*.¹⁶²

l̄ilium ‘lily’

Pre-form: **(H)leili-*, **(H)iHli-* | PItal. **leilio-*, **l̄ilio-*

Comp.: **lei(h₁)ri-* | PGK. **leirio-* | Gk. *λείριον* ‘lily’

Coptic vars. *hr̄eri*, *hl̄ēli*, *hr̄ere* < Egypt. *ḫrr.t* ‘flower’

?**Hol-* | PANat. **ʔol-* | Hitt. *alēl-*, *alil-* ‘flower, bloom’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, flower

WH (I: 801), EM (358), DV (341)

Meillet (1908: 163), Schuchardt (1918: 26-7), Cohen (1931: 37-8), Benveniste (1954: 43), Hubschmid (1960a: 38), Hemmerdinger (1968: 240), Holton Pierce (1971: 105), Furnée (1972: 369-70), Vycichl (1983: 310), Puhvel (I: 32-3), Vycichl (1990: 94), Biville (I: 365, II: 23), Orel (1998: 234), Trask (2008: 29), EDG (845), Schrijver (2018: 362), Weiss (2020: 507), Simon (fthc.)

The *l ~ r* alternation between Lat. *l̄ilium* and Gk. *λείριον* was early on remarked upon as an indicator of Mediterranean substrate origin (Meillet 1908: 163) and this interpretation continues (WH I: 801, EM 358, Weiss 2020: 507). Neither form is likely borrowed from the other (EM 358, Biville I: 365, II: 23). Further relationships outside of Latin and Greek are complicated, but there is frequent consensus that Egypt. *ḫrr.t* ‘flower’ via its Coptic descendants *hr̄eri*, *hl̄ēli*, *hr̄ere* are related (EM 358, DV 341), perhaps as the source (WH I: 801; uncertainly Hemmerdinger 1968: 240, Simon fthc.). Hubschmid (1960a: 38) reports that the Egyptian word is only attested from the 18th dynasty onwards, and its status as an inherited Afroasiatic word does not seem to be guaranteed. On the other hand, Cohen (1931: 37-8) notes Berber *alili*, *ilili*, etc. ‘oleander’, *ilillii* ‘flower’ in Oromo,¹⁶³ and ‘*elēdī*’ ‘flower’ in Harari.¹⁶⁴ With attestation in Berber,

¹⁶¹ In *Paideia: rivista letteraria di informazione bibliografica* 22 (*non vidi*, apud Puhvel III: 20).

¹⁶² Simon (fthc.) argues for a Hattic origin of the Hittite word. He defends the connection with *l̄ens*, adding Hung. *békalencse* ‘duckweed (literally ‘frog-lentil’)’ as another example of a comparison between lentils and duckweed. However, if the Hittite word refers to other types of ‘surface growth on stationary water’ like algae, leaves, etc., then duckweed cannot be ruled out as a secondary semantic development. The formal resemblance to *l̄ens* would be coincidental.

¹⁶³ Updated from his spelling *ilili* and referring to the language as Galla.

¹⁶⁴ Updated from his spelling *elad*.

Egyptian, Cushitic, and Semitic, the word is present several Afroasiatic families. Inherited status in Egyptian would make it more likely that an Egyptian source is the ultimate origin of the Latin and Greek words. The lexeme is, however, quite isolated within each of the families. Schuchardt (1918: 26-7) even interprets the Berber words are loans from Latin. Thus its native status within Afroasiatic remains unclear.

Hemmerdinger (1960: 38) had suggested Hitt. *alel-*, *alil-* ‘flower, bloom’ could be a more proximal source of the Greek word. There is wide consent that the Hittite word is related to the Egyptian word (Benveniste 1954: 43, Furnée 1972: 269-70, Puhvel I: 32-3) as a Mediterranean Wanderwort. Simon (fthc.) rejects a comparison between the Egyptian and Hittite forms because of the initial vowel of Hittite against Egypt. *hrr.t*. However, Egyptian is transcribed without vowels, and the initial *hr-* does represent a consonant cluster. Instead, Vycichl (1990: 94) reconstructs */harīra.t/. It is thus easier to get the Hittite word from Egyptian than to do so with the Latin and Greek forms (cf. the reservations on the relationship between the Egyptian forms and Latin/Greek forms in Holton Pierce 1971: 105, Vycichl 1983: 310, *pace* Simon fthc.). Schrijver (2018: 362) suggests the initial *a-* of Hittite in comparison to the Latin and Greek forms might be an example of the substrate *a-* prefix.

The semantic difference between Latin and Greek ‘lily’ on the one hand and ‘flower’ elsewhere does not seem problematic to interpret as a semantic narrowing. Perhaps it occurred in the donor language. There exist widespread lookalikes with the meaning ‘flower’ from Estonian *lill* to Basque *lili* and Alb. *lúle*.¹⁶⁵ Orel (1998: 234) is probably correct in doubting that Alb. *lúle* ‘flower’ is loan from Latin, as it would require the assumption of *i > u / L_*. Hubschmid (1960a: 37) notes the difficulty in assuming that Basque *lili* is a loan from Latin or Romance. Vasconic */l/ was rhotacized intervocalically in the early medieval period (cf. Trask 2008: 29 for the date), so it would have to be a late loan. But then one expects the meaning ‘lily’. Instead it suggests a pre-form with fortis */L/. On the other hand, Basque *lora* ‘flower’ alongside Bearnaise *lole* and Tarbes *lolo* ‘flower’ suggest that there was also a Vasconic */lola/ ‘flower’ with lenis */l/ (Hubschmid 1960a: 38). Whether these further forms are to be counted as independent comparanda seems unclear.

What is clear is that Lat. *līlium* and Gk. *λείριον* are independent loans from a third source. If Egyptian *hrr.t* is the ultimate source (and not itself a loan), it is unlikely to be the most proximal source (i.e. there was an intermediary). Hitt. *alel-*, *alil-* is a relative, but may be a borrowing from Egyptian.

malva ‘mallow’

Pre-form: **ma/Hl-Vu/g^{w(h)}*- | PItal. **malVwā-*

Comp.: **ma/o/Hl-a/og^h*- | PGk. **mVIVk^hā-* | Gk. *μαλάχη*, *μολόχη* ‘mallow’

¹⁶⁵ There is even Turk. *lale* ‘tulip’, though this and several other forms in surrounding languages are presumably from MoP *lāle* ‘tulip’ (Hubschmid 1960a: 39).

**ma/h₂l-b/g^wa-k-* | PGk. **malbak-* | Gk. μάλβακα [acc.] ‘mallow’

Hebr. *mallūah* ‘Atriplex’ < **mallūh*

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, wild

WH (II: 17-18), EM (380), DV (361)

Cuny (1910: 162), Lafon (1934: 40), Leumann (1977: 214), Klein (1987: 349), Nussbaum (1997: 190), Rata (2008), EDG (896), Rosol (2013: 104-5, 109-11), Beekes (2014: 69), Fenwick (2016: 453), Weiss (2020: 175)

The most secure comparandum for Lat. *malva* is the group of Greek words including μάλαχη ‘mallow’, often suggested to be independent borrowings from a Mediterranean substrate since a common pre-form cannot be reconstructed (Cuny 1910: 162, WH II: 17-18, EM 380, DV 361, EDG 896).¹⁶⁶ The Greek variants have differing vocalism and aspiration amongst themselves. If the variant μάλβακα has its β from **g^w*, the situation would be somewhat reminiscent of the δάφνη/δάυχνα (**kw*/**wk*) pair (s.v. *laurus*). No labial element appears in μάλαχη but the vocalism of μολόχη could suggest a rounded vowel was original. The Latin form does not attest to a final velar, for which there are several possible explanations. If it represents **malawa-*, then it fits into the δάφνη-δάυχνα-*laurus* pattern. (Cf. also the **g^w(^h)* ~ **μ* alternation in some of the comparanda of *fungus*.)

There are several other possible explanations, all necessarily *ad hoc*: **g^h* or non-IE **k^h* could have weakened to an *h* and then have been lost in word-final position or have been obscured by the development **malV_uak^h/g^ha* > **malV_uaha* > **malVwā*. Or **χ* attracted the **w* to yield **malwaχa* > **malax^wa* > **malava* > *malva*. For the latter, compare the reconstruction of PGk. **mal^wak-*, with Pre-Greek labialized *l*, proposed by EDG (896) and Beekes (2014: 69). All Proto-Italic reconstructions require a vowel between **l* and **μ*; otherwise, it must have entered Latin recently enough that the **l_μ* was not assimilated into a form like ***malla*.^{167,168}

All sources also mention a connection with Hebr. *mallūah* ‘saltbush/orach (genus *Atriplex*)’ or ‘a lettuce-like vegetable’. Several species of both orach and mallow are consumed as leaf vegetables and the leaves of both plants share some general similarities, so the comparison is not without reason. The Hebrew word is a hapax,

¹⁶⁶ An Armenian form *balbak* ‘a plant, watercress, dill, or mallow’ looks close to the Greek variant μάλβακα, but its *-ak* could have been added within Armenian. Thus it could be a loan from Georg. *balba* ‘mallow’, which Lafon (1934: 40) asserts is itself a loan from Latin. The initial *b* for *m* in the Georgian form needs further investigation.

¹⁶⁷ On the change **lw* > *ll*, cf. Leumann (1977: 214), Nussbaum (1997: 190), Weiss (2020: 175). Fenwick (2016: 453) reconstructs **mh₂l-u-eh₂-* to a root **meh₂l-* ‘type of cultivated plant or herb’ that would also underlie Gk. μάλον ‘apple’ and μῶλον ‘magical herb’. But this should also have given Lat. ***malla*.

¹⁶⁸ Hubschmid (1960a: 60) proposed a Mediterranean *-ua* suffix, which would have to have developed from earlier **-Vua*.

occurring in Job 30:4, which is known for its poeticized language (cf. Rata 2008). The word ends in a *pataḥ gnuva* and goes back to **mallūh*.¹⁶⁹ Klein (1987: 349) derives Hebr. *mallūaḥ* from *melaḥ* ‘salt’. There is an adjective Hebr. *mallūaḥ* meaning ‘salty, saline’, cognate with Arab. *melih* ‘salty’, from *milḥ* ‘salt’ (cf. an alternative name for orach in English, namely ‘saltbush’). If the Latin and Greek forms are indeed related to the Hebrew, such a Semitic etymology would suggest that they derive ultimately from a Semitic source. There remains the possibility, since the Hebrew adjective has a good Semitic etymology but the noun is a hapax describing a plant in a text that is already known for using unusual words, that the two otherwise homonymous words do not actually originate from the same source and that the link with the salt family is folk etymological.¹⁷⁰ If Latin and Greek have the word from a Semitic source, they both underwent the same semantic shift (orach > mallow). Otherwise, all three are from another source.

menta ‘mint’

Pre-form: **m(e)nt-* | PItal. **mentā-*

Comp.: **mind^h-* | PGk. **mīnt^h-* | Gk. μίνθη ‘mint’, vars. μίνθη, μίνθος

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, wild

WH (II: 72), EM (398)

Meillet (1908: 162), Hester (1964: 360), Ačařyan (1979 IV: 623), Biville (I: 145), EDG (995), Weiss (2020: 148), Kroonen (ftbc.)

Lat. *menta* is often considered borrowed along with Gk. μίνθη from a third source. Despite the substitution of aspirates being normal in early loans from Greek, the correspondence of Lat. *e* ~ Gr. ι is irregular (Meillet 1908: 162, Hester 1964: 360, EM 398, EDG 955, Biville I: 145), *pace* WH’s (II: 72) *ad hoc* suggestion of a “replacement” of *-int-* through *-ent-*. While Lat. *e* raises to *i* in the sequence *m _ nV* (cf. *minor*), it is blocked in the sequence *m _ nC* (cf. *mentis*) as maintained by Weiss (2020: 148). But the inverse situation, with *i* being lowered to *e* in a sequence *m_nC* does not occur (cf. *mingō*

¹⁶⁹ Greek does borrow Semitic *h* and *ḥ* as *χ* (cf. Rosol [2013: 104-5, 109-11] Gr. χαλβάνη ‘galbanum’ < Hebr. *helbinā* ‘id.’; Gr. χρυσός ‘gold’ < Phoen. *ḥ[u]r[ḥ]i[s]* ‘id.’). That the Greek forms are so diverse (μαλάχη, μολόχη, μολάχη, and μάλακα) means that the easy explanation from something like **mallūahā* is no longer as elegant.

¹⁷⁰ More evidence for this is that the meaning ‘a plant name: sea orach(?)’ occurs otherwise only in Aramaic *mallūh/mallūhā*. *Mallūah* was mistranslated as ‘mallows’ in the King James Bible, and a close link between the words is indeed suggested by other factors. In the Septuagint, μολόχη occurs but it does not translate *mallūah*. Instead, in Job 30:4, ἄλιμα (clearly related to salt) translates *mallūah* and μολόχη occurs in what seems like an extra clause inserted in the Greek for Job 24:24: ἐμαράνθη δὲ ὡσπερ μολόχη ἐν καύματι ‘they are withered like mallows in burning heat’. This clause does not occur in the Hebrew. Could it be that the Greek word had been loaned from Hebrew and was therefore associated with the verse, but it was no longer clear to which Hebrew word it referred because its meaning had changed?

and *mintriō*). A possible explanation is a loan from Greek through Oscan.¹⁷¹ EDG (955) considers the Greek forms to be of Pre-Greek origin because of the attestation of the alternate ending in *-ā*. The words look like an example of the Pre-Greek *vθ*-suffix. This would leave little more than *m* (or **sm*) as the root, but we cannot exclude the possibility of such a phonologically simple root in a substrate language.

Outside of Latin and Greek there are no comparanda. The comparison mentioned by WH (II: 72) with Georg. *p'it'na* 'mint' is phonologically too far removed and is better explained with a view toward unrelated MoP *pūdina* (also the source of Arab. *fūḍanaj* > Arm. *fōtanj*, cf. Ačāryan [1979 IV: 623] on the Persian origin of these forms) and a large number of Indic and other Iranian forms.

merula 'blackbird'

Pre-form: **(H)mes-(a/e/o/u)l-* | PItal. **mesa/e/o/ulā-*

Comp.: **(H)mes(a)l-* | PCelt. **mesal-(s)ka-* | W *mwyalch*, Bret. *moualc'h* 'blackbird'

**h₂/ems-lo-* | PGm. **amslōn-* | OE *ōsle*, OHG *amsala* 'blackbird'

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: animal, bird

Pokorny (35-6), WH (II: 77-8), EM (400), DV (375)

Demiraj (1997: 264-5), Schrijver (1997: 307-11), Kloekhorst (2008: 292), Matasović (2009: 268), Kroonen (2013: 25), Neri (2017: 565-568), Thorsø & Wigman et al. (2023: 109)

Latin *merula* 'blackbird' is convincingly linked with Celtic and Germanic words for 'blackbird' that all attest to a multi-syllabic "root" with a non-IE ablaut pattern **aCC-* ~ **CVC-*, a classic indicator of a non-IE word (Schrijver 1997: 307-311, DV 375, Matasović 2009: 268, Kroonen 2013: 25).

Most recently, Neri (2017: 565-568) defends an IE etymology for this family of words, reconstructing Hitt. *hanazana-* of disputed meaning (but seeming at least once to mean

¹⁷¹ Alessio (1944a: 139-41) gives a similar case that does not have attestation in Latin but rather in several Italian dialects. He compares PRom. **plenta* 'clod of earth' to Gk. *πλινθος* 'brick'. He argues that this is an additional case of a Mediterranean substrate word shared between Greece and the Italian peninsula. Interestingly, he writes "[le forme] possono risalire ad una base *PLENTA, che riterremo di origine preindoeuropeo mediterraneo." It must be traced back to the Mediterranean substrate because it is not part of the Oscan substrate. But Cid Swanenvleugel (p.c.) has suggested to me that, since the reflex PItal. **i* is lowered in Oscan and Umbrian until it becomes similar to the reflex of PItal. **ē*, the *e* of Lat. *menta* and PRom. **plenta-* might be the result of a borrowing from Oscan, which would have changed the *i* of an early loaned Gk. *μίνθη* and *πλινθος* to *ē*. Alessio presumably rejects that these could be from the "Oscan substrate" in Latin because they are not attested in Oscan and are unlikely to be inherited. But theoretically, nothing is stopping Oscan from being a mediator of substrate vocabulary into Latin.

‘black, dark’) and Skt. *ásita-* ‘dark-colored, black’ as **h_{2/3}m̥s-i-to-* ‘having a dark color’ < **h_{2/3}m̥s-i-* ‘dark coloration’ to a root **h_{2/3}ems-* ‘dark’. According to him, Germanic **amslōn-* reconstructs to **h_{2/3}éms-lah₂-* ‘the black one’, a vřddhied and feminized derivation from an Eigenschaftsadjektiv **h_{2/3}m̥s-ló-* ‘dark, black’. The Italic and Celtic reflexes would start from **h_{2/3}mes-elo-* as either a substantive use of an adjective of the shape Gk. μεγάλη, PGM. **mekila-* or as a vřddhied form of (potentially diminutive) **h_{2/3}m̥s-élo-*. This seems to be the best treatment of this group as potentially inherited, but it requires a Schwebelaut-like difference between **h_{2/3}éms-leh₂-* and **h_{2/3}mes-elo-*.

Further circumstantial arguments make this less compelling. Hittite *hanazana-* and Skt. *ásita-* can also be reconstructed to a root with syllabic **ŋ* (cf. for Hittite Kloekhorst 2008: 292) and may thus be a different root with **m* in the forms that all mean specifically ‘blackbird’. Nor is the **aCC-* ~ **CVC-* alternation limited to this word (see §3.3.2). Given the potential of this non-IE pattern against the problems in the IE etymologies offered for Lat. *merula* etc., this family likely represents loanwords from a non-IE language of Europe.¹⁷²

mūlus ‘mule’

Pre-form: **mu(g^(h)/k)s-lo-* / **mus(g/k/?g^h)-lo-* | PItal. **mus(k)lo-* / **mu(k)slo*

Comp.: **musk-* | PSlav. **mъskъ* | ORu. *mъskъ*, RuCS *mesk* ‘mule’

**muk-lo-* | PGk. **muklo-* | Gk. μύκλος ‘lascivious, lewd’

**mug^(h)/k(s)-lo-* | PGk. **muk^hlo-* | Phocian (Hsch.) μυγλός· σκολιός.
ὄχευτής, λάγνης, μοιγός, ἀκρατής. Φωκεῖς δὲ καὶ ὄνους τοὺς ἐπὶ
ὄχειαν πεμπομένους ‘crooked, lewd, lecherous, uncontrolled’ and
‘stud donkey’

**musk-lo-* | PGk. **musklo-* | Hsch. μύσκλοι· σκολιοί ‘crooked, unrighteous’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: animal, domestic

WH (II: 125-6), EM (420), DV (394)

Frisk (1960-72 II: 267-8), Orel (1998: 279), EDG (978), Furnée (1972: 133, 299)

The diminutive Lat. *muscellus* is often taken as metathesized from **muxellus*, suggesting a preform in **-ks-*, though Gk. μύσκλοι and the Slavic forms show that a variant with **-sk-* was in circulation. In any case it proves the erstwhile existence of a velar and sibilant (WH II: 125-6). Alb. has *mushk* ‘mule’, which Orel (1998: 279) calls an areal Balkan word along with the Slavic forms. It seems likely that the Albanian is a loan from Slavic.

¹⁷² Demiraj (1997: 264-5) suggests that, if Alb. *mëllënjë* ‘blackbird’ has the suffix *-(V)një*, its base *mull-* could through PAIb. **mē/āIV-* be from a similar pre-form to Latin *merula*. But several alternative etymologies also exist.

The Greek forms seem semantically remote, but *μύκλος* ‘lascivious, lewd’ is at least once used as an epithet of a pack-mule. Taken along with the Phocaeian meaning furnished by Hesychius, it seems that Greek words do indeed have (literally) asinine semantics, and are generally taken as comparanda (DV 394, EDG 978, Furnée 1972: 133, 299). The χ might suggest the reconstruction of $*g^h$, however in light of the other comparanda (including Gk. *μύσκλοι*) with a non-aspirated velar sibilant, the aspiration could be the result of a following sibilant. Thus the $\chi \sim \kappa$ alternation need not, as EDG (978) concludes, be a Pre-Greek feature no matter how frequent that alternation appears in other Pre-Greek words. Instead, it is likely a Wanderwort (Frisk 1960-72 II: 267-8, WH 125-6, EM 420, DV 394), like *asinus* (s.v.), from the homeland of donkeys in North Africa or the Levant, arriving in the forms **musk-(lo-)*, **muks-(lo-)*, and **muk-(lo-)*.

This **sk ~ *ks* alternation occurs also in the comparanda of *viscum* ‘mistletoe’, where the metathesis is likewise unexpected. Šorgo (2020: 459) notes this as a feature of the Germanic substrate, identifying at least one further example outside of Latin (PGm. **pahsu-* ‘badger’ vs. PCelt. **tazgo-*, **tasko-*, **taks-* ‘badger’). It seems like a non-IE feature rather than *ad hoc* metathesis.

nux ‘nut’

Pre-form: **(k)nu-k-* | PItal. **(k)nuk-*

Comp.: **knu(H)-* | PCelt. **knū-* | OIr. *cnú* ‘nut’
**kn(e/o)u(H)-* | PCelt. **knows-* | MW *cneu*, MBret. *cnou* ‘nuts’, etc.
**knu-d-* | PGm. **hnut-* | ON *hnot*, OE *hnutu*, OGH *nuz*, etc. ‘nut’

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: plant; nut

Pokorny (558-9), WH (II: 191-2), EM (453), DV (418, 420)
 Pedersen (1893: 251), Hirt (1907: 173), Otrębski (1939: 173), Thurneysen (1946: 31), Martinet (1955), WH (II: 185-6), Georgiev (1971: 273), Strunk (1993: 425-9), Schrijver (1995: 326-33), Kroonen (2012a: 248), Kroonen (2013: 237), Matasović (2009: 212, 250), van Sluis (fthc.)

The likelihood that this family of comparanda is of non-IE origin comes from the peculiarity of the different suffixal extensions (for this morphological analysis, cf. WH II: 191-2 with lit.), different in each branch (**H* in Celtic, **k* in Italic, and **d* in Germanic).¹⁷³

DV (420) argues that the root shape **knu-* with no full grade looks non-IE of itself.

¹⁷³ Previously frequently mentioned was that *nux* is metathesized from the **dnuk-* in Germanic (Pedersen 1893: 251, Hirt 1907: 172, Georgiev 1971: 273; the latter further compares Gk. *ἀγνύς* ‘(nut-shaped?) weaving stones’). Otrębski (1939: 173) compared Lat. *nux* and Gk. *κάρυον* ‘nut’ via an *r ~ n* alternation and metathesis.

Schrijver (1995: 330) indeed reconstructs MW *cneu* et al. to Late Proto-Brythonic **know-* < **knuu-*. However, he (pp. 326-33) also provides evidence to show that **eu* (and **euH*), **ou* (and **euH*), and **uu* (including when from **uHV*) all became Late Proto-Brythonic **ow*. Thus the Brythonic forms obscure the difference between a full-grade and zero-grade, even in a laryngeal-final root, and so they could indeed derive from a full grade **kneu(H)-* / **knou(H)-*. But a laryngeal is not actually required in the Celtic forms. OIr. *cnú* does not require the reconstruction **knu-H-*, because final vowels in open monosyllables are regularly lengthened in Old Irish (Thurneysen 1946: 31). Thus, within Celtic, we find what could be construed as an IE ablaut pattern to an IE root **kneu-*.

It is the the Latin and Germanic forms which make the family look non-IE. Latin *nux* derives from a zero-grade of the root in question with a **-k* suffixal extension.¹⁷⁴ DV (418, 420) further adduces *nūgae* ‘worthless things, nonsense’ from a form **knūg-*. If indeed related to *nux*, the differing vowel length and voiced as opposed to unvoiced velar would yield a non-IE pattern within Latin. However, besides the argument that words for ‘trifle’ are sometimes formed from lexemes for nuts or seeds (cf. English *peanuts*), the semantics are not close enough to connect these two words within Latin.¹⁷⁵ Germanic **hnut-* derives from a zero-grade of the root with a **-d* suffixal extension (Kroonen 2009: 221-2, 2013: 237). Kroonen (2013: 237) notes that it inflects as a root noun, which is an archaic, non-productive noun category within Germanic and might point to non-IE origin (cf. Kroonen 2012a: 248).

As to the discrepancy between the suffixes, Kroonen (2012a: 248) suggests it might be the reflex of something like a glottal stop. The option of reconstructing **-H* for Celtic would fit into this scenario, but interestingly, in the two other cases of this phenomenon that van Sluis (fthc.) identifies (PGm. **bīōn* ‘bee’ and the *caput* family), its reconstruction is not required either. In any case, the mismatching suffixes¹⁷⁶ within Italic and Germanic, otherwise without explanation, and the fact that the Germanic noun

¹⁷⁴ One might suggest that a pre-form **knuH-s* might yield *nux*, related to the way that e.g. *-trīx* might have arisen from **tr-iH-s* (proposed by Martinet 1955). But Schrijver (1991: 148-54) summarizes arguments as to why **-ks* is not likely to have developed from **-Hs*, and we have seen that the presence of a laryngeal is not actually required by the Celtic forms.

¹⁷⁵ Instead, *nūgae* is very similar in meaning to *naucum* ‘trifle, worthless thing’. WH (II: 185) finds it difficult to connect them, and indeed it would require accepting an **ū ~ *au* or perhaps **eu ~ *au* alternation as well as a voicing alternation. Thus this might be an unrelated substrate root. Strunk (1993: 431) argues that there is evidence that *naucum* referred to the inedible parts of nuts, that Latin speakers considered *naucum* and *nux* related, and that the whole family is inherited. He additionally argues for IE **au/u* ablaut, with support from *pau-cus* ‘small’, *pau-per* ‘poor’, vs. *pu-sillus* ‘tiny’, *pu-er* ‘boy’. But 1) **au/u* ablaut is not the only explanation for such a distribution; a full-grade/zero-grade ablaut (**-Hu-* / **-eHu-*) also works (cf. DV 450, 496). 2) Even this latter explanation cannot account for the long *ū* of *nūgae*, ruling out the possibility that all three words are inherited cognates. 3) In any case, the semantics of *nūgae* and *naucum* are closer to each other than either is to *nux* and are best kept separate from it.

¹⁷⁶ Note that the designation of these elements as suffixes is itself biased toward an Indo-European interpretation. In part to explain how the consonants are different, proposing a suffix also keeps PItal. **knuk-* from going back to an illegal **C₁eC₂-* root structure. But the element could simply be part of a non-IE root **knu₂-*.

inflects as a root noun suggest that this lexeme is of non-IE origin.

orca ‘large-bellied vessel, butt, tun, esp. for storing fish’

Pre-form: **H(o)rk-* | PItal. **orkā-*

Comp.: **H(o/u)rk-* | PItal. **urkejo-* | Lat. *urceus* ‘pitcher, water-pot, ewer’
 **H(o/u)rk-n-* | PItal. **urknā-* | Lat. *urna* ‘vessel for drawing water, urn’
 **Hurg^h-* | PGk. **urk^hā-* | Gk. ὕρχη ‘earthen vessel used for salting fish,
 etc.’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: vessel

WH (II: 220, 838-9, 841), EM (467, 754, 755)

Curtius (1858 I: 315), LS (s.v. *orca*), Cuny (1910: 160), Ribezzo (1934b: 124), Bertoldi (1939a: 290), Ernout (1946: 49), Chantraine (1968-80: 821), de Simone (1968-70 I: 138, II: 271-5, 287), Furnée (1972: 137, 361), Breyer (1993: 219), Biville (I: 231-2), EDG (1537), Beekes (2014: 67), Weiss (2020: 153, 195)

LS (s.v. *orca*) take Lat. *orca* ‘large-bellied vessel’ to be a transferred meaning of *orca* ‘whale’, but they are certainly separate words (WH II: 220, EM 467).

Orca in the cetaceous sense has been suggested to be a loan from Gk. ὄρυξ (acc. ὄρυγα) ‘pickaxe, type of whale (probably narwhal)’ (WH II: 220, EM 467) but Biville (I: 231-2) notes problems with this explanation. The borrowing of Gk. γ as Lat. *c* is not usual, and Etruscan intermediation must be proposed to explain it (Ernout 1946: 49, Breyer 1993: 219). There is however no attested Etruscan form to prove this. Biville further notes upon a close semantic inspection of the source material that, while Gk. ὄρυξ likely means ‘narwhal’ in Strabo, the descriptions of Lat. *orca* in Pliny and Paul the Deacon do not mention its long, single tusk but rather its many sharp teeth and voracious appetite. Thus Lat. *orca* almost certainly refers to a predatory whale like an orca (killer whale) and may be related to or even borrowed from Gk. ὄρκυς ‘tuna’ (itself suspected of being a substrate word, cf. Chantraine 1968-80: 821, EDG 1104). The Atlantic bluefin tuna can exceed three meters in length (National Research Council 1994: 1).

On the other hand, *orca* in the meaning of vessel is difficult to separate from Gk. ὕρχη ‘earthen vessel used for salting fish’. WH (II: 220), EM (467), and EDG (1537) all suggest that *orca* ‘large-bellied vessel’ can have been borrowed from Greek, but this is not regular. Lat. *u* > *o* before *r* followed by a vowel, not by a consonant (cf. Weiss 2020: 153), and the Greek form attests to no vowel that could have disappeared by Latin syncope. It perhaps hints at Etruscan mediation (cf. Breyer 1993: 219-20 and de Simone 1968-70 I: 138, II: 271-5, 287, with the idea that the quality of Etruscan *u* was between Latin *o* and *u*), but again there is no Etruscan form attested. The same sources alternatively suggest that both the Latin and Greek are independently borrowed from a

Mediterranean language (cf. also Ribezzo 1934b: 124, Bertoldi 1939a: 290). Furnée (1972: 137, 361) adduces it as an example of irregular $\chi \sim k$ and $v \sim o$ correspondences.

Lat. *orca* cannot be separated from two other vessel names *urceus* and *urna*, for which the precise relationship with Gk. ὕρχη is unclear (Cuny 1910: 160, WH II: 838-9, 941; EM 754, 755).¹⁷⁷ In light of the other forms, *urna* is plausibly derived from **urk-na*.¹⁷⁸ Given the endings *-eus* and *-na* for these words respectively, they are not borrowed from Greek ὕρχη. Their *u*-vocalism means they could be borrowed from unattested Greek forms, but they could also be loans from the same non-Greek, non-Latin source as ὕρχη and *orca*.

pirum ‘pear’

Pre-form: **(H)pir/s-* | PItal. **pir/so-*

Comp.: **h₂pis-o-*, **h₂pi-uo-* | PGk. **apis/wo-* | Gk. ἄπιον ‘pear’

?Shina *pisō* ‘small pear’, Burushaski *pheso* ‘pear’

?Khinalug *bzi* ‘pear’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, tree; fruit

WH (II: 309-10), EM (510), DV (467)

Tomaschek (1880: 791), Bailey (1924), Kretschmer (1933a), Hubschmid (1960: 59), Berger (1965), Berger (1966), Leumann (1977: 51), Parker (1988), Steinbauer (1989: 69), Kibrik & Kodzasov (1990: 216), Ganieva (2002: 68), EDG (116), Weiss (2020: 153)

The strange correspondence between Lat. *pirum* ‘pear’ and Gk. ἄπιον ‘pear’ is widely considered to be evidence of a Mediterranean non-IE origin (WH II: 309-10, EM 510, DV 467, EDG 116). WH (II: 310) follow Kretschmer (1933a: 89) in assuming that the Gk. ἄ is a prefix from the non-IE donor language, similar to the *Amsel-merula* phenomenon but without the concomitant root vowel gradation we would expect. Steinbauer (1989: 69) suggests a derivation from **h₂pis-o-*, but DV (467) notes that this is an unusual root shape, as PIE roots usually show decreasing sonority to the right and left borders, listing LIV’s **h₂teu(ǵ)-* ‘to spread terror’ as the only exception so far.¹⁷⁹

A major peculiarity of *pirum* regardless of its source is the fact that it is not ***perum*. In this environment, it is widely agreed that we expect *i > e / _rV*, with the Paradebeispiel

¹⁷⁷ Earlier attempts linked *urna* with *ūrere* ‘to burn’ because ceramics are made of baked (fired) clay (e.g. Curtius 1858 I: 315), but this must be folk-etymological.

¹⁷⁸ For the loss of **k* in this position, cf. also *quernus* < **k^werknos* ‘oaken’ as opposed to *quercus* where the velar remains (Weiss 2020: 195). Technically, it could be from PItal. **urχna* as if from **g^h*, which would better match Gk. ὕρχη, but this cannot explain the *c* of *orca* and *urceus*.

¹⁷⁹ LIV2 also gives **h₂ǵ^{er}-* ‘to gather’, attested only in Greek, and **h₃peus-* ‘to increase, abound in’, attested only in Indo-Iranian and perhaps Greek; both roots preceded with a question mark.

being *serō* ‘to sow’ < **sisō* (Leumann 1977: 51, EM 510, Parker 1988, Weiss 2020: 153). As Weiss (2020: 153 fn. 39) notes, if **pisom* entered Latin early enough to be rhotacized, then it was present early enough to undergo the expected change *i* > *e*. This remains unexplained, unless perhaps the other possible exception, namely *vireō* ‘to flourish’ and its relatives, shows that a labial blocks the change (Michael Weiss, p.c.).¹⁸⁰

Berger (1956: 15) suggests that Burushaski *phešo* ‘pear’ is related, which EDG (116) finds improbable. While Berger argues that the Burushaski form is the source in such cases where it exists, it is often much more likely that it attests to loans from Indo-Iranian languages. In this case, the situation is complicated. An Indo-Aryan reflex of this word exists in Shina, a language of the Gilgit valley of Pakistan from which Burushaski seems to have borrowed extensively (Berger 1966: 79). The Shina word is *pisō* ‘small pear’ (Bailey 1924: 158), which shows a startling similarity to the form **pisom* reconstructed for Lat. *pirum*. But when Burushaski borrows from Shina, it seems to faithfully reflect the quality of the sibilant (Berger 1966: 83).¹⁸¹ Thus the relationship of the Burushaski and Shina forms is irregular. Perhaps this shows that Burushaski borrowed the Shina word at an earlier stage, Shina borrowed it from Burushaski,¹⁸² or both are independent loans from a third source. The similarity in form and meaning puts this case outside the realm of coincidence, at least for the Burushaski and Shina forms.

Khinalug, a Nakh-Dagestanian language, has *bzi* ‘pear’ (Kibrik & Kodzasov 1990: 216), sometimes rendered with a schwa (cf. бзы́ in Ganieva 2002: 68). This is otherwise isolated amongst the Caucasian pear words. If this is the same lexeme as that which occurs in Burushaski, and if that is in turn the same as the one that occurs in Greek and Latin, then it looks like the remnant distribution of a once more widespread word with its origins in the East. If it is only Latin and Greek that are related, then we have a non-IE word with what looks like a Mediterranean distribution. See §3.3.2 for a discussion of the distribution of the *a*-prefix.

plumbum ‘lead’

Pre-form: **plō/undʰt-* | PItal. **plumbo-*

Comp.: **ple/oudʰ-* | PCelt. *(*ϕ*)*loudio-* | Mlr. *lúaide* ‘lead’

**moliwdo-* | PGk. **moliwdo-* | *mo-ri-wo-do* /*moliwdos*/ (Myc.), μόλιβος,
μόλυβδος (Homeric) ‘lead’, vars. μόλιβδος, μόλυβος, βόλυβδος,

¹⁸⁰ Alternatively, *serō* may never have had *i*-reduplication to begin with (cf. its reconstruction in the LIV2 as **sé-s(o)h₂-*). No other reduplicated presents show *e*-reduplication, but given that PIE had both *e*- and *i*-reduplication, it is not easy to rule out that *serō* represents an archaism. With the Paradebeispiel gone, perhaps the lowering rule does not exist, and nothing is preventing the shape of *pirum* after all.

¹⁸¹ Cf. Burushaski *sújo* ‘pure, sacred, holy’ < Shina *sujo* < Skt. *sujāta-* ‘well-born’, Burushaski *bašá* ‘turban’ < Shina *pašō* < Skt. *praśna-* ‘wickerwork, basket; turban’, Burushaski *šan* ‘awake, aware’ < Shina *šoñ* etc. < Skt. *śankā* ‘apprehension, care, fear’ (Berger 1966: 81-3).

¹⁸² Berger (and Hubschmid 1960a: 59) thought that Burushaski was the source of the Shina words, but for the wrong reason. They seem to have been influenced by Tomaschek (1880: 791) giving the Shina words as *phěšo* and *phīšo*, but these are the Burushaski forms.

βόλιμος, βόλιβος

**mlīwo-* | PGm. **blīwa-* | ON *blý*, OS *blī*, OHG *blīo* ‘lead’

?PVasc. **bl(e)un(P)-?* | Basque *berun* ‘lead’

?PBerb. **βaldūn* / *βāldūn* / *būldūn* / *βaldūm* | Kabyle *aldun*, Mزاب *buldun*, etc. ‘lead’

??Georg. *brpeni*, *prpeni* ‘lead, tin’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: metallurgy

WH (II: 325-6), EM (516), DV (474)

Meillet (1908), Lafon (1934: 43), Bertoldi (1939b: 94-7), Furnée (1972: 272), Beekes (1999), Boutkan & Kossmann (1999: 92-3), Melchert (2008), EDG (964), Huld (2012: 336), Matasović (2009: 135), Kroonen (2013: 69), Šorgo (2020: 460), Weiss (2020: 507), Thorsø & Wigman et al. (2023: 116-17)

The closest match for Lat. *plumbum* ‘lead’ is MIr. *lúaide* ‘lead’ via PCelt. **ϕludio-* (Matasović 2009: 135). Huld (2012: 336) connected them via a reconstruction **plou-d^h(H)om* ‘solder’ from a root **pleu-* ‘to flow, float’, a logical Bnennungsmotiv for a metal with such a low melting point. But there is no regular way to conjure the nasal in Latin from a form like this. Strictly speaking, the Celtic pre-form does not require a final aspirate, but to produce Latin *b* corresponding to Celtic *d*, there must have been a sequence **Nd^h(u)-* as the nasal would have blocked the RUBL Rule being activated from the left. This produces a reconstruction **ple/oud^h-* behind the Celtic form and **plo/ud^h-* for the Italic (Thorsø & Wigman et al. 2023: 116-17).

To this pair have been adduced a series of Proto-Berber reconstructions, including **βaldūn*, **βāldūn*, **būldūn*, and **βaldūm* ‘lead’, between which the large amount of variation means that this family is not native to Berber (Boutkan & Kossmann 1999: 92-3). Some compare Basque *berun* ‘lead’ (Lafon 1934: 43, Bertoldi 1939b: 94-7, WH II: 326, Boutkan & Kossmann 1999: 92). This could be from a pre-form like **bl(e)un(P)-* (Thorsø & Wigman et al. 2023: 116-17), but could also potentially be borrowed from a Romance source. Both of these groups contain a nasal like Latin, but the Berber forms have it in a much different place. If the Basque form is a borrowing from Romance, then it is not clear how heavily the nasal should feature in a reconstruction of the source form.

Georg. *brpeni*, *prpeni* ‘lead, tin’ has also been compared (Meillet 1908, Bertoldi 1939b: 94-7, WH II: 326, neutrally Weiss 2020: 507). If all these forms are related, it is clear that we are dealing with a Wanderwort. The variation seems extreme at first, but possible Greek and Germanic comparanda might fill the gap.

Despite WH (II: 326), Furnée (1972: 272, etc.), and EM (516), after Beekes (1999) it is

currently in vogue to reject a connection between Lat. *plumbum* and Gk. μόλιβος/μόλυβδος. Beekes cannot accept a connection with the West because the use of lead in Greece is very old. Supporting this is Melchert (2008), who proposes that the Greek is borrowed from Lyd. *mariwda-* ‘*dark’, attested as a theonym (cf. *plumbum nigrum*). DV (474) and Matasović (2009: 135) follow, but Kroonen (2013: 69) who connects PGm. **blīwa-* ‘lead’ through a pre-form **mlīuo-*, does not.

If we suggest that this family represents a non-IE word that achieved a widespread European distribution, then the divergence in the attested forms is not so unexpected. Beekes (1999) suggests that Myc. *mo-ri-wo-do* as a spelling for /moliwdos/ could be behind both of the oldest Homeric forms, with different treatments of the non-IE sequence **-iwd-* surfacing as *-ib-* in μόλιβος and *-udb-* in μόλυβδος. On the other hand, Pre-Proto-Germanic **mlīwo-* matches μόλιβος/βόλιβος well,¹⁸³ but cannot have originally had a dental. Thus the dental element in some of the Greek forms, present also in Celtic (and perhaps Berber, with metathesis) might represent an alternate suffixed form. If this is so, then the *b* of Lat. *plumbum* could be original.¹⁸⁴ All together, the comparanda support a grossly simplified pre-form like **M(V)lVw(n)(-d-)*, perhaps **M(V)lVw(-d-)*, with **M* representing a bilabial.

racēmus ‘bunch, cluster esp. of grapes’

Pre-form: **u/Hrak-* / **(H)rHk-* | PItal. **rak-*

Comp.: **s/ureh₂g-*, **s/uroHg-* | PGk. **rāg-*, **rōg-* | ῥᾶξ, ῥᾶγός; ῥῶξ, ῥῶγός
‘grape’

?**Hreġ^(h)-* | PIr. **raza-* | MoP *raz* ‘vine, grapes, vineyard, garden’, etc.

??**u/Hrus-*, **u/Hro/Hg-?* | PAlb. **rus-*, *raguša-?* | Alb. *rrush* ‘grape’

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: viticulture

WH (II: 414), EM (562), DV (511)

Meyer (1883: 295), Tedesco (1943), Furnée (1972: 126), Çabej (1976 II: 102-3), Katičić (1976: 109), Abaev (II 1979: 398-99), Schrijver (1991: 305-9, 314), EWAia (II: 441-2), Sihler (1995: 207), Demiraj (1997: 144), Orel (1998: 391), Hamp (2000), EDG (1274), Schumacher & Matzinger (2013: 213), Weiss (2020: 382)

Lat. *racēmus* ‘bunch (of grapes)’ is widely suspected of being a word from a Mediterranean substrate (WH II: 414, EM 562, Furnée 1972: 126, Schrijver 1991: 306, DV 511, EDG 1974) based on its viticultural semantics and geographic restriction but

¹⁸³ The sequence **ml-* would have become **bl-* in Proto-Germanic, but the variation within Greek shows that a form with **b* was also in circulation.

¹⁸⁴ Šorgo (2020: 460) interprets the Greek variant βόλιμος as also having a nasal in this position. Alternatively, the nasal in Latin is the result of an *m/b* alternation **-iwd-* / **-ub-*.

also on its irregular correspondences. The Greek forms reconstruct to a voiced velar. And while Lat. *racēmus* could theoretically represent original ***rax, ragis* > ***rax, racis* with leveling of the unvoiced velar from the nominative (cf. a similar case for *fracēs*, s.v.), this is not regular. At face value, the Latin and Greek forms attest to a **k ~ *g* alternation.

Latin *a* against Greek *ā* and *ō* is difficult to account for in an inherited way in this root.¹⁸⁵ Greek could theoretically preserve the full-grade and full *o*-grade of a root with **h₂* starting with **s* or **u*. But in Latin, **sr* yields *fr* and a root shape like **urHk-* would yield ***rāk-*. Schrijver (1991: 305-9, 314) finds evidence that **HRHC* yields Lat. *raC-*, but this initial laryngeal would vocalize in Greek and thus cannot be reconstructed for *racēmus*. Thus we must either reconstruct original *a*-vocalism or a root-initial **r*, neither of which forms a good PIE root.

Alb. *rrush* ‘grape’ is often reconstructed to PALb. **rāgušā-* based on toponymic evidence: the Dalmatian city *Ragusium* is given in an Albanian source as *Rushë* (Çabej 1976 II: 102-3, Orel 1998: 391, Hamp 2000: 9). Hamp suggests a pre-form **rāg-ūs-V-*. Between this and his reconstruction of **ῥρωγούμ* for the Salentine Greek *to rukūmi ~ ragūmi*, the *-ēmus* of Lat. *racēmus* might reconstruct to **-esmo-* (in an inherited example, cf. its development in the superlative ending like in **eksterisomo* > **ekstresmo-* > *extrēmus*, Weiss 2020: 382). Katičić (1976: 109) compares Hsch. *ράματα·βοστρύχια, σταφυλῖς. Μακεδόνες* ‘bunch of grapes, Macedonian’, which EDG (1274) proposes is from **ράγμ-* and clearly related to *ῥᾶξ, ῥᾶγός*. The sequence **rag-s-mo-* should produce PGk. **rak^hmo-* (cf. *λελέχθαι* < **lelek-st^hai* < *λέγω*). But Sihler (1995: 207) mentions forms like *δράγμα* ‘handful’ for *δράχμα*, where *γ* occurs for expected *χ* due perhaps to dialect mixture. If *ράματα* can really be from **ράγματα*, then it could potentially attest to a base **rag-s-mo-* against **rag-es-mo-* behind Lat. *racēmus*. The forms with the labial suffix have the collective meaning ‘bunch of grapes’ beside the basal meaning ‘grape’ of the forms without the suffix (PGk. **rālōg-s-* and PALb. **rāg-us-*).

But while PIE **b^(h)* seems to disappear intervocally in Albanian, and intervocalic *d* seems to disappear in loans post-dating the change **-Vd^(h)V-* > **-VðV-* (Demiraj 1997: 62), there is little indication that such was true for **g^(h)*. A more straight-forward reconstruction for Alb. *rrush* is **rus-*,¹⁸⁶ which is no longer easy to compare to the Latin and Greek forms.

¹⁸⁵ Hamp (2000: 7) reconstructs in essence a paradigm nom. **urōHg-s-*, acc. **urōHg-m-*, obl. **urHg-*. By ignoring the Latin form, he can reconstruct an initial **u* based on Salentine Greek *vrāva, vrā, grā* < **ῥράγα* perhaps < **ῥράγα*. But his source (Gerhard Rohlfs, 1962, *Neue Beiträge zur Kenntnis der unteritalischen Gräzität [non vidi]*) also compares Salentine Greek *to rukūmi ~ ragūmi*. Hamp uses this to reconstruct **ῥρωγούμ*, which seems to make a connection with Lat. *racēmus* even more inevitable.

¹⁸⁶ Schumacher & Matzinger (2013: 213) consider *rrush* a borrowing from Gk. *ῥῶξ* (PALb. **rušša-*) at a time after Gk. *ō* and *o* had fallen together. The only other example they give however is Alb. *i kuq* ‘red’ < Lat. **cocceus*.

Meyer (1883: 295) suggested a connection between Alb. *rrush* and Persian *raz* 'vine, grapes, vineyard, garden'.¹⁸⁷ Abaev (II 1979: 398-99) takes *raz* from Plr. **raza-* along with Oss. *ræzæ* 'fruit, fruits, vegetables', Tajik *raz* 'vine, vineyard', Kurdish *rāz*, *rez* 'garden', Zazaki *rāz* 'vineyard', etc. Tedesco (1943) tried to connect MoP *raz* with Slavic **lozā* (OCS *loza* 'vine, Ru. *lozā* 'vine, rod', etc.) through a pre-form **loǵā-*, but this would yield Plr. ***rāza-* with Brugmann's Law. Only **laǵā-* with a PIE **a* could yield the correspondence. Otherwise Plr. **raza-* presupposes a reconstruction Plr. **raj^(h)-* as if from *(*H*)*reǵ^(h)-* or **leǵ^(h)-*. If it is non-IE, **raǵ^(h)-* is a possibility, and it fits into the **raǵ^(h)-* ~ **rāǵ^(h)-* ~ **rōǵ^(h)-* alternation established for Latin, Greek (and perhaps Albanian) both formally and semantically. Given alternative reconstructions, its appurtenance is not completely certain, but it does not contraindicate a non-IE word for grape of the shape **rVG*.

Connections between Gk. *ῥᾶξ* and Lat. *frāga* 'strawberry' via **sṛāg-* (cf. DV 239) are not as attractive as deriving *frāga* along with Alb. *dredhë* 'strawberry' from a pre-form **ḍ^hrHg^(h)-* (s.v. *frāgā*).

rāpum 'turnip'

Pre-form: **Hl_ureh₂p-* | Pltal. **rāpo-*

Comp.: **sl_uǵ^(h)alHp-*, **sl_uǵ^(h)alHb^h-* | PGk. **rap-*, **rap^h-* | Gk. *ῥάφους*, *ῥάπυς* 'turnip', *ῥάφανος* 'cabbage, radish'

**Hreh_{2/3}b^h-* / **Hreh_{2/3}p^h-* | PGM. **rōbjōn-* | MDu. *rove*, OHG *ruoba*, *ruoppa* 'turnip', etc.

**Hreh₁p-* / **Hroip-* | PSlav. **rēp-* / **roip-* | RuCS *rěpa*, Ru. *rěpa*, SCr. *rěpa* 'turnip', etc.

**Hreh₂p-* | PBalt. **rāp-* | Lith. *rópė* 'turnip'

**h₁erb^(h)-* | PCelt. **arbīno-* | OBret. *erbin*, *W erfin* (pl.) 'turnip'

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: plant, domestic

Pokorny (852), WH (II: 418), EM (564), DV (514)

Schrijver (1991: 310), Demiraj (1997: 349-50), EDG (1277, 1283), Zohary, Hopf & Weiss (2012: 159), Kroonen (2013: 415)

From Pokorny (852) onwards (WH II: 418, Schrijver 1991: 310, EM 564, DV 514, EDG 1277, Kroonen 2013: 415), Lat. *rāpum* 'turnip' and its comparanda¹⁸⁸ have been viewed

¹⁸⁷ He also suggests a link with Skt. *rasā* 'raisin', but EWAia (II: 441-2) translates this much differently, as 'plant juice, juice, liquid, viscous fluid, essence, pulp'.

¹⁸⁸ Alb. *rrépë* 'beet, radish' < PALb. **rap-* cannot be a loan from Latin. But its status as an independent comparandum is difficult to verify; it could potentially be a Greek loan (Demiraj 1997: 349-50).

as likely Wanderwörter because of the irregularly corresponding vocalism (even if PIE **a/ā* existed, the comparanda require the reconstruction of non-IE **ā/ē* ablaut¹⁸⁹) and lack of a prothetic vowel in Greek (initial **s* or **y* could be reconstructed for Greek to avoid an invalid *r*-initial root structure, but neither option works for Germanic).

EDG (1277) considers the $\pi \sim \phi$ alternation within Greek to be a Pre-Greek feature, but attestations of this word are far too widespread to have their origins in Beekesian Pre-Greek. (Note also that the Germanic forms attest to what in native words would be reconstructed as **b^h* just like Greek ϕ , suggesting that the Pre-Greek-like variation in the donor forms was not limited to the East.) The Celtic forms attest to an *a*-prefix with zero-grade root (Kroonen 2013: 415). Metathesis from **rabīno-* would be unconditioned.

Comparanda possibly extend beyond Europe. Furnée (1972: 313 fn 35) compares the Semitic family **lapt-* ‘turnip’, finding it even in the Hsch. *λάψα: γογγυλίζ. Περγαῖοι*. Cross-linguistically, *l ~ r* alternation is not rare, and other substrate examples include *līlium ~ λείριον*. Further similar is Sumerian **lub* ‘turnip.’ The Semitic and Sumerian words cannot be adduced with nearly as much certainty, but would suggest that this family of words was widely distributed amongst the agricultural populations of Europe and Western Asia.

raudus ‘lump of copper used as currency’, vars. *rōdus, rūdus*

Pre-form: **H/ureh₂ud^(h)-* | PItal. **raudo-*

Comp.: **h₂erud-* | PGM. **arut-* ‘ore’ | ODu. *arut*, OHG *aruz, ariz* ‘ore’, etc.

?**Hrut-* | PCelt. **rutu-* | W *rhwd* ‘rust’

OSum. *aruda* ‘copper’ > Sum. *uruda, urudu* ‘copper’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: metallurgy

Pokorny (872-3), WH (II: 420-1), EM (565), DV (515)

Schrader (1883: 62), Schrijver (1991: 265), Schrijver (1997: 308), Stifter (1998: 214), Hill (2003: 196-202), Jagersma (2010: 60-1), Kroonen (2013: 37), Schrijver (2018: 361-3), Koch (2020: 110), Thorsø & Wigman et al. (2023: 109)

Lat. *raudus* ‘lump of copper used as currency’ is traditionally compared to PIIr. **Hraud^ha-* (cf. Skt. *lohá-* ‘reddish metal, copper-colored, reddish, made of iron’, MP/MoP *rōy* ‘copper, brass’, etc.), ON *rauði* ‘bog iron ore’, and OCS *ruda* ‘ore metal’ from PIE **h₁reyd^h-* ‘red’. The problem is well known and it is widely admitted that the combination **-ud^h-* should yield Lat. *-ub-* as it does in *ruber* ‘red’ < **h₁rud^h-ro-*. The solution has been to assume that the Latin word is borrowed from another IE language

¹⁸⁹ This alternation is especially remarked on by WH (II: 418) who compare it to that between Lat. *nāpus* and Arm. *niv* (s.v. *nāpus*).

(WH II: 420-1, DV 515, Schrijver 1991: 265).

WH (II: 420-1) reject a comparison to PGm. **arut-* ‘ore’ because of the initial *a*. It is clear that **arut-* cannot be reconstructed back to **h₁reyd^h-*, and if Lat. *raudus* is indeed adduced, it produces a perfect example of the substrate *a*-prefix phenomenon, creating the alternation **arud-* ~ **raud-* (Schrijver 1997: 308, Kroonen 2013: 37). Furthermore, explaining *raudus* from the perspective of a known phenomenon seems preferable to the *ad hoc* solution of a borrowing from another IE language. Thorsø and Wigman et al. (2023: 109) consider PIIr. **Hraud^ha-*, ON *raudī* ‘bog iron ore’, and OCS *ruda* ‘ore, metal’ < IE **h₁royd^h-o-* (to the root **h₁reyd^h-*) as an unrelated group coincidentally similar to Lat. *raudus* and PGm. **arut-* < non-IE **arud-* ~ **raud-* ‘ore’ (even though EM 565 suspects that the ‘red’ derivatives might actually have been remodeled based on folk etymology). PCelt. **rutu-* (cf. W *rhwd* ‘rust’) might also belong to this group. Despite the reddish color of rust, **rutu-* cannot derive simply from **h₁reyd^h-* ‘red’ (Koch 2020: 110). But as alternative etymologies exist (PCelt. **ruddo-* < **h₁reyd^h-* ‘red’ + **d^heh₁-* ‘to put’ [Stifter 1998: 214] or **sed-* ‘to sit’ by [Hill 2003: 196-202]), its connection is much less certain.

Establishing the existence of a non-IE word **arud-* ~ **raud-* ‘ore’ allows it to be linked to Sumerian *uruda*, *urudu* ‘copper’ from Old Sumerian *aruda* (Schrader 1883: 62, 118; WH II: 421; Schrijver 2018: 363; Thorsø & Wigman et al. 2023: 109; Jagersma 2010: 60-1 on the Old Sumerian form). Schrijver (2018: 361-3) takes this as evidence of a Hatto-Sumerian agricultural substrate, but the word need not be native to Sumerian.

rosa ‘rose’

Pre-form: **uroS-* | PItal. **rosā-*

Comp.: **ur(o)d-* | PGk. **wrod-* | Gk. ῥόδον, Aeol. βρόδον ‘rose’

**urd^ho-* | PIr. **urda-* | MoP *gul* ‘rose’, etc.
> Arm. *vard* ‘rose’

Arab. *ward* ‘rose, flower, blossom’, etc.

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: plant, flower

WH (II: 443), EM (577)

Buck (1904:48, 83), Kretschmer (1923a: 115), Schwyzer (1939-50 I: 344 fn. 2), Alessio (1946b: 26), Mayrhofer (1950: 74), Mayrhofer (1961: 185), de Simone (1968-70 II: 165-6), Sihler (1995: 172), Untermann (2000: 464), EDG (1289), Sims-Williams (2016: 206), Weiss (2020: 162 fn. 12), van Beek (2022: 319-21)

The intervocalic *s* of Lat. *rosa* is problematic. If it is the result of a post-rhotacism borrowing, it must have been borrowed after the middle of the fourth century BCE. Weiss (2020: 162 fn. 12) notes that this would be unexpected if it is a loan from a

substrate language and therefore proposes the interference of the initial *r* (though he notes the counterexample *rōs, rōris* ‘dew’), an effect which otherwise occurs only when an *r* occurs in the *next* syllable (Sihler 1995: 172 e.g. *caesariēs* ‘bushy-haired’, *miser* ‘wretched’, *aser* ‘blood’, etc. but *aurora* < **ausōsā*-).

Evidence points to *rosa* being a Wanderwort with comparanda stretching far to the East. It is undoubtedly related to Gk. ῥόδον ‘rose’, whose Aeolic variant βῥόδον (and appearance as Myc. *wo-do-we* ‘rose-scented’) shows it originally began with **w*. Thus it, like Arm. *vard* ‘rose’, is quite likely a loan from an Iranian source (WH II: 443, Mayrhofer 1950: 74, EM 577, EDG 1289 Sims-Williams 2016: 206). A form like Plr. **ur̥da*-¹⁹⁰ would yield e.g. Sogd. *wrd* and MoP *gul* ‘rose’. If Gk. ῥόδον represents an artificial epic reflex of **ur̥do-* (explaining the otherwise irregular reflex of the syllabic resonant, van Beek 2022: 319-21), then the Greek proto-form is very similar to that of the Iranian forms.

It is clear that there are Semitic comparanda (Mayrhofer 1961: 185, EM 577, EDG 1289), but Mayrhofer (1950: 74-7) makes a case that there are so many Semitic forms that it does not look to be a loanword in Semitic: Arab. *ward* ‘rose, flower, blossom’, Aram. *wardā* ‘rose, rose-colored; lobe of the lung’, Akk. *murdēnu, murdennu, amurdenu, amurdennu* (for *wurdēnu*) ‘a flower with thorns’. Arabic has further *warada, warrada* ‘to bloom’, *waruda* ‘to be red’, *warrada* ‘to color red’, *tawarrada* ‘to blush or flush’, *word* ‘malaria’, and *warīd* ‘jugular vein’. Given that the floral meaning is old, well-integrated, and not semantically streamlined in the Semitic languages that attest it, it does not seem obvious as a loanword. An ultimate Semitic origin seems more likely than the link with a PIE root only attested in Iranian. In any case, the rose word in Latin is a Wanderwort from the East.

If the word entered Latin through Greek (WH II: 443, EDG 1289), there must have been an intermediary, as there is no foolproof way from Gk. δ > Lat. *s*. EM (577), followed by Alessio (1946b: 26), suggest Etruscan. Etruscan seems to have borrowed Gk. -δ- as *z* (/ts/) (cf. in two names: *Arxaze* and vars. < Ἀρκαδία and *Zimaite* and vars. < Διομήδης, de Simone 1968-70 II: 165-6). So a form like **ροδία* could theoretically have been the source. But this Greek form is unattested (and in Modern Greek *ροδιά* is the pomegranate tree), and the lookalikes Etr. *ruze, rusi* are of unknown meaning. Kretschmer (1923a: 115) suggests that Lat. *rosa* might be from Gk. ῥοδέα ‘rosebush’ through a form **rodia* that passed through “sabinisch” (given the name of Sabine statesman Appius Claudius, said by Livy to have been called Att(i)us Clausus before he moved to Rome). Alessio (1946b: 26) mentions the possibility too, with the understanding that Latin *medius* corresponds to Oscan **meso-*. But this is incorrect; the Sabellic cognate of Lat. *medius* is Osc. and SPic. *mefi-* (Untermann 2000: 464). Buck

¹⁹⁰ WH (II: 443) further reconstruct for the Iranian form an IE **ur̥dʰo-*, but this is otherwise unattested. They connect what they give as OE *word* ‘thornbush’. But the form is actually *word* ‘enclosure (created with thorny shrubs)’ < PGm. **wurþa-* with PIE **t*.

(1904: 66) notes that a change $*d\dot{i} > z$ and $*t\dot{i} > s$ is restricted to Bantia. The closest to a workable solution is the Umbrian change of intervocalic $*d > \check{r}$, *rs*. Buck (1904:48, 83) notes that the *r* of both inherited and *d*-derived *rs* was weakly pronounced and is sometimes not written in the Umbrian inscriptions in the Latin alphabet.¹⁹¹ But there are problems with this solution too. Beyond there being no other cases of $*d > \check{r}$ in a Greek loan to compare (and no attestation of the *rosa* word in Sabellic to confirm), an *r* elsewhere in the word seems to block the change $*VdV > V\check{r}V$ (Buck 1904: 82, Untermann 2000: 816 with lit.). Cf. U *Coredier* ‘Coredii’ (with the same *rVdiV* sequence as Kretschmer’s proposed pre-form $*rodia$) and U *utur* ‘water’ ($< *ud\ddot{o}r$). The change in gender (Greek neuter to Latin feminine) is also without a good explanation.¹⁹² Thus, if Lat. *rosa* was indeed mediated from Gk. (β)ρόδον, the mediating language is still unknown.

sabulum ‘sand’

Pre-form: $*sa/Hb^{(h)}/d^h-lo-$ | PItal. $*sab/f/plo-$

$*(p)sa/h_2m/b^h-mo-$ | PGk. $*(p)sam/p^h-mo-$ | Gk. ψάμμος, ἄμμος ‘sand’

$*(p)sa/h_2m-\check{n}d^h-o-$ | PGk. $*(p)samat^h-o-$ | Gk. ψάματος, ἄματος ‘sand’

$*sa/o/HM-(a)d^h-$ | PGM. $*sammada-$ ‘sand’ | ON *sandr*, OE *sand*, MHG *sampt* ‘sand’, etc.

$*sap/b^h-a\acute{g}^h(/d^h?) -o-$ | Arm. *awaz* ‘sand, dust’

?Abkhaz *saba* ‘dust’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: geography

Pokorny (146), WH (II: 458), EM (585), DV (531)

MacBain (1911: 321, 323), Güntert (1914: 119-20), Boisacq (1916: 48, 1074), Schwyzer (1939-50 I: 328-9), Alessio (1944a: 144-6), Deroy (1956a: 183-4), Kuiper (1956: 218), Frisk (1960-72 I: 84, II: 1129-30), Ačařyan (1971-79 I: 351), Furnée (1972: 209), Schrijver (1991: 103), Kuiper (1995: 67), EWAia (II: 198), Garnier (2006), Martirosyan (2009: 149, 247), EDG (1660), Kroonen (2013: 425), Kroonen (ftthc.), Thorsø (ftthc.)

A connection between Lat. *sabulum* ‘sand’ and a slew of Greek forms of similar semantics is widely accepted (WH II: 485, EM 585, DV 531, Kroonen 2013: 425, EDG 1660, etc.), but the details of their relationship are complicated. Boisacq (1916: 48, 1074), followed in part by Frisk (1960-72 I: 84, II: 1130), concluded that ψάμμος and ἄματος were two originally unrelated forms, with ἄματος and ψάματος originating as crosses. He takes ψάμμος $< *psap^h-mo-$ as related to Lat. *sabulum* while ἄματος would be related to MHG *sampt* and several Sanskrit forms. (Güntert 1914: 119-20 likewise

¹⁹¹ It seems to have introduced compensatory lengthening in the preceding vowel when the *r* was lost.

¹⁹² Unless Gk. ροδέα ‘rosebush’ served as the ultimate source.

separated the Greek words due to the inexplicability of double Anlaut reflex.) Schwyzer (1939-50 I: 328-9) supports a connection between Greek and Skt. *psāti* ‘consumes’ to the root *bhas-* ‘to crunch, chew’, but the development of Gk. *ψαθος* would require the change of the cluster **b^hs- > *s* without becoming **ps-*, which seems strange.¹⁹³ WH (II: 458), who do not separate the Greek forms, explain the change of **b^hs- > *s* as *vorgriechisch* (cf. also Boisacq 1916: 1074, who calls it *préhellenique*), presumably in the Pelasgian sense.

EWAia (II: 198) questions the relationship of the Sanskrit forms (< PIE **b^hes-*, **b^hs-eH-*) to the Greek forms on semantic grounds. With the tenuous link to the Sanskrit forms gone, we can consider alternative explanations for the Greek forms. Deroy (1956a: 183, and 183-4 fn. 3) explained the Greek variation in Anlaut as due to a borrowing from a non-IE language beginning with a sibilant that was variously interpreted as **s* (and thus later lost) or as a stronger sibilant that was reflected as **ps*.¹⁹⁴ This allows for the comparative analyses that followed (cf. Furnée 1972: 209, Schrijver 1991: 103, DV 531; EDG 78, 89, 1660).

Several Germanic forms suggest a reconstruction of PGm. **samda-* < **samd^h-o-* (Kuiper 1995: 67, Kroonen 2013: 425). MHG *sambt*, *sampt* as well as Bavarian and Yiddish forms have resisted the change **md > *nd*, leading Kroonen (fthc.) to reconstruct **samm(a)d^h-o-*. This justifies a comparison to Gk. *ψαθος* and favors for it the reconstruction of **sam-ad^h-* (potentially an unnasalized variant of the *vθ*-suffix, Kuiper 1956: 218) over **sam-ṅd^h-*.

The *b* in Lat. *sabulum* can go back to **b*, **b^h*, or **d^h*. Without the nasal element however, a reconstruction **sad^h-(u)lo-* looks quite aberrant. Instead, reconstructing **sab^(h)-* and establishing a **b^(h) ~ *m* alternation with the Greek and Germanic forms finds probable support in Armenian.¹⁹⁵ The labial of Arm. *awaz* ‘sand, dust’ reconstructs to **p* or **b^h*, allowing the reconstruction **sab^had^ho-* (Ačařyan 1971-79 I: 351) or **sab^had^hs* (Thorsø fthc. fn.), remarkably similar to the Greek and Germanic forms.¹⁹⁶ Thus we have

¹⁹³ Garnier (2006) starts from a formation **b^hos-mó-* ‘the action of rubbing’, postulating a collective **b^hs-m-eh₂* ‘powder, grating, sweepings’ that was complemented with **d^heh₁-* to produce **b^hs-ṁ-h₂-d^hh₁-* ultimately behind Gk. *ψάματος*. His explanation requires several assumptions and complexities. To explain PGm. **samda-* he must propose descent from the same zero-grade pre-form with analogical full-grade **samda-* arising from analogy to **mulma-* ~ **malma-* ‘friable’. Lat. *sabulum* would be from a univerbation with a different light verb **b^hs-éh₂* **b^huH-*. He does not have an explanation for why **b^hs-* yields *ψ* in *ψάματος* but *ᾤ* in *ᾤμμος*.

¹⁹⁴ Alternatively, a borrowing into Greek both before and after the loss of initial *s* in Greek (cf. Kroonen 2013: 425). But Guus Kroonen (p.c.) notes that the abundance of unetymologized Greek words with *s₂* makes this unlikely.

¹⁹⁵ Gk. *ψήφος* ‘pebble’ and *ψαφαρός* ‘loose, rotten, crumbled’ attest to **b^h*, but are semantically more remote. If the *μ* of *ψάμμος/ᾤμμος* is from **φμ* like in Gk. *γράμμα* ‘letter, writing’ < *γράφ-μα*, it too could attest to **b^h*.

¹⁹⁶ Martirosyan (2009: 149) prefers a loan from Iranian (cf. MoP *āwāze* ‘swamp’) requiring the semantic shift ‘swamp’ > ‘silt’ > ‘sand’, which seems dubious in light of the sandy semantics of the other comparanda. Old Armenian *awazan* ‘pool, bath, basin’ could have been borrowed from Iranian, but then it is then a separate lexeme.

evidence of a non-native alternation **samadh-* ~ **sab^(h)adh-*, whose foreign **s* left a double reflex in Greek and whose dental ending is not reflected in Latin.

A further indication of the non-native origin of Latin *sabulum* is the word *saburra* ‘ballast sand, grit’ whose suffix is distinctly non- or pre-Latin, perhaps Etruscoid (WH II: 458, followed by Deroy 1956a: 184, Furnée 1972: 209) but whose root seems to be the same as *sabulum* (cf. additionally Schrijver 1991: 103, DV 531, *pace* EM 585).¹⁹⁷ Abkhaz *saba* ‘dust’ might be related, but this is difficult to confirm.

simila ‘fine flour’

Pre-form: **semil-* | PItal. **semil-*

Comp.: **Semidāl-* | PGk. **Semidāl-* | Gk. *σεμιδάλις* ‘fine flour’

Aram. *ṣamīdā*, Akk. *samīdu* ‘a kind of groats’ < Akk. *samādu* ‘to grind into groats, to be ground into groats’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: culinary

WH (II: 538), EM (626)

Lewy (1930: 28-9), Güntert (1932: 21 fn. 1), CAD (S 1984: 107, 115-6), EDG (1320)

The Semitic semolina words convincingly have their origin in the Akk. verb *samādu* ‘to grind/be ground into groats’ (Lewy 1930: 28, cf. CAD S: 107, 115-6). Gk. *σεμιδάλις* ‘fine flour’ has been borrowed directly from a Semitic language (EDG 1320), presumably after the loss of **s* (thus the reconstruction into Proto-Greek or PIE is for the sake of consistency). It attests an additional *l*-suffix.¹⁹⁸ Lat. *simila* ‘fine flour’ is no such direct borrowing, as it has *l* for *d*.¹⁹⁹ The derived form *similāgo* is already found in Cato the Elder (contrary to EM 626 asserting it was borrowed during the Empire). Lewy (1930: 28-9) ascribes the change to the phenomenon often called the “Sabine *l*”, with examples in internal position including *oleō* ‘I smell’ vs. *odor* ‘smell’ and *solium* ‘seat’ vs. *sedeō* ‘I sit’. Solid examples of this poorly understood phenomenon are few, and all of them are in inherited material. It has been proposed that there is a separate *d* ~ *l* alternation in non-IE words (s.v. *laurus*), which Güntert (1932: 21 fn. 1) thought had

¹⁹⁷ In the same vein, Alessio (1944a: 144-6) proposed that Etr. *zamaθi* ‘gold’, *zamθic* ‘golden’ is related, with the understanding that its semantics would have changed from ‘sand’ > ‘gold’ in the context of mining placer deposits, either as a loan from Gk. *ψάμαθος* or from the same substrate source. This should be kept in mind for considerations of Etruscan’s role in the Italic substrate.

¹⁹⁸ Its origin and purpose is unclear. EDG (315) notes the similarity of *δενδαλις* ‘barley-cake’ but - *αλις* and -*αλον* otherwise frequently occur in animal names (*όρταλις* ‘hen’, *σικαλις* ‘fig-pecker’, *πάρδαλις* ‘panther, leopard’, *δάμαλις* ‘young cow’ (certainly inherited), *κνώδαλον* ‘wild or harmful animal’, *ἔταλον* ‘yearling’ (certainly inherited), *ἰξαλος* ‘castrated’ he-goat’.

¹⁹⁹ This is really an alternation and not somehow a borrowing from Greek. Syncope of a form like **semidala* > **semidla* would be unusual. But even if the word arrived from Greek via some intermediary, **semidla* would yield **semilla*. From there, there is no regular way to simplify the geminate.

something to do with Asia Minor. In the case of Lat. *simila*, an *l* replaces a *d* in a word of Semitic origin. Since the change has not affected the Greek borrowing from the same source, it is likely that word was mediated to Latin via another language.

sirpe ‘silphium or the juice thereof’

Pre-form: **sirp-* | PItal. **sirp-*

Comp.: **Silp*^h-, **Selp-* | PGk. **Silp*^h-, **Selp-* | Gk. σίλφιον, Hsch. σέλπρον· σίλφιον
‘silphium’

?Berber *azlaf*, *azelaf*, *aselbu*, etc. ‘the sea rush *Juncus maritimus*’²⁰⁰

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, wild

WH (II: 547), EM (629)

Schuchardt (1918: 16), Nehring (1927: 274), Bertoldi (1937a: 144), Alessio (1944a: 124), Ernout (1946: 49), de Simone (1968-70 I: 140), Furnée (1972: 163), CAD (Š Part 1: 247), Breyer (1993: 225), Parejko (2003), EDG (1332)

Silphium was an economically important crop grown in Cyrenaica, famous for being impossible to cultivate, that was highly prized for its flavorful sap (*laser* or *lasserpīcium*²⁰¹). Popularly believed to have been exploited to extinction, descriptions in texts and on coins suggest that it was a species of giant fennel (cf. Parejko 2003).

Given that the origins of the plant are in North Africa, it is not surprising that it is without a doubt a word of non-IE origin. Crucially, Lat. *sirpe* against Gk. σίλφιον, σέλπρον points to the word entering Greek with *s* after the loss of inherited **s* and attests to an *l* ~ *r* alternation (as well as an alternation in aspiration) that we find in other words of non-IE origin. The source of these words remains unidentified. EM (629) highly suspect Etruscan, and Ernout (1946: 49) elaborated that the Latin is a borrowing of the Greek via Etruscan mediation due to the nominative in *-e*. De Simone (I: 140) finds no evidence of this and Breyer (1993: 225) notes that none of the changes that are purported to have occurred have any parallels in other examples of Etruscan-mediated Latin borrowings from Greek. WH (II: 547) correctly reject Schuchardt’s (1918: 16) suggestion that the Latin and Greek forms were borrowed from Berber, but Furnée (1972: 163) and EDG (1332) still consider it possible that the Berber forms are an

²⁰⁰ All forms are found in Central Morocco (Ba₆, Ba₁₄, and Ba₁₅ respectively, as per Schuchardt’s 1918: 16 notation). *Azlaf* is also found in Tunisia (Hu₂), and *aselbu* in North Algeria (De).

²⁰¹ WH (II: 547) and EM (342, 629) both take *lasserpīcium* from a collocation of *lac* + *serpicium*, but this smacks of a folk etymology; especially because of Gk. λάσαρον of the same meaning. EDG (835) says it is of unknown etymology. Perhaps it is a borrowing of Lat. *laser*, shortened from *lasserpīcium*, but the vowels do not match. The CAD does not list Assyrian *lasirbitu*, which e.g. Nehring (1927: 274) claims is the source of the Latin but of which WH (II: 547) is doubtful that the reading is correct. CAD (Š Part 1: 247) does however list *šallapānu* ‘a plant’, which Lévy (1900: 339) used to suggest a Semitic origin for the *sirpe* family. Its meaning is too poorly known to be able to adduce it with any certainty.

independent borrowing from the same source (cf. also Bertoldi 1937a: 144, Alessio 1944a: 124). It should be noted that the Berber forms denote a different plant. If related, because silphium was a North African plant with a comparandum in Berber (a North African language), the language(s) responsible for the *l ~ r* alternations we find might have something to do with North Africa. Bertoldi (1937a: 144) purported to notice a similar alternation between Basque *zaldi* ‘horse’ and Berber *a-serdun* ‘mule’, where the Berber form is preceded by an *a* and there is an *l ~ r* alternation. However, the inclusion of Basque does not help to more precisely locate the source of alternation; it at least still limits it to the Mediterranean.

sōrex ‘shrew’

Pre-form: **s(u)ōr-Vk-* | PItal. **sōrV-k-*

Comp.: **syo/ur-ak-* | PGk. **surak-* | Gk. ὄραξ ‘shrew’

**sur-(V)g-* | PGm. **s(w)ur(V)ka-* | OSw. *surk* ‘mole, vole, shrew’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: animal, wild

Pokorny (1049-50), WH (II: 563), EM (647), DV (576)

Kock (1909: 84), Hellquist (1922: 827), Chantraine (1933: 376-83), Ernout (1946), Vine (1999a: 572-3), EDG (1563), Beekes (2014: 32), Kölligan (2017: 369-70), Wigman (fthc.)

The traditional explanation is to connect Lat. *sōrex* and Gk. ὄραξ ‘shrew’ to a PIE root **syer-* ‘to resound’, cf. also *susurrus* ‘whisper’ and *surdus* ‘deaf, silent’ (WH II: 563, Pokorny 1049-50). As to the irregular vocalic correspondence between Latin and Greek, Vine (1999a: 572-3) treats the Greek form as an example of Cowgill’s Law in the environment of *(-)TyoR- > *(-)TyuR-. Thus he begins with an original root noun to the root **syer-* with **ō/o* ablaut rather than unmotivated **ō/φ* ablaut (cf. *ardea*, s.v.). The semantic argument smacks of a folk etymology, but Latin literature contains references to ‘singing’ shrews (in Pliny’s *Nat.Hist.* 8.82: 223, they interrupt the auspices). Shrews in reality are quite vocal, with evidence that they use their voice for echolocation.

Some are unconvinced (EM 647) and prefer a substrate origin (DV 576, EDG 1536). The Greek -αξ indeed occurs frequently on words of obscure etymology, many of which are likely not native to Greek (EDG 1536 and Beekes 2014: 32 consider it a Pre-Greek suffix). But it also appears on inherited bases (cf. Chantraine 1933: 376-83²⁰²). Thus, not every word with an -ak suffix must be Pre-Greek. (Cf. the proposed pathway in Kölligan [2017: 369-70] whereby -αξ can be inherited, when secondary to -ᾶξ < **eh₂-k-s.*) A similar situation occurs for the Latin suffix -ex (Wigman fthc. with lit.).

²⁰² Cf. κόραξ ‘raven’ and δέλαραξ ‘sow’ where it was added to an IE root perhaps due to its frequency in animal names.

The Germanic comparanda (OSw. masc. *surker*, neut. *surk*, Sw. *sork* ‘mole, vole, shrew, ODan. *syrycha mych* ‘rat excrement?’) suggest that this family is not inherited. Assuming a loan from *sōrex* has phonological problems (Kock 1909: 84, Hellquist 1922: 827). As independent comparanda, they would have to stem from the zero-grade of **suer-*, requiring the suspicious **ō ~ ø* ablaut mentioned above. Furthermore however, in stemming from a PGM. **s(w)ur(V)ka-* (p.c. Guus Kroonen) they attest to the same **k ~ *g* alternation of the velar suffix as seen in *filix* and *fulica* (s.v.). The *u* vocalism of the Greek form is thus likely original and in irregular alternation with *ō* of the Latin.

taeda ‘pine; pine branch; torch’

Pre-form: **th₂eid-* | PItal. **taidā-*

Comp.: **deh₂u-* | PGk. **daiwid-* | Gk. δαίς, -ίδος ‘torch’

?Berber *tayda* ‘Aleppo pine’

■ Irreg. correspondences □ Remarkable phonotactics

Semantics: plant, tree / tool

WH (II: 642), EM (673)

Wood (1910: 307), Charpentier (1917: 46), REW (no. 8520), Pfiffig (1969: 37), Biville (I: 221), Breyer 1993: 229-30, Rix (2008: 145-6), EDG (298), El Arifi (2014: 468)

WH (II : 642) and EM (673) take Lat. *taeda* as a borrowing from the accusative of Gk. δαίς ‘torch’, having gone through Etruscan to account for the initial devoicing. The Greek word potentially has a good etymology, derived from the verb δαίω ‘to kindle’ < **deh₂u-* ‘to burn’ (EDG 298).

Given that the primary meaning of *taeda* seems to have been a resinous species of pine (cf. EM 673), some have preferred a derivation from **teih₁-* ‘to become warm’ (**tāi-* in Wood 1910: 307, Charpentier 1917: 46), especially in comparison to OE *hīnan* ‘to become moist’, having undergone a semantic development ‘to become warm’ > ‘to melt/thaw’ > ‘to become moist’. This would be a parallel for the running, flammable pitch from the tree. However, EDG (298) lists several Greek forms that also refer to pine and its resin: cf. δάδινος ‘pertaining to the torch, made of pine-wood’ and δαδώδης ‘resinous’; thus the Latin and Greek words could still be related, albeit irregularly.²⁰³

De Simone (1970 II: 102 fn. 49) rules out Etruscan intermediation, presumably for the same reason that Biville (I: 221) questions it: the voiced word-internal *d* of the Latin form. Etruscan is often touted to have had no voiced consonants. However, there are cases where Etruscan consonants in names were perceived by Latin speakers as voiced

²⁰³ Lat. *daeda* is attested in a late gloss. Biville (I: 221) interprets it simply as a late transcription of the Greek word. But this seems difficult to reconcile with the fact that it is this form that made it into some Romance forms like Rom. *zadă* and Sicilian *deda* (cf. REW no. 8520). The form behind the Romance languages might be a re-borrowing directly from Greek, or attest to a *t ~ d* alternation within Latin.

(cf. Pfiffig 1969: 37,²⁰⁴ Rix 2008: 145-6), so the shape *taeda* is not so problematic after all. More problematic is the lack of any attested Etruscan forms that resemble this that could be the source form (cf. Breyer 1993: 229-30). In any case, the Latin is not a regular borrowing from Greek and, if borrowed, has undergone mediation; whether this was by Etruscan simply cannot be confirmed.

Berber forms of the shape *tayda* meaning ‘Aleppo pine’ (cf. e.g. El Arifi 2014: 468) are identical to Latin in form and (presumed original) meaning. Within Berber, the lexeme seems to comprise the feminine *ta*-prefix, but this could be a reanalysis. Neither a borrowing from Latin nor a borrowing from the same substrate source as Latin can be ruled out. Theoretically, the Greek words, like the Latin, only secondarily came to mean torch from an original sense of resinous pine tree, making them only coincidentally similar to the verb δαίω ‘to kindle’. In sum, there is a chance that the Latin, Greek, and Berber words are from a substrate source. If not, then Lat. *taeda* at least has been indirectly transmitted from Greek through an unknown language.

turdus ‘thrush *vel sim.*’

Pre-form: **t(o/u)r(s)d^(h)*- | PItal. **to/ur(z)do*-

**trosd^(h)*- | PCelt. **trozdi*- | Mlr. *truit, truid* ‘starling’, etc.

**drosd^(h)*- | PSlav. **drozdъ* | Ru. *drozd* ‘thrush’

**trosd*- | PGm. **prastu*- | ON *prǫstr* ‘thrush’

**tr(u)st/d^(h)(s)k*- | PGm. **prusk(j)ōn*- | OHG *thrōsca, drōsca* ‘trush’,
OE *þrysce* ‘thrush’

**strosd^(h)*- | PBalt. **strozdo*- | Lith. *strāzdas*, Latv. *strazds* ‘thrush,
blackbird’

**stroud^h*- | PGk. **strout^ho*- | Gk. στροῦθος, στρουθός ‘sparrow *vel sim.*’

**droud*- | PArm. **artout*- | Arm. *artoyt* ‘lark’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: animal, bird

Pokorny (1096), WH (II: 718), EM (708), DV (634)

Hamp (1981: 81), Matasović (2009: 392), Meiser (2010: 63), Kroonen (2013: 545), Derksen (2014 s.v. *strazdas*), Zair (2017: 263, 266, 285), Matasović (2020: 335), Weiss (2020: 104), Stifter (fthc.), Thorsø (fthc.)

²⁰⁴ CIE 832 AR·PABASSA / ARNTHAL·FRAVNAL spells Etr. *ar(nθ).papasa / arnthal fraunal*. CIE 959 THANNIA TREBO spells Etr. *thania trepu*. While the Pyrgi bilingual spells Etr. *θefarie[i] velianas* with Phoen. TBRY? WLNS̄, the Etruscan name itself is from Lat. *Tiberius*. Thus it cannot be ruled out that the Phoenician version reflects the more common form of the name.

The first vowel of Lat. *turdus* can reflect **u* as well as **o*, or **ɣ* via the relatively irregular change **o > u/_rC* (cf. *furnus* ~ *forinus* ‘oven’). Given that several comparanda reconstruct to **-ro-*, DV (634) prefers **ɣ*.²⁰⁵ It could then be interpreted as the zero-grade of a root **(s)terd^h-* found elsewhere in the *o*-grade. But West Germanic forms may also attest to *u* vocalism (Thorsø fthc., cf. Kroonen 2013: 545) and other attestations of this root offer problems.

Italic, Celtic, and Germanic attest to initial **t*, but the Slavic form starts with **d*. Lithuanian and Latvian attest to initial **s*, as if with *s* mobile, but OPr. has *tresde* ‘thrush’ (cf. Derksen 2014 s.v. *strazdas*). Likewise beginning with **s* are Gk. στρουθός, στρουθός ‘sparrow *vel sim.*’ (cf. Hamp 1981: 81, EDG 1415). While Hamp links the problem with the vocalism to the shift in meaning, this is not an explanation. Nor is it simply a matter of vocalism. The Greek forms lack the internal sibilant,²⁰⁶ a situation reminiscent of *fracēs* (s.v.). The Greek forms also attest to **d^h*. While the Celtic, Slavic, and Baltic forms can reconstruct to **d^h* or **d*,²⁰⁷ ON *þrǫstr* requires **d*.

Hamp (1981: 81) and Kroonen (2013: 545) compare Arm. *tordik* ‘thrush’ < **dorzd^h-*, but this form is suspicious. Since the form occurs only in one dictionary compiled in Italy, it might be a loan from a Romance form like Italian *tordo* (Thorsø fthc. fn. 7). More likely to be an independent comparandum is Arm. *artoyt* ‘lark’ < **droud-* (Thorsø fthc.). Like Italic, Celtic, Germanic, and Slavic, it lacks the initial sibilant and like Greek, it also lacks the internal sibilant. The quality of the dentals also matches various other branches. The irregularities between attestations of this lexeme makes it likely to be of non-inherited origin (cf. Matasović 2020: 335, Stifter fthc.).

2.2.2.2 Non-inherited Origin is Possible

adeps, -ipis ‘fat, lard’

Pre-form: **h₂edH/ep-* | PItal. **ada/ep-*

Comp.: **h₂elH/ep-* | PRom. **ala/ep-* | Middle French *auve* ‘lard’, etc.

**h₂le/oib^h-* | PGk. **ale/oip^h-* | Gk. ἄλειφα(ρ) ‘unguent, oil’, etc.

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: culinary

²⁰⁵ Another pathway to *u*-vocalism from **ɣ* would be an irregular development of **ɣ* to *ur* that is proposed to underlie e.g. *currō* < **k₁ɣs-* and *curtus* < **k₁ɣtos-* (Meiser 2010: 63, Weiss 2020: 104). Zair (2017: 263, 266, 285), who also prefers **ɣ*, groups *turdus* with the words that show **ɣ > ur* due to borrowing from Umbrian.

²⁰⁶ Within Celtic, the Brythonic forms (W *trydw*, OCo. *troet*, etc. ‘starling’) cannot reconstruct to a proto-form with an internal sibilant either. They could reflect PCelt. **troddi-*, with a strange geminate, or be loans from Irish (Stifter fthc.).

²⁰⁷ It is unclear if Lat. *turdus* can reflect **-rzd^h-*. The chronology of the changes **rzd > rd* and **rd^h > rb* is unknown because both are pre-literary, so it is difficult to rule out the possibility that **t(o/u)rzd^h-* should have yielded **to/urd^h- > **to/urb-*.

WH (I: 12), EM (9), DV (24)

Buck (1904: 69), Sperber (1917: 541), Brück (1919b: 196-7), REW (no. 161), Bottiglioni (1943: 321), FEW (XXIV: 138), Meiser (1986: 216-18), Giacomelli (1994: 31-2), Untermann (2000: 360), Weiss (2010a: 284-94), EDG (64)

Lat. *adepts* is often compared to U **ařepes** [dat.abl.pl.], whether borrowed from it or cognate with it.²⁰⁸ But a close reading by Weiss (2010a: 284-94) of the passages in which **ařepes** appears shows that there is no reason to assume it means fat at all, and it more likely means something like ‘prayers’. Thus, it probably has nothing to do with Lat. *adepts* (cf. also DV 24).

Otherwise *adepts* is suspected of being a loan from Gk. ἄλειφα(ρ) ‘unguent, oil’, ἀλοιφή ‘anointing, ointment, grease’ (Sperber 1917: 541, Brück 1919b: 196-7, REW no. 161, Bottiglioni 1943: 321, WH I: 12 with lit., FEW XXIV: 138, Giacomelli 1994: 31-2, EDG 64), but it cannot have been direct given the Latin *d* and short monophthong. Latin variants with *l* are found in the *Appendix Probi* and several Romance descendants (e.g. Old French *awe*, Middle French *auve*, Logudorese *abile* [metathesized] ‘lard’, etc.). This has either been interpreted as remnants of the original Greek form in the face of a change to *d* in *adepts* (Sperber 1917: 541, Brück 1919b : 196-7, REW no. 161) or simply late/vulgar (WH I: 13). The poorly understood Latin **d > l* change labeled the “Sabine *l*” cannot be responsible in either case (the change goes in the opposite direction of the former and occurred too early for the latter). If the Romance forms are taken at face value, they indicate an original *l ~ d* alternation within Italic. If the Greek forms are inherited, this represents an example of a Greek word that was mediated to Latin indirectly via a third language. But as the Greek words do not have a bulletproof etymology (EDG 64 considers, but is not fully convinced of substrate origin), both the Italic and the Greek forms could be independent loans. Notably, this is the opposite of the correspondence in Lat. *laurus* ~ Gk. δάφνη.

alaternus ‘buckthorn (*Rhamnus alaternus*)’

Pre-form: **h₂elH/V-ter-(i)no-* | PItal. **alater(i)no-*

Comp.: ?MoGk. (Cretan) ἐλαίτρινος ‘buckthorn’

■ Irreg. correspondences □ Remarkable phonotactics

Semantics: plant, wild

WH (I: 26), EM (19)

Billerbeck (1824: 53), REW (no. 312), Niedermann (1916: 152), Bertoldi (1928: 233 fn.

²⁰⁸ Meiser (1986: 216-18) proposed **ad-lipa-* < **lejp-* ‘to stick’ underwent Sabellic developments and was borrowed into Latin (cf. also EM 9). Given a few cases of Umbrian *ř* for inherited **l* (**kařetu** ‘to call’ < **kalē-tōd*, **fameřias** for *familiae*, cf. Buck 1904: 69, Untermann 2000: 360), WH (I: 12) consider transmission of Gk. ἄλειφα (perhaps via Etruscan) through Umbrian to Latin. This would also account for the monophthongization of Gk. ει.

3), Battisti (1931: 648 fn. 4), Alessio (1941b: 183), Ernout (1946: 30), Holmes (1947), Carnoy (1959: 114-15), Battisti (1960: 370, 373-4), Wagner (1960-4 I: 67), Paulis (1992: 417), Breyer (1993: 404-5), Weiss (2020: 128-9)

Lat. *alaternus* is often considered a classic example of a potentially Etruscan borrowing through a combination of its religious semantics,²⁰⁹ its *-rn-* suffix, and its lack of a good IE root etymology (cf. Niedermann 1916: 152, Ernout 1946: 30).²¹⁰ There are however reasons to doubt that *-rn-* is Etruscan everywhere that it appears (cf. Holmes 1947). Nor is any similar word attested in Etruscan.²¹¹

Alessio (1941b: 183) suggests a pre-form **alater* because, while many Romance languages reflect the form with *-rn-* (Perugian *laterno*, Prov. *aladern*, Sp. *aladierno* etc., REW no. 312), It. *ilátro* ‘*Rhamnus alaternus*’ attests to **alater*. This is reminiscent of the situation in Lat. *calpar* ~ PCelt. **kelqurno-* (s.v.). Beyond Latin and Romance, comparanda are difficult to confirm. Bertoldi (1928: 233 fn. 3) notes several plant names with obscure morphemes beginning with **al-*,²¹² including Sicilian *alastra* and Mortala Ligurian *la lastra* < **alastra* ‘broom’ (note the *-str-* element). To this, Wagner (1960-4 I: 67) adds Barbagian *aláse* ‘holly’, and Urzulei *alaθūli*, recorded as meaning ‘laurel’, but which he argues also likely means ‘butcher’s broom’. For the same reason (several dialects that call broom and holly ‘spiny laurel’), Paulis (1992: 417) suggests it means ‘holly’. They all have in common the thorny excrescences on their leaves “questo è tutto ciò che si può dire per il momento”. Battisti (1960: 373-4) compares the *alastra* group to *alaternus* directly, then (1960: 370) suggests a case could be made for a Mediterranean word if it is linked with Gk. ἀτάλμυος ‘plum tree’. This latter point would however require metathesis.

Instead, the best comparison is Alessio’s (1941b: 185) Cretan Greek ἐλαίτρινος ‘*Rhamnus alaternus*’ (cf. Billerbeck 1824: 53). Between it and *alaternus*, the semantics are identical but neither can easily be a borrowing from the other. The vocalic alternation suggests independent loans from a third source. The Greek word might show that *alaternus* does not have the Etruscoid *-erna* suffix at all (or an *n-*suffix like PCelt. **kelqurno-*), but rather a sequence syncopated from **-erino* suffix (cf. inherited *hībernus* < **ǵhejmr-ino-*). If both words are independently borrowed and yet both have the suffix, it was either present in the donor language or both coincidentally added the same

²⁰⁹ Tarquiti Priscus *apud* Macrobius (*Saturnalia* 3.20.2-3): *arbores, quae inferum deorum avertentiumque in tutela sunt, eas infelices nominant: al(a)ternum, sanguinem filicem, ficum atram, quaeque bacam nigram nigrosque fructus ferunt, itemque acrifolium, pirus silvaticum, pruscum rubum sentesque quibus portenta prodigiaque mala comburi iubere oportet.* ‘Trees that are under the protection of the gods of the underworld and apotropaic ones that they call ‘unlucky’: buckthorn, blood-red(?) fern, and those that bear a black berry or black fruits, also holly, wild pear, broom (if **ruscum* for *pruscum*), briar, and the brambles with which one should order that bad portents and prodigies be burnt.’

²¹⁰ Breyer (1993: 404-5) also suggests that the lack of weakening of *a > e* is irregular and might be a form of vowel harmony, but this could simply be due to the *alacer* rule (cf. Weiss 2020: 128-9).

²¹¹ The form *alθia* given by Battisti (1931: 648 fn. 4) is a ghostword (Breyer 1993: 404).

²¹² Carnoy (1959: 114-15) interprets this as a lexeme meaning ‘red’ as found in many tree names (*almus*, *ulmus*, etc.), but this is impossible.

inherited suffix.

ālium ‘garlic’

Pre-form: *aG^h-jō- | PItal. *axoljo-

>? PVerb. *agVlum | Awjila *agflum*, Ghadames *aǧelum* ‘garlic’

Comp.: *geg^l-iHd^(h)- | PGk. *geg^līd/t^h- | Gk. *γέγλις > γέλις ‘garlic’
*a-Gl-iHd^h- | PGk. *aglīt^h- | Gk. ἄγλις ‘garlic’

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, domestic

Pokorny (33), WH (I: 30), EM (21), DV (33)

Frisk (1960-72 I: 295), Chantraine (1968-80: 214-15), Furnée (1972: 194, 390), Weiss (2010b), EDG (13, 265), Kroonen (2012b), Schrijver (2018: 362), Weiss (2020: 169, 142-3, 176-7), Kroonen (fthc.)

Lat. *ālium* has resisted analysis do to the lack of well understood comparanda. Pokorny (33) linked it and *ālum* ‘comfrey’ to Skt. *ālu-* ‘the edible root of *Amorphophallus paeoniifolius*’ (EWAia III: 25 is skeptical) through **ālu-*, **ālo-* ‘bitter plant’. This connection is given as the most probable so far by WH (I: 30), although EM (21) are suspicious and suggest that a word of this sort might not be inherited. DV (33) proposes a derivation within Italic from *āla* ‘wing’. Kroonen (2012b) instead suggests a connection with two Greek words for garlic: ἄγλις and γέλις.

The two Greek words are likely from the same root, with γέλις < *γέγλις via metathesis. It could have been formed via reduplication (Frisk 1960-72 I: 295, Chantraine 1968-80: 214-15, EDG 13, 265), but Kroonen (2012b) proposes a borrowing from Akk. *gidlu* ‘braided string, string of garlic’ with *-δλ- > -γλ- like in γλυκύς ‘sweet’. Gk. ἄγλις would be an *a*-prefixed form (**a-gdl-* or **a-ggl-*, cf. also Schrijver 2018: 362), suggesting that *gidlu* reached Greek through a substrate language. That Semitic is the source rather than an independent borrowing from a third source is indicated by Akk. *gidlu* being a specific semantic derivation of the Semitic root *gdl* ‘to braid.’ The oblique forms of the two Greek words could be variants of the Pre-Greek *vθ*-suffix, and the alternation between -īθ- and -īδ- in the oblique of γέλις led Furnée (1972: 194, 390) and EDG (13, 265) to propose a Pre-Greek origin for the word (further adducing σκελλίς, -ίδος ‘garlic’ and therefore a **g* ~ **k* alternation). But Kroonen (fthc.) notes that some of the cases of -ίς, -ίθος nouns, which look like non-nasal variants of the *vθ*-suffix, are secondary, triggered by analogy after the Attic-Ionic merger of -ivθ- and -ī- stems (e.g. ὄρνις, -ίθος ‘bird’). Thus they may not attest to original Pre-Greek morphology after all.

While Gk. γέλις < *γέγλις could derive from a root shape *GeDL and ἄγλις could be from *aGDL (if we assume **gdl* > Gk. γλ), it is difficult to get Lat. *ālium* from *GDL. There are no otherwise known examples of the reflex **gdl* in Italic, but if we assume

**gdl* > **dl*, then **agdlio* could yield attested Lat. *allium* (cf. *sella* < **sed-la-*). Weiss (2010b) finds no certain cases of the *littera* rule occurring with *ā* followed by *l* and suggests the spelling *ālium* might actually represent **alljum*. Thus the explanation of *allium* < **adlio-* < **agdlio-* could be sufficient.

But Berber forms point to the persistence of a velar rather than a dental. Marijn van Putten (apud Kroonen fthc.) reconstructs **agVlum* ‘garlic.’ In loans from Latin, an *-m* is usually never preserved, probably because they were borrowed at time when it was no longer pronounced in Latin. Unless from a different source entirely, this requires a very old loan into Berber (Maarten Kossmann, p.c.)—perhaps old enough to preserve a trace of the Italic velar.²¹³

Thus an alternative focuses on an explanation of *ālium*. The spelling *allium* occurs in inscriptions from the 1st c. CE onwards, whereas *ālium* seems to be the more correct, older spelling (cf. TLL s.v. *ālium*). Given that the Greek forms can also have developed from **GeGL* and **aGGL*, perhaps there was no dental involved.²¹⁴ In that case, Gk. ἄγλις is from **aggl-* with geminate simplification. The same formation can yield Lat. *ālium* if it entered Proto-Italic with a voiced aspirate. From there it can have undergone the development **ag^h(g^h)_l-* > **aχ(χ)ol-* > **aol-* > *āl* (for the vowel contraction cf. *Māvors* > *Mārs*).

We can cautiously propose that Gk. γέλιγλις ~ Gk. ἄγλις, Lat. *ālium* constitute an example of the *a*-prefix and attest to a **g* ~ **g^h* alternation.

aper ‘boar’

Pre-form: **h₂ep-ro-* | PItal. **apro-*
h₂ep-r-ōn-* | PItal. **aprōn-* | U **abrunu [acc.sg.], etc. ‘boar’

Comp.: **h₁ep-r-* | PGm. **ebura-* | OE *eofor*, OHG *ebur* ‘boar’, etc.

**h₁ep-er-* | PGk. **epero-* | Aeol. ἔπερος ‘ram’

**μep-r-* | PBSI. **weprios-* | Latv. *vepris* ‘castrated boar’, OCS *veprь* ‘boar’, etc.

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: animal, wild

Pokorny (323), WH (I: 56), EM (38), DV(46)

Skutsch (1901-3: 67), Meillet (1925: 9), Chantraine (1933: 221), Chantraine (1968-80: 324-5), Frisk (1960-72 I: 468), Schrijver (1991: 29-30), Untermann (2000: 44-6),

²¹³ Loans of this age from Latin into Berber are otherwise unknown, as are any non-Latin Italic loans into Berber (Maarten Kossmann, p.c.).

²¹⁴ If indeed ultimately from Akkadian, perhaps the sequence GDL had been simplified to GGL in the substrate donor language (which, given the *a*-prefix, was the more proximal source of the words in Latin and Greek). Note that in γλυκός, the development of **δλ* > *γλ* is considered *ad hoc*.

Derksen (2007: 515), Kroonen (2013: 114, 457, 589), EDG (438), Barrios-Garcia & Ballari (2012: 2284), Šorgo (2020: 461)

Italic **apro-* is explained to be from **epro-* with contamination from *caper* ‘goat’ (Skutsch 1901-3: 67, followed in e.g. WH I: 56, Schrijver 1991: 30, DV 46). Given the Umbrian derived forms (cf. Untermann 2000: 44-6), the contamination would have to have occurred in Proto-Italic. This is reminiscent of the suggestion that OIr. *gabor* has its *g* from **g^haid^h-* (s.v. *caper*), but at least in this case, the proposed form is actually attested. Kroonen (2013: 114) suggests taking the vocalic alternation at face value, and in light of the irregularity of some of the other comparanda, I agree this is the best way forward. Balto-Slavic attests to boar words of a very similar shape except that they have an otherwise unexplained initial *v* (Derksen 2007: 515).²¹⁵

Aeol. Gk. ἔπερος ‘ram’ looks like a reflex of this boar word, and is adduced into the family by Meillet (1925: 9). EDG (438) however follows Chantraine (1968-80: 324-5) and Frisk (1960-72 I: 468) in strictly rejecting it, connecting it rather to εἶρος ‘wool’ via a compound with ἐπι, thus ‘who carries wool’. Given the attestation of Hsch. ἔβρος: τράγος βατής ‘a he-goat that mounts’, often suspected of being related to the boar word family (Pokorny 323, Schrijver 1991: 29, DV 46), it seems quite likely that ἔπερος ‘ram’ and κάπρος ‘boar’ are simply relatives of Lat. *aper* and *caper* that have switched meanings.

If all of these words indeed belong together, they attest to an *a ~ e* vocalic alternation. The *v*-element in Balto-Slavic is strange, but has been compared to an element **wi-* analyzed as a prefix in PGM. **wisund-* ‘wisent’ and Gaulish *uisumarus* ‘clover’ (Kroonen 2013: 457, 589; Šorgo 2020: 461). Its rarity and lack of a clear distribution make this difficult to confirm. It is interesting that a word for boar should have been borrowed from a non-IE language, as the range of the wild boar extends across Europe into the steppe (cf. Barrios-Garcia & Ballari 2012: 2284).

arāneus ‘spider’

Pre-form: **h₂erh₂(k-)s-n-* | PItal. **ara(k)snejo-*

Comp.: **h₂erh₂k-s-n-* | PGk. **arak^hnā-* | Gk. ἀράχνη etc. ‘spider’

□ Irreg. correspondences

■ Remarkable phonotactics

Semantics: animal, wild; insect²¹⁶

Pokorny (55-61), WH (I: 61-2), EM (42-3), DV (49)

Curtius (1894: 398), Lewy (1895: 121-2), Lidén (1905: 507-8), Walde (1910: 54-5), Ogle (1945: 132), Gil Fernández (1959: 24-6), Beekes (1969: 34), Biville (I: 813),

²¹⁵ This in part led Meillet (1925: 9) followed by EM (38) to propose that a root ‘goat/boar’ **aper* sometimes received a *k*-prefix (Lat. *caper*, Gk. κάπρος ‘boar’, etc. cf. also Chantraine 1933: 221) and in Balto-Slavic a *v*-prefix.

²¹⁶ In modern biological taxonomy, not an insect but rather an arachnid.

Martirosyan (2009: 270), EDG (123), Rosoł (2013: 18, 162), Cunningham (2018-20 I: 317), Weiss (2020: 183), Höfler & Nielsen (2022)

Lat. *arāneus* (also occurring as fem. *arānea*) is sometimes suspected of being a loanword from Gk. ἀράχνη (cf. EM 42-3), but this cannot be the case. Early loans from Greek substitute *c* for χ (Biville I: 183), and before *n* this should probably have given *gn* (cf. *dignus* < **deknos*, Weiss 2020: 183), which would not disappear. Additionally, Gk. χμ was borrowed into Lat. as *-cum-* (cf. *dracuma* < δραχμή, DV 49) and in later loans, even Gk. χν was borrowed with an anaptyctic vowel (cf. τέχνᾰ ‘trick’ > Lat. *techina*, Weiss 2020: 183). Thus we expect some remnant of borrowed χ regardless of the age of the loan. Instead, the most likely scenario is that Latin and Greek go back to the same pre-form like **araksnā-* (WH I: 61-2, Gil Fernández 1959: 25, EDG 123), cf. environment in **l(ēō)uk-sn-* > Gk. λύχνος ‘lamp’, Lat. *lūna* ‘moon’ (Biville I: 183).

The pre-form **araksnā-* is difficult to reconstruct to PIE, however. An inherited root shape **HerHk-* does not seem possible, so Gil Fernandez (1959: 24-6) suggests the velar element is a **k*-enlargement.²¹⁷ But on an otherwise unattested root, this is suspicious as well. The difficulties suggest an originally non-IE disyllabic root.

Höfler and Nielsen (2022) have most recently argued for a root **h₂reh₂g-* ‘to weave’ behind the Latin and Greek forms. Gk. ῥώξ in the meaning ‘venomous spider’ could be an agentive root noun **h₂roh₂g-s-* ‘weaver’ with initial laryngeal loss due to the de Saussure Effect. They propose that the original *s*-stem of which **araksnā-* is a double-zero-grade derivative is still present in Gk. ῥῆγος ‘rug, blanket’ < *(*h₂*)*reh₂g-os*. The pre-form **araksnā-* would have arisen in both Latin and Greek via the *palma* rule form **h₂ʰh₂g-s-neh₂-*. However, as seen from ῥῆγος, to accept this etymology, we must also accept (1) a rule for Greek where **#h₂RVh₂C-* > **#RVh₂C-* and (2) that a root reconstructible for Latin and Greek alone can be projected back to PIE. Thus I remain not fully convinced that an Indo-European etymology has been found.

Interpretations of Lat./Gk. **araksnā-* as a loan from Hebr. *arāg* ‘to weave’ (hesitantly Lewy 1895: 122) are rejected by Rosoł (2013: 162) on semantic grounds. On the theme of spinning/weaving, Curtius (1894: 398) compared Gk. ἄρκυς ‘net’ and Hsch. ἀρκάνη: τὸ ῥάμμα, ᾧ τὸν στήμονα ἐγκαταπέκουσι διαζόμεναι ‘thread with which the warp is intertwined when they are setting it up in the loom.’ Lidén (1905: 507-8) rejected the link with **araksnā-* in favor of a connection to Gk. ἄρκευθος ‘juniper’ and Balto-Slavic words for willow. EDG (132-3) prefers keeping all forms separate for semantic and morphological reasons. Walde (1910: 54-5) mentioned a possible relationship with OE *reng*, *rynge* ‘spider, spider’s web’ (cf. Beekes 1969: 34, whose reconstruction does not work due to Kluge’s Law). Proto-Germanic **rengjo-* could reconstruct to **Hrŋg^h-jeh₂-* alongside Gk. ἀράχνη < **h₂rŋg^h-neh₂-*, but the Latin form cannot

²¹⁷ He notes Hsch. ἄρασιν: ἀράχνην in Latte’s edition of Hesychius, taking it at face value against ἄρασιν elsewhere to suggest it represents a *si*-suffixation of the root *ara-*. As Cunningham (2018-20 I: 317) notes however, the actual codex unicus of the manuscript has ἄρασιν.

accommodate this pre-form. The Germanic form, if we assume metathesis, would look similar to Martirosyan's (2009: 270) explanation for Arm. **ernjak* 'spider' as a form with regular prothetic *e* before original initial *r* in **ra(K)nj-* < **raKn-jeh₂-*. But as Rasmus Thorsø (p.c.) has pointed out, **ernjak*, corrected from attested *ērnjak*, occurs only in the Erzurum dialects, and is almost certainly a loan from Turkish *örümcek*, *erimcak* 'spider' (from *örmek* 'knit, weave').²¹⁸ In the end, OE *reng*, *rynge* is likely simply borrowed from Old French. Once attested is *reingne*, which seems to be a variant dialectal form for *araigne*, *iraigne*, and *yrainne* etc. attested elsewhere (Ogle 1945: 132).

The Latin and Greek forms remain isolated. If the full root is indeed **arak-*, a reconstruction of **h₂erh₂k/g-* or perhaps **h₂ʔh₂k/g-* does not look to be a valid PIE root structure. Instead, they are likely loans.

ardea 'heron'

Pre-form: **H(e)rd-* | PItal. **ardeja-*

Comp.: **h₁rōd-*, **h₁roHd-*, **h₁reh₃d-* | PGk. **erōd-* | Gk. ἐρῳδιός 'heron'

?* *h₂3erd-*, *(*H*)*ord-* | PGm. **artō(n)-* 'teal/garganey/wagtail'

?**Hrod^h-* | PSlav. **rodā-* | SCr. *roda* 'stork'

?**h₁reh₃d-* | Arm. *arat* 'stork'

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: animal, bird; aquatic

Pokorny (68), WH (I: 64), EM (45), DV (52)

Cuny (1910: 160), Frisk (1960-72 I: 572), Ciorănescu (1958-66 s.v. *rátǎ*), André (1967: 33), Chantraine (1968-80: 337), FEW (X: 420-1), Puhvel (I: 176), Schrijver (1991: 65, 314), Schrijver (1997), Orel (1998: 374), Tótfalusi (2001 s.v. *réce*), Derksen (2007: 437), Martirosyan (2009: 126), EDG (464, 468), Matasović (2020: 339)

Greek ἐρῳδιός 'heron' is the most secure comparandum for Lat. *ardea* 'heron', both semantically and formally.²¹⁹ In light of the well-attested variant ἐρῳδιός, the *iota subscriptum* is likely secondarily built on other endings in -ίδιος (WH I: 64, EDG 464, etc.). EDG takes the Greek variants ἀρῳδιός and ῥῳδιός at face value, as did Cuny (1910: 160), consequently proposing Pre-Greek origin. The former variant is late (from

²¹⁸ The first to notice this seems to have been Vahagn Petrosyan on Wiktionary (wiktionary.org/wiki/Էրնյակ, accessed Feb. 7, 2022).

²¹⁹ Vennemann (2003: 325-6) instead proposes a connection with Sp. and Pt. *garza* 'heron', taking both as borrowings from Basque. PVasc. **gardea* by regular sound change would have lost its initial **g*, resulting in Lat. *ardea*. The Basque word is attested as *koartza* with its initial velar intact. While initial velar loss is sporadic (Trask 2008: 27), it seems difficult to reject the conclusion of Corominas & Pascual (1984-91 III: 116) that the borrowing went the other way, from Spanish into Basque. Instead of comparing Gk. ἐρῳδιός, Vennemann adduces Gk. χαρῳδιός 'name of a bird, perhaps plover' making his argument doubly dubious.

the Septuagint, cf. Schrijver 1991: 65), making it suspicious. But that the latter represents a secondary loss of ϵ (Chantraine 1968-80: 337) seems *ad hoc*.

The only root shape that can be reconstructed to produce the Latin and Greek forms is $*h_1red-$ with unusual but not unattested $*\bar{o} \sim *\emptyset$ ablaut: Gk. $\acute{\epsilon}\rho\omega\delta\acute{\iota}\omicron\varsigma < *h_1r\bar{o}d-$ (although accepting the validity of $\acute{\alpha}\rho\omega\delta\acute{\iota}\omicron\varsigma$ would contradictorily require an initial $*h_2$) and Lat. *ardea* $< *Hrd-$ (DV 52). Since the Greek form shows that the root vocalism is in the second syllable, the first vowel must be from a laryngeal. Thus, we cannot have root vocalism before the resonant in Latin without proposing unconditioned Schwebeablaut, and a pre-form like $*h_2erh_3d-$ for Latin is not possible. Nor can Latin represent a zero-grade of a Greek pre-form $*h_1roHd-$. In a form like $*HrHd-$, if we assume that vocalization took place from the right, the assignment would yield $*Hr\overset{\circ}{H}d-$. Thus the initial laryngeal would be lost before the sequence *CV* yielding $**rad-$ (cf. also Schrijver 1991: 314).

The appurtenance of Germanic $*art\bar{o}(n)-$ (reconstructed based on several daughter forms: ON *arta* ‘teal, garganey’ and dim. *ertla* ‘wagtail, Icel. *urt, ört* ‘teal’, Sw. *årta* ‘garganey’, etc. [Kroonen 2013: 36]) is questioned by some (cf. Frisk 1960-72 I: 572, Schrijver 1991: 65, EDG 468, Matasović 2020: 339), while others (cf. WH I: 64, André 1967: 33, EM 45, Kroonen 2013: 36) adduce it nonetheless. If related, a pre-form $*ard-$, $*ord-$ $< *h_1ord-$ for Germanic (Schrijver 1991: 65), would create the exact problem that we needed to avoid for Latin: root vocalism in front of the resonant, creating unconditioned Schwebeablaut variation.

SCr. *róda* ‘stork’ is frequently adduced as a comparandum to Lat. *ardea*, with more certainty that the Germanic even (e.g. WH I: 64, Chantraine 1968-80: 377, Frisk 1960-72 I: 572, André 1967: 33, EDG 468). Though it is semantically closer to the Latin and the Greek, its attestation (almost) exclusively in the Štokavian dialects (Matasović 2020: 339) is highly suspicious. If related, its dental must reconstruct to a voiced aspirate $*d^h$ (cf. Kroonen 2013: 36), as $*d$ would yield the Winter’s Law outcome $**r\bar{a}da$ (Schrijver 1991: 65). Due to its isolation a loan from Greek or (unattested) Romance has been suspected (cf. Schrijver 1991: 65, Matasović 2020: 339).

Arm. *arat* ‘stork’ is a hapax, occurring as gen.sg. *aratay* in Vardan Areveltsi’s commentary on *Psalms*. The interpretation is complicated by the extreme rarity of this genitive formation and its appearance next to a word that seems to be an Armenian transcription of the Greek word for stork. If *arat* itself indeed means stork, it is attractive to adduce it as a comparandum, but requires the reconstruction $*h_1reh_2d-$ in Indo-European (Martirosyan 2009: 126), which, as demonstrated above, cannot be reconciled with the Latin form.

If the Germanic group (the only group outside of Latin and Greek with secure enough attestation to be reconstructible to a proto-form) is related to the Greek and Latin forms, then the resulting fluctuating ablaut creates a problem for the reconstruction of a common proto-form. We end up with lengthened *o*-grade $*h_1r\bar{o}d-$ against zero-grade

**hird-* against unconditioned Schwebeablaut full *o*-grade **hjord-* (if we at least wish to be able to reconstruct all with the same quality laryngeal). Given the difficulties provided by the reconstruction of the initial syllable, Kroonen (2013: 36) suggests PGm. **artō(n)-* might be a case of *a*-prefixation (cf. Schrijver 1997). SCr. *rōda* would seemingly fit into this pattern if it belongs here. The Greek forms disrupt the classic distribution in that the prefixed forms (prefixed with both *a-* and *e-*) maintain full root vocalism. Arm. *arat*, if it belongs here, also requires full root vocalism in a prefixed form. Lat. *ardea* and its comparanda thus do not represent a Paradebeispiel of the *a*-prefix, but the discrepancy between the Latin, Greek, and Germanic places it amongst the lexemes of likely non-IE origin.²²⁰

Puhvel (I: 176) links Hitt. *arta-* ‘a bird-name,’ but without any further indication of the type of bird this represents, it must be left out.

bāca ‘berry, fruit, nut’

Pre-form: **behzk-* | PItal. **bākā-*

Comp.: **ba/hzk-* | PCelt. **bak-* | W *bagad, bagwy* ‘cluster, bunch, troop, flock’,
OBret. *bacat* ‘berry’, LCo. *bagaz* ‘bush’

?PPerb. **bqā* ‘blackberry, mulberry’

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: plant, berry

WH (I: 91), EM (63), DV (67)

Havet (1911: 219), Juret (1918: 195 fn. 1), FEW (I: 196), Wilamowitz-Moellendorff (1931-2 II: 63), REW (no. 859), Battisti & Alessio (1950-57 I: 392), Deroy (1956a: 188-9), Frisk (1960-72 I: 212), Chantraine (1968-80: 159), Boutkan & Kossmann (1999: 89), Weiss (2020: 82), van Sluis (fthc.)

While the variant *bacca* is poorly attested²²¹ and most Romance languages continue

²²⁰ Very likely unrelated but worth mentioning due to the semantic change assumed to have occurred within Germanic is a family of words for ‘duck’. These include Alb. *rosë*, Rom. *rață*, SCr., Slov. *raca*, Serb. (dial.) *race*, and Bulg. *rjaca*. Orel (1998: 374) assumes that PAlb. ***anājā-*, the expected reflex of the inherited duck word, was contaminated to **arājā-* and that Rom. *rață* was borrowed from Proto-Albanian. Ciorănescu (1958-66 s.v. *rățā*) however considers a borrowing from a Slavic source to be more obvious, ruling out a connection to “Dacian” which seems to be often proposed as a source. Tótfalusi (2001 s.v. *réce*) compares Hungarian *réce* ‘duck’ as an independently developed onomatopoeic animal call word. FEW (X: 420-1) says the same about Occitan *rit* ‘duck’. Further similar duck words include Friulian *raze* and German *Rätsche*. I am suspicious of proposals of widespread onomatopoeias and especially of etymologies that conclude words began as calls for animals. However, given this widespread duck word of the shape **rVts/ʃ-*, it does not seem necessary to follow Orel (1998: 374) in deriving Alb. *rosë* from a contamination of the inherited duck word.

²²¹ In manuscripts, it seems to occur only in Priscian. The earliest, like those of Vergil, all have *baca*. Thus it has been suggested to be scribal error (Havet 1911: 219, Juret 1918: 195 fn. 1), due to assimilation to *vacca* ‘cow’ in a tradition where the difference between *b* and *v* was neutralized.

**bāca/bācus* (FEW I: 196, REW no. 859), Italian attests *bacca* ‘juniper berry, fruit without seeds’ (Battisti & Alessio 1950-57 I: 392). To have entered Italian means it was in actual use. If the *littera* rule only applies to high vowels (Weiss 2010b), then *bacca* represents a true alternation with a geminate. Battisti and Alessio (1950-57 I: 391, 392) consider it a loan from a substrate.

Another potential indication of a non-native origin of this word is its relationship to PCelt. **bak-*. Most share a dental suffix that dates to Proto-Brythonic, but the suffix of *W bagwy* is obscure (van Sluis fthc.). This could be interpreted beside PItal. **bāk-* as an IE alternation between a zero-grade **bh₂k-* and a full-grade **b_hzk-*, but **b* is extremely rare in IE roots.²²² It seems unnecessary to reconstruct a root with **b* to PIE on the basis of Italic and Celtic alone. If the geminate in Italian is original, it strengthens the case for a substrate loan.

Proto-Berber **bqā* ‘blackberry, mulberry’ is unlikely to be borrowed from Latin due to the absence of the first long vowel (Boutkan & Kossmann 1999: 89). If it is related, it is an independent comparandum. Further connections to *bāca* are difficult to substantiate. WH (II: 91, cf. also Deroy 1956a: 188-9, EM 63) consider it a Mediterranean loan with original viticultural semantics,²²³ and Varro says that wine in Spain is called *bacca*. The Latin word and the word from Iberia could well be related, but whether *bacca* ‘wine’ is a semantic development from *bāca* ‘berry’ (i.e. *bacca* is Iberian Latin) or whether they both continue a non-IE lexeme (i.e. *bacca* is non-IE Iberian) is difficult to say without further comparanda.

badius ‘brown, chestnut-colored (of horses)

Pre-form: **ba/Hd^h-io-*, **bh₂ed^h-io-* | PItal. **baḫjo-*

Comp.: **b^(h)h₃ed^(h)-io-*, **b^(h)(h₂)od^(h)-io-* | PCelt. **bodyo-* | OIr. *buide* ‘yellow’²²⁴

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: color, equestrian

Pokorny (92), WH (I: 92), EM (64), DV (67)

²²² There is perhaps only one other PIE root that begins with **b*, **bel-* ‘strong/strength’ (Skt. *bala* ‘power, strength’, Gk. βελτίων ‘better’, *dē-bil-is* ‘weak’, Rus. *bol’šoj* ‘big’ [cf. Weiss 2020: 82]), and even here its reconstruction is debated. Alexander Lubotsky (p.c.) adduces PSlav. **debelb-* ‘fat, strong’. The lack of Winter’s Law shows that it is from **d^(h)eb^h-el-*. The other forms would be from an old comparative of this root **db^hel-ios-* > **bel-ios-*.

²²³ But they certainly go too far when they connect it to Βάκχος ‘Bacchus’. In any case, the origin of the Greek theonym is unclear. A Lydian-Greek bilingual inscription where Lyd. *Bakivalis* translates Gk. Διονυσικλέους leads DV (67) to follow interpretations like those of Wilamowitz-Moellendorff (1931-2 II: 63) in suggesting that the Greek word is borrowed from a Lydian source. Chantraine (1968-80: 159) finds a borrowing from Greek into Lydian possible here, and Frisk (1960-72 I: 212) finds it more likely even.

²²⁴ The only non-onomastic representative. Otherwise placenames like *Baiocasses/Bodiocasses* (Bayeux of tapestry fame) might comprise this lexeme (Delamarre 2003: 63).

Meyer-Lübke (1903: 92), Thurneysen (1946: 50), Wagner (1953: 388), Schmidt (1966: 160-1), Corominas and Pascual (1984-91 I: 550), Schrijver (1991: 454-65), Delamarre (2003: 63), Matasović (2009: 70)

The Latin and Irish words reconstruct to proto-forms with differing vocalism: *a* for Latin and **o* for Irish. Beginning from **bHd^h-jo-* > **badio-*, OIr. *buide* could be the result of **a* raised between a labial and a palatal consonant (Thurneysen 1946: 50).²²⁵ But in cases like this, both alternates are usually preserved in Irish (e.g. *moirb/mairb*, *muig/maig*), and there is no such by-form of *buide* (DV 67).²²⁶ Laryngeals and ablaut could produce the alternation (**bHd^h-* ~ **bHod^h-*, **bh₃d^h-* ~ **bh₃ed^h-*, or **bh₂ed^h-* ~ **bh₂od^h-*), but Lat. *badius* requires the reconstruction of **b*. As with *bāca* (s.v.), it seems unreasonable to reconstruct an additional PIE root beginning with **b* based on comparanda attested exclusively in Italo-Celtic.²²⁷ The pair is likely not inherited (cf. Pokorny 92, DV 67, Matasović 2009: 70), and the *a* ~ *o* alternation is original. If indeed with a suffix **io*,²²⁸ Latin *badius* seems to show the reflex of **d^h* (as an original **d_i* would have yielded *ii* (cf. *peīior*, DV 67), but this change occurred before the development of **d^h* > *d* (Weiss 2020: 172). Thus, if loaned into Proto-Italic, the shape was **bad^hio-*; if later, **badyo-* is possible.

barba ‘beard’

Pre-form: **ba/Hr(s?)d^h-* | PItal. **bar(z?)pā*

Comp.: **b^(h)a/ord^(h)-* | PBSI. **bordā?* | OCS *brada*, Ru. *borodá*, OPr. *bordus*, etc. ‘beard’

**b^(h)a/or^(s)d^(h)-* | Lith. *barzdà*, Latv. *bârzda* ‘beard’

**b^ha/or(s)d^h-* | PGM. **bar(z)da-* | ON *barð* ‘brim, prow; beard’, OE *beard* ‘beard’, etc.

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: body part

Pokorny (110), WH (I: 96), EM (66), DV (69)

²²⁵ Thus Schmidt (1966: 160-1) considers the possibility that Latin *badius* is a loan from Gaulish, given that several other Latin borrowings from Celtic are in the equestrian sphere. He still considers the Celtic word to be of non-IE origin.

²²⁶ Schrijver (1991: 454-65) details the unrounding of **o* to Latin *a* after labial consonants, but there are no examples of this occurring after **b* or, more fatally, **p*. Thus it seems unlikely that a pre-form **bod^h-jo-* could produce *badius* (pace DV 67).

²²⁷ WH (I: 98-9) take Sp. *bazo* ‘brown, almost yellow’ as an independent comparandum to Lat. *badius*. Corominas and Pascual (I 1984: 550) instead suggest that *basus* and *bazo* are reflexes of *badius* (cf. its potential attestation in a Latin gloss as *basus*: *rufus*, *niger*, Meyer-Lübke 1903: 92). This is difficult to believe, given that Sp. *bayo* ‘bay (of a horse)’ exists and is the regular reflex of Lat. *badius*. Thus Sp. *bazo* ‘brown’ is probably the same as *bazo* ‘spleen’ (cf. an explanation by Wagner 1953: 388) and therefore unrelated.

²²⁸ Rather than **iyo* (**iHo*).

Pedersen (1895: 72-3), Schrijver (1991: 488), Kuiper (1995: 66), Derksen (2007: 55), Kroonen (2011: 150-1), Pronk-Tiethoff (2012: 242-4), Kroonen (2013: 54), Derksen (2014 s.v. *barzdā*), Weiss (2018: 439-40), van Beek (2022: 365-6)

If taken at face value, the initial *b* of Lat. *barba* can only reflect PIE **b*. It is clearly related to Baltic, Slavic, and Germanic words for beard, but the details of the relationship are complex. Baltic forms like Lith. *barzdā* and Latv. *bārzda* ‘beard’ have a sigmatic element that does not appear in OPr. *bordus* or the Slavic forms. Derksen (2007: 55, 2014 s.v. *barzdā*) reconstructs PBSI. **bordā?*, since *-z-* would not be lost in Slavic (Pedersen 1895: 72-3), but this means that the East Baltic forms require a different pre-form.

Kroonen (2011: 150-1, 2013: 54) reconstructs the Germanic beard words as an *o*-grade of the root **b^hresd^h-* that in the *e*-grade and zero-grade elsewhere produces words for ‘board,’ ‘edge,’ and ‘tip.’ ON *barð* < **barzda-* means both ‘edge, prow’ and ‘beard’. Evidence that this root is inherited is van Beek’s (2022: 365-6) proposal that it is present in Gk. (epic and poetic) *πέρθω* ‘to raze, pillage’, with some attestations pointing to an original meaning ‘to cut off, shave’. Kroonen (2011: 150) interprets the position of the Germanic vocalism as a secondary development on the result of the zero-grade reflex in PGM. **burzd-*. This allows him to propose that the Baltic forms like Lith. *barzdā* owe their vocalism and sigmatic element to a Germanic borrowing. The Balto-Slavic forms without a sibilant could be borrowed from West Germanic, though they have mobile accentuation, which does not seem to occur in loans from Germanic (Pronk-Tiethoff 2012: 242-4). A Germanic borrowing into Latin would explain the *a*-vocalism there, but it requires a borrowing into Proto-Italic, which seems remarkably early. An alternative reconstruction for the Germanic (cf. Kuiper 1995: 66) keeps it separate from **b^hersd^h-* and thus does not include the **s*.

If not very early borrowings from Germanic, then Lat. *barba* is a representative of a substrate lexeme for which the only vocalic reconstruction that fits all comparanda is **a* (cf. Schrijver 1991: 488, Kuiper 1995: 66, Derksen 2007: 55, DV 69). Within Balto-Slavic there is the alternating presence of a sigmatic element (cf. the same in *fracēs* and *turdus*). Whether an **s* would have blocked the change PItal. **r^h* > *rb* is unclear (cf. fn. 207). All forms except for the Latin can be reconstructed to initial **b^h*. The **b* required by the Latin has been interpreted as an assimilation of **farba* > *barba* (WH I: 96, EM 66, recently Weiss 2018: 439). Given the other irregularities in this word, this need not be the case; it could instead be the result the borrowing process.

bolunda ‘wild, immature fig’

Pre-form: **bol-und^(h)-* | PItal. **bolundā*

Comp.: **(u)ol-und^h-* | PGk. **(w)olunt^ho-* | Gk. ὄλωνθος, ὄλυνθος ‘wild, unripe fig’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, fruit

Loewe (1884: xiv), Rönsch (1886: 317-18), Alessio (1944a: 138-9), André (1956: 55), Furnée (1972: 198), Biville (I: 89-90), EDG (1074), Kroonen (fthc.)

Lat. *bolunda* is found three times in glosses, twice for Gk. ὄλυθος (*CGILat.* II 382.40²²⁹; 517.40²³⁰) and once itself explained with *grossi primari fuci* (read: *fici*)(*CGILat.* II 570.16²³¹). The more widely attested Greek word clearly has the Pre-Greek *vθ*-suffix.

There have been several explanations proposed for the correspondence between the Latin and Greek words. Rönsch (1886: 317-18) considered the Latin word borrowed from the Greek with folk etymological changes (comparison to words from the root *bol-* ‘to throw, fall’ and interpretation as a future participle in *-unda*). Several suggest *bolunda* is a borrowing from a Doric Greek dialect with an original digamma (Alessio 1944a: 138-9, André 1956: 55, Biville I: 90, EDG 1074). While Alessio proposes that the gloss be corrected to **volunda*, Biville thinks that the late attestation might allow for *ɸ* pronounced as /β/ or /v/ to have been taken into Latin at a time when *b* was on the way to changing into /β/ then /v/. There are extremely few parallels for this.²³² And while Alessio and Biville note that the voicing of *nt* to *nd* is common in southern Italy, this is in the modern Italian dialects. Thus we would have to assume that Gk. *vθ* was borrowed as Lat. *nt* (the expected result) and that this was voiced to *nd* dialectally before being recorded in the glosses. For this reason Furnée (1972: 198) instead takes Greek *-vθ-* against Lat. *-nd-* as a substrate alternation, with both words independently borrowed from a third source (cf. also Kroonen fthc.).

Given that *bolunda* is attested in the 8th c. Cyrillus Glossary, and assuming that no part of it is the result of scribal corruption,²³³ it is not certain that it was acquired late enough to show the changes from Greek postulated by Alessio and Biville. Thus Furnée’s analysis cannot be ruled out, and *bolunda* might show that the Gk. *-vθος* suffix occurs in the substrate of Latin as *-unda* (s.v. *harundō* and *hirundō*).

calx, -cis ‘limestone, chalk’

Pre-form: **ka/Hlk-* | PItal. **kalk-*

Comp.: **g^ha/hzl-ik-* | PGk. **k^halik-* | Gk. χάλιξ ‘small stone, gravel, rubble’

²²⁹ Cyrillus Glossary (8th c.); the Stephanus manuscript has *bolundum*.

²³⁰ Glossae Servii Grammatici.

²³¹ Glossae Nominum.

²³² Cf. discussion in Biville (I: 78, 88): In Laconia, β was used to write digamma from the end of the 5th c. BCE. And Latin grammarians seem to sometimes have called it *bau* instead of *uau* (like Marius Victorinus, Keil *GL* VI 15.4-5). That *belena* in Quintilian (*Instituto Oratoria* 1.4.15) spells Ἑλένη ‘Helen’ (which can be presumed to have originally had a digamma based on e.g. the spelling *Velena* in Sergius’ commentary on Donatus [Keil *GL* IV 476.16-17]) does not seem certain; based on the context, it may be a spelling of *ballaena* ‘whale’.

²³³ Loewe (1884: xiv) notes for the *Glossae Nominum* that corrupt lemmata are not rare, noting importantly on line 258 *bafer* for *afer*. The *b* seems to have appeared *ex nihilo*, which would solve the lesser of two problems for *bolunda*.

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: geography

WH (I: 145), EM (89), DV (86)

Cuny (1910: 160), Alessio (1941b: 219), CAD (K: 62-4), Furnée (1972: 137, 384), Biville (II: 144-5), EDG (660, 1610), Rosoł (2013: 212)

Lat. *calx* ‘limestone, chalk’ is certainly related to Gk. *χάλιξ* ‘small stone, gravel’, and so close to it that it is often considered a loan from it (WH I: 145).²³⁴ The reflection of Gk. *χ* with Lat. *c* is expected, but the syncope is not (*pace* Cuny 1910: 160 who writes the opposite). This is further complicated by the presence of the vowel *i* in the verb *calicāre* ‘to whitewash (paint with lime)’, leading some to suggest that both represent independent forms and, with no good IE etymology (cf. EDG 1610), a non-IE Mediterranean origin (Biville II: 144-5, EM 89, DV 86).

The *i* did not syncopate in e.g. *calix* ‘vessel for food or drink’ (s.v.), nor does *calicāre* require an anaptyctic vowel in light of *calcāre* ‘to trample’. On the other hand, the verb *calicare* is rare, attested in an inscription and otherwise only in lexicographical texts that gloss it with the more usual *albāre*. Biville (II: 144-5) argues convincingly that it is harder to explain the syncope in *calx* than it is to assume independent (yet related) origins of *calx* and *χάλιξ*, with a later derivation of the verb *calicāre* based on Greek.²³⁵

If the Latin and Greek represent independent forms, then we have a non-IE *k ~ k^h* alternation like in *orca ~ ὄρχη* (s.v)(Alessio 1941b: 219, Furnée 1972: 137, 384). Despite Furnée’s (1972: 137, 384) comparisons of the family to Sum. *kalga* and Akk. *kalakku*, both purportedly meaning ‘limestone’, this is a mistake. Rosoł (2013: 212) shows that Akk. *kalakku* instead means ‘excavation; silo’ (cf. CAD K: 62-4 ‘excavation, truncated pyramid (as a geometrical term); storehouse, storeroom, silo; a container, a box, a vessel; a specific kind of chair; raft’).²³⁶

caput ‘head’Pre-form: **ka/Hp-ut-* | PItal. **kaput-*

Comp.: **ka/Hp-ut-* | PGm. **habuda-* | ON *hǫfuð*, OE *hafud* ‘head’
 **ka/oup-ut-* | PGm. **haubuda-* | ON *haufuð*, OE *hēafod* ‘head’
 **ka/oup-et-* | PGm. **haubeda-* | Go. *haubiþ*, OHG *houbit* ‘head’
 **ka/o/Hp-u(t)-lon-* | PGm. **hafulan-* | OE *hafola*, *-ala*, *-ela* ‘head’

²³⁴ Lat. *calculus* ‘pebble’ is either a diminutive of *calx* or an independent reduplicated formation **kal-kal-o-*. The idea that *calculus* is reduplicated rather than simply a diminutive stems from comparison with Gk. *κάλλιξ* ‘small stones, river gravel’, which EM (89) support, WH (I: 145) reject, and EDG (660) does not even mention.

²³⁵ An alternative explanation of *calicāre* is a dissimilation from **calcicāre* (Michael Weiss, p.c.).

²³⁶ WH (I: 145) had already supported rejecting their comparison on historical grounds: apparently limestone was not used in Greece until after Themistocles, and the technology of lime burning spread to Greece from Carthage where it originated.

*ka/Hp-uk- | PCelt. *kaφuko- | Ir. *cuäch*, W *cawg* ‘cup, dish’

*ka/Hp-ut- | PCelt. *kaφuto- | Ir. *cuäd* ‘cup, mug’

?*kap-o/ā/ēlo- | PIr. *kapālo- | Skt. *kapāla-* ‘cup, jar, dish; skull’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: body part

Pokorny (529-30), WH (I: 163-4), EM (98-9), DV (91)

Nussbaum (1986: 214), Schrijver (1991: 100-1), EWAia (I: 300), Boutkan (1995: 2-3), Beekes (1996: 218-20), Schrijver (1997: 295), Boutkan (1998: 111), EDG (658), Kroonen (2013: 215), van Sluis (fthc.)

Lat. *caput* is related to several Germanic words for head, between which Beekes (1996: 218-20) demonstrated irregularities including an *a* ~ *au* vocalic alternation like in *caupō* ~ *κάπηλος* (s.v.). The form with the diphthong has been explained through the presence of *u* in the following syllable (Boutkan 1998: 111, DV 91) and via metathesis in the oblique cases from an original proterodynamic **kh₂p-ut*, **kh₂p-uet-os* > **hafuf*, **habwehaz* (Kroonen 2013: 215).²³⁷ The former explanation is *ad hoc*, and as to the latter, Boutkan (1995: 2-3) has argued that suffixal ablaut in *t*-stems had been leveled, such that no trace should have remained.

An additional difficulty for reconstruction is the suffix of the attested forms. Several of the Germanic forms as well as Lat. *caput* seem to show a suffix *-*ut* whereas other Germanic forms show *-*et* and *-*ut*-. PGM. **hafulan-* attests *-*ulo-* or perhaps *-*utlo-*. While Boutkan (1998: 111) remained uncertain, van Sluis (fthc.) adduces the Celtic forms that show *-*ut* and *-*uk* suffixes to this root.²³⁸ Schrijver (1997: 295) proposes that, instead of a series of suffixes *-*ut-*, *-*uk-*, *-*ul-*, this represents a lexeme **kapu-* with suffixes *-*t-*, *-*k-*, *-*l-*. PGM. **haubeda-* with the suffix *-*et-* then looks particularly irregular in an already non-IE paradigm. DV (91) offers a slightly different interpretation (Italic, Germanic: **kap-ut-*; Celtic: **kapu-k-*, Germanic **kapu-l-*) to the same root in *capiō* ‘to seize’, interpreting it as a substrate root (s.v. *capiō*). In any case, the dental suffix is difficult to analyze as the inherited particle *-*ut-* (Beekes 1996: 219) or *-*to-* (van Sluis fthc.). Instead, the alternation between **t* and **k* (and the lack of either in **hafulan-* if not from *-*ut-lo-*) is similar to that in Lat. *nux* ~ PGM. **knud* (s.v.) and European bee words (van Sluis fthc.) that are demonstrably of non-inherited origin. Given the irregularities, a substrate origin is likely for the *caput* family as well.

Finally, Skt. *kapāla-* ‘cup, jar, dish; skull’ may be related. Its vowel can only be **a*. EWAia (I: 300) favors a connection with Lat. *capiō* over *caput* and Schrijver (1991: 100-1) argues against a connection with *caput* given the implied direction of semantic

²³⁷ See also Nussbaum (1986: 214), who proposes assimilation to the vocalism of the ‘ear’ word.

²³⁸ Schrijver (1997: 295), followed noncommittally by EDG (658), suggests that Lat. *caucum* and Gk. *κάυκος* ‘cup’, attested quite late, are borrowings from Celtic.

shift. But this assumes that the meaning ‘cup’ is primary for *kapāla-*, which need not be the case. Adducing the Sanskrit word has implications for the time of borrowing, probably requiring an early date.

catulus ‘young of an animal’

Pre-form: **ka/Ht-e/o/ul-o-* | PItal. **kate/o/ulo-*

Comp.: **ka/o/Hd^h-el-*, **kHt-él-* | PGM. **hada/e/ulō-* | MHG *hatele* ‘goat’

**ka/o/Hd^h-n-*, **kHt-n’-* | PGM. **hadnō-* | ON *hadna* ‘young goat’

**ka/Hd^(h)-Vl-* | PCelt. **kadVlot-* | MĪr. *cadla* ‘goat’

?Proto-Berber **āqāq* ‘(she-)goat’, **qayd-* ‘billy-goat’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: animal

Pokorny (534), WH (I: 183), EM (106), DV (89)

Schrijver (1991: 102, 105), Boutkan and Kossmann (1991: 89), Untermann (2000: 376), Weiss (2010a)

Lat. *catulus* ‘young of an animal’ is related to U **katel** (WH I: 183, EM 106, Untermann 2000: 376, Weiss 2010a) ‘puppy’ or at least ‘a sacrificial animal’. If it is Indo-European, then it must go back to a root **kHt-*, but its potential relationship to a group of words meaning ‘goat’ puts its IE origin in doubt.²³⁹

MHG *hatele* ‘goat’ and ON *hadna* ‘young goat’ reconstruct to PGM. **had-* (Kroonen 2013: 214), which is either the Grimm’s Law reflex of **kad^h-* or Verner variant of **kat-* (favored by Schrijver 1991: 102 and not an invalid root structure). MĪr. *cadla* ‘goat’, if old, reconstructs to **kadVlot-*. The voicing alternation behind the Latin reflex of **t*, the reflex of **d^(h)* in Middle Irish, and the potential **d^h* behind Germanic cannot be accounted for from an IE perspective. It seems simple to interpret Lat. *catulus* both formally and semantically as a diminutive, but in light of a similar suffix on MHG *hatele* and MĪr. *cadla* < **kadVl-*, the full root of the substrate word might include this “suffix”. Perhaps this encouraged a semantic shift within Italic from ‘goat’ > ‘young animal’.

Boutkan and Kossmann (1991: 89) link this to Lat. *haedus* < **g^haid-* (s.v.). The voicing/aspiration discrepancies in the reconstructions for *catulus* < **kat-* and *haedus* < **g^haid-* would mirror that of another goat word: **kap-ro-* (Lat. *caper* ‘he-goat’, s.v.) ~ **g^hab^h-ro-*²⁴⁰ (OIr. *gabor* ‘goat’). While an interesting idea, it means accepting that Italic and Germanic attest to doublets of this lexeme, perhaps due to contact with etymologically related substrate dialects at different points in time. This is too

²³⁹ If related to Slavic forms like SCr. *kōt* ‘birthing, litter, breed’, Pol. *kót* (dial.) ‘place where forest animals young’, etc., the *a*-vocalism would not be due to a laryngeal (Schrijver 1991: 102, DV 89), but the forms are too semantically divergent to adduce with certainty.

²⁴⁰ Though as mentioned, other reconstructions are possible: **g^(h)ab^(h)p-*.

speculative to confirm, and thus it is best to keep the two groups separate. Boutkan and Kossmann (1991: 89) further adduce Proto-Berber **āqād* '(she-)goat' and **qayd-* 'billy-goat', two forms which cannot be regularly linked in Berber. Perhaps they were borrowed from the same source as the forms cited.

cēpa 'onion', var. *caepa*, *cēpe* (neut. indecl.)

Pre-form: **keh₁p-* | PItal. **kēpā*

Comp.: *?*ka/h₂p-* | PGk. **kapia-* | Hsch. *κάπια*: τὰ σκόροδα. Κερυνῆται 'onions amongst the Κηρυνηται'

□ Irreg. correspondences ■ Remarkable phonotactics

Semantics: plant, domestic

WH (I: 201-2), EM (114), DV (108)

Meister (1889: 203), Ernout (1965: 130), Furnée (1972: 337), Biville (II: 325)

Biville (II: 325) shows that the Romance forms descend from **ē*, demonstrating that the Latin form with a diphthong is a hypercorrect spelling. While Biville (II: 325) asserts that neut. indecl. *cēpe* is the oldest (as does Ernout 1965: 130), fem. *cēpa* is attested since Naevius (DV 108).

The only comparandum for *cēpa* is Hsch. *κάπια*, 'onion' amongst the Κηρυνηται (WH I: 201-2, EM 113, DV 108). Meister (1889: 203) takes this to refer to Kyrenia in Cyprus, but Biville (II: 325) takes it to mean Achaean Ceryneia. In the case of the former, its identification as Greek seems uncertain. Even in the latter case, a borrowing from Greek (WH I: 201, Furnée 1972: 337, Biville II: 325) requires an the Hesychian hapax to represent unattested (Achaean) Doric neut.pl. **κάπια* for Att-Ion. **κηπια*. A second assumption is that this **κηπια* entered Latin as **cēpia* and was reanalyzed as a plural before being back-formed into singular *cēpe* (cf. also WH I: 201, Furnée 1972: 337), relying on *cēpe* being the earliest form. EM (114) and DV (108) take it as an independent loan from the same unknown source as the Greek, which seems more likely. In that case, the variation in endings between *-a* and *-e* might represent the nativization of a foreign phoneme.

corbis 'basket'

Pre-form: **k(o)rb^(h)/d^{hi}-* | PItal. **korb/f/bi-*

Comp.: *?*gréb^h-ōn-* | PGm. **krebō-* | OHG *korb* 'basket', etc.
*?*kreb-* | PGm. **hrep-* | ON *hrip*, OHG *href* 'basket carried on the back'
*?*ka/Hrb^(h)-* | PCelt. **karbanto-* | OIr. *carpat* '(war) chariot'

■ Irreg. correspondences ■ Remarkable phonotactics

Semantics: tool

Pokorny (1948-9), WH (I: 272-3), EM (142), DV (135)

Kluge (1885: 443), Kuhn (1959: 39), de Vries (1962: 256-7), EIEC (52-3), Derksen (2007: 234), Matasović (2009: 190), Kroonen (2011: 179-82), Loma (2012: 155-8), Zair (2012: 37-8), Kroonen (2013: 303), Derksen (2014 s.v. *kar̃bas*)

Lat. *corbis* ‘basket’ reconstructs most straightforwardly to a pre-form containing rare **b* or an invalid **TeD^h* root structure. A root shape **skrb^h-* would potentially be allowed, but neither Latin nor any of its potential comparanda provide a trace of an initial **s*.

Secure comparanda of Lat. *corbis* ‘basket’ are difficult to verify. Several Baltic and Slavic forms that can be reconstructed to **korb^h-* (Lith. *kar̃bas* ‘basket’, Ru. *kórob* ‘box, basket’, Cz. *krabuše* ‘wicker basket’, Sln. *kraba* ‘box’, etc., cf. DV 135, Derksen 2014 s.v. *kar̃bas*) could be loans from Germanic (EIEC 52, Derksen 2007: 234).²⁴¹ But that the Germanic forms are loans from Latin (EIEC 52, Derksen 2007: 234) is made highly implausible by the variation amongst the Germanic forms (OHG *korb*, MHG *krebe*, *krepp*, *korb(e)* ‘basket’, Efris. *krääf*, *kräawe* ‘trough, crib’, etc.), which points to an ablauting *n*-stem **krebō*, **kurpaz* < **gréb^h-ōn-*, **grb^h-n-ós*, ruling out a loan after Proto-Germanic (Kroonen 2011: 179-82, 2013: 303, but cf. already Kluge 1885: 443).²⁴² A few Germanic forms reconstruct to a root PGm. **hrep-* < **kreb-* (ON *hrip* ‘pannier’, OHG *href* ‘basket for carrying on the back’; de Vries 1962: 256-7, EIEC 52). Kuhn (1959: 39) took the alternation as pointing to a late entrance into Germanic (as though peri-Grimm’s law). Kroonen (2011: 181-2) notes that the alternation would point specifically to non-IE origin, but that the meaning ‘basket’ (and thus the semantic connection to *corbis*) for the **krebō-* words can be argued to be secondary. ON *kerf*, *kjarf* means ‘bundle’ for instance. The primary meaning of the **hrep-* root is likewise difficult to establish. Further connections with Greek forms have been proposed (Gk. γῤῖπος ‘fishing basket, creel’ and γῤῖφος ‘riddle, (as adj.) obscure’, cf. Pokorny 385-90; κάρφος ‘small dry stick’, cf. EIEC 52-3), but are semantically and/or formally more aberrant (DV 135, EDG 286, Kroonen 2011: 181).

OIr. *carpat* ‘(war) chariot’ < PCelt. **karbanto-* is formally the most similar to Lat. *corbis*. Matasović (2009: 190) considers it likely to be of non-IE origin due to the *a*-vocalism of the root²⁴³ and the same problematic **TeD^h* root structure.²⁴⁴ If Celtic

²⁴¹ Smoczyński (2018: 408) considers Lith. *gur̃bas* ‘basket woven of wicker or straw’ and other Baltic forms of this shape “a var. of *kur̃bas* with voicing of the initial consonant”, providing other cases where this has occurred in loans from Polish. It is unclear whether this strengthens the case for the Baltic words being loans from Germanic or weakens it, but it certainly does not strengthen the case for a native IE origin. Loma (2012: 155-8) notes that foreign (*T*or*T* in early loans does not seem to undergo liquid metathesis should not have undergone the liquid metathesis in e.g. Sln. *kraba*. But his reconstruction of a PIE compound **(s)kor-b^hiH-* ‘removed bark’ is semantically unconvincing and relies on the exclusion of any comparison with Germanic forms.

²⁴² In fact, he considers Lat. *corbis* more likely a loan from Germanic. This would remove the need to reconstruct *corbis* to an illegal root structure, and would make it a pre-literary loan from Germanic. But there seem to be so few of these (cf. Green 1998: 182-200) that it is difficult to accept.

²⁴³ It seems technically possible to reconstruct **karb-* to **kHrb^(h)-* on the same evidence that we can potentially do so for Latin (see fn. 86), as *aRC-* seems to be the normal reflex of initial **HRC-* in Celtic as

**karbanto-* (and thus Lat. *corbis*) are non-IE, then they could attest to an **a* ~ *o* vocalic alternation. But the two words are semantically quite distant from one another.²⁴⁵

In the end, it is quite likely that Lat. *corbis*, in part due to its irregular root structure, is a loan. Several possible comparanda exist, each with their own irregularities (potential consonant alternations within Germanic, non-IE phonotactics in Celtic), but it is unclear if all or any of these are related.²⁴⁶ It seems most likely that the Germanic basket words are related, establishing **k* ~ **g*^h and **b*^(h) ~ **p* alternations as well as aberrant vocalism. But then we face a similar problem to *catulus* (s.v.) in which we must assume that Germanic for some reason attests to a doublet of this non-IE lexeme.

cucurbita ‘(bottle)gourd’

Pre-form: **ku-ko/urb*^(h)/*d*^h-*Vt-* | PItal. **kuko/urb*/*f*/*ɸ**Vtā-*

Comp.: **k*^w*erk*^w*et-* | PGM. **hwerhwetjō-* | OE *hwerhwette* ‘cucumber’, ME *hwerwette*, *werwette* ‘cucumber, gourd’

?Skt. *cirbhaṭī*, *carbhaṭa* ‘cucumber’, *cirbhiṭa* ‘gourd’

□ Irreg. correspondences

■ Remarkable phonotactics

Semantics: plant, domestic

WH (I: 300), EM (154), DV (149)

Kuiper (1948: 143-4), KEWA (I: 378), André (1978: 49-50), EWAia (III: 182), Sebastian, Schaefer, Telford & Renner (2010), Kroonen (2013: 266), Šorgo (2020: 442)

Lat. *cucurbita* ‘gourd’ is compared to Sanskrit forms like *cirbhaṭī*, *carbhaṭa* ‘cucumber’, *cirbhiṭa* ‘gourd’ (cf. WH I: 300), which Kuiper (1948: 143-4) proposes are from a Munda language. Skt. *bhaṭā-* ‘bitter cucumber’ seems to show that the *ci/ar-* element is a prefix (cf. also EWAia III: 182). Thus the Latin and Sanskrit words are not cognate, but the originally Munda lexeme may have reached Latin as a Wanderwort (KEWA I: 378). Latin may have introduced reduplication,²⁴⁷ or it may have entered Latin already reduplicated.²⁴⁸ The similarity to the Sanskrit words could also be a case of chance of resemblance (André 1978: 50, EM 154).

well (Zair 2012: 37-8). But in light of the comparanda, there is no actual reason to do so, and the best explanation is original *a*-vocalism.

²⁴⁴ He himself does not connect Lat. *corbis* because neither **korb*^(h) nor **kɽb*^(h)- can yield the Celtic vocalism. But if they are independent borrowings from a third source, this is exactly what we would expect.

²⁴⁵ Van Sluis (fthc.) further compares OE *hearpe*, OHG *harfa*, etc. ‘harp’ < PGM. **harpōn-* < **ka/orb-*.

²⁴⁶ EM (142) assume that Lat. *corbis* belongs to a group of words for woven objects that must be from a Mediterranean substrate. But the comparanda extend beyond the Mediterranean.

²⁴⁷ WH (I: 300) suggest the influence of *cucumis*. DV (149) notes the words’ similar onset and semantics.

²⁴⁸ André (1978: 50) gives parallels of African languages (as the gourd may have come to Italy from Africa) that reduplicate in lexemes for voluminous things.

A geographically closer comparandum is PGM. **hwerhwetjō-*, albeit only with reflexes in English. It cannot be cognate with *cucurbita* and instead represents an independent borrowing of the same source lexeme (Kroonen 2013: 266, Šorgo 2020: 442). If we assume **hwerhwet-* is from earlier **hwehwert-*, its **t* would show that the *b* of *curcurbita* is a reflex of **d^h* rather than **b^(h)*. Alternatively, **hwerhwet-* might correspond to the *-curbit-* element. Deciding on which interpretation is correct has implications for the irregular alternations to which the pair attests.

excetra ‘sea serpent/monster; Lernean Hydra’

Pre-form: **h₁eksketr-* | PItal. **eksketrā-*

Comp.: **h₁e(k)s(k)etr-* | PSlav. **esetrǔ-* | ORu. *jesetrǔ*, OPol. *jesiotr*, etc.
‘sturgeon’

**h₁eksketr-* | PBalt. **ešketra-* | OPru. *esketres* ‘sturgeon’, Lith. *erškėtas*
‘whale’

?**(k)stur-* | PGM. **stura/ōn-* | ON *styrja*, OE *styria*, *styriga*, etc.
‘sturgeon’

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: animal, wild; aquatic

WH (I: 425-6), EM (205)

Weise (1881: 234), Devoto (1928: 338-41), de Simone (1968-70 II: 189, 276, 287), Breyer (1993: 200-1), Derksen (2007: 145), Kroonen (2013: 488), Šorgo (2020: 459)

Lat. *excetra* sometimes refers specifically to the Lernean Hydra, but in other cases to a (sea) serpent. The murkiness surrounding the term’s semantics beyond mythology poses difficulties for an etymology. Devoto (1928: 338-41) proposed that it represents Etruscan mediation of Gk. ἔχιδνα ‘viper’, an explanation that has been relatively well received (WH I: 425-6, EM 205, de Simone 1968-70 II: 189, 276, Breyer 1993: 200-1). Etruscan sometimes changed Gk. μν to *mr* (*Memrun* < Μέμνων, *Aχmemrun* < Ἀγαμέμνων, de Simone 1968-70 II: 287). But evidence for γν > *cr* is weak,²⁴⁹ and for δν > *tr* practically non-existent.²⁵⁰ This leaves the Etruscan explanation with problems.

WH (II: 425-6 with lit.) reject several etymological attempts to achieve *excetra* via contaminations and folk etymology, but also Weise’s (1881: 234) connection to Balto-Slavic words. However, I think this stands the best chance of being accurate. A relatively robust pre-form for Latin and Baltic would be **eksketr-*.²⁵¹ The exact developments that lead to Baltic *šk* are disputed, but a **k* at least is involved (Derksen

²⁴⁹ Perhaps Lat. *grōma/grūma/croma* ‘surveying instrument’ < Etruscan < Gk. γνῶμη ‘perception, sign’, (de Simone 1968-70 II: 189). But it kept *cn* in *Cnaive* and *Cnare* < Lat. **Gnaivos* and *Gnarus*.

²⁵⁰ Hinted at by the pair of Etruscan names *Tretra* vs. *Treina* (de Simone 1968-70 II: 189).

²⁵¹ Baltic forms with *r* like Lith. *erškėtas* were likely influenced by *erškėtis* ‘thorn’ or represent metathesis.

2007: 145). Thus, despite the Slavic being reconstructible to **esetr̥-* < **h₂ek-* ‘sharp’ (Derksen 2007: 145), such a reconstruction would make them unrelated. Alternatively, Slavic **esetr̥-* could be from **h₁eksetr-* < **h₁eksketr-*.²⁵²

Kroonen (2013: 488) instead connects the Baltic and Slavic forms with PGM. **sturja/ōn-* ‘sturgeon’. He takes the Slavic forms with initial *o-* (cf. Ru. *osētr*) at face value (those that reflect *e*-vocalism can be due to Rozwadowski’s change) to reconstruct PBSL. **asetra-*. Along with Germanic, these would represent an *a*-prefix alternation **astr-* ~ **setr-* with the vocalism “reshuffled”. Given that there are potential examples of the *a*-prefix phenomenon occurring with vowels other than *a* (s.v. *ulmus*), original *e*-vocalism for Balto-Slavic is not problematic and allows the connection of Lat. *excetra*.²⁵³ If PGM. **stur-* is from **kstr-*, the Latin, Balto-Slavic, and Germanic words probably attest to a substrate word with a complex initial cluster.

faber ‘craftsman, smith’

Pre-form: **b^h/d^h/g^{wh}a/Hb/b^h/d^h-ro-* | PItal. **f/p/χ^wab/f/pro-*

Comp.: **d^ha/Hb^h-r-* | PArm. **dabr-(s)na-*²⁵⁴ | Arm. *darbin* ‘smith’

Hurrian *tabiri* ‘metal caster’

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: metallurgy

Pokorny (233-4), WH (I: 436-7), EM (208), DV (197)

Meillet (1894: 165), Kuryłowicz (1956: 194), Mann (1963: 58), Schrijver (1991: 102), Clackson (1994: 36-41), Beekes (1996: 230), Derksen (2007: 109, 110), Martirosyan (2009: 235), Kroonen (2013: 86), Derksen (2014 s.v. *dárbas*), PSD (s.v. *tibira*), Thorsø & Wigman et al. (2023: 120)

Lat. *faber* is traditionally (since Meillet 1894: 165) compared to Arm. *darbin* ‘smith’, PGM. **daban-* ‘to fit’ and PBSL. **doba²* ‘time, manner’ (to which belongs PSlav. **dobr̥* ‘good’, cf. Derksen 2007: 110). LIV (s.v. **d^heHb^h-*) reconstructs **d^heHb^h-*, where zero-grade **d^hHb^h-* > **d^hab^h-* would yield all forms. The Balto-Slavic accentuation rules out the presence of a laryngeal however, leading Derksen (2007: 109) to reconstruct **d^hab^h-* with original *a*-vocalism to account for Lat. *faber* (also Kuryłowicz 1956: 194, who considers it a loanword). Kroonen (2013: 86) instead reconstructs **d^hob^h-* to a root **d^heb^h-* ‘to fit’ for Germano-Balto-Slavic, removing the Latin and Armenian forms from consideration.²⁵⁵ The semantic connection between Germano-Balto-Slavic ‘to fit’,

²⁵² The cluster reductions produce forms that smack of the metathesis attested in the comparanda of *ascia*, *mīlus*, and *viscum* (cf. also in the substrate of Germanic, Šorgo 2020: 459).

²⁵³ Theoretically an original Lat. **axetra* could have been reshaped on analogy with the numerous words beginning with *ex*.

²⁵⁴ This reconstruction rather than in **-īno-* is argued for by Martirosyan (2009: 235).

²⁵⁵ Beekes (1996: 230) argues that the lack of an attested *e*-grade for this root, even in the verbal

‘fitting ∴ good, timely’ and Latin ‘craftsman’, Armenian ‘smith’ was not particularly strong to begin with (cf. EM 208, Schrijver 1991: 102). Keeping them separate thus solves the vocalism of the Germano-Balto-Slavic forms.

Lat. *faber* and Arm. *darbin* still require explanation. Mann (1963: 58) suggested excluding Lat. *faber* and instead connecting Arm. *darbin* with Skt. *ḍrbhāmi* ‘to weave’, Av. *darəv-* ‘to join’, Lith. *dīrbti* ‘to work’, Lith. *dárbas* ‘work’, and Latv. *darbs* ‘work, deed’ < **d^herb^h-*. Derksen (2014 s.v. *dárbas*) shows that the Baltic forms are from a root with a laryngeal **d^hrHb^h-*, which should yield Skt. ***dūr^hb^h-*. Thus the formation only works for Armenian²⁵⁶ and Baltic; a connection with Indo-Iranian would make a more compelling case for an inherited root. But connecting Lat. *faber* and Arm. *darbin* is semantically more attractive than separating them and attaching them to other roots. While they could represent an isolated reflex of a root **d^hHb^h-*, they can be further connected with Hurrian *tab/w-* ‘cast metal’, *taballi* ‘smith’, *ta/ibira/i* ‘copper-worker’²⁵⁷ (Martirosyan 2009: 235, Yakubovich *apud* Blažek 2010: 23, Thorsø & Wigman et al. 2023: 120) as a Wanderwort.

grāmiae ‘eye rheum’

Pre-form: **g^(w)r(e)H-m-* | PItal. **grām-*

Comp.: **g_l-m-* | PGk. **glamo-* | Gk. γλάμων ‘blear-eyed’

?**g^(w)rH-m-* | PSlav. **gr̥měždžb* | RuCS *gr̥b/e/oměžd̥b* ‘pus in the eye’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: body part

Pokorny (405), WH (I: 617), EM (280), DV (270)

Buecheler (1927: 369-70), Lehmann (1986: 279), Schrijver (1991: 487-8), Demiraj (1997: 306), Derksen (2007: 194), Kroonen (2013: 291, 300), EDG (274), Smoczyński (2018: 357), TLL (s.v. *grāma*)

The length of the *a* of Lat. *grāmiae* ‘eye rheum’ is not recorded in any diagnostic context. A line in Plautus’ *Curculio* is potentially crucial: although the manuscripts have *os amarum*, this has been amended to *gramarum* (Buecheler 1927: 369-70, TLL s.v. *grāma*). Found at the beginning of a line of trochaic septenarius, it can only be scanned as *grāmārum*, with a long *ā*. The form *grammō(n)sus* is also found, often interpreted to be an example of ‘expressive gemination’ (WH I: 617, EM 280).

formation preserved in Germanic, makes *a*-vocalism more likely.

²⁵⁶ And even then, only depending on one’s view of the reflex of **CrHC* in Armenian (cf. Clackson 1994: 36-41).

²⁵⁷ It was likely borrowed into Sumerian as *tibira* ‘Metallgießer’ according to Martirosyan (2009: 235) but ‘sculptor’ according to the PSD (s.v. *tibira*).

Comparison to Germanic forms like Go. *qrammiþa* ‘moisture’²⁵⁸ (WH I: 617, EM 280, Schrijver 1991: 487-8, DV 194) is semantically difficult to defend. They probably belong to Balto-Slavic words (Lith. *grĩm̃zti*, Ru. *grjáznut’* ‘to sink into something sticky, boggy’, etc.) not as cognates of *grāmiaē* (cf. WH I: 617) but as an unrelated lexeme (cf. Kroonen 2013: 300).

DV (270) connects Gk. γλάμων, -ωνος ‘blear-eyed’. When it is clearly borrowed from Greek into Latin as *glamae*, it means the same as *grāmiaē* (cf. WH I: 617 with lit.). Thus the semantic match seems quite good, but since neither *grāmiaē* nor γλάμων can be borrowed from the other, they point to an *l ~ r* alternation. Further relatives of Gk. γλάμων are complex and doubtful (EDG 274).²⁵⁹ Thus both it and the *grāmiaē* may both be loans.

Several Slavic forms have a shape and meaning similar to *grāmiaē*. Derksen (2007: 194) reconstructs PSlav. **gr̥mēždžb*, but the attestations within RuCS alone (*gr̥mēždžb*, *gremēždžb*, *gromēždžb* ‘pus in the eye’) alongside several other attestations (SCr. *krmēlj*, *křmēlj*, Sln. *krmēlj*, *krměžalj*, etc. ‘fester in the corners of the eyes’) makes the reconstruction of a single proto-form difficult. One could propose taboo deformation or changes due to child language, but this is of course *ad hoc*. On the other hand, the *g ~ k* alternation suggests repeated borrowing into dialectal Slavic, which could not have occurred until around the second half of the first millennium CE. The exact relationship of the Slavic words to the Latin and Greek forms is difficult to determine.

grūmus ‘heap of earth, hillock’

Pre-form: **gruH(-)m-* | PItal. **grūmo-*

Comp.: **kroH(-)m-* | PGk. **krōmak-* | Hsch. κρῶμαξ· σῶρὸς λίθων, Gk. κρῶμακωτός ‘heap of stones’
**kloH(-)m-* | PGk. **klōmak-* | Gk. κλῶμαξ ‘heap of stones, rock’

Pokorny (385-90), WH (I: 623), EM (283), DV (273)

Alessio (1944a: 124-5), Belardi (1950: 210), EDG (720)

■ Irreg. correspondences □ Remarkable phonotactics

Semantics: geography

Traditional etymologies for Lat. *grūmus* ‘heap of earth, hillock’ include comparison with Gk. γρῦμέα ‘bag or chest for old clothes’ and OE *cruma* ‘crumb’ (WH I: 623 with lit.) or

²⁵⁸ Though suspected of being a misspelling for **krammiþa* because of the rarity of the onset **g^wr-* (Lehmann 1986: 279), this need not be the case.

²⁵⁹ Lith. *glēmės* ‘phlegm, slime’ is probably a neo-full-grade to *glim-/gleim-* (Smoczyński 2018: 357); cherry-picked Albanian dialectal forms (cf. *ngjomë* ‘humid, fresh’) have been compared without consideration of the full variation of the evidence (Demiraj 1997: 306). WH (I: 617) adduce Engl. *clammy* ‘sticky’, but this is probably derived from PGM. **klaīma-* ‘to smear, stick’ < PIE **glei-* (on the root, cf. Kroonen 2013: 291).

a relationship with *gremium* ‘lap, bosom’ (DV 273, since OCS *gramada* < **grōm-* means ‘heap, pile’, but it requires a change the raising of **ō* > *ū* / *_mV*_[back]). None of the proposals is particularly compelling.

A better semantic match is between *grūmus* and Gk. κλωμαξ ‘heap of stones, rock’ and its variants that attest to a non-inherited *l* ~ *r* alternation (Alessio 1944a: 124-5, Belardi 1950: 201, EDG 720). It is not the only example of a Lat. *g-* for a Gk. *κ-* (cf. e.g. *gubernāre*, s.v.), and an *o* ~ *u* alternation occurs between Lat. *cotōneum* and Gk. κοδύ-. Given that the *l* ~ *r* alternation within Greek suggests that the lexeme there is already of non-IE origin, and given that the alternations required for the connection of the Latin and Greek words are paralleled in comparanda of non-IE origin elsewhere, it seems better to compare *grūmus* with κλ/ρῶμαξ than with other words of greater semantic distance.

nāpus ‘turnip’

Pre-form: *(*s*)*neh*₂*p-* / *(*sn*)*Hp-* | PItal. *(*s*)*nāpo-*

Comp.: *(*Si*)*neh*₂*p-* | PGk. *(*Si*)*nāpV-* | Gk. *vāπν*, *σίναπι* ‘mustard’

*(*s(i/u)*)*nipV-* | PArm. *(*s*)*nēpV-* | Arm. *nīw* ‘leaf vegetable’

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: plant, domestic

WH (II: 142-143), EM (429)

Bedrossian (1875-9: 530), Ališan (1895: 101), Hehn & Schrader (1911: 211), Erichsen (1954: 43), André (1956: 297), Mayrhofer (1961: 185-6), Chantraine (1968-80: 735), Lazaryan (1981: 55), (Biville II: 316), EDG (1333)

It is widely agreed that Lat. masc. *nāpus* ‘turnip’ is a loan from Gk. neut. *vāπν* ‘mustard’ (WH II: 142-143, Chantraine 1968-80: 735, EDG 1333, Biville II: 316), but this requires both a change in meaning and in gender. André (1956: 297) suggests that the difference in meaning is due to the original use of both cruciferous plants for their greens, which is plausible. The change in gender is less easy to explain.²⁶⁰ Even its similarity to synonymous *rāpum* ‘turnip’ (alongside of which it is often mentioned)²⁶¹ has not resulted it surfacing as a neuter. Thus EM (429) consider *nāpus* an independent Mediterranean loan.

A further indication that the lexeme is not of IE origin is the alternation between Gk. *vāπν* and *σίναπι*. It has been attributed to an Egyptian source based two pairs: *σίλι* ~ *σέσελι* ‘hartwort’ (said by Pseudo-Dioscorides to be the Egyptian word for *καυκαλίς* ‘hartwort’) and *σάρτι* ‘an Egyptian water plant’ ~ *σίσαρον* ‘parsnip’ (Hehn & Schrader 1911: 211, André 1956: 296, WH II: 143). Mayrhofer (1961: 185-6) disagrees, based on

²⁶⁰ Biville (II: 316) has misunderstood Chantraine (1968-80: 735); there is no masculine doublet *vāπν* attested. Instead there is a late-attested masculine variant of *σίναπι*, namely *σίνηπν* (cf. also EDG 1333).

²⁶¹ Columella (*de Re Rustica* 2.10.23) even mentions that either plant could turn into the other.

the understanding that there is no Egyptian source form. However Erichsen (1954: 43) indeed lists one Demotic attestation of *snwp.t* ‘name of a plant’, linking it to *σίνῶπι*. The final *-t* is likely a feminine suffix, but as we cannot determine the meaning further, the link remains speculative. EDG (1333) rather puts forward a Pre-Greek argument, reconstructing **s^ynāpV-* to explain the disappearing *si-* syllable. However, as the attestations of *σίνῶπι* are later than those of *νῶπι*, they need not have entered Greek at the same time.

Beyond Greek is the potential comparandum Arm. *nīw*. Its modern dialectal meaning is ‘corn salad/mâche (*Valerianella locusta*)’ (Łazaryan 1981: 55), a small leaf vegetable. In Classical Armenian, it is a hapax. Estimates of its semantics vary, with Bedrossian (1875-9: 530) giving ‘wild turnip’ and Ališan (1895: 101) ‘tarragon’. The context in which it occurs describes monks on Mount Tabor in Israel acidifying it with salt to mix with hyssop and drink on a hot day. Again, the semantics may have shifted but remain within the realm of a leafy green vegetable. If indeed related, the Armenian form can reconstruct to **(s)nīp-* or **(s)nēp-*, which WH (II: 143) note produces a non-IE *ā ~ ē* alternation akin to that in *rāpum* etc.

paelex ‘mistress’

Pre-form: **phzeil-a/ek-* / **pehzil-a/ek-* | PItal. **pailek-*

Comp.: **pa/er-ik-* | PCelt. **φα/erikā-* | OIr. *airech* ‘concubine’

**pa/HL-ak-* | PGk. **pallakā* | Gk. *παλλακή* ‘concubine’, *πάλλαξ* ‘young woman’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: economic

WH (II: 233-4), EM (474), DV (439)

Walde (1921: 85-8), Thurneysen (1924: 146-7), Thurneysen (1946: 53-4), Leumann (1977: 69), Levin (1983: 191-7), Matasović (2009: 127), EDG (1147)

DV (439) proposes a derivation of *paelex* as **paed-Vk-s* from the root of *paedor* ‘dirt’, which is difficult formally. Walde (1921: 85-8), followed by Leumann (1977: 69) proposes that Lat. *paelex* was borrowed from an otherwise unattested Gk. **παῖλαξ*, from an earlier **παλιαξ* that would also have produced *πάλλαξ*. Thurneysen (1924: 146-7) adduces MIr. *airech* ‘concubine’, though Matasović (2009: 127) speculates that it is from **peri-* ‘around’, thus **perikeh₂* is ‘a female servant, one that is around’.²⁶² The connection with Av. *pairikā-* ‘witch, demoness’ proposed by Walde (1921: 87-8) seems too semantically far to justify the long-distance link. EDG (1147) dismisses all connections beyond that of *paelex* and *παλλακή*, favoring a connection with Hebr.

²⁶² With **e > a* before a palatal consonant, a phenomenon that is not entirely consistent (Thurneysen 1946: 53-4).

pilegeš, Aram. *palqatā* ‘concubine’, as loans from a Mediterranean language. The Hebrew word has alternatively been considered a loan from Greek (cf. WH II: 234).²⁶³

It is plausible that the Latin, Greek, and Celtic words represent a Mediterranean loan (a Wanderwort according to WH II: 233). Mir. *airech*, with its *r*, is formally the most aberrant, but *r* ~ *l* alternations are not unattested in the Mediterranean (cf. *lilium*, s.v.). The forms further point to an *a* ~ *ai* vocalic alternation and all contain a velar suffix. The independence of the Semitic words remains uncertain.

pannus ‘piece of cloth, rag’

Pre-form: **pa/H-N-* | PItal. **panno-*

Comp.: **pa/o/h₂-no-* / **peh₂-nó-* | PGM. **fanan-* ‘cloth’ | Go. *fana* ‘cloth’, OE *fana*, OHG ‘flag, banner’, etc.

?**peh₂-no-* | PGk. **pāno-* | Gk. *πήνη* ‘the thread of the woof, wound around the bobbin; woof’, Hsch. *πήνος· ὕφασμα* ‘woven robe, web’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: textiles

Pokorny (788), WH (II: 247-8), EM (479), DV (443, 444)

Schrijver (1991: 218-20), LIV2 (s.v. *(s)*penh₁-*, *(s)*pend-*), Weiss (2010b), EDG (1186), Kroonen (2013: 127), Höfler (2017)

Lat. *pannus* ‘piece of cloth, rag’ is close in form and meaning to reflexes of the Proto-Germanic *n*-stem **fanan-*: cf. Go. *fana* ‘cloth’, OE *fana*, OHG *fano* ‘flag, banner’ (WH I: 247-8 with lit., DV 443, Kroonen 2013: 127). Potentially related is Gk. *πήνη* ‘the thread of the woof, wound around the bobbin; woof’ if one trusts the Hesychian gloss *πήνος· ὕφασμα* ‘woven robe, web’ (EDG 1186 are doubtful). Doric forms have *ā*, so the Greek forms reconstruct to **peh₂-n-*. There are formal problems with this set of comparanda. The two nasal consonants in each the Germanic and Latin would have separate explanations. Within Germanic, a paradigm **péh₂-ōn-*, **ph₂-n-ós* seems to have levelled the position of the *n* in the oblique to create a remodeled **ph₂-no-n-* or **peh₂-nó-n-* (with Dybo’s Law, Kroonen 2013: 127). The Latin has a short vowel and geminate consonant, which smacks of the *littera* rule (cf. Kroonen 2013: 127). But this would be one of the only occurrences of this rule involving a nasal (Weiss 2010b). Thus we cannot maintain that *pannus* is a *littera* variant of **pānus* like in Gk. *πήνος* and their correspondence of *a/ā* and *n/nn* is instead irregular, pointing to a loan (DV 443).²⁶⁴ Additionally, if Gk. *πάτος* ‘garment of Hera’ is from **pñ-to-*, then the vocalism of *πήνος*

²⁶³ Levin (1983: 191-7) formulates a narrative in which Hebrew preserves an IE *(*h*)*pi-leg^h-es*, which entered Hebrew along with the institution from the Philistines. This would make (at least) Lat. *paelex* a loan from Semitic.

²⁶⁴ Lat. *pānus* ‘spool with thread; abscess; panicle’ is probably a direct loan from Greek (cf. DV 444).

is not the result of a full-grade root containing a laryngeal.

Lat. *pannus* has also been compared to OCS *ponjava* ‘cloak, dress’ and *opona* ‘curtain’ (WH II: 247-8 with lit., EM 479). These are from a root *(s)*penh₁*- ‘to stretch, weave, spin’ (LIV2 s.v., Kroonen 2013: 127),²⁶⁵ from which PGm. **fanan-* could also descend. But it cannot explain the vocalism of the Greek or Latin forms.²⁶⁶

rādīx, *-īcis* ‘root’

Pre-form: **ur(e)h₂d(h)*- | PItal. **wrādīk-*, **wrādmo-* (Lat. *rāmus* ‘branch, twig’)

Comp.: **ur(e)h₂d-* | PGk. **wrādīk-* | Gk. ῥᾶδίξ, ῥᾶκος ‘branch, twig’

**urh₂d(V)-n₁eh₂-* | PALb. **wradn(i)ā-* | Alb. *rrënjë*, etc. ‘root; oak’

**ure/oh₂d-* | PGm. **wrōt-* | ON *rót* ‘root’, etc.

**urh₂d-i-* / **urd-i-* | PGm. **wurti-* | Go. *waurts* ‘root’, ON *urt*, OE *wyr*t ‘plant, herb’, etc.

**urad-* | PCelt. **wradi-* | MW *gwreidd* ‘root’, OCo. *gwreiten* ‘gl. radix’

**ur(i)d-* | PCelt. **wridā-* | MBret. *gruizyenn* ‘root’, etc.

**urid-ih₂* | PGk. **wridja* | Gk. ῥίζα, Aeol. βρίζα, βρισόδα, Myc. *wi-ri-za* ‘root’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant

Pokorny (1167), WH (II: 415, 416), EM (562-3, 564), DV (512)

Schwyzer (1939-50 I: 344 fn. 2), Schrijver (1991: 182-3), Schrijver (1995: 174), Demiraj (1997: 350-1), Vine (1999b), Matasović (2009: 430) EDG (1108, 1258, 1270, 1271). Kroonen (2013: 597, 601), Weiss (2020: 181), Stifter (fthc.)

Lat. *rādīx* ‘root’ and Gk. ῥᾶδίξ ‘branch, twig’ are formally identical, establishing that the full meaning of this lexeme is both ‘root’ and ‘branch’. Lat. *rāmus* ‘branch, twig’ can represent **ur(e)h₂d-mo-* (DV 513; cf. *caementum* ‘chopped stone, cement’ < **kaid-mentom*, Weiss 2020: 181). If these reflect a root **ureh₂d-*, Alb. *rrënjë* ‘root; oak’ looks like a zero-grade (Demiraj 1997: 351) and Germanic attests to both a zero-grade in **wurti-* and a full-grade in **wrōt-* (Vine 1999b).

Other forms of similar shape and identical meaning complicate the picture. Schrijver

²⁶⁵ There is a very similar root *(s)*pend-* ‘to stretch’ behind Lat. *pendeō* ‘to hang, weigh, pay’ (LIV2 s.v., cf. EM 479). LIV2 calls it a Parallelwurzel. The interchange of **d* and **h₁* is reminiscent of the expected results of the glottalic theory.

²⁶⁶ The reflex of an *n*-stem like **pnh₁-Vn-* should probably have given ***pennus* (cf. Schrijver 1991: 218-20). A *palma* rule development like **pn₁H-no-* is uncertain. The only good examples occur with *r* and *l* (cf. Höfler 2017). An example with *n* could be *antae* ‘pilaster’ if from **anatā-* < **h₂nHt-*, but **h₂nHt-* > **h₂enHt-* (cf. Schrijver 1991: 314) > **ant-* seems reasonably possible and does not require the rule.

(1991: 182-8, 1995: 174) separates OIr. *frén* ‘root’ < **urid-no-* and W gwrysg ‘branch’ < **urid-sko-* from the *rādīx* forms < **ureh₂d-*. MW *gwreidd* ‘root’ could reconstruct to **urh₂d-jo-* (if **CRHT* > *CRāt*, Schrijver 1991: 182-3, DV 512) and thus be related to the *rādīx* forms. Vine (1999b) unites the Celtic forms with the explanation that the **urid-* forms are actually neo-*aniŋ* formations²⁶⁷ from the **urh₂d-* forms, while Matasović (2009: 430) unites them in separating them from the root **ureh₂d-* entirely. He reconstructs **wrid-* (OIr. *frén*, W gwrysg, MBret. *gruizyenn*) and a secondary full-grade **wrad-* (MW *gwreidd*). In this vein, Kroonen (2013: 610) supports a reconstruction of PGm. **wurti-* not as a zero-grade of **ureh₂d-* but as laryngeal-less **urd-*. Matasović’s solution seems the most compelling, especially given the potentially non-IE suffixes attested on PCelt. **wridsko-* (cf. Stifter fthc.) and **wridnā-* (see §3.3.4). Gk. *ρίζα* ‘root’ at face value reconstructs to **wrid-* as well, though Vine (1999b) alternatively suggests this is a morphological zero-grade with *schwa secundum* triggered by the fact that both full- and zero-grades of **ureh₂d-* would have yielded PGk. **wrād-*.

The most straight-forward reconstructions from an IE perspective are a group of words to a root **ureh₂d-* and a group to a root **urid-* (and probably **urad-*).²⁶⁸ But the reflexes of both roots mean ‘root/branch’ and they are formally identical but for their vocalism. It is highly likely that they represent the same lexeme, and the incompatibility of vocalism points to a non-IE origin. It remains peculiar that there are several different reflexes per branch.

raia ‘marine fish, ray’

Pre-form: **H/ura/Hg/i-ieh₂-* | PItal. **ragjā-* / **raijā-*

Comp.: **HruG^h-* | PGm. **rugg-* | MDu. *rogghe*, *rochghe*, Du. *rog*, MLG *rugge* ‘ray’

**HreK-* | PGm. **rehhōn-* | OE (*h*)*reohhe*, ME *rezge*, *reyhhe* ‘ray’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: animal, wild; aquatic

WH (II: 415), EM (563), DV (512)

Kluge & Seebold (1989: 603), Schrijver (1991: 314), Kroonen (2009: 154)

WH (II: 415) reject a connection between Lat. *raia* ‘ray’ and the Germanic words because they can only envision a pre-form like **rgiā-* (an outdated reconstruction that would not yield *raia* anyways). Kluge and Seebold (1989: 603) suggest that the Germanic fish words are related to Ger. *rauh* ‘rough, raw’ because of the texture of the

²⁶⁷ Cf. *-sreth* (PPP of *semaid* ‘to arrange’) as if from **stj-to-* beside *srath* ‘valley’ < **strāto-* < **stj_{h3}-to-*.

²⁶⁸ If the initial omicron of Aeol. *ὀρόδαμνος* means that the word originally started with *ɸ* (EDG 1108, 1270), then Gk. *ῥόδαμνος* ‘branch, twig, shoot’ probably belongs here too. Though interestingly, the *ῥά/ῥό* alternation points to a zero-grade (since in Aeolic this is the regular outcome, cf. Schwyzer 1939-50 I: 344 fn. 2). Vine (1999b) suggested this might also be the result of a *schwa secundum*.

fish, but tentatively relate it to Lat. *raia* with both originating in an unknown language. DV (512) champions the connection between Latin and Germanic. The mismatched forms within Germanic attest to an *e* ~ *u* vocalic alternation along with peculiar gemination from different velars. Kroonen (2009: 154) notes that the attested material makes it difficult to explain this as the result of an ablauting *n*-stem. The *a*-vocalism of Latin could theoretically arise from **HrHg-ǵeh₂-*, as Schrijver (1991: 314) shows that **HRHC* seems to yield *raC-*. But the Germanic forms seem to contradict a reconstruction with a laryngeal. Together, the forms points to a non-IE root **ra/e/uK-* ‘ray’.

sappīnus ‘fir tree’

Pre-form: **sa/HP-* | PItal. **sappīno-*

**sa/HP-* | PRom. **sappo-* | OFr. *sap* ‘fir’

**sa/Hk^w-* | PCelt. **sapo-uidu-* | OCo. *sibuit* ‘fir’

?**sa/Hb^(h)-* | PItal. **sab/fīnā* | Lat. *sabīna* ‘*Juniperus sabina*’

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: plant, tree

WH (II: 478), EM (585, 594)

Walde (1910: 675), Bertoni (1925: 422-23), REW (no. 7592), Alessio (1948-9: 147), Hubschmid (1953: 98-9), Campanile (1974: 95), Delamarre (2003: 267-8), Trask (2008: 258), DV (596), Matasović (2009: 420), Smoczyński (2018: 1124), GPC (s.v. *sybwydd*)

OFr. *sap* ‘fir’ is interpreted by WH (II: 478) as from Gaulish **sapos* < **sak^w-*, the root behind Lat. *sūcus* ‘juice’, Lith. *sakaĩ* ‘resin, pitch’, and OCS *sokъ* ‘juice’. But it must instead be from **sappos* with a geminate like the Latin. REW (no. 7592) and EM (594) recognized this, and favor the idea that Lat. *sappīnus* might be the result of a compound word Gaulish **sappo-* ‘fir’ + Lat. *pīnus* ‘pine’. But this idea seems *ad hoc* and the solution irregular. Its explanatory power might be slightly greater if it accounted for the geminate *pp*, but it does not; the geminate is already there in the pre-form of French. *Sappīnus* could simply be a substantivized *-īno* adjective from **sappus*. OFr. *sap* may be a backformation from *sappīnus* (REW no. 7592, Alessio 1948-9: 147), but given that manuscripts of Pliny have *sappium*, the unsuffixed form may have actually been in circulation.

A form with a single *p* does exist in OCo. *sibuit* glossed as *abiēs* ‘fir’ (potentially a hapax in all of Celtic).²⁶⁹ Taken at face value, it reconstructs to **sapo-widu-* (Delamarre 2003: 267-8), the second element of which seems to be PCelt. **widu-* ‘wood’ (cf. Matasović 2009: 420). Thus the first element would be **sapo-*, the form that WH (II: 478) took to be from **sak^w-*. If it is, it is not related to Lat. *sūcus* ‘juice’ < **se/ouk²-* (DV

²⁶⁹ GPC (s.v.) considers W *sybwydd* a borrowing from Cornish, but Campanile (1974: 95) instead suggests that it is the Cornish form that is borrowed.

596) nor to Lith. *sakaĩ* ‘resin, pitch’ and OCS *sokъ* ‘juice’ < PBSl. **sʷak-a-* < **sʷok^w-* (Smoczyński 2018: 1124). Given its poor attestation in Old Cornish, it could presumably be a loan from Latin. EM (594) indicate that spellings with a singleton were in existence, and if legitimate, one of them could have served as the source of the Celtic; perhaps a form related to Plinian *sappium*. If the Celtic form is independent, it attests to a *p ~ pp* alternation between Latin and Celtic. If it is not, then it helps illustrate a *p ~ pp* alternation within Latin. In either case, it cannot be accounted for in inherited terms.

A further alternation might be attested within Latin in the form of *sabīna* ‘savin juniper (*Juniperus Sabina*)’ (Alessio 1948-9: 147). WH (II: 457) and EM (585) suspect that this word is related to *sa(m)būcus* ‘elder tree’ after the suggestion by Walde (1910: 675, cf. *sambūcus*, s.v.), despite not overwhelmingly agreeing with his arguments. And in fact, the evergreen, coniferous juniper is much more similar to the fir than to the deciduous flower- and berry-producing elder tree. This would establish a *p ~ pp ~ b* alternation similar to that seen in *lepus* ‘rabbit’ (s.v.).

Bertoni (1925: 422-23), then later Hubschmid (1953: 98-9) further compared Basque and Berber oak words. These are namely Basque *sapar* ‘thicket, scrub’ and *txapar* ‘kermes oak’ and Berb. *tasaft* ‘*Quercus ballota*’ (found in Chaouia, Tashelhit, etc.). While Trask (2008: 258) analyzes *txapar* as a diminutive of *lahar* ‘bramble’ from which all forms with *p* (like *sapar*) could be back-formed, ‘oak’ and ‘bramble’ are quite different. *Txapar* formally does look like a diminutive, potentially from a pre-form **tzapar-*. In the end however, since the Basque and Berber words mean ‘oak’, it is difficult to link them semantically with the Latin words meaning ‘fir/pine’.

sulpur ‘sulfur’, vars. *sulphur*, *sulfur*

Pre-form: **su(e/o)lp-(o)r-* | PItal. **so/ulpur-*

Comp.: **su(o)lf-(o)r-* | PRom. **su(l)fur-* | Catal. *sofre*, etc. ‘sulfur’

**sue(l)b^h-lo-* / **sue(l)p-ló-* | PGm. **swe(l)bla-* | Go. *swwibls*, OE *swefl*, OHG *swebal*, etc. ‘sulfur’

□ Irreg. correspondences

■ Remarkable phonotactics

Semantics: geography

Pokorny (1046), WH (II: 628), EM (664-5), DV (598)

Much (1898: 165-6), Brüch (1933), Corominas & Pascual (1984-91 I: 438-9), Breyer (1993: 453), Szemerényi (1995: 410), Kroonen (2013: 497), Šorgo (2020: 450-1)

Since Much (1898: 165-6), Lat. *sulpur* has been compared to MHG *Schwefel* ‘sulfur’, though the details have varied. The earliest attempts (cf. also Brüch 1933) took the proto-form as **sʷelk^w-*²⁷⁰ due to Germanic dialectal words seeming to derive from both

²⁷⁰ The status of the Upper Palatinate form *Schwefel* is debated. Much (1898: 165) and Kroonen (2013: 497, cited as Bavarian) take it as an archaic, undissimilated form from **sʷelplo-* (or **sʷelk^wlo-*) whereas

**swēbla-* (Go. *swibls* etc.) and **swēgla-* (OE *swegel* etc.). This assumes dissimilations within Germanic and a Latin borrowing from a Sabellic language. WH (II: 628) instead take the Germanic forms in -*g-* as recent dissimilations rather than evidence of **g^w* (but do not support uniting the Latin and Germanic forms). The most straightforward account is given by Kroonen (2013: 497), who unites the forms under **sue(l)plo-*, as the Verner variant would yield PGm. **swēbla-*, accounting for all the Germanic daughter forms (assuming unproblematically that the first *l* was lost in most languages through dissimilation and that the *b ~ f* alternation is due to von Bahder's Law²⁷¹).

Thus we can reconstruct a root **suelp-*: full *e*-grade **suelp-* for Germanic and zero-grade,²⁷² *e*-grade, or *o*-grade for Latin. The problem remaining is the *p ~ ph ~ f* alternations in (at least) spelling in Latin. WH (II: 628) explain the *ph* as a learned Hellenized spelling, with *f* being a 'bad' spelling. But the Romance languages reflect *f* in some forms (Sp. *azufre*, Port. *enxofre*,²⁷³ Catal. *sofre*), showing that it was more than a spelling variant (Corominas & Pascual 1984-91 I: 438-9). WH (II: 628 with lit.) suspect a Mediterranean word while EM (665) suspect Etruscan (cf. also Breyer 1993: 453). But as for *ferrum* (s.v.), there is very weak evidence for Etruscan being responsible for *p ~ f* alternations. Thus the *p ~ f* discrepancy of the Italic forms is a true alternation. Non-initial *f* for Latin cannot be reconstructed, and it generally points to loans from the Sabellic treatment of the voiced aspirates. Thus *sulfur* could represent a Sabellicism < **su(e/o)lb^h-(o)r*. Then a reconstruction of **sue(l)b^h-lo-* for the Germanic forms (i.e. not a Verner variant of the shape underlying *sulpur*) is also not out of the question.

Far-reaching comparanda from languages to the east (cf. Kroonen 2013: 497) are likely unrelated (cf. Šorgo 2020: 450-1), and potential IE cognates that mean 'fat, oil' (Szemerényi 1995: 410, DV 598) are semantically unattractive in light of the close semantic match with Germanic. Thus only the Italic and Germanic material can be compared with certainty. Within Italic, the Romance forms and some Latin attestations create a *p ~ f* alternation that is difficult to explain from an inherited perspective.

tamarix 'tamarisk'

Pre-form: **ta/Hm-ar-ik-* | PItal. **tamarik-*

Comp.: **mur-ik-* | PGk. **murikā-* | Gk. μωρίκη 'tamarisk'

Brüch (1933: 73) is suspicious that the original form would be maintained in only one dialect and suspects that the form in question is the result of contamination between *Schwefel* and Upper Palatinate *Schwell* 'rheinisches Gold'. The exact details do not seem to matter, as Brüch still assumes that the proto-form was **suēlk^wlós*.

²⁷¹ Described in von Bahder (1903), cf. also de Vaan (2014).

²⁷² Even following the argument that **sulp-* would have been realized as **su_lp-*, the result **suolp-* is the same as the *o*-grade.

²⁷³ Corominas & Pascual (1984-91 I: 439) write that the initial vowel of some of the forms does *not* require transmission via Arabic. The Arabic word in use for sulfur was *kibrīt* (cf. Catal. *alcrebite*).

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, tree

WH (II: 646), EM (676)

Lewy (1895: 44), Solmsen (1901: 14-15), Schuchardt (1918: 16), REW (no. 5360, 8548), Bertoldi (1937: 145), Alessio (1941b: 207), Hubschmid (1953: 81), EDG (981), Weiss (2020: 128-9)

Lat. *tamarix* looks at first to have avoided undergoing vowel weakening, but the second *a* could have been preserved via the *alacer* rule (cf. Weiss 2020: 128-9). The Greek word for tamarisk is *μυρίκη*,²⁷⁴ such that Lat. *tamarix* looks like the same root with a *ta-* prefix. Lewy (1895: 44) suggested this phenomenon had its source in Semitic, with *μυρίκη* from e.g. Hebr. *mārar* ‘to be bitter’ and *tamarix* being from a form like Hebr. *tamrūrīm* ‘bitternesses’ referencing the bitterness of tamarisk bark used in medicinal preparations. EDG (981) follows his comparison between Hsch. *μυρίκη· δυσώδης* ‘stinking’ and Aram. *mōrīqā* ‘crocus’, but crocuses are not particularly bad-smelling. Schuchardt (1918: 16) instead interprets *ta-* as the Berber feminine prefix, but there is no Berber comparandum for this word.²⁷⁵ WH (II: 646) note that the variant Lat. *tamarīcē* looks like it is taken from a Greek pre-form, but no such form is attested.

Romance descendants provide more information. Hubschmid (1953: 81) takes Apulian *támaro* ‘bushy shrub’ to represent a variant of the lexeme without the velar suffix (see §3.3.3). Most important is the variant *tamariscus*. WH (II: 646) and EM (676) consider it dubious in Classical Latin, but REW (no. 8548) notes that it underlies Romance forms like It. *tamarisco* and Prov. *tamarisc*. It would be a later variant, but seems to attest to a *cs/sc* metathesis or at least the appearance of a sigmatic element before the final consonant of **tamarik-*. Alessio (1941b: 207) identifies this with the **mariscus* element in several Romance forms for a type of rush that descend from **mariscus juncus* (Piedmontese, Lombardy *maresk*, *marask* ‘swampy land’, Lombard *brisk* ‘rushes’, Berrichon *març* ‘rushes for thatching the roof’, cf. REW no. 5360). The comparison would be better if the semantics were closer.

In the end, it seems difficult to separate Gk. *μυρίκη* from Lat. *tamarix* (with potentially original variant *tamariscus* showing *SK* metathesis, cf. §3.2.1.2.8.3) due to their identical meaning. The initial syllable of the Latin word is from an unidentified source.

tilia ‘linden tree’

²⁷⁴ The oldest attestations of Gk. *μυρίκη* (Homeric) attest to both *ī* and *ī̄*, suggesting that metrical lengthening has changed an original *-ikē* ending (Solmsen 1901: 14-15).

²⁷⁵ This recalls the case of Lat. *buda* ‘cattail (*Typha* spp.)’, which seems to have originated in African Latin and spread throughout Romance, and which is difficult to separate from Berber forms like Kabyle *tabuda* ‘*Typha angustifolia*’ with the feminine article (Schuchardt 1918: 16, Bertoldi 1937: 145, Hubschmid 1953: 26-7). Portuguese is alone amongst the Romance languages in having *taboa*, the form with the article attached.

Pre-form: *(p)te/il- | PItal. *te/ilia-

Comp.: *(p)tel- | PArm. *(p)tel- | Arm. *ʿeli* ‘elm’

*ptel- / tpel- | Gk. *ptel- | Myc. *pte-re-wa*, Gk. *πελέα* ‘elm tree’

?*h₂pel- | PGk. *apel- | Hsch. ἀπελλόν· αἴγειρος, ὃ ἐστὶν εἶδος δένδρου
‘black poplar’

?*p(t)el- / (t)pel- | PGm. *felwō- | OHG *felwa, felawa* ‘willow’

*h₂eptlV- | PCelt. *axtl/nV- | MBret. *ezlen* ‘aspen’, W *aethnen* ‘aspen,
poplar’, OCo. *aidlen* ‘abiēs’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, tree

Pokorny (847), WH (II: 340), EM (522), DV (480, 620)

Ernault (1895), Henry (1900), Bathe (1955), Hamp (1984), EWA (III: 132), Kluge and Seebold (1989: 525), Blažek (2003: 6), Deshayes (2003: 223), Gliwa (2008), Martirosyan (2009: 284), EDG (115, 1247), Meiser (2010: 81), Kroonen (2013: 140, 136), Schrijver (2015), GPC (s.v. *aethnen*), Šorgo (2020: 456-457), Matasović (fthc.)

The best comparandum for Lat. *tilia* ‘linden tree’ is Gk. *πελέα* ‘elm tree’. It could in fact be a borrowing from Greek (cf. DV 620) via the process that raises *e to i before i in the following syllable (cf. Meiser 2010: 81) after vowel weakening. The disparate semantics suggest that they are independent of one another. A variant *πελέα* of Gk. *πελέα* ‘elm tree’ suggests a pattern similar to *πτόλεμος*~*πόλεμος* and *πτόλις*~*πόλις*. As PIE **tpersneh₂*- yields Gk. *πέτρνη* ‘heel’, both **ptel-*/**tpel-* are possible reconstructions for *πελέα*. But as **tpersneh₂*- yields Lat. *perna* ‘heel’ (cf. DV 460), only **ptel-* could yield *tilia*. This could be further evidence for Latin having borrowed from Greek (cf. *tisana* ‘pearl barley’ < Gk. *πιτσάνη* ‘id.’). The status of Arm. *ʿeli* ‘elm’—that is whether it represents a loan from Greek or an independent attestation—is disputed (cf. Martirosyan 2009: 284), but if it is indeed independent, it shows **ʿel-* < **ptel-* (cf. *ʿakčim* ‘hide’ : Gk. *πήσσω* ‘to cower’). Kroonen (2013: 136) adduces PGm. **felwō-* ‘willow’, which could be from **tpel-* (cf. **fersnō-* ‘heel’ < **tpērs-neh₂*, cf. Kroonen 2013: 137) or **ptel-* (cf. **farna-* ‘fern’ < **ptorH-no-*, cf. Kroonen 2013: 129), though alternative etymologies exist (cf. EWA III: 132). There is a further possibility that Celtic comparanda exist. W *aethnen*, MBret. *ezlen* ‘poplar’, and OCo. *aidnen* gl. *abiēs* (cf. GPC s.v. *aethnen*) together allow for a reconstruction to PCelt. **axtl/nV-* followed by a feminine singular suffix.²⁷⁶ An *a-*prefixed form **a-ptlV-* would regularly yield PCelt. **axtlV-*, which in turn regularly yields MBret. *ezlen*, suggesting that the forms with *n*, i.e. W *aethnen*, Corn. *aidnen*, are innovations.

²⁷⁶ MoBret. *evl* ‘poplars’ (singularive *evlenn*) is the continuation of Old Southwest British *hob-aebl* borrowed from **ebulum* < Lat. *ebulus* ‘?elderberry’, with semantics contaminated by the *ezlen* word (Deshayes 2003: 223, Schrijver 2015) and therefore does not represent original variation.

The semantics of this group are relatively disparate, but remain within the realm of trees.

WH (II: 340) and EM (522) further compare *pōpulus* ‘poplar’.²⁷⁷ A closer semantic connection is surely Hsch. ἀπελλόν ‘black poplar’ (EM 522, EDG 115). Since the Latin word looks reduplicated, one could reconstruct **h₂pel-* for Greek against **po-h₂pel-* for Latin (EDG 115 hesitantly),²⁷⁸ but this looks similar to the **h₂pis-* rejected as the preform of ἄπιον ‘pear’ (cf. *pirum*, s.v.). Matasović (fthc.) compares *pōpulus* to PSlav. **tōpolb* ‘poplar’ < **ta/op-ol-*. Lith. *tūopa* ‘poplar’ suggests a root **toHp-* or **tōp-*, the vocalism of which can also be reconstructed for Latin *pōpulus*.²⁷⁹ This leads him to further suggest that *pōpulus* derived via assimilation from **tōp-*. WH (II: 340) had proposed the opposite development, but Matasović notes this would require independent dissimilation in both Slavic and Baltic against one assimilation in Latin. The latter is thus more likely.

The most conservative account would be to consider two separate lexemes of non-IE origin: 1) Lat. *tília* ‘linden tree’ and Arm. *t’eli* ‘elm’ as loans or independent comparanda of Gk. πετέλα ‘elm tree’ and PCelt. **axtl/nV-* < **aptl/nV-* and 2) Lat. *pōpulus* ‘poplar’ alongside PSlav. **ta/op-ol-* and Baltic **tōp-*. But comparanda from both groups could potentially be united under a substrate root PTL/TPL,²⁸⁰ with shifting vocalism perhaps due to the accentually conditioned pattern identified by Šorgo (2020: 456-7) for the Germanic substrate: CVCVC (PSlav. **ta/opol-*, Pre-PItal. **tōpol-*) ~ aCCC (Pre-PCelt. **aptl-*) ~ aCCVC (Hsch. ἀπελλόν if < **ἀπελλόν*) ~ CCVC (Pre-PGk., Pre-PGm. **ptel-/*tpel-*, PItal. **ptel-*, PArm. **ptel-*).

trabs ‘tree trunk, beam’

Pre-form: **trab-* | PItal. **trab-*

Comp.: **trēb-* | PItal. *trēb-* | Osc. **trībúm** [acc.sg.] ‘house’, **tríbarakavúm** [inf.] ‘to build’, etc.

**treb-* | PItal. **treb-ī/ē/ĭe-* | U *trebeit* [3.sg.pr.] ‘lives, dwells’, etc.

**treb-* | PCelt. **trebā-* | OIr. *treb* ‘settlement’, MW *tref* ‘town’, etc.

**trb-o-* | PGM. **þurpa-* | Go. *þaurp* ‘farmland’, ON *þorp* ‘isolated settlement’, OE *þorp* ‘crowd’, Ger. *Dorf* ‘village’

**trob-* | PBalt. **trōb-* | Lith. *trobà* ‘cottage, farmhouse’, Latv. *trāba*

²⁷⁷ Ger. dial. *Vielbaum* might be a remnant of the lexeme in Germanic before *Pappel* was borrowed from Latin (Bathe 1955, Kluge & Seebold 1989: 525), but it is unclear if it really belongs here.

²⁷⁸ Blažek (2003: 6) interprets *pōpulus* (as a backformation from adjectival *pōpulus*?) as a reduplicated formation to a stem **-pġno-* found in Gk. πάλλω ‘to swing’. The stem would also occur, prefixed with **sm-* in ἀπελλόν < **sm-pelno-*. But the reduplicated syllable *pō-* is not explained by this.

²⁷⁹ Gliwa (2008) concludes that Lith. *tūopa* cannot be an inherited word due to the discrepancies with the Slavic evidence. But rather than this meaning he thinks it is a substrate word, he thinks it is not a genuinely Lithuanian word. He finds its limited attestation, the absence of most poplar species from Lithuania, and the existence of other words for aspen that derive from ‘to tremble’ to be suspicious.

²⁸⁰ In this case, the PGM. **felwō-* ‘willow’ could belong to this root, from **ipel-* (Kroonen 2013: 136).

‘hut, hovel

□ Irreg. correspondences

■ Remarkable phonotactics

Semantics: plant / architecture

Pokorny (1090), WH (II: 696-7), EM (698), DV (626)

von Planta (1892-7 II: 1 fn. 2), Ernout (1946: 29), Schrijver (1991: 481-2), Weiss (1993: 75-89), Untermann (2000: 765-6), Matasović (2009: 388), EDG (1467), Kroonen (2013: 553), Derksen (2014 s.v. *tropa*), Weiss (2020: 168, 322)

Sabellic attests to a verbal root **treb-* ‘to dwell’, with a root noun **trēb-* ‘house’ with cognates in PCelt. **trebā-* ‘settlement’ (*e*-grade), PGM. **purpa-* ‘crowd, settlement’ (zero-grade), and PBalt. **trōb-* (*o*-grade lengthened by Winter’s Law). It is furthermore generally well agreed that Latin *trabs* ‘beam, tree trunk’ belongs here (WH II: 696-7, EM 698, Schrijver 1991: 481-2, Derksen 2014 s.v. *tropa*). However von Planta (1892-7 II: 1 fn. 2) already doubted the connection. It is conspicuously absent from Matasović (2009: 388) and Kroonen (2013: 553).

The problem is both formal and semantic. While the rest of the comparanda show IE ablaut, Lat. *trabs* requires *a*-vocalism, a neo-zero-grade (Weiss 1993: 77), or an explanation by which original **trēb-*, **trb-/trob-* ablaut was replaced with *ē/a* ‘ablaut’ found in verbal paradigms (Schrijver 1991: 482). Additionally, while all the other comparanda refer to dwellings, settlements, and communities, Lat. *trabs* sometimes refers to part of a building: the beam. EM (698) notes that *taberna*, which is generally derived via dissimilation from **traberna* (cf. Weiss 1993: 75-6 fn. 3, 2020: 168),²⁸¹ since it means ‘tavern, hut’, would suggest that *trabs* was indeed part of the dwelling word family. The ending *-erna* is often taken to be a hallmark of Etruscan origin (Ernout 1946: 29) but there are clearly cases in which it has been added to an Italic (or at least inherited) root (Weiss 2020: 322 lists e.g. *caverna* ‘a hollow’ to *cavus* ‘hollow’ and *lucerna* ‘oil lamp’ to **leuk-* ‘to shine’) making it non-diagnostic in this case.

DV (626) followed by Derksen (2014 s.v. *tropa*) suggests that the difficulty in reconstructing a single pre-form, the **b*, and the European distribution indicates that this might not be an inherited root. Another option is a connection—originally made by e.g. WH (II: 696-7) from an inherited perspective and rejected by Schrijver (1991: 482) because it cannot be explained from an inherited perspective—with Gk. *τέραμνα/τέρεμνα* [nom.pl.] ‘house, residence’ (cf. also Weiss 1993: 83-5). EDG (1467) reconstructs a pre-form **terh₂b-no-* and agree with Furnée (1972: 40, 351) in comparing *θεράπνη* ‘servant, maid; house, residence’ as evidence of a non-native origin of these words. Untermann (2000: 766) reconstructs **terh₂-mno-* which would technically work as well. He writes that Lat. *trabs* is impossible to connect to the Sabellic forms phonetically and is too distant semantically; thus it is best connected with the Greek

²⁸¹ Although the same dissimilation does not occur in e.g. *fraternus*.

words.²⁸² The Latin form is indeed the most aberrant of the non-Greek comparanda, but seeing that all the comparanda in question *including* the Greek forms can be reconstructed with **b*, a chance remains that they are all remnants of a non-IE lexeme.

ulmus ‘elm tree’

Pre-form: **h₁el(i)mo-* / **Ho/ul(i)mo-* | PItal. **e/ol(i)mo-*

Comp.: **h₁elmo-* | PGm. **elma-* | Engl. *elm*
**h₁olmo-* | PGm. **alma-* | ON *almr* ‘elm’

**(h₁)limo-* / **(h₁)lemo-* / **h₁lmo-* | PCelt. **limo-* / **lemo-* | MIr. *lem*
 ‘elm’

**(h₁)leim-* | PCelt. **lēmā-* | W *llwyf* [pl.] ‘elm’

**(h₁)limo-* / **h₁lmo-* | PSlav. **jьlьm-* | Russ. *il'm* ‘elm’

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: plant, tree

Pokorny (302-4), WH (II: 811-12), EM (744), DV (637)

Schrijver (1991: 66), Schrijver (1997: 311), Derksen (2007: 211), Matasović (2009: 237)

The Germanic forms **elma-* and **alma-* ‘elm’ show an initial full vowel, so the root, if Indo-European, must be reconstructed with an initial laryngeal to avoid an invalid vowel-initial root structure. Lat. *ulmus* can be reconstructed as **He/o/ulmo-*, but a zero-grade from a laryngeal-initial root would probably have given ***almus* (Schrijver 1991: 66). Matasović (2009: 237) argues that the Italic and Germanic forms can be syncopated from **h₁e/olimo-*, which would match the Celtic preform from a root **h₁lim-*, but this creates a disyllabic root. DV (637) skeptically derives Celtic and Slavic (cf. for the latter also Derksen 2007: 211) from a zero-grade **h₁lmo-* which, along with PGm. **elmo-* and **almo-* would require the reconstruction of three different ablaut grades, full *e*-grade, full *o*-grade, and zero-grade of the root. This does not fit into any known accentual paradigm. It also requires the Brythonic form to have secondarily developed **leimo-*. Schrijver (1997: 311) experiments with reconstructing **lemo-* for the pre-form of MIr. *lem*, which would require unconditioned Scheweablaout to arrive at PCelt. **Hlem-* vs. PItal./PGm. **Helm-* or an interpretation of the final **-m-* as an ablauting suffix producing **He/ol-m-* vs. **Hl-em-*. Neither of these explanations can accommodate the Brythonic form however. In the end, Schrijver (1997: 311) proposes that this is a case of *a*-prefixation, albeit with a vowel other than *a*, in which Latin and Germanic attest to **o/e-lm-* and Celtic to **lVm-*. Slavic might show **i-lm-* (cf. DV 637). The difficulty in reconstructing a pre-form that follows PIE rules without requiring extra assumptions

²⁸² The formatting makes it look like he attributes this to von Planta (1892-7 II: 1 fn. 2), but von Planta simply writes that the connection of Lat. *trabs* to the Sabellic forms “scheint nicht ganz zweifellos,” so I assume the vehement rejection is Untermann’s own opinion.

favors considering this family of words to be of non-IE origin.

vaccīnium ‘hyacinth, whortleberry’

Pre-form: **ya/HK-* | PItal. **wakkīnio-*

Comp.: **ya/hzk-* | PGk. **wakint^ho-* | Gk. ὑάκινθος ‘hyacinth’

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: plant, berry

WH (II: 722), EM (710)

Meillet (1908: 162), FEW (14: 106), REW (no. 9111), Deroy (1956a: 188), Furnée (1972: 242, 377), Sommer & Pfister (1977: 91), Vander Kloet (1992), EDG (1523), Matasović (2009: 27), Magni (2017), Weiss (2020: 308), Kroonen (fthc.)

Vergil uses Lat. *vaccīnium* to translate Gk. ὑάκινθος ‘blue/purple flower, probably the hyacinth’ used by Theocritus (EM: 710), but in its other attestations (Vander Kloet 1992) and the Romance languages (cf. FEW 14: 106)²⁸³ it refers to the whortleberry/bilberry. The Latin and Greek words were proposed to be independent loans from a Mediterranean language by Meillet (1908: 162). The Greek form has the textbook pre-Greek suffix -vθ- (EDG 1523) and its ὑά- (along with perhaps Cretan inscriptional Βάκινθος ‘name of a month’ and ἑάκινθος [inscription from Argos] ‘name of a Laconic festival’) lead Furnée (1972: 242) to suggest a pre-form **u-wa-*. EDG (1523) rejects the proposal of a prothetic *u-*, and thus must be suggesting that ὑά- was a spelling for the continued pronunciation of *wa-* after the loss of digamma. This does not seem like the only option.

The semantic match between Latin and Greek is not perfect, though both refer to blue clusters, the Latin of berries and the Greek of flowers. Alternatives have included a borrowing from Greek with the geminate introduced due to contamination from *vacca* ‘cow’ and *vaccīnus* ‘bovine’ (WH II: 722, Sommer & Pfister 1977: 91)²⁸⁴ and a derivation from *bacca* ‘berry’ with a substitution of *v* for *b* due to their converging pronunciation (Vander Kloet 1992). The former idea seems to have little to recommend it; Latin usually preserves the -vθος suffix in words it borrows from Greek (*absinthium*, *acanthus*, *calaminthe*, *plinthus*, *terebinthus*) and it is difficult to see what whortleberries have to do with cows. The latter idea seems better, but it cannot be ruled out that *vaccīnium*’s relationship to *bacca* (actually *bāca*, s.v.), itself probably non-IE, is deeper (cf. Deroy 1956a: 188). Thus *vaccīnium*, *bāca*, and Gk. ὑάκινθος could be borrowings from the same non-IE source.²⁸⁵

²⁸³ REW (no. 9111) gives only Sursilvan *muschin*, but cf. MFr. *vassine*, *baciet*, *vaciet*, etc.

²⁸⁴ Kroonen (fthc. with lit.) considers *vaccīnium* to derive wholesale from *vacca* ‘cow’, but the semantic motivation does not seem strong enough.

²⁸⁵ Jahowkyan (1987: 310) compares Arm. *vaz* ‘vine branch’ to Lat. *bāca* ‘berry’, suggesting a reconstruction **u/ibaǵ^h* (Rasmus Thorsø, p.c.). A form **ubaǵ^h* with its prothetic **u-* + labial is reminiscent of the **u-yaak-* Furnée (1972: 242) proposes as the pre-form of Gk. ὑάκινθος. It is also

As to the suffix of *vaccinium*, it looks on the surface like a combination of adjectival *-īno-* and *-īo-*. But this may be coincidental. In the strictest interpretation, *-īno-* forms a genitival relationship with the base (Magni 2017, Weiss 2020: 308). Thus **vaccīno-* (if the base is that of *bāca* ‘berry’) would not mean ‘like a grape/berry in shape or color’ but rather ‘of the grape/berry’. Instead, one might compare the whole suffix **-īnyo-* to that of PCelt. **agrīnyo-* ‘sloe, fruit of the blackthorn’ (cf. Matasović 2009: 27, Kroonen fthc.).

viscum ‘mistletoe; birdlime’

Pre-form: **uisk-o-* | PItal. **wisko-*

Comp.: **uiks-o-* | PGk. **wikso-* | Gk. ἰξός ‘mistletoe; birdlime, sticky substance’

**ueiks-* | PGM. **wīhsilō-* | OHG *wīhsela* ‘sour cherry’

**uei(k)s-i-* | PSlav. **višb-* | Ru. *višnja* ‘cherry’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, wild

Pokorny (1134), WH (II: 802-3), EM (741), DV (683)

Cuny (1910: 160), Chantraine (1968-80: 465), EDG (593)

Lat. *viscum* and Gk. ἰξός both refer to mistletoe and the sticky birdlime that is produced from its berries. Because birdlime is also made from the sap of the cherry tree, OHG *wīhsela* ‘sour cherry’ and Russ. *višnja* ‘cherry’ are often compared to the pair (WH II: 802-3, EDG 593, DV 683; not Chantraine 1968-80: 465, EM 741). Then one wonders what the original lexeme would have referred to: the plant or the sticky product. Otherwise, given the toxicity of the European mistletoe (*Viscum album*), the *Benennungsmotiv* of potent berries could have been extended to the sour cherry. WH (II: 803) argue that the later meanings of *viscidus* ‘bitter, pungent; powerful, concentrated’ demonstrate a link with *vīrus* ‘slimy liquid; venom, poison’ (so too does Pokorny 1134), but this is formally impossible.

All forms would reconstruct to an ablauting root **ueiks-* but for the unexpected metathesis in Latin and the fact that it does not otherwise belong to any known IE root. In fact, just such a metathesis occurs in non-IE *ascia* ‘axe’ (s.v., cf. already Cuny 1910: 160). What looks like metathesis might be an original non-IE root shape **wiKsk-* (DV 683). EDG (593) also questions IE origin. Given the parallel to another word of non-IE origin, it is more attractive to consider *viscum* of similarly non-native origin rather than inventing a PIE root **ueiks-* ‘plant with potent berries or sticky sap’.

distantly reminiscent of the shape of Lat. *ūva* ‘grape’, whose appurtenance to *e-* and zero-grade formations of a root **HeiH-u-* ‘yew’, an infamously toxic plant, is not bulletproof.

2.2.3 Comparanda only in Latin and Romance

arbutus ‘strawberry tree’

Pre-form: **H(e)rb(h)/d^h-u-to-* | PItal. **arbuto-*

Comp.: **h₂erm-ōn-* | PRom. **armōn-* | Genovese *armön*, *armún* ‘strawberry tree’

■ Irreg. correspondences □ Remarkable phonotactics

Semantics: plant, tree; fruit

WH (I: 62), EM (43)

REW (no. 610), Alessio (1941b: 188-90), Bertoldi (1942: 174), Battisti & Alessio (1950-57 I: 108, 294), FEW (XXV: 91)

Lat. *arbutus* has no good comparanda outside of Italic (WH I: 62, EM 43) but its reflexes in the Romance languages are irregular. While generally restricted to Tuscany, Corsica, and the Iberian peninsula, an adjectival derivation *arbutus* is also found in France (Battisti & Alessio 1950-57 I: 108, FEW XXV: 91). Tuscan *àlbatro* and *àrabatro* are from *arbutus*,²⁸⁶ but the Ligurian dialects attest to a *b ~ m* alternation. Lunigianese *armótolì* (and *ramótolì*, *marmótolì*) are from **armutulus* and Genovese *armön*, *armún* are from **armō*, *-ōnis* (Alessio 1941b: 189, Battisti & Alessio 1950-57 I: 294).²⁸⁷ The alternation within the Romance forms cannot be accounted for from an inherited perspective. Along with the lexeme’s restricted distribution, it suggests Mediterranean substrate origin (Alessio 1941b: 188-90, Bertoldi 1942: 174, EM 43).

cerrus ‘turkey oak’

Pre-form: **kerr/so-* | PItal. **ker/so-*

Comp.: **ga/HR-* | PRom. **garr-* | Prov. *garric* ‘oak’
 **ka/HR-* | PRom. **karr-* | Catal. *carrasca* ‘holm oak’
 **ka/Hr-* | PRom. **kar-* | It. dial. *cariglio* ‘turkey oak’

■ Irreg. correspondences □ Remarkable phonotactics

Semantics: plant, tree

WH (I: 207), EM (116)

²⁸⁶ Further from *arbutus* are Sp. *álborto* and Pt. *érvodo*. It. *arbitro*, Asturian *albédro*, Galician *érbedo*, and (derived) Corsican *arbitronu* are from **arbitus* (REW no. 610).

²⁸⁷ Alessio (1941b: 188-90) notes that the lexemes also attest an alternation between a *-to ending and a *-no ending, proposing that in this case the *-to ending forms a type of collective (cf. *arbos* ‘tree’ : *arbutum* ‘orchard’, *filix* ‘fern’ : *filictum* ‘fernbrake’, *laurus* ‘laurel’ : *laurētum* ‘laurel grove’, *pomum* ‘fruit tree’ : *pomētum* ‘place planted with fruit trees’, etc.) which might mean it is not the inherited adjectival suffix. Weiss (2020: 313-14, including fn. 48) offers an alternative, native explanation, taking the full suffix *-ētum* as the original participle of statives.

Schuchardt (1918: 18-19), FEW (II: 408-12), Bertoldi (1933a: 287), Alessio (1935), Alessio (1936), Alessio (1941: 179), Hubschmid (1953: 93-7), Hubschmid (1960: 37, 41)

Lat. *cerrus* ‘turkey oak’ has been compared to several Berber words like *akarruš* that mean ‘(evergreen) oak’ and Arabic forms like *qerrūš* along with a plethora of Romance forms of the shape *karr/garr*, especially from Iberia and southern France (Schuchardt 1918: 18-19, WH I: 207, Hubschmid 1953: 93-7, 1960: 41). Hubschmid notes that the -š rules out a Berber loan from Latin and proposes a Eurafrikan substrate with *a ~ e* vocalic alternation. But the Berber forms need not be independent loans from a substrate; instead they can be loans from Arabic < Romance (Maarten Kossmann, p.c.).

The Romance forms on their own indeed attest to *g ~ k* and *r ~ rr* alternations: cf. e.g. Prov. *garric* ‘oak’, Catal. *carrasca* ‘holm oak’, Port. *carrasca* ‘species of olive, heater, holm oak’,²⁸⁸ It. dial. *cariglio* ‘turkey oak’, Calabrian *carrigliu* ‘turkey oak’, etc. Basque *arta-karro* ‘type of oak’ is a compound of *arta* ‘oak’ and apparently this word (Hubschmid 1953: 93-7). It seems that only Lat. *cerrus* shows *e*-vocalism against *a*-vocalism everywhere else. The alternations are still not able to be accounted for in any regular way.²⁸⁹

genesta, var. *genista* ‘broom (plant)’

Pre-form: **gen-es-to-*, **gen-is-to-* | PItal. **genestā*, **genistā*

Comp.: ?

□ Irreg. correspondences ■ Remarkable phonotactics

Semantics: plant, wild

WH (I: 591), EM (270)

Sommer (1900: 336), Lehmann (1907: 391), Herbig (1917), Bertoldi (1937b: 167), Bertoldi (1942: 196), Alessio (1937: 258), Alessio (1944a: 102), Alessio (1948-9: 116), Hubschmid (1953: 29), Hubschmid (1958: 214), Battisti (1959: 327-31), Hubschmid (1960b: 145), Breyer (1993: 100-2), DV (23), van Sluis (ftnc.)

Lat. *genesta* occurs alongside *genista*, and the *e ~ i* alternation between them is without explanation. They are otherwise isolated to Italic.²⁹⁰ Romance descendants attest to further vocalic irregularity (cf. It. *ginestra*, Calabrian *yinsotra*, REW no. 3773, WH I: 591, EM 270). Sommer (1900: 336) explained the differing vocalism as contamination from *arista* ‘awn, head of grain’, but since this also occurs as *aresta*, it only moves the

²⁸⁸ Note the wide range of meanings that includes oak. In general, amongst the dozens of forms Hubschmid cites, there is a relatively wide range of arboreal semantics, often verging on scrubland plants.

²⁸⁹ While the FEW (II: 408-12) includes these words under an entry on **carra* ‘stone’ (with the same non-IE pre-forms **gar(r)a-* / **kar(r)a-*) based on the idea that Basque *haritz* ‘oak’ might be a derivation of *harri* ‘stone’ and due to the German parallel *Steineiche* (Bertoldi 1933a: 287, Alessio 1935, 1936, 1941: 179), Hubschmid (1953: 97, 1960: 37) wisely keeps them separate.

²⁹⁰ And not a derivation from *genū* ‘knee’ (pace Lehmann 1907: 391).

problem to a different lexeme.²⁹¹

Herbig (1917) argued that the word is Etruscan, on Isidore's information that *lanista* (cf. var. *lanistra* and derived forms) 'trainer of gladiators' is Etruscan.²⁹² He further proposed Etruscan origin or mediation for e.g. *lepista* (vars. *lepesta*, *lepistra*) 'goblet', *arista/aresta*, and *fenestra* 'window'. But Breyer (1993: 100-2) argues that such suffixes, where they occur in Etruscan, are coincidental conglomerations of other morphemes.²⁹³ *Lanista*, if indeed of Etruscan origin, has perhaps received the Greek ending -ίτης (Breyer 1993: 240, cf. *lanius* 'butcher' without the suffix). Thus Etruscan is not the source of these suffixes in Latin. Instead, it is often considered Mediterranean (Bertoldi 1942: 196, Alessio 1944a: 102; 1948-9: 116;²⁹⁴ Battisti 1959: 196). The best example in non-onomastic material is probably Sard. *golostru* etc. 'holly', widely attested (though not in Latin²⁹⁵) and of demonstrable non-IE origin (cf. Bertoldi 1937b: 167; Hubschmid 1953: 29, 1958: 214, 1960b: 145; recently van Sluis fthc.).

Despite being isolated, the inner-Latin *e ~ i* alternation (along with the suffix) make *genesta* quite likely to be of non-inherited origin.

lābrusca 'the wild grapevine'

Pre-form: **la/(e)Hb^(h)/d^h/s-r-* | PItal. **lāb/f|prūscā*

Comp.: **la/(e)Hmb^(h)/d^h/s-r-* | PItal. **lāmb/f|prūscā* | It. *lambrusca* 'wild grape'
etc.

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: viticulture

WH (I: 740-1), EM (334-5)

REW (no. 4814, 8281), Schwyzer (1934: 242), Alessio (1941b: 215-18), Bertoldi (1942: 171), Battisti (1960: 367, 371), FEW (V: 108-9), Furnée (1972: 272), Breyer (1993: 405)

The vowel length of the first *a* in Lat. *lābruscum* 'wild grape' is uncertain, but the *u* is given as short by REW (no. 4814) and EM (334). The ancient grammarians thought it was related to *labrum* 'lip, edge' in the sense that it grew at the edges of the fields. This smacks of a folk etymology, but the base could be a Latin word given the same suffix in

²⁹¹ The variant *aresta* is found only in a few glosses but is widespread in Romance descendants (REW no. 648). Given its poorer attestation in Latin, I do not treat it separately. See further WH (I: 67), EM (46). As to its comparanda, some compare its "root" to *arinca* 'a kind of grain' with a Ligurian or Mediterranean suffix (Alessio 1944a: 104-5, 1948-9: 113; Battisti 1960: 353-4) but Hubschmid (1960b: 175) prefers a connection to Basque (*h*)*ari* 'thread, spun plant fiber'. Neither seems convincing.

²⁹² Isid. 10.159: *lanista gladiator* i.e. *carnifex Tusca lingua*.

²⁹³ Further, the *-sta/-stra* alternation of some forms might have its roots in Vulgar Latin developments (Breyer 1993: 100-2 with lit.).

²⁹⁴ A development from Alessio (1937: 258), where he too considered it Etruscan.

²⁹⁵ In fact, Latin has *aquifolium* from earlier *ācrifolium* 'sharp leaf' (DV 23), which is a conspicuous neologism in the face of the widespread opaque word.

asinusca ‘a grape of little value’ < *asinus* (EM 334); cf. further the ampelonym *atrusca* ‘a kind of grape’, presumably a derivation of *āter* ‘black’. Numerous Romance forms attest to **lāmbriūscā-* with an additional nasal (with **ū*: MFr. *lambrusce*, Marche *lambrusca*, Piacenza *lāmbriūska*, etc.; with **ū*: Lyon *lambrochi*, Piedmontese *lambrosca*, etc., REW no. 4814, FEW V: 108-9). Since the alternation with *m* is not attested in Latin, it is not certain that it is original. WH (I: 740) and FEW (V: 108-9) consider it secondary, with the latter noting that the same phenomenon occurred with Lat. *strabus* ‘squinting, crooked (of eyes)’ < Gk. στραβός, Lat. *strambus* ‘bow-legged’ occurs in glosses and is the only form continued by the Romance languages (It. *strambo* ‘strange, contorted’, Sp. *zambo* ‘bow-legged’, Romanian *strâmb* ‘crooked’, etc., cf. REW no. 8281). Schwyzer (1934: 242), after discussing cases where late Gk. μβ has developed from original *ββ, suggests that *strambus* might reflect *στραββός with expressive geminate alternation. No such explanation can be given for **lambrusca*, as it is does not have a Greek pre-form. Thus it may represent the *b ~ mb* alternation of *sambūcus ~ sabūcus* (Alessio 1941b: 215-18; Bertoldi 1942: 171; Battisti 1960: 367, 371; Furnée 1972: 272, EM 334²⁹⁶). Alessio (1941b: 215-16) sees behind **labr-* the Mediterranean substrate word **lapa*/**laba* and compares it to Lat. *lapis* ‘stone’ (s.v.).²⁹⁷ His justification, that *labrusca* is the ‘vite della rupi’ is too imaginative to be secure. *La(m)brusca*’s deeper etymological origins remain opaque.

While often compared to *laburnum* ‘broad-leaved beantrefoil’ with varying degrees of certainty (cf. WH I: 740-1, EM 334-5) due to the similarity of the element *lab-*, there is no compelling semantic reason to link them.²⁹⁸

lepus, -oris ‘hare’

Pre-form: **lep-os-* | PItal. **lepos-*,

Comp.: **la/Hpp-Vr-* | PRom. **lapparo-* | Fr. *lepereau* ‘bunny’

■ Irreg. correspondences □ Remarkable phonotactics

Semantics: animal, wild

WH (I: 775, 783), EM (346, 351), DV (335)

Körting (1908: 231), Bruch (1914: 351-70), Schrader & Nehring (1917-23: 442), Bertoldi (1937a: 146), Hubschmid (1943), Alessio (1944a: 101), FEW (V: 175-7), Carnoy (1955b: 597-600), Wagner (1960-4 II: 22-3), Furnée (1972: 231), Trask (2008: 173), Weiss (2020: 163)

The inherited reflex of Lat. *lepus* in the Romance languages is e.g. Fr. *lievre*, It. *lepre*,

²⁹⁶ Some of these scholars also place emphasis on the fact that this is a viticultural word, but we do not need the semantics to suspect a non-IE word here.

²⁹⁷ Alessio (1944a: 104) seems to suggest that the suffix *-usco* here, like *-asco*, is Ligurian. But I am skeptical of morphological claims like these that are often based on toponyms.

²⁹⁸ Even if related, that they are of Etruscan origin due to the ending of *laburnum* is without good evidence (cf. Breyer 1993: 405).

Sp. *liebre* all meaning ‘hare’. The source of Fr. *lapin* ‘rabbit’, *lapereau* ‘bunny’ and Pt. *laparo* ‘rabbit’ is a different version with *a*-vocalism and gemination, which is likely a sister rather than a daughter of Lat. *lepus* (cf. Hubschmid 1943, FEW V: 175-7). Sardinian dialects (*lèppore*, *lèppere*, *lèppuri*, *lèppuri*, *lèppiri*) attest three preforms **leppore*, **leppere*, and **leppure* (cf. Wagner 1960-4 II: 22-3) whose gemination shows they are also not inherited from Latin. The Romance and Sardinian evidence corroborate forms in the writings of classical authors, which otherwise might not have much credence. Strabo writes λεβηρίς ‘rabbit’, later labeled as Massiliot by Erotianus. Pliny writes gives *lauricēs* (sg. presumably **laurex*) ‘rabbit fetuses’ as Balearic (WH I: 775, EM 346; cf. further Bertoldi 1937a: 146, Alessio 1944a: 101).²⁹⁹ Beyond gemination, these forms attest to a labial alternation *p* ~ *b* ~ *w* (Furnée 1972: 231).

Benveniste rejects that *lepus* is an old *s*-stem.³⁰⁰ In fact, outside of Latin, the lexeme always has an *r*. Non-neuter polysyllabic *s*-stems generally undergo levelling to *r* in the nominative (Weiss 2020: 163), which *lepus* [masc.] has not done. On the other hand, seeing as the rest of the comparanda do not have an *s*, this is not a retention but rather an analogical production. (Strangely then, the analogy is with the neuter *s*-stems that do not undergo the levelling).

Semantic explanations have been plentiful and imaginative. At least thrice, the family of words has been explained as meaning “the one with hanging ears” (Körting 1908: 231 from Germanic *lapp-*; Brüch 1914: 351-70 as original IE Ligurian words via the roots **leg*^w- and **lep-* + **ausro-*; Carnoy 1955b: 597-600 from an Indo-European substrate). FEW (V: 175-7) suggested a derivation from **lappa-* ‘stone slab’ (cf. *lapis*, s.v.). It has been called Iberian (WH I: 783), Lybico-Iberian (Bertoldi 1937a: 146), Mediterranean (EM 351), and Ligurian (Brüch 1914: 351-70).

In the end, we can conclude that all of the independent forms taken together seem to show that the whole root was disyllabic, non-IE **IVBvr-*. Strangely enough, the European hare (*Lepus europaeus*) is widespread in Europe and its native range includes the Pontic steppe. It is the European rabbit (*Oryctolagus cuniculus*), Lat. *cuniculus* (s.v.), that was foreign to Europe outside of Iberia (EIEC 258). It is thus curious why a foreign word was applied to the native species.³⁰¹

sambūcus ‘elder(berry/flower)’

Pre-form: **sa/Hmb*(^h)- | PItal. **samb/fūko-* | Lat. *sambūcus*
 **sa/Hb*(^h)- | PItal. **sab/fūko-* | Lat. *sabūcus*

²⁹⁹ Varro writes that Siculi, Acolis, and Graeci called *lepus* λέπων, but some have taken the apparent rhotacism to mean that this is a loan from Latin (EM 352, WH I: 786). The *r* may be original rather than the result of rhotacism, but this word is still so close to the Latin oblique form that I am wary of using it as independent evidence.

³⁰⁰ Apud WH (I: 786) cited as BSL 33: 53f., but I can find no such article in that volume.

³⁰¹ The forms outside Latin suggest that this word also originally meant ‘rabbit’. Perhaps it displaced *cānus* < **kHs-no-* (only ‘white, hoary, gray’ in Latin but with the additional meaning ‘hare’ in Celtic, Germanic, Baltic, and Indo-Iranian) via some sort of taboo (cf. Schrader & Nehring 1917-23: 442).

Comp.: ?

□ Irreg. correspondences

■ Remarkable phonotactics

Semantics: plant, tree; fruit

WH (II: 473), EM (592)

Cuny (1910: 158), Walde (1910: 675), Peterson (1914: 142-3), Brüch (1922: 232-41), Schwyzer (1934: 242-3), Knobloch (1955), Haas (1959: 35), Hester (1965: 364), Ahd. Wb. (I 1968: 1478), Furnée (1972: 272, 347), Puhvel (X: 106-7), EWA (II: 417), EDG (563), Simon (fthc.)

Walde (1910: 675) tried to connect *sabūcus* to *faex sabīna* ‘strong-smelling oil’ and *sabīna* ‘type of juniper’ via **sab-*, a root variant of **sap-* (cf. *sapīō* ‘to taste, perceive’), attributing the nasal of *sambūcus* to analogy with the Greek loan *sambūca* ‘stringed instrument.’ Brüch (1922: 232-41) instead began with Dioscorides’ account that for elder, the Romans say σαμβούκουμ, the Gauls σκοβύην, and the Dacians σέβα. He proposed that a PIE **(s)keb-* entered Latin as loan via a *satəm* Daco-Thracian reflex **sab-* ~ **sam-* (with the understanding that Daco-Thracian had an internal *b* ~ *m* alternation) or directly via a zero-grade **skbūko-* followed by loss of the *k* (like in the **skt* of *pastus* < *pask-tos*) and anaptyxis. These analyses rely on heavy speculation about developments within poorly understood Dacian. Furnée (1972: 347) takes the Lat. *a* vs. Dacian *e* at face value.

Walde (1910: 675) had alternatively suggested a borrowing from Gk. σάμψ(ο)υχον ‘marjoram’ with dissimilatory loss of the second sibilant. Brüch (1922: 237) found this unlikely because it does not explain the forms without the nasal, but thought that contamination with it (or its borrowed Latin form *sampsūchum*) could have led to the introduction of the *m* into original *sabūcus*. WH (II: 473) are unwilling to believe that contamination would occur from a word with such a different meaning. Cuny (1910: 158) instead suggested they are independent reflexes of a third source form. Haas (1959: 35) proposed a PIE **som-b^h(o)uǵ-* (cf. Ru. *buzina* ‘elder’ etc.³⁰²) entered IE Pre-Greek with ‘lautverschobenem’ **k*, whence it was borrowed into Latin before undergoing the Pre-Greek change **b^hu-* > ψυ-. Hester (1965: 364) notes that the normal spelling in Greek is with -ουχ-, outside the environment of the proposed change (though Haas considers that the change also occurred before diphthongs with *u*). The connection with the Greek word is semantically very weak. The connection with PSlav. **bъzъ-* ‘elderberry’ is stronger, but requires proposing a **k* ~ **ǵ^h* alternation and an analysis of the Latin word as *sam-būcus*. A similar analysis that takes *-būcus* to be a

³⁰² Peterson (1914: 142-3) had earlier rejected a connection between PSlav. **bъzъ-* and Lat. *sa(m)būcus* due to phonological difficulties. Pogodin (apud Peterson) had suggested that Slavic forms with additional initial *cha-* and *che-* elements would correspond to Lat. *sabūcus* and *sambūcus*, each with irregularities in the system (PSlav. **a* vs. **ъ*, the *m* in Latin). Peterson (1914: 143) instead proposes that this initial element is from another lexeme (PSlav. **xъbъtъ* ‘dwarf elder’) and the irregularity is due to folk etymological contamination.

separate element, is to compare it to OHG *buggila* (cf. EM 592) ‘mugwort (artemisia),’ but this is riddled with semantic problems³⁰³ and still does not explain the *sam-* element.

Secure comparanda for the Latin forms are therefore unknown, but the vacillating nasal element between *sambūcus* and *sabūcus* are irregular within Latin itself. WH (II: 473) note that the *m ~ mb* alternation is similar to cases like Gk. θύβρις / θύμβρις ‘savory (plant),’ which also looks non-IE (cf. also EDG 563). The likely explanation is that these alternating forms within Latin point to a non-native origin for this lexeme as well. The additional variants *sabuncus* and *sabbūcus*³⁰⁴ appear in glosses and, if trustworthy, suggest that this is one of the later loans into Latin.

Sometimes Hitt. *sampukki-* ‘a pot-dish (ingredient)’ as a “typical culinary culture word” is adduced (Knobloch 1955: 5-10, Puhvel X: 106-7, Simon fthc.), but its meaning is far too poorly understood to connect it with certainty.³⁰⁵

talpa ‘mole’

Pre-form: **ta/Hlp-* | PItal. **talpā*

Comp.: **da/Hrb^(h)-* | PRom. **darbo(n)-* | OProv. *darbon*, etc. ‘mole’

■ Irreg. correspondences □ Remarkable phonotactics

Semantics: animal, wild

WH (I: 324, II: 644), EM (164, 675), DV (605)

REW (no. 2473), Bertoldi (1931: 149-52), Alessio (1939: 328), FEW (III: 13-14), Hubschmid (1963-5 I: 14)

Lat. *talpa* ‘mole’ has resisted etymological analysis (WH II: 644, EM 675, DV 605), but Bertoldi (1931: 149-52, later Alessio 1939: 328, Hubschmid 1963-5 I: 14) convincingly compares it to Romance forms that attest to a voiced version of the stops and an *r* for the *l*, alternations that occur in other substrate lexemes. The word already occurs in Polemius Silvius as *darpus* ‘a four-footed animal’, which WH (I: 324) explain as underlyingly **darbus* with a *p* for *b* on the influence of *talpa*. Inherited *talpa* underlies e.g. Fr. *taupe*. The alternate **darbo-* (cf. REW no. 2473) underlies several forms including OProv. *darbon*, dauphinois *darbon*, *drabon*, *zarbō*, *darbō*, *žarbō*, etc. ‘mole’ but also Draguignan *darbou* ‘rat’, Tarn *darboun* ‘shrew’, etc. It is restricted to Frainc-Comtou,

³⁰³ OHG *buggila* refers to species of *Artemisia*. Ahd. Wb. (I 1968: 1478) suggests that OHG *buchil(e)*, *puchil* is the same lexeme, but EWA (II: 417) is cautious because these refer to ‘water hemlock’, an entirely different plant. None of the words has a secure etymology. The potential attraction of *puchil* is that it is once given as a gloss of *sambuca*. But the Prudentius passage cited in the gloss (*et varios iubet obmutescere cantus, organa, sambucas, citharas calamosque tubasque*) is clearly about musical instruments. Thus Ahd. Wb. (I 1968: 1478) has to assume that *sambūca* (the musical instrument) was somehow mistakenly given for *sambūcus* (elder) in the gloss.

³⁰⁴ Schwyzer (1934: 242-3) proposes explaining the variation in forms via metathesis or substitution of borrowed (expressive) gemination.

³⁰⁵ It could be taken into consideration, however, that elderberries are poisonous unless cooked.

Franco-Provençal, and eastern Provençal until just across the Rhone, leading FEW (III: 13-14) to propose that it is of Ligurian or Gaulish origin; in any case a pre-Latin language.

2.3 Origin Unclear

2.3.1 No Comparanda

acinus ‘berry, esp. grape’

Pre-form: **h₂ek-ino-* | PItal. *akino-*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: viticulture

WH (I: 8-9), EM (6-7), DV (23)

LS (s.v. *acinus*)

WH (I: 8-9 with lit.) and EM (6-7) reject proposed cognates and suspect that Lat. *acinus* is from a Mediterranean language based on its viticultural semantics. But on the grounds of an additional meaning ‘grape seed,’ DV (23) proposes a derivation from **h₂ek-* ‘sharp’ due to the bitter taste of grape seeds. Without comparanda, there is no way to support the claim of substrate origin. But the inherited explanation does not seem fully satisfactory. (LS s.v. *acinus* cite Cicero *de Senectute* 15.52 as an example of the words use in the meaning ‘grape seed’, but it occurs in the collocation *ex acini vinaceo* in which it is *vinaceo* that means ‘grape seed’.)

ās, assis ‘copper coin < *‘bronze plaque of one pound in weight’

Pre-form: **h₂ed-ti-* | PItal. **assi-*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: metallurgy / economic

WH (I: 71), EM (50), DV (57)

von Planta (1892-7 I: 295), Ernout (1954: 106), Breyer (1993: 123), Vine (2016: 324)

Lat. *ās* is without comparanda. There is no evidence that it is of Etruscan origin (*pace* Ernout 1954: 106, EM 50).³⁰⁶ Perhaps the most promising etymology is by von Planta

³⁰⁶ No similar word is attested in Etruscan, the assumption that semantically similar *libra* is also Etruscan is wrong (it is probably inherited, s.v. *libra*), and Etruscan numerals attest to a decimal system (Breyer 1993: 123), not a duodecimal system.

(1892-7 I: 295), followed in part by WH (I: 71) from something like **ad-ti-*,³⁰⁷ given as meaning ‘solidified’ as though to the PIE root **h₂ed-* ‘to dry out’. Semantically, this is difficult to verify.

autumnus ‘autumn’

Pre-form: **h₂eut-* | PItal. **auto/umno-*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: cosmology

WH (I: 87-8), EM (61), DV (19, 64)

Ernout (1946: 34), Breyer (1993: 411-14 with lit.), Rix (1997 with lit.), Schaffner (2014)

WH (I: 87-8) reject etymological proposals like a derivation from the root of Lat. *augeō* ‘to grow’ or OIr. *ócht* ‘cold’ because they rely on the likely folk etymological manuscript spellings with *auct-*. Etruscan origin is often claimed or considered (Ernout 1946: 34, WH I: 88, EM 61, DV 64), but Breyer (1993: 411-14 with lit.) shows that the arguments are problematic.³⁰⁸ Rix (1997 with lit.) likewise rejects all previous proposals and proposes a preform **au-tom-ino-* from **h₂ep-* + **temh₁-* ‘cutting away’. His argument that *au-* is not simply a conditioned variant of *ab* is difficult to believe however (cf. DV 19) and requires the preservation of otherwise unattested very archaic semantics and morphemes. Schaffner (2014) revives Schrader-Nehring’s connection with PGM. **auda-* ‘riches, wealth’. But this relies on the reconstruction **h₂eu-tó-* as opposed to other alternations (like **Heu-d^hh₁-o-*, allowing a connection with Lat. *über* ‘rich, abundant’ < **Hou(H)d^hri-*, cf. Kroonen 2013: 40). Thus, to my mind, Lat. *autumnus* remains without comparanda.

balteus ‘belt’

Pre-form: **ba/Hlt-* | PItal. **baltejo-*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: textiles

WH (I: 95), EM (65)

³⁰⁷ The nominative is usually given as *ās*, proposed to be from **ass* (DV 57, Vine 2016: 324), but the lengthening of the vowel seems difficult to explain. Lachmann’s Law should produce from **h₂ed-ti-* > **āss* > *ās*, but then the oblique should be ***āsis*.

³⁰⁸ That *-mno-* is a suffix of Etruscan origin is difficult to confirm. Many Etruscan forms in *-mma* and *-mne* can however be interpreted as derivations in *-na* to a stem ending in *m* (Breyer 1993: 68). No similar word is attested in Etruscan besides perhaps *avil* ‘year’, but this leaves the rest of the form unexplained.

Pfiffig (1969: 37), Bonfante (1985: 203), Breyer (1993: 428-9), Rix (2009: 145-6)

Charisius in his *Ars Grammatica* (I 77.9) says that Varro gives *balteus* as a *Tuscum vocabulum*, and so it is generally accepted as borrowed from Etruscan (WH I: 95, EM 65, Breyer 1993: 428-9). Bonfante (1985: 203) specifically removes it from consideration due to the *b*, as he considers Etruscan to have had strictly no voiced plosives. Indeed, where Latin speakers reflected Etruscan names with voiced consonants, they are word internal (see fn. 204). The assumption that *balteus* is of Etruscan origin without any attested Etruscan forms of the word is already problematic on its own. Lat. *balteus* remains without good comparanda.

cicōnia ‘stork’

Pre-form: **(ki-)kōn-* | PItal. **kikōniā-*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: animal, bird

Pokorny (525-6), WH (I: 212), EM (119), DV (113)

Walde (1910: 123), Niedermann (1919: 80-1, fn. 1), Alessio (1943: 234), André (1978: 30), Breyer (1993: 244-5), TLL (s.v. *cicōnia*)

Lat. *cicōnia* also occurs as Praenestine *cōnea* in Plautus. Niedermann (1919: 80) derives the latter via hapology, while DV (113) proposes onomatopoeic reduplication. A connection with *canō* ‘to sing’ is semantically questionable (storks do not sing, André 1978: 30) and relies on comparison with OHG *huon* ‘hen’ (Walde 1910: 123) < **koh₂n-* (DV 113), but the latter may instead be a secondary *vrddhi*-derivative of PGm. **hanan-* < **kh₂n-on-* (Kroonen 2013: 207, 240). Thurneysen (in the TLL, s.v. *cicōnia*) noted a similarity between *cicōnia* and Etruscan words like *cicu* and *cicunia*. Further similar forms attested in Etruscan include *cicui* and *cicusa* (Breyer 1993: 245), but none of them is of known meaning. A Hesychius gloss gives γνίς as the “Tyrrhenian” word for ‘stork’, but this undoubtedly simply means Italic (cf. fn. 339), and at best can be taken as another example of the un-reduplicated lemma. While Etr. *cicunia* looks like an exact match for *cicōnia*, Breyer (1993: 245) notes that, within Etruscan, this would be a feminine formation to masculine *cicu*. Thus, either Latin borrowed a less frequent, derived word for ‘female stork’ from Etruscan or, more likely, the Etruscan word is a borrowing from Latin. Claims of Mediterranean origin based on reduplication and similarity to *cicāda* (Niedermann 1919: 80-1, fn. 1, Alessio 1943: 234, WH I: 123, EM 119) are without comparanda (either for *cicōnia* or *cicāda*) to verify them.

ferula ‘giant fennel’

Pre-form: **b^h/d^h/g^{wh}es-* | PItal. **fese/olū*

Comp.: ?

- Irreg. correspondences □ Remarkable phonotactics

Semantics: plant, wild / tool

WH (I: 487, 489), EM (230, 231), DV (214, 216)

Alessio (1941b: 197-203), Battisti (1959: 154, 156-7)

Lat. *ferula* ‘giant fennel’ has no convincing external comparanda. Based on the idea that the giant fennel was named in part after its hollow stalks, it is widely accepted that *ferula* from a root **fes-* is related to *festūca* ‘stalk, straw; ram, pile-driver’ (WH I: 487, 489; EM 230, 231; DV 214, 216). But this gets us no closer to an internal etymology. Alessio (1941b: 197-203) took the root as **fis-* due to the change that may have made **si-so* into *serō*, adducing several other words: *fistula* ‘reed, tube’, *fistūca* ‘pile-driver’ (perhaps a variant of *festūca*), and *fiscus* ‘woven basket’. But the *i > e* change is probably not regular (cf. *pirum*, s.v. for discussion). It would point instead to an irregular *e ~ i* alternation, but it is not clear that the words belong together semantically.

It is even unclear if *ferula* and *festūca* belong together. The latter has the suffix *-ūca* found in e.g. μούτουκα ‘thyme, Cistus’,³⁰⁹ *sambūcus* ‘elderberry’, *lactūca* ‘lettuce’, etc., suggesting that the stem is **fest-* as opposed to the **fes-* of *ferula* (DV 216).³¹⁰ Alessio (1941b: 197-203) interprets the *-st-* as an Etruscan suffix further found in *are/ista* ‘awn, ear of grain’ and *gene/ista* ‘broom (plant)’. But this would further separate **fes-* from **fe-st-*. The internal analysis of *ferula* and its potential relatives does not satisfactorily demonstrate non-IE origin, and without comparanda, little more can be said.

fovea ‘pit, trap, cave’

Pre-form: **b^h/d^h/g^{wh}e/ou-* | PItal. **f/b/χ^wowejo-*

Comparanda: ?

- Irreg. correspondences □ Remarkable phonotactics

Semantics: magico-religious / geography

Pokorny (451), WH (I: 467-8, 538), EM (221, 250), DV (237)

Solmsen (1904: 4), Ernout (1946: 35-6), Schrijver (1991: 443-4, 448), Breyer (1993: 256-9), EDG (1618)

Lat. *fovea* ‘pit’ is without comparanda. A connection with Gk. χεῖά, Hsch. χεῖά· ἡ κατάδυσσις τῶν ὄφεων καὶ δρακόντων ‘serpent’s den’ (WH I: 538 with lit.), whose further relations inside Greek are unclear (EDG 1619), fails on inherited grounds in that **g^h* does not yield Lat. *f-* (Schrijver 1991: 448, DV 237;). Both may be loans from a common source, but the semantic link is not very strong (EM 250).

³⁰⁹ Called Etruscan by Pseudo-Dioscorides and appearing in Calabrian *mútaka* ‘*Cistus monspeliensis*’.

³¹⁰ Alessio argues that this is a Mediterranean suffix, but it occurs in native formations as well (cf. *cadūcus* ‘fallen’ : *cadō* ‘to fall’, *fīdūcia* ‘trust’ : *fīdō* ‘to trust’).

The appurtenance of *favis(s)ae* ‘cisterns?’ is considered doubtless by WH (I: 467-8). But it is doubted by EM (221) and DV (237) and all but rejected by Schrijver (1991: 443-4) on semantic grounds.³¹¹ The *-issae* ending is widely considered to be of Etruscan origin,³¹² but this gets us no closer to an etymology. Either *fovea* is also Etruscan or *favis(s)ae* is a Latino-Etruscan hybrid formation (cf. Breyer 1993: 256-9 with lit.) via Thurneysen-Havet’s Law (**fou.íssae* > *favis(s)ae*, Solmsen 1904: 4). Schrijver (1991: 444) is hesitant to explain one etymologically obscure word as a regular development from another etymologically obscure word, and his hesitation is well advised. Without further comparanda, little more can be said about *fovea*.

hircus ‘he-goat’

Pre-form: **g^her-k-* | PItal. **χi/erko-*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: animal, domestic

Pokorny (445-6), WH (I: 649-51), EM (296), DV (286)

Fruyt (1986: 242), Blažek (2005: 6-7), Kloekhorst (2008 s.v. *ualkuūa-*), Garnier (2017b)

Given that the other Latin goat words are potentially non-IE, some suggest *hircus* is too (EM 296, DV 286), but there are potential explanations for its invalid **D^{he}T^(w)* root structure (with labiovelar reconstructed to account for Sabine *hirpus* ‘wolf’, WH I: 649, Fruyt 1986: 242, EM 296). WH (I: 649-50) compare it to *hirtus* ‘rough-haired’. Perhaps *hircus* < **χerk-o-* and *hirtus* < **χerk-to-* (with dialectal raising of **e* before *rC*) to **g^her-*, an *s*-less variant of **g^hers-* (cf. *horreō* ‘to stand erect’)(DV 286). The velar element could be a *k*-suffix. Without secure comparanda, we cannot see if it is part of the root. Garnier (2017b) alternatively proposes a sound law whereby **-t-ŷ-* > **-k-ŷ-* such that some case forms of a formation **hirtuus* would yield **hirquus* > Lat. *hircus*. (Presumably, if early enough, this would allow for the development of Sabine *hirpus* as well, but the semantics are obviously problematic [cf. DV 286].) Blažek (2005: 6-7) compares HLuw. *irwa-* ‘gazelle’, which would as good as guarantee an IE origin, but Kloekhorst (2008 s.v. *ualkuūa-*) shows that a labiovelar is not lost in this environment in Luwian (cf. CLuw. *papparkuūa* ‘to cleanse’ < **prk^w-*).

Lār, Laris ‘tutelary deity’

Pre-form: **leHs-* | PItal. **lās-*

Comp.: ?

³¹¹ Its meaning is not entirely clear: either cells and cisterns under the Capitoline temple in which sacred objects are placed (Varo *apud* Gellius) or areas of enclosed water around temples (Festus).

³¹² Cf. Ernout (1946: 35-6), on e.g. the evidence of *mantissa* ‘addition, makeweight’, *lingua Tusca* according to Festus.

- Irreg. correspondences □ Remarkable phonotactics

Semantics: magico-religious

Pokorny (654), WH (I: 762-3, 766), EM (341-2), DV (327, 328, 380)
Schrijver (1991: 167-8), Breyer (1993: 42-3)

The *Carmen Arvale* has *Lases*, strongly suggesting that the root of *Lār* is not **lar-* but rather **las-* (though the form could be purposefully archaizing). DV (327) argues that the ablaut *lār- : lār-* is not a productive pattern and therefore is a secondary phonetic development or the result of a loanword. It occurs in *sāl, sālis* ‘salt’ and *mās, māris* ‘male, masculine’, but they are archaic (Schrijver 1991: 167-8). On the other hand, WH (I: 762-3) argue that the length of *Lārs* is not secure; it is not metrically secured, only explicitly called for in Priscian. Comparanda are only convincing within Latin: *Lārua* ‘evil spirit, demon, mask’ < **lās-Vwā-* is formally and semantically a good match for *Lār* (WH I: 766, EM 342, DV 328). A further connection to *lascīvus* ‘frisky, lustful’ in which the root **las-* meant ‘eager’ or ‘voracious’ (cf. Pokorny 654) is difficult to prove.

EM (341-2) suspect Etruscan origin exclusively on semantic grounds and due to the similar formation in *Minerva*. The latter is however conventionally explained as inherited **menes-wo-* < **men-os-* ‘thought’ (cf. DV 380), regardless of the existence of a morpheme *-ua* in Etruscan (Breyer 1993: 42-3). *Lār* remains without secure external comparanda to help determine its origin.

mēlēs ‘badger, marten’

Pre-form: **meH-l-* / **mēl-* | PItal. **mēl-*

Comp.: ?

- Irreg. correspondences □ Remarkable phonotactics

Semantics: animal, wild

Pokorny (118-20), WH (I: 474), EM (394)
Alessio (1944a: 138), REW (no. 5474), Schrijver (1991: 375)

The meaning of Lat. *mēlēs* is not entirely clear. In Pliny (*Nat.Hist.* 8.138), it is an animal that inflates its skin to repel the blows of men and the bites of dogs. Perhaps this might be the sturdy badger. But it is followed by a description of the behavior of squirrels. If this is any indication that Pliny thought they looked similar, then perhaps it is a marten or polecat. In Varro (*de Re Rustica* 3.12.3) both *fēlēs* and *mēlēs* are animals that can be kept out of a rabbit warren by plastering the gaps in the surrounding fence. That he uses both words suggests there is a distinction between them, but that he mentions them in rapid succession might mean that they are synonyms for weasel-like pests. In any case, the Romance descendants of *mēlēs* (like Tarentine *miloña*, Calabrian *muloña*, REW no. 5474) mean ‘badger’. Alessio (1944a: 138) notes Spanish *melandro* ‘badger’, emphasizing that the suffix is found in some substrate words; but the rest of the root

shows no irregular alternation.

Thus, despite *mēlēs* often being considered borrowed with *fēlēs* (s.v.), from an Alpine language and presumably attesting to a **b^h ~ *m* alternation (Pokorny 118-20, WH I: 474, EM 394), Schrijver (1991: 375) keeps them separate, only comparing them in that the *-ēs* declension may have been generalized to both from other animal words like *volpēs* ‘fox’. I feel it is best to keep them separate as well, which leaves *mēlēs* without comparanda. No more can be said of its origins.

nūtēla ‘kind of rodent’

Pre-form: **(k)neit-* | PItal. **nūtēlā-*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: animal, wild

WH (II: 170), EM (442), DV (410)

WH (II: 170) suggest that *nūtēla*, better attested as *nūtēdula* ‘hazel dormouse’ is from *nūtor* ‘to climb’. But as DV (410) notes, *nūtor* does not mean ‘to climb’; it means ‘to lean, exert’. He suggests a connection with *nīdor* ‘strong smell, fumes’ (its other cognates mean ‘to scratch’), which would produce a non-IE *d ~ t* alternation. But no comparanda are certain enough to confirm an origin.

puteus ‘well, pit’

Pre-form: **put-* | PItal. **putejo-*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: geography

Pokorny (827), WH (II: 393), EM (547-8), DV (502)

Breyer (1993: 378-9)

Pokorny (827) connects *puteus* ‘well, pit’ to *pavīre* ‘to thump, pound, strike’, but this is formally difficult (DV 502 derives *pavīre* from **ph₂u-ie/o-*, thus the short *u* of *puteus* rules out a direct connection). A derivation from *putāre* ‘to prune’ (from the same root) would work (WH II: 393), but is semantically arbitrary, especially given the *-eus* ending of *puteus*, which looks like the material suffix (DV 502). DV thus considers the possibility of a loanword. EM (547-8) compare the *-eus* ending to that in *balteus* (s.v.), said to be an Etruscan word, and thus suggest Etruscan origin. Breyer (1993: 378-9) rejects *-eus* as an ending indicative of Etruscan origin. The suffixes may be the same, but *balteus* has no comparanda either.

rumex ‘sorrel, dock’

Pre-form: **H/uru-m-* | PItal. **rumek-*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: plant, wild

WH (II: 450), EM (581)

Osthoff (1890: 76-8), Krogmann (1938: 133), EDG (1295), Kroonen (2013: 493), Weiss (2020: 181-2)

Lat. *rumex* ‘sorrel, dock’ has no secure comparanda. Osthoff (1890: 76-8) suggested a connection with the words for ‘sour’, but his suggestion relied on rejecting *fr* as the regular Latin outcome of **sr* as well as connecting **suH-ro* ‘sour’ (cf. Kroonen 2013: 493 on the reconstruction) to OIr. *serb*, MW *chwerw* ‘bitter’, which is untenable (cf. Krogmann 1938: 133). WH (II: 450) present his argument as though he argued for Latin attesting to an *s*-less variant of the ‘sour’ lexeme (likewise impossible, as the rhotic element is not part of the root but is rather the *-*ro* suffix), along with Gk. ῥῶτή ‘rue’.³¹³ Though the latter is likewise without etymology, *rumex* and ῥῶτή share too little semantically to suggest that they both contain a root **ru-* with non-IE length alternation. Krogmann (1938: 133) took *rumex* from **rugmex* (for **rug-*, cf. Lith. *rūgti* ‘to ferment, grow sour’) with dissimilatory loss of *g*, but this is *ad hoc*.³¹⁴ Thus the origin of *rumex* remains unclear.

sagitta ‘arrow, bolt, shaft’

Pre-form: **sa/Hg-* | PItal. **sagit(t)ā-*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: tool

WH (II: 464), EM (588), DV (534)

Defrémercy (1862: 89), Alessio (1944a: 142-3), Ernout (1946: 39), Prasse (II 2003: 882), Ritter (I 2009: 796), Weiss (2010b)

Lat. *sagitta* is without etymology (cf. DV 534) and because of this, it is often considered to have been borrowed from a non-IE Mediterranean language (Ernout 1946: 39, WH II: 464, EM 534). Plautus’ use of the word scans as *sagīta* (EM 534), with a non-geminate *t*. Alessio (1944a: 142-3) sees the ending *-itta* as Etruscoid, but finds it otherwise only in

³¹³ Krogmann (1938: 133) suggested both the Latin and Greek rue words were from a Mediterranean substrate, but there is no evidence to rule out a loan from Greek (cf. EDG 1295).

³¹⁴ The expected outcome of **gm* is probably *mm*, but is at least *gm* (cf. Weiss 2020: 181-2).

salapitta ‘a box on the ear’ and personal names like *Gallitta* and *Pollitta*.

Alessio further connects it with a widespread (modern) European Wanderwort (It. *zagaglia*, Sp. *azagaya*, Engl. *assegai*, etc. ‘iron-tipped spear, especially those of the Bantu peoples of southern Africa’). The European words are from Arab. *az-zaġāya* ‘bayonet’.³¹⁵ While Alessio argues that the origin of the Arabic word is Berb. *zayāya*, this lemma seems only to occur in the Tuareg dialects of Mali and Niger³¹⁶ and its morphophonology betrays that it is certainly a loan from Arabic rather than *vice versa* (Maarten Kossmann, p.c.). The word’s absence from Berber until the arrival of Arabic suggests its origins are not in North Africa. At best, Lat. *sagitta* shares a source with the Arabic word, but this is unlikely given the dates involved. Alessio’s (1944a: 142-3) further comparison with Gk. σαγήνη ‘large drag-net’ is semantically distant. *Sagitta* remains without etymology or comparanda.

scurra ‘urban dandy; joker’

Pre-form: **sk(u)r-s-* | PItal. **scurrā-*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: characteristic

Pokorny (933-5), WH (II: 502), EM (606), DV (548)

Breyer (1993: 275-6, 279-81), Meiser (2010: 63-4), Willi (2012: 267-9), Zair (2017: 262, 266), Weiss (2020: 104, 320)

If inherited, Lat. *scurra* ‘dandy, joker’ could represent an example of the irregular reflex *ur* < **r* (cf. further **kṛse/o-* > *currō*, **kṛtos* > *curtus* ‘short’), which might be a dialectal treatment (Meiser 2010: 63-4, Weiss 2020: 104). But attempts at an inherited etymology have not been entirely successful (Zair 2017: 262, 266). Pokorny (933-35) and Meiser (2010: 63) take it from **skers-* ‘to jump, jump around, move oneself in a turning fashion, swing’. But as DV (548) notes, this requires more imagination than evidence. Furthermore, to the root *(s)*k’er-*, LIV2 links only Gk. σκαίρω ‘to jump, hop, dance’ and W *cerddaf* ‘to walk, journey’, both with uncertainty. Willi (2012: 267-9) connects *scurra* to the root of *scīre* ‘to know’ via the *pius* and *littera* rules, but the environment is not correct for the latter (Zair 2017: 262).

The other popular explanation for *scurra* is a borrowing from Etruscan (WH II: 502, EM 606). This relies on Etruscan forms of unknown meaning, perhaps representing a root *scur-* (Etr. *scurineś*) extended with the suffix *-na* (*scurnal*, *scurnas*, etc.), as well as the

³¹⁵ Defrémery (1862: 89) wrote that in Algeria, it had the more specific meaning ‘iron hook at the end of a stick for hunting hedgehogs and porcupines’.

³¹⁶ *Tazyāyt* ‘steel of high quality, sword blade of high quality, sword of high quality (imported from Libya or Egypt)’, (Prasse II 2003: 882) or ‘blade generally of European origin’ (Ritter I 2009: 796). The *ta-* ... *-t* is a feminine circumfix.

fact that *scurra* is a masculine in *-a* (Breyer 1993: 275-6 with lit.). While such formations often refer to “low-down types” (Weiss 2020: 320) like *lixa* ‘camp-follower’ and *verna* ‘slave born into his master’s house’,³¹⁷ they are not actually good indicators of Etruscan origin. *Scrība* ‘scribe’ shows that inherited roots can build this formation, nor is it unknown to other IE languages, most notably Greek. Without a better etymology or secure comparanda, the origin of *scurra* remains unknown.

sīl ‘ochre’

Pre-form: **siHl-* / **seil-* | PItal. **sīl-* / **seil-*

Comp.: ?

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: geography

Pokorny (923-7), WH (II: 535-6), EM (625), DV (564)

LS (s.v. *sil*), LEIA (S-38), EDG (1321), Derksen (2014 s.v. *skalà*)

WH (II: 535) enigmatically state that Lat. *sīl*, *-is* ‘ochre’³¹⁸ is identical to Lat. *sil/sil(li)* ‘seselis’, but this is unlikely for semantic reasons. The latter, also attested as *seselis*, is clearly a Greek loan, corresponding to Gk. σέσελι(ς) and σίλι (WH II: 535) ‘hartwort’ (EDG 1321). Pseudo-Dioscorides ascribes it Egyptian origin (cf. *nāpus*, s.v.).

Given the geological connotation of otherwise isolated Lat. *sīl* ‘ochre’, I suggest a connection with Lat. *silex* ‘hard rock, flint, lava’. Otherwise Pokorny (923-7) and WH (II: 536) explain *silex* as dissimilated from **skelik-* to *(s)*kel-* ‘to cut’, yielding Lat. *calx* ‘limestone, chalk’, Lat. *siliqua* ‘legume pod’, Mlr. *sceillec* ‘rock, stone, crag’, and OCS *skolbka* ‘shell, mussel’. The dissimilation is not regular however (EM 625, DV 564), making the connection of *silex* with anything but *siliqua* unlikely (LEIA S-38 on the Irish form, DV 564 [cf. Derksen 2014 s.v. *skalà*] on the Slavic form). But *siliqua* is too semantically disparate to be a match.³¹⁹

Only *sīl* and *silex* may potentially belong together, but they remain without further comparanda.

2.3.2 Uncertain Comparanda

abiēs, *-ētis* ‘fir tree’

Pre-form: **h₂eb^(h)i-et-* | PItal. **ab/fiētis-*

³¹⁷ *Verna* is likewise of unclear etymology and has itself been attributed to Etruscan, though with equally limited evidence (cf. Breyer 1993: 279-81).

³¹⁸ WH (I: 535) and EM (625) give the earliest attestation as Pliny, but LS (s.v. *sil*) cite at least one occurrence in Vitruvius.

³¹⁹ Instead, for it Bertoldi (1937: 141) mentions Basque *sigil* ‘vetch’. But in searching for it in Trask (2008: 367) one finds instead *zalke* ‘*Vicia sativa*’. Remarkably similar to Lat. *siliqua*, Trask also cannot explain the alternation between *-lk-* and *-lg-* in some dialects.

Comp.: **hzebi-* | PGk. **abi-* | Hsch. ἄβιν· ἐλάτην· οἱ δὲ πέυκη [acc.sg.] ‘fir tree’

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: plant, tree

WH (I: 4), EM (3), DV (20)

Smith (1854: 3), André (1956: 13), EDG (5)

Lat. *abiēs* ‘fir’ is compared to Hsch. ἄβιν of the same meaning (WH I: 4, André 1956: 13, EM 3, DV 20, EDG 5). Further comparanda are extremely speculative (a tribe called the Abii, the region of Hylaea (Υλαία ‘woody’) also being called Ἀβική, Smith 1854: 3).³²⁰ DV (20) suggests a non-IE origin given the lexeme’s limitation to the Mediterranean, the **b*, and the fact that there is no indication that the word glossed by Hesychius is actually Greek. André (1956: 13) considers the word pre-IE but not Mediterranean, given the growth zones of the fir. The Hesychian comparandum is not strong enough to help determine the origin of *abiēs*.

aesculus ‘type of oak’

Pre-form: **hzei(g/k)s-* | PItal. **ai(k/g)sk/t(V)lo-*

Comp.: ?**hzeig-* | PGm. **aik-* | ON *eik*, OE *āc*, OHG *eih*, etc. ‘oak’

?**hzeig-* | PGk. **aig-* | Gr. αἰγίλωψ ‘type of oak tree, haver-grass’

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: plant, tree

Pokorny (13), WH (I: 20), EM (13), DV (28)

Niedermann (1909: 49), Kretschmer (1912: 335), Strömberg (1940: 137), Hubschmid (1953: 82-4), Schrijver (1991: 39), Schrijver (1997: 306), Orel (1998: 88), Derksen (2007: 388), EDG (32), Kroonen (2013: 9), Šorgo (2020: 460)

Because of the shape of the Latin word, any plosives between the vowel and sibilant that may have existed are obscured. The second element of the Latin word is not likely to be a diminutive for semantic reasons (cf. WH I: 20 **aig-s-clo-* / **aig-s-colo-*), and Schrijver’s (1991:39) alternative **hzeig-s-tlo-* is also possible. Based on the semantics, it could be related to PGm. **aik-* ‘oak’ and perhaps to Gk. αἰγίλωψ ‘type of oak tree; haver-grass’.³²¹

Pokorny (13), EM (13), and DV (28) hesitantly suggest a Mediterranean origin for the lexeme. Hubschmid (1953: 82-4) and Schrijver (1997: 306) adduce Hsch. ἄσκρα· δρῦς

³²⁰ There is no compelling evidence that the Abii actually existed and were not just a play on words by Homer.

³²¹ The meaning ‘haver-grass’ is probably due to confusion with αἰγίλος ‘haver-grass’ (Strömberg 1940: 137, EDG 32). Kretschmer (1912: 335) uses a description by Pliny of the *aegilops* to suggest that αἰγίλωψ is the cork oak, thus **aig-* ‘oak’ + *lōps* ‘*cork’ (cf. λώπη ‘mantle, cloth’, and the way that Pliny describes the tree as “bearing strips of dry cloth”).

ἄκαρπος ‘a tree without fruits’ and Basque *askar* ‘type of oak’,³²² which would attest to a non-IE *a ~ ai* vocalic alternation. Since Basque *ezkur* ‘acorn’ has an older meaning ‘tree’ preserved in a proverb (Trask 2008: 188), it may belong to the comparanda. EM (13) mentions a connection with Berber *iškir* ‘wild oak’. It and a likely related Tuareg form *ašək* ‘tree, plant, shrub’ are indeed difficult to reconstruct due to the presence of š, which is generally not reconstructible to earlier stages of Berber. Nor do they look like loans from Latin.

Comparison with Baltic forms (Niedermann 1909: 49, Kroonen 2013: 9) including Lith. *q̄žuolas*, *aižuols*, *áužuolas*, Latv. *uōzuōls*, and OPru. *ansonis* ‘oak’ is difficult. Derksen (2007: 388) reconstructs these (along with their Slavic cognates Ru. *úzel* etc. all meaning ‘knot’) to PBSl. **onʔz̥-(ō)l-*, a form with a nasal. Šorgo (2020: 460) takes this as evidence of non-IE pre-nasalization, but this means accepting several irregular alternations. Alb. *enjë* ‘English yew; stinking juniper’ can reconstruct to PAlb. **ai(g?)njā*, so Orel (1998: 88) adduces it as a comparandum despite its aberrant semantics. A more straightforward reconstruction of **e/ēnjā* is also possible, which looks more similar to Sard. *ēni* and PGm. **(j)ainja-* ‘juniper’ (cf. Lat. *iūniperus*, s.v.).

It is unclear if Lat. *aesculus* ‘type of oak’ is related to any of the forms beyond PGm. **aik-* ‘oak’ and perhaps Gk. *αιγίλωψ* ‘type of oak tree’, none of which has any blatantly non-IE features. But its lack of attested velar before the sibilant means that even a connection with these words is difficult to verify.

alga ‘algae, seaweed’

Pre-form: **Hlg-* / **h₂elg-* | PItal. **algā*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: plant, wild; aquatic

Pokorny (305), WH (I: 28-9, II: 813), EM (20, 744), DV (33)

Lidén (1897: 29-31), Schrijver (1991: 70), EWAia (I: 252), Martirosyan (2009: 32, 39), Kroonen (2013: 598)

If Lat. *alga* is connected to Lat. *ulva* ‘aquatic plants’ < **Hol-Vua-* or **Holg^{wh}-* (cf. Pokorny 305, WH I: 28-9), the consonant alternation points to a foreign origin (DV 33). But other comparisons exist. Most frequent is a comparison to several words for repulsive, slimy things, interpreted as a root **Vl-* with numerous extensions (Nw. dial. *ulka* ‘mold, slime; to feel sick, vomit’, ON *uldna* ‘to mold’, Lith. *eļmes* ‘exudate of a corpse’, Arm. *alt*, *alb* ‘dirt’, even Skt. *ṛjīśā-* ‘an epithet of Indra’, etc.; cf. Lidén 1897: 29-31, WH I: 28-9 with lit.). A stricter comparison with only the forms that reconstruct to **Hlǵ-* would remove the irregular alternations (cf. Schrijver 1991: 70). But the semantic connection is tenuous (EM 20, DV 33) and alternative etymologies for several

³²² But Basque *askar* and Hubschmid’s further comparison of Basque *gastigar*, Languedocien and Prov. *agast*, etc. are best kept separate because they mean ‘maple’.

of the proposed comparanda also exist (cf. Kroonen 2013: 598 for the Germanic, Martirosyan 2009: 32, 39 for the Armenian). It is not immediately clear which of the proposed solutions is best.

apis ‘bee’

Pre-form: **h₂ep-* | PItal. **ap-*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: animal, wild; insect; apiculture

WH (I: 57), EM (39), DV (47)

Ernout (1925: 115), Alessio (1944a: 130), Erman & Grapow (1971: 182), Schrijver (1991: 374), Gamkrelidze & Ivanov (1995: 516), Vennemann (1998: 485-6), van Sluis (2022: 9-10)

WH (I: 57 with lit.) reject most etymologies for Lat. *apis* ‘bee’. It may be irregularly related to the **b^hei-* word found elsewhere (EM 39, Gamkrelidze & Ivanov 1995: 516). The latter lexeme exhibits consonant alternations that make it likely to be of non-IE origin, and *apis* could represent an *a*-prefixed variant (van Sluis 2022: 9-10). Vennemann (1998: 485-6) proposes a loan from Egypt. *ʃj* ‘bee’³²³ at a time early enough that Italic had not yet developed /f/ and would have substituted /p/. A problem for both suggestions is that both *apium* and *apum* are attested genitive plurals, suggesting that *apis* is only secondarily an *i*-stem (Ernout 1925: 115, Schrijver 1991: 374). Alessio’s (1944a: 130) comparison to Basque *abia* ‘gadfly’ on comparison with the French collocation *mouche à miel* for ‘bee’ is semantically dubious.

aula ‘cooking pot’

Pre-form: **h₂eug^(w)(^h)/k^(w)-slo-* | PItal. **aukslā*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: vessel

Pokorny (88), WH (I: 88), EM (59), DV (62)

Lehmann (1986: 49), Schrijver (1991: 47), EWAia (I: 210), Demiraj (1997: 76-7), Vine (1999: 20-24), Kloekhorst (2008: 348), Kroonen (2013: 3, 557), EDG (596), Weiss (2020: 193)

The diminutive *auxilla* ‘small jar’ shows that *aula* is from **aukslā-* (Schrijver 1991: 47, Weiss 2020: 193), which can be reconstructed to any PIE velar (DV 62). The complete

³²³ Cf. Egypt. *ʃj-n-bj.t* ‘honey bee’ > Copt. ⲁϥⲛⲉⲃⲓⲱ /afnebiō/ ‘bee’ (Erman & Grapow 1971 I: 182).

set of comparanda is difficult to verify. Skt. *ukhā-*, *ukhā-* ‘boiler, pan’ looks closest (EWAia I: 210, DV 62). Its voiceless aspirate, if the words are inherited, must be from a following laryngeal. Others are more ambiguous. Germanic seems to show two differently shaped roots, **uhna-/ugna-* and **ufna-* (Lehmann 1986: 49 with lit.). The former could be Verner variants of **h₂uk-* (DV 62) while the latter reconstructs to **h₂up-*. But Kroonen (2013: 557) argues that all Germanic forms are reconstructible to **ufna-* < **upno-*, a Wanderwort along with Gk. ἰπνός ‘furnace; kitchen; lantern’,³²⁴ OPr. *wumpnis* ‘oven’, and Hitt. *happen-*, *hapn-* ‘baking kiln, fire-pit, broiler (oven)’. Whether Lat. *aula* is related (cf. DV 62 hesitantly) is difficult to confirm. The semantics are adjacent, but are they good enough to accept a **p ~ *k^(w)* alternation? (Lat. *aqua* ‘water’ (s.v.), PGM **ahwo-* ‘river’ < **h₂ek^w-* against **h₂ep-* elsewhere, and whether this can be regular or not.) Alternative etymologies for the forms involved also exist. (Greek from **sep-* ‘to boil, bake’ with *schwa secundum*, Vine 1999: 5-30, EDG 596; Hittite to Gk. ὀπτός ‘baked’ < **h₃ep-* [Kloekhorst 2008: 348 with lit.] or to PGM. **afla-* ‘hearth’ [Kroonen 2013: 3].) Alb. *ani*, *anë* etc. ‘vessel, kitchen appliance’ could be from **aukn-* < **h₂euk^(w)-sno-*, but alternative connections exist (e.g. Gk. ἔντεα ‘equipment’, Demiraj 1997: 76-7).

bucca ‘puffed out cheek; mouth’

Pre-form: **buK-* | PItal. **buccā*

Comp.: ?

□ Irreg. correspondences

■ Remarkable phonotactics

Semantics: body part

Pokorny (98-103), WH (I: 120), EM (77), DV (76)

Sihler (1995: 224), Matasović (2009: 60), Kroonen (2013: 400)

The gemination in Lat. *bucca* has been proposed to be expressive, perhaps hypocoristic or belong to abusive words like *gibber* ‘hunch-backed’ (Sihler 1995: 225, DV 76). Celtic origin has been suspected (EM 77) on the testimony of Suetonius (*de Vita Caesarum*, section on Vitellius). Names like *Buccus*, *Buccō*, and *Bucciō* are of Celtic origin, but even if Lat. *bucca* is from one of them, their relation to PCelt. **bekko-*³²⁵ with **e* is not clear. Otherwise, Pokorny (98-103) compared PGM. **puh/kkan-* ‘bag’, whose geminate is the result of Kluge’s Law from **buk-n-* (Kroonen 2013: 400). It begins with rare **b*, as does Latin *bucca* at face value. Nw. *poka-* ‘pigskin, sward’ might point to the original lexeme having meant ‘animal skin,’ semantically remote from Lat. *bucca*. But if Pol. *buczyć się* ‘to puff oneself up’ < **bouk-eie-* is related, the ‘inflated’ semantics match the ‘puffed out cheek’ meaning that DV (76) considers primary for *bucca*. It is unclear which explanation to accept.

³²⁴ WH (I: 88) compare Gk. ἰπνός to Lat. *aula* directly, but Myc. *i-po-no* rules out PGk. **k^w*.

³²⁵ Matasović (2009: 60) likewise analyzes the geminate in Celtic as expressive.

carbō, -ōnis ‘piece of charcoal’

Pre-form: **krH-(V)d^h/b^(h)-?* | PItal. **kar(a)b/fo-*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: tool

Pokorny (571-2), WH (I: 165-6), EM (99), DV (91)

Schrijver (1991: 194-5, 207-8), EDG (651), Kroonen (2013: 258), Derksen (2014 s.v. *kurti*)

Non-native origin of Lat. *carbō* has been suspected because of its technical (EM 99) or “specific” (Schrijver 1991: 208, DV 91) meaning. But a stronger argument is formal: at face value, it reconstructs to an invalid **TeD^h* root structure. The details of its reconstruction however rely on its cognates/comparanda, which are difficult to verify. Pokorny (571-2) connected it to Lith. *kūrti* ‘to kindle’. DV (91) rejects the connection, taking *kūrti* simply from **k^wer-* to make. But Derksen (2014 s.v. *kūrti*) notes that this does not explain the acute accent, instead proposing **krH-*,³²⁶ which could be the source of *carbō* via a suffixed **krH-eb^(h)-* > **kareb/f-* with subsequent syncope (Schrijver 1991: 207).³²⁷ Via a *palma* rule formation (**k^rH-b^h-*) the suffix need not be in the *e*-grade. PGM. **hurja-* (cf. Go. *hauri* ‘coal, burning charcoal, ember’, ON *hyrr* ‘fire’), if related (Kroonen 2013: 258 is hesitant), indicates that it really is a suffix, removing the need to reconstruct an invalid root structure for *carbō*.

Alternative connections are also possible. A connection with *cremō* ‘to burn’ (WH I: 166, Schrijver 1991: 208, DV 142) requires a root without a laryngeal. DV (142) operates with a root **krb-* (**krbn-* > **karbn-* for *carbō* and **krebm-* > *kremō*), but it is unclear why *a*-vocalism should develop. Schrijver (1991: 208) analyzes *cremō* as **kr-em-* (implying *carbō* < **ker-b^(h)/d^h* with **e* > *a* after a plain velar), but it is unclear what the suffix of the formation **kr-em-* is. Lat. *carbō* can reconstruct to the same pre-form as Gk. *κάρφω* ‘to dry up, wither’, *κάρφος* ‘arid stalk, twig, hay’, but the semantics are not a good enough match.

carīna ‘ship’s hull or keel; walnut shell’

Pre-form: **ka/Hr-* | PItal. **kar-*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

³²⁶ Kroonen (2013: 258) advocates for **h₃* to explain **r* > BSl. **ur* instead of **ir*.

³²⁷ Schrijver (1991: 194-5, 2007-8) considers a reconstruction with **-ed^h-*. But given that the change **d^h* > *b* seems to be part of the Proto-Italic treatment of the voiced aspirates, it is difficult to imagine that it could have operated after syncope in Latin. Thus, a reconstruction with **d^h* likely only works in a pre-form **kerHd^h-*, of illegal root structure, and assuming Schrijver’s change **e* > *a* after a plain velar.

Semantics: maritime

Pokorny (531-2), WH (I: 168), EM (100), DV (93)

Furnée (1972: 391), Schrijver (1991: 208), Biville (II: 32), Kroonen (2013: 211), EDG (645, 651, 772), van Sluis (fthc.)

The earliest attestations of *carīna* refer to ships and only from Pliny onwards does the nutshell meaning occur (DV 93), but EM (100) consider that this could be an artifact of preservation.³²⁸ In Greek, *κάρυον* ‘nut’ never has the maritime semantics. Perhaps *καρύϊνος* ‘of nuts, nut-brown’ through ‘like a nut shell’ meant ‘ship’s hull’ in a dialect of Magna Graecia/Sicily or Koine, whence it entered Latin as a loan (WH I: 168, Biville II: 32, EDG 651). Otherwise, Schrijver (1991: 208) compares *carīna* and Gk. *κάρυον* to *W ceri* ‘stone of a fruit’ (< **ka/e/orī*), perhaps to the alleged root **ker-* ‘hard’,³²⁹ itself perhaps of non-IE origin due to the *κ ~ χ* alternation attested between Hsch. *κάρκαρον* *τραχείς*, *καὶ δεσμοὶ* ‘coarse, rough’ and Gk. *κέρχνος* ‘raw voice; hoarseness’ (EDG 645). Whether or not the connection to the ‘hard’ root is valid, further evidence of a non-native origin of Gk. *κάρυον* is the potential *k ~ zero* alternation it shows with Hsch. *ἄρυα*: *τὰ Ἑρακλεωτικὰ κάρυα* (Furnée 1972: 391, EDG 651).

W ceri more frequently appears with the meaning ‘service tree, rowan’. If they are the same lexeme, they do not belong here (van Sluis fthc. argues that the ‘rowan’ word is a substrate word restricted to Celtic). The precise relationship of Lat. *carīna* to Gk. *κάρυον* remains unclear.

cāseus ‘(a unit of) cheese’, also *cāseum* ‘cheese’

Pre-form: **keh₂t-s* | PItal. **kāssejo-*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: culinary; dairy

Pokorny (627), WH (I: 176), EM (103), DV (96)

Schrijver (1991: 251-2), Christol (1996), Kroonen (2013: 264)

Etymologies for Lat. *cāseus* are problematic (WH I: 176 with lit., EM 103). Connection with a root **k_ueth₂-* ‘to boil, bubble’ (cf. Go. *lvapō* ‘foam’ < **k_uotH-eh₂-*, Skt. *kvāthant-* ‘fuming’ < **k_uéH-e-*, Kroonen 2013: 264) requires the unexplained loss of **_u*. Otherwise compared are Latv. *kūsāt* ‘to boil over’ < **kHus-* and Slavic words for sour (OCS *kysnōti* ‘to turn sour’ < **kuHs-* itself < **kHus-*; OCS *kvasъ*, Russ. *kvas* ‘leaven’ < secondary full-grade **kuHs-*). The development in Latin would then be **kH_u-ōs- >*

³²⁸ Biville (II: 32) notes *carīnum* and *carīnariū*, hapaxes in Plautus that refer to women’s clothing, and which might be referring to their nut-brown color.

³²⁹ Cf. further Pokorny (531-2): PGM. **hardu-* ‘hard, severe’ and Gk. *κράτος* ‘strong’ but these can be reconstructed to a root **kert-* that includes the final dental (cf. Kroonen 2013: 211, EDG 772).

**kaṵōs-* > **kaōs-* > *cās-*, serving as the basis of a material adjective **kās-ejo-* > *cāseus* (Schrijver 1991: 251-2). This assumes the lack of rhotacism in ‘rural’ words, which Christol (1996) supports where it is attested after long vowels, proposing unrhotacized *z* < **s* was mapped on to *ss* when borrowed into Latin (cf. unrhotacized *nāsus* ‘nose’).³³⁰ Thus it could apply to *cāseus*. DV (96) remains unconvinced of the contraction of **kaṵōs-* > *cās-* and I am unconvinced of the semantic link.

cicāda ‘cicada, cricket’

Pre-form: **ki-keh₂d-* | PItal. **kikādā-*

Comp.: *(*d*)*ieigara-* | PGk. **zeigara-* | Hsch. ζειγάρη· ὁ τέττιξ παρὰ Σιδήταις
‘cicada amongst the
Sidetians’

□ Irreg. correspondences

■ Remarkable phonotactics

Semantics: animal, wild; insect

WH (I: 211), EM (119), DV (112)

REW (no. 1897), FEW (II: 663), Alessio (1943: 234), André (1978: 29-30)

Lat. *cicāda* is quite plausibly onomatopoeic, which would account for the reduplication (André 1978: 29-30, DV 112). It is unclear if this could be responsible for the *d* ~ *l* ~ *r* alternation (characteristic of other suspicious lemmata, cf. *laurus*, s.v.) that occurs in the final syllable of glosses and Romance forms like It. *cigala* and Sp. *cigarra* (REW no. 1897, FEW II: 663, Alessio 1943: 234, André 1978: 29-30). The similarity of the Hsch. ζειγάρη is curious, but it is difficult to verify that it is Greek. Other potential Hesychian comparanda include σιγαλ(φ)οι ‘voiceless; wild cicadas’ and ζεγερίαι ‘a kind of mouse’, but their forms and semantics do not allow much more than speculation. WH (I: 211) suggest *cicāda* might be from a Mediterranean substrate while EM (119) see it as an expressive word (like Gr. τέττιξ), which they consider a sort of Mediterranean regional feature. Given its potential onomatopoeic origin, its possible substrate origin remains uncertain without more secure comparanda.

cicūta ‘hemlock’

Pre-form: **ki-kuH-t-* | PItal. **kikūtā*

Comp.:³³¹ **ko-kuH-t-* | PCelt. **kokūtā-* > **kokītā-* | OBret. *cocitou*, Bret. *kegid*, W
cegid, Co. *ceges*³³² ‘hemlock’

³³⁰ Schrijver otherwise uses this to explain *asinus* ‘donkey’ and *casa* ‘hut’, where it is unlikely.

³³¹ Albanian has *kakuda*, *kukutë*, and *kukuta* ‘poison hemlock’, all seemingly from Romance. It also has *kakutë* ‘black henbane [*Hyoscyamus niger*], corn stubble left in a field’ (Newmark 2005 s.v.) however. Its phonology seems too similar to the other more clearly borrowed forms for its deviant semantics to suggest an independent comparandum.

³³² From this may have been borrowed Engl. *kex* ‘hollow stalk’ > W *cecys* ‘kex, reeds, hemlock’.

- Irreg. correspondences □ Remarkable phonotactics

Semantics: plant, wild

WH (I: 213), EM (119)

Prellwitz (1905: 171), Pedersen (1909-13 I: 209), Walde (1910: 159), Meyer-Lübke (1920), FEW (II: 668), Alessio (1943: 233), Furnée (1972: 121, 371), André (1978: 19-20), Newmark (2005), DV (139), EDG (53, 815)

Lat. *cicūta* looks like it is reduplicated,³³³ but the base is unclear. Walde (1910: 159) suggested a reduplication of *cautēs* 'rough, pointed rock', a hyperurbanized old plural of *cōs* 'whetstone' < **keh₃-* 'to sharpen'. But since *cautēs* appears as a hypercorrect spelling of *cōtēs* only after Vergil (DV 139), we would instead have to assume a reduplication of *cōtes* with dialectal *ū* for *ō*. This is also semantically unlikely. WH (I: 213) suspect foreign origin for *cicūta*. EM (119) seem to follow because of the reduplication.

The Romance descendants attest to three source forms. **Cicūta* (OFr. *cēue*) is expected. **Cucūta* (MFr. *cocue*) has been explained as due to assimilation (FEW II: 668, Alessio 1943: 233). As it is present in Romanian *cucută*, the assimilation is quite old or it occurred twice. A third form, **ciccūta* (OFr. *cegue*) reconstructs to a geminate. FEW (II: 668) suggests that the geminate already existed in Latin or that this represents a case of the initial *c* blocking the lenition of the second, intervocalic *c*.

On the understanding that all forms descended from **keh₃-*, Prellwitz (1905: 171) compared Gk. κόνειον 'hemlock' and κῶνος 'pinecone'. EDG (815) considers the Greek forms loans, following Furnée (1972: 121, 371) only as far as he also compares ἀκόνιτον 'wolf's bane', another poisonous plant. But whether the Greek words are inherited or not, their nasal makes them difficult to connect to Lat. *cicūta*. Much easier to connect are the descendants of a Proto-Brythonic **kokīlā-* (cf. Pedersen 1909-13 I: 209) as if from PCelt. **kokūlā-*. On the other hand, Proto-Brythonic **kokīlā*, depending on the chronology of the sound changes, could be a borrowing from the Proto-Romance variant **cucūta*, after *ū* > *o*. The appearance of Proto-Brythonic **ī* for Latin *ū* usually points to a very early loan, but later examples are not without parallel (see fn. 432).

Each of the irregularities has a potential explanation. But if the Brythonic forms are independent, and if the vocalic alternation and geminate in the Romance forms are genuine, *cicūta* could be a substrate word.

cirrus 'a curl, a lock of hair, the fringe of clothing'

Pre-form: **kel-ir-s-* | PItal. **kirso-*

Comp.: ?

- Irreg. correspondences □ Remarkable phonotactics

³³³ André (1978: 19-20) also favored a sound-symbolic formation because of the plant's use in flutes.

Semantics: body part / textiles

WH (I: 221-2), EM (123)

Niedermann (1927: 109-10), Alessio (1943: 232), Furnée (1972: 278), EDG (695), Kroonen (2013: 220), Derksen (2014 s.v. *keras*), Weiss (2020: 149)

The vocalism of *cirrus* can be original or the result of dialectal *e > i / _rC* (cf. Weiss 2020: 149). Its geminate *rr* is likely the result of **rs*. WH (I: 221-2 with lit.) mention several Baltic words meaning ‘tree stump/bad hair’ (e.g. Latv. *cēra* ‘messy hair, *cērba* ‘lock’, *cecers* ‘tree with torn-out roots; fuzzy-wuzzy’). Derksen (2014 s.v. *keras*), without mention of *cirrus*, adduces to the Baltic material PSlav. **čermъ/b* ‘stem, stub’ as well as OIr. *cern* ‘angle, corner’ and W *cern* ‘cheekbone, side of the head’, making it very unclear what the original meaning of this root would have been. Otherwise Niedermann (1927: 109-10) suggested an independent borrowing from a Mediterranean language of Lat. *cirrus* < **cicirrus* and Gk. κίκυνοϛ ‘curly hair, lock of hair’. Existing Lat. *cincinnus* ‘curled lock of hair’ is interpreted as a borrowing of Greek κίκυνοϛ with anticipation of the nasal (WH I: 216, Alessio 1943: 232, EM 123). But EDG (695) follows Furnée (1972: 279) in postulating **κικυνοϛ*, a Pre-Greek pre-nasalized variant of κίκυνοϛ. It seems possible that *cirrus* has something to do with this word based on the semantics, but its exact relationship is unclear.

CRUX ‘wooden frame, cross’

Pre-form: **kru-k-* | PlItal. **kruk-*

Comp.: ?

□ Irreg. correspondences

■ Remarkable phonotactics

Semantics: tool

Pokorny (935-8), WH (I: 296), EM (153), DV (147)

Derksen (2007: 254), Matasović (2009: 226), Kroonen (2011: 268-70; 2013: 250), van Sluis, Jørgensen & Kroonen (2023: 216)

EM (153) take *crux* as a Mediterranean loanword, perhaps Punic, based on cultural and historical arguments. WH (I: 296) would rather see it as inherited. A root **kruk-* would be of an invalid **C₁CC₂* root structure, thus the final velar would have to be a suffix (DV 147 is skeptical). Comparanda are uncertain but have included Skt. *kruñcati* ‘to make or become crooked’, PGM. **hrauka-* ‘pile, rick’,³³⁴ and PGM. **hrugja-* ‘ridge, back’. None is semantically convincing.

curñculus ‘rabbit; rabbit burrow > underground tunnel, mine’

Pre-form: **kun-īk-VI-* / **kun-isk-VI-* | PlIta. **kuni(s)klo-*

³³⁴ PCelt. **krowko-* ‘heap, hill’ is borrowed from Germanic (Kroonen 2011: 268-70, 2013: 250; van Sluis, Jørgensen & Kroonen 2023: 216).

Comp.: ?Basque *untxi* 'rabbit'

?Mozarabic *conchair* 'hunting dog'

Irreg. correspondences Remarkable phonotactics

Semantics: animal, wild

WH (I: 308-9), EM (157)

Simonet (1888: 128-9), Bertoldi (1937a: 146), Hubschmid (1943: 267-9), Corominas & Pascual (1984-91 III: 173), EIEC (258), Trask (2008: 27, 388)

Lat. *cuniculus* looks like diminutive, but without knowing the root, this may not actually be the case (EM 157). WH (I: 308-9) show that it cannot be connected with *canis* 'dog', *cavus* 'hollow', or *canālis* 'canal'. Aelian and Pliny write that *cuniculus*, like *laurex* 'unborn rabbits cut from the womb' are Iberian words. And the European rabbit (*Oryctolagus cuniculus*) was indeed foreign to Europe outside of Iberia (EIEC 258)(cf. *Iepus*, s.v.).

The best comparison is to Basque *untxi* 'rabbit' (WH I: 308, EM 157). Hubschmid (1943: 267-9) suggests that both Latin and Basque are borrowed from unattested Gaulish **kunikō-* 'little dog', but this seems unlikely. Trask (2008: 388) considers *untxi* as possibly part of the earliest stratum of the Basque lexicon, reconstructing it as **untzi* or **untsi* with expressive palatalization (the normal Basque way of forming diminutives). This all would indicate that the word indeed entered Latin from the West, as is argued for *Iepus*. While the Basque form is not particularly similar to the Latin, Corominas and Pascual (1984-91 III: 173) reconstruct **kun-txi*.³³⁵ Simonet (1888: 128-9) gives Basque *uncharia* 'podenco',³³⁶ whose pre-form may well be the source of Mozarabic *conchair* 'podenco, hunting dog' and 'dog' in general, attesting to the initial velar. If Basque *untxi* is indeed from earlier **kun-txi*, then it could be the ultimate source of Lat. *cuniculus*. But given the uncertainty in reconstruction, it is difficult to confirm.

fibra 'fiber, lobe'

Pre-form: **b^hi-b^(h)r-o-* | Pltal. **fibl^hfro-*

Comp.: *?*b^himb^(h)r-* | Pltal. **fimb^hl^hfro-* | Lat. *frimbriae* 'fringe on a garment, fringe of curly hair'

*?*b^he-b^hr-* | PGm. **bebura-* | ON *bjórr* 'piece of skin', Far. *bjóri* 'patch, strip', etc.

Irreg. correspondences Remarkable phonotactics

Semantics: textiles

³³⁵ See Trask (2008: 27) on "initial velar loss or gain". Cf. already Bertoldi (1937a: 146) on the preform.

³³⁶ Perhaps this would represent **kuntxi-ārius* 'the rabbitier'.

Pokorny (268-271), WH (I: 491-2), EM (232), DV (217, 220)

LS (s.v. *fibra*), de Vries (1962: 40), Magnússon (1989: 60), Kroonen (2013: 57)

LS (s.v. *fibra*) define *fibra* as ‘fiber, filament, entrails’. A more careful definition seems to be ‘the root fiber of a plant’ as well as ‘plant fiber, vein’ in general and in augural terms it refers to the lobes of organs like the liver and lungs (WH I: 491-2, EM 232, DV 217). The best explanation from an inherited perspective is a link with *filum* ‘thread’ via **g^{wh}is-lo-* (WH I: 491-2, EM 232), but DV (217, 220) shows that the root behind this is **g^{wh}iH-*, with any sibilant element (whose presence or absence cannot be seen in Latin after the lengthening by the laryngeal) belonging to the suffix (ruling out **g^{wh}is-ro-*). The only other attractive option was proposed in antiquity by Festus: a connection with *fimbriae* ‘fringe’, which DV (217) sees as a specialized meaning of *fibra*. The unexplained appearance of a nasal element would point to a non-IE word, but the semantic match is not as strong as between e.g. *sabūcus* and *sambūcus* (s.v.).

ON *bjórr* ‘triangular cut off piece of skin; land; party wall’ can reconstruct to **beura-* (cf. ON *bjórr* ‘beer’), but **bebura-* is also possible (cf. ON *bjórr* ‘beaver’). The comparison to Lat. *fibra* has led to a preference for the latter (cf. de Vries 1962: 40, Magnússon 1989: 60). Kroonen (2013: 57) provides Germanic-internal evidence for this reconstruction in the form of Far. *bjarva* ‘to mend, patch; wrap’ < ON **bjafra* (with regular metathesis). If the Germanic connection is upheld, then we have what looks remarkably like a duplicate of the beaver word, down to the aberrant *i*-vocalism in Latin (cf. *fiber* ‘beaver’) but without the widespread cognates or well-established derivation from another root to back up its inherited origin. Nevertheless, the *e* ~ *i* alternation in the inherited word alongside the inexact semantic match with Germanic (itself reconstructible in different ways) makes the comparison uncertain. Lat. *fibra* may be without comparanda, and thus several other reconstructions are possible (e.g. **b^{hi}d^h-ro-* / **b^{hi}-d^hro-*).

frōns, -dis ‘foliage, leaves’

Pre-form: **slb^hid^hg^{wh}ron(-)d^(h)* - | PItal. **slflp^hχ^wrondi-*

Comp.: ?

Irreg. correspondences

Remarkable phonotactics

Semantics: plant

Pokorny (142), WH (I: 550-1), EM (255), DV (244)

Solmsen (1898: 474-6), Furnée (1972: 189), EDG (557), Kroonen (2013: 81), van Beek (2022: 84-8)

Lat. *frōns* ‘foliage’ has been linked with several groups of comparanda, but it is unclear where it actually belongs. Solmsen (1898: 474-6) connected it with Ru. *děrn* ‘lawn’ etc.

and Gk. *θρόνα* ‘herbs, flowers’³³⁷ < **θ̥r(o)n-*, but this ignores the Hesychius variant *τρόνα: ἀγάλματα. ἢ ῥάμματα ἄνθινα* ‘ornament, colorful stitchings’, which makes the Greek word look non-IE (Furnée 1972: 189, EDG 557). In fact, van Beek (2022: 84-8) argues that the meaning preserved in Hesychius is the original one. By reinterpreting the Homeric hapax *θρόνα* as ‘dyed threads’ and Myc. *to-ro-no-wo-ko* as ‘dyers or dye-makers’, van Beek suggests that the lexeme in question, glossed in antiquity as *φάρμακα* in its technical sense ‘dye’, was misinterpreted to mean *φάρμακα* in its other sense ‘medicinal herbs’. Thus Gk. *θρόνα*, originally ‘dyed threads’, is semantically a poor match for *frōns*. WH (I: 550) compare ON *brum* ‘leaf bud’ (as if < **b̥r̥m-*) and DV (244) suggests a derivation from **b̥er-* ‘to bear’. All connections require Latin to have the *-*d(i)*- suffix of *glāns* ‘acorn’. Kroonen (2013: 81) thus compares MHG *brozzen* ‘bud’ < **b̥rd-néh₂-*, which would yield Lat. *frond-*. It is unclear which if any of these suggestions is correct.

frūmen ‘larynx, throat’

Pre-form: **sl̥b̥l̥d̥l̥gʷh̥rug-(s)men-* | Pltal. **frugsmen-*

Comp.: ?

Irreg. correspondences

Remarkable phonotactics

Semantics: body part

Pokorny (145), WH (I: 551-2), EM (256)

Martirosyan (2009: 258), Derksen (2007: 65), EDG (1556), Kroonen (2013: 53, 76)

Lat. *frūmen* ‘larynx, throat’ is only found in glosses. EM (245) explains it as preserving the original meaning of *fruor* ‘to enjoy’, namely ‘to nourish’, as also found in derivatives referring to nutriment like *frūmentum* ‘grains’ and *frūctus* ‘fruit, produce’. WH (I: 551-2 with lit.) and Pokorny (145) take it as an inherited word for gullet, comparing Gk. *φάρυξ* ‘throat, larynx’ (later *φάρυγξ* with contamination from *λάρυγξ* ‘larynx’), Arm. *erbuc* ‘breast of animals’ < **b̥rug-* and ON *barki* ‘throat, larynx’ < **b̥org-*. But Kroonen (2013: 53) doubts the appurtenance of the Germanic form and EDG (1556) disagrees with the patterning of *φάρυγξ* on *λάρυγξ*, taking it instead to contain a pre-nasalized suffix of non-IE (Pre-Greek) origin. Martirosyan (2009: 258) upholds the connection between Latin, Greek, and Armenian, suggesting that, if *φάρυγξ* is a substrate word, then all three might be. But if *-υγξ* is a suffix in *φάρυγξ*, then the word is not so similar to *frūmen* or *erbuc* after all. Nor can the shape of *frūmen* guarantee a relationship with *erbuc*. Its origin remains uncertain.

gigarus ‘*Arum italicum* or *Dracunculus vulgaris*’

Pre-form: **gi-galHr-* | Pltal. **gigaro-*

³³⁷ As described by EDG (557), this refers to flowers as a woven decoration in fabrics, as a medicine and charm, and potentially more generally for colorful clothing.

Comp.: ?

□ Irreg. correspondences ■ Remarkable phonotactics

Semantics: plant, wild

WH (I: 597), EM (275)

Bertoldi (1936: 298-9), Alessio (1937), Alessio (1943: 229, 233), Alessio (1944: 112), André (1956: 148), EDG (131, 136, 147)

Lat. *gigarus* is given as a Gaulish word by Marcellus Empiricus (WH I: 597, EM 275) but as Etruscan by Pseudo-Dioscorides (cf. André 1956: 148³³⁸). Bertoldi (1936: 298-9) suggests it would be easy to trust Marcellus, as he himself was from Bordeaux, and the suffix *-aro-* appears in other Celtic plant names. But only modern Tuscan dialects of Italian preserve the word. Thus Alessio (1937, 1943: 229, 1944: 112) considers it is more likely to be of Etruscan origin after all. He proposes a relationship to Gk. ἄρον and ἄρον ‘*Arum italicum*, cuckoopint’ (Alessio 1937) and Gk. ἀρίσαρον ‘*Arisarum vulgare*, friar’s cowl’ (Alessio 1943: 229), especially based on the testimony of Pseudo-Dioscorides. He considers the reduplication to be a Mediterranean feature (Alessio 1943: 229, 233). While Lat. *gigarus* and several Greek forms do refer to similar plants, the phonological relationship between them is difficult to confirm. EDG (131, 136, 147) compares within Greek ἀρίς ‘*Arisarum vulgare*’, ἀρίσαρον, ἄρον, and perhaps ἄσαρον ‘*Asarum europaeum*, hazelwort’ without mention of the Latin forms.

Gigarus has not undergone the expected weakening of medial *a > e* before *r*, so it is indeed probably a loan. There is no indication beyond the testimony of Pseudo-Dioscorides³³⁹ that it is Etruscan. And if it is a loan from Celtic, its bearing on the substrate lexicon of Latin is greatly diminished. In any case, its origin remains unclear.

guttur ‘throat’Pre-form: **gelou-ty-* | PItal. **gūtōr-*

Comp.: ?

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: body part

Pokorny (393-4), WH (I: 629), EM (286), DV (276)

Puhvel (IV: 315)

DV (276) finds it unlikely that the *-ur* of *guttur* preserves a heteroclitic ending because

³³⁸ André (1956: 148) suggests that he trusts neither source, pointing to poorly attested *giger* ‘wild parsnip’ and its similarities to Arabic words for carrot.

³³⁹ Claims like this by Greek authors this should always been taken with a grain of salt. Hesychius calls *capra*, *dea*, and *nepos* Tyrrhenian whereas Dioscorides ascribes *apium*, *spīna*, and *sūcinum*, words of clear Italic origin, to the *Thoῦskoī* (*Tuscī*) (cf. Breyer 1993: 133). It is thus clear that in some cases, they simply meant that these words were used on the Italian peninsula, not specifically by the Etruscans.

we do not know the root lexeme and hints at non-IE origin by comparing other etymologically obscure throat words (*gula*, *glut-* and *gurguliō*). Any link with Hitt. ^(UZU)*kuttar-* ‘strength, force, power; back of the neck, top of the shoulders’ (cf. Pokorny 393-4) can indeed be rejected on formal and semantic grounds (Puhvel IV: 315). But if *guttur* is a *littera* variant of **gūtur* (rather than expressive gemination, cf. EM 286) < **gelou-ty-*, it could preserve a neuter instrument noun found also in MLG *koder*, Ger. dial. *Köderl* ‘throat, gullet’ (cf. Pokorny 393-4) < PGM. **kupra-* < **gu-tro-*. Though peculiar, it cannot be ruled out that this is a chance preservation of an archaic formation. On the other hand, DV (276) notes that the Low German attestation means the root in Germanic could instead derive from **gud^h-*.³⁴⁰ Thus the relationship of the Germanic and Italic forms remains unclear.

hāmus ‘hook, fish-hook’

Pre-form: **g^heh₂m-* | PItal. **χāmo-*

Comp.: ?

Irreg. correspondences

Remarkable phonotactics

Semantics: tool

WH (I: 633), EM (289), DV (279)

EDG (1605, 1613)

Lat. *hāmus* has been compared to Gk. *χαμός* and *χαβός* ‘curved’ (WH I: 633 with lit.) < **g^hh₂m/b-*, where the difference in vowel length rules out a direct loan and the *m/b* alternation within Greek points to a non-IE origin there (DV 279, EDG 1605, 1613). But the semantic match between ‘curved’ and ‘hook’ is too weak to confirm the connection with any certainty.

harundō, -inis ‘reed’

Pre-form: **g^halHr-olund^(h)-* | **χαrundōn-*

Comp.: ?

Irreg. correspondences

Remarkable phonotactics

Semantics: plant, wild; aquatic

Pokorny (68), WH (I: 634-5), EM (289), DV (279)

Čop (1969: 187), FEW (IV: 72), Puhvel (III: 143), EDG (131, 136), Adams (2013: 153), Kroonen (ftbc.)

Lat. *harundō* has been compared to Greek words for plants in the family *Araceae* including ἄρον ‘*Arum italicum*’ and ἀρίσαρον ‘*Arisarum vulgare*’. Along with *arista*

³⁴⁰ OE *cēod* and OHG *kiot* ‘bag’ < PGM. **keuda-* ‘bag’ might be a separate lexeme.

'awn', Pokorny (68) suggests that the family are Mediterranean loans. WH (I: 634 with lit.) note that the connection only works if the *h* of *harundō* is unetymological. They find the link with *arista* unlikely and EDG (131, 136) considers the connection between Latin and Greek unlikely; the plants involved are indeed quite different.³⁴¹ Otherwise DV (279) briefly mentions an (irregular) connection to PGM. **hreuda-* 'reed' and Toch. B *karwa*, Toch. A *kru-* 'reeds' but the latter has a good alternative etymology (Adams 2013: 153) and the former requires setting up a series of irregular alternations. Driessen (apud DV 279) suggests a connection to Gaulish **garunda-* 'shallow water-course, river, river bank' (cf. Prov. *garouno* 'drainage canal' and river names like the Garonne³⁴²), but this requires some semantic leaps. The suffix, found also in *bolunda* and *hirundō* (s.v.), looks like a potential Italic reflex of the Pre-Greek *vθ*-suffix (cf. Kroonen fthc.), but there is no corresponding Greek form. Certain comparanda and thus the origin of *harundō* remain elusive.

hirūdō, -inis 'leech'

Pre-form: **gʰir-* | PItal. **χirūdōn* | **χiruzdōn*

Comp.: **gʰ(h)elir-* | PCelt. **gelirūddo-* | Olr. *giritán* 'edible periwinkle'

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: animal, wild; aquatic

WH (I: 652), EM (296), DV (280, 286)

Deroy (1956b), Breyer (1993: 351-4), Stifter (fthc.)

It seems quite likely that *hirūdō* is related to Olr. *giritán*, Molr. *gioradán* 'edible periwinkle' < **gelirūddo-* (Stifter fthc.). A common pre-form is reconstructible all the way down to the *-ūd(d)o-* suffix. The geminate **dd* of the Celtic reconstruction is suspicious (although it is in a suffix and might not bear on the origin of the root), but can apparently be the result of **zd* (Stifter fthc.). In fact, the sequence **uzd* would also yield the *ūd* of *hirūdō*. Thus *giritán* and *hirūdō* can both reconstruct to identical **gʰiruzd-*.³⁴³ If kept on its own, this Italo-Celtic formation shows no clear signs of being borrowed. But DV (286) supports a comparison with other Latin words that mean 'intestines'. *Haruspex* 'diviner who inspects the internal organs of sacrificial animals' has been suspected to be of Etruscan origin for semantic reasons, but it seems at best to be a calque of the attested Etruscan word *nekʷis* (Deroy 1956b, cf. further Breyer 1993: 351-4 with lit.). The *haru-* can be the reflex of inherited **ǵʰyH-u-* 'intestines' (cf. Skt. *hirā-* 'vein', Lith. *žarnā*

³⁴¹ Based on this connection, Čop (1969: 187) connected Hitt. *hariuzzi-* 'Tisch aus Rohrgeflecht (?)'. Puhvel (III: 143), translating *hariuzzi-* as 'wickerwork table' seems to reject the connection of the Hittite because he disagrees with the comparison of the Latin and Greek material. I cannot tell for certain, but he seems to be making a pun when he calls the pair "one of the weakest reeds in Pokorny's compendium".

³⁴² FEW (IV: 72) alternatively derives these from the Celtic word for 'crane' cf. W, Co., Bret. *garan*.

³⁴³ WH (I: 652) and EM (295) compare the formation of *testūdō* 'tortoise' < *testu-* 'pot', suggesting that *hirūdō* is from another otherwise unclear **hiru-*.

'intestine', DV 280), but Italo-Celtic **hir-* cannot, given its *i*-vocalism. Nor can Lat. *hīra*, another word for intestines of similar shape. It seems attractive to connect these three formations, setting up an irregular vocalic alternation, but it is not certain that they belong together.

īlex, -icis 'holm oak, ilex'

Pre-form: *(H)*īl-elak-* / *(H)*eil-alk-* | PItal. **īlelak-* / **eilelak-*

Comp.: ?(H)*slū**il-ak-* | PGk. **ilak-* | Hsch. ἱλαξ: ἡ πρίνος, ὡς Ῥωμαῖοι καὶ Μακεδόνες, 'holm oak amongst the Romans and Macedonians'

Irreg. correspondences Remarkable phonotactics

Semantics: plant, tree

WH (I: 678), EM (308), DV (298)

Cuny (1909: 21-6), REW (no. 4259), FEW (IV: 545), Wagner (1960-4 I: 487-8), EDG (32)

Cuny (1909: 21-6) interpreted the Romance descendants of Lat. *īlex* (e.g. It. *elce*, Prov. *euze* > Fr. *yeuse*) as attesting to **īlex*, and proposed that it is related to the second element in Gk. αἰγίλωψ 'kind of oak', demonstrating that it is a Mediterranean word. Alternative etymologies of the Greek word exist (EDG 32). WH (I: 678) and EM (308) generally agree with Cuny (1909: 24) in connecting Hesychius's 'Roman and Macedonian' ἱλαξ. Lat. *-ex* could be from **-aks* with vowel weakening, but we know too little about Macedonian to use it to inform us about vowel correspondences. DV (298) does not even mention the form.

While both **ē* and **ī* are possible for West Romance, Logudorese *élige* can only be from **ēlex* (cf. Wagner 1960-4 I: 487-8). Thus it is likely that all the Romance forms go back to **ēlex* rather than **īlex*.³⁴⁴ REW (no. 4259) and FEW (IV: 545) explain PRom. **ēlex* as the Umbrian reflex³⁴⁵ of PItal. **eileks*, which in Latin would have monophthongized to attested *īlex*. This is more plausible than a contamination with *ēligō* 'to choose' (*pace* WH I: 678). Without stronger evidence of a *bona fide* vocalic alternation, the origin of Lat. *īlex* remains unclear.

lanx, -cis 'metal dish, tray'

Pre-form: **l(a)nk-* | PItal. **lank-*

Comp.: ?**lek-* | PGk. **lek-* | Gk. λέκος, λέκις, λεκάνη 'dish, pot, pan'

Irreg. correspondences Remarkable phonotactics

Semantics: vessel

³⁴⁴ Except for Campidanese *ilīzi*, which continues Lat. *īlex*.

³⁴⁵ EM (308) label **ēlex* simply "d'origine dialectale".

Pokorny (307-9), WH (I: 761), EM (340), DV (326)

Schroeder (1930-31: 111), von Soden (1965-81 I: 527), Schrijver (1991: 488-96), Mastrelli (2002), EDG (847, 853)

The vocalism of *lanx* can perhaps have arisen from the complex cluster **lnks* (as per Schrijver 1991: 488-96 on **CaCCC*). Pokorny (307-9) and WH (I: 761) compare it to Gk. λοξός ‘slanting’, λέγχιος ‘slanting, crosswise’, but only the vessel names like λέκος, λέκις, and λεκάνη (var. λακάνη) are plausible. EDG (847) takes the alternation between λεκάνη and λακάνη as indicative of a non-IE origin, despite it elsewhere being taken as a late assimilation (cf. Furnée 1972: 352). DV (326) follows EM (340) in conceiving of the Latin and Greek forms as loans from a Mediterranean language.³⁴⁶ In the end, the nasal of *lanx* is in the wrong place to secure the comparison with the Greek forms beyond a doubt.

larix, -icis ‘larch tree’

Pre-form: **lalHr-* | Pltal. **larik-*

Comp.: ?

Irreg. correspondences

Remarkable phonotactics

Semantics: plant, tree

Pokorny (214-17), WH (I: 765), EM (342), DV (328)

Stokes (1885: 88), Terracini (1921: 409-10), Brück (1923), Alessio (1941b: 221-3), Bottiglioni (1943: 319-20), LEIA (D-12), Trask (2008: 265), Matasović (2016: 704-5), Weiss (2020: 504, fn. 63)

Stokes (1885: 88) suggested that Lat. *larix* ‘larch’ was borrowed from Celtic oak words < PIE **doru-*, specifically OIr. gen. *darach* < **darix*.³⁴⁷ The change **d* to *l* as part of the “Sabine *l*” phenomenon was ruled out because the larch does not grow in historically Sabine areas;³⁴⁸ Brück (1923) proposed that the *l* arose via contamination with *lacrima* ‘pitch/resin’. Others have suspected the mediation of a substrate language (Terracini 1921: 409-10, Bottiglioni 1943: 319-20). But the semantic match is not perfect to begin with (DV 328). Matasović (2016: 704-5) proposes a connection with **ǵ^helh₃-* ‘yellow, green’, which requires transmission through Sabellic and is semantically not compelling. Alessio (1941b: 221-3) proposes an *e* ~ *a* alternation on comparison with Basque *ler* and *leher* ‘pine’, which Trask (2008: 265) suggests is the original Vasconic word for ‘pine’ (elsewhere replaced by loans from Lat. *pīnus*). But *leher* is likely the original form and its medial consonant (which could be from an original **n*, Trask 2008: 25) already makes

³⁴⁶ For a review of the link to Gk. λάγυνος, λάγηνος ‘flask, pitcher’, Hitt. *lahan(n)i* ‘vessel’ and their potential Semitic and Sumerian sources (cf. Schroeder 1930-31: 111, von Soden 1965-81 I: 527), see Mastrelli (2002 with lit.); even the link between the λάγυνος and λακάνη remains uncertain.

³⁴⁷ Its existence as a guttural stem is an innovation within Celtic (LEIA D-12).

³⁴⁸ But note that “Sabine” is a misnomer (Weiss 2020: 504 fn. 63).

it look quite different from *larix*. Thus *larix* remains without certain comparanda.

legūmen ‘pulse, legume, bean’

Pre-form: **leg*(^w)- | PItal. **legūmen*

Comp.: ?

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: plant, domestic

Pokorny (658), WH (I: 781), EM (350), DV (332)

Vaniček (1881: 230), Schwyzer (1913: 196-7), Reichelt (1914: 348-9), Frisk (1960-72 II: 94), Leumann (1977: 103, 370), Puhvel (V: 37-8), EDG (839, 847-8, 871)

The etymology given by Varro, that Lat. *legūmen* is from *legō* ‘to gather, collect’, has been partially accepted since Vaniček (1881: 230, still in DV 332). WH (I: 781) and EM (350) are rightly skeptical of what looks like a folk etymology.

EM (350) suspect a non-native connection to Gk. λέβινθοι ‘ἐρέβινθοι’ and λεβηρίς ‘snakeskin, bean shell’. Even from a substrate perspective, it is difficult to connect the Latin and Greek forms with certainty though. The β of Gk. λέβινθοι could reflect **g*^w, but that of λεβηρίς cannot (unless secondary; **g*^w before *e* yields δ). In *legūmen* a **g*^w would delabialize before *u*. But the form *legarica* mentioned in Varro (*de Re Rustica* 1.32.2) must have **g* (unless secondary; a pre-form with **g*^w would give ***levarica*).³⁴⁹ Thus the Greek forms reconstruct to **b* (Reichelt 1914: 348-9, WH I: 781)³⁵⁰ and the Latin forms reconstruct to **g*.

In meaning, *legūmen* is closer to Gk. λέκιθος ‘gruel of pulse or cereals’. Puhvel (V: 37-8) suggests deriving both (along with Gk. λέκος and λεκάνη ‘dish, pan’) and Hitt. *lak(k)arwant-* ‘poddled leguminous vegetable, legume’ from a PIE root **lek-*. The Hittite word would be an *o*-grade **lókʷ-* while *legūmen* would be from **lekm̥-*, yielding **legumen*, then *legūmen* via tribrach elimination (i.e. metrical lengthening) or analogy to *frūmen*. But the assumed development of **lekm̥* > **legumen* is based on one, very irregular example (*tegimen/tegumen* for *tegmen*, cf. Leumann 1977: 103, 370). Gk. λέκος and λεκάνη at best belong to Lat. *lanx* (s.v.), and EDG (847) suggests that the suffix of λέκιθος makes it Pre-Greek. More likely, if related, Lat. *legūmen* and Gk. λέκιθος represent a *g* - *k* alternation. The appurtenance of the Hittite word, whose meaning I am not convinced can be specified further than ‘edible vegetable’, remains uncertain.

lemurēs ‘evil spirits of the dead’

Pre-form: **lem*-(*u*)*r-* | PItal. **lemolur-*

³⁴⁹ Even if it is potentially Celtic (Varro writes that *alii, ut Gallicani quidam* use this word), *g* reflects **g*^(h).

³⁵⁰ They are probably further related to λοβός ‘lobe, pod’ (Schwyzer 1913: 196-7, Frisk 1960-72 II: 93-4, EDG 867).

Comp.: ?

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: magico-religious

Pokorny (675), WH (I: 781-2), EM (351), DV (333)

Fraenkel (1962-5: 354), Furnée (1972: 216), Schrijver (1991: 218), Breyer (1993: 212-3), EDG (830)

Lat. *lemurēs* is often compared to a family of Greek words including λαμυρός ‘voracious, eager’ and λάμια ‘chasm, a man-eating monster’ (WH I: 781-2, EM 251). The vocalism can be explained as an Italic full-grade against a Greek zero-grade (DV 333, skeptically) or as a non-IE *a* ~ *e* alternation (Furnée 1972: 216). The latter interpretation has led some to suggest a loan from Anatolia or Etruscan (Furnée 1972: 216, DV 333), but this seems unlikely.³⁵¹ EDG (830) instead doubts the connection, and it is indeed semantically difficult to justify. Further comparanda (Lith. *lemóti* ‘to long for’ or Latv. *lamāt* ‘to badmouth, scold’, W *llef*, Bret. *leñv* ‘voice’, cf. Pokorny 675, Schrijver 1991: 218 through a sense like ‘bigmouth’) are semantically equally dubious (DV 333). In the end, Lat. *lemurēs* may well be isolated.

mantum ‘short coat’, var. *mantus*Pre-form: **ma/Hnt-* | PItal. **manto-*Comp.: **ma/Hnd-* | PGk. **mandua-* | Gk. μανδύα ‘a woolen garment’

■ Irreg. correspondences □ Remarkable phonotactics

Semantics: textiles

WH (II: 32-3), EM (385)

Alessio (1950: 45-6), Alessio (1955: 331-2, 560), Furnée (1972: 186), EDG (900)

Furnée (1972: 186) compares Lat. *mantum* ‘short coat’ and Gk. μανδύα ‘woolen garment’, proposing that they demonstrate a non-IE *d* ~ *t* alternation.

EDG (900) calls Gk. μανδύα an unexplained foreign word, with indications by ancient authors that it is from Persian (or Liburnian). Brust (2008: 424) suggests that the word might simply have been known to be foreign, despite the source no longer being known. But there do exist potential Iranian donor forms (like Saka *maṇḍūla-* ‘coat’). Lat. *mantum* ‘short coat’ along with *mantellum* ‘shroud, blanket’ and *mantica* ‘a sack that hangs down on both sides’ on the other hand, are said to be Spanish by Isidore (not rejected by EM 385). Alessio (1950: 45-6) gives as support for an Iberian origin Sp., Pt., Cat. *manto* along with Basque *mantar* ‘shirt’, ‘deck of a boat’ (cf. further Alessio 1955:

³⁵¹ An *a* ~ *e* alternation in Etruscan seems to be the result of umlaut (Breyer 1993: 212-13), so we might expect a variant with *a* in to appear in Latin. Nor does it occur in Greek loans transmitted through Etruscan to Latin. Furthermore, no potential Etruscan source form is attested.

331-2). Trask (2008: 282) however considers Basque *matar* ('gaiter, legging, many other meanings') a loan from Spanish. The other forms are simply reflexes of inherited *mantum*. WH (I: 33) suspects a Celtic origin for the Latin forms.

If Gk. *μανδύα* is actually a Greek word, or at most Liburnian, it can be (irregularly) connected to the Latin words. If it is Persian, it seems less like a substrate alternation and more like chance resemblance.

mergae 'pitchfork'

Pre-form: **h₂merg-* | Pltal. **merg-*

Comp.: ?**h₂merg-* | PGk. **amerg-* | Gk. ἀμέργω 'to pluck (as of flowers)'

Irreg. correspondences Remarkable phonotactics

Semantics: tool

Pokorny (738), WH (II: 76), EM (399), DV (375)
EDG (86)

It is possible that Lat. *mergae* 'pitchfork' (and derived *merges* 'sheaf of grain') is related to Gk. ἀμέργω 'to pluck (as of flowers)' given their reconstructability to a unified pre-form (WH II: 76, DV 375). Further connections with Skt. *marj-*, Av. *marz-* 'to wipe' are considered possible yet difficult by EDG (86) and more uncertain by WH (II: 76) and DV (375). The problem is that the Greek reflex ὁμόργυμι 'to wipe' requires reconstruction of the root with **h₃* as opposed to the **h₂* required by ἀμέργω. The semantics of 'to wipe' and 'to pluck' seem quite distant from each other, and separating them yields two more or less 'tight' proto-forms: **h₃merg-* 'to wipe' and **h₂merg-* 'to pluck'. It thus seems best to keep *mergae* and ἀμέργω separate from the other forms. EM (399) consider the comparison between *mergae* and ἀμέργω possible at best, noting the technical semantics of the Latin word and the fact that the pair is otherwise without an IE etymology. DV (375) mentions the possibility of a loan from non-IE for this reason as well. If ἀμέργω is related, there is nothing non-IE about a root **h₂merg-* beyond its restriction to Italic and Greek.

mūtulus 'corbel, rafter head'

Pre-form: **muHt-* | Pltal. **mūtululo-*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: architecture

WH (II: 138, 139), EM (426, 427), DV (398)
Bertoldi (1936: 309-16), Bertoldi (1942: 156), Alessio (1948-9: 132), Hubschmid (1953: 80), Alessio (1955: 583), Furnée (1972: 218-19), Kortlandt (1981), EDG (987)

Lat. *mūtulus* is an architectural term, referring to decorative features that jut out from ceilings and walls. Its technical semantics leads EM (427) to suggest it must be loaned from Etruscan. Bertoldi (1936: 309-16) builds a case that it continues an Etruscan root **mut-* ‘jutting out’. This involves connecting Lat. *mūtō* ‘penis’, the priapic deity names *Mūtīnus/Mūtīnus*, μούτουκα ‘thyme, Cistus’ (called Etruscan by Pseudo-Dioscorides, attested as Calabrian *mūtaka* ‘*Cistus monspeliensis*’, further forms in Alessio 1948-9: 132), several Etruscan words of the shape *mut/muθ* without known meaning, and several toponyms and personal names. Furnée (1972: 218-19) further adduces to this family Hsch. μυττός: τὸ γυναικεῖον and βύττος: γυναικὸς αἰδοῖον ‘female genitalia’. Bertoldi (1942: 156) and Hubschmid (1953: 80) compare Basque *mutur* ‘extremity, snout’, which Trask (2008: 273) notes alternates with *mustur*. It is unlikely that all of these forms belong together, and of those that do, there is little evidence of Etruscan origin.

Lat. *mūtō* ‘penis’ (cf. DV 398) does not have a secure IE etymology, but there are several compelling options including a comparison with OIr. *moth* ‘penis’.³⁵² Even if it is not inherited, little speaks to a connection with μούτουκα ‘thyme, Cistus’, the only form with any convincing potential Etruscan pedigree. The connection between *mūtulus* ‘corbel’ and *mūtō* ‘penis’ is imaginative at best, but if it holds then it is likewise potentially inherited (and without any evidence of Etruscan origin). Trask (2008: 273) explains Basque *mutur/mustur* as an expressive formation, of the shape *mVCVR* (e.g. *makur* ‘twisted’, *motel* ‘insipid’, *makar* ‘scrawny’, *moker* ‘hard’, *mukur* ‘clumsy’). The Greek forms μυττός and βύττος convincingly show a non-IE *m ~ b* alternation (EDG 987, cf. further examples in Alessio 1955: 583). But their relationship to *mūtō* ‘penis’ and then further *mūtulus* is far from secure.

palātum ‘roof of the mouth; dome, vault’

Pre-form: **plh₂-V-* | PlItal. **palāto-*

Comp.: ?

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: body part / architecture

WH (II: 237), EM (475-6), DV (440)

Pfiffig (1969: 38, 42), Breyer (1993: 292-5)

EM (475-6) and WH (II: 237) mention that *palātum* is etymologically obscure and settle on suggesting Etruscan origin based on Festus’ account: *falae dictae ab altitudine, a falado, quod apud Etruscos significat caelum*.³⁵³ This may well explain Lat. *fala* ‘siege tower’ and its derivatives (cf. Breyer 1993: 292-5 with lit.). But *p ~ f* alternations are not

³⁵² This pair could be due to pre-tonic shortening (Dybo’s Phenomenon, cf. Kortlandt 1981) with **mūH-to-* behind Italic and **muH-tó-* behind Celtic, or “pretonic absence of lengthening” (DV 398, cf. Schrijver 1991: 248-9) with *mHú-to-* behind Italic and **mHu-tó-* behind Celtic.

³⁵³ ‘Siege towers are named from their height, from *falado*, which amongst the Etruscans means ‘sky’.’”

easy to explain even with Etruscan (cf. *ferrum*, s.v.) and the semantic link is tenuous. While Ennius uses *caeli palātum* to mean ‘the vault of the sky’, it is *caeli* that means sky, not *palātum*.³⁵⁴ While parallels exist (cf. Du. *gehemelte* ‘palate’, collective formation to *hemel* ‘sky, heaven’), it simply seems too imaginative to suggest that Ennius, rather than simply wanting to express the concept of the sky as a vault, chose *palātum* because he knew it also meant ‘sky’.

DV (440) proposes an elegant solution: Lat. *palātum* ‘roof of the mouth, dome, vault’ and perhaps *Palātium* ‘the Palatine Hill’ are from IE **p_hh₂-* ‘flat’. This of course requires the assumption of a semantic change ‘flat’ > ‘vaulted’, which is not obvious, but it seems like a better option than Etruscan origin.

palla ‘long outer garment, particularly for women; curtain’

Pre-form: **pa/Hl-d/n/s/y-* | PItal. **pald/n/s/wā*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: textiles

Pokorny (803-4), WH (II: 238-9), EM (476), DV (440)

Matasović (2009: 240), Höfler (2017)

Lat. *palla* is sometimes suspected of being a (Mediterranean) loanword (EM 476, DV 440). Comparanda are difficult to ascertain. The most straightforward reconstruction **p_hl-d/n/s/y-* (DV 440) is not otherwise attested. It can semantically be linked to several known IE roots, but formal problems remain.

A connection with Lat. *pellis* ‘skin, hide’ < **pelni-* (WH II: 238-9 with lit.) is semantically attractive (cf. further Lith. *plėnė* ‘membrane’, OCS *pelena* ‘band for swathing children, PGm. **fella-* ‘membrane, skin, hide’). But the *a* vocalism is difficult to motivate. DV (440) suggests a secondary full-grade in *a*. Höfler (2017) proposes a derivation from a *set*-root **pelH-* ‘to cover’ in Gk. *πέπλος* ‘women’s garment’ < **pé-pl(h₂)-o-*. If an *s*-stem is preserved in U *pelsa-* ‘to bury’, then a derived formation like **p_hh₂-s-eh₂* > **palasā* > **palsā* could be behind Lat. *palla*. This relies on the *palma* rule, which is not universally accepted.

pērō, -ōnis ‘military and work boots made of rawhide’

Pre-form: **pēr-ōn-* | PItal. **pērōn-*

³⁵⁴ WH (II: 237) say it is in imitation of Gk. οὐρανός in its meaning ‘the vault of heaven’, suggesting that Ennius wanted to express more than just ‘sky’ and so added the extra word that meant vault. On a related note, Battisti (1960: 34) and Breyer (1993: 294) mention Hsch. βαλόν· τὸν οὐρανόν. Furnée (1972: 231) considers it Pre-Greek because of the attestation of φάλος ‘part of a helmet, perhaps a protrusion’, but it is not at all clear that these belong together. The connection of the Etruscan word with βαλόν is likewise unclear. None of this bears on the origin of Lat. *palātum* however.

Comp.: ?

- Irreg. correspondences Remarkable phonotactics

Semantics: textiles (leather)

WH (II: 290), EM (499)

Furnée (1972: 151-2), Leumann (1977: 363), EDG (1187)

Comparanda for Lat. *pērō* 'rawhide boot' are unclear. WH (I: 290) and EM (499) suspect it must be connected with Lat. *pēra* 'sack, bag', convincingly from Gk. *πήρα* 'leather bag, knapsack'. Then *pērō* would attest to an unattested Gk. **πήρων* with the unattested meaning 'boot' (Leumann 1977: 363). Other possible alternations within Greek suggest it is not native there: Hsch. *βηρίδες: ὑποδήματα, ἃ ἡμεῖς ἐμβάδας λέγομεν* 'sandals, which we call ἐμβάδες', Hsch. *περι-βᾶρίδες: ὑποδημάτων εἶδος γυναικεῖον* 'women's shoes' (Furnée 1972: 151-2 followed by EDG 1187). Thus it is unclear if Greek must be the direct source. PRom. **barr-* 'small vessel' (Furnée 1972: 152) is probably unrelated.

pūlēium 'pennyroyal'Pre-form: **puH-* | PItal. **pūlje-*

Comp.: ?

- Irreg. correspondences Remarkable phonotactics

Semantics: plant, wild

WH (II: 384-5), EM (544)

REW (no. 6815), Furnée (1972: 152)

Lat. *pūlēium* appears in some manuscripts as *pulegium* and *puledium*, but WH (II: 384-5) suspect that these are secondary. They further doubt any direct connection with *pūlex* 'flea' (cf. also EM 544) as it leaves the suffix unexplained. Furthermore, the length of the *ū* is only confirmed metrically in Martial (*Epigrams* 12.32.19). There it may have been folk etymologically influenced by *pūlex*, given that the Romance languages continue **ū* (REW no. 6815). Further contamination with *pūlex* may have given rise to forms like *pūlicāria* 'fleabane *Plantao indica*',³⁵⁵ which should presumably look more like Gk. *ψόλλιον*, the word it is translating (WH II: 384-5). Beyond this, Furnée (1972: 152) compares it to *πόλιον* 'felty germander (*Teucrium polium*)', which, given the variant *βόλιον* in Pseudo-Dioscorides, is unlikely to be related to *πολιός* 'gray'. This could represent a substrate lexeme; pennyroyal and germander are vaguely similar and are both used medicinally.

rēte 'net', var. *rētis*Pre-form: **HlureH-* | PItal. **rēti-*

³⁵⁵ Given as *Plantago psyllium* in Liddell and Scott.

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: tool

Pokorny (332-3), WH (II: 431), EM (572), DV (521)

Schrijver (1991: 17-8, 314), Rosén (1995), Derksen (2007: 434)

WH (II: 431) and Pokorny (332-3) link *rēte* to *rārus* ‘with wide interstices, far apart, rare’, and further to Gk. ἐρήμιος ‘lonely’ and Baltic forms (Lith. *rė̃tis* ‘sieve’, etc.). Most of the connections fail however. Gk. ἐρήμιος is formally incompatible with *rārus*³⁵⁶ as is *rārus* with *rēte*, and their semantics are not close enough to justify proposing an irregular alternation. The connection between *rēte* and ἐρήμιος is semantically gratuitous (Schrijver 1991: 17-18). DV (521) notes that the Baltic forms (Lith. *rė̃tas* ‘rare, thin, slow’, *rė̃tis* ‘sieve’, *rė̃sti* ‘to become rare’) < BSl. **rēto-* and **ret-* are semantically similar to Slavic forms < BSl. **reʔd-* (cf. OCS *rědъkъ* ‘rare’, Derksen 2007: 434), which perhaps attests to an irregular alternation. But given the semantic difference, Lat. *rēte* is likely unrelated to these either. Given its isolation Rosén (1995) suggested *rēte* could be borrowed from Canaanite **reθt-* ‘net’ (cf. Biblical Hebrew *rešet*). Epenthesis did not occur in roots where the last two consonants were similar or identical. Thus he proposes that **reθt-* (cf. Ugaritic *rθt*) could have entered Latin as such whereupon the **θ* was despirantized (or perhaps the despirantization happened in an intermediary language) yielding **rett*. The latter situation, in which *rēte* is a borrowing from a Mediterranean language that in turn had borrowed the Semitic form seems more plausible, but, without further forms borrowed this way, remains uncertain.

saepēs ‘hedge, fence’

Pre-form: **sh₂eip-* / **seh₂ip-* | PItal. **saipi-*

Comp.: *?*sh₂eim-* | PGk. **haim-* | Gk. αἰμασιά ‘wall around a terrain’,
Hsch. αἰμοί: δρυμοί ‘copses, thickets’

Irreg. correspondences Remarkable phonotactics

Semantics: tool

Pokorny (878), WH (II: 461-2), EM (588), DV (533)

Furnée (1972: 223), EDG (39)

DV (533) would see Lat. *saepēs* as deriving from **sh₂ei-* ‘to bind’ except that no **p-* suffix is known that can produce the derivation. Thus we are left to assume that the full root was **sh₂eip-* or **seh₂ip-*. To get Lat. *saepēs* to match Gk. αἰμασιά, Pokorny (878) tentatively reconstructs **saip-mṛtiā-* (and **saip-mo-* for αἰμοί) with which WH (II:

³⁵⁶ Gk. ἐρήμιος requires **h₁rēh₁-mo-* (cf. Myc. *e-re-mo*, EDG 456), whose ablaut grade would yield Lat. ***rērus*. Zero-grade **h₁rēh₁-ro-* should have given ***rārus* (Schrijver 1991: 17).

461-2) agree, ignoring the fact that the regular outcome of **-pm-* is *-mm-* in Greek. Thus EDG (39) follows Furnée (1972: 223) in taking the *m ~ p* alternation at face value and evidence of non-IE origin. The semantic match between the Latin and Greek forms is not perfect however, and given the additional formal problems it is not clear that they are actually related.

sēcale 'rye'

Pre-form: **seHkAl-* | PItal. **sēkal-*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: plant, domestic

WH (II: 504), EM (607)

Huld (1990: 405)

Reconstructing a PIE pre-form is artificial for this word. The length of the vowels is known from Romance descendants. WH (II: 504) and EM (607) are both convinced it is a loan. Its source is unknown, but it looks suspiciously similar to several Caucasian words including Rutul *sikil*, Tsakhur *sikil* 'rye' and Khinalug *silgli* 'oats' (Huld 1990: 405).

sorbus 'service tree'

Pre-form: **s(o)rb^(h)/d^h-* | PItal. **sorbo-*

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: plant, tree; fruit

Pokorny (910-11), WH (II: 562), EM (636), DV (567)

Furnée (1972: 230), EDG (1373)

Lat. *sorbus* is often connected to Lith. *sā̃tas* 'red, brown (of horses)', Latv. *sārts* 'red, pink' to a root **ser-* 'red, reddish' (WH II: 562 with lit., Pokorny 910-11, EM 636), but the semantics of the root are questionable. Several other Balto-Slavic forms could be related: Ru. *sorobalina* 'rose hip, blackberry', Lith. *serbentā*, *serbeñtas* 'redcurrant, blackcurrant', *sīrbti* 'to ripen'. DV (567) suggests that if they are related, they point to a non-IE **sVrb-* 'berry', but all forms can be reconstructed to IE ablaut grades of an (otherwise unknown) root **serb^h-*. The semantic difference between the Balto-Slavic forms and Lat. *sorbus* makes the link difficult to confirm in any case. *Sorbus* may be without comparanda.

Furnée (1972: 230) followed by EDG (1373) proposes that Hsch. σπόρα παλι[v]ούρου εἶδος 'a kind of Christ's thorn (*Paliurus spina-christi*)' is meant as a spelling of *σopρα

producing a *w* ~ *b* alternation with *sorbus*. But these trees have very little in common.

spiōnia ‘a sort of grapevine’, var. *spīnea*

Pre-form: **spiH-(i)ōn-* / **spī-(i)ōn-* | PItal. **spi/ī(i)ōniā*

Comp.: ?

Irreg. correspondences

Remarkable phonotactics

Semantics: viticulture

WH (II: 575), EM (642)

Pedersen (1909-13 I: 68)

WH (II: 575) mention that, like *acinus*, *spiōnia* could descend from a Mediterranean-Aegean language, but this is solely due to its viticultural semantics and lack of a better etymology. Pedersen (1909-13 I: 68) compares Celtic forms like Mlr. *sían* and W *ffion* ‘purple foxglove’. They reconstruct to PCelt. **sφī(i)on-*, practically identical to the pre-form of Latin. But the semantic distance is large, and it remains unclear if *spiōnia* has any comparanda.

sūber ‘cork oak’

Pre-form: **suHb^(h)-* | PItal. **sūb/fer-*

Comp.: ?

Irreg. correspondences

Remarkable phonotactics

Semantics: plant, tree

WH (II: 617), EM (661), DV (595)

Cuny (1910: 158), EDG (1425)

Lat. *sūber* ‘cork oak’ is compared to Gk. *σῦφαρ* ‘wrinkled skin; old person; milkskin’. A borrowing from Greek into Latin should have yielded ***sūpar*, and the initial *s* of Greek rules out a reconstruction to a common root **suHb^(h)-*. If they are related, they are not of IE origin (Cuny 1910: 158, WH II: 617, EM 661DV 595, EDG 1425). But the semantic difference is too great to secure the comparison and assume an irregular alternation.

tamīnia ‘a common plant amongst hedges with red berries, black bryony (*Dioscorea communis*)’ or ‘a type of wild grape’

Pre-form: **ta/Hm-* | PItal. **tam-*

Comp.: ?

Irreg. correspondences

Remarkable phonotactics

Semantics: plant, wild / viticulture

Pokorny (1063), WH (II: 646-7, 657), EM (676, 679-80)

Bertoldi (1942: 165, 169-70), Alessio (1944b: 414), Ernout (1946: 35), Alessio (1948-9: 135), Battisti (1960: 373), Hubschmid (1960a: 63), Furnée (1972: 200), Breyer (1993: 390-2), EDG (533)

Lat. *tamīnia* ‘black bryony’ or ‘a type of wild grape’ is generally taken as related to *tamnus* ‘the vine of the *taminia*’ or ‘black bryony’ or ‘wine made from *taminia*’ (WH II: 646-7, uncertainly EM 676, etc.). But the exact definitions differ amongst scholars.

Bertoldi (1942: 165) and Alessio (1944b: 414) further connect these forms with Lat. *tēmētum*, which they define as ‘a kind of rustic wine made of wild grapes’, thus showing that the stem alternates between **tam-* and **tem-* with a Mediterranean *a ~ e* alternation. *Tēmētum* is more traditionally defined as ‘intoxicating liquor’ (cf. Lat. *abstēmius* ‘abstaining from wine’), and seems to have an IE etymology (**tēmH-* cf. Arm. *ʿmrim* ‘to become stunned’, Ger. *dämisch, dämlich* ‘stupid’; Skt. *tāmyati* ‘to be dazed’ has secondary **ā*, Pokorny 1063, WH II: 657, DV 609; EM 679-80 finds the connection arbitrary).

Ernout (1946: 35) proposes that the pair *taminia, tamnus* are of Etruscan origin due to the suffixes *-mmo-* and *-mnia-/mina-*. Bertoldi (1942: 169) follows because of some toponyms and the attested Etruscan forms *tannia* and *taminai*. Breyer (1993: 392) discusses different concatenations of Etruscan morphology that could result in the Latin forms, but she continues to work with the assumption that *tēmētum* is related. As the Etruscan look-alike forms are of unknown meaning, there is no solid evidence of an Etruscan origin for the Latin words. Tuscan *tamaro, tamarro* ‘*Dioscorea communis*’, etc. seem to attest to a root form with an *r* instead of *n*, which Alessio (1948-9: 135) proposes is either due to dissimilation or Etruscan *r* for *n* replacement like in *Memrun* for Μέμρων and (proposed) **cruma*³⁵⁷ for γρόμα ‘mark’ — further evidence that the lexeme is Etruscan (cf. also Alessio 1944b: 414, Battisti 1960: 373). The Etruscan proposal would seem to indicate however that the word is *not* Etruscan. All examples show a change from *n* > Etr. *r*, not the other way around. Etruscan origin would not explain *taminia*. Hubschmid (1960a: 63) takes the form with *r* as a Mediterranean substrate variant and suggests connecting Lat. *tamarīx* ‘tamarisk’ and its Romance descendants. This is a different plant however, and so the connection is not secure.

Hubschmid (1960a: 63) does mention several irregular looking descendants of the Latin words. These include Bergamo *tam* < **tamus*, lacking the *n* as well as Istrian *dāmi* with initial *d* instead of *t*. This situation is reminiscent of Lat. *talpa* ‘mole’ versus PRom. **darbo-* (s.v. *talpa*), but the irregular forms are not nearly as widespread as with that case. To *tamnus* Furnée (1972: 200), following Alessio (1944b: 414), further compares θάμνος ‘bush, shrub’. He is dissatisfied with the IE etymologies proposed for the word and argues that it belongs with forms like θαμά ‘often’ as a Pre-Greek lexeme. If the

³⁵⁷ Argued to be the source of Lat. *grōma, grūma, crōma* ‘field surveying instrument’.

form $\theta\acute{\alpha}\mu\nu\eta$ ³⁵⁸ really does mean ‘wine from pressed grapes’, it would be closer in meaning to *tamnus*; but still not quite the same. EDG (533) rejects the connection with *tamnus*, but it is not clear exactly why. “With its ending in $-\acute{\alpha}\mu\nu(\omicron\varsigma)$, the word seems Pre-Greek; its meaning makes this quite possible.” It seems they reject it in part because, as Pre-Greek was spoken in Greece only, it should not have comparanda outside of Greece. The second part of the rejection stands, however. “Bush” and the grape or bryony vine are not similar enough meanings to compare. Lat. *tamīnia* remains without certain comparanda to elucidate its etymology.

unēdō ‘strawberry tree and its fruit (*Arbutus unedo*)’

Pre-form: **un-eh₁d-ōn-* | PItal. **unēdōn-*

Comp.: *(H)*oHI-id-ōn-* | PRom. **ōlidōn-* | Sard. (*o*)*liðōne*, etc. ‘*A. unedo*’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: tree, wild; fruit

WH (II: 818), EM (747)

REW (no. 9068), Wagner (1960-4 II: 185)

Pliny famously folk etymologized the word: *pomum inhonorum, ut cui nomen ex argumento sit unum tantum edendi* (*Nat.Hist.* 15: 98).³⁵⁹ But the fruit of *A. unedo* is not particularly awful in odor or taste, and the *u* of *unēdō* is short as opposed to *ūnum*. The Romance forms attest to a very different form (Meyer-Lübke 1911 no. 9068): Piedmontese (*Iurion* < **ūlidone*, Saintongeais *olon* < **ol(id)one*, Gascon (Landes) *auledun* < **olidone*, Guyenne *leduno* < *(*o*)*lidone*. Sardinian has (*o*)*liðōne* < **ōlidone* (Wagner 1960-4 II: 185). Between the Romance forms and Latin, there seem to be several vocalic and consonant alternations. However, given the aberrance of Latin alone and the folk etymology given by Pliny, it cannot be ruled out that the Latin word has been deformed somehow.

2.3.3 Conflicting Possibilities

2.3.3.1 Non-inherited vs. Inherited

acer ‘maple tree’

Pre-form: **h₂ek-r-i/o-* | PItal. **akri/o-*

Comp.: **h₂ek-r-no-* | PGm. **ahurna-* | OHG *ahurn*, *ahorn*, *acharn* ‘maple tree’

³⁵⁸ It is attested once, in the *Geoponica* 6.13.2, where it is recorded that, after draining the must from the winepress, the remains are put into casks and used to make inferior wine “which provincially they call *thamna*” (translation from Owen 1805: 209). The passage is attributed to Anatolius, who might be the same as the 4th century author Vindonius/Vindanionius Anatolius of Beirut. But this is uncertain. Otherwise, the *Geoponica* was compiled in the 10th century.

³⁵⁹ ‘[A] dishonorable apple, such that its name is from the evidence of only eating one’.

**h₂ek-r-* | PGm. **ah(i)ra-* | ODan. *ær*, Upper German *Acher* ‘maple tree’

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, tree

Pokorny (18-22), WH (I: 6-7), EM (6), DV (21)

Hubschmid (1953: 80-2), Furnée (1972: 343, 371), Puhvel (III: 304-5), Schrijver (1991: 37-8), Trask (2008: 115), EDG (50), Kroonen (2013: 7)

Lat. *acer* and the Germanic forms can be reconstructed to the same root **h₂ek-*, and it cannot be excluded that this is simply **h₂ek-* ‘sharp’ (Pokorny 18-22, WH I: 7, Schrijver 1991: 37-8) named after the shape of the leaves. Some Germanic forms show an additional *n*-suffix, which might be the one that occurs in substrate words (see §3.3.4).

Further comparanda that could support a substrate origin are uncertain. Hsch. ἄκαρνα δάφνη ‘laurel-tree’ is a formal match for the Germanic even down to the *n*-suffix (Schrijver 1991: 37), but is semantically aberrant leading EDG (49) to consider it isolated within Greek. Hsch. ἄκατος ἢ σφένδαμνος ‘maple’ is often compared under the assumption that it lost its **r* from *ἄκαρ-στος (WH I: 7, DV 21). Perhaps the *r* was never there however: cf. Gk. κάστων ‘wood’, Basque *gastigar* ‘maple’ Furnée (1972: 343, 371), Nuorese *kóstike*, Logudorese *kóstige*, Languedocien and Prov. *agast*, etc. ‘maple’³⁶⁰ (Hubschmid 1953: 80-2, though he connects them to Lat. *aesculus*). The Greek and Basque forms without *r* might represent a separate lexeme. Puhvel (III: 304-5) adduces Hitt. *hiqqar-* ‘name of a tree, perhaps maple’, but that it might mean ‘maple’ is based in part on the fact that it is attested as being used to make tables. Nor do the formal details work very well.

apex ‘top, point; (part of) a priest’s hat’

Pre-form: **h₂ep-ek-* | PItal. **apek-*

Comp.: ?

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: architecture / textiles

Pokorny (50-1), WH (I: 57), EM (38), DV (46)

Breyer (1993: 333-4)

Because of its technical use in architecture as well as its referring to an article of priestly attire, Lat. *apex* has been suggested to be of Etruscan origin (e.g. EM 38, cf. Breyer 1993: 333-4 with lit.). But there exists no non-onomastic Etruscan word of known meaning to compare it to. The word can theoretically be PIE, from **h₂ep-* ‘away’, cf. Gk. ἄπιος ‘far off’, Skt. *ápāra-* ‘next, further, more to the back’ (DV 46) or related to *apiō*

³⁶⁰ The Basque Linguistic Atlas (EHHA, map 468) lists several variants of a word for *Acer campestre*: *askar*, *astiar*, *astiger*, etc.

‘to tie, bind’ (Pokorny 50-1, WH I: 57), but neither seems like a perfect fit. The *-ex* suffix is often found on words of murky etymology (DV 46), but it occurs on inherited bases too (cf. *vertex* ‘whirl, eddy’). Lat. *apex* is either of IE origin or it is isolated.

cancer, -t̄ ‘crab; cancer’

Pre-form: **kan-kr-* / *?*kar-kr-* | PItal. *kankro-*

Comp.: *?*kr-kr-ino-* | PGk. *?*karkrino-* | Gk. *καρκίνος* ‘crab; ulcer’

■ Irreg. correspondences □ Remarkable phonotactics

Semantics: animal, wild; aquatic

Pokorny (531-2), WH (I: 151), EM (91-2), DV (86)

KEWA (I: 169), Furnée (1972: 129-30), Vycichl (1983: 246-7), Schrijver (1991: 428), EWAia (III: 64), EDG (646), Meiser (2010: 127)

The etymological explanations for Lat. *cancer* ‘crab’ cannot be accepted as certain. Most frequently, it is assumed that PItal. **kankro-* was dissimilated from **karkro-* (WH I: 151, EM 91, DV 87). However this would be the only example of such dissimilation in Latin.³⁶¹ WH and EM take this as a dissimilation that would already have occurred at an Indo-European date, to a reduplication of the root **ker-* ‘hard’ (cf. also Pokorny 531-2).³⁶² It cannot have occurred in PIE if **kr-kr-* is also the root behind Gk. *καρκίνος*. There the dissimilation has either occurred differently or, if it is from **karkrino-*, not at all (cf. Schrijver 1991: 428, EDG 646). The root **ker-* in question has poor evidence to support it (s.v. *carīna*).

Further evidence of an inherited word is often given as Skt. *karkaṭa-* ‘crab’, if a Middle Indic development of original **karkṛta-* (cf. Schrijver 1991: 428), but EWAia (III: 64) argues that it is unlikely to be inherited (cf. further KEWA I: 169). Vycichl (1983: 246-7) mentions Egyptian Arabic *karkand* ‘crayfish’ as a potential comparandum for the Sanskrit if it is not inherited. EDG (646) still connects the Sanskrit word as a comparandum for Gk. *καρκίνος* because Furnée (1972: 129-30) demonstrated a *k ~ kh* alternation through Hsch. *κάρχαι· καρκίνοι, καὶ <κ>όγχοι. Σικελοί* ‘crabs and snails amongst the Sicels’, making it non-inherited.³⁶³ If the best comparandum for the Latin word itself might be of non-IE origin, then the Latin would be as well. The details are not sufficiently clear and the number of assumptions too high to accept either an inherited or a substrate origin.

³⁶¹ The opposite development, by which **n* has become *r*, is found in *carmen* < **kan-men-* (cf. Meiser 2010: 127).

³⁶² DV (86) favors dissimilation from **karkros* ‘enclosure’ (cf. *carcer*) and reconstructs **kr-kr-o-* ‘circular’ because of the ring formed by the pincers.

³⁶³ Furnée (1972: 129-30) also happens to doubt the appurtenance of Lat. *cancer* here and instead considers it to be from another non-inherited lexeme in alternation with *γάγγραινα* ‘gangrene, flesh-eating illness’. There is no reason to separate *cancer* from the other words meaning ‘crab’ if only to attach it to another word for which an origin meaning ‘crab’ must be theorized.

capiō ‘to take’

Pre-form: **ka/h₂p-i-* | PItal. **kapi-*

Comp.: **ka/o/h₂p-* | PGM. **habēn-* | Go. *haban* ‘to have’, etc.

**ka/o/h₂p-i-* | PGM. **haf/bjan-* | Go. *hafjan* ‘to heave, lift’, etc.

**ka/h₂p-i-* | PGk. **kapy-* | Gk. κάπτω ‘to gulp down’

**koh₂p-* | PGk. **kōp-* | Gk. κοπή ‘grip’

**ka/o/h₂p-* | Alb. *kap* ‘to grab, seize, reach’

**g^ha/Hb^(h)-* | PItal. **hab/f-* | Lat. *habēō* ‘to have, possess’

**g^ha/Hb^h-i-* | PCelt. **gab-yo-* | OIr. *gaibid* ‘to take, hold’

**g^(h)ā/ōb^(h)-* | PSlav. **gabati-* | Ukr. *hábaty* ‘to seize’, Sln. *gábati* ‘to be in need, starve, be lost, die’, etc.

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: action

Pokorny (407-9, 527-8), WH (I: 159-60, 630-1), EM (95-7, 287-8), DV (89, 277)

Collitz (1912: 86-8), Lehmann (1986: 167), Gysseling (1987: 60), Schrijver (1991: 92-3), EWAia (I: 463-4), Kortlandt (1992: 104), Demiraj (1997: 212-3), Boutkan & Kossmann (1999: 89, fn. 3), Untermann (2000: 311-16), Schrijver (2003: 67), Boutkan & Siebinga (2005: 155), Derksen (2007: 159), Matasović (2009: 148), EDG (640, 815), Kroonen (2013: 173, 210), Derksen (2014 s.v. *gebéti*)

If kept separate, **kh₂p-* and **g^hHb^h-* are two independent roots whose IE origin is difficult to rule out. But an idea is in circulation that they represent variants of a substrate lexeme, with an alternation similar to that between Lat. *caper* and OIr. *gabor* (Gysseling 1987: 60, Boutkan and Kossmann 1999: 89, fn. 3, Boutkan & Siebinga 2005: 155, DV 89, 277). Since Collitz (1912: 86-8), an alternative idea, that the two originally separate roots have contaminated each other, has been in circulation (supported e.g. recently in Lehmann 1986: 167, Untermann 2000: 313).

There are some problems with the reconstruction, especially of the *habēō* comparanda (Osc. *hafiest* [3sg.fut.] points to **g^hHb^h-* but U **habian** [3sg.pres.subj.] to **g^hHb-*, Untermann 2000: 313-16; Balto-Slavic forms do not support a reconstruction with a laryngeal, Derksen 2007: 159, Derksen 2014 s.v. *gebéti*), but factors that speak in favor of an inherited origin include: 1) Italic, Germanic, and Greek forms all reconstruct to the same *i*-stem present **kHp-i-* (Schrijver 2003: 67, Kroonen 2013: 198). 2) Italic and Germanic would attest to doublets of this root, which is difficult to explain in a borrowing scenario. 3) The root might also be present in PIIr. **gab^ha-* < **g^hab^ha-* (cf. Skt. *gábhasti-* ‘hand, forearm’, though it suggests original *a*-vocalism unless from **g^hHeb^h-*) giving it a very broad, IE-looking distribution. Thus, while an interesting idea,

the evidence does not seem strong enough to securely assign these roots a non-IE origin.

cicer, -eris ‘chickpea’

Pre-form: **ki-ker-* | PItal. **kiker-*

Comp.: **ke/oi-ker-n-* | PArm. **sēsern-* | Arm. *siserñ* ‘chickpea’

**(ki-)ker-* | PAlb. **θier-* | Alb. *thjer* (vars. *thierr, thjérr*, etc.) ‘lentil’

□ Irreg. correspondences

■ Remarkable phonotactics

Semantics: plant, domestic

Pokorny (598), WH (I: 212), EM (119), DV (113)

Strömberg (1937: 50), Alessio (1943: 233), Hubschmid (1953: 114-15), Berger (1956: 4-8), Battisti (1960: 380), Chantraine (1968-80: 585), Ačařyan (1971-79 IV: 218), Schmalstieg (1976: 264), André (1978: 80), Greppin (1981: 6), Tischler (1983: 570), Jahowkyan (1987: 49, 601, 612), Demiraj (1997: 398-9), Orel (1998: 479), Beekes (2000: 29) Martirosyan (2009: 576), EDG (781, 1684), Mikić and Vishnyakova (2012: 220), Zohary, Hopf & Weiss (2012: 87-8), EWA (V: 510), Cunningham (2018-20 II: 604)

Lat. *cicer* is a neuter *r*-stem closest in form and meaning to Arm. *siserñ* ‘chickpea’. The unsyncopeated *i* in Armenian should technically go back to **ē* < **ei/oi*, producing disyllabic **ke/oi-ker-n-* as a pre-form (Ačařyan 1971-79 IV: 218, DV 113). OPr. *keckers* ‘chickpea’ points to a root without palatovelars (EM 119, Beekes 2000: 29) and Alb. *thjer* looks like it lacks the reduplicated syllable (Orel 1998: 479). Thus Jahowkyan (1987: 49, followed by Martirosyan 2009: 576) takes it as non-IE, with Clackson (1994: 143) even calling it Mediterranean. Alessio (1943: 233) includes it as an example of a word showing Mediterranean substrate reduplication.

But several of the irregularities can be explained. Arm. *siserñ* seems to follow a normal pattern of reduplication in Armenian (though generally in the semantic category of animals and expressive words, not in plant names), where the *i* of the reduplicated syllable was immune to syncope (p.c. Rasmus Thorsø). Greppin (1981: 6) for instance reconstructs *i*-vocalism for the first syllable. OPr. *keckers* is a borrowing from German (Schmalstieg 1976: 264 with lit.); cf. OLG *kekera*, a loan from Latin (EWA V: 510). Demiraj (1997: 398-9) supports a pre-form **ki-ker-* for the Albanian, perhaps through dissimilation (cf. the variant *thírqe*).

Other comparanda are uncertain. The codex unicus of Hesychius gives as Macedonian κίβερροι· ὄχροι, the latter word meaning ‘pale’. But with two emendations of the text (followed by Cunningham 2018-20 II: 604, EDG 1684), we get κίκερροι· ὄχροι with the latter word meaning ‘Cyprus vetch (*Lathyrus ochrus*)’. Beekes (2000: 29) reconstructs the amended κίκερροι to **ki-kerjo-*. Gk. κρῖός ‘ram’ but also ‘chickpea’ is sometimes reconstructed to *(*ki*)*krio-* (Pokorny 598, WH I: 212, Beekes 2000: 29), but the iota is

long. Chantraine (1968-80: 585) and EDG (781) therefore follow Strömberg (1937: 50) in assuming that ‘ram’ is the original meaning, with ‘chickpea’ being metaphorical after the curved shape of the pods.³⁶⁴ Neumann (*apud* Tischler 1983: 570) notes the similar shape of hapax Hitt. *kikri-*, but all that is known of its meaning is that it occurs as a modifier of BA.BA.ZA ‘porridge’. Thus the connection is too uncertain. Most uncertain are several North (-west and -east) Caucasian forms cited by Jahowkyan (1987: 601, 612). Several of the Dagestanian forms especially look similar to the (unreduplicated) base of *cicer* (cf. Akusha, Chiragh, Dargi *qara*, Aghul *xur*, etc. ‘peas’, updated via Mikić and Vishnyakova 2012: 220). This is similar to what Hubschmid (1953: 114-15) and Battisti (1960: 380) purported to find (a root **gar-* / **ger(g)-*) behind Romance forms, Hsch. γάλινοι and γέλινοι ‘ἐρέβινοι’, Georg. *gorwela* ‘type of pea’, and Burushaski *gark* ‘peas’.³⁶⁵

André (1978: 80) is unsure whether to consider *cicer* reduplicated. But the most securely related forms can be reconstructed to **ki-ker-*, avoiding an invalid **C_ieC_i-* root structure or a disyllabic root. The root itself has been proposed to be **kerh₃-* ‘to feed’ (WH I: 212, DV 113), though this need not be the case. Methodologically, the ability to reconstruct a common pre-form generally means an inherited origin cannot be ruled out. But a reduplicated noun formation like this should be archaic, and there is unlikely to have been an Indo-European word for the chickpea, a crop domesticated in Anatolia (Zohary, Hopf & Weiss 2012: 87-8). It remains unclear whether the geographically and formally more disparate comparanda, whose exact relationship to the *cicer* group is unknown, provide positive evidence in favor of a non-IE origin.

cubō, -āre; -cubō, -ere ‘to lay down, recline’

Pre-form: **kub^(h)(H)-* | PItal. **kubāje/o-* | *cubō, -āre*
**kub^(h)-n-h₂-* / **ku-m-b^(h)-* | PItal. **kumbe/o-* | *-cumbō, -ere*
kub(h₂)-* | PItal. **kuba-* | SPic. **qumat, Fal. **cupat** [3sg.pres.] ‘lies’, etc.

Comp.: **kub^(h)-* | PCelt. **kuφ-ske/o-* | MW *kyscu*, MCo. *koska*, MBret. *cousquet* ‘to sleep’

□ Irreg. correspondences

■ Remarkable phonotactics

Semantics: action

Pokorny (588-92), WH (I: 298), EM (153-4), DV (152)

Schumacher (2004: 424), Bakum (2009 I: 78), Matasović (2009: 228), LIV (s.v. *?(^k)eubh₂-*)

The root of *cubō* can be reconstructed as *sef*,³⁶⁶ but need not be (LIV s.v. *?(^k)eubh₂-* with

³⁶⁴ Chantraine notes Lat. *cicer arietinum*, a type of legume, which is an apt parallel.

³⁶⁵ Cf. Berger (1956: 4-8) for its reconstruction to **kiker*.

³⁶⁶ The LIV suggests *?(^k)eubh₂-* only works for Italic, but Schumacher (2004: 424) asserts that the loss of the laryngeal in this sequence is expected in pre-Proto-Celtic.

lit.). Instead, it is the nasal infix present of *-cumbō* that makes the verb look to be of IE pedigree. A non-IE *b ~ mb* alternation is present in e.g. Lat. *sabūcus ~ sambūcus*, but the preservation of both variants makes it look like it entered Latin quite late. Thus such an explanation does not seem to be able to supplant the assumption of native nasal infixation for *-cumbō*. Still, the Faliscan and Sabellic forms require the reconstruction of **b* (WH I: 298), a rare phoneme in PIE.³⁶⁷ Additionally, the root is restricted to Italo-Celtic.³⁶⁸ Thus DV (152) is not certain if it should be considered of PIE origin (cf. *badius* and *bāca*, s.v.).

dulcis ‘sweet’

Pre-form: **d̥k-* / **dulk-* | PItal. **dolki-* / **dulki-*

Comp.: **dluku-* | PGk. **dluku-* | Gk. γλυκύς ‘sweet’

■ Irreg. correspondences □ Remarkable phonotactics

Semantics: culinary

Pokorny (222), WH (I: 379-80), EM (186-7), DV (182)

Stanley (1982: 578), Sihler (1995: 96), EDG (277), Weiss (2020: 335 fn. 66)

Reconstructing a common pre-form for the Latin and the Greek forms is impossible. Lat. *dulcis* can reconstruct to **d̥k-*, but this would not give Gk. γλυκ-. Myc. *de-re-u-ko* (probably corresponding to later γλεῦκος ‘grape must/sweet wine’, cf. Stanley 1982: 578), if indeed a related lexeme, shows that the γλ of γλυκύς is from earlier **δλ* (cf. EDG 277). But this sound change is otherwise unparalleled in Greek.³⁶⁹ WH (I: 379-80 with lit.) suggest it was triggered by assimilation to the κ, but this and their proposal that **d̥* > λν because of υ in the next syllable are *ad hoc* (EM 187, DV 182). From the other side, the Greek reconstructs to a *u*-stem **dluku-*. But as Latin turned *u*-stems into *i*-stems, it should have given ***dulquis* (with the assumption of metathesis; **dluku-* should actually have given ***luquis*, Weiss 2020: 335 fn. 66). Thus perhaps it does not descend from the *u*-stem, or perhaps one of these irregular developments indeed occurred. Otherwise, given that a root **d̥k-/ *dluk-* is isolated to Latin and Greek, DV (182) proposes that both words were borrowed from a third, unknown source.

falx, -cis ‘sickle’

Pre-form: **dʰalHlk-* | PItal. *ɸalk-*

Comp.: ?

³⁶⁷ Faliscan did not have an orthographic *b* and represented it with <p> (Bakkum 2009 I: 78).

³⁶⁸ There is no reason to reconstruct **skeubʰ-* to make up for an illegal root structure and link it with PGm. **skeubanan-* ‘to throw’ as suggested by Matasović (2009: 228), despite the interesting parallel between *iacere* ‘to throw’ and *iacēre* ‘to lie’.

³⁶⁹ Sometimes proposed in the pre-form of Lat. *lac*, Gk. γάλα ‘milk’ (cf. Sihler 1995: 96); otherwise in γέλυγλις ‘garlic’ if ultimately from Akk. *gidlu* (Kroonen 2012b).

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: tool

Pokorny (247), WH (I: 449-50), EM (214), DV (200)

Mikkola (1899: 74), Niedermann (1918: 17-36), Gamillscheg (1920: 517-18), Brück (1921: 583-4), Gamillscheg (1922: 86-9), REW (no. 2458, 2458), FEW (III: 2-3), Alessio (1946a: 165), Fraenkel (1962-5: 81), Matasović (2009: 94), Derksen (2014 s.v. *dilgē, dilgti*), Smoczyński (2018: 193, 229), Weiss (2020: 178)

Lat. *falx* 'scythe, sickle' reconstructs to an invalid *D^heT* root structure, but is similar in form and meaning to two different groups of inherited lexemes. Its *a*-vocalism is difficult to account for.

On the one hand, it is formally most similar to reflexes of a root **d^helg-* (PCelt. **delgo-* 'pin, needle', ON *dálkr* 'pin, dagger', Lith. *dilgti* 'to sting', cf. Matasović 2009: 94, Derksen 2014 s.v. *dilgti*). On the other hand, it semantically more similar to Lith. *dālgis* 'scythe' (Mikkola 1899: 74, Alessio 1946a: 165), generally reconstructed to **dolg^h-* to a root **delg^h-* 'to hew, split' (cf. ON *telgja* 'to hew, cut short' < **dolg^h-eje-*, OIr. *dluige* 'splitting', cf. Derksen 2014 s.v. *dalgis*, Smoczyński 2018: 193). WH (I: 450) are thus suspicious, since this seems to represent a Baltic semantic development of a root with an originally broader meaning. However, Lith. *dālgis* has been reconstructed to the other root, **d^helg-* (cf. Fraenkel 1962-5: 81) perhaps via *métatonie douce*.³⁷⁰

In any case, even a derivation from a root **d^helg-* cannot explain the *a*-vocalism or **k* of PItal. **palk-*. Thus Niedermann (1918: 17-36) made use of purportedly Sicilian ζάγκλη 'sickle'³⁷¹ and Hsch. δάγκλον δρέπανον 'sickle' to suggest that some pre-Italic but still IE language provided a form **dalkla-* (< **dal-tla-*) that entered Latin as ***falcula*. This would have been interpreted as a diminutive whence *falx* was back-formed. Thus the velar element of *falx* would be part of the suffix, removing it from comparison with any of the inherited forms mentioned. DV (200) instead supports that *falx* is indeed a reflex of **d^helg-*, but transmitted through "a non-Latin Indo-European language of Italy". (Cf. a potentially similar scenario for the Greek and Armenian comparanda of Lat. *hordeum* 'barley'.) Given the semantics, this is an attractive hypothesis, but non-IE origin also cannot be ruled out (cf. EM 214).

Several Romance forms (OProv. *dalh*, MFr. *dail*, etc. 'sickle') reflect **dacula*. REW (no. 2456, 2458) considers it the diminutive of **daca* 'dagger, Dacian knife'. Gamillscheg (1920: 517-18) instead proposed that the forms represent the Gaulish reflex of PCelt. **delgo-*, with *l* palatalized through its position in front of *g*,³⁷² a development that Brück

³⁷⁰ Smoczyński (2018: 229) argues the opposite, that Lith. *dilgti* 'to sting' is actually from **delg^h-* with secondary acute accent.

³⁷¹ Thucydides 6.4 says the Sicilian town of Zancle was named after the shape of its harbor (EDG 495).

³⁷² He further considered this to be the source of forms with an *r* in more southerly dialects, but Brück (1921: 583-4) and FEW (III: 2-3) show that the forms reach as far North as Picard *dard* and also mean 'spear', so that the cases where these forms mean 'sickle' is due to contamination.

(1921: 583-4) argued is unparalleled. Gamillscheg (1922: 86-9) then proposed that the hypothetical **ðalkla-* behind ***falcula* could have given **daklo-* via dissimilation and then produced *dail* regularly. Matasović (2009: 94) follows, proposing that PCelt. **delgo-* could represent a different metathesis. This would seem to favor the existence of ***falcula* from which *falx* was back-formed, but it is not the only option. PRom. **dācula-* can alternatively represent **d(e)h₂-tleh₂-* < **deh₂-* ‘to cut off’; cf. Skt. *dātra-* ‘sickle’ (Guus Kroonen, p.c.), thus an unrelated lexeme.

fax, -cis ‘torch, a light’

Pre-form: **ǵʰuok^(w)-* | PItal. **χwok-*

Comp.: **ǵʰuok^(w)-* | PBalt. **žvakijā-* | Lith. *žvākė* ‘candle, icicle’

Irreg. correspondences

Remarkable phonotactics

Semantics: tool

Pokorny (495), WH (I: 438), EM (209), DV (207)

Budenz (1859: 289), Grassmann (1863: 88), Schwyzer (1939-50I: 302), Schrijver (1991: 464-5), EDG (1143, 1551, 1603) Derksen (2014 s.v. *žvakė*), Kroonen (2017: 106), Smoczyński (2018: 1762-3)

Lat. *fax* ‘torch’ is difficult to analyze. An early idea was a connection to *foveō* ‘to make/keep warm’ (cf. Budenz 1859: 289) via **dʰog^{wh}-s-* (cf. *nix, nivis* ‘snow’), but this cannot account for the Latin *a*-vocalism. Instead, a compelling comparison is with (isolated) Lith. *žvākė* (WH I: 438, Pokorny 495, DV 207). Both can be reconstructed to the same pre-form **ǵʰuok^(w)-* (DV 207, Derksen 2014 s.v. *žvakė*), with **o* unrounding to *a* in Latin in an open syllable after **u* (Schrijver 1991: 464-5). Original *a*-vocalism is also a possibility (Smoczyński 2018: 1763, Weiss 2020: 280). Hsch. φῶψ· φάος ‘light’ as if < **ǵʰuōk^s-* along with Hsch. διαφάσσειν· διαφαίνειν ‘to show through’ have been compared (Schwyzer 1939-50 I: 302 gives them as an example of Gk. φ < **ǵʰu-*), but φῶψ might represent a remodeling of φῶς ‘light’ (on ᾠψ ‘eye’, EDG 1603) < **bʰeh₂-* ‘to shine’ (EDG 1551). Whether the Greek is related or not, the root **ǵʰuok-* or **ǵʰuok^w-* is of an invalid **DʰeT* structure, leading DV (207) to suggest it is a loanword.

Kroonen (2017: 106) alternatively suggests that Lat. *fax* might be a back-formation from attested *facula* ‘torch’, perhaps an old instrument noun to the root **bʰh₁-tleh₂* to the root **bʰeh₁-* ‘to make warm’ (cf. already Grassmann 1863: 88 on the root etymology) akin to PGM. **bēla-* (< **bʰeh₁-tlō-*, cf. ON *bál* ‘campfire’). This solves the problem of the invalid root structure for Latin, but requires Lith. *žvākė* to be unrelated. It is unclear which solution to choose.

glārea ‘pebble, gravel’

Pre-form: **g^(h?)lH-ro-* | PItal. **glārejo-*

Comp.: **ǵʰlar-* | PGk. **kʰlaro-* | Hsch. χαράρον· κόχλαξ ‘pebble’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: geography

Pokorny (390-1), WH (I: 605), EM (276), DV (264)

Alessio (1944: 132), EDG (1636), Zair (2013), Weiss (2020: 177, fn. 26), Kroonen et al. (2022: 8)

The traditional etymology of Lat. *glārea* ‘gravel’ relates it to *grānum* ‘grain’ < **grH-no-*. It would develop from **grH-ro-* > **grāros* >> **grārejos* with dissimilation to **glārejo-* (WH I: 605, Pokorny 390-1). DV (264) writes that this etymology relies on the original meaning of *grānum* being ‘small piece’ rather than ‘ripened, aged’, and Kroonen et al. (2022: 8) indeed argue in favor of this while EM (276) are uncertain. Zair (2013) remains open to the root etymology but adduces PCelt. **grāwā-* (cf. MW *gro* ‘gravel, shingle’, OCo. *grou* ‘sand, gravel’) along with Friulian *grava* ‘gravel’ as cognate from **grā-ūā-* beside Lat. *glārea* << **grā-ro-* (with the same metathesis).

Alessio (1944: 132) instead compares *glārea* to Hsch. *χλαρόν· κόχληξ* (= *κάχληξ*) ‘pebble’, followed by EDG (1636), and further assigns it Mediterranean substrate status. It is uncertain if **g^hl-* should yield **gl-* or **l-* in Latin (the same question posed by Zair 2013; cf. Weiss 2020: 177, fn. 26), but see the entry cf. *laena* (s.v.). The *ā* of Latin would be irregular against a short *a* in Greek, but can we be certain that the *a* is short if it appears only in Hesychius? If the word is **χλαρόν*, both comparanda can be regular reflexes of **g^hlh₂-ro-*. The traditional etymology relies on dissimilation, which is inherently *ad hoc*. But it seems drastic to reject this in favor of a Hesychius gloss.

haedus ‘young goat, (goat) kid’Pre-form: **g^hheid-/g^heh₂id-* | PItal. **χaido-*Comp.: **g^hheid-/g^heh₂id-* | PGM. **gait-* | Go. *gaits*, ON *geit*, OHG *geiz* ‘goat’?PSem. **gadī-* | Akk. *gadū*, Arab. *jady*, Hebr. *gdī* ‘(goat) kid’?PBERb. **āqāḏ-* ‘(she-) goat’; **qayd* ‘billy-goat’

□ Irreg. correspondences

■ Remarkable phonotactics

Semantics: animal, domestic

Pokorny (409-10), WH (I: 632), EM (288), DV (278)

Möller (1911: 128), CAD (G: 9), Schrijver (1991: 269), Demiraj (1997: 341), Boutkan & Kossmann (1999: 89), Kroonen (2012a: 242, 245-7), Blažek (2013: 46), Kroonen (2013: 163)

An Italo-Germanic **g^haid-* does not look non-IE except for its *a*, leading Schrijver (1991: 269) to reconstruct **g^heh₂id-* or **g^hheid-*. Kroonen (2012a: 245) notes that both root structures are unusual and further that the Germanic forms inflect as a root noun, a feature of old borrowings. Semitic comparanda, first adduced but incorrectly used by

Möller (1911: 128, cf. later CAD [G: 9], Kroonen [2012a: 246 with lit.]), and Berber (cf. Boutkan & Kossmann 1999: 89), if they belong, potentially hint at a non-IE origin. It is conceivable that **g^haid-* entered Latin and Germanic from an agricultural substrate (cf. DV 278, Kroonen 2013: 163), though it is difficult to prove.³⁷³

īnsula ‘island’

Pre-form: **in-sVl-* | PItal. **īnsVlā-*

Comp.: **e/ine/istī-* | PCelt. **e/ine/i-stī-* / **ine/issī-* | OIr. *inis*, W *ynys* ‘island’

**(s)neh₂t/k-ǵo-* | PGk. **nās(s)o-* | Gk. *νήσος* vars. Doric *vāσος*, Rhodes *vāσσος* ‘island’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: geography

Pokorny (878-9), WH (I: 707), EM (319), DV (306)

Prellwitz (1897: 123), Rozwadowski (1907: 348-0), Cuny (1910: 157), Derksen (2007: 379), Matasović (2009: 116), EDG (1019), Derksen (2014 s.v. *salà*)

The ancient etymology for Latin *īnsula* ‘island’ derives it from *in* + *sal*, ‘in the salt (water)’ (WH I: 707, EM 319, DV 306) but this is quite possibly a folk etymology. Islands are not in salt, but rather in (salt)water. Gk. *ἐναλος* is formed similarly (*ἐν* ‘in’ + *ἄλς* ‘salt’) but means ‘maritime’. EM (319) note that the Slavic (cf. OCS *ostrovъ* ‘island’ < **ob* ‘around’ + **strovъ* < **strujà* ‘to stream’, Derksen 2007: 379) and Indo-Iranian words for island suggest river islands rather than oceanic ones.

Early on, linguists like Prellwitz (1897: 123) compared Lith. *salà* ‘island’. Rozwadowski (1907: 348-9) was skeptical of the connection, as Lat. *īnsula* would then have to mean ‘in the island’. Derksen (2014 s.v. *salà*) suggests that Lith. *salà* ‘island’ maybe have developed from **ap(i)salā* with the second element being *sālti* ‘to trickle, flow’ (a formation parallel to OCS *ostrovъ* ‘island’). Otherwise, it reconstructs to **sol-eh₂-*, where the root **sol-* is similar to ON *sol* ‘sea’ (< PIE **sH/ol-u-*, Guus Kroonen, p.c.). If this represents an inherited word for ‘sea’, then Lat. *īnsula*, instead of being derived from **in-sal-o-* ‘in the salt’ could instead be derived from **in-sol-o-* ‘in the sea’. On the other hand, de Vries (1962: 578) derives *sol* from the salt lexeme.

An alternative explanation is to consider potential Celtic and Greek comparanda. For the Celtic island words such as OIr. *inis*, Matasović (2009: 116) prefers the reconstruction of PCelt. **eništī* because it allows for a PIE interpretation **eni-sth₂-ih₂* ‘that which stands in (the water)’. The Bennungsmotiv is similar to the aforementioned Latin explanations,

³⁷³ Further similar words for goat cannot be adduced with any certainty. PGM. **kidja-* (ON *kið* > Engl. *kid* ‘(goat) kid’) as if from **gid^h-* would introduce further alternations to Germanic (cf. Blažek 2013: 46). Pokorny (409-10) suspects it is derived from a call for goats. He likewise explains Alb. *qith* ‘young goat’ this way. Demiraj (1997: 341) takes it as a dialectal variant of *kedh* ‘(goat) kid’, itself perhaps a contamination of Turk. *keçi* and Alb. *edh* ‘billy-goat’.

but is morphologically and semantically quite different. The Greek forms, **nāso-* and **nāssō-* with their geminate alternation (Furnée 1972: 387), cannot be explained in this way (nor by derivation from **sneh₁-* ‘to swim’, cf. EDG 1018). But the similarity of their consonantism makes linking them to the Italic and Celtic potentially attractive. This family could represent non-IE loans (EM 319, DV 306, EDG 1018) of the *Amsel-merula* pattern of *a*-prefixation, either suggesting that other vowels could fulfill this role (cf. *ulmus*, s.v., where Schrijver [1997: 311] proposes a non-*a* vowel in the phenomenon on comparison with PGm. **elma-* ‘elm’) or that the Latin and Celtic forms were subjected to change due to old folk etymology. The Italic forms would represent **i-ns(-elo)-* against Greek **nās-* (cf. Cuny 1910: 157), but the Celtic would require **i-nVs-* without the zero-grade we expect in the prefixed forms. The PCelt. reconstruction with geminate *s* suggested by DV however also parallels the Greek variations with a geminate *s* quite well.

iūniperus ‘juniper’

Pre-form: **(H)jojn-i-* / **(H)juH-n-i-* **-pVr/s-* | PItal. **yoīnipVr/so-* / **yūnipVr/so-*

Comp.: *?(H)(i)ojn-i-* | PGm. **(j)ainja-* | ON *einir*, Dan. *ene-bær*, etc. ‘juniper’

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, wild

Pokorny (513), WH (I: 729-31), EM (328), DV (313)

Brüch (1922: 224-232), Pisani (1935: 37-8), FEW (V: 75), Falk & Torp (1960: 194), Kroonen (2013: 12)

Lat. *iūniperus* ‘juniper’ is often assumed to be related to *iuncus* ‘reed, rush’ (s.v.), but the comparison is based on faulty semantics and does not explain the *-perus* element of *iūniperus*.³⁷⁴ If it is to be reconstructed as **yoīni-*, it is not the same as the **yoīni-* of *iuncus*. Another option would be **yūni-* < **(H)juH-n-*.

PGm. **ainja-* ‘juniper’ is an attractive comparandum. Most Germanic forms are North Germanic (ON *einir*, Norw. *eine*, Dan. *ene-bær*, etc. ‘juniper’). Kroonen (2013: 12) rules out a reconstruction with PGm. **j* because of the West Germanic forms (Low German *ēn(e)ke* and Ger. dial. *Einbeerbaum*). But there is a chance that these West Germanic forms are loaned from North Germanic and/or have undergone folk etymological contamination with the numeral ‘one’ (cf. Brüch 1922: 226, Falk & Torp 1960: 194). In this case, the Latin and Germanic words could reconstruct to a common pre-form, at least in the *iūni-* element.

³⁷⁴ Older attempts at an explanation included: **joīni-d^hro-* (purportedly with the suffix of *combrēnum* ‘rush’) > Lat. **iūnibro-*, which was interpreted as Sabellic and hyper-Latinized to *iūniperus* (Brüch 1922: 227-30); an original *s*-stem composed of **jojn-ik-yos-*, which yielded Sabellic **iūnipes-*, borrowed into Latin as **iūnipeso-* with rhotacism to *iūniperus* (Pisani 1935: 37-8). It is also difficult to image any relation to *pirum* ‘pear’ or *pariō* ‘to give birth, beget’.

If related to *iuncus* ‘reed’, one might expect the element *-perus* to provide the meaning ‘juniper’, but it is of obscure etymology. The comparison with Germanic on the other hand, if valid, suggests that the *imi-* element means juniper, leaving *-perus* with unknown meaning and function. EM (328) mention the form *iupicellos* given as Gaulish by Pseudo-Dioscorides. The Romance languages descend from **ieniperus*, but this is probably regular (Brüch 1922: 230-2, FEW V: 75). The origin of *iūniperus* and its relationship to the Germanic forms remains obscure.

labium ‘lip’

Pre-form: **la/o/Hb-io-* | PItal. **labijo-*

Comp.: **leb-io-* | PGm. **lepjan-* | OE, OFri. *lippa*, MDu. *lippe*, etc. ‘lip’

□ Irreg. correspondences

■ Remarkable phonotactics

Semantics: body part

Pokorny (655-7), WH (I: 738), EM (333-4), DV (319)

Schrijver (1991: 479), Sihler (1995: 146), EDG (867), Kroonen (2013: 331)

The only secure comparanda for Lat. *labium* and *labrum* ‘lip’ are in Germanic: **lepjan-* < **leb-ion-* and an *s*-stem **lepaz-*. This points to **b* rather than **b^h* for *labium*. DV (319) and Kroonen (2013: 331) further compare Gk. *λοβός* ‘lobe, lap, slip’ < **lob-*, though the semantics are not as close and EDG (867) thus compares it elsewhere.

The **b* is suspicious. As Sihler (1995: 146) notes, *labium* follows the classic pattern of a root with **b* in that it is 1) nominal, not verbal and 2) restricted to two or three (usually) adjacent languages. But he is too quick in noting that Lat. *-a-* cannot match e.g. OE *-i-*. The Germanic forms are from an *e*-grade. Schrijver (1991: cf. 479) proposes that the consonant cluster in **l**j**io-* could yield **labio-*. Kroonen (2013: 331) takes the Latin from an *o*-grade, delabialized to *a* after **l* like in *lacus* < **lok-u-*. The Latin could also be from **lHb-* if the Germanic is from full-grade **lh_veb-*. In any case, the vocalism does not preclude IE origin: both forms can reconstruct to more or less the same pre-form. It is the **b* and the geographic restriction (unless Greek is related) that might.

līnum ‘linen, flax’

Pre-form: **liHno-* / **leino-* | PItal. **līno-* / **leino-*

Comp.: **lino-* | PGk. **lino-* | Gk. *λίνον* ‘linen’

**lino-* | BSl. **līnum-* | OPr. *lynno* ‘flax’, Lith. *linas* ‘flax (plant)’, Latv. *lini* ‘flax’, CS *льня*, Ru. *лён* ‘flax’

?**liHno-* | PCelt. **līno-* | OIr. *lín* ‘flax’, etc.

?**liHno-* | PGm. **līna-* | Go. *lein* ‘canvas’, etc.

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, domestic / textiles

Pokorny (691), WH (I: 810-11), EM (361-2), DV (344)

Biville (II: 23), Derksen (2007: 298), EDG (863)

Whether irregular vocalic alternations exist in this lexeme depends on the independence of the Celtic and Germanic forms. If they are borrowings from Latin (WH I: 810, EDG 863), then Lat. *līnum* presents a full-grade **lein-o-* against a zero-grade **lin-o-* elsewhere.³⁷⁵ Derksen (2007: 298) takes the forms as independent, creating a non-IE alternation **i ~ ī*. It does not seem possible to decide.

DV (344) gives the Latin forms in **lint-* (e.g. *linteum* ‘piece of linen cloth, towel, sail’, *linteolum* ‘piece of linen’, and *linteō* ‘weaver of linen’) as additional evidence of a non-IE source of the lexeme. But this does not seem necessary. Its short *i* is the expected result of Osthoff’s Law. WH (I: 811) and EM (361)³⁷⁶ explain the suffix as a secondary innovation, potentially via contamination with reanalyzed forms like *spart-eus* > *spart-teus* ‘made of broom’. However, an *-eus* derivation of a *-tus* derivation seems like a more simple solution (cf. *rōbur* ‘oak’ > *rōbustus* ‘oaken, hard’ > *rōbusteus* ‘oaken’).

lōrum ‘leather strap, thong’

Pre-form: **(H/s/y)loH-ro-* | PItal. **lōro-*

Comp.: **h₁uleh₁-ro-* / **h₂e-h₂ul-eh₁-ro-* | PGk. **eulēro-* / **āulēro-* | Gk. εὔληρα, Dor. αὔληρα ‘reins’

**(h_{1/2})ulh₁-ro-* | PArm. **ularo-* | Arm. *lar* ‘rope, cord, rein’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: tool

Pokorny (1140-4), WH (I: 822), EM (366-7), DV (349)

Lidén (1906: 100-1), Cuny (1910: 158), Beekes (1988: 71), Schrijver (1991: 74-5, 122-3), Clackson (1994: 39), Olsen (1999: 30), Martirosyan (2009: 305), EDG (480, 569)

Since at least Lidén (1906: 100-1) Lat. *lōrum* ‘leather strap’ has been compared to Gk. εὔληρα ‘reins’ and Arm. *lar* ‘rope, cord, rein’. A pre-form is difficult to reconstruct that works for both the Latin/Armenian and the Greek. The Greek forms require an initial laryngeal. Beekes (1988: 71) favors εὔληρα and reconstructs **h₁ulēr-* while Olsen (1999:30) takes εὔληρα as assimilation from αὔληρα and reconstructs **h₂uleh₁r-*. Arm.

³⁷⁵ Biville (II: 23) even suggests that, since this length alternation occurs in the initial syllable as it does for example with *mōrum* vs. μόρον ‘mulberry, blackberry’, there is a chance that the length was induced by the addition of primary stress after borrowing from Greek. This is uncertain.

³⁷⁶ EM (261) further propose a *-teo* material suffix, otherwise attested only in *robusteus* ‘oaken’ (but this is easy to explain as doubly derived) or Etruscan factors based on the shape of *balteus* ‘belt, girdle’, often thought to be Etruscan.

lar could be from **h₁ulh₁-ro-* (Martirosyan 2009: 305) or **h₂ulh₁-ro-* (Olsen 1999: 30). However it is unlikely that a matching reconstruction of **h₁ulō-* for Lat. *lōrum* (Schrijver 1991: 74-5, 122-3; Olsen 1999: 30) would yield anything other than **ulō-*; **H* or **u* are possible, but not both. Clackson (1994: 39) reconstructs **ulh₁-* for the Armenian and **uloh₁-* for the Latin, but **uleh₁-* does not yield the proper Greek forms.³⁷⁷ Given the formal difficulties, Martirosyan (2009: 305) proposes a Mediterranean substrate word. EDG (480) too takes the two Greek forms (along with Hsch. ἄβληρα: ἡνία ‘rein’) to attest to a non-IE *a* ~ *e* alternation. DV (349) prefers an IE explanation due to the *ē* ~ *ō* (or perhaps *e* ~ *o* ~ *∅*) alternation of the suffixes, which looks like IE ablaut. He provides an alternative reconstruction for the Greek (and suggests that εὔληρα is *metri causa*, in fact Osthoff’s shortening, for *ἄυληρα < **āulēra* < **h₂e-h₂ul-ēr*), proposing a loan from an extinct IE language. This is of course difficult to prove.

Since Varro (*de Lingua Latina* 5.116), Lat. *lōrīca* ‘cuirass’ has been taken as a derivation from *lōrum* (cf. WH I: 822), denoting that cuirasses were made of leather. However *lōrīca* is also compared to Gk. θώραξ ‘cuirass; torso, chest’, in which the *āk*-suffix is indicative of non-IE origin (e.g. Cuny 1910: 158, EDG 569). Any relationship between them would have to be entirely irregular, and the connection between *lōrum* and *lōrīca* would have to be folk etymological.

palumbēs ‘wood pigeon’, var. *palumbus*

Pre-form: **pa/Hl-e/o/umb^(h)-* / **p_lH-e/o/umb^(h)-* | PItal. **pale/o/umb/f-*

Comp.: *?*pel-ej-* | PGk. **peleja-* | Gk. Gk. πέλεια, πελειάς ‘wild pigeon’

*?*poh₂l-* | PBalt. **pōli-* | OPr. *poalis* ‘pigeon’

*?*p_lh₂-b^h-ōn-* | PArm. **alawun-* | Arm. *alawni* ‘dove’

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: animal, bird

Pokorny (804-5), WH (II: 242), EM (478), DV (126, 442)

Bugge (1893: 1), Prellwitz (1897: 102), Ernout (1965: 15-16, 23), Greppin (1978: 131-2), Klingenschmidt (1982: 165), Lockwood (1990: 262-3), Schrijver (1991: 375), Olsen (1999: 508), Witczak (1999: 177-8), Martirosyan (2009: 29, 565), EDG (1166), Jakob (fthc.)

Lat. *palumbēs* occurs contemporaneously with *palumbus* (the former in Plautus, the latter in Cato). The suffix *-umb-* occurs nowhere else in Latin except for the other dove word *columba* (s.v.). Interpretations of its origins vary, and none is entirely without problems.

Given Gk. πέλεια, which lacks the suffix but which can formally quite easily derive from

³⁷⁷ EM (367) simply call the vowel ‘prothetic’.

PIE **pel-* ‘gray, pale’, it is often proposed that *palumbēs* contains the root of *palleō* ‘to be pale’ with the suffix of *columba* (Prellwitz 1897: 102, Pokorny 804-5, Ernout 1965: 23, WH II: 242, Lockwood 1990: 262-3 [from a pre-form ***palēs*], EM 478, EDG 1166, ambivalently DV 442). However the wholesale transfer of such a rare suffix seems difficult to motivate. Additionally, it is *columba*, not *columbus* that is the earliest and most securely attested form (Ernout 1965: 15-16, Schrijver 1991: 375).³⁷⁸

Klingenschmitt (1982: 165) proposed a connection with Arm. *alawni* ‘dove’³⁷⁹ via a pre-form **p_lh₂-b^h-nih₂-*. Martirosyan (2009: 29) instead proposed an original **p_lh₂-b^h-ōn*, *-b^h-n-os*, since dialectal variation points to *alawni* being a secondary formation from an original **alawun*. But **p_lh₂-b^h-n-os* should give Lat. ***plāmnus*; *palumb-* requires something like **p_lh₂-Vn-b^h-*. He alternatively proposes a Mediterranean origin for the Armenian and Latin forms, noting the similar pair Lat. *columba* ‘dove’ ~ Arm. *salamb* ‘francolin’. But the order of the nasal and labial in the suffix is not the same between *salamb* and *alawni*. Could there have been a metathesis? Additionally this would probably rule out the appurtenance of the Greek and Old Prussian forms.

pīnus ‘pine tree’

Pre-form: **pi(C)s-no-* | PItal. **pi(C)sno-*

Comp.: **pit-* | PGk. **pitu-* | Gk. πίτυς ‘pine, fir, spruce’

**pi(t?)s-* | PALb. **pishā-* | Alb. *pishë* ‘fir, pine’

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, tree

Pokorny (793-4), WH (II: 308), EM (509), DV (467, 469)

Furnée (1972: 260), Schrijver (1991: 231-2), EWAia (II: 137-8), Demiraj (1997: 321-2), Derksen (2007: 426), EDG (1198), Derksen (2014 s.v. *pikis*), Smoczyński (2018: 959)

Because of Latin sound developments, Lat. *pīnus* can be reconstructed with any (or no) stop before a sibilant + *n*, making it difficult to confirm which words are comparanda and therefore whether or not they are irregular.

The strictest semantic approach, comparing *pīnus* only to other words for pine, gives a non-IE impression. Gk. πίτυς ‘pine, fir, spruce’ < **pit-u-* can match with *pīnus* if the latter is from **pit-sno-*.³⁸⁰ Explanations for Alb. *pishë* ‘fir, pine’ vary greatly. The *sh*

³⁷⁸ The *a*-vocalism of *palleō* means it probably continues a different root than πέλεια and OPr. *poalis* (DV 440 suspects it itself might be a loanword). This would mean that the Latin is a separate formation from the Greek and Prussian.

³⁷⁹ Previously linked to the root **h₂elbo-* ‘white’ by Bugge (1893: 1). Greppin (1978: 131-2) doubts the connection because white doves do not seem to appear until the 5th c. BCE and there are some formal difficulties (cf. Olsen 1999: 508). Witzak (1999: 177-8) tried to connect *alawni* with Lith. *balandis* and Ossetic *balon* ‘pigeon, dove’, but it requires strange metathesis.

³⁸⁰ Furnée (1972: 260) claims an *s* ~ *t* alternation behind Gk. πίτυς ‘pine, fir, spruce’, but only has

could be due to simple intervocalic *s in *pis-ā. Clusters in *Cs do not seem to be well understood, and Demiraj (1997: 321-2) lists without rejection several other proposals: *peukā, *pit-s-ja, *peit-s-eh₂, and *pikso- before settling on perhaps a root *pī- with a collective ending *-sjo. The forms with *t would make this group look regular, but *pisā would produce a sibilant alternation with Gk. πῖτος.

Schrijver (1991: 231-2) mentions the possibility that *pīnus* is simply from *piH-no-, or at least has a long vowel *per se* against the short vowel of Gk. πῖτος. This is in comparison with Skt. *pītūdāru-*, *pūtūdru-*, etc. ‘a resinous tree’. Assuming that the forms starting with *pītu-* are secondary, EWAia (II: 137-8) writes that the similarity to πῖτος and *pīnus* can hardly be coincidental. Such a length alternation is difficult to explain from an inherited perspective.

DV (467) notes that a reconstruction *pik-sno- would link it to *pix* ‘pitch’, which is semantically not unimaginable. The comparanda of *pix* do not require any irregularities in reconstruction such that there is no reason to assume non-IE origin, but are of limited distribution (Gk. πίσσα ‘pitch’ < *pit/k-ja,³⁸¹ OCS *pъcbъbъ* ‘pitch’ < *pik-i/ul,³⁸² cf. DV 469).

porrum ‘leek’

Pre-form: *p̥rso- | PItal. *porso-

Comp.: *p̥rso-? | PGk. *praso- | Gk. πράσον ‘leek’

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, domestic

Pokorny (846), WH (II: 343), EM (523), DV (481)

Cuny (1910: 157), Schulze (1933: 116-7), CAD (G: 142), Vycichl (1963), CAD (K: 212-4, 567), Puhvel (IV: 274), EWAia (II: 101), Orel (1998: 344), Wachter (2006: 139-44), EDG (1179, 1229), Rosól (2013: 16, 202), Garnier & Sagot (2017: 34, 47-8), van Beek (2022: 386-8, 394-5)

EM (523) asserts that Lat. *porrum* and Gk. πράσον are independent borrowings from a third source, and DV (481) generally agrees. Even Pokorny (846) suggests it is a Mediterranean loanword. Their arguments are mainly semantic, but there are potential formal inconsistencies as well (cf. already Cuny 1910: 157). EDG (1229) notes the disputedness of the retention of Gk -s- after a syllabic resonant. Schulze (1933: 116-7) only gives three examples, one of which is this very word. The best is Gk. δασύς ‘hairy’,

toponyms as evidence (EDG 1198).

³⁸¹ As EDG (1197) notes, πίσσα (Attic πίττα) could also instead be linked to πῖτος ‘pine’.

³⁸² East Baltic forms are loans from Low German (Smoczyński 2018: 959, *pace* DV 469), and Old Prussian *pyculs* ‘hell’ might be from Polish (Derksen 2014 s.v. *pikis*). The Slavic forms seem to be independent (Derksen 2007: 426).

which if related to Lat. *dēnsus* ‘thick’, is from **d̥ns-u-*. The semantics are not as good as they are for the best counter-example: Gk. γράω ‘to gnaw, eat’, Skt. *grásate* ‘devours’. Here, the Greek vocalism must reflect **gys-* or **gr̥ns-*,³⁸³ but in neither case is the **s* preserved by the resonant. It is thus not entirely clear if the Latin and Greek forms can reconstruct to the same pre-form **p̥rso-* (cf. also van Beek 2022: 394-5).

Wachter proposes that the first element in the early variant *περσόφαττα* of the name Persephone/Proserpina is **perso-* ‘ear of grain’ or ‘sheath’. As a PIE root, this would find support only in Indo-Iranian, where Skt. *parśá-* is a hapax occurring at RV 10.48.7 and must mean ‘sheaf or bundle of grain’. Otherwise it occurs in YAv. *parša-* ‘ear of grain’ (EWAia II: 101). Weiss *apud* Wachter suggests a connection with *porrum*/πράσον, but DV (481) questions the semantics of the comparison. The connection would require that both Latin and Greek innovated the meaning of leek from what is otherwise a very poorly preserved grain root.³⁸⁴

Vycichl (1963) argues that **p̥rso-* is a loan from Sumerian via Semitic. But while he gives Sum. *guraš* ‘leek’ and Akk. *kurāšu*, *karāšu* ‘leek’, an updated spelling of these is *garaš* and *karašu*. The other examples that Vycichl gives of Sem. *k* to Gk. *p* are actually of Sem. *gu-* to Gk. *bu-*: βύρσα ‘skin, hide’ from (the same source as) Akk. *kursinnu/gusānu* ‘leather sack’ (CAD K: 567, G: 142) and Hitt. *kursa-* ‘skin, hide, fleece’ (cf. also Puhvel IV: 274); Βύβλος ‘the city of Byblos’, cf. Akk. *Gubla*, Hebr. *Gebal*. It therefore seems like Semitic *gu-* was interpreted in Greek as **g^wu-*, presumably after **u* stopped delabializing **g^w* (cf. EDG 246³⁸⁵). Rosół (2013: 16, 202) indeed rejects the connection between the Sumerian-Akkadian material and Gk. πράσον due to the unparalleled phonological matches in comparison to the rest of his data.

sapa ‘grape juice or new wine boiled down to a syrup’

Pre-form: **sa/Hp-* | PItal. **sapā-*

**sa/o/Hp-on-* | PGM. **saf/ban-* | ON *safi*, OHG *saf* ‘sap, juice’, etc.

?**sa/Hb^(h)-?* **sa/Hp-?* | PArm. **sab-mo-?* **sap-mo-?* | Arm. *ham* ‘juice, taste’

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: culinary; viticulture

Pokorny (880), WH (II: 476-7), EM (594), DV (538)

Jahowkyan (1987: 189), Schrijver (1991: 104), Olsen (1999: 27), EWAia (II: 701),

³⁸³ The latter is more likely (van Beek 2022: 386-8, cf. *grāmen*, s.v.).

³⁸⁴ Garnier and Sagot (2017: 34, 47-8) propose an alternate etymology by suggesting the existence of an IE substrate that underwent changes such as **b^h > p*, making it possible to connect **p̥rso-* to the root **b^hers-* ‘to point, burst, bud’.

³⁸⁵ He alternatively suggests (for βύβλος) that an assimilation *g-b > b-b* may have taken place. But this solution is more *ad hoc*.

Kroonen (2013: 420)

Lat. *sapa* has a specific and technical meaning, referring usually to unfermented grape juice ('must') boiled into a syrup.³⁸⁶ Thus *sapa* was a sort of artificial honey, used in part as a preservative.

Some Germanic forms seem to reconstruct to PGm. **sap-* < **sab-*, which DV (538) takes at face value to identify an irregular *b ~ p* alternation. However both Italic and Germanic reconstruct to **sHp-* if the root in Germanic formed an *n*-stem (Kroonen 2013: 420). Additionally some of the forms seem to have been borrowed from Latin (EM 594, DV 538, Schrijver 1991: 104, though WH II: 476 reject this on semantic grounds).

Arm. *ham* 'juice, taste' can be related (EM 594, DV 538) and is sometimes derived from **sHp-mo-* (Pokorny 880, Olsen 1999: 27, Kroonen 2013: 420).³⁸⁷ Jahowkyan (1987: 189) suggests reconstructing **sab-mo-*. If **-pn-* yields Arm. *wn*, would we not expect **-pm-* to yield Arm. *wm* and therefore ***hawm*? There are unfortunately no clear examples of the outcome **-pm-* or **-bm-* to confirm.

The only remaining evidence for a *p ~ b* alternation in the root comes from Indo-Iranian, but the relation of these forms to the rest is questionable. Pokorny (880) and WH (II: 467) connect Av. *vīšapa-* 'whose juices are poison' from **viš-sāpa-*, but Schrijver (1991: 104) is right to consider it too uncertain. Given we must assume that the *s* of ***sāpa-* is hidden by the sibilant of *vīš-*, the second element may just as well be *ap-* 'water'. Also compared are OAv. *hābuuant-* meaning something like 'juicy' < **sab-uant* and the first element in Skt. *sabardūh-*, an epithet especially of a dairy cow (WH II: 477, EWAia II: 701, DV 538). Given that the evidence for a root **sab-* in Europe is already unclear, the Indo-Iranian forms are best left out. There is a chance that Latin, Germanic, and Armenian all regularly attest to a root **sap-*.

simpvium 'earthenware ladle used in religious ceremonies'

Pre-form: **simp-* | PItal. **simpu-*

Comp.: ?

Irreg. correspondences

Remarkable phonotactics

Semantics: tool; magico-religious

WH (II: 540-1), EM (627), DV (554, 565)

Masson (1967: 44-5), Furnée (1972: 272, 286), Leumann (1977: 136), Untermann (2000:

³⁸⁶ So too was *dēfrutum* but one was boiled down to half of its original volume and the other further to one third. (Which one was which depends on the source: the proportions are given in various orders in e.g. Pliny *Nat.Hist.* 14.80, Columella *de Re Rustica* 12.19, and Varro apud Nonius *Comp. Doc.* 18.551M.)

³⁸⁷ WH II: 476 reconstruct **sap-no-*, but **pn* produces Arm. *wn*, cf. *tawn* 'feast' < **dh₂p-ni(h₂)-* (Martirosyan 609). EM (594) use the meaning 'flavor' for Arm. *ham* to propose that Lat. *sapa* is related to Lat. *sapiō*, *-ere* 'to taste, to know'. Pokorny (880) agrees, but DV (538) thinks the semantic range including 'juice' and 'sap' of the comparanda would make equating these roots strange.

668), Rix (2005: 569), EDG (1335), Meiser (2010: 81), Rosoł (2013: 205)

Lat. *simpuvium* also occurs as *simpulum*. Meiser (2010: 81) derives the lexeme from **semH-* ‘to scoop’, which would seem to require the *p* to have arisen via epenthesis. But *simpulum* may be a remodeling of *simpuvium* (Leumann 1977: 136, DV 565) or its *l* may be a misreading of *i* (Untermann 2000: 668).³⁸⁸ This removes the environment for epenthesis, and the root is probably **simp-*, with a suffix found in another ritual word *atannuium*³⁸⁹ ‘an earthenware bowl used in offering sacrifices’ (Leumann 1977: 136).

The Latin forms are compared to Gk. σιπύη, σιπύα ‘box for keeping flour and bread’ (WH II: 540, EM 627, DV 565). The Hesychian variant ιπύα is strange. If it is meant to be *ιπύα < **sjp-*, it suggests a loanword that entered Greek before and again after the loss of **s*. A direct loan into Latin is difficult given the difference in meaning and the additional nasal element. Furnée (1972: 272, 286) suggests independent loans from a third source, further comparing Gk. σίμβλος ‘beehive, larder’.³⁹⁰ The nasal of the Latin form is present in σίμβλος, but its semantics, as well as those of all the Greek forms, are so distant that its appurtenance is difficult to verify.

Alternatively, Rix (2005: 569) takes *simpuvium* as a loan from the Sabellic reflex of **seik^w-* ‘to pour’ (cf. Skt. *śīncāti* ‘pours’, Gk. ικμάς ‘wetness’). It is unclear which of these solutions, if any, correctly explains the origins of Lat. *simpuvium*.

termes, -itīs ‘branch of a tree, especially one cut off’

Pre-form: **ter(H)(b)-m-* | PItal. **ter(V)met-*

Comp.: **terh₂/b-mn-* | PGk. **teramno-* | Gk. τέραμνα [pl.] ‘house’

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant

Pokorny (1070-1), WH (II: 670), EM (686), DV (615)

Bertoldi (1939b: 92), Bertoldi (1942: 180-1), Alessio (1944a: 109), Frisk (1960-72 II: 877), Furnée (1972: 219), Weiss (1993: 84), Untermann (2000: 766), EDG (1467, 1469)

Because it sometimes specifically refers to an olive branch, Bertoldi (1939b: 92, 1942: 180-1) suspected pre-Latin origin and compared it to Gk. τέρμινθος, τρέβινθος, τρέμιθος ‘the turpentine tree’. Its *-es* ending led EM (686) to hesitatingly follow, finding this morphology in other words of obscure (or purportedly Etruscan) origin like *cocles* ‘one-eyed’, *miles* ‘soldier’, *satelles* ‘attendant, bodyguard’, etc. The Greek words seem to show a *b ~ m* alternation (Furnée 1972: 219), but EDG (1469) follows the

³⁸⁸ WH (II: 540) suggest that the existence of the form with *l* is supported by U *seples* of the same meaning. EM (627) reject this, and Untermann (2000: 668) shows that *seples* meant ‘nail’, not ‘ceremonial ladle’.

³⁸⁹ Interestingly, it also has the variant *atanulum*.

³⁹⁰ The Greek words have been suspected to be loans from Semitic (Masson 1967: 44-5, DV 565, EDG 1335) but Rosoł (2013: 205) rejects this.

interpretation that, if not *m...n* dissimilation, they were remodeled on analogy with ἐρέβινθος. In any case, the semantic match between the Latin and these Greek forms is not good enough to warrant a comparison (cf. WH II: 670; Alessio 1944a: 109, fn. 110, who still entertains the possibility of a borrowed origin due to the ending).

Connections within Latin such as with *terminus* ‘end, limit’ are theoretically possible. But a quite compelling connection is in fact with Gk. τέραμνα ‘house’. Gk. τέραμνα is often compared to Lat. *trabs* ‘beam’ (s.v. *trabs* for discussion), so the semantics of the comparison are basically the same. But in this case, the formal comparison works much better: Gk. < **terh₂-mn-* (Untermann 2000: 766) and Lat. < **terh₂-m-*. In fact, while Gk. τέραμνα is widely translated as ‘house’, Euripides (*Hippolytus* 418) writes τέραμνα οἴκων, where it must mean something like ‘beams of the house’ (Weiss 1993: 84). If the aberrant vowel of variant τέρεμνα can be explained Greek-internally (Frisk 1960-72 II: 877) and the connection to θεράπων ‘servant, maid’ (EDG 1467) rejected, then Lat. *termes* can be inherited. If not, then the Greek forms attest to non-IE *t~t^h*, *a~e*, and *b~p* alternations and Gk. τέραμνα, Lat. *termes* reconstruct to a **tera/eb-* of foreign origin, extended with inherited suffixes.

tībia ‘reed pipe (flute); shinbone’

Pre-form: **teib^(h)-* / **tīHb^(h)-* | PItal. **teib/fiā* / **tīb/fiā*

Comp.: ?

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: tool

Pokorny (1102), WH (II: 680, 712), EM (691, 705), DV (88, 619, 632)
 Froehde (1889: 108), Solmsen (1898: 477), Boisacq (1916: 867), Schwyzer (1931: 205),
 REW (no. 8964), EWAia (II: 759-60), Derksen (2007: 472), EDG (1338), Garniet &
 Sagot (2017: 48-9), Weiss (2020: 174)

WH (II: 680 with lit.) compare Lat. *tībia* ‘flute; shinbone’ to reflexes of BSl. **stib-* ‘stem, stalk, trunk’ (Ru. dial. *stebλό* ‘stem, stalk’, SCr. *stáblo* ‘tree, tree trunk’, Lith. *stiebas* ‘stem, stalk, mast’, etc.) including importantly Lith. *stibýna* ‘shin, calf’ (Derksen 2007: 472 on the forms and reconstruction). This would be a Latino-Balto-Slavic isogloss in which Latin has preserved the full-grade **teib^(h)-* and Balto-Slavic the zero-grade with *s* mobile **stib^(h)-* of a root meaning ‘(hollow) stem’. Semantically, it is an excellent match, but relying on the poorly understood *s* mobile is not ideal.³⁹¹

An alternative semantically sound comparison is with Gk. σῆφων ‘tube, siphon’.³⁹² It was

³⁹¹ WH (II: 680) come to the same conclusion, but only upon comparing Lat. *stīpō* ‘to compress, surround’ and Skt. *stibhi-* ‘clump, tuft’, noting that the **s* is always present. These forms however are likely unrelated (DV 88 on the Latin, EWAia II: 759-60 on the Sanskrit).

³⁹² EDG (1338) considers the word onomatopoeic.

early on thought to be regular, with both from **t̥uībʰ-* (Froehde 1889: 108, Boisacq 1916: 867, WH II: 680). But just as early on, it was doubted (Solmsen 1898: 477, Schwyzler 1931: 205). Indeed, there is no evidence for **t̥u > Lat. t*.³⁹³ The pair could thus actually attest to an irregular alternation, and DV (619, 632) takes it as indicative of substrate origin. This explanation is potentially bolstered by the further comparison of Lat. *tuba* ‘trumpet’ and *tubus* ‘tube’. Romance reflexes with *f* (REW no. 8964) can be explained as the Sabellic reflex of **bʰ*, but the *i ~ u* alternation points to a loanword (DV 632).^{394,395} A similar alternation appears with *supparus* (s.v.). It remains difficult to decide to which of the two groups of potential comparanda *tibia* belongs.

trahō, -ere ‘to pull, drag, haul’

Pre-form: **tragʰ-* | PItal. **traxelō-*

Comp.: ?

□ Irreg. correspondences

■ Remarkable phonotactics

Semantics: action

Pokorny (257), WH (II: 698-9), EM (698-9), DV (626)

Walde (1906: 106), Sommer (1914: 50-64), Hamp (1978: 186), Schrijver (1991: 188-9), Nielsen (2004: 194), Schumacher (2004: 635-6), Derksen (2007: 122-3), Matasović (2009: 387), EDG (352, 1506), Kroonen (2013: 99, 510, 522, 544), Derksen (2014 s.v. *dirginti, drugʷs*), Stifter (2017: 1190), Weiss (2018)

It is difficult to determine with certainty what the best comparanda for Lat. *trahō* ‘to pull, drag, haul’ are due to its neutralizing phonetic environment. Its *a*-vocalism is difficult to account for in any reconstruction; it almost certainly cannot be due to a laryngeal.³⁹⁶

The most straightforward reconstruction is to **tregʰ-*, a root with an invalid structure that nonetheless may underlie several Celtic forms (cf. DV 626). OIr. *tráigid* ‘ebbs, recedes’ is probably denominal from **trāgʰ-* ‘beach, low tide’ (Matasović 2009: 387, cf. Schumacher 2004: 635), while OIr. pret. *tethraig* ‘ran away, receded’ might be the original **treg-i-ti* (Weiss 2018: 440-1); together they can represent IE *elō* ablaut. Further reconstructed to this root is Gk. *τρέχω* ‘to run, hurry’. While the Celtic could reconstruct to **treg-* and the Greek to **dʰregʰ-* (cf. EDG 1506), PGM. **pragjan-* ‘to run’ (cf. Go. *pragjan* ‘to run’) semantically unites the verbs and confirms the invalid root shape

³⁹³ In fact, there is only evidence for **t̥u > p*, cf. Lat. *pariēs* ‘wall’, Lith. *tvėrti* ‘to seize, enclose’ < **tuerH-* (Weiss 2020: 174).

³⁹⁴ WH (II: 712)’s proposal of **i > *ü > u* before a labial + non-high vowel is *ad hoc*.

³⁹⁵ Garnier and Sagot (2017: 48-9) suggest that *tubus* is from the root **dʰeubʰ-* via an IE substrate language whose reflex of **dʰubʰ-ó-* was **núβo-*.

³⁹⁶ Schrijver (1991: 188-9) reconstructs the root with a laryngeal on the evidence of *trāgula* and *trāgum* ‘net’; thus *trahō* could be evidence of the sporadic outcome *CRāC < *CRHC* that otherwise only occurs in the environment **CRHTC*. But actually, only short *a* is securely attested. Nielsen (2004: 194) shows that the vowel length of *trāgula* ‘dragnet, sledge’ and *trāgum* ‘net’, despite almost always being given as long, is actually indeterminate.

**treg^h*- (cf. Kroonen 2013: 544, Weiss 2018: 441).

However, Weiss (2018) argues that a connection with PGm. **dragan*- ‘to draw, pull, carry’ (ON *draga*, OE *dragan*, OHG *tragan*, etc. ‘to draw, pull carry’) < **drōg^h*-e- is semantically better. Kroonen (2013: 99) considers this root isolated to Germanic and DV (626) notes it is formally impossible unless the result of a loan (cf. also Kroonen 2013: 99). Weiss (2018) accounts for it via “Limited Latin Grassmann’s Law” that occurs in root shapes **D^hreD^h* due to an aspirated quality of the *r*.³⁹⁷ Weiss finds only two secure examples of this limited Grassmann’s Law: *trahō* and *glaber*, though it is a very elegant explanation for the latter form.

A third possibility is a connection to a root **dreg^h*- or (invalid) **dreg*-.³⁹⁸ Either shape could underly Gk. δράσσομαι ‘to grasp, take handfuls’ (cf. EDG 352) and PGm. **trekan*-, **trekkan*- ‘to pull’ (if the **k* is secondary from **g^h*, Kroonen 2013: 522). Balto-Slavic comparanda seem to favor the reconstruction with **g*. Derksen (2014 s.v. *dirginti*) takes Lith. *dirginti* ‘to pull a trigger’, Ru. *děrgat* ‘pull, tug’ < **drHg^h*-, but Kroonen (2013: 522) argues instead for a Winter’s Law affected **drg*-. BSL. **drug*- ‘to tremble, shake’ (cf. Ru. *drógat* ‘to shake’, *drožát* ‘to tremble, shiver’ cf. Derksen 2007: 122-3) would imply **dorg^h*- (Kroonen 2013: 522), but on semantic grounds this could be a different lexeme. For both this etymology and the one involving post-Grassmann’s Law **drag^h*-, a further change **dr*- > *tr*- must also have occurred.³⁹⁹

In the end, the etymology of Lat. *trahō* and the identity of its comparanda are difficult to secure. One gets the impression that *trahō* and similar forms might even be iconic/sound symbolic (cf. Engl. *jerk* ‘to yank’).

tūber, -eris ‘swelling, tumor’

Pre-form: **te/ou(H)b^(h)-es*- | PItal. **te/oub/fes*-

Comp.: ?

Irreg. correspondences Remarkable phonotactics

Semantics: body part

Pokorny (1080-5), WH (II: 712-13), EM (705), DV (632, 633)

Johansson (1890b: 444), Reichelt (1906: 74), EDG (521)

Lat. *tūber* is not well understood. It has been compared to Oic. *púfa* ‘knoll, hillock’ < **pūbaz*- < **tūb^h*- / **tūp*- and Gk. τόφη (vowel length unknown) ‘a plant used for filling

³⁹⁷ Walde (1906: 106) earlier called this *Aspiratendissimilation*, followed by an initially positive but skeptical discussion by Sommer (1914: 50-64). Weiss’ treatment shows much more restraint.

³⁹⁸ Stifter (2017: 1190) in fact derives the Celtic forms mentioned above from **dreg^hg^h*- with the sporadic devoicing seen also in **tang^wāt*- ‘tongue’ < **dnǵ^wueh₂* and **keng*- ‘to go, step’ < **ǵ^heng^h*-.

³⁹⁹ It is difficult to find proof of this sound law in initial position. Hamp (1978: 186) suggests that words like Lat. *truncus* ‘(tree)trunk, thorax, torso’ and *trudis* ‘pole, pike’ might represent reflexes of **doru*- ‘tree’ in Italic, but other examples are less convincing.

pillows and beds' < **tu(H)b^heh₂* (cf. WH II: 712-3, EM 705, DV 632). A root **tūb^h*- is of an invalid root structure; thus all forms could have as their root **teuH-* 'to swell' (Pokorny 1080-5, DV 632, EDG 521),⁴⁰⁰ but the explanation of one root with several root extensions is old-fashioned. WH (II: 712 with lit.) write of a "Parallelwurzel" **tūb^h*- to **tu-m-* in *tumeō*, but this is not a real explanation. There is some evidence for the root being **tu-* with suffixation *-*h₂-* and *-*m-* as in **g^mm-* and **g^mh₂-* 'to come' (cf. DV 633; cf. *tumulus* s.v.), and some therefore reconstruct **tūm-r-* for *tūber* with regular *-*mr-* > -*br-* (cf. Johansson 1890b: 444, 1906: 74). But the shape *tūber*, -*eris* and its neuter gender make it look very much like the reflex of a neuter *s*-stem. Thus it is difficult to understand how the proposed reconstructed **m* and **r* would ever have formed a cluster (cf. WH II: 713).⁴⁰¹

tumulus 'knoll, burial mound'

Pre-form: **tum-e/olo-* | PItal. **tume/olo-*

Comp.: **tum(-)b^(h)-* | PCelt. **tumbo-* | Mlr. *tomm* 'bush, bunch, hillock', MW *tom* 'dung, heap of dung, mound'

**tum(-)b^h*- | Arm. *tumb* 'embankment, earthen wall'

**tum(-)b-* | PGk. **tumbo-* | Gk. τῦμβος 'mound, burial mound, grave'

**tuHm-* | PGk. **tūmo-* | Gk. (Corcyra) τῦμος 'mound, burial mound, grave'

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: geography

Pokorny (1080-85), WH (II: 716), EM (707), DV (633)

Georgiev (1941: 70), Hester (1965: 379), Ačařyan (1971-79 II: 206), Matasović (2009: 392, 394), EDG (1517), Garnier & Sagot (2017: 49)

Lat. *tumulus* 'knoll, burial mound' is often derived from *tumeō* 'to swell, be swollen' (WH II: 716, EM 707, DV 633). Pokorny (1080-85) gave the root as **tēu-*, *təu-*, *teuə-*, *tūō-*, *tū-* with extensions in: *bh*, *g*, *k*, *l*, *m*, *n*, *r*, *s*, and *t*. On the parallel of two variants of the verb 'to come' (**g^mm-* and **g^mh₂-*), DV (633) proposes two suffixations of a root **tu-*: **tu-m-* (e.g. Lat. *tumeō*) and **tu-h₂-* (e.g. PSlav. **tŭti-* 'to become fat'). The form with *m* is well attested as a stative **tum-eh₁-* (Lat. *tumeō*, PCelt. **tumī-*, Lith. *tumėti* 'to become thick', Matasović 2009: 394), proving that it is old.⁴⁰²

The gist of Pokorny's analysis, a root with a suffix chain -*m-b^h*-, is followed by Ačařyan

⁴⁰⁰ Cf. recently Imberciadori (2022) who adduces Toch. A *tpär*, B *tapre* 'high' to *tūber* as *ér*-locative derivatives of a(n extended) root **teuHb^h*-, thus **tuHb^h-ér* 'swelling, highness'.

⁴⁰¹ This also rules out interpretations of an original *nomen abstractum* **touH-d^hro-*.

⁴⁰² The attestations also rule out any reconstructions that take the *m* form as a suffix to the form with a long vowel (like **tuh₂m-*): Dybo's Law does not operate on Lithuanian and the de Saussure effect on an *o*-grade **touH-m-* does not operate in Celtic or Lithuanian.

(1971-79 II: 206) for Arm. *fumb* and Matasović (2009: 392) for PCelt. **tumbo-*. Gk. τῦμβος would fit into this system,⁴⁰³ but a suffix comprised of the rare phoneme **b* would be unparalleled. Indeed, the proposal of multiple “root extensions” is today lacking in explanatory power. Furthermore, the Corcyrian form τῦμος with the same meaning as τῦμβος points to an *m ~ mb* alternation (EDG 1517).⁴⁰⁴ It is therefore possible to link the forms meaning ‘hillock’ together as reflexes of a substrate lexeme. But this requires separating Lat. *tumulus* from inherited *tumēō*.

verbascum ‘mullein’

Pre-form: **uerb(h)/dʰ-* | PItal. **werb/ɸbasko-*

Comp.: ?

Irreg. correspondences

Remarkable phonotactics

Semantics: plant, wild

WH (II: 756), EM (722)

Alessio (1939), Alessio (1944a: 103), André (1956: 326-7), DV (664), EDG (1269), Falk & Torp (1960 I: 802), Frisk (1960-72 II: 636-7), Derksen (2014 s.v. *virbas*), Smoczyński (2018: 1672-3)

The interpretation of Lat. *verbascum* ‘mullein’ is complex. Alessio (1939: 326) sets up a *b ~ p* alternation by connecting Lat. *verpa* ‘penis’, but this feels gratuitous. More complex are a series of spellings from glosses. André (1956: 326-7) gives *berbascum*, *barbascum*, and *vervasca*. Alessio (1939: 327-8) notes several others that may be variations of the same word, but the analysis is not as clear-cut. Several glosses list *belbe* as meaning *fellenis* or *fellonis*, themselves of unclear meaning. But one gloss says that *fellenis* means *lupicuda*. Another says that *lupicuda* means *flomus*. Some of the alternate spellings that André gives are in fact glosses that give the meaning of *flomus* (namely as *barbasco* and *vervasca*). Furthermore, *flomus* is probably a borrowing of Gk. φλόμος ‘mullein’. Thus through a concatenation of steps, it seems that *belba* = *fellenis* = *lupicuda* = *flomus* = *verbascum*. Alessio (1939: 327) even suggests correcting *fellenis/fellonis* to *flomis* = Gk. φλομῖς. This is uncertain. But if *belba* is indeed related to the *verba-* of *verbascum*, and is not due to some later Romance developments, it could attest to an *l ~ r* alternation as well as a relationship between *b* and *v* within Latin such as that proposed for *bolunda* (s.v.).

Alessio (1939, 1944a: 103) interprets the *-asco-* of *verbascum* as a Ligurian suffix based largely on placenames. If it is a suffix, and if we consider the variants in the glosses as

⁴⁰³ Gk. τῦμβος, as it has the same meaning as Gk. τάφος ‘grave’ (< **dʰmbʰ-* on comparison with Arm. *damban* ‘tomb’) was used by Pelasgianists (cf. Georgiev 1941: 70) to demonstrate that Pelasgian, after aspiration dissimilation, exhibits *um* < PIE **m* and *b* < PIE **bʰ*. But Hester (1965: 379) notes that this cannot explain the Corcyrian form.

⁴⁰⁴ Garnier and Sagot (2017: 49) propose different Greek reflexes of an IE substrate **túmbo-*, in turn from **dʰubʰ-nó-* ‘deep’.

uncertain, then *verbasum* can perhaps be derived from a root **uerb(h)-* ‘stick’. It is elsewhere found in Lat. *verbera* ‘twigs for flogging’, *verbēna* ‘leafy branch or twig’ (cf. DV 664), Lith. *vir̃bas* ‘stick, twig; type of willow’, Ru. *vérbā* ‘willow twig’ (cf. Derksen 2014 s.v. *virbas*), and PGm. **wurba-* ‘scythe handle’ (cf. Falk & Torp 1960 I: 802).⁴⁰⁵ While all comparanda reconstruct to a root **uerb(h)-*, DV (664) notes that the Slavic forms point to **b*. Thus the inherited status of this root is not guaranteed, but nor is the relationship of *verbasum* to it.

2.3.3.2 Non-inherited vs. Loan from a Known Language

ātriplex ‘orach, saltbush (*Atriplex* spp.)’

Pre-form: PItal. **ātriplek-*

Comp.: PRom. **atra/ipek-* | OFr. *arrace*, It. *atrepice*, etc. ‘orach’

PGk. **at/drap^hak-* | Gk. ἄτράφαξ/ις, ἄδράφαξ/ις, ἀνδράφαξ/ις ‘orach’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, wild or cultivated

WH (I: 76), EM (54)

Keller (1891: 61), Niedermann (1905/6: 74-5), REW (no. 759), Alessio (1955: 706), André (1956: 46), Chantraine (1968-80: 135), Frisk (1960-72 I: 181), Furnée (1972: 179), EDG (164), FEW (XXV: 684)

Gk. ἄτράφαξ/ις ‘orach’ occurs with several variations, the hallmark of a non-inherited word (Alessio 1955: 706), which cannot be easily explained by folk-etymological contamination with other words (Furnée 1972: 179, EDG 164). Lat. *ātriplex* ‘orach’ is widely accepted as a loan from Greek (WH I: 76, Alessio 1955: 706, Frisk 1960-72 I: 181, EDG 164, EM 54). This partially builds on the assumption that the Romance descendants represent a more original situation than the Latin, since they look closer to the Greek (**atrapex*, **atrapica*, **atrapicu*, etc., REW no. 759, FEW XXV: 684).

Keller (1891: 61) suggests folk etymological reanalysis after borrowing from Greek with *āter* ‘black’ and *-plex*. Niedermann (1905/6: 74-5) prefers a series of developments: Lat. *ātriplex* would have developed from **atrapex* (an analogically produced nominative from the oblique ***atrapacis*⁴⁰⁶) > **atripex* >> **atripex* >> *atripex* through dissimilation. One wonders why, if the Romance languages preserved a form more similar to the original Greek into the present day, Classical Latin would have undergone so many changes. André (1956: 46), followed by Chantraine (1968-80: 135), mentions the possibility that both Latin and Greek are independent loans from a non-IE source. This seems quite likely, but it is difficult to rule out the effects of dissimilation as the

⁴⁰⁵ Frisk (II 1960-72 II: 636-7), Alessio (1939), and Smoczyński (2018: 1673) compare Gk. ῥάβδος ‘rod, stick, wand’, but EDG (1269) removes it from comparison due to its suffix *-δ-*.

⁴⁰⁶ There is no reason for Gk. ξ to be borrowed as /k/.

cause of the aberrant Latin formation.

bardus ‘stupid’

Pre-form: **b/g^wa/Hrd-* | PItal. **bardo-*

Comp.: ?

□ Irreg. correspondences

■ Remarkable phonotactics

Semantics: pejorative

WH (I: 96-7), EM (66-7), DV (69)

Nehring (1928: 117-27), Ernout (1946: 27), Breyer (1993: 241-4), Meiser (1998: 63), Untermann (2000: 530), Zair (2018: 311-18)

Lat. *bardus* has several hallmarks of being non-IE: an invalid **DeD* root structure, **b*, and presumably **a*. Nehring (1928: 117-27), followed mostly by WH (I: 96-7) and EM (66-7), argued that it is Etruscan (cf. DV 69). He adduced several other words of similar form and meaning (*barginna*, *bargena* ‘barbarian’, *bargus/barcus* ‘without intellect’, *barō* ‘dumb idiot’, cf. further Ernout 1946: 27, Breyer 1993: 241-4). But as for *balteus* (s.v.), initial voiced stops in an Etruscan borrowing are suspicious.

Zair (2018: 311-18) is correct in his criticism of an Etruscan origin theory. He instead takes *bardus* as either a loan from Gk. βραδύς ‘slow’ < **g^wɣd-u-* with unparalleled metathesis or, more likely in his eyes, a loan from the Sabellic reflex of **g^wɣd-u-*.⁴⁰⁷ Lat. *gurdus* ‘stupid’ would be the native reflex of this root found further in Balto-Slavic (e.g. Lith. *gurdūs* ‘slow’, OCS *grъdbъ* ‘proud, haughty’)(DV 275, Zair 2018: 315-16). The phonological details are not fully clear however. Zair (2018: 316) prefers **g^word-o-* > *gurdus*, but **g^worh₃-* gave *vorāre* ‘to devour’ (DV 690, unless from an *e*-grade). After a labiovelar, **ɣ* perhaps gives Latin *ur* instead of *or* (Meiser 1998: 63, DV 275), so perhaps **g^wɣd-o-* > *gurdus*. But then we have to assume that **u* delabialized **g^w* (whereas it seems to have led to the loss of **k^w* in *ubi*, Zair 2018: 215-16, Weiss 2020: 86).

The Sabellic hypothesis is an attractive way to link *bardus* and *gurdus*. In both cases, we must accept a semantic shift ‘slow/heavy’ > ‘stupid’, which is not without parallel. But in combination with Quintilian’s report (*Inst.Orat.* 1.5.57) that *gurdus* is from Spain along with the fact that both words reconstruct to an invalid **DeD* root structure (**g^wɣd^h-o-* would give Lat. ***gurbus*), the inherited status of these words remains unclear.

burgus ‘fort, castle, watchtower’

Pre-form: **burg-* | PItal. **burgo-*

⁴⁰⁷ This requires that the Oscan reflex of **ɣ* can sometimes be *-ar-* (Untermann 2000: 530, Zair 2018: 313).

Comp.: ?

□ Irreg. correspondences ■ Remarkable phonotactics

Semantics: architecture; military

Pokorny (140-1), WH (I: 124), EM (78)

Kretschmer (1934: 100-3), Georgiev (1941: 60, 69, et alib.), van Windekens (1952: 7-8 et alib.), Biville (I: 235-7), EDG (1262), Kroonen (2013: 85), Garnier & Sagot (2020: 184)

Lat. *burgus* appears late, but a derivative *burgarii* ‘soldiers guarding a *burgus*’ is found in inscriptions from ca. 140 CE. On its own, *burgus* reconstructs to an invalid **DeD* root structure and is thus plausibly a loan. The question is from what language. It is similar formally and semantically to Gk. πύργος ‘tower, wall-tower’ and to PGm. **burg-* (cf. Go. *baurgs* ‘fortified town’, etc.). The Germanic forms are inherited from a root **b^herg-* ‘to guard’ or **b^herǵ^h-* ‘high’ (cf. Kroonen 2013: 85). The idea that πύργος reconstructs to the same root played an important role in Pelasgian hypotheses (cf. e.g. Georgiev 1941, van Windekens 1952). But they are probably only coincidentally similar. Greek by-forms with consonant alternations (Hsch. φύρκος· τεῖχος ‘wall’, φ<ο>ύρκος· ὀχύρωμα ‘stronghold’) suggest it is an early (since πύργος is in Homer) loan in Greek (EDG 1262).⁴⁰⁸

The second century attestation of Lat. *burgus* is very early for a military loan from Germanic (cf. Biville I: 235-7), but the borrowing would be formally regular. If it is a loan from Greek, Lat. *b* for Gk. *p* implies mediated borrowing via an unknown language.

carpisculum ‘type of shoe’Pre-form: **ka/Hrp-* | PItal. **karp-*

Comp.: ?

■ Irreg. correspondences □ Remarkable phonotactics

Semantics: textiles

Pokorny (581), WH (I: 172), EM (101-2)

Furnée (1972: 146), Beekes (2000: 28), Matasović (2009: 189), EDG (643, 778), Zair (2012: 83), Kroonen (2013: 244), Derksen (2014 s.v. *kurpē*)

Lat. *carpisculum* appears in the 4th century, and its late attestation has led to it being accepted as a loan (WH I: 172, EM 101). Whether it is from an unknown language or more directly from Greek is difficult to verify.

Gk. καρβάτινος ‘made of skins’ attests to a *p* ~ *b* alternation with Hsch. καρπάτινον·

⁴⁰⁸ Kretschmer (1934: 100-3) suggested a Balkan-mediated loan from Germanic, but this seems extremely unlikely. Garnier and Sagot (2020: 184) propose it is a loan from the Lydian reflex of **b^herǵ^h-* ‘high’.

ἀγρο<ι>κικὸν ὑπόδημα μονόδερμον ‘one-layer farmer’s sandal’ (Furnée 1972: 146, EDG 643).⁴⁰⁹ Other comparanda include OIr. *cairem* ‘shoemaker’ < **karafyo-mon-*, Lith. *kūrpe* ‘loafer, wooden shoe’ < PBSl. **kūr?p(i)a?*, and ON *hriflingr* ‘type of shoe’ < **hreflinga-*. Along with Gk. κρηπῖς, -ῖδος ‘man’s high boot, half-boot’, these forms seem to reconstruct to a PIE root **krh₂p-i-* ‘shoe’ (Matasović 2009: 189, EDG 778 hesitantly, Kroonen 2013: 244; but see Zair 2012: 83). Beekes (2000:28, followed by Derksen 2014 s.v. *kurpe*) instead keeps Gk. κρηπῖς separate due to its slightly different semantics and connects all the others to the root with non-IE alternations. It seems strange that Greek would have a καρβ/π- of non-IE origin beside a κρηπ- of inherited origin, both coming to mean ‘shoe’. More likely, all the shoe words are from the same non-IE source.⁴¹⁰

Furnée (1972: 146 fn. 20) proposes a Latin-internal derivation of a loaned (otherwise unattested) Gk. *καρπίς, -ίδις. It would have the same suffix as *acisculus* ‘stone mason’s hammer’ and *portisculus* ‘baton or hammer of the master of rowers’. Given the late attestation of Lat. *carpisculum*, which suggests a late borrowing, perhaps this is more likely than the suffix being added to an independent borrowing of the root **karp/b-*.

cibus ‘food’

Pre-form: **kib(h)-o-* | PItal. **kib/fo-*

Comp.: ?

□ Irreg. correspondences ■ Remarkable phonotactics

Semantics: culinary

WH (I: 210-11), EM (118), DV (112)

Thurneysen (1907: 797), EDG (693)

DV (112) notes that a root **kib^h-* violates PIE root structure constraints, but **b* is a rare phoneme. Thus in and of itself, Lat. *cibus* looks non-IE. Otherwise, it may be a borrowing from a dialect of Greek (Thurneysen 1907: 797, WH I: 210). Paulus *ex Festo* wrote *cibus appellatur ex graeco, quod illi peram in qua cibum recondunt, cibus<im> appellat*, suggesting that this was already suspected by the grammarians. Greek forms include κῖβωτός ‘wooden chest, box, cupboard’, κίβισις ‘sack’, and κίβος (or κῖβος), the further etymology of which is disputed (EDG 693). The semantic match is not perfect however, so it may be an ancient folk etymology.

conger ‘conger eel’

Pre-form: **kong-er-* | PItal. **konger-*

Comp.: **gong-er-* | PGk. **gongro-* | Gk. γόγγρος ‘conger eel; tubercular disease in olive trees’

⁴⁰⁹ Thus Lat. *carpatinus* ‘of raw leather’ could be a direct borrowing from the variant with π.

⁴¹⁰ Cf. also Gk. ἄρπις, ῥαπίς ‘a kind of shoe’, though it may be a different lexeme.

- Irreg. correspondences
- Remarkable phonotactics

Semantics: animal, wild; aquatic

Pokorny (379-80), WH (I: 260), EM (137)

Schwyzler (1930: 261), REW (no. 2144), Biville (I: 232), EDG (281), Weiss (2020: 133)

Variants in Latin as well as the Romance descendants attest to *gonger* (Pliny) and *gongrus* (It. *gongro*, *grongo* < **grongus*, REW no. 2144) alongside better-attested *conger*. If Lat. *conger* ‘conger eel’ is borrowed from Gk. γόγγρος ‘conger eel’ (WH I: 260, EDG 281), we need to explain the devoicing.⁴¹¹ Thus EM (I: 260) suggest it might be independently borrowed from the same Mediterranean source as γόγγρος.

There are several ways to explain the devoicing, but they are *ad hoc*. *Conger* may have been remodeled on the numerous other words beginning with *con-*. Or perhaps there was a dissimilation; the opposite occurs in *clucidatus* < *glucidatus* ‘sweetened’ (Biville I: 232).⁴¹² In light of these possibilities, Lat. *conger* cannot be ruled out as a loan from Greek as easily as other cases like *ballaena* and *cupressus*.

ibiscum ‘marsh mallow’, vars. *hibiscum*, *hibiscus*, *ebiscum*, *ebiscus*

Pre-form: **g^hib^(h)-* | PItal. *(*h*)*ib/fisco-*

Comp.: Gk. ἰβίσκος, var. ἐβίσκος ‘a kind of mallow’

- Irreg. correspondences
- Remarkable phonotactics

Semantics: plant, wild

WH (I: 670), EM (293)

Furnée (1972: 355), EDG (575)

Lat. (*h*)*ibiscum* ‘marsh mallow’ is clearly related to Gk. ἰβίσκος, var. ἐβίσκος ‘a kind of mallow’. Because the Latin forms are attested earlier, the Greek may be a borrowing from Latin, whose suffix has been suspected to be of Celtic origin (WH I: 670, EM 293, EDG 575). On the other hand, Furnée (1972: 355) takes the Greek variants in *i/ê* to indicate a non-IE alternation and asserts that the Latin forms are borrowed from Greek. Since both variants appear in both languages, it is difficult to determine in which direction the borrowing went.

īdus ‘the middle day of the month (13th or 15th)’

Pre-form: **Heid^(h)-* | PItal. **eidu-*

⁴¹¹ That it has *-er* for **-ros* makes it quite early, though not pre-literary (cf. inscriptional SAKROS from before the change, Weiss 2020: 133).

⁴¹² The dissimilation required here of *g—g* > *c—g* is very rare in Latin. It is much more often kept, especially in foreign words. In fact, there is a Late Latin tendency to go in the opposite direction *c—g* > *g—g* (Schwyzler 1930: 261, esp. fn. 1 and 2).

Comp.: ?

- Irreg. correspondences □ Remarkable phonotactics

Semantics: magico-religious

WH (I: 672), EM (306-7), DV (295)

Fay (1917: 213), Breyer (1993: 297), Untermann (2000: 203-4, 563-5), EDG (1053), Martzloff (2019: 305 fn. 42)

Varro (*de Lingua Latina* 6.28) and Macrobius (*Saturnalia* 1.15.14) report that Latin *īdūs* is of Etruscan origin, (claiming the Etruscan word is *itis*). Varro further tells us that *īdūs* in its current form is from Sabine. It is difficult to ignore such a straightforward statement from (in the case of Varro) an author who would have had access to Etruscan (WH I: 672-3, EM 306-7, DV 295). Breyer (1993: 297) mentions Etruscan *eitva* ‘perpetual, continual’. Its meaning seems to be known and the semantics do not make it impossible that it is at least a derivation of the donor form. Osc. **eīdūís** [abl.pl] is likely the same word as *īdūs*, differing only in stem class (a fem. *o*-stem, Untermann 2000: 203-4). This could indirectly attest to the Sabine form that Varro mentions. Without the Etruscan word itself, which one might expect to be attested given the semantic category of the surviving Etruscan sources, this cannot be confirmed.

Attempts at an IE etymology rely on what seems to be the Umbrian word for the Ides, **plenasier**, which clearly derives from **plēno-* ‘full’, presumably referring to the stage of the moon (cf. Untermann 2000: 563-5). This would indicate a similar semantic source for *īdus/eīdūís*. Fay (1917: 213) compared it to Gk. οἰδέω ‘to swell’, from the same root (**h₂eid-*) as Lat. *aemidus* ‘swollen’. Martzloff (2019: 305 fn. 42) has recently proposed a pathway by which **h₂eid-* could yield *īdus*. From a lengthened grade **h₂ēid-o-*, Eichner’s Law would prevent *a*-coloring of **ē*. The resulting **ēid-o-* is shortened by Osthoff’s Law where it yields **eid-o-*. Osc. **eīdūís** substantivized this directly whereas Latin substantivized it via conversion to a *u*-stem. The lengthened grade starting point is not attested elsewhere, so solid evidence for this pathway is lacking.

supparus ‘women’s garment of linen’, later var. *supparum*

Pre-form: **suP-Ar-* | PItal. **suppAro-*

Comp.: **S(e)ib^h-r-* | PGk. **S(e)ip^haro-* | Gk. σίφαρος, σείφαρος
‘topsail, topgallant sail’

- Irreg. correspondences ■ Remarkable phonotactics

Semantics: textiles

WH (II: 633), EM (668-9)

von Planta (1892-7 I: 236, 542, 544), Conway (1897: 220), Walde (1910: 756), Alessio (1955: 537-40), Frisk (1960-72 II: 712), Furnée (1972: 163), Biville (I: 165-7), EDG (1337), Zair (2016: 301-3), Flemešad & Olsen (2017: 214), Weiss (2020: 153)

Despite semantic differences, Furnée (1972: 163) and Biville (I: 165-7) defend the comparison of Lat. *supparus* to Gk. σίφαρος ‘topsail’ (later borrowed as *sīpharus* ‘sail’). Festus and Nonius tell us that the *supparus* was made of linen, and a few other lexemes attest to the same double meaning (Gk. φόσσων ‘thick linen garment’ and ‘sailcloth’; Gk. κάρπασος ‘fine linen, cotton’, and Lat. *carbasus* [s.v.] ‘fine linen’ and ‘sail’). Lat. *supparus* must be a loan, as it has not undergone expected vowel weakening to **supperus*, but it cannot be a direct loan from Greek. It has thus long been suspected that Oscan was the intermediate source (WH II: 633, EM 668, Flemested & Olsen 2017: 214). The Oscan word is not attested, but its existence is hinted at in an etymology by Varro (*de Lingua Latina* 5.131) when he writes *nisi id quod item dicunt Osce* ‘unless [*supparus* is called that] because they say the same in Oscan.’

Conway (1897: 220) claims that the gemination of *p* before *r* is a typical Oscan feature. But he cites von Planta (1892-7 I: 542), who shows that the gemination before *r* occurs almost exclusively with *t*, and never when a vowel separates the cluster.⁴¹³ There is only one example of an unetymological geminate *pp* in a loanword in Oscan: *Appelluneis* (von Planta 1892-7 I: 544). Flemested and Olsen (2017: 214) propose that the *u* for Gk. ι is due to the following labial, but (beyond in clitics) this only occurs in non-initial syllables as part of Oscan vowel weakening (e.g. Zair 2016: 301-3). Instead, an *i ~ u* vocalic alternation appears in other Mediterranean lexemes (cf. Alessio 1955: 537-40)⁴¹⁴ such as *frīgō* ~ φρύγω ‘to fry’, and *fīcus* ~ σῦκον ~ Arm. *tʰuz* ‘fig’.⁴¹⁵ Thus the only real argument in favor of Oscan transmission is the lack of vowel weakening. And the Oscan word, if indeed it existed, does not certainly seem to be a regular loan from Greek either. In that case, it cannot be ruled out that both Latin and Oscan received their word from Greek (cf. Walde 1910: 756) via a Mediterranean intermediary at a relatively recent date (i.e. after Latin vowel weakening). If the Greek word is not native (cf. EDG 1337, who points to vocalic alternation between σίφαρος ‘topsail’ in Arrian and Hesychius vs. σείφαρος ‘theater curtain’ in an inscription from Ephesus⁴¹⁶ alongside the similarity in shape to Akk. *šuparraru* ‘to spread out’ [cf. also Frisk 1960-72 II: 712]), then the comparanda attest to a Mediterranean lexeme with **pp ~ *bʰ* and *i ~ u* alternations.

prūnus ‘plum tree’, *-um* ‘plum’

Pre-form: **pru(C)s-no-* | PItal. **pru(C)sno-*

Comp.: Gk. προύμνη ‘plum tree’

⁴¹³ Von Planta (1892-7 I: 236) is himself skeptical of Oscan origin.

⁴¹⁴ Though he claims it arose from a substrate vowel *ū*. Some of his examples are now understood to be regular, like Celtic *i* for Lat. *ū* in loanwords and the change Lat. **u > ī* between *l* and a labial. The latter occurs in e.g. *clupeus/clipeus* ‘shield’ of unclear etymology, but also in inherited *libet* against early inscriptional LVBENS (cf. Weiss 2020: 153).

⁴¹⁵ Biville (I: 166) notes that C. Brandis, in his 1881 dissertation (p. 24, *non vidi*), mentions similar *ū:ī* pairs but she seems to suggest he explained them all via Oscan. She reports Lat. *plūrima* : Osc. *plīsima*, but the “Oscan” form is an archaic Latin form reported by Festus. As for *sūbulonem* : *sībulum*, neither form is attested in Oscan either; both again are Latin.

⁴¹⁶ Biville I: 165 vehemently rejects the existence of the variant σίπαρος.

- Irreg. correspondences □ Remarkable phonotactics

Semantics: plant, tree; fruit

WH (II: 379), EM (541)

Schmidt (1895: 131), Furnée (1972: 243, 247), Sommer & Pfister (1977: 175), EDG (1241)

EM (541) consider Lat. *prūnus* and Gk. προύμνη parallel borrowings, while WH (II: 379) follow Sommer (Sommer & Pfister 1977: 175) in considering Lat. *prūnus* to be directly borrowed from Gk. προύμνη via dissimilation of the *m*, though this would be irregular.⁴¹⁷ EDG (1241) instead follows Furnée (1972: 243, 247) in proposing that *prūnum* is a borrowing from unattested *προ(φ)νον.⁴¹⁸ The Greek word may well be of a non-IE origin based on historical arguments, but it seems unclear whether Latin is borrowed directly from the Greek or not.

taurus ‘bull’

Pre-form: **th₂eur-o-* | PItal. **tauro-*

Comp.: **th₂eur-o-* | PCelt. **tarwo-* | OIr. *tarb*, MW *tarw* ‘bull’, etc.

**th₂eur-o-* | PGk. **tauro-* | Gk. ταῦρος ‘bull’

**th₂eur-o-* | PBSI. **taurós-* | Lith. *taūras*, Latv. *tāurs* ‘aurochs’, OCS *туръ* ‘bull’, Ru. *tur*, SCr. *tūr* ‘aurochs’, etc.

**t(h₁)eur-o-* | PGM. **peura-* | ON *þjórr*, etc. ‘castrated bull’

Etr. *θevru-* ‘bull?’ (In *Θevrumines* = Ταῦρος Μινώτιος)

PSem. **tawr-* | Akk. *šūru*, Arab. *tawr*, Hebr. *šōr* ‘steer’, etc.

- Irreg. correspondences ■ Remarkable phonotactics

Semantics: animal, domestic

Pokorny (1080-5), WH (II: 650-2), EM (677), DV (607)

Lewy (1895: 4), Cuny (1910: 162), Goldmann (1938: 411), Bartholomae (1961: 1590-1), de Simone (1968-70 II: 25-6, 34-9, 95-6), Demiraj (1997: 46, 384), Orel (1998: 452), Untermann (2000: 777-8), Militarev & Kogan (2005: 309-310), Anthony (2007: 147), Derksen (2007: 500), Matasović (2009: 371), EDG (1455), Kroonen (2013: 478, 540), Weiss (2020: 170)

Lat. *taurus* ‘bull’ has not undergone the usual metathesis **auRV* > *aRūV* (**pauros* >

⁴¹⁷ The cluster *mn* remains in e.g. (*con*)*temnō* and *alumnus*. After a long vowel, the *n* seems to be lost, e.g. *sublīmis* < *līmen* (Schmidt 1895: 131).

⁴¹⁸ Furnée’s argument is that the word is Pre-Greek, and that Pre-Greek attests to a change *m* > *w*. However, if Lat. *prūnus* really attests to *προ(φ)νον, then alongside προύμνη it might rather be an example of a Pre-Greek *b ~ w* alternation (since **bn* > Gk. *mn*).

parvus 'small' but *pauci* 'few', *nervus* < **nēuros*, *alvus* < **aulos* Weiss 2020: 170), which points to a loanword (DV 607, Weiss 2020: 170).⁴¹⁹ Given that Lat. *bōs* 'cow' (instead of what should be inherited ***ūs/vōs*) is most likely borrowed from Sabellic (e.g. DV 74), I wonder if the irregularly unmetathesized *taurus* could also have come from Sabellic. U **turuf**, *toru* [acc.pl.], Osc. *ταυρομ* [acc. sg.] (of not completely certain meaning⁴²⁰) seem to attest to an unmetathesized Proto-Sabellic **tauro-*, but no other Sabellic forms with this phonetic environment exist to allow us to determine if this is regular. The Sabellic forms could just as well be loans from Latin.

The Balto-Slavic forms can be reconstructed to the same proto-form as all the rest of the comparanda so far (cf. Derksen 2007: 500). It is the Germanic evidence that shows the first problem in terms of a PIE reconstruction. It shows a diphthong **eu* rather than **au*, which cannot be accounted for in PIE terms. This mirrors Etr. *θevru-*, attested twice, in *Θevrumines* (4th c.) and *θevruclnas* (5th c.) (de Simone II 1968-70: 95-6). The former is very clearly a representation of Minotaur, but the elements are switched. This is not lacking in attestation. Kretschmer (1931: 216) names a Greek vase from Etruria with the Greek inscription Ταῦρος Μινώϊος, showing the same order of the elements. However, Etr. *θevru-* is unexpected if borrowed from Gk. ταῦρος. Etruscan also has *Taure*, a name probably borrowed from Gk. Ταῦρος (albeit very late, between the 3rd and 1st centuries) as well as *Clauce* < Γλαῦκος and *Autu* perhaps < Αὔτων (de Simone 1970 II: 25).⁴²¹ De Simone (1968-70 II: 26) remarks that this is otherwise only similar to the Germanic form, and Kroonen (2013: 540) agrees, taking the Germanic and Etruscan as independent witnesses to a **peur-* alternant of the **taur-* root.

Alb. *ter* 'bull' has been explained as a singularized plural of **tar* < PALb. **taura-* (Orel 1998: 452) or the result of umlaut from the plural (Demiraj 1997: 384), which would make it either an additional independent reflex of this word or a borrowing from Latin or Greek. Interestingly enough however, on the same page where he gives the explanation that Alb. *ter* is the result of umlaut, Demiraj (1997: 46) shows that **eu* yields Alb. *e*. Thus *ter* might attest to another language with **teur-* rather than **taur-*.

Germanic further complicates the picture by attesting to a form **steura-* 'bull' (Go. *stiur*, OHG *stior*, OE *stēor*, Engl. *steer*, etc.). It has the same *eu* diphthong as **peura-* but begins with a sibilant. Explanations include *s* mobile or the result of a non-IE phoneme

⁴¹⁹ The reflexes of Celtic **tawro-* have indeed undergone this metathesis, but there are actually no other attested forms with this phonetic environment in Celtic. Thus it may well be regular (Matasović 2009: 371). Alternatively, the Celtic reflexes have been remade on analogy to **karwo-* 'deer'.

⁴²⁰ The context only allows us to conclude it refers to a type of sacrificial animal or the quality of a sacrificial animal (Untermann 2000: 777-8). It seems to be used as an attribute of *vitlu* 'calf', leading e.g. Goldmann (1938: 411) to suggest U **turuf** could mean 'plump, fattened' < **teuhz-* 'to swell'.

⁴²¹ While the Greek diphthong *ai* occurs in Etruscan loans as both *ai* (Α(κ)ταίων > Etr. *Ataiun*) and *ei* (ἐλαίῤῥα > Etr. *eleivana*), this might represent a process of monophthongization *ai* > *ē* that was complete by the second half of the 5th century (de Simone II 1968-70: 45-6). We have no examples of a parallel change from *au* > *eu* (de Simone 1968-70 II: 26). There seems to otherwise be a strong preference to maintain the quality of *a*: there are examples of Etruscan reflecting Gk. *a*, *ā*, and even *ε* as *a* (Ἄτταλος > Etr. *Atale*, Αἰσκληπίος > Etr. *Esplace*, Πενθεσίλεια > Etr. *Pentasila*, de Simone II 1968-70: 34-9).

like **p* (cf. Kroonen 2013: 478). It is difficult to take this as an independent form (and thus evidence that Germanic borrowed a non-native phoneme as **p*~**st*) because of the existence of an almost certainly inherited **stōra-* ~ **stura-* ‘big’ < PIE **stéh₂uro-* ~ **sth₂uró-* (ON *stórr*, OSw. *stōr*, *stur* ‘big’, OE *stōr* ‘giant (adj.)’, Du. *stoer* ‘tough’, etc., Kroonen 2013: 482). Cognates⁴²² include Skt. *sthávira-* ‘broad, thick’ (**stéuh₂r-o-* with laryngeal metathesis, with the indication that the *r* is part of the root provided by the root accentuation, unusual for a *ro*-derivation adjective, cf. Kroonen 2013: 482) as well as Skt. *sthūrā*⁴²³ ‘big, strong, thick, massive’ and Av. *stūra-* ‘strong’ (zero-grade **stuh₂r-ó-*). Av. *staora* means most properly ‘heavy livestock (camel, horse, cow, donkey)’ (Bartholomae 1961: 1590-1). Thus the bovine semantics are a coincidental secondary semantic development of this unrelated root. It is possible that PGm. **steura-* is the result of contamination between **peura-* and **stōra-*.

An additional detail concerning this family of words is the question of its precise relationship to Semitic. While Cuny (1910: 162) considered both groups borrowed from a third source, comparanda of PSEM. **ṣawr-* ‘steer’ are widespread and the lexeme may even reconstruct to Proto-Afro-Asiatic (Militarev & Kogan 2005: 307-10). Some have therefore proposed that the IE family is borrowed from Semitic, given the variation in IE reflexes (Lewy 1895: 4) and the origin of domestic livestock (Anthony 2007: 147).

This lexeme is similar to *caper* and *porca* (s.v.) in that most attestations allow for the reconstruction of a common pre-form. But there seems to be just enough variation that, in combination with the possibility of ultimate Afroasiatic origin, here there is a stronger indication of a Wanderwort. The widespread attestation of the lexeme with minor variations suggests that it may have entered the Indo-European languages at an early date,⁴²⁴ and it entered Proto-Celtic (where both Goidelic and Brittonic attest to the metathesis of **-wr-*). But regardless of its antiquity elsewhere, the fact that it has resisted metathesis in Latin suggests that it entered later, at a post-Proto-Italic date.⁴²⁵ DV (607) finds it unlikely on the grounds of the semantic field, but it cannot be completely ruled out that the Latin and Sabellic forms are loaned from Greek or another IE language.

2.3.4 Core-Periphery Cases

caper ‘goat’

⁴²² The root may also be behind Lat. (*m*)*staurō* ‘to repeat, restore’ (DV 305), but then we again wonder why it did not undergo the usual metathesis.

⁴²³ EWAia (II: 768) prefers to consider this formation a *u*-extension of **steh₂-*, and is skeptical of *sthávira-* being considered a primary formation.

⁴²⁴ Rasmus Björn (p.c.) takes examples like this with extensive Afro-Asiatic matches and minor differences between the daughter languages as evidence of very old loans, borrowed just after the initial splitting of PIE.

⁴²⁵ Given the possibility that the metathesis is regular in Celtic (see fn. 419 above), and since it can neither be confirmed nor rejected for Sabellic, it is theoretically possible that the metathesis is of Italo-Celtic date. If so, then the lexeme can have entered Italic (presumably replacing the existing word) earlier, between Italo-Celtic and Proto-Italic.

Pre-form: *ka/Hp-ro- | PItal. *kapro-

Comp.: *ka/Hp-ro- | PGM. *hafra- | ON *hafr*, OE *hæfer* ‘goat’

*ka/h₂p-ro- | PGk. *kapro- | Gk. *κάπρος* ‘wild boar’

?*ka/Hp-ero- | PCelt. *kaφero- | W *caer-iwrch* ‘roe buck’, OIr. *cauru*,
cáera ‘sheep’

*g^(h)a/Hb^(h)/p-ro- | PCelt. *gabro- | W *gafr*, OIr. *gabor*, etc. ‘goat’

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: animal, domestic

Pokorny (529), WH (I: 157), EM (94-5), DV (89)

Foy (1896: 297), Johansson (1902: 312), Pedersen (1909-13 I: 92), Thurneysen (1921: 107), Meillet (1925: 9), Chantraine (1933: 221), Wagner (1957: 72 fn. 2), Frisk (1960-72 I: 783), Campanile (1974: 48), Schrijver (1991: 99), EWAia (I: 302), Gamkrelidze & Ivanov (1995 I: 435), Untermann (2000: 368), Matasović (2009: 148), EDG (438, 639), Kroonen (2013: 198), Stifter (2020: 31-4)

Latin *caper* and Umbrian forms (**kaprum**, **kabru** etc. cf. Untermann [2000: 368]) along with PGM. *hafra- (cf. Kroonen 2013: 198) and Gk. *κάπρος* ‘wild boar’ can all be reconstructed to *káp-ro-. We must assume that Greek has undergone a semantic change.⁴²⁶ In light of a lack of Balto-Slavic accentological evidence, *h₂ cannot be rejected as the source of the *a*-vocalism (Schrijver 1991: 99, EDG 639).

PCelt. *gabro- ‘goat’ looks related, but while PCelt. *-br- is the regular outcome of PIE *-pr-, the initial voiced guttural is unexpected. It could irregularly have assimilated the voicing of -b- or, even less likely, have been reshaped on analogy with the root *g^hajd- (cf. Lat. *haedus*) otherwise unattested in Celtic (Matasović 2009: 148). This irregular reflex is in contrast to a potentially regular one, still within Celtic. While Thurneysen (1921: 107, followed by WH I: 157) connected W *caer-iwrch* ‘roe buck’ to OW *caru* ‘stag’ < *k^r-yo (cf. Lat. *cervus* ‘deer’ from the full-grade) and OIr. *cáera* ‘sheep’ to OIr. *cáer* ‘clump, grapes’ “nach seinen Exkrementen benannt”,⁴²⁷ further Brythonic words for ‘roe deer’ (OW *iurgchell*, Corn. *yorch*, and OBret. *yorch*) show that it is the *jork- element that means ‘deer’ (cf. Pedersen 1909-13 I: 92, recently Stifter 2020: 32). On comparison with Lat. *capreolus* ‘roe deer’, also a derivative of *caper*, it is plausible that the *caer* in *caeriwrch* is from *kapero- (p.c. Michael Weiss). Schrijver

⁴²⁶ Meillet (1925: 9) followed by Chantraine (1933: 221) Wagner (1957: 72) hypothesized, in light of Aeolic *ἔπερος* ‘ram’, that the goat words were the result of a *k- prefixation of the boar word found in Lat. *aper* and OHG *ebur* etc. Gamkrelidze and Ivanov (1995 I: 435) suggest regular loss of *q^h- in some branches from an inherited *q^hwep^h-. It seems more likely that *τράγος* simply displaced *κάπρος* as the word for goat (Frisk 1960-72 I: 783; EDG 438, 639). Other evidence suggests that the boar word is a separate non-IE lexeme (s.v. *aper*).

⁴²⁷ Wagner (1957: 71 fn. 2) is not the only one who finds this to be a stretch.

(1991: 96) considered **kap-ero-* in Celtic against **kap-ro-* in Italic, Germanic, and Greek to point to archaic *r*-stem ablaut. While DV (89) suggests this is more likely to be analogical, Stifter (2020: 31-4) provides additional evidence that the formation is old within Celtic. The oldest Old Irish attestations have nom. *cauru*, which, against e.g. gen. *cáerach*, indicates an original PCelt. nom. **kaferūxs*, obl. **kaferāk-* with vocalic alternation explained by a reconstruction to PIE **kaperō-h₃k^w-* ‘having the appearance of a goat’, therefore ‘sheep’. The order of the sound changes involved is unparalleled but not impossible. Foy (1896: 297) used a proposed connection with poorly attested Skt. *káprth-*, *kaprthá-* ‘penis’ to support an inherited origin (followed or mentioned by Pedersen 1909-13 I: 92, Pokorny 529, EWAia I: 302, Gamkrelidze & Ivanov 1995 I: 435, Matasović 2009: 148), but this seems unconvincing and several have rejected the link (Johansson 1902: 312, WH I: 157, Frisk 1960-72 I: 783).⁴²⁸

This provides a dilemma: why would Celtic have both a regular and an irregular reflex of a word for goat if the root itself is inherited? The semantic distance of the **kaperō-* forms and the fact that they could phonologically go back to **kasero-* as well is potentially suspicious. The voicing/aspiration discrepancies, if taken at face value, are reminiscent of those found in other lexemes of non-IE origin, and a European substrate origin is suspected by several (Campanile 1974: 48, DV 89, EDG 639, Kroonen 2013: 198). But then why would all branches but the Celtic reconstruct to the same proto-form if the word were of non-IE origin (in light of the numerous cases where each branch attests to an irregularity)?

Briand (1997: 91-115) proposed deriving the forms from an old adjective to the root **kap-* ‘to take’ that would have been used to describe a snatching way of eating, then coming to denote several different animals. Wagner (1957: 73-4) noted the similarity of **kap-* (Lat. *capere* ‘to take, seize’, PGM. **hab-* ‘to have’) to **g^habh-* (Lat. *habēre* ‘to have’, OIr. *gaibid* ‘to take, seize’), which in the end (and in light of similar lexemes outside of the Indo-European languages) might hint at ultimately onomatopoeic or sound symbolic origin. But such deep-time semantic derivations cannot be proven. In the end, the existence of PCelt. **gabro-* beside **kapro-* elsewhere (and the limited confirmed extent of this and other terms for goat) might point to a different type of contact scenario than those that led to the more irregular loanwords.

⁴²⁸ EWAia (I: 302) also notes MoP *kahra* ‘kid’. The root is also suggested to go back as far as Old Persian, but these suggestions are made based on two personal names, attested in the Elamite Persepolis archives, which are suggested to represent Old Persian: *qa-pu-ra* /*Kapura*/ = OP **kafra-* and *qa-ap-ri-ya* /*Kapriya*/ = OP **kafriya-*. Gershevitch (1969: 199) writes that, if the interpretations as Old Persian are correct, then both “may belong to [MoP] *kahra* ‘kid’, which Henning used to relate to Lat. *capere*.” Hinz (1973: 114), though he agrees that *qa-ap-ri-ya* transcribes OP **kafriya-*, finds it difficult to believe that the name means ‘young goat’ and rather interprets it as **ka-friya-* “wie lieb!”. Hinz and Koch (1987 I: 413) take *qa-pu-ra* (given under *qa-bu-ra*, as Elamite does not consistently differentiate voicing) as representing Aram. *kabbūra* “the stout one.” Thus the root’s antiquity in Iranian is uncertain. Besides MoP *kahra* ‘kid’ exists at least Zazaki *kavir* ‘sheep’, but the forms are seemingly restricted to the Western Iranian languages. Thus a later loan within Iranian is theoretically possible, and these forms are best not considered independent evidence.

hordeum ‘barley’Pre-form: *g^h(o)r(s)d^(h)- | PItal. *χor(s)d-ejo-Comp.: *g^hersd- | PGm. *gerstō- | OS, OHG *gersta* ‘barley’*g^hers(d^(h))- | Hitt. *karaš-* ‘wheat, emmer wheat’*ǵ^hrsd^(h)- / *ǵ^hrid^(h)- | PALb. *drisdā- | Alb. *drithë* ‘cereal, grain’?*g^hriHd^h- | PGk. *k^hrīt^h- | Gk. κριθή, epic nom. sg. κρῑ ‘barley’?*g^hrijo- / *g^heritV- | Arm. *gari* ‘wheat’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, domestic

Pokorny (446), WH (I: 656-7), EM (299), DV (289)

Huld (1983/4: 149), Demiraj (1997: 145-6), Olsen (1999: 439), Rieken (1999: 63-5), Kloekhorst (2008: 444), Martirosyan (2009: 199), EDG (779), Kroonen (2013: 175), Schumacher & Matzinger (2013: 261), Thorsø (2020), Kroonen et al. (2022: 7)

Latin *hordeum* and PGm. *gerstō- reconstruct to different ablaut grades of a root *g^hersd- (cf. Kroonen 2013: 175, *pace* EM 299 who gives *t as an option for Germanic), with Latin in the zero or o-grade and Germanic in the e-grade. A final *d^h is also possible for the Latin pre-form. Hitt. *karaš-* ‘wheat, emmer wheat’ requires a derivation from a root shape *g^hersT- to undergo the lowering of *e to a, which occurs before *RCC, after which a word-final dental would regularly be lost (Kloekhorst 2008: 444).⁴²⁹ Thus Kloekhorst reconstructs *ǵ^hersd^h- for *karaš-*, though DV (289) gives *ǵ^hersd-. The latter would be a perfect match for the Italic and Germanic forms.

These forms are traditionally explained as dental enlargements to the PIE root *ǵ^hers- ‘to bristle up, stiffen’ (LIV2 s.v.) e.g. in Lat. *horreō* ‘to stand on end, tremble’ (for Hittite, Rieken 1999: 63-5; for Latin, WH I: 656-7, EM 299). The Bannengsmotiv would be barley’s long, bristly awns. As to Alb. *drithë* ‘cereal, grain’, if *th* is specifically from PIE *sd (as opposed to *dh* < *sd^h), we should reconstruct PALb. *drisdā from something like *ǵ^hrsd-eh₂- (Huld 1984: 149). Otherwise, if *sd and *sd^h both became *dh*, which was devoiced in word-final position (Schumacher 2013: 261), *th* could have been leveled from a paradigm like *dridh, pl. *dridhā (Thorsø 2020: 257) < *ǵ^hrsd-eh₂- or *ǵ^hrsd^h-eh₂-. The latter matches the Anatolian, Italic, and Germanic pre-forms. Further connections however have led many to suspect that this lexeme is a loanword from a non-IE language (e.g. Demiraj 1997: 146, DV 289, Martirosyan 2009: 1999, Kroonen 2013: 175).

The Albanian form can alternatively be reconstructed as *ǵ^hriḍ^(h)- (cf. Demiraj 1997:

⁴²⁹ Kroonen (2013: 175) takes the lack of dental at face value and instead connects *karaš* to PGm. *hersja(n)- ‘millet’.

146), making it look strikingly similar to Gk. κριθή ‘barley’. The Greek form lacks any trace of an internal sibilant, and therefore fits better with Arm. *gari* ‘wheat’, especially given its epic by-form κριθ̄. But both Greek forms can reconstruct to PGk. **krīth-* (EDG 779) < **g^{hr̄}īd^h-*. While Arm. *gari* could reconstruct to a Lindemann variant of **g^{hr̄}iom* (Olsen 1999: 439), a reconstruction with **t* is also possible (Thorsø 2020: 256-8).

Kroonen et al. (2022: 7) keep the Greek and Armenian forms separate from the Anatolian, Italic, Germanic, and Albanian ones due to their formal aberration. The presence of a formally matching Hittite cognate indeed makes the latter group look inherited. Accepting the aberrant Greek and Armenian forms as part of a non-IE lexeme does not require accepting irregular correspondences that are without parallel (cf. the aspiration alternation of *lēns* ~ λάθροϛ and the vacillating presence of a sibilant in the *fracēs* group), but it does make the unity of the rest of the forms difficult to explain. Instead, if related, they may be seen as peripheral forms. Their reflex of the inherited formation was mediated to them by another language (whether IE or not), implying they had lost the inherited formation or had it replaced.

porca ‘ridge of soil between furrows’

Pre-form: **p(o)rk-* | PItal. **porkā-*

Comp.: **p_{r̄}rk-* | PCelt. **prikā-* | W *rhych* ‘furrow’, etc.

**p_{r̄}rk-* | PGm. **furh-* | OHG *furuh*, OE *furh* ‘furrow’, etc.

**b^(h)r̄g^h-* | Lith. *biržė* ‘row, furrow; timber tract; border mark’, Latv. *birze* ‘furrow, row’

**b^(h)o/ark/ǵ^h-d^(h)-* | PSlav. **borzdà* | OCS *brazda* ‘furrow’, Ru. *borozdá* ‘furrow, harrow, canal’, etc.

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: geography; agriculture

Pokorny (821), WH (II: 340), EM (522), DV (481)

ESSJa (II: 220), Lühr (1988: 318-19), Holzer (1989: 51-4), Olsen (1999: 953), LIV2 (s.v. **perk-*), EWAia (II: 100), Derksen (2007: 59), Kroonen (2011: 137), Kroonen (2013: 160), Matasović (2013: 79)

The Latin, Germanic, and Celtic words are usually derived from a root **perk-* (LIV2: 475; DV 481, Kroonen 2013: 160). The only potential evidence for this group descending from a verbal root is semantically remote Lith. *peršėti* ‘to itch’, which leads LIV2 (s.v. **perk-*) to reconstruct a meaning ‘graben, aufreißen’, but this feels like too great a compromise—especially in light of the fact that the only comparandum outside this group is Rigvedic *pársāna-*, perhaps ‘a low sunken place’, but whose meaning is not entirely known (EWAia II: 100). On the other hand, Nw. dial. *fere* ‘ridge between two furrows’ < **ferhan-* is important because it 1) looks like an *n*-stem formation almost

identical to the Rigvedic form and 2) suggests an *ablauting* paradigm that would make it inherited (Kroonen 2013: 160), whether or not the *n*-stem itself has been inherited as ablauting (cf. Lühr 1988: 318-19, Kroonen 2011: 137).

Lith. *bir̃žė* as if < **br̃ǵh-* is so similar in form and meaning to the Italo-Celto-Germanic **pr̃k-*, differing only in voicing/aspiration, that Holzer (1989: 51-4) argues that they are connected via his IE Temematic language. But this alternation in voicing and aspiration occurs in several other lexemes of non-IE origin. Further Slavic comparanda point to non-IE origin. A connection between PSlav. **borzdā*, and e.g. Skt. *bhr̃ṣṭī-* ‘point, top, spike, tooth’ (ESSJa II: 220) requires an element with **-d^(h)-* (a Temematic reflex of **-t-*, Holzer 1989: 51-4; **-d^heh₂-*, Matasović 2013: 79) not reflected in the Baltic forms (Derksen 2007: 59, Matasović 2013: 79). Perhaps, rather than Lith. *bir̃žė* representing a quasi-PIE **ǵ^h* that has undergone satōmization, it corresponds to a sigmatic element that alternates with **zd* in Slavic. A potentially similar situation occurs between *fracēs* and its comparanda.

The appurtenance of the aberrant Balto-Slavic forms need not necessarily prove that the quite Indo-European-looking Italic, Celtic, and Germanic forms are not inherited (although it remains a possibility). Instead, the inherited lexeme could have been mediated to Balto-Slavic via indirect means, IE (like Holzer’s Temematic) or not.

Arm. *herk* ‘fallow land just broken up’ < **perg-* is similar on semantic and formal grounds. While the Balto-Slavic forms hint at satōmization, the Armenian form cannot reconstruct to a palatovelar. Similar to the Balto-Slavic words, several interpretations are possible. Olsen (1999: 953) suggests it could be related to the **pr̃k-* forms but from a centum substrate within Armenian. Otherwise it shows the alternations we expect to see in non-inherited words. However, it would also be the most semantically distant comparandum, given that all the other comparanda attest a specific meaning ‘furrow’ or ‘ridge between furrows’. Thus I leave it out for now.

2.3.5 Methodologically Difficult to Delimit Comparanda

campus ‘flat land, field’

Pre-form: **ka/Hmp-o-*, **kh₂e-n-p-* | PItal. **kamp-*

Comp.: **ka/h₂mp-*, **kh₂e-n-p-* | PGk. **kamp-* | Gk. κάμπτω ‘to bend, curve’,
καμπή ‘curve, curvature’

**ka/omp-*, **kh₂e-n-p-* | PBalt. **kamp-* | Lith. *kāmpas* ‘corner, angle,
bend’, etc.

**ka/omp-t-* | PSlav. **kǫtъ-* | OCS *kǫtъ* ‘corner’

**kump-* | PBalt. **kump-* | Lith. *kuĩpas* ‘crooked, bent, hooked’, etc.

?**ka/e/omp-*, **kh₂e-n-p-* | PIr. **kamp-* | Sogd. *nk’np* ‘to bend; subdue’,

etc.

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: geography

Pokorny (525), WH (I: 148-9), EM (90-1), DV (86)

Schrijver (1991: 424-35), Cheung (2007 s.v. *kamp*), Derksen (2007: 244), Matasović (2009: 186), EDG (632, 1341), Kroonen (2013: 207), Derksen (2014 s.v. *kaĩmpas*, *kuĩmpas*), Smoczyński (2018: 630), Pronk (2019)

Lat. *campus* ‘field’ has been connected to Gk. καμπή ‘curve, curvature’ through the assumption that it originally referred to a depression or curvature of the earth (Pokorny 525, WH I: 148). With this being the case, it is difficult to know where to draw the line for including forms as comparanda.

The most likely to belong are several Balto-Slavic forms. PBalt. **kamp-* can be from **kamp-* or **komp-* which either confirms the *a*-vocalism of the Latin and Greek forms or establishes an irregular *a* ~ *o* alternation with them. Alternatively, all the forms so far could be from a root which LIV2 reconstructs as **kamp-*. A reconstruction not requiring PIE **a* would be **kh₂emp-*, though EDG (632) is suspicious of this root structure. Perhaps this would be solved by a root **kh₂ep-* with a nasal infix **kh₂e-n-p-* > **kh₂emp-*. Derksen (2014 s.v. *kaĩmpas*) notes that it is difficult to separate the inherited forms from Germanic borrowings. But PSlav. **kǫtǫ-* ‘corner’, if from **ka/omp-to-*, seems to attest to a Balto-Slavic root. On the other hand, the Slavic form cannot contain the onset **kh₂-*, as that would yield PSlav. **x-* (Pronk 2019: 149). Between the Baltic and Slavic forms, the meanings are similar enough to consider them related, but this either removes them from the Latin and Greek forms or rules out **kh₂emp-* as a pre-form.

As to Lith. *kuĩmpas* ‘crooked, bent, hooked’, which reconstructs at face value to PBalt. **kump-*, Derksen (s.v. *kuĩmpas*) argues that it is secondary, its *u*-vocalism coming from semantically similar words like *kuprà* ‘hump, hunch, back’. Kroonen (2013: 207) followed by Smoczyński (2018: 630) however takes this from zero-grade **kmp-*. Additionally compared is PGm. **hamfa-* ‘maimed’ (Go. *hamfs*, OS *hāf*, OHG *hamf*), reconstructed to **kómp-o-* by Kroonen (2013: 207). Followed partially by Pronk (2019: 149), he considers the possibility that all forms can be from a root **kemp-*. Lith. *kaĩmpas* < **komp-o-*, Lith. *kuĩmpas* < **kmp-o-*, and Gk. κάπτω < **kmp-je-* with secondary nasal infix (cf. χανδάνω ‘to hold’ < **g^hnd-* with nasal reintroduced from full-grade **g^hend-*). This rules out the Latin unless it the result of *e* > *a* after a plain velar in **kemp-* (Schrijver 1991: 424-35),⁴³⁰ an uncertain development. Pronk (2019: 149) defends the omission of *campus* from this group on semantic grounds and further considers the appurtenance of the Germanic forms questionable.

Given the difficulty in reconstructing the vocalism of this root in an Indo-European way

⁴³⁰ He does not use *campus* as an example due to its murky etymology.

while maintaining a connection with all the comparanda, it is possible that we are dealing with a non-IE lexeme (DV 86, EDG 632). This conclusion is surprising in light of Plr. **kamp-* ‘to bend’ (Parth. *nkmb-*, BSogd. *nk’np-* ‘to bend’, etc.).⁴³¹ Cheung (2007 s.v. *kamp*) does not seem to reject a connection between the Iranian and European forms, but does seem to suggest their connection is not strong enough to invoke IE origin given the irregularities at hand. The alternative is to isolate Lat. *campus* from the rest of the forms that can derive from **kemp-* or to remove the Balto-Slavic forms from a group that can reconstruct to **kh₂e-n-p-*.

cūpa ‘cask, tub, barrel’

Pre-form: **kuHp-* / **ke/oup-* / **koip-* | PItal. **kūpā-* / **koupā-* / **koipā-*

Comp.: ?

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: vessel

Pokorny (588-92), WH (I: 310-11), EM (158), DV (155)

Furnée (1972: 176-7, 284), EWA (V: 872-4), EWAia (I: 370, 385), Schrijver (1991: 245-6), Beekes (1996: 223-7), Lubotsky (1998: 76), Deshayes (2003: 390), EDG (29, 801, 804), Kroonen (2013: 308), (2020: 444), Weiss (2020: 155), van Sluis (fthc.)

Latin *cūpa* ‘cask, tub, barrel’ is the same as *cuppa* ‘cup’, the latter being a *littera* rule variant (Weiss 2020: 155).⁴³² There are several options for reconstruction, but selecting one depends on the comparanda that are accepted. The decision is difficult to make.

A comparison with semantically more distant Skt. *kūpa-* ‘well, pit, hole’ (widely compared, though hesitantly by EWAia I: 385) and PGm. **hūfa-* ‘hull, hive’ would require **kuHp-* (cf. Weiss 2020: 155). Comparison with Hsch. κύπη- τρώγη ‘gap, hole; type of ship, hut’⁴³³ and semantically closer Gk. κύπελλον ‘bulbous drinking vessel, goblet’ would rule out a laryngeal (cf. Schrijver 1991: 245-6). Some (DV 155; EDG 801, 804) compare both groups, which requires the reconstruction of a non-IE **u ~ ū* alternation.

⁴³¹ The Indo-Aryan root **kamp-* means ‘to shake, tremble’, and Cheung (2007 s.v. *kamp*) is unsure if it belongs together with Iranian **kamp-*.

⁴³² Celtic forms including W *cib* ‘vessel, coffer’, Bret. *kib* ‘drink, cup’, W *cibell* ‘skin, hide, shell’, and Bret. *kibell* ‘tub, container’ are often taken as loans from Lat. *cūpa* (cf. Deshayes 2003: 390), but Latin *ū* is not usually borrowed into Brythonic as **ī*, suggesting that something more indirect occurred. The expected reflex of Lat. *ū* is Brythonic **ū* (cf. Lat. *mūtus* ‘mute’, Bret., W *mud* ‘dumb, mute’). Other potential examples of the unexpected outcome are not straightforward. Lat. (*ferrum*) *dūrum* yields Bret. *dir* ‘steel’ but W *dur* ‘steel’. Lat. *scrīpulus* ‘sharp stone; 1/24 of an ounce’ yields W (y) *sgrubl* and (y) *sgribl* ‘work animal, livestock; unit of currency’ along with Mr. *screpul(l)*, but within Latin there is also *scrīpulum* ‘small unit of measure’. Thus the exact correspondences and their chronology require further work.

⁴³³ I cannot locate the source of the meanings ‘type of ship, hut’. Hesychius only seems to give τρώγη (Cunningham 2018-20 II: 696).

Greek, Indo-Iranian, and Germanic attest to further similar forms. Greek has forms with β (Hsch. κύβος: τὸ τρύβλιον ‘bowl’) and a nasal (κύμβη, κύμβος ‘cup’) as well as an aspirated variant of κύπελλον, namely κύπελλα ‘hollows of the ears’ (EDG 801, 804). There is even a κύμβη (appearing once as κύβη) ‘head’ (EDG 802) that seems to have undergone the same semantic shift as Ger. *Kopf* ‘head’ < *‘vessel’. Skt. *kumbhá-* ‘jar, pitcher’ and YAv. *xumba*, MoP *xumb* ‘jar’, must also be related to these Greek forms (EWAia I: 370, Lubotsky 1998: 76), though from a form **k^humb^h-* with aspirates.⁴³⁴ PGm. **kumb/pan-* ‘basin, bowl’ belongs at least with the Indo-Iranian forms and Gk. κύμβη (EWA V: 872-4). In various combinations, they are often considered to be non-IE ‘culture words’ (Furnée 1972: 176-7, EWAia I: 370, Beekes 1996: 223-7, Kroonen 2013: 310, Šorgo 2020: 444).⁴³⁵

WH (I: 310-11) and Beekes (1996: 223-6) make much wider comparisons, including Gk. γύπη ‘cavity in the earth, den, corner’ and PGm. **kuban-* ‘shed’ < **gub^h-on-*. While Kroonen (2013: 308) sees the Germanic material as unrelated to the Greek due to the *u* ~ *ū* discrepancy, EDG (292) takes this as evidence of a substrate lexeme. Beekes (1996: 227) saw it as a widespread Wanderwort, perhaps even sound symbolic. WH (I: 310-11) go even further in comparing Germanic words for hill all the way to Lat. *campus*, seeing behind all of the material a primordial meaning ‘concave depression, convexity, bend.’ At this point, we would require vocalic alternations even beyond *u* ~ *ū*. Intellectually, it is interesting to speculate on the existence of a timelessly ancient substrate lexeme **KV(m)B-* behind all of this material, but it is beyond any empirical ability to demonstrate.

From the very beginning, it was difficult to draw the line in terms of where Lat. *cūpa* fit in relation to the several groups of lookalike forms, many with the basic or derived meaning ‘vessel’ and together or separately themselves considered of non-IE origin. It is likely that *cūpa* fits somewhere in this spectrum. But because the most accurate comparison remains elusive, the exact irregular alternations and the geographic distribution of the substrate lexeme cannot be accurately determined. Without those, the lexeme loses its methodological value.

glēba, var. *glæba* ‘lump of earth, clod’

Pre-form: **g^(h)leHb^h-* | PItal. **glēfā*

Comp.: ?

■ Irreg. correspondences

■ Remarkable phonotactics

Semantics: geography

⁴³⁴ Given that PIE did not have **k^h*, one could reconstruct **kHumb^h-* to yield the Iir. base, but it could instead represent a non-IE phoneme.

⁴³⁵ Mlr. *comm* ‘vessel’, W *cwm* ‘deep narrow valley’, Bret. *komm* ‘trough’, Gaul. *cumba* ‘bottom of a ship’ look at face value to continue PCelt. **kumbā-* (cf. Beekes 1996: 224), but could be loans from Gk. κύμβη or from Lat. *cymba* ‘bark, small vessel’, itself from Greek.

Pokorny (356-64), WH (I: 606-9), EM (277-8), DV (264-5)

Vaniček (1881: 83), Rohlfs (1972: 19), Stang (1972: 22), EWA (V: 556-8), Matasović (2009: 161), Kroonen (2013: 293-4), Derksen (2014 s.v. *klēbti*, *glaūbti*, *glēbti*), Weiss (2020: 181)

It seems reasonable that Lat. *glēba* ‘lump of earth, clod’ is related to Lat. *globus* ‘round, compact mass.’ But frequently, Lat. *glomus*, *-eris* ‘ball-shaped mass’⁴³⁶ is connected as well. Vaniček (1881: 83) derived *glomus* from **glob-mo-*, but this is uncertain. Presumably it should give **glommo-*, but examples (cf. *glūma* < **glou̯b^h-mā-*, Weiss 2020: 181) admittedly involve a long vowel and are assumed to have undergone subsequent degemination. WH (I: 606-9) take them as different extensions of a root **gel-* ‘to ball up’ (cf. Pokorny 356-64): **gle-b^h-* beside **gle-m-*, but additional evidence for such a root is lacking. If indeed connected, the Latin material suggests a non-inherited *b^h ~ m* alternation (cf. DV 264-5).⁴³⁷ The connection is semantically attractive but difficult to confirm. Also difficult is where to draw the line in terms of comparanda.

DV (264-5) lists several Baltic, Germanic, and Celtic potential matches for the Latin family. The closest matches are the Baltic, which suggest that the original meaning of the root referred to squeezing together. Matasović (2009: 161) is suspicious of the connection of OIr. *glomar* ‘bridle-bit, muzzle’ and OHG *klamma* ‘trap, gorge’, OE *clam(m)* ‘tie, fetters’. He writes that the semantics for the Celtic form would need to go from ‘ball-shaped mass’ > ‘gag’ > ‘bridle’, although it seems possible that an original ‘squeeze together’ > ‘tie’ (yielding the Germanic forms) > ‘muzzling a horse’ is possible. It is indeed speculative.

As to the Baltic forms, Lithuanian shows a voicing alternation in verbs for ‘to embrace’: standard *glēbti*, Žemaitian *klēbti*, which would require the reconstruction of a **g^(h) ~ *k* alternation if it is not secondary. Lith. *glaūbti* ‘to clasp one’s bosom’ < **g^(h)lou̯b^h-* requires a form with a **u* whose absence elsewhere cannot be explained from an inherited perspective (Derksen 2014 s.v. *glaūbti* compares it to the **a ~ *ai* alternation found in substrate words).⁴³⁸

Within the semantic sphere of *umarmen* and likely similarly requiring the reconstruction of a **g ~ *k⁴³⁹* alternation are OHG *klāftra* and MHG *lāftra* ‘fathom, length of the outstretched arms’. The latter is poorly attested, but would through PGm. **hlēftrō-* reconstruct to **klēp/b^h-* if it is not somehow secondary. While DV (264) does

⁴³⁶ In part due to Romance forms continuing **glem-*, *glomus* is often considered to be from an earlier **glemos-* (‘sogenannte o-Umlaut’ [WH I: 609], rounding from velar *l* [Schrijver 1991: 468] and non-front vowel in the next syllable [DV 265]).

⁴³⁷ The reconstruction of **b^h* seems to be required by e.g. Calabrian *gliefa* and Salentine *ghiefa*, *gnifa*, plausibly reflexes of an Oscan reflex of **gleb^h-* (Rohlfs 1972: 19).

⁴³⁸ For the semantics, cf. further Lith. *glėbys* ‘embrace, armful’.

⁴³⁹ Stang (1972: 22) notes this remarkable circumstance of both variants in both Baltic and Germanic, suggesting a “Parallelwurzel”.

not find *klāftra* semantically close enough, EWA (V: 558) defends the connection through a comparison with Gk. ὄρῳια ‘fathom’ < ὀρέγω ‘to stretch’. EWA (V: 556-7) takes *klāftra-* as an instrumental construction **glēb^h-treh₂-* with a lengthened grade paralleled only by Lat. *glēba* and with possible *o*-grade **klaban-*: ON *klafi* ‘yoke, packsaddle’, etc. Kroonen (2013: 293-4) reconstructs **glēmb^h-(n)-* with Kluge’s Law effects for OE *climban/climman*, MDu. *climmen/clemmen* ‘to climb’, MHG *klimmen/klimpfen* ‘to squeeze, to climb’, arguing that the climb meaning must be secondary to the clasp/clamp meaning based on other formations (crucially zero-grade **klumpan-* ‘lump’). While he takes these as isolated to Germanic, they look like they could morphologically and semantically belong to the root of **k/glēb^h-treh₂-*.

But it is clear that this is a slippery slope. EWA (V: 557) further adduces OHG *kolbo* ‘cob, cudgel’, which Kroonen (2013: 309) takes from a separate root PGM. **kulba(n)-* ‘round object’ < **gleh₁b^h-*. It is to this latter group that he compares Lat. *globus* (with pretonic shortening in **gloh₁b^h-ō-*), *glēba*, and the Baltic forms. The semantics of the group are vague enough to allow either of these interpretations. But other Germanic words (Du. *klont* ‘lump’ and Engl. *clod* itself) from a root like **glud^h-* fit semantically and differ phonologically only slightly. Perhaps, like for some of the forms under *cūpa*, there is an element of sound symbolism to formations of this type. In the end, a conservative approach is to keep Lat. *globus* and *glēba* separate from *glomus*, and to consider the *k ~ g* alternations in both Baltic and Germanic as secondary. If this is too conservative, it is methodologically difficult to decide how many additional forms to compare given the vague semantics of the words involved.

mōrum ‘mulberry’

Pre-from: **moHr-* / **mōr-* | PItal. **mōro-*

Comp.: **mor-* | PGk. **moro-* | Gk. μόρον ‘black mulberry, blackberry’

?**mor-* | PArm. **mor-* | Arm. *mor* ‘blackberry’

?**mor-* | PCelt. **mor-* | W *mer-wydd* ‘blackberry’

?PU **mura-* | Finn. *muurain*, Tundra Nenets *məraŋka* ‘cloudberry’, etc.

?PCelt. **smi(y)ar-* | W *mwyar*, OCo. *moyr*, MBret. *mouar* ‘blackberry’

?PCelt. **smeyir-* | OIr. *smér* ‘blackberry’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant; fruit

Pokorny (749), WH (II: 114), EM (415)

Berger (1956: 22-6), Ciorănescu (1958-66 s.v. *zméurǎ*), Vasmer (1959-61 s.v. смород), Battisti (1960: 380), Frisk (1960-72 II: 256), Turner (1966-9 I: 562), Hamp (1973: 167), Rédei (1988: 287), Sammallahti (1988: 538), Campbell (1990: 165-6), Schrijver (1991: 123-4), Biville (II: 23), Orel (1998: 245), Trask (2008: 285), Martirosyan (2009: 474),

Matasović (2009: 347), EDG (968), Topalli (2017: 939), Cunningham (2018-20 II: 866), GPC (s.v. *morwydd*)

Lat. *mōrum* ‘mulberry’ differs in vowel length from but is indeed often suspected of being a loan from Gk. μόρον ‘black mulberry, blackberry’ (WH II: 114, Frisk 1960-72 II: 256, EM 415, EDG 968). This is in part due to Hsch. μῶρα: συκάμυνα, corrected to μόρα already by Marcus Musurus but maintained by Cunningham (2018-20 II: 866). Biville (II: 23) suggests that the vowel was lengthened upon borrowing into Latin due to primary syllable accentuation, but the only other potential cases of this (*rāpum* and *līnum*) are not certainly loans from Greek. As to whether Arm. *mor* is borrowed, Martirosyan (2009: 474) notes that it would have to have been prehistoric, since the word is widespread in the dialects. W *merwydden* < **mor-* cannot be from Latin due to the vowel length (Hamp 1973: 167)⁴⁴⁰ and thus would have to be a loan from Greek. If the *ō* ~ *o* alternation between Latin and the other forms is taken at face value, it is reminiscent of PIE ablaut. If Hsch. μῶρα is legitimate, it remains possible that all attested forms are loans from Greek.

Another Celtic blackberry word of similar shape shows discrepancies between Brythonic and Goidelic: OIr. *smér* ‘blackberry’ ~ W *mwyar*, OCo. *moyr*, MBret. *mouar* ‘blackberry’. Matasović (2009: 347) reconstructs PCelt. **smēro-*, with the Brythonic forms potentially being collectives with an *-*aro-* suffix to account for their hiatus. But PCelt. **ē* should yield OIr. *ía* unless this is prevented by a high vowel in the next syllable. Hamp (1973: 168-9) instead reconstructs **smi(i)ar-*, but OIr. *smér* requires a pre-form like **smejir-*. It is semantically plausible that these discrepant Celtic forms represent the same lexeme as *mōrum* (Hamp 1973 168-9, Schrijver 1991: 123-4). Formally, they verge on being too dissimilar.⁴⁴¹

Alb. *mare* ‘strawberry tree’ is of similar shape to the *mōrum* group. Alb. *a* can be from **o*, but the date at which this change would need to have occurred rules out a borrowing from Gk. μόρον.⁴⁴² Topalli (2017: 939) regards it as a borrowing from MoGk. κουκουμάρα ‘strawberry tree’ (< Gk. κόμαρος, Hsch. κύμαρος: κόμαρος), but it seems unlikely that the initial syllables would simply be deleted. North Caucasian forms of similar shape to κόμαρος but with the same meaning as μόρον (cf. Chechen *komar* ‘mulberry’, Ingush *komar* ‘raspberry, mulberry’, and Batsbi *kumel* ‘raspberry’) could again point to a substrate lexeme with the vacillating presence of a prefix.

⁴⁴⁰ W *mor-* and *mōr-wýdd* can be late loans from Lat. *mōrum*, though interestingly they are attested around a century earlier than *merwydd* (GPC s.v. *morwydd*).

⁴⁴¹ Hamp further adduced Rom. *zmeură*, MoGk. σμέουρο ‘raspberry’, but these forms are complicated. Ciorănescu (1958-66 s.v. *zméură*) asserts that the Greek is borrowed from Romanian. The Romanian itself may represent original neuter **smeu* with the *r* introduced from the plural *smeuri*. He alternatively proposes a connection with Slavic words for ‘currant’ (cf. Ru. *smoródina*), but these are more plausibly derived from PSlav. **smordъ-* ‘stench’ (cf. Vasmer 1959-61 s.v. смороди).

⁴⁴² Orel’s (1998: 245) analysis as a borrowing from Lat. *marum* ‘cat-thyme’ (in the germander family) can be rejected as it is an entirely different kind of plant.

Martirosyan (2009: 474 with lit.) collects several other potential comparanda including Gk. *μυρίκη* ‘tamarisk’, Lezgič *mer* ‘raspberry, blackberry’, Kartv. **marc’q’w-* ‘strawberry’ (Georg. *marc’q’wi*, Svan *bäsq’i*), Lak *mamari* ‘blackberry’, Darwa **mVmVrV* ‘raspberry’, Chechen *mürg* ‘guelder rose’, PU **mora* ‘raspberry, cloudberry’ (more accurately **mura-*, Redei 1988: 287, Sammallahti 1988: 538),⁴⁴³ even Hitt. *mu-uri-uš* ‘grape’. In the end he finds a widespread non-IE word **mor-/mōr-/mur-* ‘mulberry; blackberry; tamarisk’ > ‘raspberry, strawberry; grapes’.⁴⁴⁴

In the end, it is highly unlikely that all the comparanda mentioned are actually related. But it is difficult to draw the line in terms of appurtenance. We may be dealing with a widespread lexeme of the shape **(s)mVr-* and the general meaning ‘berry’. But the most conservative scenario, in which Lat. *mōrum* ‘mulberry’ is borrowed from a poorly attested Greek word, cannot fully be ruled out.

2.4 Non-IE Origin in Latin Rejected

2.4.1 No Positive Evidence of Borrowing

aqua ‘water’

Pre-form: **h₂ek^w* | PItal. **akwā*

Comp.: **h₂ek^w* | PGM. **ahwō-* | Go. *ahva* ‘body of water, river’, OHG *aha* ‘river’, etc.

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: geography

Pokorny (23), WH (I: 60), EM (41), DV (48)

Krahe (1962: 294), Beekes (1998: 459-61), Kroonen (2013: 7)

Lat. *aqua* and its comparanda in Germanic are considered potentially substrate due to their limited distribution and opposition to the other, similar yet widespread inherited water word **h₂ep-* (Beekes 1998: 459-61,⁴⁴⁵ DV 48). If only these two comparanda are considered, then there is an exact formal match restricted to Italic and Germanic, with no morphophonological features pointing to a non-IE origin (cf. Kroonen 2013: 7).

arcus, -ūs ‘bow; arch’

⁴⁴³ See further Campbell (1990: 165-6). He however would further link these words to the apple word (*mālum*) which is certainly going too far.

⁴⁴⁴ Berger (1956: 22-6) followed by Battisti (1960: 380) adduces Burushaski *biranč* ‘mulberry’, but the reconstruction to **moron-š* is forced. It is a loan from a Pamir language like Khowar *mrač* or Shina *marōč*, which continue **madhuravrkša* ‘a tree with sweet fruit’ (cf. Turner 1966-9 I: 562). Nor is it clear that the numerous, formally aberrant Basque forms which they both adduce (cf. *masusta, marzuza*, etc.) are related (cf. Trask 2008: 285 on the forms).

⁴⁴⁵ In part due to the numerous European hydronyms of the shape **aC(a)-* noted by Krahe (1962: 294).

Pre-form: **h₂erk^w-o-* / **h₂erk-uo-* | PItal. **arkwo-*

Comp.: **h₂erk^w-ō-* / **h₂erk-uō-* / **h₂erk-uh₂-* | PGm. **arhwō-* | ON *qr*, OE *earh*,
etc. ‘arrow’

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: weapon

Pokorny (67-8), WH (I: 64), EM (44), DV (52)

Derksen (2007: 375), EDG (132), Kroonen (2013: 34)

EM (44) suspect that the Latin and Germanic bow/arrow words are not of IE origin, especially given their semantic field. But they can be reconstructed to a common pre-form **h₂erk^w-/*h₂erk-u-*, where Germanic has either formed a possessive derivative ‘belonging to the bow’ > ‘arrow’ (DV 52) or forms an ablauting **uh₂-*stem (nom. **h₂érk-uh₂* > **arhū* > ON *qr*, gen. **h₂erk-uéh₂-s* > **arwōz* > ON *qrvar*, Kroonen 2013: 34).⁴⁴⁶ Connection with Greek and Balto-Slavic juniper/willow words (Gk. ἄρκευθος, Hsch. ἄργετος ἢ ἄρκευθος. Κρητες ‘juniper’, PSlav. **ork̑ta* ‘brittle willow’, Latv. *ērcis* ‘juniper’) that would point to a substrate origin (Derksen 2007: 375, DV 52, EDG 132) is semantically unnecessary (WH I: 64).

caelum ‘sky’

Pre-form: **kh₂ei-lo-* / **keh₂i-lo-* | PItal. **kailo-*

Comp.: *?*kh₂ei-lo-* / **keh₂i-lo-* | PCelt. **kaylo-* | OW *coil(i)ou* ‘omens, auguries’, OBret. *coel* ‘haruspicem’

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: geography; magico-religious?

Pokorny (916-17), WH (I: 130-1), EM (83-4), DV (80)

Solmsen (1894: 184), Gray (1902: 300-1), Schrader & Nehring (1917-23: 500), Schrijver (1991: 267-8), Untermann (2000: 363), Derksen (2007: 75), Matasović (2009: 197), Weiss (2016)

The comparanda of Lat. *caelum* ‘sky’ are not certain. It has been connected via **kaid/t-(s)lo-* to e.g. OE *hādor* ‘clear sky’ (Solmsen 1894: 184, Schrader & Nehring 1917-23: 500), Lith. *skāistas* ‘clear’, and Skt. *citrá-*, Av. *čithra-* ‘clear, conspicuous’, etc. (WH I: 130-1, Pokorny 916-17). The dental of the Germanic and Baltic forms is not the same, so these would have to be suffixes. Furthermore the Indo-Iranian vowel length rules out a laryngeal (Schrijver 1991: 267-8) making the *a*-vocalism of *caelum* difficult to account for this way.

A better alternative might be that proposed by Gray (1902: 300-1), based Osc. *kāila*,

⁴⁴⁶ Go. *arhwazna* ‘arrow’ has a different suffix (cf. Kroonen 2013: 34) and so is derivationally secondary.

which probably means ‘temple’ (cf. Untermann 2000: 363). Proposed independently by Schrijver (1991: 268), Lat. *caelum* can reconstruct to the same pre-form, **keh₂i-lo-* or **kh₂ei-lo-* as PCelt. **kaylo-* ‘omen’, with the semantic link perhaps lying in the field of augury. Further connection with PGM. **haila-* ‘whole’ and PBSI. **kailo-* ‘whole, healthy’ (DV 80, Matasović 2009: 197), perhaps in the sense that the sky was the ‘whole’ as opposed to *templum* ‘the part’, is unlikely. Semantically the Germanic and Balto-Slavic forms are much closer to each other (Derksen2007: 75, Kroonen 2013: 200) and the Baltic accentual paradigm probably rules out the laryngeal required for the Italo-Celtic *a*-vocalism. WH (I: 131) called the Italo-Celtic proposal phonologically flawless but semantically difficult. DV (80) briefly mentions the possibility that Italo-Celtic **kailo-* could be a non-IE loan. If the Italic and Celtic forms are in fact related, their common pre-form that is reconstructible to a valid IE root structure provides no positive evidence of a non-native origin.

A final possibility is that mentioned by EM (83-4), revived by Weiss (2016). From **kaid-(s)lo-* to the root in *caedō* ‘to cut’ would be derived *caelum* ‘sky’ and its synonym *caelum* ‘chisel’, the latter an instrument noun and the former a result noun. Weiss (2016) argues that, since the plural *caeli* is masculine (peculiar for neuter noun like *caelum*), it was originally a dual referring to the twain **keh₂id-(s)loih₁* ‘divided parts’, earth and sky.

caīnus ‘deep vessel, bowl, dish; cavity in rocks’

Pre-form: **kh₃-t* | PItal. **kaīno-*

Comp.: **kh₃-t* | PGk. **kotulo-* | Gk. κοτύλη, κότυλος ‘bowl, dish’

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: vessel

Pokorny (587), WH (I: 182), EM (105), DV (98)

Furnée (1972: 205 fn. 14), Giacomelli (1994: 40), EDG (763)

WH (I: 182) rightfully doubt the connections that Pokorny (587) suggests (e.g. CS *koťcb* ‘cell, nest’ and Go. *hēþjō* ‘chamber’) on semantic grounds. The only potential comparandum for Lat. *caīnus*, despite the reservations of EM (105),⁴⁴⁷ is Gk. κοτύλη (also κότυλος) ‘bowl, dish’. DV (98) and EDG (763) argue that the deviation in vocalism and the different suffix, along with the semantic category of vessel names, suggests that the two words might be independent loans from a third language.⁴⁴⁸ But the correspondence of the vowels can be explained via a common pre-form **kh₃-t*. While Gk. -ύλη is often found attached to Pre-Greek lexemes (Furnée 1972: 205 fn. 14, EDG

⁴⁴⁷ They also compare OE *heden* which, although it does seem to reconstruct to a similar **kHt-en-*, does not mean ‘cooking dish’ but rather ‘cloak, mantle’ (cf. Kroonen 2013: 214).

⁴⁴⁸ Giacomelli (1994: 40) proposes considering the vocalic alternation the result of lower register variation in a population with widespread Greek-Latin diglossia, but this idea has been criticized by e.g. Ruijgh (1986).

783), it does not necessarily grant Pre-Greek status to the root. Lat. *-inus* is productive. Thus there is no positive evidence of a substrate origin for *catinus*.

colus ‘distaff’

Pre-form: **ke/olh₃*- | PItal. **kolo*-

Comp.: **k_lh₃*- | PGk. **klō-* | Gk. κλώθω ‘to spin’

Irreg. correspondences Remarkable phonotactics

Semantics: textiles

Pokorny (639-40), WH (I: 250), EM (134-5), DV (127)

Osthoff (1892: 302), Frisk (1960-72 I: 879), Chantraine (1968-80: 545), Schrijver (1991: 469), EWAia (I: 316), EDG (720)

Lat. *colus* ‘distaff’ is usually taken from the root **k^wel-* ‘to turn, spin’ (Pokorny 639-40, Schrijver 1991: 469, EM 134-5, DV 127). WH (I: 250) crucially note that the distaff does not spin however. It is a staff on which the unspun fibers are kept, and drawn off while being twisted into thread by use of a drop spindle. (Note that the modern polysemy of *to spin* is secondary.) A better semantic match is Gk. κλώθω ‘to spin’ and derivatives (already Osthoff 1892: 302).⁴⁴⁹ While EDG (720) suggests the Greek verb is Pre-Greek, he does not provide any arguments. The pair can be reconstructed to a root **kelh₃*- (Lat. *colus* < **ke/olh₃-o-*, Gk. κλώθω < **k_lh₃-C-*), although it is unclear where else this root is attested.⁴⁵⁰

cornus ‘cornelian cherry tree’

Pre-form: **k_r-no-* | PItal. **korno-*

Comp.: **k_r-no-* | PGk. **krano-* | Gk. κρίνεια, κρίνον, etc. ‘cornelian cherry tree’

Irreg. correspondences Remarkable phonotactics

Semantics: plant, tree; fruit

Pokorny (572-3), WH (I: 276-7), EM (143-4), DV (137)

Boisacq (1911-12: 57-9), CAD (K: 122), Furnée (1972: 346), Demiraj (1997: 393), Orel (1998: 472), EDG (677, 770), Rosoł (2013:), Beekes (2014: 32), Blažek (2014: 44), van Beek (2022: 285-6), van Sluis (fthc.)

Lat. *cornus* and Gk. κρίνον can go back to the same pre-form **k_r-no-*, though van Beek (2022: 285-6) notes that the Attic-Ionic outcome of **k_r* should in fact be αρ; a borrowing

⁴⁴⁹ Often considered related to καλάθος ‘basket’ (Frisk 1960-72 I: 879, WH I: 250), not all agree (Chantraine 1968-80: 545, EDG 720). It is semantically distant enough to keep separate.

⁴⁵⁰ Osthoff (1892: 302) had considered Skt. *kṛṇāti* ‘to spin, draw fibers’, but EWAia (I: 316) disagrees; the **t* is part of the root.

from Epic Greek could account for the $\rho\alpha$, but is not very likely. From a similar formation can also be derived Lith. *Kìrnis* ‘the divine protector of the cherry’ (< * $k̑ȓn-io-$), but as onomastic evidence it is much less certain (DV 137, EDG 770). That the root is the same as in *cornū* ‘horn’ (cf. WH I: 276-7, EM 143-4) is unlikely, since the formation * $k̑ȓn-$ already meant ‘horn’ in PIE (cf. Skt. *śṛṅga-* ‘horn’, DV 136).⁴⁵¹ Boisacq (1911-12: 57-9) adduces *κέρασος* ‘sweet/bird-cherry’, suggesting that the intervocalic *s* is a borrowing from an Anatolian language or Thracio-Phrygian. EDG (677) agrees that it must be Anatolian or Pre-Greek (cf. Beekes 2014: 32 on the Pre-Greek nature of the suffix *-ασο-*), adding that the improved cherry seems to have originated in the area of the Pontos. We should remain cautious of assigning non-IE origin to a root based on the origin of a suffix; *κέρ-* could be the *e*-grade of the root behind *cornus* and *κράνον*. Alternatively, it represents a different lexeme entirely.

Furnée (1972: 346) compares Assyrian *karšu* ‘sweet cherry’, but the word does not exist (Rosol 2013: 179). Blažek’s (2014: 44) proposal of an intermediated loan from Akk. *kamīle/aššaru* ‘pear tree, pear’ (cf. CAD K: 121) requires formal changes *a* and semantic shift. In the end, the most secure comparanda, *cornus* and *κράνον*, might reconstruct to the same pre-form with little formal indication of a foreign origin.

corulus ‘hazel tree’

Pre-form: **kos-e/o/ul-o-* | PItal. **kose/o/ulo-*

Comp.: **kos-(V)l-o-* | PCelt. **koslo-* | OIr., OW *coll* ‘hazel’

**kos-l-o-* | PGM. **hasla-* | ON *hasl*, OHG *hasal*, etc. ‘hazel’

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, tree

Pokorny (616), WH (I: 280), EM (145), DV (138)

Schrijver (1995: 433), EIEC (260), Matasović (2009: 218), Kroonen (2013: 213), Smoczyński (2018: 504)

Lat. *corulus* ‘hazel’ reconstructs to **kosVlo-*, as **koslo-* would yield ***cōlus*. For Celtic, Schrijver (1995: 433) reconstructs **koslo-* to PIE **kos-lo-*, though Matasović (2009: 218) asserts that the Celtic forms could be derived by syncope from **kos-Vlo-*. It does not seem likely that the Germanic forms can be from anything other than **kos-lo-* (Kroonen 2013: 213). Even if this is so, the pattern that emerges is an inherited ablauting *l*-stem.

Lith. *kasūlas* ‘hunter’s spear’ has been compared (WH I: 280, Pokorny 616) with EIEC (260) noting the historic use of hazel for spears, spits, and poles. Because the Lithuanian form reconstructs to **kosulo-* with suffix vocalism that is aberrant from the perspective of IE ablaut, non-IE origin has been suspected (EM 145, DV 138). But Smoczyński

⁴⁵¹ Alb. *thánë* ‘cornelian cherry’ has been compared but it is difficult to make it work formally and several alternative etymologies exist (cf. Demiraj 1997: 393 with lit., Orel 1998: 472 with lit.).

(2018: 504) shows that the Lithuanian word is a deverbal derivative of *kàsti* ‘to hew wood with an axe’, with a suffix *-ul-* like that of *krātulas* ‘sieve’ < *kratýti* ‘to shake, make litter’. Thus it is unrelated, and all comparanda of *corulus* can be reconstructed to a root with vocalic alternation within the realm of IE ablaut.

crātis ‘construction of wickerwork, hurdle’

Pre-form: **kr(e)h(2)-ti-* | PItal. **krāti-*

Comp.: **kr(H)-ti-* | PGM. **hurdi-* | Go. *haurds* ‘(lattice) door’, ON *hurð* ‘door’, OHG *hurt*, *hurd* ‘hurdle, grate, railing’, etc.

?**korH-to-* | OPr. *corto* ‘heyn’

?**korh₁-et-* | PCelt. **koret-* | Mlr. *cora*, ‘palisade, wall’, MW *cored* ‘weir, dam’, etc.

Irreg. correspondences

Remarkable phonotactics

Semantics: tool

Pokorny (584-5), WH (I: 285-6), EM (147-8), DV (141)

Frankel (1962 I: 178), Puhvel (IV: 277-9), Schrijver (1991: 176, 191), EWAia (I: 290), Matasović (2009: 216, 225, 228), EDG (808), Kroonen (2013: 258)

WH (I: 285) place *crātis* ‘wickerwork, hurdle’ under *crassus* ‘thick, fat’, but DV (141) dismisses the connection on semantic grounds.⁴⁵² Much better is the connection with semantically close PGM. **hurdi-* ‘wickerwork door’. The Germanic forms do not require the presence of a laryngeal (cf. Schrijver 1991: 176), so they could derive from **krt-* to **kert-* ‘to turn, twist’ (cf. Skt. *ṛtāti* ‘to bind, attach’, *ṛṇātti* ‘to twist’).⁴⁵³ But the semantic relationship between PGM. **hurdi-* and Lat. *crātis* is good enough that it warrants disconnecting the Germanic root from **kert-* and instead reconstructing for it and Lat. *crātis* a pre-form **krH-ti-* (DV 141, Kroonen 2013: 258). That the vocalism can be explained with laryngeals makes a non-IE origin unlikely.

Amongst the numerous comparanda proposed (cf. WH I: 285-6 with lit.), OPr. *corto* ‘heyn’ could work formally (cf. also EM 148) < **korH-to-*. Matasović (2009: 216) compares PCelt. **koret-* ‘palisade, wall’, for which **korh₁-et-* seems possible. Neither need be related, given the semantic differences. In the end, Italic and Germanic have the

⁴⁵² There are formal difficulties as well. *Crātis* points to **krHt-i-* while *crassus* points to **krHt-to-*. The latter form might be expected to yield ***crāsus*, though Schrijver (1991: 191) has proposed **CRHTC* > *CRaTC*.

⁴⁵³ From this root has also been derived Skt. *kāṭa-* ‘woven mat’, though EWAia (I: 290) notes it requires Middle Indic developments to be from **kṛta-*. Puhvel (IV: 277-9) connects Hitt. *kurtal(l)(i)-* ‘crate, hamper, basket’, assuming it meant originally ‘wicker crate’. He also mentions Gk. *κυρτία* ‘wicker shield’. But as EDG (808) notes, there are semantic problems with the connections and there is no way to connect all of these forms in an inherited way, as Puhvel does in hopes of seeing an inherited PIE word for wickerwork.

closest semantic match. Despite some suspicions of non-IE origin (DV 141, Matasović 2009: 216), there are no indications of a substrate origin for Italo-Germanic **krH-ti-* beyond its limited distribution.

crēta ‘fine white clay’

Pre-form: **k^wreh₁-ie/ot-* | PItal. **krēt-* / **krējVt-*

Comp.: **k^wreh₁-ie/ot-* | PCelt. **k^wrīyet-* | OIr. *cré*, W *pridd*, etc. ‘mud, clay’

?**k^wreh₁-* | Toch. B **kw(ä)riye*, ?Toch. A *tukri* ‘clay’

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: geography; ceramics

WH (I: 290-1), EM (150), DV (144)

Pedersen (1909-13 I: 68), Pinault (2000: 105-8), Mallory and Adams (2006: 121), Matasović (2009: 182), Adams (2013: s.v. *kwraiññe*)

WH (I: 290-1) give two possibilities for the etymology of Lat. *crēta* ‘fine white clay’. Firstly, it could be from the PPP of *cernō* in the sense *terra crēta* ‘sifted earth’. DV (144) calls this semantically unconvincing, as clay is not sifted; but in fact it often is to ensure finer particle sizes. In any case, it is more attractive to follow Pedersen (1909-13 I: 68) in comparing several Celtic words for mud/clay. The Latin and Celtic forms can be reconstructed to the same pre-form **k^wreh₁-i-e/ot-*, albeit with unclear morphology (DV 144). This common pre-form, despite its limitation to Italo-Celtic and its technical semantics (DV 144, Matasović 2009: 182), does not provide any phonological indications of a non-IE origin.

Positive evidence for inheritance may come in the form of Toch. B **kw(ä)riye* (based on the adj. *kwraiññe* ‘pertaining to clay’). Mallory and Adams (2006: 121, also Adams 2013 s.v. *kwraiññe*) reconstruct **tk^wreh₁yot-* for the Latin and Celtic forms (assuming the ‘thorn cluster’ would resolve to *k* in a triconsonantal cluster) as well as Toch. B **kw(ä)riye* and Toch. A *tukri* ‘clay’. Pinault (2000: 106) shows that there is no trace of the dental in the Tocharian, and that its **i* element can be secondary. While Matasović (2009: 182) suggests that Italic, Celtic, and Tocharian might at least share the same root **k^wreh₁-*, Pinault (2000: 107) prefers comparing the Tocharian to a different group of words (OIr. *coire*, OHG **(h)wer*, Skt. *carú-* ‘cauldron’). The appurtenance of the Tocharian forms thus remains questionable.

dōlium ‘large earthenware vessel’

Pre-form: **deh₃l-* / **doHl-* / **dōl-* | PItal. **dōlijo-*

Comp.: **d_l-* / **dol-* | PItal. **dol-* | Lat. *dolāre* ‘to hew wood, shape/fashion’

**del-* | PCelt. **delwā-* | OIr. *delb*, OW *delu* etc. ‘form, appearance, image’

**dġ-* | PSlav. **dbly-* | MBulg., RuCS *дѣly* '(clay) cask', etc.

- Irreg. correspondences □ Remarkable phonotactics

Semantics: vessel

Pokorny (194-6), WH (I: 364), EM (181), DV (176)

Schrijver (191: 123), Derksen (2007: 134), Matasović (2009: 95), LIV (s.v. **delh₁-*)

The comparanda of Lat. *dōlium* are not entirely secure. WH (I: 364 with lit.) derive Lat. *dōlium* from a root **del-*, also behind Lat. *dolāre* 'to hew wood, shape/fashion', PCelt **delwā-* 'form; appearance, image', and PSlav. **dbly* 'cask'. EM (181) and Schrijver (1991: 123) are suspicious of the link with *dolāre* on semantic grounds seeing as it refers to woodwork, not ceramics. The semantic change is not so problematic if the meaning 'to shape/fashion' is original. But it may be unrelated, instead belonging to **delh₁-* (cf. LIV2 s.v. **delh₁-*). Matasović (2009: 95) connects the Celtic, Slavic, and Lat. *dolāre* without *dōlium*, and despite Derksen (2007: 134) considering the semantics of the Celtic material to be too far from the Slavic, EM (181) note that several Slavic forms (RuCS *delva*, *delъ* 'cask', Bulg. *délva* 'big jug with handles') have a **w* element reminiscent of that in PCelt. **delwā-*.

The semantic field of ceramics has led to the suspicion of a non-IE loanword (EM 181, Schrijver 1991: 123, DV 176). But regardless of which comparanda belong to *dōlium*, none of the forms requires vocalism outside of unusual but not unattested **ō* ~ **∅* ablaut.

fēlēs 'small carnivore, perhaps marten'

Pre-form: **b^heH- / - / *b^hē / -* | PItal. **fē / -*

Comp.: **b^hH- / - / *b^he / -* | PCelt. **bal-* | W *bele* 'wolf; marten, weasel; predatory beast'

- Irreg. correspondences □ Remarkable phonotactics

Semantics: animal, wild

Pokorny (118-20), WH (I: 474), EM (223-4), DV (209)

Johansson (1890a: 351), Pedersen (1909-13 I: 98), Kluge & Seebold (1989: 84), Schrijver (1991: 375), Schrijver (1995: 123), Matasović (2009: 187), GPC (s.v. *belau*, *bele*, *bela*, *bala*, *balh*)

Earlier editions of Kluge's etymological dictionary compared OHG *bilih* 'doormouse' and W *bele* 'wolf; marten' (cf. Johansson 1890a: 351), though the Germanic word was later asserted to be a borrowing from Slavic (see now Kluge & Seebold 1989: 84).⁴⁵⁴ The semantics were never a good match to begin with, and Johansson (1890a: 351) instead

⁴⁵⁴ Originally, PSlav. **pъlxъ-* 'mouse' was thought to have been borrowed from Germanic. But because they seem to relate to Baltic mouse words like Lith. *pelė* from **pelH-* 'gray' (cf. Smoczyński 2018: 937), the relationship is now understood to have gone in the other direction.

connected Lat. *fēlēs* to *W bele*, a comparison still generally supported (WH I: 474, DV 209). The old connection with OHG *bilih* however shaped Pedersen's (1909-13 I: 98) reconstruction of *W bele* as **b^(h)eleg^(h)-*, which remains in circulation (WH I: 474, Schrijver 1991: 375, DV 209). But the GPC (s.v. *belau*, *bele*, *bela*, *bala*, *bal*) suggests its original inflectional pattern was sg. *belau*, pl. *balawon* (Paulus van Sluis, p.c.), making it similar to *cenau*, *canawon* 'whelp' and thus continuing an original *u*-stem of PCelt. **bal-* (cf. Schrijver 1995: 123). Assuming that the *a*-vocalism in Celtic is original, the Latin and Celtic pre-forms could precariously be linked via a root with a laryngeal; though a root shape **b^heHl-* is suspicious. Thus the *l* might be a suffix, though this makes the further derivation within Celtic difficult. Otherwise, PCelt. **balawon-* can be the result of Joseph's Rule < **bel-* (cf. Matasović 2009: 187) such that both Latin and Celtic attest to a root **b^hel-*, with Latin preserving a lengthened grade (cf. Johansson 1890a: 351).

The traditional link between reconstructed **b^hel-* and the root **b^hel-* 'to shine' is semantically tenuous and was already doubted by Pokorny (118-20). However, his own suggestion that *fēlēs* and *mēles* 'badger' were related and borrowed from an Alpine substrate language (followed by WH I: 474,⁴⁵⁵ EM 224) is not convincing either. Despite the occurrence of such **b^h ~ *m* alternations in other loans, there is no semantic reason to assume one here.

follis 'bag, sack; ball, testicles'

Pre-form: **b^(h)(o)l-n-* | PItal. **folli-*

Comp.: **b^hl-n-* | PGk. **p^hallo-* | Gk. φαλλός 'penis'
 ??**b^hl-n-* | PGk. **balla-* | Gk. βαλλάντιον 'purse', var. βαλάντιον
**b^hol-n-* | PGm. **ballan-* | ON *bǫllr*, OHG *ballo*, *bal*, etc. 'ball'
**b^(h)l-n-* | PCelt. **ballo-* | OIr. *ball* 'penis', *W balleg* 'sack'

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: body part

Pokorny (120-2), WH (I: 524), EM (244), DV (230)

Matasović (2009: 53), EDG (196, 197, 1550), Kroonen (2013: 50), van Beek (fthc.)

While the semantic match between 'sack' and 'penis' is not perfect, all comparanda can remarkably be reconstructed to an *n*-stem of a root **b^hel-*, otherwise unattested⁴⁵⁶ but perhaps with the meaning 'to swell' (WH I: 524 with lit., EM 244 [who consider the geminate expressive], Matasović 2009: 53, Kroonen 2013: 50). The only aberrant forms

⁴⁵⁵ WH (I: 474) suggest that the source language had a nasalized labial spirant, showing that they think this was a non-IE language (though they alternatively point to a discussion on a Ligurian *m ~ v* alternation in Kretschmer [1905: 114]).

⁴⁵⁶ WH (I: 524 with lit.) and EM (244) suggest **b^helǵ^h-* (LIV2 s.v.) is an extension with **ǵ^h*.

are Greek. EDG (196) compares *follis* to Gk. βαλλάντιον ‘purse’ with the variant βαλάντιον suggesting a Pre-Greek origin. The semantic match is admittedly better. But since Gk. φάλλος also attests to variants with single λ and initial β (EDG 197, 1550), in light of the agreement of all other forms, we could be dealing with taboo deformation or reflexes of another language in Greek (cf. EDG 197 on the latter).

An even more promising comparison is by van Beek (fthc.),⁴⁵⁷ who compares Lat. *follis* to several Germanic forms < **bʰolǵʰ-i-* (Go. *balgs* ‘skin bag’, ON *belgr* ‘skin; bellows’, etc.) and Celtic forms < **bʰolǵʰ-o-* (Olr. *bolg*, *o*-stem, ‘bag; belly; bellows’, etc.) and **bʰolǵʰ-éh₂-* (Olr. *bolg*, *ā*-stem, ‘blister; ball; pouch’). It relies on the argument that **h* > *ll* in Latin, such that Lat. *follis* would reconstruct to **bʰolǵʰ-i-* like the Germanic forms.

In any case, Lat. *follis* can be furnished with a compelling IE etymology.

frāga ‘strawberries’

Pre-form: **dʰrHǵ-o-* | Pltal. **prāgo-*

Comp.: **dʰrHǵ-o-* | PAbI. **drað-* | Alb. *dredhë* ‘strawberry’

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: plant, fruit

WH (I: 540), EM (251), DV (239)

Schrijver (1991: 177), Demiraj (1997: 144)

EM (251) link Lat. *frāga* through **srāg-* to Gk. ῥάξ/ῥώξ ‘grape’ as a word from a Mediterranean language, followed in large part by Schrijver (1991: 177). Instead, *frāga* goes back to the same pre-form **dʰrHǵ-o-* as Alb. *dredhë* ‘strawberry’ (Demiraj 1997: 144 with lit.).⁴⁵⁸ There is nothing non-IE about them except for their restricted distribution.

frutex ‘shrub, bush; shoot’

Pre-form: **bʰru-t-* | Pltal. **flpʰχʷutek-*

Comp.: ?**bʰru-t-* | PCelt. **bruto-* | Mlr. *broth* ‘awn, ear’

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: plant, wild or domesticated

Pokorny (169), WH (I: 554), EM (257), DV (245)

LEIA (B-98), Kroonen (2013: 76)

WH (I: 554 with lit.) support deriving Lat. *frutex* via **bʰru-to-* from a root **bʰreu-*, to

⁴⁵⁷ “Latin *follis*, *vellō* and *ille* as evidence for a sound change **ly*, **lh* > *ll*” to appear in *Glotta*.

⁴⁵⁸ East Ghag has *drathe*, which leads Demiraj to suggest that **drað-* has undergone umlaut from the plural **dradi*.

which other branches would attest a suffix **-d-*: PGm. **breutan-* ‘to break open, bud’ and Mlr. *broth* ‘awn, ear’). But Kroonen (2013: 76) suggests **breutan-* might be backformed to an iterative **brut(h)ōn-* ‘to bud’ < **b^hṛd-néh₂-* (cf. Lat. *frōns* s.v.). Thus the **d* is part of the root.⁴⁵⁹ Mlr. *broth* can reconstruct to **b^hru-to-* (LEIA B-98, DV 245), thus Lat. *frutex* could go back to an Italo-Celtic **b^hru-to-* to which Latin later added the suffix *-ex*. (This prevents having to reconstruct *frutex* to an invalid **D^heT* root structure.) While DV (245) finds PIE origin of *frutex* uncertain, there does not seem to be morphophonological evidence to reject it.

iuncus ‘reed, rush’

Pre-form: **(H)jojn-i-ko-* | PItal. **yoiniko-*

Comp.: **(H)jojn-i-* | PCelt. **yoini-* | OIr. *áin* ‘reed, rushes’

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, wild

Pokorny (513), WH (I: 729-31), EM (328), DV (313)

Brüch (1922: 224-232), Wagner (1960-4 I: 489), Matasović (2009: 437), Kroonen (2013: 12)

Lat. *iuncus* ‘reed, rush’ is close in form and meaning to OIr. *áin* ‘reed, rush’. Both can be reconstructed to **H₁jojn-*, with the expected ***ū* in Latin (resulting from the monophthongization of **oi* in an initial syllable) being shortened by Osthoff’s law. Latin has added a **-ko-* suffix (Pokorny 513, WH I: 729-30, DV 313, Matasović 2009: 437).

Lat. *iuncus* is often presumed to be related to the etymologically obscure *iūniperus* ‘juniper’ (WH 730-1), leading to comparisons with the Germanic juniper word **ainja-*. There is chance that the Germanic material reconstructs to a pre-form like **H₁jojn-i-*. Semantically however, the comparison between the reed and juniper words is problematic, relying on the idea that both reeds and juniper branches are used for weaving (cf. WH I: 731). Thus, it seems best to keep the Latin words separate from the comparanda beyond Celtic. Despite the restricted distribution (DV 313, Matasović 2009: 437), there are no formal indications of a non-IE origin.

lībra ‘scale, pound’

Pre-form: **liH-d^hro-* | PItal. **līprā*

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: economic; metallurgy

WH (I: 795-6), EM (356), DV (339)

⁴⁵⁹ LIV2 however hesitantly reconstructs the formation at **b^hreuH-de-*. A connection between this root and *frutex* is still ruled out unless the latter has been shortened from **frūto-*, which is unlikely (DV 245).

Schulze (1895: 223), Walde (1910: 428), Niedermann (1918: 32), Güntert (1933: 20), Ribezzo (1934a: 91), Furnée (1972: 182), Szemerényi (1991 II: 655-672), Lejeune (1993: 2), Willi (2008: 22), EDG (867), Weiss (2021)

Lat. *lībra* is widely compared to Gk. λίτρα ‘Sicilian coin’. They are often proposed to reflect a pre-form **lībra-* of Mediterranean substrate origin, either as independent loans (Niedermann 1918: 32 fn. from previous page, Güntert 1933: 20)⁴⁶⁰ or with the Greek having been borrowed from an Italic pre-form (Schulze 1895: 223, Walde 1910: 428, DV 339). In a similar vein, the pair might be the result of a Sichel-Ausonian substrate (Ribezzo 1934a: 91, Szemerényi 1991 II: 655-672).

The solution is probably quite simple. The shape of the Latin word strongly suggests a PIE instrument noun suffix **-dʰro-* (cf. DV 339, Weiss 2021). Given that the Greek word refers to a Sicilian coin, it is probably a loan from the Sichel cognate of Lat. *lībra* < **liH-dʰreh₂* (Weiss 2021, cf. Lejeune 1993: 2), since **dʰ* > *d* with devoicing before sonorant consonants is probably regular in Sichel (Willi 2008: 22). The PIE root from which this Italic formation derives is not clear (cf. WH I: 796 with lit.; most recently Weiss 2021 argues for a root **leiH-* ‘to pour’), but nothing so far requires the rejection of inherited origin.

mālus ‘pole, mast’

Pre-form: **mh₂sd-o-?* / **mh₂sd-lo-?* | PItal. **mas(d)lo-*

Comp.: **mh₂sd-jo-* | PCelt. **mazdyo-* | Mlr. *maide* ‘post, stick, bundle, wood’
*?*mh₂sd-lo-* | PCelt. **mazdlo-* | MW *meithlyon* ‘masts?’

**mh₂sd-o-* | PGm. **masta-* | OE *mæst* ‘mast’, OHG *mast* ‘stick, pole, mast’

■ Irreg. correspondences

□ Remarkable phonotactics

Semantics: tool; maritime

Pokorny (701-2), WH (II: 19), EM (381), DV (361)

Bottiglioni (1943: 318), Adams (1985), Schrijver (1991: 167), Matasović (2009: 260), Kroonen (2013: 357), Prospér (2019), Koch (2020: 88)

DV (361) and Matasović (2009: 260) suggest on the basis of geographic restriction and technical semantics that Lat. *mālus* ‘pole, mast’ and its comparanda may be non-IE loans, but otherwise there are several paths to reconstructing a common pre-form.

Lat. *mālus* is easily comparable to Germanic and Celtic⁴⁶¹ forms but for the *d* ~ *l* alternation this produces. We do not fully understand the “Sabine *l*” phenomenon by

⁴⁶⁰ EDG (867) follows Furnée (1972: 182) in further adducing Hsch. λιδρίων· τρύβλιον ‘cup’. But the semantic difference makes this unnecessary.

⁴⁶¹ Kroonen (2013: 357) opposes adducing Lat. *mālus* to what he considers a Celto-Germanicism.

which inherited **d* becomes Lat. *t*⁴⁶² so it is potentially risky to apply it as an explanation. Alternatively, Koch (2020: 88) reconstructs **mazdlo-* on the evidence that MW *meithlyon* ‘masts?’ similarly preserves a **lo-* derivation. The meaning of the Welsh word is not certain, nor do we have comparanda for the outcome of **-zdl-* in Italic, but it does not seem impossible. Otherwise, contamination with *pālus* ‘pole’ has been suggested (Bottiglioni 1943: 318, EM 381).

Schrijver (1991: 167) favors a connection with *mās*, *maris* ‘man’ from an inherited root **meh₂-(o)s*, **mh₂-(e)s-*. As Adams (1985) argues, *mas-* in Latin *masturbari* might mean ‘penis’. Thus *mās* ‘man’ would be a metaphorical extension of the meaning ‘penis’, as would *mālus*. With or without the root etymology, a common pre-form is probably reconstructible for Italic, Celtic, and Germanic.

miliu ‘millet’

Pre-form: **mel(H)-* | PItal. **me/ilio-*

Comp.: **mel(H)-* | PGk. **melinā-* | Gk. *μελίνη* ‘millet, esp. foxtail millet’

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, domestic

Pokorny (716-19), WH (II: 87-8), EM (403), DV (379)

Nieminen (1956: 167-8), Furnée (1972: 246), Leumann (1977: 101), Witczak (2003: 78), EDG (926), Meiser (2010: 81), Kroonen et al. (2022: 24)

Lat. *miliu* can derive from earlier **melium* via *i*-mutation (Leumann 1977: 101, Meiser 2010: 81), so it matches the root of Gk. *μελίνη* ‘millet’. Lith. *mālnos* ‘floating sweetgrass’ has been compared (Pokorny 716-19, WH II: 87-8, EM 403, DV 379), but is rather a loan from Polish *manna* ‘floating sweetgrass’ with a dissimilation of the geminate (Nieminen 1956: 167-8). Kroonen et al. (2022: 24) point out that the Latin and Greek, despite the same meaning, do not reconstruct to the same inherited formation, and are thus at best independent formations to the same root. That root might be **melh₂-* ‘to grind’, though millet certainly does not seem to be the grinding grain *par excellence*. Alternatively, it could be related to Gk. *μέλας* ‘black’ (cf. Skt. *śyāmāka-* ‘type of millet’ to *śyāmā-* ‘black’, Witczak 2003: 78).

Furnée (1972: 246) suggested that *ἔλωμος* ‘millet’ and Hsch. *ἐλίμαρ· κέγγρω ὄμοιον [ἐλίνη] ἢ μελίνη ὑπὸ Λακόνων* ‘proso millet or foxtail millet among the Laconians’ continue **ʔel-* and thus attest to an *m ~ w* alternation, but EDG (926) rightly calls this too far-fetched. In the end, even if a fitting IE root cannot be identified as the source, there are no irregular alternations between the Latin and Greek forms that point to a non-IE root.

⁴⁶² Prospér (2019) proposes that it at least occurred with initial **da-*, and **masdo-* does not fit this phonetic environment.

olor 'swan'Pre-form: **h₁el-* | PItal. **elōr-*Comp.: **h₁el-* | PCelt. **eIV-* | OIr. *elu*, MW *alarch*, OCo. *elerhc*, etc. 'swan'

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: animal, wild; aquatic

Pokorny (302-4), WH (II: 207), EM (461), DV (427)

Schrijver (1991: 37), Derksen (2000: 84), Derksen (2007: 365), Matasović (2009: 114), EDG (404), Kroonen (2013: 20), Jakob (fthc.)

Latin and Celtic words for 'swan' can reconstruct to derivations from the same *e*-grade⁴⁶³ root **h₁el-* (Schrijver 1991: 37, DV 427, Matasović 2009: 114), restricted to Italo-Celtic.⁴⁶⁴ Nothing else about the pair prohibits inherited origin.

Semantically attractive are comparisons to Slavic and Germanic words for 'swan', but they are actually formally difficult. Traditionally reconstructed as **h₂elb^h-* 'white' with a suffix, this is precluded by the Slavic accentuation (Derksen 2007: 365, DV 32, Kroonen 2013: 20). In fact, the Slavic forms more accurately reconstruct to **lebedb* and **albodb* (Jakob fthc. with lit.) and Germanic to **albut-* (Kroonen 2013: 20), similar to the second of the two Slavic forms but without the nasal element. Derksen (2000: 84) proposes it is an example of *a*-prefixation. In the end, this leaves little similarity between them and the Italo-Celtic swan words; thus they are best kept separate (cf. Kroonen 2013: 20).

ornus 'ash tree'Pre-form: **Hh₃-es-* | PItal. **osVno-*Comp.: **Hh₃-es-* | PCelt. **osno-* | OIr. *uinnius*, MW *onn*, etc. 'ash-tree'**Hh₃-es-k-* | Arm. *hac^ci* 'ash-tree'**Hh₃-es-ko-* | PGM. **aska-* | ON *askr*, OE *æsc*, etc. 'ash-tree'**Heh₃-s-* | PBalt. **oʔs-io-* | OPr. *woasis*, Lith. *úosis*, etc. 'ash-tree'**Heh₃-s-* | PSlav. **oʔs-en-* | Ru. *jásen'*, Cz. *jasan*, etc. 'ash-tree'?**Hh₃-es-k-* | PAIb. **osk-* | Alb. *ah* 'beech'?**Hh₃-es-k-* | PGK. **oks-* | GK. ὀξύα 'beech'

□ Irreg. correspondences □ Remarkable phonotactics

⁴⁶³ Though Lat. *olor* can theoretically represent an *o*-grade (DV 427).

⁴⁶⁴ The appurtenance of Gk. ἐλέα 'singing bird, perhaps reed warbler' is doubtful, in part because of its semantic remoteness and in part because of variants like ἐλεία and ἔλατος that make a reconstruction difficult within Greek (EDG 404).

Semantics: plant, tree

Pokorny (782), WH (II: 223), EM (469), DV (435)

Schrijver (1991: 77-8), Derksen (2007: 29), Martirosyan (2009: 399, 641), Matasović (2009: 300), EDG (1088), Kroonen (2013: 38)

There are indications that *ornus* and its comparanda are not inherited, for example the *n*-suffix of the Celtic forms (see §3.3.4). But beyond the suffixes, though not all agree (cf. Matasović 2009: 300, EDG 1088), the reconstructible alternation **Heh₃-s-*, **Hh₃-es-* looks remarkably like an inherited *s*-stem (Schrijver 1991: 77-8, Derksen 2007: 29, DV 435, Martirosyan 2009: 399, Kroonen 2013: 38). If it is of non-IE origin, there are no irregular alternations that show it.

For Gk. ὄξυα, we might have to assume metathesis, which in *ascia* and *viscum* was a sign of non-IE origin. However there are potential explanations for this. Despite EDG's (1088) disagreement, that ὄξυα can also mean 'spear' makes it possible that this lexeme was contaminated by ὄξύς 'sharp'. On the other hand, its different meaning might show that it is unrelated (cf. Martirosyan 2009: 641).

salix 'willow'

Pre-form: **sIH-ik-* / **sh₂el-ik-* | PItal. **salik-*

Comp.: **sIH-ik-* / **sh₂el-ik-* | PCelt. **salik-* | OIr. *sail*, MW *helyg*, etc. 'willow'

**solH-ik-* / **sh₂el-ik-* | PGM. **salihōn-* | ON *selja*, OHG *salaha*, etc.
'willow'

**selH-ik-* | PGk. **helikā-* | Myc. *e-ri-ka*,⁴⁶⁵ Arcad. *ἑλικης* 'willow'

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: plant, tree

Pokorny (782), WH (II: 223), EM (469), DV (435)

Frisk (1960-72 I: 494), Chantraine (1968-80: 338), Friedrich (1970: 53-7), Schrijver (1991: 77-8), Derksen (2007: 29), Martirosyan (2009: 399, 641), Matasović (2009: 300), EDG (1088), Kroonen (2013: 38)

Methodologically, there is little to reject an Indo-European origin for this word beyond its distribution and its arboreal semantics. Italic and Celtic reconstruct to the same proto-form **salik-*. If from a zero-grade formation **sIH-ik-*, PGM. **salihōn-* could be from a full *o*-grade **solH-ik-* (cf. EM 591). Though it is often suggested that the Germanic forms do not all attest to an *i* vowel in the suffix (WH II: 469, Schrijver 1991: 103, EM 591), with DV (536) pointing to the **-ik* ~ **-k* alternation as a non-IE feature,

⁴⁶⁵ We should perhaps be cautious of this form however. Myc. *e-ri-ka* is a descriptor of wheels (Chadwick & Baumbach 1963: 190), so the assumption would be that they are wheels of willow wood.

Kroonen (2013: 424) reconstructs **-ik-* for all Germanic forms.⁴⁶⁶

While Kroonen (2013: 424) suggests that the Mycenaean form < PGk. **helik-* attests to the irregular vocalic alternation **selik-* ~ **salik-*, the Greek could simply be from the full *e*-grade **selH-ik-* of the root in question (Matasović 2009: 319). The Greek evidence is complicated due to the appearance of Boeot. *ῥελικόν* (in Korinna), a hill otherwise called *Ἐλικόν* and understood as ‘willow-mountain’ akin to the Viminal Hill in Rome. DV (536) notes that it is a toponym, and thus does not certainly contain the same word. If they do represent the same word, then EDG (410) removes them from comparison.⁴⁶⁷

Despite some suspicion of non-IE origin (cf. DV 536, Matasović 2009: 319, Kroonen 2013: 424), an inherited origin cannot be rejected for this lexeme. This has important implications for the analysis of the *-ik* suffix (see §3.3.3).

scutra ‘shallow dish, pan’

Pre-form: **sku-treh₂-* | **skutrā*

Comp.: **skeu-* | **skeuso-?* | Gk. *σκεῦος* ‘vessel, implement’

Irreg. correspondences Remarkable phonotactics

Semantics: vessel

WH (II: 503), EM (606), DV (548)

Frisk (1960-72 II: 727), Matasović (2009: 342), EDG (1348)

WH (II: 503) compare Lat. *scutra* ‘shallow dish, pan’ and its diminutive *scutella* ‘small shallow dish’ (EM 606) to *scūtum* ‘shield’ based on the idea that they could both have been made of leather. But the length of the vowel is problematic, especially if *scūtum* is from **skoī-to-* (cf. OIr. *sciath*, OCS *štīť* ‘shield’, Lith. *skiētas* ‘reed’, etc.: WH II: 503, EM 607, Matasović 2009: 342, DV 548). DV (548) is further suspicious of the suffix *-ra* to derive *scutra* from *scūtum* and suggests a loanword. But it is formally and semantically attractive to compare Gk. *σκεῦος* ‘vessel, implement’, often held to be without good cognates (EDG 1348). The preservation of its diphthong is strange but might suggest original **σκεῦσος* (Frisk 1960-72 II: 727). Lat. *scutra* could be an instrument noun **sku-treh₂* from the same root **skeu-*.

viola ‘violet, stock (*Matthiola spp.*)’

Pre-form: **u(H)i(i)H-ol-el-* | Pltal. **wiolā-*

⁴⁶⁶ Matasović (2009: 319) suggests that the Germanic word could be a prehistoric borrowing from Celtic, but there does not seem to be any compelling reason to assume this.

⁴⁶⁷ Boeot. *ῥελικόν* is suspiciously similar to **wel-ik-* (cf. Frisk 1960-72 I: 494), a willow word otherwise restricted to West Germanic (e.g. OE *welig*, Engl. *willow*, OS *wilgia*, etc.). Some (cf. Chantraine 1968-80: 338, Friedrich 1970: 53-7) have tried to connect both to the *salix* family by proposing an inherited **swel-/*sel-*, but this is untenable (cf. EDG 410). It does not solve the problem of the appurtenance of Germanic, as it would yield PGm. ***swel-ig-*.

Comp.: **uī-* | PGk. **wio-* | Gk. ἴον ‘violet’, Hsch. γία: ἄνθη (= **φία*) ‘flowers’

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: plant, wild; flower

WH (II: 795), EM (738)

Vaniček (1881: 256), Meillet (1908: 162), Cuny (1910: 157), Walde (1910: 840), Schrijver (1991: 245), DV (677, 680), EDG (594, 605), Weiss (2020: 300)

Older etymological explanations relied on e.g. Pliny’s description (*Nat.Hist.* 21.14) that violets were the premier flowers used in wreaths to derive their name as a diminutive of *viēō* ‘to plait, weave’ (Vaniček 1881: 256) < PIE **ueih₁-*, **ueh₁-i-* (LIV2 s.v. **uīeh₁-*, DV 677).⁴⁶⁸ But this root, even in the zero-grade, does not seem to be able to produce the Gk. ἴον (**uih₁-o* > ***īon*, **uh₁-i-o* > ***eion*). Nor would such a derived meaning from such an undervived formation be likely. It seems that both Latin and Greek simply reconstruct to a root **uī-*, but since the root is otherwise unknown, the pair is widely considered to represent independent borrowings from a Mediterranean language (Meillet 1908: 162, WH II: 795, Biville I: 246, EM 738, EDG 594). While Walde (1910: 840) took the Latin as a diminutivized borrowing from Greek, Cuny (1910: 157) saw in *viola* the same suffix as in *īnsula* against Gk. (Doric) *vāsoς*.

While a non-IE origin seems quite likely, there are no formal indications that the root **uī-* is non-IE. Nor does Latin need to have borrowed from Greek; if *viola* is a diminutive formation, it could have been produced within Latin.

2.4.2 Best Explained as Inherited

(Comparanda are listed in the header only when they too have been proposed to be of non-IE origin).

agna ‘ear of grain’

Pre-form: **h₂ek₁-(o)n-* | PItal. **ak(o)nā*

Comp.: **h₂ek₁-on-* | PGm. **ahanō-* ~ **aganō-* | Go. *ahana*, ON *ogn* ‘chaff’, etc.

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: plant, domestic

Pokorny (18-22), WH (I: 22-3), EM (15), DV (29)

Thurneysen (1882), Furnée (1972: 362), Kroonen (2013: 5), EDG (184), Kroonen et al. (2022: 23)

That Lat. *agna* ‘ear of grain’ did not undergo nasal metathesis (cf. Thurneysen 1882)⁴⁶⁹

⁴⁶⁸ Schrijver (1991: 245) shows that Russ. *vilá* with final accentuation has not undergone Hirt’s Law and thus suggests that the vowel preceded the laryngeal in a formation **uīHi-leh₂-*.

⁴⁶⁹ It only seems to occur sporadically with velars however (cf. de Vaan 1999: 22). Lat. *pangō* ‘to fix’ in

suggests that it is from PItal. **akVnā-* (cf. Kroonen 2013: 5) from the same PIE formation (**h₂ek-on-eh₂* < **h₂ek-* ‘sharp’) as the PGM. Verner variants **ahanō-* / **aganō-* ‘chaff’ (cf. most recently Kroonen et al. 2022: 23).⁴⁷⁰ Gk. ἄχνη ‘foam, froth; chaff’ has similar semantics to the Germanic forms and could be from **h₂ek-s-neh₂-* (WH I: 22-3, DV 29). But ἄχνη is certainly related to Gk. ἄχυρον ‘chaff’ (Furnée 1972: 362, EDG 184), which reflects **g^h*. Despite the semantic match between Greek and Germanic, the formal match between Italic and Germanic is so close that it cannot be ruled out that the Greek forms are unrelated and that *agna* represents an Italo-Germanic retention of an inherited or dialectal PIE formation.

anguilla ‘eel’

Pre-form: **h₂eng^{wh}-īn-leh₂* | PItal. **anguīnlā*

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: animal, wild; aquatic

Pokorny (43-5), WH (I: 48), EM (33), DV (42)

Hirt (1907/8: 65-8), REW (no. 461), Strodach (1933: 38), Corominas and Pascual (1984-91 I: 271-3), Katz (1998: 321-9), Driessen (2005: 42-3), Derksen (2007: 386), EDG (372)

EDG (372) notes that Lat. *anguilla*, Gk. ἔγχελυς, and Lith. *ungurys* ‘eel’ do not reconstruct to a common PIE pre-form, pointing to non-IE origin. But since Lith. *ungurys* can be a reflex of inherited **h₂eng^{wh}-* ‘snake’ with East Lithuanian **an-* > *en-* (Derksen 2007: 286), it is attractive to derive *anguilla* from *anguis* ‘snake’ as well.

Its geminate *ll* means that it is not simply a diminutive. Katz (1998: 321-9) thus follows Hirt (1907/8: 65-8) in analyzing *anguilla* as a compound of *anguis* + the hapax *illa* ‘worm’ < **ēlu-ā* (purportedly in PGM. **ēla-* ‘eel’ as if from a root **(H)elo-*), mirrored in Gk. ἔγχ-ελυς.⁴⁷¹ Driessen (2005: 42-3) supports the analysis of a compound, but instead takes the Plautine variant *anguīla* as primary,⁴⁷² interpreting *illa* as a diminutive of **īlā* < **h₁i-h₁l-eh₁* (with PGM. **ēla-* < **h₁e-h₁l-o-*). Neither of these proposals is certain. Instead, *anguilla* could reflect **anguīn-lā* to *anguīnus* ‘pertaining to a snake’ (Strodach 1933: 38, DV 42). No element of *anguilla* need be of non-IE origin. Gk. ἔγχελυς would

light of the stem type of Gk. πηγνυμι ‘to fasten’ could reflect metathesis from original **g-n-*. But it has not occurred in e.g. Lat. *dignus* ‘worthy’ < **dek-no-*.

⁴⁷⁰ OPr. *ackons* ‘awn’ is startlingly similar (cf. Kroonen 2013: 5, Smoczyński 2018: 15). There are further Baltic relatives with different suffixes (cf. Lith. *akiotas* ‘awn, fishbone, bristle, etc.’). It is difficult but potentially not impossible for these to derive from **h₂ek-*.

⁴⁷¹ His further connection of (in reversed order) Hitt. *Illuyankaš* ‘mythical snake’ is too far-fetched.

⁴⁷² Sp. *anguila* and Opt. *anguia* ‘eel’ seem to descend from *anguīla* (REW no. 461), but could be borrowed from Catalan where **-īll-* > *-il-* (Sp. *anguila* replaced *anguilla* in the 17th c. but the Portuguese loan would have to be earlier, Corominas & Pascual 1984-91 I: 271-3). Katz (1988: 321-9) takes the Plautine *anguīla* as secondary, from the sporadic avoidance of an extra-heavy syllable.

thus be unrelated.

cōnīveō, -ēre ‘to be tightly closed, to close (the eye)’

Pre-form: **kom-sne/oiḡ^{wh}*- | PItal. **komsne/oiχ^{wē}*-

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: action

Pokorny (608), WH (I: 261), EM (137-8), DV (130)

Corssen (1863: 21), Sommer & Pfister (1977: 194, 196), Leumann (1977: 218), Meiser (1988: 70-1), Untermann (2000: 417), Kroonen (2013: 236), Weiss (2020: 130 fn 16), Kroonen, Wigman & Thorsø (2021)

The traditional explanation of Lat. *cōnīveō* takes it from PIt. **kneig^{wh}-ē-* ‘to blink, to draw together’ along with U **kunikaz**, *conegos* [nom.sg.masc. PPP] ‘kneeling?’ and Germanic **hnīwan-* ~ **hnīgan-* ‘to bow (down)’ (Corssen 1863: 21, Sommer & Pfister 1977: 194, 196, WH I: 261, Pokorny 608, Leumann 1977: 218, EM 137-8, DV 130, Kroonen 2013: 236). It is also however widely acknowledged that this would require the reconstruction of an invalid **TeD^h* root structure, which DV (130) takes to suggest a loanword.

The Latin form can instead be derived from **sneig^{wh}-* (Kroonen, Wigman & Thorsø 2021). If the original meaning of **sneig^{wh}-* was ‘to sink/fall down’, then PIt. **kom-* in the sense ‘together’ added to a causative **snoig^{wh}-eje-* ‘to make fall’ closely matches the meaning ‘to close the eye’ attested for *cōnīveō*. The outcome of **oi* in medial syllables is not fully resolved (cf. Weiss 2020: 130 fn 16), but *pōmērium* < **post-moir-io-* (since *mūrus* is from **moi-ro-*) seems to show that non-initial **oi* > *ī* (lowered to *ē* before *r*) (Meiser 1988: 70-1). This proposal eliminates the potential Sabellic match. Untermann (2000: 417) shows that the exact meaning of the Umbrian words is unknown, the context being a ritual behavior performed by the sacrificing priest upon bringing the sacrificial cakes in the **skalçe**-vessel. Given that we do not know its meaning, it is a small cost to pay.

culleus ‘leather sack’

Pre-form: **kol-u-ejo-* | PItal. **kolweyo-*

Comp.: **kol-ey-* | PGk. **kolewo-* | Gk. *κολέον* ‘sheath of a sword’

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: textiles (leather)

Pokorny (553-4), WH (I: 303), EM (155), DV (250)

Meyer (1887: 163), Muller (1926: 108), Ernout (1946: 44), EDG (735)

The link between Lat. *culleus* and Gr. *κολέον/ς* ‘sheath of a sword’ is widely agreed

upon, as is the idea that they are independent loans from a substrate language (non-IE according to DV: 150; Etruscan according to Ernout 1946: 44; Mediterranean according to WH I: 303, EM: 155, EDG: 735). WH (I: 303) would prefer a loan from Greek, but the geminate *l* in Latin seems to preclude it. They also rule out inherited cognancy between the forms. However, the two forms seem easy to link to PIE **kel-* ‘cover’ (cf. Muller 1926: 108). Greek *κολέον* < **κολεφόν* < **kol-ey-* can be a suffixal full-grade to **kel-u-*, whence Italic could have produced a derivative **kol-u-ejo-* > **colleus* (cf. already Meyer 1887: 163). It requires us to make the small assumption that *culleus* is a non-urban form of **colleus* rather than the very large assumption that these forms must be borrowed from a non-IE language.

fūnis ‘rope, cable’

Pre-from: **g^{wh}oiH-ni-* | PItal. **χ^woini-*

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: tool

Pokorny (272), WH (I: 567-8), EM (262), DV (220, 222, 250)

Bezenberger and Fick (1881: 239), Thurneysen (1888: 351), Osthoff (1892: 303), Niedermann (1930: 7), Alessio (1944a: 108), Bertoldi (1948), Furnée (1972: 391), Bammesberger (1990), EDG (1438), Smoczyński (2018: 109)

Lat. *fūnis* ‘rope’ and *fīnis* ‘boundary’ have been linked as substrate words with an *i ~ u* alternation in comparison with Gk. *σχοῖνος* ‘rush, reed, rope of plaited rush; a land measure’ (cf. Alessio 1944a: 108, Bertoldi 1948). The latter can be reconstructed to **sg^hoiH-no-*⁴⁷³ whereas both Latin forms require **g^{wh}*. Rather than an irregular alternation however, the Greek word is probably unrelated. The semantically closer of the two Latin words, namely *fūnis* ‘rope’ (already rather distant if the primary Greek meaning is ‘rush, reed’), has a convincing IE etymology.

Since the alternation between Lat. *ī* and *ū* can be reconstructed to two ablaut grades of a root containing a diphthong: **ei* and **oi*, some connect *fīnis* and *fūnis* as inherited forms (cf. Niedermann 1930: 7). But *fūnis* ‘rope’ on semantic grounds is more likely the *o*-grade to *fīlum* ‘thread, line’ < **g^{wh}iH-* (cf. MW *gieu* ‘sinew, nerves’, Lith. *gýsla* ‘vein, sinew’, etc., EM 262, DV 22, 250 *pace* WH I: 498).⁴⁷⁴ The etymology of Lat. *fīnis* ‘boundary, limit, territory’ is not completely clear,⁴⁷⁵ but as no inscriptional forms attest

⁴⁷³ With the assumption of the de Saussure effect. It is considered Pre-Greek by Furnée (1972: 391) and EDG (1438) due to the Hesychian form *κοίνα* ‘fence’.

⁴⁷⁴ Some alternatively compare it to Gk. *θῶμυξ* ‘cord, string; bowstring’ (Pokorny 272, WH I: 567-8), potentially not of IE origin (EDG 569), but this does not work as well.

⁴⁷⁵ Proposals include **fīg-snis* to *fīgō* ‘to drive in, implant’ (WH I: 503 with lit., EM 237); derivation from **b^hiH-* ‘to hit’ (Thurneysen 1888: 351, Osthoff 1892: 303, DV 222); from **b^hiH-n-* otherwise attested in PGM. **baina-* ‘bone, leg’ (Bammesberger 1990); relationship to Lith. *baigti* ‘to finish’ (Bezenberger & Fick 239), though the semantic development from ‘to break’ > ‘to finish’ is isolated to Baltic and potentially late (cf. Smoczyński 2018: 109).

to a diphthong, its \bar{i} is probably original (WH I: 503, Bammesberger 1990: 264). In any case, it is semantically distant enough to be unrelated to Gk. $\sigma\chi\omicron\iota\nu\omicron\varsigma$ or Lat. *fūnis*.

grāmen, -inis ‘grass’

Pre-form: **g^hrh₁-(s-)mn-* | PItal. **grāsmen-* / **grāmen-*

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: plant, wild

Pokorny (404), WH (I: 616-7), EM (280), DV (269)

Schrijver (1991: 487), Kroonen (2013: 187), van Beek (2022: 386-8)

A connection with Skt. *grāsate* ‘devours’ and Gk. $\gamma\rho\acute{\alpha}\omega$ ‘to gnaw, eat’ (Pokorny 404, EM 280; WH I: 616 are skeptical) is unlikely. The Sanskrit and Greek forms can reconstruct to **gr^hs-* (van Beek 2022: 386-8), which would probably give PItal. **grens(-men-)* > Lat. ***grēmen*.⁴⁷⁶ It is also semantically very weak.

Otherwise the closest match for *grāmen* is PGm. **grasa-* ‘grass’, though the Germanic cannot reconstruct to **g^hrh₁s-* like Latin as this would yield ***gurs-* (pace Schrijver 1991: 487). DV (269) therefore suspects a substrate origin. Kroonen (2013: 187) explains the Germanic form as a secondary *s*-stem to the verb **grōan-* ‘to grow’ < **g^hróh₁-e-* in the way that the *s*-stem **glasa-* ~ **glaza-* ‘glass’ was formed from **glōan-* ‘to glow’.⁴⁷⁷ Instead of demonstrating an irregular correspondence, it seems Lat. *grāmen* and PGm. **grasa-* are two independent treatments of the same IE root.

mūrex ‘the purple-fish (a mollusk used to make purple dye)’

Pre-form: **muh₂s-* | PItal. **mūsV_k-*

Comp.: **muh₂s-* | PGk. **mūsak-* | Gk. $\mu\acute{\omicron}\alpha\zeta$ ‘sea mussel’

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: animal, wild; aquatic; economic

Pokorny (752-3), WH (II: 129), EM (422-3), DV (395)

EDG (973)

EM (422-3) and DV (395) consider the correspondence between Lat. *mūrex* and Gk. $\mu\acute{\omicron}\alpha\zeta$ to point to a substrate word. However, WH (II: 129 with lit.) and EDG (973) take both words as built on PIE **muh₂s-* ‘mouse’. Despite the $-\alpha\zeta$ suffix being often found in non-native words (EDG 973), Lat. *musculus* ‘little mouse’, ‘muscle’, and ‘mussel’ and

⁴⁷⁶ Even if there is a change **CCCC* > **CaCCC* in Latin, this does not occur in forms containing a syllabic nasal (Schrijver 1991: 496).

⁴⁷⁷ It is tempting to adduce *herba* as a full-grade form of the **g^hrh₁-* root behind *grāmen*, thereby bolstering the vegetable semantics of the root, but Germanic seems to show that the root is **g^hreh₁-*, not **g^herh₁-*.

mūs marīnus ‘salt-water fish, shellfish’ show that there was a preexisting semantic connection between mice and shellfish (WH II: 129).

pollen, -inis ‘flour, powder’

Pre-form: **polH-uen-* | PItal. **polwen-*

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: culinary

Pokorny (802), WH (II: 331-2), EM (519), DV (474, 498)

Schrijver (1991: 25-7), Nussbaum (1997: 197), Derksen (2007: 395), EDG (1220), Pronk (2011: 187)

DV (498) proposes a connection between Lat. *pollen* ‘flour, powder’ and Lat. *puls, -tis* ‘porridge’ via a Mediterranean loan or an otherwise unknown PIE root. But both of these can be derived from PIE **pelH-* ‘to swing’, despite some semantic concerns (Nussbaum 1997: 197, DV 474), which I think are not impossible to overcome in the context of agricultural processing. That the root ends in a laryngeal is guaranteed by the Balto-Slavic chaff words: Lith. *pėlūs* [nom.pl.], Ru. *polóva*, etc. < PBSl. **pelʔus*, **pelʔuaʔ* < **pelH-u-* (Derksen 2007: 395). Skt. *palāva-* ‘chaff’ seems to represent an ablaut grade of this *u*-stem as **pelH-óu-*. From an *o*-grade of this *u*-stem was formed a heteroclitic **ur/uen* stem of which **polH-uen-*, having lost its laryngeal to the de Saussure effect, is preserved in Lat. *pollen* (cf. Nussbaum 1997: 197). Without the effect, the laryngeal could have been lost in the full-grade of an *n*-stem paradigm **polH-en-* ~ **polH-n-*, producing **polen* ~ **poln-*. The latter would become *pollis* (an attested by-form), from which the geminate *ll* was generalized to the nominative, after which **pollen, pollis* > *pollen, pollinis* (Schrijver 1991: 25-7, DV 474, Pronk 2011: 187). *Puls* < **polti-*, if not a loan from Greek *πόλτος*, can be from **polH-ti-* (cf. WH II: 387-8 with lit., DV 498, EDG 1220).

sarp(i)ō, -ere ‘to prune’

Pre-form: **sHrp-* | PItal. **sarp-(j)e-*

Comp.: **sʔp-eh₂* | PGk. **sarpā-* | Gk. *ἄρπη* ‘sickle’

**sʔp-* | BSl. **sʔrpb-* | Latv. *sirpis*, OCS *srǫpъ*, Ru. *serpb* ‘sickle’

**sorp-nó-* | PGM. **sarpa-* | OHG *sarf*, etc. ‘sharp, severe’

?**s(o)rp-o-* | PAnat. **sarpa-* | Hitt. *sarpa-* ‘harrow’

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: action

Pokorny (911-12), WH (II: 480), EM (595), DV (540)

Puhvel (X: 149, 195), Schrijver (1991: 493), Karulis (1992 II: 187-8), Kloekhorst (2007

s.v. *šārr-ⁱ*), Matasović (2009: 330), EDG (138), van Beek (2022: 425-6), Kroonen et al. (2022: 17-18)

Gk. *ἄρπη* ‘sickle’ matches well formally and semantically with the reflexes of Balto-Slavic **syrp-* (Karulis 1992 II: 187-9), both of which can represent the zero-grade of a root **serp-* (cf. van Beek 2022: 425-6). PGM. **sarpa-* ‘severe, sharp’ can be from an *o*-grade derivation **sorp-nó-* via Kluge’s Law (Kroonen et al. 2022: 17-18).⁴⁷⁸ The *a*-vocalism of Lat. *sarp(i)ō* is difficult to analyze. Schrijver (1991: 493) suggests it may have taken its *a* from *sarrīre* ‘to hoe, weed’ or that the PPP **syp-to-* yielded *a*-vocalism in a cluster **C_CCC*. EDG (138) is suspicious and suspects a substrate word.

Puhvel (X: 149) compares Hitt. *sarr-*, *sar(r)a-*, *sarriya-* ‘separate, sever, etc.’, but Kloekhorst (2007: s.v. *šārr-ⁱ*) reconstructs an otherwise isolated *seṭ*-root **serh₁-*. Puhvel (X: 195) further compares Hitt. *sarpa-* ‘harrow’ to the sickle words. The semantic distance is not small, but the comparison can be preserved if Hitt. *sarpa-* represents an independent derivation from the root behind Latin, Greek, Germanic, and Baltic **serp-*, perhaps with the original meaning ‘sharp’ having been preserved only in Germanic. On the other hand, Germanic reconstructs with a derivational **-no-* suffix, suggesting that ‘sharp’ is the derived rather than basal meaning.

If the problems with the Latin vocalism are solved in one of the aforementioned ways, then perhaps the European IE languages are descended from an inherited root **serp-* whose meaning shifted after the split of Anatolian. Its link to a verb **ser-* ‘to cut’ and thus relationship to the verb *sarrīre* is attractive given Skt. *srñī-* ‘sickle’ (WH II: 480 with lit., EM 595, Schrijver 1991: 493) but remains uncertain.

testa ‘earthenware vessel, tile, sherd, shell’

Pre-form: **te(k)-s-teh₂-* | PItal. **te(k)stā*

Comp.: **te-tk-to-* / **tek-s-to-* | PIIr. **taštā-* | Av. *tašta-* ‘bowl, cup’

□ Irreg. correspondences

□ Remarkable phonotactics

Semantics: vessel

Pokorny (1058-9), WH (II: 675-6), EM (688-9), DV (617, 619)

LIV2 (s.v. **tek-*, **tek^k-*), Meiser (2010: 96-7), Weiss (2020: 197)

Lat. *testa* ‘earthenware vessel’ may be built on the same root as *texō*, *-ere* ‘to weave, construct’ (WH II: 675-6, Pokorny 1058-9), via PItal. **tekstā* (cf. **Sekst-ijos* > *Sēstius* with secondary lengthening, Weiss 2020: 197).⁴⁷⁹ DV (617) doubts the connection on

⁴⁷⁸ The independence of OIr. *serr*, OW *serr* ‘sickle’ < PCelt. **serrā-* is unclear. Matasović (2009: 330) takes it from the full-grade **serp-* but a borrowing from Latin cannot be ruled out (cf. recently Kroonen et al. 2022: 17-18).

⁴⁷⁹ DV (619) argues that the root of *texō* is **tek^k-* ‘to build’ via **tek^k-s-*. Others (Meiser 2010: 96-7, LIV2

semantic grounds, instead suspecting a loanword given the semantic field of vessels. But it does not seem inconceivable that a word for pottery could develop from a verb for ‘to build/construct’. *Testa* is further compared to Av. *tašta-* ‘bowl, cup’. While the latter can derive from **teṭk-to-*, it cannot be ruled out that both Lat. *testa* and Av. *tašta-* reflect an original **teḱ-s-to-* to the root **teḱ-* ‘to build’.

vīnum ‘wine’

Pre-form: **uīh₁-no-* | PItal. **wīno-*

□ Irreg. correspondences □ Remarkable phonotactics

Semantics: viticulture

Pokorny (1120-22), WH (II: 792-3), EM (737-4), DV (680)

Meillet (1908), Bertoldi (1939b: 86), Bertoldi (1942: 162), Alessio (1944a: 108), Battisti (1960: 351, 367), CAD (I/J: 152), Gamkrelidze & Ivanov (1995 I: 557-61), Agostiniani (1998), Greppin (1998), Klimov (1998: 227), Fähnrich (2007: 486, 501), Martirosyan (2009: 214), Zohary, Hopf & Weiss (2012: 121-6), Gorton (2017), Lipp (2020)

Because wine is understood to have been developed in the Pontic regions, and because it was a crucial Mediterranean trade item, nearly all early scholarship assumed that Lat. *vīnum* and its many relatives were loans from a Mediterranean language (Meillet 1908, Bertoldi 1942: 162, Alessio 1944a: 108, Pokorny 1120-22, WH II: 792-3, Battisti 1960: 367, EM 737-4, etc.). So convinced were they, that e.g. Bertoldi (1939b: 86) and Battisti (1960: 351) even show how *vītis* ‘vine’ is inherited but still remain convinced that the pair *vīnum* ~ οἶνος are from the Mediterranean substrate. The word’s widespread presence in Semitic as well as the fact that the Sabellic attestations ruled out the *ī* of Latin originating from a diphthong seemed to prolong the confusion.

But Lat. *vīnum*, U ***vinu***, Gk. (Ϝ)οἶνος, Hitt. *wīyan-*, Arm. *gini*, Alb. *verë/venë* are all inherited. The Celtic, Germanic, and Balto-Slavic forms were likely loaned from Latin. While Agostiniani (1998) argues that the Italic family has **wīno-* from Etruscan *vinun*, *vinum* (in turn from Greek), Lipp (2020: 208-11) shows that the opposite direction is just as likely. Gorton (2017) and Lipp (2020: 205-11) show that the IE forms can be derived from **ueh₁i-* ‘turn, twist’ (the same source as *vītis* ‘vine’) through different ablaut grades of an *n*-stem formation with the meaning ‘grapevine’ (cf. also Gamkrelidze & Ivanov 1995 I: 557-61).⁴⁸⁰ Given that the meaning ‘wine’ exists for the *n*-stem in Hittite, it seems that the lexeme had this meaning before the split. Additionally, since the *n*-stem is athematic in Anatolian but is thematic everywhere else, the meaning ‘wine’ could have

s.v. **tek-*, **teṭk-*) reconstruct it to a separate **tek-* ‘to weave, braid’ on the strength of Arm. *fek'em* ‘twist, warp’ < **tek-* and given that most continuants of **tek-* actually reflect reduplicated **te-tk-* (which probably would not yield *texō*, DV 619).

⁴⁸⁰ *Vitis vinifera*, the grape vine, grows not only in the Mediterranean, but also across southern Europe, SW Asia, and across the southern Caspian belt (Zohary, Hopf & Weiss 2012: 121-6).

been solidified in Core PIE with the thematicization being a genitive derivation; thus **μ(ο)ῖη-*n-o-** ‘of the vine’ > ‘wine’. This means that PKartv. **γvin(i)* ‘wine’ is a loan (Klimov 1998: 227) from a pre-stage of Armenian (Greppin 1998, Martirosyan 2009: 214).⁴⁸¹ The Semitic forms, absent from East Semitic (Akk. *īnu* is a borrowing from a non-East Semitic dialect, CAD I/J: 152, p.c. Benjamin Suchard) were borrowed from Indo-European languages.

2.4.3 Loan from a Known Language

alcēdō ‘kingfisher’

Lat. *alcēdō* and Gk. ἀλκυών have long been considered a Tyrrhenian-Aegean pair borrowed from the Mediterranean substrate (e.g. Alessio 1941c: 149). EM (20) mention the possibility that both are borrowed from a Mediterranean language, but otherwise follow Pokorny (302-4), WH (I: 27-8), and EDG (71) in suggesting a loan from Greek. The Greek word has no etymology, but the variant ἀλκυδών attested in Herodianus could easily have served as the base of the Latin form, with the suffix nativized to *-ēdō* (cf. EDG 71).

alica ‘emmer groats’

Lat. *alica* was likely borrowed from the oblique of Gk. ἄλιξ ‘wheat groats’, perhaps in the context of Greek medicine (EM 21). WH (I: 29) takes Gk. ἄλιξ from ἀλέω ‘to grind’ on analogy with *ptisana* ‘barley groats’ < Gk. πτισάνη ‘peeled barley, barley groats’ < πτίσσω ‘to shell/grind grains by stamping’. EDG (69) is not convinced, and considers the source of the Greek word unknown. Thus DV (33) suggests the Latin could still be an independent loan from a substrate language. A relationship with Hitt. *ḫalki-* ‘grain, barley’ is sometimes proposed (cf. Polomé 1952: 451, Puhvel III: 39), nor is the inherited status of the Hittite word guaranteed (cf. Kloekhorst 2007 s.v. *ḫalki-*). The deeper origin of the Greek word, whether it is connected to Hittite in some way, does not actually bear on the possibility that Latin *alica* was borrowed from it however.

anīsum ‘anise’, *anēthum* ‘dill’

Though EM (32) call this a Mediterranean word, it cannot be ruled out that Lat. *anīsum*, *anēthum*, and variants are borrowed directly from Gk. ἄν(ν)ησ(σ)ον ‘anise’ (with variant ἄν(ν)ισον) and ἄν(ν)ηθον ‘dill’ (WH I: 846).

brīsa ‘skins of pressed grapes’

⁴⁸¹ If, as e.g. Fähnrich (2007: 486) argues, the Kartvelian material can be derived within Kartvelian from a verbal root **γun-* ‘to bend, wind’ then we are faced with an interesting predicament. The Indo-European material is also internally derived from a root ‘to turn, twist’. The ablaut gradation within the IE languages and the fact that Lat. *vītis* ‘vine’ and English *withe* are further derived from the verb seems to tip the balance in favor of a PIE origin. In Kartvelian the root **γun-* seems to have as its primary meaning ‘to bend’ (cf. Svan *u-γwn-a* ‘elbow’, Old Georgian *romel vals iγunal* ‘who goes around bent over’, etc. in Fähnrich 2007: 501).

The intervocalic *s* of Lat. *brīsa* points to a recent loan. But it is difficult to decide how it entered Latin.

It is widely connected with forms attested in Greek: Gk. βρῦτος, Hsch. βροῦτος: ἐκ κριθῶν πόμα, Hsch. βρότιον: πόμα ἐκ κριθῆς ‘barley beer’; Gk. βρύτεια ‘refuse of olives or grapes’ (WH I: 116, EM 76). The variation between the Greek words shows that they are loans there, and the source is often assumed to be Thracian (Frisk 1960-72 I: 273, Chantraine 1968-80: 199, EDG 245), though the evidence is not overwhelming.⁴⁸² In any case, Lat. *brīsa* cannot be a direct loan from Greek (Biville I: 275). If the words are in fact related—it is strange that most of the Greek forms refer to barley beer, not grapeskins—Latin has *brīsa* from a different source.

Thus it has been proposed that *brīsa* entered Latin via Illyrian (Brüch 1922: 244-5, Krahe 1955: 117). Through a more modern lens, we can wonder if a form ancestral to Alb. *bërsí* ‘remains of pressed grapes, plums, olives’ is the more proximal source of Lat. *brīsa*. Demiraj (1997: 98 with lit.) notes that *bërsí* cannot be a direct borrowing from the Greek forms or their putative Thracian source (*pace* Orel 1998: 23). It represents a metathesis from PALb. **brīšā-* < **b^(h)rīšā-*. Its pre-form has the *ī*-vocalism of Latin against the *ū*-vocalism of Greek, an alternation that also occurs between *ficus* and σῦκον ‘fig’. It seems like metathesis in Albanian can occur quite late, after contact with Latin (cf. Alb. *tërffil* ‘clover’ < Lat. *trifolium* ‘clover’). Given that PALb. **brīšā-* is all but identical to Lat. *brīsa* both formally and semantically, and since Albanian can have produced the sibilant from a cluster (**tī*) that includes the dental present in the Greek attestations, it is difficult to rule out that it is the most proximate source of Lat. *brīsa*. Cf. a similar situation for Lat. *sīca* (s.v.).

cēra ‘wax’

Baltic words for ‘honeycomb’ reconstruct to **kār-* (Lith. *kor̃ys*, Latv. *kāre*). This forms a non-IE *ā* ~ *ē* alternation with Gk. κηρός ‘wax’.⁴⁸³ The suffix of κήρινθος ‘bee-bread’ further points to a non-IE origin (Alessio 1944a: 130, Alessio 1946a: 161-2, EDG 689). Lat. *cēra* ‘wax’ is certainly related, but it cannot be ruled out that it is a loan from Greek (cf. WH I: 202 with lit.) The change in gender has been explained as due to influence from *tabella* ‘tablet’ (in the sense of wax writing tablets) or *crēta* ‘clay’ (from its use for

⁴⁸² Chantraine (1968-80: 199) says the best evidence is from Archilochus. But the (rather lewd) passage reads: ὥσπερ ἀλῶδι βρῦτον ἢ Θρεΐτζ ἀνήρ ἢ Φρυγὸς ἔμυζε ‘like a Thracian or Phrygian man sucks βρῦτος through a straw’. The peculiarly Thracian (or Phrygian) aspect may be the straw rather than the beer. Hellicanus (*apud* Athenaeus) writes πίνουσι δὲ βρῦτον ἐκ τινῶν ῥιζῶν, καθάπερ οἱ Θρεῖτες ἐκ τῶν κριθῶν ‘they drink βρῦτος made of certain roots, similar to the Thracians who make it of barley’. Rather than βρῦτος being a Thracian drink, it seems like a drink that Thracians make a certain way. An origin in an IE language whose reflex of **b^h* was *b* would allow this word to be connected to **b^hru-* (cf. Lat. *defrūtum* ‘grape juice reduction’). But we know too little about Thracian phonology to confirm that it had this treatment.

⁴⁸³ The *ē* is genuine. Fick (1890-1909 I: 378) claimed the existence of a Doric κᾶρός, but it does not exist (Osthoff 1901: 21-2, WH I: 202, Frisk 1960-72 I: 844, EDG 689). In fact, κήρινος ‘of wax’ is attested in Alcman, a Doric writer.

sealing). Van Sluis (2022: 18) proposes Etruscan mediation.

gabata ‘bowl, wooden vessel’

Lat. *gabata* ‘bowl, wooden vessel’ along with several Greek forms with consonant alternations (Gk. καθάθα⁴⁸⁴, Hsch. γάβαθον· τρυβλίον ‘bowl’, Hsch. ζάβατος· πίναξ ἰχθυηρὸς παρὰ Παφίας ‘trencher for fish’) and Romance forms that refer to geological features (PRom. **gabatro-*: OProv., Prov. *gaudre* ‘ravine, torrent’; PRom. **gabara/o-*: Béarnaise *gabe* ‘torrent’, OProv. *gaura* ‘canal’) are often considered independent loans (either from a Mediterranean substrate [Hubschmid 1950a: 39, Furnée 1972: 116] or a Semitic source [WH I: 575, EDG 253]). But the appurtenance of the semantically dissimilar Romance forms is uncertain. In any case, given the Greek variants with initial γ and feminine gender, we cannot rule out that Lat. *gabata* is a loan from Greek, with regular *t* for θ.

matula ‘pot, vase, chamberpot’

The etymology of Lat. *matula* ‘pot, vase, chamberpot’ is uncertain (WH II: 53, EM 391), but Furnée (1972: 194, 212) convincingly compares it to Gk. μαθαλίς ‘type of cup, measure of volume’. The Greek word itself shows evidence of not being inherited (cf. also EDG 891), but the correspondence of Lat. *t* to Gk. θ in loans is not unexpected. Nor is the Latin form strange if we, like Biville (I: 153) follow André (1959: 87) in supposing it was borrowed from an unattested Gk. *μαθάλη. Despite the form being unattested, other such -ίς, -ίδος / ἦ, -ῆς pairs are known (cf. at *clapar*, s.v.: κάλπις, -ίδος ‘jug, urn’ vs. κάλπη ‘pitcher’).

nepeta ‘catnip’

Bertoldi (1936: 300-4) uses the Etruscan city names Νέπετα and Νέπιτα as well as the personal names *Nepius* and *Neponia* to propose that Lat. *nepeta* is from an Etruscan root **nep-* ‘damp’, as catnip thrives in damp places. The deity *Neptūnus* would then be an Etrusco-Latin name of a river god, whence Lat. *neptūnia* ‘a kind of mint’. Alessio (1941b: 224) gives some Italian words from Marche that seem to continue the root **nep-* (*nebbi* ‘*Sambucus racemosa*’).⁴⁸⁵ André (1956: 218) follows the Etruscan etymology while EM (437) are noncommittal and WH (II: 160) are suspicious, seeing as *Neptūnus* has a good IE etymology. Since both νέπετος and νέπιτα are attested in Greek, we cannot rule out that Latin borrowed a Gk. *νέπετα (Furnée 1972: 257, EDG 1010), even if the *e ~ i* alternation within Greek points to a word of substrate origin.⁴⁸⁶

olīva ‘olive’, ***oleum*** ‘oil’

⁴⁸⁴ Accent technically unknown, attested in papyri and the Edict of Diocletian.

⁴⁸⁵ His identification of Libyan *nepa* ‘crab, scorpion’ (reported by Festus) with ‘crayfish’ certainly goes too far.

⁴⁸⁶ That *nepeta* has not undergone vowel weakening to ***nepita* could point to a recent loan or be due to the *alacer* rule (cf. Weiss 2020: 128-9 on the rule).

Despite interpretations of a loan from the same Mediterranean source as Gk. ἔλαια ‘olive’ < *ἐλαίφα (e.g. Terracini 1929: 214, Bertoldi 1942: 162) or via Etruscan (e.g. Battisti 1959: 360), Lat. *olīva* ‘olive’ and *oleum* ‘oil’ are perfectly regular early loans from Greek (WH II: 205-6, Biville I: 86-7, EM 460). From **elaiwa*: **e* > *o*/_p^{inguis}, internal **ai* weakened to **ei* and was fully monophthongized to *ī*. For neut. **elaiwon*: **e* > *o*/_p^{inguis}, internal **ai* weakened to **ei* and began to undergo monophthongization to *ī* through **ē*. The *w* was lost before *o*, making the **ē* antevocalic and triggering its loss of length before the completion of the change to *ī*, resulting in *oleum* (Biville I: 87). The origin of the Greek word is potentially obscure, but does not change the fact that the Latin words were borrowed from it.

sīca ‘dagger’

Despite how semantically attractive it is to connect *sīca* ‘dagger’ to *secāre* ‘to cut’, it is formally impossible (cf. DV 561); the solutions mentioned by WH (II: 505 with lit.) are outdated,⁴⁸⁷ and already Pokorny (895-6) questioned the appurtenance. A connection with Lith. *šỹkis* ‘time, occasion; blow’ (cf. WH II: 505) is semantically unlikely (cf. the alternative etymology in Smoczyński 2018: 1163). Brugmann’s (1894: 260-1) connection via a *-*k* extension to the root in Skt. Skt. *śēnā-* ‘missile; battle line’, *śāyaka-* ‘missile, arrow’, *prāsītī-* ‘line of fire’ is unlikely since the root seems to have meant ‘to throw’ (cf. EWAia II: 186, 725, 746). Romance reflexes of *sīcīlis* ‘spearhead’, probably related to *sīca* (WH II: 533, EM 623), have short vowels (**sīcīlis*, REW no. 7900). DV (561) and EM (623) both consider *sīca* a potential loanword, the latter from Thrace.

The source is probably closer: *sīca* is very likely related to Alb. *thikë* ‘knife’ (Orel 1998: 477-8 with lit.). A loan from Latin is ruled out, as Lat. *s-* is borrowed as Alb. *sh-*. On the other hand, (Pre-)PALb. **tsīkā* could plausibly have yielded *sīca*. The source of the Albanian form is unclear. Orel proposes a reconstruction to a root to sharpen (cf. Skt. *śā-* ‘to sharpen’, Arm. *sowr* ‘sharp’, Lat. *cōs* ‘whetstone’), but neither his reconstruction **kēi-* nor more commonly reconstructed **kēh₃-* (cf. Mayrhofer II: 627, DV 139; though Schrijver 1991: 91 reconstructs **kēh₁-*) can explain the vocalism or the second velar. Thus it cannot be fully ruled out that both forms are borrowed from a third source. But a borrowing from a form ancestral to Albanian does not seem problematic. (Cf. a similar situation for *brīsa*, s.v.)

turba ‘commotion, upheaval’

Pokorny (1100-1) and Meiser (2010: 63) explain Lat. *turba* and Gk. τὺρβη/σύρβη from

⁴⁸⁷ Also often linked with *secō* is poorly attested (though borrowed into Old Irish as *scían* ‘knife’) *s(a)cēna* ‘sacrificial axe’ (EM 585), often along with *saxum* ‘stone’ (WH II: 459). There are again formal difficulties (cf. DV 440, 541). Breyer (1993: 272-3) supports Etruscan origin because of the presence of a root *sac-* ‘to consecrate, sanctify’. But this could easily be a borrowing from Latin (cf. *sacer* ‘holy’). Rosén (1994) instead compares Hebr. *śakkīn*, Aram. *sakkīn* ‘slaughtering knife’, which she considers independent borrowings from a third source.

an inherited root **t̥mer-* (cf. Skt. *tvarate* 'to hurry'), but this requires the problematic assumption of a root extension **b* (DV 634).⁴⁸⁸ Given the formal and semantic closeness of the Lat. *turba* to the Greek variant τὺρβη, it is potentially a loan (Biville II: 271, EDG 1520). EDG favors this on the understanding that the variation within Greek points to Pre-Greek origin there. Otherwise, Vasmer (1959-61 s.v. τὺρβη) notes the similarity of the Greek to PSlav. **turgb* 'merchandise, market'. If related, the velar is **gʷ*, from which Latin could not have produced *b* (i.e. its source must be Greek). DV (634) notes that *turba* has well-developed variations already by Plautus, suggesting that it has been in Italic for a long time and was thus borrowed independently from the same non-IE source as the Greek words. While *turma* 'small squadron, company' may point to a non-IE *b ~ m* alternation in this lexeme, its imperfect semantic match makes its appurtenance uncertain. In the end, beyond the potential difficulty of having to propose quite an early date for the borrowing, it does not seem possible to reject that Lat. *turba* is a loan from Gk. τὺρβη.

⁴⁸⁸ He looks for a root of the shape **(s)terb^h*- but finds only isolated Gk. στρέφω 'to turn'. While a zero-grade of this root could in fact yield the *u*-vocalism of *turba* (cf. *turdus* and *scurra*), the *s* mobile explanation and **TeD^h* root structure behind the Latin form are suspicious.

2.5 Latin Index for the Data Section

—Used in the Analyses—

2.2 Non-inherited Origin in Latin Accepted

2.2.1 Phonotactic Reasons

2.2.1.1 Isolated to Latin but with Unrhotacized S

asīlus 42 *asinus* 43 *casa* 45

2.2.1.2 Isolated to Latin but with an Invalid Root Structure

faex 46 *farciō* 47 *pampinus* ... 48 *tabānus* 48

2.2.2 Comparanda in Other Branches

2.2.2.1 Non-Inherited Origin is Probable

<i>alnus</i> 49	<i>citrus</i> 71	<i>fungus</i> 90	<i>orca</i> 113
<i>ascia</i> 51	<i>columba</i> 72	<i>gubernō</i> 93	<i>pirum</i> 114
<i>avēna</i> 53	<i>cotōneum</i> ... 75	<i>hasta</i> 95	<i>plumbum</i> ... 115
<i>baculum</i> 55	<i>cucumis</i> 76	<i>hedera</i> 96	<i>racēmus</i> ... 117
<i>bāiulus</i> 56	<i>cupressus</i> ... 77	<i>hirundō</i> 96	<i>rāpum</i> 119
<i>ballaena</i> 58	<i>ervum</i> 78	<i>lacerna</i> 98	<i>raudus</i> 120
<i>burrus</i> 58	<i>faba</i> 80	<i>laena</i> 99	<i>rosa</i> 121
<i>buxus</i> 59	<i>far</i> 81	<i>lapis</i> 100	<i>sabulum</i> ... 123
<i>caballus</i> 60	<i>fascinus</i> 82	<i>laurus</i> 102	<i>simila</i> 125
<i>calix</i> 63	<i>ferrum</i> 83	<i>lēns</i> 103	<i>sirpe</i> 126
<i>calpar</i> 64	<i>ficus</i> 85	<i>līlium</i> 105	<i>sōrex</i> 127
<i>cant(h)ērius</i> 65	<i>filix, felix</i> ... 86	<i>malva</i> 106	<i>taeda</i> 128
<i>carbasus</i> 66	<i>fracēs</i> 87	<i>menta</i> 108	<i>turdus</i> 129
<i>carpinus</i> 68	<i>frīgō</i> 88	<i>merula</i> 109	
<i>caulae</i> 69	<i>fulica</i> 89	<i>mūlus</i> 110	
<i>caupō</i> 70	<i>funda</i> 90	<i>nux</i> 111	

2.2.2.2 Non-inherited Origin is Possible

<i>adeps</i> 130	<i>bāca</i> 139	<i>catulus</i> 146	<i>grāmiaie</i> ... 152
<i>alaternus</i> .. 131	<i>badius</i> 140	<i>cēpa</i> 147	<i>grūmus</i> 153
<i>ālium</i> 133	<i>barba</i> 141	<i>corbis</i> 147	<i>nāpus</i> 154
<i>aper</i> 134	<i>bolunda</i> ... 142	<i>cucurbita</i> .. 149	<i>paelex</i> 155
<i>arāneus</i> ... 135	<i>calx</i> 143	<i>excetra</i> 150	<i>pannus</i> 156
<i>ardea</i> 137	<i>caput</i> 144	<i>faber</i> 151	<i>rādīx</i> 157

<i>raia</i> 158	<i>tamarix</i> 161	<i>ulmus</i>166
<i>sappīnus</i> ... 159	<i>tilia</i> 162	<i>vaccīnium</i> .167
<i>sulpur</i> 160	<i>trabs</i> 164	<i>viscum</i>168

2.2.3 Comparanda only in Latin and Romance

<i>arbutus</i> 169	<i>genesta</i> 170	<i>lepus</i>172	<i>talpa</i>175
<i>cerrus</i> 169	<i>lābrusca</i> ... 171	<i>sambūcus</i> ..173	

—Not Used in the Analyses—

2.3 Origin Unclear

2.3.1 No Comparanda

<i>acinus</i> 176	<i>cicōnia</i> 178	<i>Lār</i>180	<i>rumex</i>183
<i>ās</i> 176	<i>ferula</i> 178	<i>mēlēs</i>181	<i>sagitta</i>183
<i>autumnus</i> .. 177	<i>fovea</i> 179	<i>nūtēla</i>182	<i>scurra</i>184
<i>balteus</i> 177	<i>hircus</i> 180	<i>puteus</i>182	<i>sīl</i> 185

2.3.2 Uncertain Comparanda

<i>abiēs</i> 185	<i>cirrus</i> 193	<i>īlex</i>201	<i>pūlēium</i>208
<i>aesculus</i> ... 186	<i>crux</i> 194	<i>lanx</i>201	<i>rēte</i>208
<i>alga</i> 187	<i>cuñīculus</i> .. 194	<i>larix</i>202	<i>saepēs</i>209
<i>apis</i> 188	<i>fibra</i> 195	<i>legūmen</i> ...203	<i>sēcale</i>210
<i>aulla</i> 188	<i>frōns</i> 196	<i>lemurēs</i>203	<i>sorbus</i>210
<i>bucca</i> 189	<i>frūmen</i> 197	<i>mantum</i>204	<i>spiōnia</i>211
<i>carbō</i> 190	<i>gīgarus</i> 197	<i>mergae</i>205	<i>sūber</i>211
<i>carīna</i> 190	<i>guttur</i> 198	<i>mūtulus</i>205	<i>tamīnia</i>211
<i>cāseus</i> 191	<i>hāmus</i> 199	<i>palātum</i>206	<i>unēdō</i>213
<i>cicāda</i> 192	<i>harundō</i> ... 199	<i>palla</i>207	
<i>cicūta</i> 192	<i>hirūdō</i> 200	<i>pērō</i>207	

2.3.3 Conflicting Possibilities

2.3.3.1 Non-Inherited vs. Inherited

<i>acer</i> 213	<i>falx</i> 219	<i>līnum</i>225	<i>termes</i>232
<i>apex</i> 214	<i>fax</i> 221	<i>lōrum</i>226	<i>tībia</i>233
<i>cancer</i> 215	<i>glārea</i> 221	<i>palumbēs</i> ..227	<i>trahō</i>234
<i>capiō</i> 216	<i>haedus</i> 222	<i>pīnus</i>228	<i>tūber</i>235
<i>cicer</i> 217	<i>īnsula</i> 223	<i>porrum</i>229	<i>tumulus</i>236
<i>cubō</i> 218	<i>iūniperus</i> .. 224	<i>sapa</i>230	<i>verbascum</i> 237
<i>dulcis</i> 219	<i>labium</i> 225	<i>simpuvium</i> .231	

2.3.3.2 Non-Inherited vs. Loan from a Known Language

<i>ātriplex</i> 238	<i>carpisculum</i>	<i>conger</i>241	<i>prūnus</i>244
<i>bardus</i> 239 240	<i>ibiscum</i>242	<i>supparus</i> ...243
<i>burgus</i> 239	<i>cibus</i> 241	<i>īdus</i>242	<i>taurus</i>245

2.3.4 Core-Periphery Cases

<i>caper</i> 247	<i>hordeum</i> ... 250	<i>porca</i>251
------------------------	------------------------	-----------------------

2.3.5 Methodologically Difficult to Delimit Comparanda

<i>campus</i> 252	<i>cūpa</i> 254	<i>glēba</i>255	<i>mōrum</i>257
-------------------------	-----------------------	-----------------------	-----------------------

2.4 Non-IE Origin in Latin Rejected

2.4.1 No Positive Evidence of Borrowing

<i>aqua</i> 259	<i>corulus</i> 263	<i>frāga</i>268	<i>olor</i>272
<i>arcus</i> 259	<i>crātis</i> 264	<i>frutex</i>268	<i>ornus</i>272
<i>caelum</i> 260	<i>crēta</i> 265	<i>iuncus</i>269	<i>salix</i>273
<i>catīnus</i> 261	<i>dōlium</i> 265	<i>lībra</i>269	<i>scutra</i>274
<i>colus</i> 262	<i>fēlēs</i> 266	<i>mālus</i>270	<i>viola</i>274
<i>cornus</i> 262	<i>follis</i> 267	<i>mīlium</i>271	

2.4.2 Best Explained as Inherited

<i>agna</i> 275	<i>culleus</i> 277	<i>mūrex</i>279	<i>testa</i>281
<i>anguīlla</i> 276	<i>fūnis</i> 278	<i>pollen</i>280	<i>vīnum</i>282
<i>cōnīvēō</i> 277	<i>grāmen</i> 279	<i>sarp(i)ō</i>280	

2.4.3 Loan from a Known Language

<i>alcēdō</i> 283	<i>brīsa</i> 283	<i>matula</i>285	<i>sīca</i>286
<i>alica</i> 283	<i>cēra</i> 284	<i>nepeta</i>285	<i>turba</i>286
<i>anīsum</i> 283	<i>gabata</i> 285	<i>olīva</i>285	

3 Feature Analysis

3.1 Introduction to the Feature Analysis

As detailed in §1.4, it is the irregular alternations between comparanda that allow us to classify lexemes as originating from a source other than PIE. Karl Oštir was one of the first to list, in dense and copious detail, comparanda by irregular correspondence in order to give the evidence for a widespread Alarodian group of languages. In *Drei voroslavisch-etruskische Vogelnamen* (1930) for instance, he collected and classified alternations that he thought betrayed a *Alteuropäisch-Altkleinasiatisch* layer. These included (with his symbol ∞) voiced ∞ unvoiced, simplex ∞ geminate, metathesis, unaspirated ∞ aspirated, $l/r \infty$ dental or sibilant, $l/r \infty n$, $l/r \infty j$, dental ∞ sibilant, velar ∞ sibilant, labial ∞ nasal, as well as vocalic alternation, loss of plosives, liquids, sibilants, nasals, and the semi-vowels etc. His style was too dense (Meillet 1922b) and his analysis was critically lacking in rigor (Schuchardt 1922: 80). Nevertheless, many of the categories that he identified occur in the data presented here. They are not, however, indicative of origin in a single, common substrate.

A similar endeavor was made by the Pelasgianists and others looking for an Indo-European substrate amongst the Indo-European languages. Starting approximately with Georgiev, the Pelasgianists had identified alternations in aspiration (*mediae aspiratae* ~ *mediae*, *tenues* ~ *tenues aspiratae*) and voicedness (*mediae* ~ *tenues*) as well as labial quality (PIE $*k^w$, g^w , g^{wh} > Pelasgian k^h , k , g) and some vocalic alternation ($a \sim o$). Their explanation was that the irregularities are actual regular, borrowed from an IE language in which these are the regular reflexes of PIE material. Holzer's *Thematic* and Ribezzo and Szemerényi's Ausonian via the same explanation found different alternations in aspiration (*mediae aspiratae* ~ *tenues*), while the former found also voicing (*mediae* ~ *tenues*) alternations. Georgiev (1941: 111-44) identified some "angeblich unindogermanische Präfixe und Suffixe" including *a-*, *le-*, *-ss-*, *-nth-*, and *-l-*, which he explains as Pelasgian. Van Windekens (1952: 34-57) found for Pelasgian suffixes deriving from PIE $*-t-$, $*-k^w-$, $*-g-$, $*-p-$, $*-l-$, $*-n-$, $*-n-t-$, $*-mn-$, $*-nd^h-$, $*-rn-$, $*-s-$, $*-tj-$, etc. Of course, a major catalyst of the movement to study the substrates, IE or not, underlying the Indo-European languages, had been Kretschmer's (1896: 401-9) treatment of the Gk. $-\nu\theta\omicron\varsigma$ suffix.

The Italian scholars including Bertoldi, Battisti, and Alessio also found recurring irregular alternations amongst the words they assigned to the Mediterranean substrate. In Battisti's (1959: 385) index for example, he lists *a/e*, *b/p*, *d/l*, *d/t*, *e/i*, *f/v*, *i/u*, *k/g*, *l/ll*, *p/f*, and *r/rr*. They too purported to locate morphological features, mainly in the form of suffixes that they ascribed to different strata and locations, often based on the evidence of placenames. As examples, Alessio (1939, 1944a: 103) interprets the *-asco* suffix as evidence of a Ligurian origin. Bertoldi (1942: 196), Alessio (1944a: 102), and Battisti (1959: 196) gave evidence of a Mediterranean *-st-* suffix.

Furnée (1972) had gone through the literature on Pre-Greek and was dissatisfied with the previous scholars having assigned material to the substrate without detailing the phonological aspects. He listed the alternations he found, but interestingly proposed that they were already present in Pre-Greek, created as expressive formations. This led Beekes in several publications (esp. his 2014 *Pre-Greek: Phonology, Morphology, Lexicon*) to detail the phonological alternations and pieces of morphology that he considered indicative of Pre-Greek origin.

Several scholars between and after those mentioned have also sought to study the substratal lexicon of the Indo-European languages in a similar way: through listing the phonological alternations and non-inherited morphology. In this chapter, I will do the same for the dataset comprised of lexemes in §2.2 (*Non-inherited Origin in Latin Accepted*). What follows is an analysis of all of the irregular phonological alternations between Latin words and their comparanda that I have been able to identify. Following that is a discussion of some of the morphological features of these words that I consider diagnostic. The list is not exhaustive. Instead, the consideration of the morphology is secondary, a result of identifying words of non-IE origin by means of their irregular phonological correspondences. Suffixes and morphological phenomena that recur in relation to lexemes that can be identified as loans for other reasons may then themselves originate in the substrate languages; especially those that themselves attest to irregular phonological alternations.

In the tables, the Latin lexemes and their comparanda are sorted by which reflex of a *quasi-* (= “as if”) PIE phoneme they attest to (with non-IE languages in parentheses). When it cannot be determined due to sound laws which reflex is present, the word is listed in both places but inside of square brackets. The cells highlighted in gray show which reflex is attested in Latin. (Lighter gray marks the cases where Latin could reconstruct to either of two categories due to its medial treatment of the voiced aspirates). QPIE reconstructions follow those given in §0, and less certain comparanda (those marked with ? and ??) are left out.

3.2 Phonological Alternations

3.2.1 Consonants

3.2.1.1 Alternations between PIE Rows

Proto-Indo-European is reconstructed as having three “rows” of plosives: the labials, dentals, and velars. Within each was a further phonological interplay between two features, reconstructed as either voicing and aspiration (traditionally) or glottalization and fortition/lenition (in the glottalic theory). The velars could show a further distinction between palatalization and labialization. The combinations of features produced a series of phonemes whose reflexes in the daughter languages are well understood. Several of the irregular consonant alternations that allow the identification of lexical material as

non-native in origin exist within these rows.

3.2.1.1.1 Non-Velars

In the labials and dentals, voicing and aspiration produces the traditional *mediae aspriatae* (**b^h* and **d^h*), *mediae* (**b* and **d*), and *tenuis* (**p* and **t*). In the glottalic theory, the contrasts are instead between fortition/lenition and glottalization (various presentations in Hopper 1973, Gamkrelidze & Ivanov 1973, Salmons 1993, Beekes 2011: 128-9, etc.). One need not decide in favor of one or the other, but the choice has implications for the sort of substrate phonemes or dialectal variation underlying the different reflexes in the IE daughter languages. Conclusions will be different depending on whether one understands the alternation to be between **b^h* and **p* or between **p*: and **p*. This caveat is of course relevant for all the upcoming categories.

From a quasi-PIE perspective, there are for the non-velars only four possible combinations of irregular correspondence. Besides a alternation between all three types, the remaining three combinations have all been noticed; each has been explained in the context of the sound laws of a lost Indo-European language: *D^h ~ D* (cf. Pelasgian), *D^h ~ T* (cf. Temematic, Ausonian), *D ~ T* (cf. Pelasgian, Temematic). As will be seen, there is at least one Latin lexeme which, in comparison with its comparanda, fits into each of these alternations. The significance of this, and how it bodes for the stratificational power of these features will be discussed, as will the legitimacy of describing these alternations in terms of PIE phonology.

3.2.1.1.1.1 Labials

3.2.1.1.1.1.1 Voicing

QPIE * <i>b</i>	QPIE * <i>p</i>
QPIE * <i>burso-</i> : Lat. <i>burrus</i>	QPIE * <i>p(h₂)ur-s(-y)o-</i> : Gk. <i>πυρρός</i>
QPIE * <i>bukso-</i> : Lat. <i>buxus</i>	QPIE * <i>pukso-</i> : <i>πύξος</i>
QPIE * <i>karb-</i> : Lat. <i>carbasus</i>	QPIE *QPIE * <i>karp-</i> : Gk. <i>κάρπασος</i> QPIE * <i>karp-</i> : Skt. <i>karpāsa-</i>

Table 3.1 Alternations between **b* and **p*

Technically, the *b* of *carbasus* could reconstruct to **b^h*, but it entered Latin after rhotacism, much too late to be affected by the development of the voiced aspirates.

3.2.1.1.1.1.2 *Aspiration*

QPIE * <i>b^h</i>	QPIE * <i>b</i>
QPIE * <i>b^ha/h₂L-</i> : Gk. φάλλαινα	QPIE * <i>ba/HL-</i> : Lat. <i>ballaena</i>
QPIE * <i>b^ha/oer(s)d^h-</i> : PGm. * <i>bar(z)da-</i> [QPIE * <i>b^(h)a/ord^(h)-</i> : PBSl. * <i>bordá?</i>] [QPIE * <i>b^(h)a/orsd^(h)-</i> : Lith. <i>barzdà</i>]	QPIE * <i>ba/Hr(s?)d^h-</i> : Lat. <i>barba</i> [QPIE * <i>b^(h)a/ord^(h)-</i> : PBSl. * <i>bordá?</i>] [QPIE * <i>b^(h)a/orsd^(h)-</i> : Lith. <i>barzdà</i>]
QPIE * <i>b^ha/Hsk-</i> : Lat. <i>fascinus</i>	QPIE * <i>ba/h₂sk-</i> : Gk. βάσκανος
QPIE * <i>b^helik-</i> : Lat. <i>felix, filix</i>	QPIE * <i>blē/eh₁g^h-</i> : Gk. βλήχρον
QPIE * <i>b^hreg-</i> : PGm. * <i>brekna(n)-</i>	

Table 3.2 Alternations between **b^h* and **b*

To this group seems also to belong Lat. *fascis*, especially on comparison with Gk. φάκελος, φάσκαλος ~ Hsch. βάσκιοι. But if Lat. *baiulus* is indeed also related, then there is a **b^h* ~ **b* alternation attested within Latin as well.

3.2.1.1.1.1.3 *Voicing and Aspiration*

QPIE * <i>b^h</i>	QPIE * <i>p</i>
QPIE * <i>kub^h-</i> : Gk. κυφαρίσσινος	QPIE * <i>kup-</i> : Gk. κυπάρισσος QPIE * <i>kup-</i> : Lat. <i>cupressus</i> (Hebr. <i>gofer</i>)
QPIE * <i>Silb^h-</i> : Gk. σίλφιον (Berb. <i>azlaf, aselbu</i> , etc.)	QPIE * <i>sirp-</i> : Lat. <i>sirpe</i> QPIE * <i>Selp-</i> : Hsch. σέλπον
QPIE ? <i>*su(o)lb^h-</i> : PRom. * <i>su(l)fur</i> [QPIE * <i>sue(l)b^hlo-</i> : Go. <i>swibls</i>]	QPIE * <i>su(e/o)lp-</i> : Lat. <i>sulpur</i> [QPIE * <i>sue(l)p-ló-</i> : Go. <i>swibls</i>]
QPIE * <i>g^(h)ra/ob^h-</i> : PSlav. * <i>grabrъ-</i>	QPIE * <i>ka/Hrp-</i> : Lat. <i>carpinus</i>
QPIE * <i>h₂le/ob^h-</i> : Gk. ἄλειφα(ρ)	QPIE * <i>h₂edep-</i> : Lat. <i>adeps</i> QPIE * <i>h₂elep-</i> : PRom. * <i>ala/ep-</i>

Table 3.3 Alternations between **b^h* and **p*

For *cupressus* and *sirpe*, the **b^h* ~ **p* alternation exists within Greek. It is notable that in all cases, Latin attests to the unvoiced variant (but note Romance *sulfur*). In 2 cases, Italic treatment of the voiced aspirates obscures the original quality of the medial plosive, and it is unclear whether they represent **b* ~ **p* or **b^h* ~ **p* alternations:

QPIE * <i>b^(h)</i>	QPIE * <i>p</i>
QPIE * <i>da/Hrb^(h)-</i> : PRom. * <i>darbo-</i>	QPIE * <i>ta/Hlp-</i> : Lat. <i>talpa</i>
QPIE * <i>sa/Hb^(h)-</i> : Lat. <i>sabina</i>	QPIE * <i>sa/HP-</i> : Lat. <i>sappīnus</i> QPIE * <i>sa/HP-</i> : OCo. <i>sibuit</i>

Table 3.4 Alternations between **b^h* or **b* and **p*

In at least one case, alternations between all three qualities are attested:

QPIE *b ^h	QPIE *b	QPIE *p
QPIE *ka/ob ^h - : PSlav. *ka/ob- [QPIE *ka/Hb ^(h) - : Lat. <i>caballus</i>] [QPIE *kab ^(h) /p- : MoP <i>kawal</i>]	QPIE *ka/h ₂ b- : Gk. καβάλλης [QPIE *ka/Hb ^(h) - : Lat. <i>caballus</i>] [QPIE *kab ^(h) /p- : MoP <i>kawal</i>]	QPIE *ka/HP- : PCelt. *kappe/ilo- [QPIE *kab ^(h) /p- : MoP <i>kawal</i>]

Table 3.5 Alternations between *b^h, *b, and *p

Additionally, Lat. *rāpum* beside Gk. ῥάφους, ῥάπτως securely attests to a *b^h ~ *p alternation. Whether PCelt. *arbīno- reconstructs to *b^h or *b is obscured by Celtic sound laws.

3.2.1.1.1.2 Dentals

3.2.1.1.1.2.1 Voicing

QPIE *d	QPIE *t
QPIE *kudo-, *kodu- : Gk. κυδώνια, κοδύ-	QPIE *koto- : Lat. <i>cotōneum</i>
QPIE *da/Hrb ^(h) - : PRom. *darbo-	QPIE *ta/Hlp- : Lat. <i>talpa</i>
QPIE *deh ₂ - : Gk. δαίς, -ίδος	QPIE *th ₂ eid- : Lat. <i>taeda</i>
QPIE *drosd ^(h) - : PSlav. *drozdъ QPIE *drou ₂ d- : Arm. <i>artoyt</i>	QPIE *t(o/u)r(s)d ^(h) - : Lat. <i>turdus</i> QPIE *trosd ^(h) - : PCelt. *trozdi- QPIE *trosd- : PGm. <i>prastu-</i> QPIE *strosd ^(h) - : PBalt. *strozdo- QPIE *stroud ^h - : Gk. στρουθός

Table 3.6 Alternations between *d and *t

A further example of this alternation may be Lat. *citrus* ~ Gk. κέδρος, but devoicing of -dr- to -tr- is possibly regular in Latin. It would fit the pattern in which Latin attests to the unvoiced variant (but note Romance *darbo-).

3.2.1.1.1.2.2 Aspiration

QPIE *d ^h	QPIE *d
QPIE *stroud ^h - : Gk. στρουθός [QPIE *t(o/u)r(s)d ^(h) - : Lat. <i>turdus</i>] [QPIE *trosd ^(h) - : PCelt. *trozdi-] [QPIE *strosd ^(h) - : PBalt. *strozdo-] [QPIE *drosd ^(h) - : PSlav. *drozdъ]	QPIE *trosd- : PGm. <i>prastu-</i> QPIE *drou ₂ d- : Arm. <i>artoyt</i> [QPIE *t(o/u)r(s)d ^(h) - : Lat. <i>turdus</i>] [QPIE *trosd ^(h) - : PCelt. *trozdi-] [QPIE *strosd ^(h) - : PBalt. *strozdo-] [QPIE *drosd ^(h) - : PSlav. *drozdъ]

Table 3.7 Alternations between *d^h and *d

For this lexeme, the alternation between *d^h and *d (at the end of the root) is secured by Gk. στρουθός and PGm. *prastu-*, Arm. *artoyt*. Whether Lat. *turdus* reconstructs to *t(o/u)r(s)d^h- or *t(o/u)r(s)d- is unclear. The former is only possible if an intervening sibilant blocks the change *rd^h > rb.

3.2.1.1.1.2.3 *Voicing and Aspiration*

QPIE <i>*d^h</i>	QPIE <i>*t</i>
QPIE <i>*kh₂end^h-</i> : Gk. καθήλια	QPIE <i>*ka/Hnt-</i> : Lat. <i>cant(h)ērius</i>
QPIE <i>*l^hd^h-</i> : Gk. λάθυρος	QPIE <i>*l(e)nt-</i> : Lat. <i>lēns</i>
QPIE <i>*mind^h-</i> : Gk. μίνθη	QPIE <i>*m(e)nt-</i> : Lat. <i>menta</i>

Table 3.8 Alternations between **d^h* and **t*

Like with the **b^h ~ *p* alternations, Latin attests to the unvoiced, unaspirated variant. But unlike that category, where sometimes both variants were attested in Greek, the alternation here is more exclusive. Lat. *hasta* < QPIE **g^ha/Hst-* beside PGM. **gazda-* < QPIE **g^ha/o/Hzd^h-* belongs here as well. Celtic sound laws obscure the whether the dental of PCelt. **gazdo-* was borrowed as **d^h* or **d*. Such is also the case for Lat. *catulus* < QPIE **ka/Ht-* against Ml. *cadla* < **ka/Hd^(h)-* (where the Germanic forms could reflect **ka/o/Hd^h-* or **ka/o/Ht-* with Verner's Law).

The nature of the dental alternation between Lat. *raudus* and its comparanda is unclear. Its dental can reconstruct to **d^h* or **d*. If PCelt. **rutu-*, whose appurtenance is uncertain, is not compared, and if *raudus* was borrowed with **d* like PGM. **arut-*, then there is no alternation attested.

Theoretically, a **d^h ~ *t* alternation could exist between the comparanda of Lat. *fīcus* with QPIE **d^h*, where Gk. τῦκον, σῦκον could attest to **t_i/u-* or **d^h_i/u-* and Arm. *fuz* mechanically reconstructs to **t*. But it is more likely that these words were borrowed with **t^h* or **θ*.

3.2.1.1.1.3 *Interim Conclusion on Labials and Dentals*

So far, an interesting pattern emerges amongst the cases where the quality of the consonants can be verified (i.e. it has not been obscured due to sound laws). Firstly, alternations involving all possible combinations of quality are attested. While the category of **b^h ~ *b* is mixed (twice Latin reflects **b^h*, twice **b*), in each of the others, Latin patterns consistently. Between the categories, however, it is not consistent. For **b ~ *p* alternations, Latin reflects **b* but for **d ~ *t* alternations, it reflects **t*. The distributions of the attested comparanda show that these alternations are not the result of one monolithic contact situation; more on this follows in §4. However, even amongst words with a Mediterranean distribution, the pattern of Latin reflexes is difficult to reconcile with the two IE substrates proposed that might be expected to affect Latin (Ausonian and Pelasgian). If Pelasgian is responsible for **D > *T*, then we must assume for Lat. *buxus* ~ Gk. βύξος, the Greek has borrowed the Pelasgian reflex but for Lat. *cotōneum* ~ Gk. κυδώνια, Latin instead has the Pelasgian reflex. One wonders why it is always Latin that has the Pelasgian reflex when a dental is involved. Beyond this, very few of these cases can be etymologized to an IE root.

It is more likely that these alternations are not of Indo-European origin. While they may in part be due to dialectal differences within the substrate languages, the nativization of

foreign phonemes must certainly have played a large role in producing irregularity. Such is also the case for the velars (see below), where the additional parameter of labial-/palatalness creates even more possibilities for nativization.

3.2.1.1.2 Velars

The velar consonants as reconstructed for PIE do not only differ in aspiration and voicedness, but have the added (mutually exclusive) aspects of labial- and palatalization. These are traditionally given as **k, *k̑, *kʷ, *g, *g̑, *gʷ, *gʰ, *gʰh* and in the glottalic theory as e.g. **k:, *k̑:, *kʷ:, *'k, *'k̑, *'kʷ, *k, *k̑, *kʷ* (from Beekes 2011: 128-9, other presentations elsewhere as cited above). By the time of the separate daughter branches, when loanwords would be entering, centumization and satəmization would have been developing or would have already occurred. Thus from an inherited perspective, we should not expect to find a palatovelar reflex in a *satəm* language corresponding to a labiovelar reflex in a *centum* language. But a phenomenon like this cannot be ruled out *a priori* for loans from a non-IE language. This adds a layer of difficulty to the analysis. As with the labials and dentals, below are the data that show the distributions. I only include alternation between labial-/palatalness when there is explicit reason to do so.

3.2.1.1.2.1 Voicing

QPIE <i>*g</i>	QPIE <i>*k</i>
QPIE <i>*ga/HR-</i> : Prov. <i>garric</i>	QPIE <i>*kerr/so-</i> : Lat. <i>cerrus</i> QPIE <i>*ka/Hr-</i> : Ital. dial. <i>cariglio</i> QPIE <i>*ka/HR-</i> : Catal. <i>carrasca</i>
QPIE <i>*b^ha/ol-ig-</i> : PGM. <i>balikōn-</i>	QPIE <i>*b^hul-Vk-</i> : Lat. <i>fulica</i> QPIE <i>*b^(h)o/ul-a/oK-</i> : SGael. <i>bolachdan</i>
QPIE <i>*sur-(V)g-</i> : PGM. <i>*s(w)ur(V)ka-</i>	QPIE <i>*s(y)ōr-Vk-</i> : Lat. <i>sōrex</i> QPIE <i>*syo/ur-ak-</i> : Gk. ὕραξ
QPIE <i>*gruHm-</i> : Lat. <i>grūmus</i>	QPIE <i>*kroHm-</i> : Hsch. κρωμαξ QPIE <i>*kloHm-</i> : Gk. κλωμαξ

Table 3.9 Alternations between **g* and **k*

Lat. *corbis* belongs here if it is indeed related to PGM. **krebō-* < QPIE **gréb^h-ōn-*. For Lat. *grāmiaie*, there exists a **k ~ *g* alternation amongst the Slavic comparanda, but this must be the result of a post-Common Slavic (i.e. during the first millennium CE) borrowing into Slavic. Thus its bearing on earlier substrate features, at least in terms of this alternation, seems dubious.

3.2.1.1.2.2 Aspiration

Interestingly, there is one uncertain case of an aspiration alternation involving velars. This is the case of Lat. *ālium* ~ Gk. ἄλις, γέλις. PVerb. **agVlum-* of similar shape to the Latin form suggests that it once had a velar like Gk. ἄλις, and one explanation for its disappearance is that it was **g^h*, undergoing development to **h*.

3.2.1.1.2.3 *Voicing and Aspiration*

QPIE *g ^h	QPIE *k
QPIE *tu/ūg ^h - : Arm. <i>t^cuz</i>	QPIE *d ^h rk- : Lat. <i>fīcus</i> QPIE *d ^h /tj/ūk- : Gk. τῦκον, σῦκον (Hebr. <i>šiqmā</i>)
QPIE *Hurg ^h - : Gk. ὕρχη	QPIE *H(o)rk- : Lat. <i>orca</i>
QPIE *g ^h a/hz-l-ik- : Gk. χάλιξ	QPIE *ka/Hlk- : Lat. <i>calx</i>

Table 3.10 Alternations between *g^h and *k

In 1 case, it cannot be determined whether the alternation is *g ~ *k or *g^h ~ *k:

QPIE *g ^(h)	QPIE *k
QPIE *g ^(h) ra/ob ^h - : PSlav. *grabrъ-	QPIE *ka/Hrp- : Lat. <i>carpinus</i>

Table 3.11 Alternations between *g^h or *g and *k

In 2 cases, alternations between all three qualities are attested:

QPIE *g ^h	QPIE *g	QPIE *k
QPIE *blē/eh ^h - : Gk. βλήχων	QPIE *b ^h reg- : PGm. *brakna(n)-	QPIE *b ^h elik- : Lat. <i>felix, filix</i>
QPIE *HruG ^h - : PGm. *rugg-	QPIE *H/ura/Hg- : Lat. <i>raia</i>	QPIE *HreK- : PGm. *rehhōn-

Table 3.12 Alternations between *g^h, *g, and *k

Lat. *felix, filix* against PGm. *brakna(n)- shows the same alternation as between Lat. *fulica* ~ PGm. *balikōn- and Lat. *sōrex* ~ PGm. *s(w)urka- above; but the Greek comparanda of *felix* show the reflex of a voiced aspirate. For Lat. *raia*, Germanic shows two variants, both geminates.

3.2.1.1.2.4 *Palatalization*

Given that palatovelars are a class reconstructed for PIE, it is valid to question whether non-IE languages of Europe would have had such a feature. One case seems to indeed suggest that something akin to palatovelars was indeed present. Lat. *cucumis* matches Hsch. κύκων as if from *ku-ku-. Arm. *sex* could reconstruct to *k^hek^h- with an unvoiced aspirate not reconstructible for PIE and in opposition to the plain unvoiced consonant of the Latin and Greek forms. That its *s* is from something akin to *k^h is supported by Gk. σικύα, Hsch. σεκούα < QPIE *k^hiiek-. The Greek reflex of inherited *k^h is simply κ, thus this lemma was potentially borrowed with a palatal element that seems to have been interpreted in Armenian as a palatovelar. In one case, the *satəm* languages allow us to see that there was an alternation in palatalization. In the comparanda of *columba*, Arm. *salamb* attests to palatalized *k^hol- while OCS *golqbb* is from unpalatalized *gol-.⁴⁸⁹

Alb. *dallëndyshe*, in light of the velar of Gk. χελιδών (and Lat. *hirundō* < *g^h),

⁴⁸⁹ Such an alternation potentially also exists in the suffix of the hawk word (cf. *capyys*, whose circulation in Latin is suspect), where PSlav. *kobuzъ attests to *-uǰ^h- against *-ig- in Arm. կ'օւփիճ 'male hawk or falcon' (cf. Thorsø fthc.).

reconstructs to palatovelar *ǵ^h. Less straightforward are the cases of Lat. *excetra* and *avēna*. The former, taken at face value, reconstructs to a cluster *-ksk-. For the Baltic comparanda, an option exists to reconstruct this cluster with one palatovelar (*-k̑sk-). For the latter, a reconstruction of *aweCsnā is possible, but so is simply *awesnā. Its Baltic and Slavic comparanda reconstruct to an alternation between the reflexes of *k̑ and *ǵ^h, but all forms might have been borrowed with an affricate (Kroonen et al. 2022: 19-20) or (esp. on comparison with West Uralic *wešnā and PGM. *hab(a)zan-) a “spirant of indeterminate voicing” (Huld 1990: 404).

3.2.1.1.2.5 Labialization

Like with palatovelars, labiovelars are a distinctly PIE reconstructed feature. But five cases include comparanda that securely reconstruct to labiovelars. Lat. *fungus* and its comparanda attest to alternation between *g^{w(h)} and *ɥ (Latin and Greek: *(s)g^{w(h)}ong- ~ Germanic : *sɥong^{w(h)}-) that seems to suggest that the velar component of the original phoneme could be “overtaken” by the labial component. In Germanic, the final velar seems to have been labialized as well. A similar situation may occur in Lat. *malva* and its Greek comparanda where Gk. μάλβακα reconstructs to *malg^wak- against Gk. μάλᾱχη, μολόχη. The latter probably continue *malɥak^h-.⁴⁹⁰

Lat. *laurus* forms a Greco-Italic isogloss with Gk. δάφνη, δάϋνα. The most straightforward way of accounting for the φ ~ χ alternation is via a reconstruction with *g^{w(h)}, which also works for Latin. The vocalism of the Greek forms produces two further possibilities, neither of which allows the group to be of Indo-European origin. The first possibility is that there was an irregular *a ~ *au vocalic alternation that triggered the *boukolos* rule in Greek and Latin. QPIE *dag^{w(h)}nā- would yield Gk. δάφνη while *daug^{w(h)}na- would yield Gk. δάϋνα. QPIE *laug^{w(h)}ro- would yield PItal. *lauχro- > *laurus*. But given the possibility of a non-native labialized velar with a strong labial component, the vocalic alternation could be seen as the result of different interpretations of the placement of the labial element: *K^w vs. *^wK. In this way, *dag^{w(h)}nā- would yield Gk. δάφνη and *da^wg^hna- Gk. δάϋνα. A pre-form *la^wg^hro- would yield PItal. *lauχro- > *laurus*.

Go. *aqizi* requires a reconstruction with *g^w, which is not possible for Gk. ἄξιτι or (if the metathesis is not secondary, which I argue it is not) Lat. *ascia*, producing an alternation in labialization. Similarly, PGM. *hwerhwetjō- reconstructs to *k^w. In Lat. *cucurbita*, the internal velar could theoretically have been a labiovelar (unrounded before *u*), but an initial *k^wu- seems to have been deleted (cf. *ubi* < *k^wu-d^he-i). Thus this lexeme too points to an alternation in labialization.

⁴⁹⁰ The preform *malɥak^h- at first glance looks like it could also yield Lat. *malva*, but in order to not produce ***malla*, it would have to have been borrowed too late for *k^h > *h*. One solution to this is to have the form enter Latin as *malwaxa, and have the labialization be attracted to the *χ producing *malax^wa-. This would then give *malava and with later syncope the attested *malva*.

3.2.1.1.2.6 *Interim Conclusion on Velars*

For the labials and dentals, it was possible to see that all possible combinations of the 3 qualities are attested. This is much more difficult to say for the velars. Theoretically, given 9 reconstructible velars, there are a total of 502 possible combinations. This is an inflated estimate however. While a *g* in both a *centum* and *satəm* language could be hiding alternations between palatovelars and labiovelars ($*\acute{g} \sim *g$, $*g \sim *g^w$, $*\acute{g} \sim *g^w$), we find no examples of the opposite phenomenon—a visible *satəm* palatovelar ~ *centum* labiovelar alternation. A more realistic number of possible combinations can be approximated from a set consisting of $*k$, $*g$, $*g^h$, $*k^w$, $*\acute{g}^w$, $*\acute{g}^{wh}$ (in which labial-/palatalness is collapsed). And indeed, there are examples of alternations in palatalness and labialness in the languages that can preserve them. This is the most meaningful conclusion from the data: there is evidence for both palatalized and labialized phonemes in the substrate languages of Europe (cf. Beekes 2014: 4 on Pre-Greek).

3.2.1.1.3 *Conclusions on the QPIE Plosive Rows*

The stratificational power of alternations like this on their own is hampered by the fact that there is only a limited possibility for nativization of foreign sounds. That is, a phoneme that did not exist in Proto-Italic, upon being borrowed, would have to be mapped onto one of the reflexes of an existing PIE phoneme. Theoretically, two different foreign sounds could end up being mapped onto the same reflex, masking their originally separate origins. On the other hand, a singular sound in a foreign language can be borrowed with phonological variation even within the same borrowing language (cf. Meester fthc.).

As remarked on above, the sound laws proposed by IE Pelasgianists seem artificial, as when all the evidence of irregularities is taken together, there are more alternations than can be explained by discrete sets of chain shifts. While Latin exhibits tendencies (it reflects an unvoiced, unaspirated reflex in cases of $*b^h \sim *p$, $*d^h \sim *t$, $*d \sim *t$, $*g^h \sim *k$, and $*g \sim *k$ alternation) they are not always fully consistent, they are not all due to the same contact situations, and the variation outside of Latin is generally without a pattern (on this latter aspect, cf. Beekes 2014: 4, Šorgo 2020: 459). Beekes (2014: 4-5) uses this to conclude that (for Pre-Greek in his context), voice and aspiration were not distinctive features (cf. also Palmer 1963: 39, Furnée 1972: 115-200). What is at least clear from the alternations is that such features must not have worked the same way as in Indo-European. Labialization and palatalization do indeed seem to be phonetic features of one or more of the substrate languages of Europe, at least to the point where they could be perceived by speakers of PIE (cf. cases like Gk. σικῶα Lat. *laurus* ~ Gk. δάφνη, δάυχνα) or mapped onto the phonologized contrast between plain and palato-/labiovelars.

In general, describing alternations in terms of PIE phonology is also artificial, since the words in which they occur were borrowed at a post-PIE date and at various points in

time. A foreign /f/ borrowed into Proto-Italic at a time before the production of *f from the IE voiced aspirates would almost certainly have been mapped differently than an /f/ that entered after the development of Italic *f. The latter case, which presumably would have been preserved as Latin *f*, would nevertheless be mechanically reconstructed to PIE *b^h, obscuring the fact that it was actually borrowed as a fricative. QPIE reconstructions provide a useful shorthand for being able to show that lexemes are not inherited, but a more detailed distributional analysis (see §4) is needed to be able to identify distinct contact scenarios.

3.2.1.2 Alternations Beyond the Plosive Rows

A phonological justification must exist for considering that two different phonemes were diachronically underlyingly the same. Even with this caveat, alternations within the same place of articulation are not limited to the PIE plosive rows. In fact, there are even several cases of alternations between different places of articulation. Cases like the latter are not typologically unparalleled (cf. the allophonic variation of *k ~ t* in Hawai’ian).

3.2.1.2.1 Labial Plosive ~ Labial Nasal Alternation

The following alternations share a place of articulation (labial), differing in manner of articulation:

QPIE *b ^(h)	QPIE *m
QPIE *h ₂ erb ^(h) /d ^h - : Lat. <i>arbutus</i>	QPIE *h ₂ erm- : PRom. *armōn-

Table 3.13 Alternation between *b^(h) and *m

This alternation occurs between Latin and Ligurian dialects, suggesting that the source of the alternation was close to the Italian peninsula.⁴⁹¹ It is present within Greek, between Gk. κυβερνάω and Cypriot *ku-me-re-na-i*. Lat. *gubernāre* seems to be an indirect borrowing of κυβερνάω, and thus is not independent evidence of a *b^(h) ~ *m alternation. (The potential Baltic forms that suggest a further alternation with *mb^(h) are likewise too insecure to include as independent evidence.)

QPIE *p	QPIE *b (and *b ^h)	QPIE *m
QPIE *plo/ũNd ^h - : Lat. <i>plumbum</i>	QPIE *bolubdo- : Gk. βόλυβδος	QPIE *moliwdo- : Gk. μόλιβδος
QPIE *ple/oud ^(h) - : PCelt. *(ϕ)loudio-	[QPIE *b ^h liHwo- : PGm. *blīwa-] (Basque <i>berun</i>) (PBerb. *βaldūn etc.)	[QPIE *mliHwo- : PGm. *blīwa-]

Table 3.14 Alternation between *p, *b (*b^h), and *m

Germanic *blīwa- reconstructs to QPIE *b^h, but it can also be the reflex of *ml, like in many of the Greek forms. A QPIE *b is attested by Greek variants with β (and is

⁴⁹¹ Cf. the “alternanza mediterranea” of *m* and *b* listed by e.g. Alessio (1946a: 154). Bertoldi (1933b) finds a widely distributed (Sardinia, Iberia, Etruria, Gaul) lexeme “bush” (represented by Gk. βάτος and μανρία) with a *b ~ m* alternation.

probably the source of the Basque and Berber forms).

The $*b^{(h)}$ of Lat. *sabulum* alternates with $*m$ in Gk. ψάμαθος and PGm. $*sammada-$. Arm. *awaz* can be reconstructed to $*sab^h-$ or $*sap-$; thus it is uncertain if this represents a $*b^{(h)} \sim *m$ or a $*p \sim *b^{(h)} \sim *m$ alternation.

QPIE $*b^{(h)}$	QPIE $*mb^{(h)}$
QPIE $*la/Hb^{(h)}/d^h/sr-$: Lat. <i>lābrusca</i>	QPIE $*la/Hmb^{(h)}/d^h/sr-$: It. <i>lambrusca</i>
QPIE $*sa/Hb^{(h)}$ - : Lat. <i>sabūcus</i>	QPIE $*sa/Hmb^{(h)}$ - : Lat. <i>sambūcus</i>
QPIE $*gul-ub^h-$: PGm. $*kulubrōn-$	QPIE $*ke/ol-o/umb^{(h)}$ - : Lat. <i>columba</i> QPIE $*g^{(h)ol-omb^{(h)}}$ - : OCS <i>golqъbъ</i> QPIE $*kol-(o)mb^h-$: Arm. <i>salamb</i> (Copt. бροомπε)

Table 3.15 Alternations between $*b^{(h)}$ and $*mb^{(h)}$

The case of Lat. *columba*, in which the nasal is lacking in Germanic,⁴⁹² has widespread comparanda.

Otherwise the cases of $*b^{(h)} \sim *mb^{(h)}$ alternation are restricted to Latin and Romance. The case of *lābrusca* is more uncertain, since its *b* can be reconstructed in several ways. While there are numerous cases of etymological nasals being dropped before consonants due to their weak pronunciation (cf. Väänänen 1981: 63), nasal epenthesis is a more complex phenomenon in Romance. Cases often cited included *sa(m)būcus*, *la(m)brusca*, *stra(m)bus* ‘squinting, crooked (of eyes)’, and *sambatium/sabbatum* ‘sabbath’. The latter are clear loans from Greek (with some evidence for the form with *m* existing already in Greek), while the former two have no good etymology. Nor do the former fit into many of the more easily explained cases of Romance nasal epenthesis (leveling of the nasal infix in verbs, anticipation of an upcoming nasal, blends like $*rendō < reddō$ modelled on *prēndō*, hypercorrection based on cases of restitution of lost nasals, cf. Malkiel 1984⁴⁹³). This seems to be indicative of non-native origin. In Greek, original voiced stops were fricativized, but new voiced stops appeared when this process was blocked by a preceding nasal or when a voiceless stop was voiced by a preceding nasal (Holton et al. 2020: 114). Thus voiced stops appeared only after nasals, and loans of voiced stops subsequently appear as μπ, νδ, and γγ. If the Latin voiced stops were fricatives (cf. Kortlandt 2007: 150-1), then perhaps the cases of the appearance of a nasal was due to the borrowing of a voiced stop. But the only concrete examples in the dataset involve *b*.⁴⁹⁴

⁴⁹² The reconstruction $*kulumfrōn-$ seems to be ruled out. Given the restriction of epenthesis to Northumbrian, it is more likely that *culfre* is the result of syncope from *culufre*. Such syncope is not expected in a vowel that is the result of $*ium$ (cf. Jakob fthc.).

⁴⁹³ Malkiel further includes “phonosymbolic” i.e. expressive cases (including *strambus* for *strabus*) and notes cases of *mb* for both original *b* and *m*, the latter mainly being a feature of Sardinian (see also Wagner 1941: 223-4).

⁴⁹⁴ Though cf. potentially Lat. *lēns* < QPIE $*l̥ti-$ ~ Gk. λάθυρος < QPIE $*l̥d^h-$. We might consider it

Lat. *plumbum* attests to an *mb*, where the nasal is difficult to account for in comparison to the Celtic, Germanic, and Greek comparanda. On the one hand, it could be the result of specifically Italic nasal insertion: since an argument can be made that the *b* is from **d^h* and since a nasal would almost certainly have blocked the change **d^h > b* unless we reconstruct an otherwise unparalleled **plo/uNd^h-uo-*, the nasal was introduced after the Italic development of the voiced aspirates. On the other hand, Greek variants like μόλυβος and βόλιβος attest to a *b* in the same position as in Lat. *plumbum*. If the *b* is original in Latin, then so too might be the *m*, cf. the pre-nasalization in the substrate vocabulary of Germanic (Kuiper 1995: 68-72, Šorgo 2020: 459-60) and Greek (Beekes 2014: 14-15).

3.2.1.2.2 Labial Plosive ~ Labial Approximant Alternation

Beekes' (2014: 15) discussion of nasalization in Pre-Greek included alternations between a labial stop and *w*.⁴⁹⁵ Furnée (1972: 230-1) noted examples beyond Greek, and Šorgo (2020: 460-1), who considers it a separate phenomenon from nasalization, suggests it represents an attempt to render a foreign phoneme like [β] or [v]. In cases involving Italic, Latin attests to both variants.

QPIE <i>*b^h</i>	QPIE <i>*b</i>	QPIE <i>*μ</i>
QPIE <i>*b^(h)a/ob^h-</i> : PBSI. <i>*ba/ob-</i>	QPIE <i>*b^hab-</i> : Fal. <i>haba</i>	QPIE <i>*ba/ou-n-</i> : PGm. <i>*baunō-</i> (PBerb. <i>*ā-b/βāw</i>)

Table 3.16 Alternation between **b^h*, **b*, and **μ*

In the case of Lat. *faba*, Fal. *haba* points to **b* while Balto-Slavic requires **b^h*.

QPIE <i>*b</i>	QPIE <i>*μ</i>
QPIE <i>*h₁orh₃b-</i> : Gk. ὄροβος	QPIE <i>*h₁er(H/V)μ-</i> : Lat. <i>ervum</i>
QPIE <i>*h₁erh₁b-</i> : Gk. ἐρέβινθος	QPIE <i>*h₁orμ-</i> : PGm. <i>*arwīt-</i>
QPIE <i>*h₁orVb-</i> : Arm. <i>arowoyt</i>	

Table 3.17 Alternation between **b* and **μ*

Beyond these two secure cases with a relatively wide distribution are other more complex examples. While Hubschmid (1953: 63, fn. 2) purported to find no cases of such an alternation with a Mediterranean distribution, Lat. *bolunda* ~ Gk. ὄλυνθος seems to be one such example (with the assumption of original Gk. **Fόλυνθος*). If Lat. *bāca* is related to *vaccīnium*, they too hint at a **b ~ *w* alternation.

unlikely that the non-IE source language would have a phoneme akin to a PIE syllabic resonant, in which case the nasal of the Latin form might be explained as intrusive, albeit before unvoiced *t*.

⁴⁹⁵ There may be one case of an alternation like this within Italic. Lat. *Mārs* is from a form with **μ*, preserved in *Māvors*. But the Sabellic languages and the *Lapis Satricanus* attest to a stem *mamart-*. Thus **māmart-* has somehow changed to something like **māwart-*, an otherwise unparalleled change that suggests it might not be native to Italic (cf. DV 366).

Another complex case occurs between PGm. **blīwa-* and Greek variants μόλιβος, μόλυβος, βόλιβος. The variant with **w* seems also to appear in PCelt. **ϕloudio-*. The analysis of Lat. *plumbum* is complicated. If its *b* is from **d^h*, then perhaps its *u* participates in the **b ~ *w* alternation. If however its *b* is original (and it lacks the dental suffix of e.g. μόλυβδος), then it patterns with the Greek forms with *b*.

3.2.1.2.3 *L ~ R Alternations*

Amongst the liquids, laterals and rhotics are articulated in much of the same space. The two classes, even within Italic and its descendants, are prone to metathesis (cf. Spanish *milagro* ‘miracle’ < Lat. *mīrāculum*, *palabra* ‘word, speech’ < Lat. *parabola*, *peligro* ‘danger’ < Lat. *perīculum*, cf. e.g. Straka 1979: 400-22, Schmid 2016: 481) and dissimilation (cf. *-ālis* > *-āris* when attached to a base containing *l* like *vulgāris*, Väinänen 1981: 70, Weiss 2020: 168). Their alternation can easily come about in contact situations: several languages have one underlyingly liquid phoneme that surfaces as rhotic or lateral depending on the environment or is used in free variation (Korean and Japanese for example, cf. Ladefoged and Maddieson 1996: 182, 243).

Several cases exist in which Latin words and their comparanda attest to *l ~ r* alternations. I present them in three groups. In Group A, the alternation exists between Latin and Greek. In Group B, Latin and Greek agree against an alternation in other comparanda. In Group C, there is no Greek comparandum.

QPIE <i>*l</i>	QPIE <i>*r</i>
QPIE <i>*kh₂end^heHL-</i> : Gk. κανθήλια	QPIE <i>*ka/HnteHr-</i> : Lat. <i>canthērius</i>
QPIE <i>*kloHm-</i> : Gk. κλωμαξ QPIE <i>*kroHm-</i> : Hsch. κρωμαξ	QPIE <i>*gruHm-</i> : Lat. <i>grūmus</i>
QPIE <i>*g^heliHd-</i> : Gk. χελιδών QPIE <i>*g^ho(u)l(H)-(o)nt/d^(h)-</i> : Alb. <i>dallëndyshe</i>	QPIE <i>*g^hiro/und^(h)-</i> : Lat. <i>hirundō</i>
QPIE <i>*silb^h-</i> : Gk. σίλφρον QPIE <i>*selp-</i> : Hsch. σέλπον (Berber <i>azlaf</i> , <i>aselbu</i> , etc.)	QPIE <i>*sirp-</i> : Lat. <i>sirpe</i>
QPIE <i>*g_lmo-</i> : Gk. γλάμων	QPIE <i>*greHm-</i> : Lat. <i>grāmia</i> QPIE <i>*g/krHm-</i> : PSlav. <i>*k/gr_ɛm-</i>
QPIE <i>*leili-</i> : Lat. <i>līlium</i> QPIE <i>*Hol-</i> : Hitt. <i>alēl</i> (Copt. <i>hlēli</i>)	QPIE <i>*leiri-</i> : Gk. λείριον (Copt. <i>hrēri</i>)

Table 3.18 Group A alternations between **l* and **r*

For Lat. *grūmus*, the **l ~ *r* alternation exists within Greek, such that it could theoretically belong to Group B.

QPIE * <i>l</i>	QPIE * <i>r</i>
QPIE * <i>b^helVk-</i> : Lat. <i>felix</i> QPIE * <i>bleHg^h-</i> : Gk. βλιγγρον, βλιγγρον	QPIE * <i>b^hreg-n-</i> : Dan. <i>bregne</i>
QPIE * <i>ph₂eil-a/ek-</i> : Lat. <i>paelex</i> QPIE * <i>pa/HL-ak-</i> : Gk. παλλακή	QPIE * <i>pa/erik-</i> : OIr. <i>airech</i>

Table 3.19 Group B alternations between **l* and **r*

QPIE * <i>l</i>	QPIE * <i>r</i>
QPIE * <i>ta/Hlp-</i> : Lat. <i>talpa</i>	QPIE * <i>da/Hrb^(h)-</i> : PRom. * <i>darbo-</i>

Table 3.20 Group C alternation between **l* and **r*

Given that the appurtenance of Gk. κόλυμβος to Lat. *columba* is uncertain for semantic reasons, the group, for which all forms beside Copt. ⲩⲣⲟⲟⲙⲡⲎ attest to the variant with **l*, most likely fits into Group C. For the stratificational power of this alternation, see §4.2.2.4.1.

3.2.1.2.4 *N ~ M Alternation*

There is one lexeme that seems to show a nasal alternation of *n ~ m*: Lat. *laena* vs. Gk. χλαῖνα, χλανίς, χλαμύς. As explained by Rosol (2013: 107-9), the alternation likely has its source in Semitic: Hebr. *glōm* against Late Babylonian *gulēnu*. Thus the alternation is likely not a result of the borrowing process into Latin/Greek but instead attests to this word being borrowed more than once.

3.2.1.2.5 *L ~ D Alternation*

A lateral can develop into a dental or *vice versa* because of an overlap in the place of articulation (cf. Sardinian, Sicilian, Calabrian retroflex *ɖɖ < ll*, cf. NavigAIS).

There exists within Latin a phenomenon referred to as the “Sabine *l*”, in which Latin attests to *l* as the reflex of inherited **d* in some words. The phenomenon was attributed to Sabine by Conway (1893) via dubious methodology. No ancient source attributes the phenomenon to Sabine, but the idea was followed by Petr (1899) who found 17 Sabinicisms and others like Schrijnen (1914) who argued on the basis of historico-political and -social evidence. Bottiglioni (1943: 316-17) shows that the words we have as purportedly Sabine prove that it was not responsible for this change. His best example is a passage of Varro, where the deities called *Novensiles* by Livy are said to be called *Novensides* by the Sabines. Weiss (2020: 504 fn. 63) adds *fedus = haedus* ‘goat’ and *īdūs*, both purported to be Sabine, to show that Sabine allowed *-d-*.

Nor is the phenomenon itself well understood. There is a short list of generally accepted cases, some problematic, and some other unclear cases. The best cases are *odor* ‘smell’ < **h₃ed-* ~ *olere* ‘to smell’, *sedere* ‘to sit’ < **sed-* and *solium* ‘chair, throne; bathtub; sarcophagus’. Lat. *lacrima* ‘tear’ might be a loan from Gk. δάκρυμα ‘tear’ or might represent an inherited formation. Lat. *lēvir/laevir* ‘husband’s brother’, attested in glosses,

could have *-vir* for *-ver* from contamination with *vir* ‘man’ and thus descend from **dehzi-uer-* (cf. Gk. δᾶήρ ‘husband’s brother’, etc.; DV 336). Lat. *lingua* ‘tongue’ seems to be attested also as *dīngua*; in any case it descends from **dnǵ^h-uh₂* ‘tongue’, but its *l* might be from contamination with *lingere* ‘to lick’ (DV 343, Weiss 2020: 504). Thus it cannot be adduced as an example of the phenomenon with certainty. Lat. *lautia* ‘entertainment provided for guests’ is generally taken as the development of *dautia*, a hapax in Paulus *ex Festo* (cf. DV 161). It would be from **douH-ó-* ‘giving, bestowing’, with Thurneysen-Havet’s Law (Vine 2006: 238). However, since *lautia* always occurs in the collocation *locus lautiaque* or *loca lautia*, it is sometimes suspected that the *d > l* is due to alliteration (cf. DV 161). Prósper (2019: 463) even proposes that *dautia* is an artificial archaism. In any case, the status of *lautia* as an example of the phenomenon is uncertain.

That the phenomenon of the “Sabine” *l* occurs only in inherited lexemes could theoretically be due to the fact that it is only visible in cases with secure comparanda. Lat. *simila* is, along with Gk. σεμίδαλις, most likely originally a loan from Semitic (cf. Aram. *samīdā*, Akk. *samīdu*). Its *l* for *d* could theoretically be due to the “Sabine” *l* affecting loanwords. Lat. *laurus* has been borrowed from the same source as Gk. δάφνη, (Thess.) δάφνη. Its *l* for *d* may also be due to the phenomenon at hand, but Hsch. λάφνη, said to be in currency in Pergamon, suggests that the alternation here had a wider distribution. Lat. *adepts* occurred alongside PRom. **ala/ep-* but in this case its relationship to Gk. ἄλιφα(ρ), from which it cannot be a direct borrowing, rules out the “Sabine” *l*; the change is in the opposite direction. In this latter case, it should be mentioned that Lat. *clueō* ‘to be known’ with inherited **l* occurs in South Picene as **kduúú** [1sg.pres.] with <d>. Thus perhaps it could have been involved in some **l > d* changes.

3.2.1.2.6 *S ~ D Alternation*

There is one instance of Lat. *s* for a *d* in all other comparanda, namely the case of Lat. *rosa*. As discussed (s.v. *rosa*), not even recourse to the Umbrian change of intervocalic **d > ř*, *rs* provides a convincing explanation. Biville (I: 257-6) suggests that the same alternation between a dental and a sibilant might be found between Lat. *rēsīna* ‘resin’ and Gk. ῥητίνη ‘resin’, but that it is more likely that it is borrowed from a Greek by-form **ῥησίνα* (cf. also EDG 1284 who only accepts this latter possibility). Biville also mentions Lat. *asinus* ~ Hebr. *‘ātōn* ‘donkey’. It is possible that they are connected, and an alternation of this sort might be behind both lemmata. The route of transmission of the two words must have been different however; Gk. ὄνος, if related to *asinus*, can only be explained via the erstwhile presence of a sibilant whereas in ῥόδον it has the dental.

3.2.1.2.7 *D ~ K ~ Ø Alternation*

The comparanda of Lat. *nux* each end in a different consonant: Lat. *nux* < QPIE **(k)nuk-* ~ PGM. **hnut-* < QPIE **knud-* ~ PCelt. **knū-*, **knowes-* < QPIE **knu(H)-*, **kn(e/o)u(H)-*. Whether the Celtic comparanda attest to a laryngeal is obscured by

regular developments within Irish and British Celtic. The phenomenon of a dental-velar-zero alternation is found by van Sluis (fthc.) in two other lemmata, one of which has a secure Latin comparandum. But in both cases, the dental is QPIE **t*. These are the ‘head’ word Lat. *caput* < PQIE **ka/Hput-* ~ W *cawg* < QPIE **ka/Hpuk-* (~ OE *hafola* < QPIE **ka/o/Hp-ulo-* or **ka/o/Hput-lo-*) and the ‘bee’ word Lith. *bitė* < QPIE **b^{hi}t-* ~ OIr. *bech* < QPIE **b^{hek-}* ~ OE *beó* < QPIE **b^{hi}-on-*. If these all represent the same phenomenon, then there is no need to reconstruct a laryngeal for the Celtic comparanda of Lat. *nux*. Kroonen (2012a: 248) suggested that the pattern is the result of the nativization of a substratal glottal stop, which could be the case whether or not a laryngeal is reconstructed. Otherwise, the alternation might be due to existing dialectal or paradigmatic differences in the source languages. In North Saami, for example, stem-final *-*g* regularly (albeit rarely) becomes *-t* in absolute final position (Aikio and Ylikoski 2010 course handout, page 37⁴⁹⁶).

The **k* ~ **t* alternation within Lat. *caput* and its comparanda have often been explained as belonging to a suffixal element (cf. Boutkan 1998: 111, Schrijver 1997: 295). It cannot be ruled out that the vacillating stop of the ‘bee’ word and the ‘nut’ word are suffixes as well, but they certainly look like part of the root. If Skt. *kapāla-* ‘cup, jar, dish; skull’ is only coincidentally similar to the rest of the comparanda for the ‘head’ word, then all three lemmata share a relatively similar distribution. Beyond these three “canonical” examples are a few potential others. Anthony Jakob (fthc. thesis) notes the phenomenon potentially between ON *síld* and Lith. *siĺkė* (OPr. *sylecke*, > Fin. *silakka*) ‘herring’. If a laryngeal truly is involved, perhaps the alternation is found between **natr-ik-* (Lat. *natrix* ‘water-snake’, OIr. *nathir* ‘adder, snake’) ~ **natr-iH-* (W *neidr* ‘snake’) ~ **natr-* (Co. *nader*, etc.; Go. *nadre* [gen.pl.], etc. ‘adder’, cf. with a long vowel OHG *nātara*, etc. ‘adder’), despite usually being etymologized as a derivation from **sneh₂-* ‘to swim’.⁴⁹⁷ Finally, Lat. *pix* ‘pitch, tar’ and *picea* ‘pine/spruce’ < **pik-* might show a *k* ~ *t* alternation with Gk. *πίτυς* ‘pine’ < **pit-*. Lat. *pīnus* could be from either **pik-s-no-* or **pit-s-no-*. This would however introduce a Greek comparandum into the phenomenon, which so far has importantly excluded Greek. On the importance of the lack of any Greek comparanda to the stratification of this alternation, see §4.3.2.3.1.

3.2.1.2.8 Further Irregular Involvement of a Sibilant

3.2.1.2.8.1 *S mobile*

Some Indo-European roots appear with and without an initial *s*, even within the same branch (cf. Skt. *páśyati* ‘sees’, *spāśáyate* ‘makes seen’), a phenomenon called *s mobile* (cf. Forston 2010: 76-7). It is poorly understood in inherited roots, but may have something to do with re-bracketing of endings. It is *a priori* strange to be able to find

⁴⁹⁶ Available online: https://www.academia.edu/36836577/The_Structure_of_North_Saami

⁴⁹⁷ Stifter (fth.) alternatively suggests that this could be an example of the non-native suffix spreading to an inherited base. At the same time, this smacks of the sort of laryngeal hardening proposed by Martinet (1955).

such a phenomenon occurring in non-inherited material unless it occurred independently in the daughter branches and could thus occur after loanwords were taken up. This could only be the case if it was a very trivial phenomenon, i.e. that re-bracketing was a frequent tendency for roots with certain onsets. This question requires further investigation, but there seems to be one good example of something akin to *s* mobile in a lexeme of non-IE origin.

The lexeme in question is Lat. *turdus* and its widespread comparanda. An initial **s* is not present in Italic, Celtic, Germanic, Slavic, or Armenian. But it does appear in Baltic **storzdo-* and Gk. στρούθος.

For two other lexemes, it is questionable whether the forms with initial **s* really belong with those that lack it. To Lat. *calix* ~ Gk. κύλιξ might belong U *skalçeta* and Hsch. σκάλλιον, but these may be coincidentally similar.⁴⁹⁸ Lat. *carpinus* is semantically best matched by PSlav. **grabrǫ*. It has been compared to OPr. *skerptus* and Lith. *skī̃pstas* ‘hornbeam, elm, alder buckthorn, honeysuckle, beech’. But given the broad semantics of the Baltic forms with initial **s*, they too may be only coincidentally similar.

Finally, twice in the dataset Latin has *f-* for Gk. σφ(ι/π)-: Lat. *fungus* ~ Gk. σπ/φόγγος and Lat. *funda* ~ Gk. σφενδώνη.⁴⁹⁹ These are not inherited from PIE and they cannot be direct loans from Greek (in light of Lat. *spinter* ‘bracelet’ < Gk. σφιγκτήρ ‘something that binds’ and Lat. *spaeritam* ‘dumpling’ < a derivation of σφαῖρα ‘sphere, ball’, attested since Plautus and Cato respectively, cf. Hiersche 1964: 229-31; Biville I: 152-7, 197-9). It is unclear if these forms entered Latin without initial **s*. It is present in all secure⁵⁰⁰ comparanda. Initial *sf* is not an allowed cluster in (Classical) Latin; thus perhaps a foreign sequence borrowed as something akin to **sb^h* or **sg^{wh}* would have lost its **s* along with the development of the voiced aspirates to attested *f*.

3.2.1.2.8.2 *S Insertion*

In several cases, comparanda attest to the vacillating presence of a sibilant before a consonant word-internally. In the two most secure cases, the phonetic environment is such that it cannot be determined if the Latin word was borrowed with or without the sibilant. Such is the case for Lat. *turdus*, which could continue PItal. **to/urzdō-* or **to/urdo-*. Pre-forms with an internal sibilant are reconstructible for Celtic,⁵⁰¹ Slavic, Germanic, and Baltic but not for Greek or Armenian. The second *b* of Lat. *barba*, in light of comparanda, reconstructs to **d^h*. It is not clear if the sequence **-rsd^h-* would also yield *-rb-*, but in any case a sibilant cluster is present in Lith. *barzdà*, Latv. *bārzda*

⁴⁹⁸ Cf. potentially OHG *scāla* ‘drinking bowl’ etc. (Untermann 2000: 684) < PGm. **skēlō-* (Philippa et al. 2003-9 s.v. *schaal* 1). The Germanic verb is from **(s)kel-* ‘to cut’, but LIV2 also gives **skelH-* ‘to slit open, split’. U *skalçeta* < PItal. **skalik-* could continue **sklH-ik-*.

⁴⁹⁹ A third case I have not mentioned is that of Lat. *fidēs* ‘stringed instrument, gutstring’ ~ Hsch. σφίδες χορδαὶ μαγειρικαὶ ‘intestines for cooking’, σφίδη χορδὴ ‘intestines’. While Gk. χορδὴ can also mean gutstring (thus matching Lat. *fidēs*), in light of the other gloss, it more likely means ‘sausage’.

⁵⁰⁰ Not the insecure Slavic comparanda of *fungus*.

⁵⁰¹ The Brythonic forms cannot have an internal sibilant, but they may represent a loan from Goidelic.

‘beard’ whereas the sibilant is absent in PSlav. **bordā*.

The case of Lat. *fracēs* is similar in many ways (German and Baltic reconstruct to **d^hra/og^h*- but Slavic to **d^(h)rasg^h*-), but Alb. *dra* attests to a slightly different pattern. The velar should probably not be lost in Albanian, but **s* can be lost, at least before a back vowel (cf. Schumacher & Matzinger 2013: 262-3); thus the best reconstruction for Alb. *dra* is **d^(h)ras-*. It is unclear if the **s* is the same as the sigmatic element of the Slavic form (but lacking the velar), or if it should instead be interpreted as an irregular correspondent of the velar.

A similar alternation seems to occur between Lat. *fascis* ~ Gk. φάσκιος ~ Hsch. βάσκιος on the one hand and Gk. φάκελος on the other. Given this latter form without the sibilant, it is tempting to compare PGM. **pakk-*, **bagg-* (and thus Lat. *baiulus*). Alb. *bāshkē* also has the variant with the sibilant while Celtic comparanda can be reconstructed with or without it. Finally, Sörgo (2020: 459) calls attention to the pair Lat. *aesculus* ~ PGM. **aik-*. I have included it under §2.3.2 *Uncertain Comparanda*, but if the two words are related, then they would indeed exhibit the same pattern shown here.

3.2.1.2.8.3 SK Metathesis

There are several cases of irregular metathesis involving a sibilant and velar.

QPIE * <i>sk</i>	QPIE * <i>ks</i>
QPIE * <i>h₂esk-</i> : Lat. <i>ascia</i>	QPIE * <i>h₂eg^(h)/ks-</i> : Gk. ἀξίνη QPIE * <i>h₂eg^wes-</i> : PGM. * <i>akwesi-</i>
QPIE * <i>mus(g^(h)/k)lo-</i> : Lat. <i>mūlus, muscellus</i> QPIE * <i>musku-</i> : PSlav. * <i>mъskъ</i> QPIE * <i>musk-lo-</i> : Hsch. μύσκλοι	QPIE * <i>mug^(h)/k(s)lo-</i> : Hsch. μυγλός
QPIE * <i>wisko-</i> : Lat. <i>viscum</i>	QPIE * <i>wikso-</i> : Gk. ἰξός QPIE * <i>weiks-</i> : PGM. * <i>wīhsilō-</i> QPIE * <i>wel(k)si-</i> : PSlav. * <i>višb</i>

Table 3.21 Alternations in the order of sibilant and velar in clusters

Otrębski (1939: 133) lists 10 cases of (simple) *s*-metathesis,⁵⁰² 6 of which are still generally compared today. His list includes Lat. *ascia* and *viscum*. Additionally he gives 1) PGM. **fahsa-* ‘hair of the head’, Gk. πέκος ‘fleece’ ~ Gk. πέσκος ‘skin, rind’ (perhaps **po^k-so-*, **pek-es-* ~ **pek-sk-o-*, though EDG 1180 is not fully certain), 2) PBSI. **woško-* ‘wax’ ~ PGM. **wahsa-* ‘wax’ (perhaps through metathesis or dissimilations of clusters, cf. Kroonen 2013: 566, or a borrowing from a substrate language, cf. Philippa et al. 2003-9 s.v. *was*), 3) Lat. *vespa* ‘wasp’ ~ PCelt. **waxs-V-*, PBSI. **wóps(w)a?*, Germanic (cf. OHG *wafsa*, OE *wæsp*, etc.), perhaps PIr. **wabža-* ‘wasp’ (potentially from PIE *(*h*₁)*uebh-* ‘to weave’), and 4) PGM. **aspō-*, **apsō-* ‘aspen’ ~ PBSI. **aps-* ‘aspen’ (also in Turkic languages and Finnic, considered non-IE by i.e. Kroonen 2013:

⁵⁰² Note however that the goal of his 1939 *Indogermanische Forschungen* was to explain these as regular.

39).

A further example might include Lat. *tamarix*, some of whose Romance descendants attest to **tamarisk-*. Šorgo (2020: 459) notes the Celto-Germanic isogloss (potentially a substrate borrowing, cf. Kroonen, van Sluis & Jörgensen 2023: 212) PGM. **pahsu-* ~ PCelt. **tazgo-*, **tasko-* ‘badger’. The phenomenon thus seems to be quite limited, with the only case that demands an inherited origin being that of the wasp word given its widespread distribution in combination with a reasonable IE root etymology.

This phenomenon might represent different attempts at simplifying a foreign complex cluster. For example, beside Hsch. μύσκιλοι and μυγλόος, there is also Gk. μύκλος, lacking the sibilant entirely. A more exemplary case is that of Lat. *excetra*, with the cluster **-ksk-*. Its Baltic comparanda reconstruct to a cluster **-k̑sk-*, while the Slavic forms seem at most to reflect **-ks-*. Theoretically, **ks-* could also be behind Germanic comparanda, in which the cluster has been simplified further to PGM. **stur-*.

3.2.1.2.9 *Gemination*

Indo-European roots are not reconstructed with geminate consonants (except for nursery words like **atta-*), and gemination that should have occurred via morphological processes was reduced or interrupted (cf. 2sg. **h₁esi* ‘you are’ for expected ***h₁es-si*, nom/acc dual **h₂usih₁* ‘two ears’ for expected ***h₂uss-ih₁*, and the insertion of **s* in **TT* clusters; examples from Ringe 2006: 18). Various processes occurred in the individual daughter languages that produced gemination, often the assimilation of clusters produced by the addition of derivational morphology. In Latin, the most important source of gemination that is not the result of assimilation is the *littera* (or *Iuppiter*) rule, in which a long vowel followed by a voiceless stop resulted in either a long vowel plus single stop or a short vowel plus a geminate (thus *lītera* vs. *littera* and *Iūpiter* vs. *Iuppiter*). Weiss (2010b, 2020: 155) suggests that the long vowel must be of diphthongal origin and the consonant involved must be a voiceless stop. A similar phenomenon occurs for *r* after *a* (the best example being *narrō* ‘to say, tell’). There are thus some cases of Latin geminates for which the *littera* rule may not provide an adequate explanation and a non-native origin can be considered.

The opposite situation has occurred in the research on the substrate of Germanic. Kuiper’s (1995: 68-72) A2 layer is labelled the “language of the geminates” by Schrijver (2001: 420-1) based on its most peculiar feature. It is credited as the source for much of the non-IE vocabulary in Germanic (Boutkan 1998, Boutkan & Siebinga 2005: xvi-xvii). Kroonen (2009: 60-2) however importantly noted that this feature was overused. Many of the geminate forms occur in the *n*-stems and iteratives, such that they can easily be explained by assimilation of **n* to the preceding consonant (Kluge’s Law). Thus the importance of geminates in Germanic for the study of the pre-IE substrate had been overestimated.

An explanation that has often been offered to explain the presence of geminates is that

they are expressive. In certain semantic categories, including onomatopoeias, this seems plausible. But as Kroonen (2009: 59) notes (in defense of Kluge’s Law, but the argument transfers), “it is *a priori* implausible that a completely new range of phonemes (i.e. geminates) could be introduced into a linguistic system by extra-linguistic factors such as charged semantics,” comparing this to the outdated idea of spontaneous generation in biology.

Taking these considerations into account, when the appearance of a geminate in a lexeme can be excluded as a regular development, it may be indicative of a non-native origin for the lexeme in which it occurs. We should keep in mind that this does not automatically presuppose that the source language had geminates however. Once geminates arose via sound changes in the daughter languages, a geminate consonant became a native phoneme onto which a foreign sound could be mapped. There are other phonological properties that the original sounds could have had that resulted in their appearance as geminates. Kroonen (2009: 62) even adds: “One could even speculate, for instance, Kluge’s Law was triggered by the absorption of speakers of this substrate language [that had long stops] into the PIE dialect that ultimately became known as Germanic.” But to assume that geminate stops appeared as features of the daughter languages because they borrowed words from a substrate language that contained them is perilously close to the *ex nihilo* argument that Kroonen warned against. In any case, the indication of non-native origin that non-*lautgesetzlich* geminates provide is strengthened, as with all the categories above, when the geminate is not consistently present (i.e. it alternates) between comparanda and especially when it occurs in a lexeme that attests to other irregular alternations.

Singleton	Geminate
QPIE *ba/Hk- : Lat. <i>baculum</i> QPIE *ba/h ₂ k- : Gk. βάκτρον QPIE *ba/o/Hk’ / *ba/o/Hg ^h - : PGm. *pagjō-	QPIE *ba/HK- : PRom. *bakkillo- QPIE *ba/HK- : PCelt. *bakko-
QPIE *beh ₂ k- : Lat. <i>bāca</i> QPIE *bh ₂ k- : PCelt. *bak-	QPIE *beh ₂ K- : It. <i>bacca</i>
QPIE *ba/Hg- : Lat. <i>baiulus</i>	QPIE *ba/o/HG- (or *ba/o/Hg-nó-) : PGm. *pakka- QPIE *b ^h a/o/HG ^h - : PGm. *bagg-
QPIE *b ^h ul-Vk- : Lat. <i>fulica</i> QPIE *b ^h a/ol-ig- : OHG <i>belihha</i>	QPIE *b ^(h) o/ul-a/oK- : SGael. <i>bolachdan</i>
QPIE *la/Hk- : Lat. <i>lacerna</i>	QPIE *la/h ₂ K- : Gk. λάκκος QPIE *lh ₂ K- : Hsch. λόκκη
QPIE *la/Hp- : Lat. <i>lapis</i> QPIE *le/h ₁ p- : Gk. λίπας QPIE *le(h ₁)p- / *līp- : OIr. <i>líe</i>	QPIE *la/HP- : PRom. *lappa
QPIE *lep-os- : Lat. <i>lepus</i>	QPIE *la/HP-Vr- : PRom. *lapparo-

QPIE * <i>ph₂eil-a/ek-</i> : Lat. <i>paelex</i> QPIE * <i>pa/er-ik-</i> : OIr. <i>airech</i>	QPIE * <i>pa/HL-ak-</i> : Gk. <i>παλλακί</i>
QPIE * <i>Hra/Hg-</i> : Lat. <i>raia</i>	QPIE * <i>HruG^h-</i> : PGm. * <i>rugg-</i> QPIE * <i>HreK-</i> : PGm. * <i>rehhōn-</i>
QPIE * <i>sa/HP-</i> : OCo. <i>sibuit</i> QPIE * <i>sa/Hb^(h)-</i> : Lat. <i>sabīna</i>	QPIE * <i>sa/HP-</i> : Lat. <i>sappīnus</i>
QPIE * <i>ua/h₂k-</i> : Gk. <i>ύάκτινθος</i>	QPIE * <i>ua/HK-</i> : Lat. <i>vaccīnium</i>

Table 3.22 Alternations in gemination

A remarkable pattern emerges in which attested Classical Latin almost always has the variant *without* the geminate. (In *caballus*, it indeed has a geminate *l* despite a singleton *b* however). The robustness of this pattern is questionable however, as in 4 of these cases, Romance forms continue a geminate. This perhaps explains the geminate of Lat. *vaccīnium* (especially if related to *bāca*) and the gemination alternation between Lat. *sappīnus* and Lat. *sabīna*. Classical Latin mainly records forms with a singleton, but it seems to mask the variation that actually existed in Italic. Besides being a sociolinguistic phenomenon, this could also be a chronological issue. The forms with a singleton may have been borrowed before the phonemicization of gemination in Latin, with the geminate forms reborrowed upon the later expansion of Latin.

The geminate *r* of Lat. *cerrus* (and PRom **karr-*, **garr-*) could technically be the result of a suffix like **-so-* against It. dial. *cariglio* < **kar-*. But given the gemination alternations within Italic attested above, it may simply be original. The gemination of Lat. *pannus* is also likely to be genuine; it would be one of the only examples of the *littera* rule involving a nasal and, while it could be syncopated from **pan-ino-*, its meaning seems too basal for such a derived formation.

3.2.2 Vowels

The accent-ablaut system of Proto-Indo-European is relatively well understood. Morphemes with ablaut can appear in the zero grade, the full *e-* or *o-*grade, and the lengthened *ē-* or *ō-*grade. The semivowels **i* and **u* did not participate in ablaut, but could form diphthongs with ablauting vowels (appearing as stand-alone *samprasāraṇa* vowels in the zero-grade). The vowel **a*, if it is reconstructible for PIE outside of the influence of laryngeal coloring, was extremely rare. There are only a very limited number of cases that might suggest that it also participated in ablaut, and of them, only quantitative ablaut seems visible (cf. Forston 2010: 81, Melchert 2022). This thesis generally follows the idea that **a* is not reconstructible (cf. Lubotsky 1989, Pronk 2019) and if it is required, it is indicative of the non-native origin of the lexeme in which it occurs. Given this information, we can rule out certain vocalic alternations as inheritable from the outset. Cases of *i* ~ *u*, *e* ~ *u*, *o* ~ *u*, *e* ~ *i*, *o* ~ *i*, *u* ~ *ū* and *i* ~ *ī* are not reconstructible and are indicative of non-native origin. So are, in this thesis, alternations involving *a* of non-laryngeal origin.

In the most general terms, the semantic value of the different ablaut grades is unknown (Fortson 2010: 80), though they are understood to occur in a number of relatively predictable ways in the form of, for the nouns, the various reconstructed accent-ablaut classes (cf. Fortson 2010: 119-22, Beekes 2011: 190-216, Weiss 2020: 276-81 for overviews). Thus in the root of nouns, allowable ablaut is between *e*, *o*, and zero, with complications introduced by the effects of laryngeals. Lengthened \bar{e} -grade of the root seems to be reconstructible at least for $*H_1\check{e}k^{w-}r-$ ‘liver’. Lengthened grade roots may also have arisen in the form of Narten-type nominal roots (Schindler 1994) and through the process of creating $v\check{r}ddhi$ derivations (though that the latter dates to PIE is doubted by i.e. Beekes 2011: 181-2). Given these considerations, a question that arises is whether all cases that look like they could be ablaut are in fact of PIE origin. Without any other change in derivational morphology or semantics for example, vocalic alternation between comparanda that could be reconstructed as e.g. the zero-grade and lengthened \bar{o} -grade of a root does not fit into any of the understood accent-ablaut classes. Thus it looks inherited superficially, but the morphology does not behave in inherited ways, in turn suggesting that it may not be an inherited pattern after all. Both the clear-cut cases mentioned above and cases like this will be considered here.

3.2.2.1 Clearly non-IE Alternations

3.2.2.1.1 *E ~ I*

Vocalic alternation between *e* and *i* has often been given as a Mediterranean, frequently specifically Etruscan alternation but also Anatolian (e.g. Bertoldi 1939b: 89, Battisti 1959: 154-7). Breyer (1993: 16) notes that, because Etruscan underwent a sound change from *i* > *e*, Latin words that show an alternation (like *vespillō/vispillō*) might represent borrowings from Etruscan at different times or regions, or have to do with Etruscan vocalic phonology. When alternations occur with comparanda outside of Italic, it seems like Etruscan can have little to do with the alternation unless as a mediator of vocabulary to Latin—a difficult hypothesis since none of the forms is attested in Etruscan.

The alternation is indeed found with an almost exclusively Mediterranean distribution (see §4.2.2.4.2):

QPIE $*e$	QPIE $*i$
QPIE $*kedro-$: Gk. κέδρος	QPIE $*kitro-$: Lat. <i>citrus</i>
QPIE $*g^heI-iHd-\bar{o}n-$: Gk. χελιδών	QPIE $*g^hir-o/und^{(h)}-\bar{o}n-$: Lat. <i>hirundō</i>
QPIE $*gen-es-to-$: Lat. <i>genesta</i>	QPIE $*gen-is-to-$: Lat. <i>genista</i>
QPIE $*kup(V)r-et-to-$: Lat. <i>cupressus</i>	QPIE $*kupa-r-it-jo-$: Gk. κυπάρισσος
QPIE $*m(e)nt-$: Lat. <i>menta</i>	QPIE $*mind^h-$: Gk. μίνθη
QPIE $*g^hed^{(h)}-a/er/s-$: Lat. <i>hedera</i>	QPIE $*k/g^hid^h-ar-$: Gk. κίθαρα
QPIE $*h_2eu-e(C)s-n-$: Lat. <i>avēna</i>	QPIE $*h_2eu-ik/s-$: PSlav. $*ovъsъ$
	QPIE $*h_2eu-iǵ^h/S-$: PEbalt. $*(a)vi\check{z}a\check{z}$

Table 3.23 Alternations between $*e$ and $*i$

A less clear example is that of Gk. σικύα ~ Hsch. σεκούα, Arm. *sex* (where the Lat. comparandum is *cucumis*). For Lat. *menta*, a loan from Greek through Sabellic could theoretically explain Lat. *e* for *i* (see fn. 171), perhaps in the other cases with this distribution as well. Without attestations of the word in Sabellic, this explanation is no better than suspecting Etruscan as an intermediary. The pair Lat. *hirundō* ~ Gk. χελιδών has a further comparandum in Alb. *dallëndyshe*, whose root vocalism cannot reconstruct to **e* or **i*. The singular non-Mediterranean case is that of Lat. *avēna* against Baltic and Slavic forms with **i*.

3.2.2.1.2 I ~ U

Vocalic alternation between *i* and *u* has also often been considered characteristic of the Mediterranean substrate (e.g. Hubschmid 1953: 28, Alessio 1955: 375, 537-40 suggesting a substrate vowel *ū*; Battisti 1959: 155), potentially in relation to Etruscan (cf. Bertoldi 1948: 70). Some cases of this alternation within Latin can be explained as regular however. For example, *lubet* ~ *libet* ‘it pleases’ and *clipeus* ~ *clupeus* ‘shield’ attest to the change **u* > *i* between *l* and a labial (cf. Weiss 2020: 153). While Bertoldi (1948) suggested that *fūnis* and *fīnis* were related via borrowing from a substrate, such inner-Latin relationships can be explained via IE ablaut, with *e*-grade **ei* > *ī* and *o*-grade **oi* > *ū* (though in this case, the words are not related). Additionally, Sabellic seems to have undergone a change *ū* > *ī* in monosyllables (Cf. Buck 1904: 41). There are several cases of *i* ~ *u* alternation that cannot be explained in these ways; all indeed show a Mediterranean distribution of attestation:

QPIE <i>*i</i>	QPIE <i>*u</i>
QPIE <i>*g^hel-iHd-ōn-</i> : Gk. χελιδών	QPIE <i>*g^hir-o/und^(h)-ōn-</i> : Lat. <i>hirundō</i>

Table 3.24 Alternation between **i* and **u*

As for the **e* ~ **i* alternation in Lat. *hirundo* ~ Gk. χελιδών, Alb. *dallëndyshe* reconstructs to different vocalism. An additional case seems to include Lat. *cucumis*, Hsch. κύκουιν in alternation with Gk. σικύα (though compare also Hsch. σεκούα, Arm. *sex* in the **e* ~ **i* alternation above). Finally, even if *supparus* is a loan from Oscan, its *u*-vocalism against Gk. σίφαρος remains irregular.

QPIE <i>*ī</i>	QPIE <i>*ū</i>
QPIE <i>*d^hīk-</i> : Lat. <i>fīcus</i> (Hebr. <i>šiqmā</i>)	QPIE <i>*d^htī/ūūk-</i> : Gk. τῦκον, σῦκον QPIE <i>*tu/ūg^h-</i> : Arm. <i>tūz</i>
QPIE <i>*b^hrīg-</i> : Lat. <i>frīgō</i>	QPIE <i>*b^hrūg-</i> : Gk. φρῦγω

Table 3.25 Alternations between **ī* and **ū*

There is an additional *ī* ~ *ū* alternation present in Lat. *brīsa*. Its most proximal source might be a pre-form of Albanian (or a relative thereof), but the vocalism of PALb. **brīsā-* is in irregular alternation with forms attested as Gk. βρῦτος, βρῦτεα.

3.2.2.1.3 $E \sim U$

QPIE * <i>e</i>	QPIE * <i>u</i>
QPIE * <i>sb^h/g^{wh}end-</i> : Gk. σφενδόνη	QPIE * <i>b^h/g^{wh}und^(h)-</i> : Lat. <i>funda</i>

Table 3.26 Alternation between **e* and **u*

There is no regular way to get a change from **o* > Lat. *u* in this position, thus the *e* ~ *u* alternation is original.

3.2.2.1.4 $O \sim U$

Alternations within Latin words between *o* and *u* can occur for several reasons. Some are regular (like **o* > *u* / _ IC). Because Etruscan lacked a graphical distinction between *o* and *u* and Latin transcriptions of some Etruscan names use Lat. <*o*> for Etr. <*u*>, the Etruscan vowel's quality may have been phonetically between Lat. *o* and *u* (cf. Breyer 1993: 14-15). Indeed, there are several cases where Etruscan is suspected to have mediated Greek words with *o* to Latin where they show up as *u* and *vice versa*: e.g. Lat. *ancora* < Gk. ἄγκυρᾶ 'anchor', Lat. *sporta* < Gk. [acc.] σπορίδα 'basket', Lat. *amurca* < Gk. ἀμόργη 'watery part of pressed olives' (cf. Alessio 1941a: 551 fn. 2, de Simone I: 132-42). Without attested Etruscan forms, this is difficult to verify. Notably, Messapic too lacked a distinction between *o* and *u* (de Simone 2018: 1844). Several lexemes from the dataset show this alternation:

QPIE * <i>o</i>	QPIE * <i>u</i>
QPIE * <i>ke/ol-o/umb^(h)-</i> : Lat. <i>columba</i> QPIE * <i>g^(h)ol-omb^(h)-</i> : OCS <i>golq̃bь</i> QPIE * <i>k̃ol-(o)mb^h-</i> : Arm. <i>salamb</i>	QPIE * <i>gul-ub^h-</i> : PGm. * <i>kulubrōn-</i>
QPIE * <i>kotōn-</i> : Lat. <i>cotōneum</i>	QPIE * <i>kudōn-</i> : Gk. κυδώνιον QPIE * <i>kodu-</i> : Gk. κοδύμαλον
QPIE * <i>H(o)rk-</i> : Lat. <i>orca</i> [QPIE * <i>H(o/u)rk-</i> : Lat. <i>urceus</i>]	QPIE * <i>Hurg^h-</i> : Gk. ὕρχη [QPIE * <i>H(o/u)rk-</i> : Lat. <i>urceus</i>]
QPIE * <i>b^ha/ol-ig-</i> : OHG <i>belihha</i> [QPIE * <i>b^(h)o/ul-a/oK-</i> : SGael. <i>bolachdan</i>]	QPIE * <i>b^hul-Vk-</i> : Lat. <i>fulica</i> [QPIE * <i>b^(h)o/ul-a/oK-</i> : SGael. <i>bolachdan</i>]

Table 3.27 Alternations between **o* and **u*

\bar{o}	\bar{u}
QPIE * <i>kroHm-</i> : Hsch. κρῶμαξ QPIE * <i>kloHm-</i> : Gk. κλωμαξ	QPIE * <i>gruHm-</i> : Lat. <i>grūmus</i>

Table 3.28 Alternation between \bar{o} and \bar{u}

\bar{o}	<i>u</i>
QPIE * <i>s(u)oHr-e/ak-</i> : Lat. <i>sōrex</i> [QPIE * <i>ṣuo/ur-ak-</i> : Gk. ὕραξ]	QPIE * <i>sur-Vg-</i> : PGm. * <i>s(w)ur(V)ka-</i> [QPIE * <i>ṣuo/ur-ak-</i> : Gk. ὕραξ]

Table 3.29 Alternation between \bar{o} and *u*

Intermediation via Etruscan or Messapic seems particularly unlikely for *columba*⁵⁰³ and *fulica* due to their distribution of attestation. While *cotōneum* has been suspected of being an Etruscan-mediated Greek loan, especially given its apparent devoicing (cf. de Simone I: 134, II: 271-2, 279), the preservation of *ō* is suspicious as Etruscan does not seem to have distinguished vowel length.

3.2.2.2 *a*-Vocalism

3.2.2.2.1 *Reconstructed a-Vocalism*

As mentioned above, there is reason to be suspicious of roots for which original *a*-vocalism must be reconstructed. There are number of ways that *a*-vocalism can arise in inherited roots beyond the effects of laryngeals. In Latin for instance, Schrijver (1991: 505) summarizes five sound laws that he finds can lead to Latin *a*: 1) **e > a* after a pure velar, 2) Thurneysen-Havet's Law of **ou̯ > au*, 3) **o > a* /*m, ŋ* _ CV and /*m* _ *r* + velar, 4) **RDC > RaDC*, and 5) epenthesis of *a* in sequences C_CCC. Not all of these are universally accepted. There are nevertheless several lexemes for which the *a*-vocalism in Latin has no regular explanation, suggesting it was present at the time of borrowing.

Since **CRHC* yield Lat. *CRāC*, the short *a* of *fracēs* is straightforwardly reconstructed as having entered Italic as **a*.⁵⁰⁴ As to Lat. *trabs*, Schrijver (1991: 376, 482) suggests that the Latin form, on comparison with **trēb-* in Sabellic, if it does not represent a regular development from **trb-*, could represent a morphological zero-grade **trēbs*, **trābes* in which the expected **torb-* < **tṛb-* was replaced with a form with *a*-vocalism on comparison with the more frequently seen *ē/ā* pattern amongst verbs (like *agere*, *ēgī* and *frangere*, *frēgī*). That *trabs* could have developed from **trb-* is not clear; though Schrijver (1991: 483-4) finds evidence that tautosyllabic **RD* (where **D* is an unaspirated voiced stop) yielded *RaD*, he prefers the idea that *trab-* developed in the oblique cases of **trēb-* and thus would not have been tautosyllabic. But the idea that the oblique stem of a nominal paradigm was reshaped on analogy with verbal ablaut seems very strange. Given the problems, I think the most straightforward reconstruction is with **a*. Such seems also to have been the case for (at least) the second *a* of Lat. *caballus* and *tamarix*. Potential Sanskrit comparanda for Lat. *caput* and *calix* must reconstruct to *a*-vocalism, since they have not palatalized the preceding *k* or undergone Brugmann's Law, but their appurtenance is uncertain.

For none of the words given here is the reconstruction of *a*-vocalism the only peculiarity; other alternations confirm a non-inherited origin. This seems like good evidence for the existence of this vowel in the contact languages.

⁵⁰³ The Coptic forms, lacking the first vowel, show *oo*, *o*, *a*, and *aa*, interestingly suggesting *a*-vocalism in the Egyptian parent form (Allen 2020).

⁵⁰⁴ This would be the case for *grāmia* as well, if its vowel (whose length is indeterminate) were short.

3.2.2.2.2 Alternations Involving A

In the cases above, attested *a*-vocalism has no internally reconstructible source and is likely to have been original. There are additionally several cases where *a*-vocalism can be reconstructed to valid IE pre-forms though only for individual daughter languages; that is, alternations between comparanda rule out the feasibility of these reconstructing representing anything besides original *a*-vocalism. In other cases, irregular alternations that indicate non-native origin make the reconstruction of specifically IE phonemes like laryngeals highly dubious.

3.2.2.2.2.1 A ~ \bar{A}

While Lat. *racēmus* can reflect **HrHk-* (since **HRHC* > Lat. *RāC*, Schrijver 1991: 314), Greek comparanda lack a prothetic vowel and thus rule out an initial laryngeal. While ** μ re/oHg-* could be behind the Greek forms (ῥᾶξ, ῥώξ), ** μ rHg-* in Latin should have given ** $\bar{r}āc-$* . Thus here too, the *a*-vocalism of Latin was likely present upon borrowing.

The short **a* of Gk. ῥάφους against the long ** \bar{a}* of Lat. *rāpum* and Lith. *rópė* could be seen as ablaut grades of a sequence **eh₂*. The Germanic preforms can reconstruct to ** \bar{a}* (as if **eh₂*) or ** \bar{o}* (as if *oh₂*). But the root must be reconstructed with invalid initial **r-* and Greek evidence provides variation in the quality of the plosive. The Slavic vocalism reconstructs to more aberrant ** \bar{e}* or **oi* (or **ai*), and the Celtic comparanda instead show something akin to the *a*-prefix phenomenon. Unlikely to be inherited, the attested *a*-vocalism is thus unlikely explainable via PIE laryngeals.

The long ** \bar{a}* of Lat. *bāca* stands in contrast to the short **a* of Celtic forms like W *bagad*. These could be reconstructed as ablaut grades of a sequence **eh₂*, but the root begins with **b*, suspicious if inherited. It. *bacca*, if its geminate is original, further points to non-native origin and thus original *a*-vocalism. Similar is the case of Lat. *pannus* alongside PGm. **fanan-* against Gk. πήνη < ** \bar{a}* . Its geminate *n* is unlikely to be due to the *littera* rule and points to non-native origin.

3.2.2.2.2.2 A ~ E

Even amongst those who work with the existence of a quantitatively ablauting PIE **a*, there is uncertainty about its participation in qualitative ablaut (cf. Melchert 2022: 202). Šorgo (2020: 457-8) lists several cases of *a* ~ *e* alternations in what he identifies as the substrate lexicon of Germanic, proposing it resulted from treatments of a substrate vowel intermediate to **a* and **e*. An *a* ~ *e* alternation has also been proposed to be characteristic of the Mediterranean substrate (cf. Battisti 1943: 146; Alessio 1946a: 165; Hubschmid 1953: 48, Battisti 1959: 130, 147, 284).

An *a* ~ *e* alternation occurs in the initial vowels of Lat. *alaternus* against Cretan Greek ἐλαίτρινος. While ** \bar{h}_1lC-* could have yielded the forms, ** \bar{h}_1lV-* cannot. Romance comparanda for Lat. *cerrus* attest to *a*-vocalism. While this could be the result of ablaut within a root ** $\bar{k}h_1er-$* (full *e*-grade for *cerrus* and zero-grade for the Romance forms),

further alternations in consonant voicing and gemination demonstrate non-native origin of this lexeme. These two cases are attested with a Mediterranean distribution.

The *a* of Lat. *lapis* against the *e* of Gk. λέπας could reconstruct to **h₁*. But Romance forms attest to an irregular geminate *pp* (potentially even more aberrant P_{ROM}. **lībb-*). OIr. *līe* (if it reconstructs to **p* rather than **μ*) can be reconstructed to either **ē* like Greek or **ɾ* like the most aberrant Romance form.

For the Slavic comparanda of Lat. *alnus*, an *a ~ e* alternation does not seem to be easily explained away. The same is probably true for the **e* of ON *jǫlstr* against **a* elsewhere. (It cannot be ruled out that the *a ~ e* alternation of the Baltic forms is due to Rozwadowski's change.) While the *a ~ e* alternation in Germanic, Baltic, and Slavic can be explained as **o ~ *e*⁵⁰⁵ and thus QPIE **h₁e ~ *h₁o*, Lat. *alnus* would require **h₂e*. Thus a laryngeal cannot account for the alternation. Lat. *aper* and its Umbrian cognates have *a*-vocalism against *e*-vocalism in Germanic and Greek. Some have suspected contamination from the *a* of *caper*. But it can also be taken at face value as a substrate alternation. That Balto-Slavic has an additional element before the vowel is further evidence of this.

The long *ē* of Lat. *cēpa* cannot be reconstructed to the same pre-form as Hsch. κάπια. If the latter is truly Greek, then it attests to an irregular alternation. If it is not, then Hesychius has recorded a foreign word that appears in Latin with different vocalism. Finally, Lat. *nāpus* corresponds to Gk. νάπυ, whence it may or may not be a borrowing. If independent, it stands in irregular alternation with Arm. *nīw*, whose vocalism can be reconstructed to **i* or **ē*.⁵⁰⁶

3.2.2.2.2.3 A ~ O

The analysis of *a ~ o* alternation is made difficult by the fact that both phonemes merge in Albanian, Germanic, Baltic, Slavic, and Indo-Iranian. Thus the true extent of *a*-vocalism and *a ~ o* alternation is hidden in these branches.⁵⁰⁷ Šorgo (2020: 458-9) even suggests that the substrate language(s) of Germanic may not have had a phonemically rounded low vowel, noting that in several cases of P_{Gm}. **a* against **o* in other branches, the phonetic environment between a labial and resonant may be responsible for conditioning the rounding.

Within Latin, an *a ~ o* alternation seems to exist between *caulae* and *cohūm*. There are also a few examples between Latin and other branches that do not merge **a* and **o*. It likely appears between Lat. *corbis* ~ OIr. *carpat* (though the appurtenance of the latter is uncertain), unless this represents a *ø ~ a* alternation (**krb^(h)-* for Latin, **karb^(h)-* for

⁵⁰⁵ In light of this, we can reconstruct either **a ~ *e ~ *o* alternation or simply **a ~ *e* alternation. The latter seems preferable (cf. Šorgo 2020: 459 fn. 38).

⁵⁰⁶ Lat. *cēra* cannot be ruled out as a loan from Gk. κηρός. But the latter, whose vocalism is almost certainly to be reconstructed as **ē*, stands in irregular alternation with PEBalt. **kār-*, attesting to an **ā ~ *ē* alternation.

⁵⁰⁷ And given that the languages where the merger does not occur are Celtic, Armenian, Greek, and Latin, there exists the risk of over-Mediterraneanizing the presence of *a*.

Celtic). While some reconstruct this alternation for PIE (cf. Melchert 2022: 198, Forston 2010: 81), an Indo-European origin for this lexeme is made unlikely by the further alternation introduced by PGm. **krebō-* < QPIE **grebʰ-*. The vocalism of Lat. *lacerna* matches that of Gk. λάκκος, though the latter itself alternates with Gk. λόκκη. This is similar to the case of Lat. *racēmus* above, which alternates with Gk. ῥάξ and ῥώξ. The *a* ~ *o* alternation of Lat. *badius* and OIr. *buide* is less straightforward, as mechanisms have been proposed within each language that could result in the change: Celtic **a* raising between a labial and a palatal consonant (Thurneysen 1946: 50), Italic **o* unrounding after a labial consonant (Schrijver 1991: 454-65). In neither case is it fully certain that the mechanism can have occurred.

3.2.2.2.2.4 A ~ AU

There are two cases of an *a* ~ *au* alternation in the dataset. The first is the Mediterranean pair Lat. *caurō* ~ Gk. κάπηλος. Even if the ultimate source is Hitt. *hāppar-* (cf. Puhvel III: 127), it has not entered into Latin directly. Second, PGm. **haubuda-* and **haubeda-* attest a diphthong **au* against **a* in PGm. **habuda-* and **hafulan-* as well as in all other comparanda for the word (including Lat. *caput*). This has been interpreted as *u*-infection (Boutkan 1998: 111, DV 91) or metathesis from oblique forms (Kroonen 2013: 215), but in light of the possibility of non-IE origin, this could represent a genuine alternation.

3.2.2.2.2.5 A ~ U

The best semantic match for Lat. *calix* (whose *a*-vocalism may be matched by Skt. *kaláśa-*) is Gk. κύλιξ, attesting to an *a* ~ *u* alternation. Lat. *raia* could reconstruct to **HrHg-*, but its Germanic comparanda (PGm. **rugg-* and **rehhōn-*) do not reconstruct to a laryngeal. Thus the *a* of Latin alternates with both **u* and **e* in Germanic. Lat. *tamarix* has either been borrowed very late (after the effect of vowel weakening) or owes its unweakened second *a* to the *alacer* rule. In any case, its original quality was **a*, whether of laryngeal origin or not. It stands in alternation to the *u* of Gk. μωρίκη, which appears without the initial *ta-* of the Latin form (and has therefore been analyzed as a prefix, perhaps of Berber or Semitic origin).

3.2.2.2.2.6 A ~ AI

Schrijver (1997: 306) notes a pattern in which British Celtic *a* corresponds to Germanic *ai* in eight lexemes, proposing that it is the result of the branches nativizing a foreign substrate phoneme like /aə/. In none of the lexemes does a comparandum exist in Latin. He does however mention a similar phenomenon in the Mediterranean, noting for example Gk. χλανίς against Gk. χλαῖνα (cf. also Lat. *laena*) as well as Lat. *aesculus*,⁵⁰⁸ for which he considers the comparison with Gk. ἄσκρα and Basque *azkaí* and *askaí* certain (and not mentioning the possibility of a relationship with PGm. **aik-*). Because

⁵⁰⁸ Cf. also Bertoldi (1942: 191 but only between Basque and Greek, the Basque forms somehow with *e-*), Alessio (1948-9: 148), Hubschmid (1953: 84, fn. 1).

of the differences in patterns of attestation, he is rightly hesitant to attribute both patterns to the same phenomenon. Šorgo (2020: 457-8) suggests expanding this alternation pattern to include the cases of *a ~ e* alternation within Germanic and between Germanic and other branches. He similarly proposes that this was the result of a foreign phoneme (transcribed as *æ), perhaps one that was perceived by PIE-speakers as similar to both *a and *e. Including this as part of the same phenomenon as the *a ~ ai* alternation involves the inclusion, beyond Germanic and Celtic, of Italic, Greek, Baltic, and Slavic.

I would rather follow Schrijver's more cautious approach in considering these different alternations as potentially separate phenomena, especially in light of the large number of more or less perfect Celto-Germanic correspondences he is able to locate. Thus I here list the *a ~ ai* alternation separate from the *a ~ e* alternation above. I find only two examples (excluding Lat. *laena*, as its diphthong seems to suggest an indirect borrowing from Greek, where the diphthong may be the result of Greek sound laws) of an *a ~ ai* alternation involving Latin. The first is Lat. *alaternus*, whose second *a* vowel stands against α in Cretan Greek ἐλαίτρινος. Second is Lat. *paelex* ~ Gk. παλλακή. At first glance, it seems to fit the secondary Mediterranean type that Schrijver mentions, but this is complicated by the additional comparison with OIr. *airech*. Its *a* most easily reconstructs to *a*-vocalism, though before a palatalized consonant, **e* is also a possibility. It feels like a stretch to link the Celto-Germanic *a ~ ai* alternation to the Mediterranean *a ~ ai* alternation on the basis of one form whose vocalism is not even guaranteed. Thus it really seems like these are two separate phenomena, with OIr. *airech* potentially representing the partial participation of Celtic in the Mediterranean substrate.

3.2.2.3 Wider Variation

There are several cases for which the wider amount of variation in reconstructible vocalism makes it difficult to categorize alternations.

A root of the shape **ureh₂d-* could yield PGM. **wrōt-* in the full-grade and PALb. **wradn(i)ā-* and PGM. **wurti-* in the zero-grade (along with Lat. *rādīx* and Gk. ῥᾰδίξ) in either. No ablaut grade of a root of this shape can give PCelt. **wradi-* however (as both *e-* and zero-grades would give ***wrādi-*). Nor can a root with a laryngeal give PCelt. **wridā-* or Gk. ῥίζα. The original thus variation seems to be between **ā*, **a*, **ɹ*, and **i*. Given that the presence of an IE feature like a syllabic *r* in a non-IE language is dubious, perhaps the vowel of PGM. **wurti-*, PCelt. **wridā-* and Gk. ῥίζα reflects something akin to a schwa (cf. Šorgo 2020: 456 fn. 25).

The *a* of Lat. *trabs* alternates with **ē* in Oscan and **e* in Umbrian and Celtic as well as **o* (with Winter's Law and metathesis) in Baltic and what reconstructs to a syllabic resonant in Germanic (perhaps original **u* or a schwa as above).

Lat. *tilia* can be reconstructed to original *e*-vocalism like PGM. **felwō-*, Arm. *ʿeli*, and Gk. πτελέα. If PCelt. **axil/nV-* can be reconstructed as such, it points to the vocalic reduction triggered by an *a*-prefix (see more below). If Lat. *pōpulus* is connected, then

this reducible *e*-vocalism additionally alternates with **ō* and **a/o*.

The case of Lat. *plumbum* is complex, likely representing a Wanderwort. The whole family seems to reflect a non-IE diphthong. PGM. **blīwa-* against Gk. μόλιβδος, μόλυβος points to an **ī ~ *u* alternation in a diphthong with **b ~ *w* alternation. Lat. *plumbum* matches μόλυβος quite well, if perhaps the vowel of the diphthong was nasalized. PCelt. **(φ)loudio-* < QPIE **ple/oud^(h)-* is difficult to analyze. If the dental element is a suffix like μόλυβδος, then the **u* of its diphthong might correspond to the **b ~ *w* element. Its **e/o*-vocalism would then be in alternation with the **ī ~ *u* of the other forms.

3.2.2.4 Ablaut Phenomena

It is important to note, as has been mentioned above, that several lexemes attest to vocalic alternation that looks similar to acceptable Indo-European ablaut. In many cases, further irregular correspondences between the comparanda show that the words are not of IE origin, and therefore that the resemblance to IE ablaut is coincidental.⁵⁰⁹ In other cases, the vocalic alternation is between vowels that participate in ablaut, but its appearance does not correspond to a known accent/ablaut pattern. In these cases, we must ask, as will be done in §3.3 about morphological features, whether Indo-European features behaving in non-Indo-European ways might not be Indo-European after all.

3.2.2.4.1 Ablaut Unparalleled in IE

Lat. *almus* and its comparanda can be interpreted as a root **a/el-* with a sigmatic suffix, but the alternations in vocalism of that suffix do not follow a known PIE pattern. Szemerényi (1960: 228) proposed **-is-* as the basis for all comparanda, but Latin *almus* can only be reconstructed to **-s-*. The *i ~ u* alternation within the suffix in PGM. **aluz-* beside **alis/zo-* is the expected reflex of an inherited PIE *s*-stem (cf. Schrijver 1991: 41), but the *i*-vocalism is present in Slavic as well, where it cannot be explained in such a way (cf. Derksen 2007: 370). Thus its incorporation into the inherited *s*-stems seems to be a uniquely Germanic development. The absence of the *i* in Baltic may or may not be due to sporadic syncope. However, put all together, the evidence shows the alternating presence of a vowel in a sigmatic suffix **-s-* (guaranteed by Latin) ~ **-is-* (guaranteed by Slavic). This alternation cannot be understood in terms of inherited ablaut and, in the face of the *a ~ e* alternation also present in the comparanda, is a feature of the substrate language itself or the borrowing process from that language.

Inherited ablaut is likewise unable to account for, on the one hand, Lat. *ulmus* and the Germanic variants **elma-* and **alma-* and, on the other hand, PCelt. **limo-/*lemo-* (behind MIr. *lem*) and **lēmā-* (behind W *llwyf*) < QPIE **(h₁)leim-*.

⁵⁰⁹ Cf. Stifter's (fthc.) "linguistic pareidolia".

3.2.2.4.2 *Ablaut Difficult to Motivate from an IE Perspective*

The most secure comparanda for Lat. *ardea* is Gk. ἔρφιδίος. Even if the variants ἀρφιδίος and ῥοιδίος are explained away, *ardea* can only be reconstructed to a zero-grade **Hrd-* to which ἔρφιδίος represents the lengthened *o*-grade (as **HrHd-* ought to yield Lat. **radea*, cf. Schrijver 1991: 314). It is immediately suspicious that there is no recognized accent/ablaut pattern that results in **ō ~ *φ* ablaut. The appurtenance of PGm. **artō(n)-*, which requires a full vowel to the left of the resonant further points the conclusion that the unparalleled ablaut is not PIE at all.

The same pattern occurs for Lat. *sōrex* beside Gk. ὄραξ and OSw. *surk*. Working from a root **suer-*, Vine (1999a: 572-3) explains the Greek vocalism as an original *o*-grade with Cowgill's Law. This cannot apply to Germanic however, which can only reconstruct to a zero-grade **sur-*. Again we are faced with unparalleled **ō ~ *φ* ablaut, this time in roots of identical structure (i.e. none of the forms can be argued to be derivational). The aberration in voicedness of the velar suffix between the forms allows us to conclude that the ablaut looks so strange because it is not PIE.

3.2.2.4.3 *Vocalic Alternations That Can Occur in Ablaut Paradigms*

There are several cases where vocalic alternation can be reconstructed as relatively unproblematic IE ablaut, but for which other irregular correspondences provide sufficient evidence of a non-IE origin. In these cases, the ability to reconstruct inherited ablaut grades must be due to coincidence.

The following lexemes can be reconstructed to alternations in *e-* and *o*-vocalism. In inherited lexemes these could continue an old acrostatic paradigm, but here they are not inherited:

QPIE <i>*e</i>	QPIE <i>*o</i>	Irregularities
QPIE <i>*h₁er(H/V)_u-</i> : Lat. <i>ervum</i> QPIE <i>*h₁erh₁b-</i> : Gk. ἐρέβινθος	QPIE <i>*h₁orh₃b-</i> : Gk. ὄροβος QPIE <i>*h₁oru-</i> : ON <i>ertr</i> QPIE <i>*h₁orVb-</i> : Arm. <i>arowoyt</i>	<i>b ~ w</i> Disyllabic root
[QPIE <i>*su(e/o)lp-</i> : Lat. <i>sulpur</i>] [QPIE <i>*su(e/o)lF-</i> : PRom. <i>*su(l)fur-</i>] QPIE <i>*s_ue(l)b^hlp-</i> : PGm. <i>*swe(l)bla-</i>	[QPIE <i>*su(e/o)lp-</i> : Lat. <i>sulpur</i>] [QPIE <i>*su(e/o)lF-</i> : PRom. <i>*su(l)fur-</i>]	<i>*b^h ~ *p</i>

Table 3.30 Alternations between **e* and **o*

Several cases can be reconstructed to alternations in *e*-grade and zero-grade vocalism, as if perhaps (leveled) continuants of a proterokinetic stem. Again, here they are not inherited:

QPIE * <i>e</i>	QPIE * <i>ø</i>	Irregularities
QPIE * <i>b^hers-</i> : Lat. <i>ferrum</i>	QPIE * <i>b^hros-</i> : PGm. * <i>brasa-</i>	Schwebeablaut
QPIE *(<i>H</i>) <i>mes-Vl-</i> : Lat. <i>merula</i> QPIE *(<i>H</i>) <i>mes-(a)l-</i> : PCelt. * <i>mesal-</i>	QPIE * <i>h₂/s_{ems-lo-}</i> : PGm. * <i>amslōn-</i>	Schwebeablaut
QPIE * <i>yeiks-</i> : PGm. * <i>wīhsilō-</i> QPIE * <i>yei(k)s-</i> : PSlav. * <i>višb-</i>	QPIE * <i>yisk-</i> : Lat. <i>viscum</i> QPIE * <i>yiks-</i> : Gk. ἰξός	SK metathesis

Table 3.31 Alternations between **e* and **ø*

For Lat. *ferrum*, a derivation from **b^her-s-* looks like a proterokinetic *s*-stem, but the **b^hr-os-* with zero-grade of the root for PGm. **brasa-* is unexpected, even from a neuter *s*-stem. Thus it seems rather that the **s* was part of the root. Instead of an *e*-grade ~ zero-grade alternation, if the **s* is part of the root, we have *e*- and *o*-vocalism, but on opposite sides of the resonant. Its presence in Italic and Germanic alone in the face of a similar-looking Wanderwort of Luwian origin makes it very unlikely that the Italic and Germanic words are inherited. For Lat. *merula*, the forms without an initial vowel show a full-grade root whereas Germanic, with an initial vowel, shows a zero-grade root. This could be construed as Schwebeablaut in root **h₂ems-*, but it fits much better into the pattern of *a*-prefixation that will be discussed below.

3.2.3 Phonological Conclusions

The overall trend that emerges is that phonological alternations on their own, while being the gold standard for identifying non-inherited lexemes, are not very useful for stratificational purposes. A combination of at least four factors has been interacting to produce the complicated picture we have received. 1) Latin has interacted with an unknown number of other languages. 2) Latin has interacted with these languages at different points in time and for different lengths of time; thus both it and the languages with which it was interacting were undergoing changes during the periods of contact. 3) Latin had a closed set of phonemes which it used to reflect all foreign sounds present in the words it borrowed. Another language's larger phoneme inventory would have been collapsed. But even if another language had a smaller phoneme inventory, if some of those phonemes were perceived as intermediate to native sounds, they could have been borrowing different ways (cf. the ideas proposed for the borrowing of Etruscan *u*). These effects of the borrowing process are likely behind some of the various alternations in plosive voicedness and aspiration. 4) The languages from which Latin borrowed could themselves have had dialectal variation or have been related more distantly at the family level to other languages, resulting perhaps in some of the same alternations as in (3) but also perhaps more drastic alternations (like some of those beyond the plosive rows). The effects of each of these factors has been collapsed down to one dimension: that of the attested Latin lexicon. And it is therefore perilous to conclude that every example of an irregular alternation is due to contact with the same language. In the other direction, it is difficult to know which alternations are related to one another. Are the cases of *s*-insertion related to the cases of SK metathesis for instance?

In fact, describing the alternations in terms of voicing and aspiration is almost certainly inaccurate. It is a useful shorthand to show that alternations exist by reconstructing proto-forms to PIE and pointing out their irreconcilability. But this obfuscates the original nature of the phonology of the contact languages. The traditional reconstruction of *tenues*, *mediae*, and *mediae aspiratae* is in competition with the various reconstructions under the glottalic theory. Thus a reconstructed $*b^h \sim *p$ alternation could instead have been a $*p: \sim *p$, $*p \sim *p^h$, or $*p \sim *b$ alternation. Which of these represents the truth, we may never know. (Though in some cases, like *ficus*, our guesses can be refined.)

A clearer stratificational picture will be provided when the alternations are examined in combination with the distributions of the lexemes in which they are attested (see §4).

3.3 Morphological Alternations

An interesting contrast to the phonological alternations is provided by morphological alternations. As opposed to the large number of different factors that can all have the same phonological result, there is less of a chance that two different languages have identical morphemes (or morphological phenomena). It is by no means bulletproof however (sometimes unrelated languages *do* have otherwise identical morphemes or phonological consequences of the borrowing process might merge two originally different morphemes).

Much of the work that has been done on non-IE morphological features to date has been on affixes. For instance, Ernout (1946: 21-51, reprint of 1930) proposed several suffixes in Latin that he thought could represent borrowings from Etruscan. Bertoldi, Alessio, Battisti, and Hubschmid defined several different substrate suffixes that recurred in Latin and Romance languages. Beekes (esp. 2014) lists 149 suffixes that he attributes to his version of Pre-Greek. With the understanding that suffixes of any origin can be added to bases of any origin, I looked for morphological features secondarily. Some morphological patterns factor into the primary evidence, such as the *a*-prefix phenomenon and polysyllabic roots, because they cannot be regularly reconstructed to PIE. Otherwise, considered in this section are a few cases of recurring morphemes on lexemes whose non-native origin is indicated by other features, making them potentially also of non-native origin themselves.

3.3.1 Pre-Greek Suffixes

While the following sections will discuss morphological features that appear between comparanda of Latin words, there are several cases of Latin words that themselves contain suffixes otherwise suspected to be Pre-Greek (in the Beekesian sense of restricted to Greece).

3.3.1.1 Latin *-essus*

On the basis of Greek placenames in $-(\sigma)\omicron\varsigma$ (and $-(\tau)\omicron\varsigma$) that matched placenames in Asia Minor, Kretschmer (1896: 405-6, further e.g. 1923a: 69) proposed influence from a language of Asia Minor, with the caveat that such a sequence also appears in inherited formations. With the discovery of the Anatolian languages, the possibility that these suffixes represented vestiges of previous Anatolian-speaking (more specifically Luwian) inhabitants developed (cf. recently Finkelberg 2006: 52, West 2007: 8, who names it “Parnassian”). The variant $-(\tau)\omicron\varsigma$ has led some to be suspicious of this explanation (cf. Morpurgo Davies 1986: 119-120). Beekes (2009: 192-3) explains it as the reflex of Pre-Greek palatalized velars (adducing as evidence the Hsch. var. $\delta\alpha\lambda\acute{\alpha}\gamma\chi\alpha\nu$ of Gk. $\theta\acute{\alpha}\lambda\alpha\sigma\sigma\alpha$, Attic $\theta\acute{\alpha}\lambda\alpha\tau\tau\alpha$) and (in Beekes 2014: 39) considers it a non-IE Pre-Greek suffix.

Whether of ultimate IE origin or not, the suffix is at home in Greek. In Latin, the sequences *-issalus* and *-essalus* appear, when not inherited (e.g. compounds of *missus* and *gressus* etc.), in direct loans from Greek (e.g. *narcissus* < $\nu\acute{\alpha}\rho\kappa\iota\sigma\sigma\omicron\varsigma$). In one case, *cupressus*, the suffix appears on a Latin word that has not been directly borrowed from its Greek comparandum $\kappa\upsilon\pi\acute{\alpha}\rho\iota\sigma\sigma\omicron\varsigma$.

3.3.1.2 Latin *-undo*

Likewise on the basis of placenames, matches between Gk. $-\nu\theta\omicron\varsigma$ and *-anda* in Asia Minor led the Greek suffix to be early on considered a relic of a Pre-Greek substrate (cf. Kretschmer 1896: 402-5, but already Pott 1853: 451). Like $-(\sigma)\omicron\varsigma$, attempts have been made to give it an IE origin (cf. discussion in Kroonen fthc.; additionally, Finkelberg 2006: 52 and West 2007: 8 consider it Luwian). And like $-(\sigma)\omicron\varsigma$, it is still widely considered to be of non-IE Pre-Greek origin (cf. Beekes 2014: 37, Kroonen fthc.).⁵¹⁰

The representation of the Pre-Greek suffix in Latin is difficult to analyze. Firstly, there are several cases of Greek lexemes with the suffix $-\nu\theta-$ that have Latin comparanda in which the suffix is not present; even when something similar to it is present in other non-Greek branches. The best and most curious example is that of Gk. $\omicron\rho\beta\omicron\varsigma \sim \epsilon\rho\acute{\epsilon}\beta\iota\nu\theta\omicron\varsigma$, beside which PGM. **arwīt-* < QPIE **orw-īd-*, Arm. *ařowoyt* < QPIE **HrVb^h-oud-*, and maybe even Iranian forms < QPIE **Hreb^(h)-e/ont/d^(h)-* contain a suffix that looks comparable to $-\nu\theta-$ / $-\nu\theta-$. Lat. *ervum* however lacks the suffix completely.

In contrast to this, Lat. *bolunda* seems to suggest that *-und-* is the Latin reflex of Pre-Greek $-(\nu)\nu\theta-$, as compared with Gk. $\omicron\lambda\nu\nu\theta\omicron\varsigma$ whence it almost certainly cannot be a regular borrowing. There are two other cases of obscure Lat. *-und-*, but they remarkably do not correspond to Greek forms with $-\nu\theta-$. Lat. *harundō* has no secure relatives while Lat. *hirundō* corresponds to Alb. *dallëndyshe* (whose root ends in QPIE **(o)nt/d^(h)-*) but Gk. $\chi\epsilon\lambda\acute{\iota}\delta\acute{\omega}\nu$ (with no trace of the nasal or aspirated consonant). Perhaps this suggests

⁵¹⁰ Against a Luwian origin, Morpurgo Davies (1986: 120) rightly wonders why it surfaces as $-\nu\theta-$ rather than $-\nu\tau-$ or $-\nu\delta-$.

that the Latin words are borrowed independently of Greek (i.e. from a relative of Pre-Greek on the Italian peninsula), but they are so few that this is unlikely. Additionally, Kroonen (fthc.) notes the peculiarity that Lat. *menta* ~ Gk. μίνθη and PRom. (Southern Italian) **plenta-* ~ Gk. πλίνθος potentially present conflicting outcomes of the same suffix.

3.3.1.3 Latin *-ara

Beekes (2014: 32), following Furnée (1972: 256-7, esp. 257 fn. 36), gives evidence of -ar- as a Pre-Greek suffix. Gk. κισσός ~ κισσαρος, κιθάρα has a Latin comparandum in *hedera*. The latter reconstructs to a pre-form identical to that of κιθάρα but for an *e* ~ *i* alternation, interestingly the same as that between Lat. *cupressus* ~ Gk. κυπάρισσος, Lat. *hirundō* ~ Gk. χελιδών, Lat. *menta* ~ Gk. μίνθη and PRom. **plenta-* ~ Gk. πλίνθος above.

3.3.1.4 Conclusion on Pre-Greek Suffixes in Latin

The paucity of the reflexes of these suffixes in Latin lexemes in comparison with their frequent occurrence in Greek seems to point away from a Pre-Greek-speaking population in Italy. If this is the case however, then the words must have been transmitted from Greek to Latin via an intermediary language, perhaps via sea trade. This is plausible for *cupressus*, *bolunda*, and **plenta-*, which may be considered items of economic importance. It seems more difficult to understand this for *harundō*, *hirundō*, *menta*, and *hedera*. The first two, as mentioned, do not certainly contain a Pre-Greek element. For the Greek comparanda of *menta* (and **plenta-* for that matter), although they end in a sequence containing -ivθ-, it is difficult to confirm that they indeed attest to the -ivθος suffix (cf. Kroonen fthc.). The explanation of *hedera* remains elusive. It is *a priori* plausible that a language related to Pre-Greek would have been spoken on the Italian peninsula (i.e. that it was not exclusive and unique to Greece), but it seems dubious to confirm its presence on the basis of one word. In the end, what we find might be a combination of words from a Mediterranean substrate language borrowed independently by Latin and Greek and Pre-Greek words mediated into Latin by a (and in some cases surely the same) substrate language.

3.3.2 The *a*-Prefix

Schrijver (1997: 296-7, 307-12) collected examples of a phenomenon in which some languages attest to a lexeme with an initial *a-* against others without the *a-*. In the best examples, the *a*-prefixed words shows a concomitant reduction in vocalism. The first three he accepted were **amsl-* (OHG *amsla*) ~ **mVsl-* (Lat. *merula*, W *mwyalch*) ‘blackbird’, **alaud-* (Gaulish *alauda*, documented in Latin) ~ **laiwaz-* (OE *lāwerce*) ‘lark’, and **arud-* (OHG *aruz*) ~ **raud-* (Lat. *rauda*) ‘ore’. Further examples exhibit the vacillating presence of an initial *a-* but lack the vocalic reduction, perhaps representing leveling of the vocalism: W *garr* ‘leg, shank’ ~ Hsch. ἄκαρα ‘leg’, PCelt. **strabi-* (OIr. *straiph*, *sraib* ‘sulphur’, *sraiftene* ‘lightning’) ~ PGk. *(*a*)*st(e)rVp-* (ἀστεροπή, στεροπή,

ἀστραπη, etc. ‘lightning’). But Schrijver was unsure of the many examples given by Furnée (1972: 368-74) of vacillating *prothetisches a* belonged to the same phenomenon.⁵¹¹ Iversen and Kroonen (2017: 518) added several further examples:

- *r(̄)ʰd- (Gk. ῥωδιός, Serb. *róda* ‘stork’) ~ *ard- (ON *arta* ‘teal, garganey’, Lat. *ardea*, Gk. ἐρωδιός, ἀρωδιός ‘heron’)
- *reʔk- (ON *rækja* ‘shrimp’) ~ *arʔk- (Ru. *rak* ‘crayfish’, Lith. *ėrkė* ‘tick, mite’)
- *gedl- (Gk. γέλγις ‘garlic’) ~ *agd- (ἄγλις, Lat. *allium* ‘garlic’), cf. already Kroonen (2012b)
- *raiʔs- (Lith. *riėšas* ‘nut’) ~ *arʔs- (Ru. *orėx*, Alb. *arrė* ‘walnut’)
- *sak-/*se-sk- (OE *secge*, OIr. *seisc* ‘sedge’) ~ *as(a)k- (Ru. *osóka* ‘sedge’)
- *setr- (OHG *sturio* ‘sturgeon’) ~ *as(e)tr- (Ru. *osėtr* ‘sturgeon’), cf. already Kroonen (2012a: 240, 256)
- *rap- (Gk. ῥάφ/πυς, Lat. *rāpum*, OHG *ruoba* ‘turnip’) ~ *arb- (W *erfin* ‘turnips’)
- *sker- (OHG *sker* ‘mole’) ~ *askr- (Gk. ἄσκαρίς ‘worm’, Ru. *jáččer* ‘lizard’)

Schrijver (2018: 362) adds *aleil- (Hitt. *alēl*) ~ *leil- (Lat. *līlium*, Gk. λείριον) ‘lily’.

Schrijver (2017: 362) suggested that the *a*-prefix and its effects on vocalism was similar to that of the Hattic nominal prefix *ha-*,⁵¹² as part of the evidence that the language of the first European farmers may have been related to Hatto-Sumerian. Šorgo’s (2020: 457) explanation on the other hand is that it is not actually an *a*-prefix at all. Instead it is a subset of a broader substrate feature of accentually conditioned alternation from an original (C)VCVC ~ (C)VCVC.⁵¹³ As examples of other types of cases, he provides:

VCC- ~ VCVC- (PGm. *arwīt-, Lat. *ervum* ~ Gk. ὄροβος, ἐρέβινθος)

CVCR̄- ~ CCVR̄- ~ CCR̄- (Gk. μόλυβδος etc. ~ PGm. *blīwa- ~ Lat. *plumbum*)

CVCC ~ CCVC (PGm. *maldjo- ‘orach’ ~ Gk. βλίτον, βλήτων ‘amaranth’), (PCelt. *se/immr- ‘clover’ ~ PGm. *smai/ēr- ‘sourgrass’), (PGm. *waldō- ~ Lat. *lūtum* < *uloūt-/*ulūt- ‘dyer’s rocket’)

CVCC ~ CVCV (PGm. *samda- ~ Gk. ἄμαθος ‘sand’)

CVR̄CC ~ CR̄CVC (PGm. *waizda- ~ Gk. ισάτις ‘woad’)

If Lat. *ervum* ~ Gk. ὄροβος is a part of this phenomenon, then the numerous cases of the

⁵¹¹As a general substrate feature, cf. e.g. Alessio (1944a: 149, fn. 242), Hubschmid (1950b: 291). Furnée (1972: 368 with further lit.) finds this phenomenon distributed from Iran and Mesopotamia, through Asia Minor, Greece, and the Balkans, to the Western Mediterranean.

⁵¹²A similar source had been suspected of producing the prothetic *a*- in some toponyms in Asia Minor (e.g. Kretschmer 1933a: 86).

⁵¹³Recently, Schrijver (fthc.) along similar lines has instead proposed that the source of the *a*-prefix alternation is an East Caucasian language.

a-prefix only coincidentally have *a*-vocalism. Potentially in favor of this is the case of the elm word (Lat. *ulmus*), which Schrijver (1997: 311) was uncertain about adding because Germanic has both *e*- and *a*-vocalism of the prefix (and Latin probably has *e*- or *o*-vocalism). Perhaps the Greek form ἐρῳδιός with its initial *e* shows this pattern as well. If Kroonen's addition of PGm. **sturja/ōn-* is correct, it should be noted that the Slavic forms all regularly reconstruct to *e*-vocalism. Thus this would be another example of an *e*-prefix. If the vowel does not have to be *a*, then Šorgo's proposal seems attractive. It is strange that there are so few cases with non-*a* vocalism, but perhaps this has something to do with the distribution of vowels in the substrate language(s).

Regardless of which explanation is correct, there are indeed several Latin lexemes that participate in the phenomenon of *a*-prefixation. From the dataset, *ardea*, *merula*, *rāpum*, and *raudus* (as well as *ervum* and *plumbum*) have already been mentioned. *Ālium* fits, with some difficulties, into the pattern including Gk. γέλυγς ~ ἄγλις, as does *excetra* to the pattern including Baltic, Slavic and Germanic sturgeon words. I am not sure that the *līlium* lexeme belongs to this group, because it seems plausible that the initial *h* of the Egyptian comparandum, if it represents the source, could have affected the anlaut in the borrowings.

To the list can be added Lat. *tilia*, whose Celtic comparanda from PCelt. **axtlV-* can be reconstructed further to **aptlV-* (cf. the preserved initial cluster in Gk. πελέα). Lat. *pīrum* ~ Gk. ἄπιον might represent an example where the vocalism of the root has been leveled. Restricted to Latin and Greek (and perhaps languages much further East), its distribution makes its relevance to this widespread European phenomenon uncertain. Finally, if the vowel of the *a*-prefix indeed does not have to be *a*, then an example from the uncertain cases might include Lat. *īnsula* (s.v.) against Gk. νῆσος. The other evidence of a non-*a* vowel involves *e*, but *īnsula* (with regular lengthening from **in-*) has perhaps undergone the same change **en-* > *in-* that occurred in the preposition, preverb, and negative prefix.

3.3.3 The Velar Suffix⁵¹⁴

Numerous Latin words ending in *-ax*, *-ex*, *-ix*, and *-ox* lack a good IE etymology and fall into semantic categories including animals, biting insects, trees, plants, and body parts (cf. Ernout 1946: 133-63, Leumann 1977: 375-6, Weiss 2020: 326-7). While some lexical bases to which this suffix is attached can be argued to be inherited (Ettmayer 1926: 23, Ernout 1946, Specht 1947: 40-1, Martinet 1955, Olsen 2009, Matasović 2016, Weiss 2020: 326-7), many of the etymologically obscure forms have long been suspected of being non-IE in origin (Ettmayer 1926: 23; Terracini 1929: 212-14; Bertoldi 1937: 157; Gerola 1942: 364; Alessio 1944a: 104; Hubschmid 1953: 84, 1960: 97; Leumann 1977: 375; DV 299). Likewise, numerous Greek words in *-αξ*, *-ακος* belong to similar semantic categories and are also suspected of being loanwords (Nehring 1925,

⁵¹⁴ This section follows and is developed out of Wigman (fthc.).

Chantraine 1933: 376-83, Beekes 2014: 32, 44).

A suffix **-k-* is reconstructible for PIE in many forms and functions (cf. Brugmann 1906: 472-506), but an interesting pattern emerges as concerns especially the *-ix/-ex* suffixes of Latin.⁵¹⁵ While thematic velar suffixes and athematic velar suffixes with a long vowel have good parallels in other branches and have relatively well-understood sources (cf. Wigman fthc.), the athematic, short-vowel *-ix/-ex* suffixes are often isolated to Latin. There is little lexical overlap with other branches. One case might be Lat. *natrix* 'sea serpent' against OIr. *nathir* < **natrik-* but this is complicated by the Brythonic forms reconstructing to **natrī-* and Germanic forms similarly lacking the velar element. In Greek, Kölligan (2017: 369-70) suggests that no cases of Greek formations in *-αξ* beside a velar element in another branch need be interpreted as anything but individual parallel developments. There are indeed remarkably few cases of Latin *-ex* corresponding to Gk. *-αξ* (suggesting that Lat. *-ex* is, at least in some cases, the result of weakening from **-ax*). This includes Lat. *mūrex* ~ Gk. *μούραξ* 'murex' and, rather uncertainly, Lat. *īlex* ~ Macedonian (Hsch.) *ἰλαξ* 'holm oak'.⁵¹⁶ Crucially, this same correspondence occurs in one lexeme that has a Germanic comparandum where the velar is voiced (1). This irregular alternation between Germanic **k* < QPIE **g* against Lat. **k* < QPIE **k* occurs in two further lexemes (2, 3):

- (1) Lat. *sōrex* < QPIE **s(y)ōrVk-* ~ Gk. *ῥοραξ* < QPIE **syo/urak-* ~ PGM. **s(w)ur(V)ka-* < QPIE **sur(V)g-*
- (2) Lat. *fulica* < **b^hulVk-* ~ SGael. *bolachdan* < QPIE **b^(h)o/ula/oK-* ~ OHG *belihha* < QPIE **b^ha/olig-*
- (3) Lat. *filix, felix* < QPIE **b^helik-* ~ Gk. *βλήχρον, βλήχρον* < QPIE **blēg^h-n/r-* ~ PGM. **brekna(n)-* < QPIE **b^hreg-n-*

In (3), the Greek and Germanic forms have an additional suffix added, which may be responsible for the aspiration in Greek (see §3.3.4). This irregular alternation within the suffix itself indicates that some examples of Lat. *-ix/-ex* have been borrowed. On the other hand, this correspondence pattern is not the only source of Lat. *-ix/-ex*. Lat. *salix* beside PCelt. **salik-*, PGM. **salihōn-*, and potentially Gk. (Arcad.) *ελικης* can be reconstructed to ablaut grades of a root **selH-ik-*. All cases reconstruct to the same unvoiced velar in the suffix, and there is thus no positive evidence indicating that this lexeme is not inherited. That not all examples of Lat. *-ix/-ex* are from the same source makes it difficult to determine whether the *sōrex-fulica-filix* type truly represents a suffix or whether it is simply the final velar of a disyllabic substrate root. Some weak evidence against suffix status is that other suffixes have been added (e.g. the *-n-* suffix of the Greek and Germanic comparanda of *filix*). I suggested a relationship between Lat. *sīl* 'ochre'

⁵¹⁵ Due to the possibility of leveling of the vocalism from the oblique, some cases of *-ix* might represent originally the same suffix as *-ex*.

⁵¹⁶ Lat. *latex* 'liquid, fluid' ~ Gk. *λάταξ* 'drop of wine' looks like an example, the Greek oblique forms have *γ* rather than *κ*.

and *silex* ‘flint’, but without a Germanic comparandum, it is difficult to know if this represents the same “suffix”. Such is also the case for Lat. *cerrus* ‘holm oak’ against Romance forms like OProv. *garric* ‘kermes oak’.⁵¹⁷

Beyond these cases, there are few lexemes with a velar element in what looks like a suffix position that can be shown to be of non-inherited origin due to irregular correspondences in the root material. Such is the case for Lat. *calix* ~ Gk. κύλιξ. The velar element is present in U *skalçeta* but is lacking in Hsch. σκάλλιον, though the appurtenance of the latter two is not secure. Nor is it for Skt. *kalása-*, which would indicate that the velar involved is a palatovelar, a feature that is not normally visible due to attestations in *centum* languages. For Lat. *tamarix* ~ Gk. μύρικη, the Greek form is thematic as it is in Lat. *paelex* ~ Gk. παλλακή (also attested is *πάλλαξ*, but EDG 1147 suggests it is a backformation from *παλλακή*) ~ PCelt. **φα/erikā-*. The latter Celtic form reconstructs to QPIE **k*, making this group more similar to the *salix* type of velar suffix than the *sōrex-fulica-filix* type where the Celtic comparandum of *fulica* reconstructs to a geminate **kk*.

In Lat. *calx* ~ Gk. χάλιξ, the relationship to the velar suffix(es) is unclear. Given the productivity of *-ix* in Latin, it is highly unexpected that the vowel would be syncopated. But is it likely that the Greek form has added an anaptyctic vowel? Finally, the suffix *-ix*, which can have a good native etymology (the inherited *devī*-suffix **-ih₂-*, probably with the addition of **-k-* [Schrijver 1991: 148-54, Weiss 2020: 325], though arguments can be made that it is the result of laryngeal hardening [Olsen 2009]), seems to have been added to Lat. *rādīx* and Gk. ῥᾶδιξ in light of the fact that no other comparanda have the suffix. It is curious that the exact same suffix should be added, but the semantic difference between the Latin and Greek forms makes it unlikely that either is loaned from the other.

In the end, there is evidence for a Lat. *-ix/-ex* of non-inherited origin, as well as cases of Lat. *-ix/-ex* attached to lexemes of non-inherited origin. It is not possible in all cases to know if 1) the element is the same everywhere it appears (and in fact, there are at least two separate sources) and 2) if the element functioned as a suffix in the language where it originated. It seems easy enough for Latin to have analyzed it as a suffix upon borrowing on comparison with inherited suffixes of the shape **-Vko-* and **-V̥k-* and nominal compounds in *-fex < facere*.

3.3.4 The *n*-Suffix

Kuiper (1995: 80) noted for PGm. **baunō-* that “suffixation of *-no-/-nā-*, whatever their origin may be, occurs in several loanwords.” Unlike the above case of the velar suffixes, due to the nature of the dental nasal in the daughter languages, there is no possible phonological alternation that could suggest irregular correspondences; only the vacillation of its presence. Proto-Indo-European is reconstructed to have had a deverbal suffix of this shape (Fortson 2010: 131). Leumann (1977: 320) also notes that stems of

⁵¹⁷ FEW (II: 411) considers this an example of the “wohl ebenso iberisches” suffix *-ico-*.

this shape can be the result of derivation of inherited *n*-stems. Given that there was an inherited *n*-suffix, its presence on a word cannot be any indication of non-native origin. But when it appears on words of non-native origin, we must wonder how it appeared there. Was its usage transferred from inherited cases as a strategy for nativizing foreign lexical material? Or was it borrowed in place because the substrate language(s) also had a nasal suffix?

There are a few indications that this is not an inherited *n*-suffix. It is a derivational morpheme, but between the comparanda that contain it and those that do not, there is little derivation in meaning. The vacillation of its presence itself is suspicious; if it is deverbal, then the forms without it should still be verbal. There is no way to know whether the borrowed lexeme had a verbal function (and therefore that the *n*-suffix could represent the addition of native morphology), but the semantic similarity of the forms without it to the forms with it suggest that they were borrowed in nominal function. Unless we propose that the substrate language(s) had an IE-esque *n*-stem construction, this cannot be the explanation either. This is a potential example of IE morphology behaving in non-IE ways turning out not to be IE after all. But one point merits caution: In Latin, an additional source of words with an *n*-suffix are the material adjectives in *-(i)no-* (either **-no-* or always-syncopated **-ino-*), a derivational shape that also occurs in Greek *-ῖνος*.⁵¹⁸ In certain cases, a substantivized material adjective could come to be synonymous with the underived nominal base.

The words that can be determined to be of non-IE origin due to other factors and which contain an *n*-suffix are:

With <i>n</i> -suffix	Without <i>n</i> -suffix
Lat. <i>alnus</i> < PItal. <i>*alsno-</i> PBalt. <i>*(a)el(i)snio-</i>	PGm. <i>*aluz</i> , <i>*alis/zo-</i> , <i>?*elustrō-</i> PSlav. <i>*o/elbxa-</i>
Lat. <i>avēna</i> < PItal. <i>*awe(C)snā</i> ?West Uralic <i>*we/äšnä</i>	PSlav. <i>*ovъsъ</i> PBalt. <i>*(a)vižaʔ-</i> ?PGm. <i>*hab(a)zan-</i>
Lat. <i>urna</i> < PItal. <i>*ur(k)nā</i>	Lat. <i>orca</i> , <i>urceus</i> Gk. <i>ὄρχη</i>
PGm. <i>*hadnō-</i>	Lat. <i>catulus</i> < PItal. <i>*kate/o/ulo-</i> PGm. <i>*hada/e/ulō-</i> PCelt. <i>*kadVlot-</i>
PGm. <i>*baunō-</i>	Lat. <i>faba</i> < PItal. <i>*b^habā</i> PSlav. <i>*bòbъ-</i> PBalt. <i>*babō-</i>
Gk. <i>βλήχρον</i> , <i>βλήχρον</i> PGm. <i>*brekna(n)-</i>	Lat. <i>filix</i> , <i>felix</i> < PItal. <i>*felik-</i>

⁵¹⁸ Lat. *-īnus*, Gk. *-ῖνος*, and PGm. *-īnaz* have a long vowel that would not syncopate in Latin.

Gk. δάφνη, δάχνα	Lat. <i>laurus</i> < PItal. * <i>lauro-</i>
------------------	---

Table 3.32 Alternating presence of an *n*-suffix

The suffix looks also to be present in Lat. *alaternus*, but Cretan ἐλαίτρινος suggests that the *-rno-* sequence is secondary from *-rino-*. One Celtic comparandum for Lat. *rādīx* (OIr. *frén* < **wridnā*) has an *n*-suffix as opposed to several other forms, even within Celtic, that do not have it. Therefore I am uncertain if it represents the same phenomenon. But a very similar case, albeit without a Latin comparandum, is the holly word. It has an *n*-suffix in PCelt. **kolinno-* < QPIE **kolis-no-* that is lacking in PCelt. **kelastr-* < QPIE **kelas(-)str-* and in all the other comparanda (Germanic, Greek, Armenian, and Romance forms including Sardinian, cf. van Sluis fthc.).

Two other forms that have been classified as uncertain because they have no other features pointing to a non-native origin (or competing plausible etymologies) include Lat. *acer* against PGm. **ahurna-* (with the suffix, and PGm. **ah(i)ra-* without it) and Lat. *pīnus* against either Lat. *pix*, *picea* or Gk. πίτυς (or both, if this is a *t ~ k* alternation, see above).

Between nearly all sets of words, the meaning is identical. The explanation of a syncopated adjectival **-ino-* being substantivized does not work for the cases where it is not Latin that attests to the suffix (like PGm. **baunō-*). In the case of *avēna*, it is even present on the Uralic forms, suggesting that they borrowed the word *with* the suffix (and from the source, as both Baltic and Slavic lack the suffix showing they cannot have been the source of it in Uralic). This all seems to indicate that in these words, the suffix is of non-IE origin. And unlike with the velar suffix, there is enough vacillation amongst the comparanda to suggest that it functioned as a suffix (in that it could be added or removed) in the source language.

The Greek forms attest to an interesting pattern. In all cases where Greek attests to an *n*-suffix, when this suffix is attached to a stop, that stop is aspirated (κλίχνη; βλήχρον; δάφνη, δάχνα; ἀράχνη). This pattern occurs elsewhere within Greek, for example between Gk. πέλιξ and πελλίχνη ‘bowl’. This led Beekes (2014: 37) to follow Furnée (1972: 132, fn. 64, 65) in suggesting that the *n*-suffix may have been responsible for aspirating a Greek κ. But Indo-European also had derivations in **-sno-*, potentially from diverse sources, which, after a stop in Greek, produced aspiration (cf. **louk-sn-eh₂* > Lat. *lūna* ‘moon’, Av. *raoxšna* ‘lantern, bright light’, **luk-sn-o-* > Gk. λύχνος ‘lamp’). Mawet (2008: 43) notes that the resulting ‘aspirated consonant + νη’ spread as a derivational pattern within Greek. Thus Gk. κλίχνη may represent a regular derivation of Gk. κύλιξ. In part for the same reasons as given above for the *n*-suffix in general, I do not think that all of these cases, especially when they occur on words that can be demonstrated to be of non-IE origin, can be explained as an inherited phenomenon.

In βλήχρον/βλήχρον, the aspiration occurs not only before the **n* but also the **r* in the alternate form. (Notably, the *n*- and *r*-suffix distribution is also found between Lat.

laurus ~ Gk. δάφνη, δάχνα). In the Germanic comparandum **brekna(n)-*, no **s* occurs. And in fact, as opposed to Greek, which, as just noted, always aspirates the stop before this suffix, the three cases of this suffix in Germanic (**hadnō-*, **baunō-*, and **brekna(n)-*) where an **s* would be preserved never attest to an **s*. One might propose that this pattern of ‘aspirated consonant + νη’ became ubiquitous in Greek and that perhaps the *r*-suffix of βλήχρον replaced the *n*-suffix that had already triggered the analogical change of **-κν-* > *-χν-*. But the pair Gk. ἀράχνη ~ Lat. *arāneus* shows that not all cases are analogical. Here the Latin long vowel shows that both forms go back to a true cluster **-ksn-*. The βλήχρον/βλήχρον case is further interesting because it looks like the *n*-(/r)-suffix has been added to a velar ‘suffix’, attested in its plain form in the Latin comparandum *filix/felix*. The Greek forms could hypothetically represent the *n*-suffix added to a nominative formation in *-ξ*. This certainly cannot be the case, but it raises the question of how to segment this cluster. Does βλήχρον represent **blēk-sno-* or **blēks-no-*? Given that Germanic attests to the same suffix without the sibilant, perhaps the latter is more likely. In this case, some quality of the foreign velar was interpreted in Greek as an affricate.

Evidence that Latin has interpreted a foreign velar as an affricate is found in *avēna* < **awe(C)snā-*. Against the reflex of **k̥* in Slavic and **g^h* in Baltic, all forms may simply have been borrowed as some sort of sibilant (cf. West Uralic **we/äšnä* and, if indeed related, PGM. **hab(a)zan-*). But **Ks* is also possible for the Latin. Direct evidence of this is perhaps found in Lat. *pīx* (whose status as a loanword I consider uncertain, s.v. *pīx*). If it is from the same source as Georgian *pič’vi* ‘pine’ (borrowed into Armenian as *p^ciči*, cf. Furnée 1979: 28 for details), the latter has an affricate. Lat. *pīnus*, thought to represent either **pik-sno-* or **pit-sno-* might therefore represent **piks-no-*, the reflex of this foreign velar (which would then also have produced the sibilant of Alb. *pishë*). Some of the *-ix/-ex* suffixes thus may not have just entered Latin as **-Vk-* but rather **-Vks*, which was nativized into the consonant stem declension.

In any case, at least one of the substrate languages of Europe seems to have had an *n*-suffix. The trio Lat. *filix/felix* ~ Gk. βλήχρον/βλήχρον ~ PGM. **brekna(n)-* shows that it and the confirmed irregular velar “suffix” (§3.3.3) occurred in the same language. It may well have occurred in more than one language, seeing as a few of the examples (Lat. *laurus, urna*) are words with a Mediterranean distribution.

A discussion by Kretschmer (1921: 277-8, fn. 1) surrounding the potentially non-IE origin of Greek ethnonyms in *-ηνός* (cf. also Beekes 2003: 30) concluded, as Nehring (1925: 189) similarly would for the *-αξ* suffix of Greek, “Aus allem dem folgt, daß sowohl die indogermanischen wie die nichtindogermanischen Sprachen Kleinasien mit *n*-Suffix gebildete Ethnika besaßen, ein Zusammentreffen, das nicht verwunderlich ist, wenn man sich erinnert, daß auch das Etruskische mit dem Lateinischen in mehreren Suffixen zusammentrifft.” And as he would further note, Etruscan has a suffix *-na*. In fact, in Etruscan, the suffix *-na* is extremely frequent, perhaps one of the most productive suffixes of Etruscan (Steinbauer 1999: 121). If we are looking for a non-IE language

with a systematic *n*-suffix, and if Etruscan represents one of the substrate languages of Europe, then perhaps Etruscan might represent a relative/descendant. But within Etruscan, *-na* seems to produce derivational changes in meaning: the substantives on which it occurs are seemingly substantivized adjectives, cf. *θi* ‘water’, *θina* ‘(water) jug’ < *‘pertaining to water’ (Steinbauer 1999: 107, Wallace 2008: 53). And it was in part the lack of derivational changes in meaning that led me to propose that the *n*-suffix on several substrate words was not of IE origin. But this is a mystery we should expect to have to solve. If the non-IE *n*-suffix really was a suffix, which its vacillating presence amongst the comparanda seems to suggest, then it must have had some derivational function in the source language. Perhaps derivational differences in semantics were bleached by the borrowing process, but it is also possible that the suffix’s function in the source language was more subtle than that of the Etruscan *-na* suffix.

3.3.5 Reduplication

Latin attests to several types of reduplication in nouns, many of which (like the archaic pattern of *fiber* < **b^he-b^hru-*) are inherited (cf. André 1978, Weiss 2020: 287). But Latin, as well as many other Indo-European languages, also has more isolated instances of reduplication. Alessio (1943) collected several Latin and Greek words with CV-reduplication (with *e* and *i* vocalism of the reduplicated syllable) in nouns, which he suspected to be of Mediterranean origin and used this to propose that the Mediterranean substrate therefore had this type of reduplication.⁵¹⁹

As per his methodology, many of the cases he finds are isolated; as per my methodology, these get classed as uncertain. I therefore consider many of the words he proposes as evidence to be non-diagnostic (*cicāda*, *cicōnia*, *cicūta*, *gigarus*).⁵²⁰ The isolated words are suspicious because, as DV (500) notes, Latin frequently uses reduplication in affective words, explaining *cicāda* and *cicōnia* as likely onomatopoeic (DV 112, 113). I am generally suspicious of labelling words as ‘affective’ or ‘onomatopoeic’. André (1978) was also dissatisfied with these labels and set out to better classify the cases, with comparative material from a much wider typological perspective. He classes Latin reduplicated formations, not all necessarily of IE origin, as *impressifs* of sound (like *bambalō* ‘stutterer’, *cucurru* ‘the cry of the cock’, and words in *gurg-* and *garg-* relating to swallowing and the throat), of movement (like words for back-and-forth in *pal-*, the repetitive circular motion of *furfurāculum* ‘wood-boring tool, auger’), of form (like *circus* ‘circle’, several round fruits and vegetables including *cucurbita*), and of quantity

⁵¹⁹ André (1978: 12) calls into question what exactly it is about the phenomenon that makes it Mediterranean.

⁵²⁰ There are several as well that I simply did not treat. This includes (poorly attested) *biblax* ‘rododaphne’; *cicendula*, *cicindēla* ‘firefly’ for which he doubts the connection with *candēla* ‘candle’; *cicimalindrum*, *cicilindrum* ‘a made-up spice name in a Plautus play’, *giger* ‘wild parsnip’, *siser* ‘skirret’ potentially borrowed from Greek σίσαρρον ‘parsnip’ (though André 1978: 49 agrees on a Mediterranean substrate origin); *cicūlus* for *cucūlus* ‘cuckoo’, onomatopoeic; *sisarra* ‘sheep older than one year’, cf. Calabrese *sarra* ‘fat old woman’, perhaps a non-IE word after all; *vīverra* ‘ferret or weasel’, which DV (685) treats as inherited.

(like *grex* ‘herd, crowd’, *populus* ‘people’, and *calculus* ‘pebble’), as well as nursery words. From Alessio’s list, he classifies *cicōnia* and *cicāda* as expressive of sound (along with *cicūta* for its use in making flutes), and *cucurbita* and *gigarus* as expressive of form (the latter apparently due to its root; though while this could apply to *giger* ‘wild parsnip’, it does not feel particularly fitting for *gigarus*). Given that he includes cases like *siser* ‘skirret’, which he believes to be of non-IE origin, under reduplicated formations due to impression of form, a word’s appearance in his list does not necessarily mean that it is of inherited origin. If non-IE words show the same motivations for reduplication that inherited words do, one might wonder if the tendency within Latin for expressive (André’s “impressif”) formations was influenced by language contact. But given the frequency of such a phenomenon in the world’s languages, an explanation like that does not seem necessary.

On the other hand, it is likely that substrate languages were a source of multi-syllabic roots. This is for instance probable in Lat. *arāneus* ~ Gk. ἀράχνη and for the Greek comparanda of Lat. *ervum* (ἐρέβινθος, ὄροβος). Some of them may have resulted in what looks like reduplication. In fact, André (1978: 12) makes this explicit in his analysis: since the two consonants of an IE root must differ in nature, roots with shapes like **bab-* and **pip-* are automatically suspected of being reduplicated “except for substrates (of languages whose root may be of a different structure), like *sisarra*, loans (*popina*, *lalisio*), phonetic treatments (*cocus*, *barba*, *quinque*).”⁵²¹ It also means that, without positive evidence one way or the other, an isolated word with a reduplicated shape could be classed as more or less permissible Latin reduplication or a non-IE root shape produced by a substrate disyllabic root.

Two words stand a chance of representing some sort of reduplication, but neither is simple. Lat. *cucumis* has Greek, Armenian, and probably Slavic comparanda. Arm. *sex* reconstructs to something like **kek^h-* where the consonants are not the same. Thus, despite other forms like Hsch. κύκουον looking like reduplication, we cannot ascertain that it would have existed this way in the source language. Lat. *cucurbita* looks like a reduplicated form a group of lexemes attested in Sanskrit (*cirbhaṭī*, *carbhaṭa*, *cirbhiṭa*) but this might be due to chance. A geographically closer match is PGm. **hwerhwetjō-*, but its relationship to *cucurbita* is not straightforward. Is it a metathesis of **hwehwert-* (thus implying the reconstruction of **kuko/urd^h-* for Latin with a **d^h ~ *t* alternation) or does it correspond to the base *-curbit-* (implying the reconstruction **kuko/urb^(h)-Vt-* for Latin with an otherwise unattested labial ~ dental alternation)?

Lat. *cicer* looks like a reduplicated formation as do its Armenian and (possibly) Albanian comparanda. But because they can reconstruct to the same pre-form, it cannot be ruled out that the formation is old. Even if it is considered a substrate word, without a secure simplex form, it cannot be used to confirm that the source language had morphological

⁵²¹ “Exception faite des substrats (de langues dont la racine peut être de structure différent), comme *sisarra*, des emprunts (*popina*, *lalisio*), des traitements phonétiques (*cocus*, *barba*, *quinque*).”

reduplication (cf. this as a requirement for confirming reduplication in André 1978: 13).

On the evidence of Fal. *haba*, Lat. *faba* does not represent true reduplication as its initial consonant is the reflex of a voiced aspirate but its medial *b* must be from a plain **b*. Lat. *pōpulus* is probably related to Slavic and Baltic words, none of which can securely represent reduplicated formations. Lat. *cicōnia* is isolated, but is attested once as *cōnia* (also perhaps in the Hesychius gloss γνίς), and a few more cases of simplex-reduplicated alternations seem to occur within Greek (τιθύμαλος, θύμαλος ‘euphorbia’; κίκνωψ, κνώψ ‘wild beast’; σέσηλος, σελάτης ‘snail’, Alessio 1943; Beekes 2014: 27 on potential reduplication in Pre-Greek). But this does not seem like certain enough evidence to propose that the substrate language had reduplication.

3.3.6 Morphological Conclusions

There are a very limited number of Latin lexemes of non-inherited origin that seem to contain suffixes otherwise widespread in the Greek substrate vocabulary. These likely reflect a combination of indirect loans from Greek (and its substrate languages) and Mediterranean substrate words borrowed independently into Latin and Greek. The phenomenon of *a*-prefixation (with concomitant vowel reduction) and the Lat. *-ix/-ex* suffixes that show irregular correspondences with comparanda are difficult to explain from an inherited perspective and meet the criteria of non-inherited features. The *n*-suffix seems likely to have its source in non-IE languages due to its vacillating presence on words with other irregular correspondences. These last three features appear with a quite widespread distribution. The implications of this will be explored in §4. Finally, though reduplication has been reported to be a feature of the substrate lexicon of Latin, it is difficult to confirm that disyllabic roots actually reflect reduplication in the source language.

4 Distribution Analysis

4.1 Introduction to Distribution Analysis

There is a strong pattern amongst the Latin words of non-inherited origin. Of the 109 words from §2.2, 93 have comparanda in Celtic, Germanic, Greek, or a combination of the three. Of the 16 words without comparanda in one or more of these three branches, there are:

- 15 words isolated to Italic:
 - 3 words without certain comparanda (*asīlus*, *asinus*, *casa*) that represent recent loans due to their intervocalic *s*,
 - 5 words without certain comparanda (*faex*, *farciō*, *focus*, *pampinus*, *tabānus*) that represent loans due to their invalid root structure,
 - 7 words (*arbutus*, *cerrus*, *genesta*, *lābrusca*, *lepus*, *sambūcus*, *talpa*) with irregular alternations clearly attested within Latin and/or Romance, and
- 1 word (*faber*) with a comparandum only in Armenian, probably a Wanderwort from an ultimately Hurrian source.

There are additionally two additional words amongst the 93 whose comparanda in Celtic/Germanic/Greek are not fully secure. *Carpinus* may only have secure comparanda in Slavic and *avēna* in Baltic and Slavic.

The 93 words with Celtic/Germanic/Greek comparanda pattern as follows (Figure 4.1). Words in non-italics are from §2.2.2.1 and those in italics are from §2.2.2.2. Words in green have no comparanda outside of the constellation of Celtic/Germanic/Greek where they are listed. Four words are preceded by a question mark because some comparanda are uncertain;⁵²² not enough to greatly influence the results.

⁵²² On *carpinus* and *avēna*, see above. The Celtic comparandum for *raudus* is not as secure as the Germanic one. *Tilia* has matches in Greek and Celtic, possibly Germanic. *Columba* is quite likely not actually found in Greek.

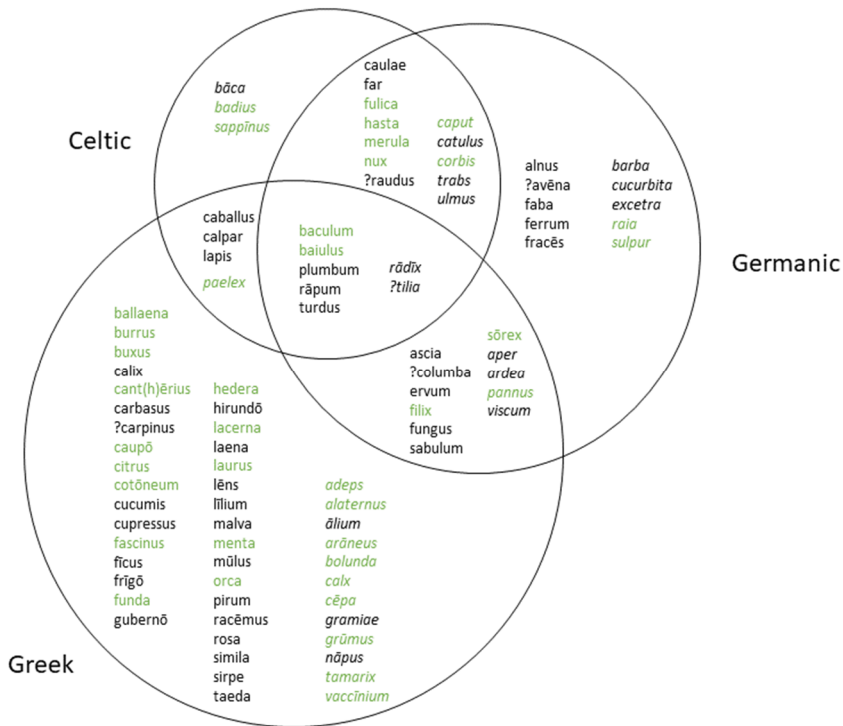


Figure 4.1 Latin words of loanword origin distributed by existence of comparanda in Celtic, Germanic, and Greek

The importance of Greek is immediately clear. 68 of 109 (62%) Latin words of non-inherited origin have Greek comparanda. For 46 of those 68 (68%), Celtic and Germanic do not attest comparanda. And 23 of those 46 (50%) are exclusive Latin-Greek isoglosses. Many of these must have been acquired in the Mediterranean. For example, 2 or 3 (*sirpe*, *taeda*, *?lēns*) have independent comparanda in Berber; 6 (*cupressus*, *ficus*, *laena*, *malva*, *rosa*, *simila*) in Semitic. The comparanda for *laena* and *simila* are well-attested in Semitic such that it might be the ultimate source of the Latin and Greek words. *Rosa* is also deeply entrenched in Semitic, but seems to have its ultimate source further East and cannot be the direct source of the Latin. *Cupressus*, *ficus*, and *malva* have a much more limited distribution in Semitic. There is a good chance that they represent independent borrowings. Hebr. *gofer* is a hapax, but attests to the lexeme behind Gk. κῶπαρισσοϛ without the suffix. The Latin form has the suffix but cannot be a direct borrowing from Greek. This suggests a mode of transmission like in Figure 4.2.

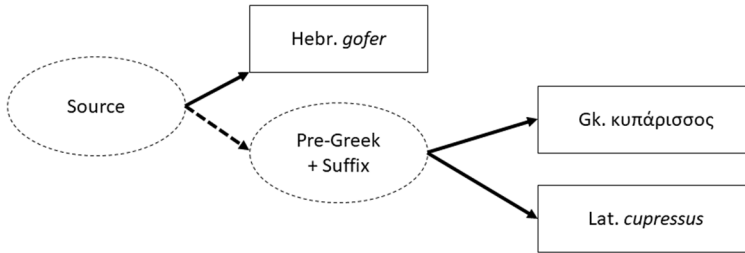


Figure 4.2 The transmission of Lat. *cupressus* and its comparanda

Hebr. *šiqmā* ‘sycamore fig’ is not as isolated, but occurs only in Hebrew and Aramaic (the latter likely being the source of Arabic *suqūm*, Klein 1987: 679). Lat. *ficus* mechanically reconstructs to PIE $*d^h$, but in light of Gk. *τῦκον* and Arm. *ʿuz*, there are two possibilities: 1) $*t^h$ can be reconstructed for Italic, Greek,⁵²³ and Armenian, suggesting that the word was borrowed with the same initial consonant or 2) the word entered Latin with $*θ$, which is also one of the Proto-Semitic sources of Hebr. *š*. In either case, the form borrowed into Latin (whether $*θ$ of $*t^h$) entered Latin’s treatment chain of the voiced aspirates, showing it was borrowed still into Proto-Italic.

On the other side of the spectrum are the 25 words for which there are Celtic and Germanic comparanda to the exclusion of Greek. 6 of these 25 (24%) are Latin-Celtic-Germanic isoglosses. This likely represents words borrowed in a part of Europe whose substrate languages did not affect Greek. The possibility of an Italo-Celtic subnode has interesting implications for this group, especially given the small number of Latin-Celtic isoglosses (*badius* and *sappīnus*, though *bāca* has only Berber additionally).

The other significant portion is at the center—the 7 words that have comparanda in Celtic, Germanic, and Greek. 2 of these 7 (29%) are Latin-Celtic-Germanic-Greek isoglosses. Others like *plumbum*, *rāpum*, and *turdus* are quite widespread. Their distributions are likely the result of different types of contact situations, as will be discussed below.

The picture that emerges based on the distribution of loanwords within these four daughter branches must be an oversimplification in many respects. Thus the following sections treat the stratification using further evidence. But a clear distribution is already apparent: the most recent stratum of loanwords in Latin was acquired in the Mediterranean. It is quantitatively the largest, and contains several of the words that cannot be reconstructed beyond Latin even. In contrast, the number of words with a pan-European distribution is much smaller. This suggests that these are the oldest loans,

⁵²³ Recall that Att-Ion. *σῦκον* points to the presence of a glide that would neutralize the difference between PGk. $*t$ and $*t^h$ (s.v. *ficus*).

borrowed at a time when the daughter languages were closer to one another followed by lexical replacement.

4.2 Partial Stratification Based on Phonology: The Most Recent Borrowings

Certain: *asinus*, *asīlus*, *casa*, *carbasus*, *ballaena*

Possible: *caballus*, *rosa*, *tamarix*, (*cāseus*, *gabata*, *nepeta*, *supparus*)

We know that PItal. **s* was rhotacized in Latin in the historical period, ca. 350 BCE (cf. Weiss 2020: 208). Thus Latin words with unrhotacized *s* not following a long vowel/diphthong must have entered Latin after rhotacism took place. Schrijver (1991: 252) points to *asinus* and *cāseum* to suggest that rural dialects that did not undergo rhotacism might be the source of some of these unrhotacized forms. Christol (1996) proposes a stricter version of this, where non-rhotacizing dialects preserved intervocalic *z*, sometimes borrowed back into Latin as *ss*. Geminated *s* is also the result of borrowed Gk. ζ, and it explains e.g. Plautine *nassus* otherwise *nāsus* ‘nose’ beside *nāris*. *Asinus*, *asīlus*, *casa*, and *carbasus* cannot be explained according to this stricter framework and would thus be post-rhotacism loans. Weiss (2020: 162) suggests that the *s* of *rosa* remained unrhotacized due to the dissimilatory effects of the *r*, like in *miser* ‘wretched’. *Cāseus*, given its long vowel, might attest to a simplified geminate rather than a true unrhotacized *s*; its analysis is hampered by a lack of secure comparanda.

Vowel weakening began to occur sometime after 500 BCE. The earliest change was the weakening in non-initial syllables of *a* to *e*, and in open syllables further to *i* (cf. Weiss 2020: 130). *Ballaena* should have weakened to ***balleina* and been monophthongized to ***ballīna*, so it was borrowed after vowel weakening. The same could be the case for *caballus* and *tamarix*, but it cannot be ruled out that their medial vowels resisted weakening due to the *alacer rule* whereby a short vowel in a medial syllable is preserved if it contains the same vowel as that of the initial syllable and is separated by only one consonant (cf. Weiss 2020: 128-9). For the same reason, it is difficult to prove when *gabata* and *nepeta*, most likely loans from Greek, were borrowed. The unweakened *a* of *supparus* cannot be explained in this way, but there exists the possibility that it is a loan from Oscan.

Importantly, *asinus*, *asīlus*, and *casa* have no secure comparanda, but *ballaena* is a Latin-Greek isogloss. The Latin and Greek forms are very similar, but slight differences mean that the Latin word cannot be a direct borrowing from Greek. This suggests the existence of (an) unknown language(s) of the Mediterranean that served as an intermediary of Greek words into Latin or from which both Latin and Greek borrowed. *Ballaena* could be seen as preserving the *αι* diphthong of Gk. φάλλαίνα, but it cannot be ruled out that the source word had this diphthong. In any case, this language must have been in existence at a very late date: after vowel weakening. *Carbasus* is somewhat more

complex in light of its other comparanda, but it is a crucial case. Its unrhotalized *s* places it amongst the most recent Latin loanwords. It is a close match with Gk. *κάρπασος*, but the alternation in voicing shows it, like *ballaena*, is not a direct loan. If *carbasus* and *ballaena* belong to this most recent stratum of loanwords into Latin, then probably so too do other Latin-Greek isoglosses that show similar types of alternations.

4.2.1 Latin-Greek Isoglosses with Recurring Irregular Alternations

4.2.1.1 Voicing and Devoicing

Like Lat. *carbasus* ~ Gk. *κάρπασος*, there is a group of Latin-Greek isoglosses that differ in the voicedness of their consonants.

Lat. *buxus* is identical to Gk. *πύξος* in form and meaning but for the irregular *b* ~ *p* alternation. It occurs also between Lat. *burrus* ~ Gk. *πυρρός* (as well as between the same forms in use as personal names). It may be present in Lat. *burgus* ~ Gk. *πύργος*, but there is a chance that the Latin form is a loan from Germanic (Biville I: 235-7).

A similar voicing alternation occurs between Lat. *gubernō* ~ Gk. *κυβερνάω*. Since Greek has variants like Cypriot *ku-me-re-na-i*, the closeness of the Latin form to the Greek form might mean that, regardless of the source of the Greek form, it was the ultimate source of the Latin, mediated by an unknown language. Such is also the case for Lat. *grūmus* ~ Gk. *κρῶμαξ*, *κλωμαξ* with its *l* ~ *r* alternation.⁵²⁴

The opposite voicing correspondence occurs in Lat. *cotōneum* ~ Gk. *κυδόνια* and Lat. *taeda* ~ Gk. *δαίς*. It may also occur in Lat. *citrus* ~ Gk. *κέδρος*, but here the devoicing may be regular.

4.2.1.2 Aspiration Alternations

Like Lat. *ballaena* ~ Gk. *φάλλαινα*, several Latin-Greek isoglosses attest to aspiration alternations. The inherited source of Gk. *φ* is PIE **b^h*, so one might propose that the pair *ballaena* : *φάλλαινα* represents an older loan, closer to the time at which PGk. **p^h* was developing from **b^h*. But we have already seen that the lack of vowel weakening in *ballaena* requires it to be quite a recent loan in Latin.

Otherwise, Greek generally has a form that looks like the reflex of a voiced aspirate against a Latin *tenuis*. Because this is the normal outcome in Latin loanwords from Greek (cf. Biville I: 142), these cases critically show other irregularities that demonstrate that they are not direct loans. The alternations include **d^h* ~ **t*: Lat. *menta* ~ Gk. *μίνθη* (with additional *e* ~ *i* alternation), Lat. *lēns* ~ Gk. *λάθυρος* (vocalism points to syllabic *n*, too early for borrowing), Lat. *cant(h)ērius* ~ Gk. *κανθήλιος* (with *l* ~ *r* alternation in

⁵²⁴ Other cases of this correspondence are difficult to analyze due to the occurrence of unvoiced variants in Latin (see fn. 75). The opposite correspondence might occur between Lat. *conger* ~ Gk. *γόγγρος*, but the *c* might be the result of remodeling in Latin on other words beginning with *con-*.

what is otherwise looks to be the same suffix); **g^h* ~ **k*: Lat. *orca* ~ Gk. ὄρχη (with additional *o* ~ *u* alternation), Lat. *calx* : Gk. χάλιξ (with unexplained syncope); **b^h* ~ **p*: Lat. *adeps* ~ Gk. ἄλειφα(ρ) (with *d* ~ *l* alternation). Cf. also Lat. *sirpe* ~ Gk. σίλφιον (with additional *l* ~ *r* alternation), which has clear Berber comparanda. The Hesychian variant σέλπιον attests to an aspiration alternation within Greek for this lexeme. This is also the case for Gk. κυάρισσος, which has an aspirated variant preserved in κυφαρίσσινος.

4.2.2 Wider Implications: A Mediterranean Substrate

Thanks to the phonological features of *ballaena* and *carbasus* that require them to be recent loans in combination with the alternations they attest to with their Greek comparanda, we can conclude that a number of Latin-Greek irregular isoglosses can potentially be dated to quite recently. Nor is it surprising that Mediterranean contact situations should have been the most recent; both Latin and Greek ended up being spoken in the Mediterranean.

But while some loanwords seem to have entered Latin in the historical era, others must have entered at a Proto-Italic date. This includes *ficus*, as mentioned above, but also probably the Latin-Greek isogloss *fascinus* ~ βάσκανος, which seems to have entered Italic either as **b^h* or an intermediate stage of this phoneme's development to attested *f* (see further under §4.3.2.1). Given this time span, it is important to consider the question of how many different languages Italic was in contact with in the Mediterranean and whether any of these was actually exclusive to the Mediterranean region. In short, is there really a Mediterranean substrate in Latin?

4.2.2.1 Other Words with a Mediterranean Distribution

A problem with the search for a Mediterranean substrate is (and has been) the difficulty in delineating which areas count as Mediterranean. Here I will examine the distribution of the words that have comparanda in Greek to the exclusion of Celtic and Germanic. This entails, including the cases discussed above:

Several Latin-Greek isoglosses: *adeps*, *alaternus*, *arāneus*, *ballaena*, *bolunda*, *burrus*, *buxus*, *calx*, *cant(h)ērius*, *caupō*, *cēpa*, *citrus*, *cotōneum*, *fascinus*, *funda*, *grūmus*, *hedera*, *lacerna*, *laurus*, *menta*, *orca*, *tamarix*, *vaccīnium*

Some Latin-Greek-Berber isoglosses: *sirpe*, ?*ālium*, ?*lēns*, ?*taeda*

Some Latin-Greek-Semitic isoglosses: *cupressus*, *laena*, *malva*, *simila*

A Latin-Greek-Armenian isogloss: *nāpus*

A Latin-Greek-Armenian-Semitic isogloss: *ficus*

A Latin-Greek-Albanian isogloss: *hirundō*

And word with comparanda in Latin, Greek, Egyptian, perhaps Anatolian:

lilium.

Without any assumptions about time-depth, these distributional clusters are all arguably geographically Mediterranean. The only exception is Armenian, but it attests to other loanwords with a Mediterranean distribution (e.g. Arm. *xstor*, Gk. σκόροδον, Alb. *hurdhë* ‘garlic’, Clackson 2017: 1123).

There are also words with distributions whose relationship to the Mediterranean is not immediately clear.

4.2.2.2 Balkan Connections

The Latin-Greek-Albanian isogloss *hirundō* indicates a Mediterranean loanword with attestation in Albanian, which is not surprising given the geographic position of the Balkan Peninsula. There are however several words with an otherwise Mediterranean distribution that are also attested in Slavic. These are:

PSlav. **k/gr̥m-* ~ Lat. *grāmiae* ~ Gk. γλάμων

PSlav. **tyky-* ~ Lat. *cucumis* ~ Gk. σικύα, Hsch. κύκνον, σεκούα ~ Arm. *sex*

PSlav. **mъskъ-* ~ Lat. *mūlus* ~ Gk. μύκλος, Hsch. μυχλός, μύσκλοι ~ Alb. *mushk*

Since the migration of Slavic peoples into the Balkan Peninsula occurred in the sixth to seventh centuries CE (cf. Kobyliński 2005: 532-6), these words either entered Slavic after its speakers appeared in Balkans (and thus may still represent words from a Mediterranean language) or they attest to a different, earlier contact situation. The variation within Slavic of (roughly reconstructed) **k/gr̥m-* shows that it entered after the Common Slavic era and is thus an example of the former case. The *l ~ r* alternation between Lat. *grāmiae* ~ Gk. γλάμων is well-attested in other Mediterranean words (see §4.2.2.4.1). The same could be true for **tyky-*, whose comparanda are otherwise distributed similarly to *nāpus* and *fīcus*. It is more difficult to decide for **mъskъ-*. It seems to attest to the phenomenon of *SK* metathesis that otherwise sees a broader European distribution (cf. Lat. *ascia* ~ Gk. ἀξίτη ~ PGM. **akwes(i)-*). Its status as a Balkan word might be artificially bolstered by the Albanian comparandum; it is likely a loan from Slavic. On the other hand, both variants are present in Greek, such that Slavic might have borrowed its word after arrival in the Balkans. Since donkeys were introduced to Italy and Greece from the East in the first millennium BCE (EIEC 33-4, Todd et al. 2022), it is quite unlikely that **mъskъ-* represents a loan alongside Italic and Greek from a time before the latter languages entered the Mediterranean region.

Attestation in Slavic therefore does not seem to rule out the possibility of a lexeme’s origin in a substrate language of the Mediterranean when there are other indications that this is the case. Nor does it seem possible to extract from the Latin data a group of specifically Balkan origin.

Alessio (1946a) thought he found traces of a Mediterranean substrate in Baltic. Most of his evidence is almost exclusively toponymic, relying on the ability to see the same roots between e.g. hydronyms in the East and the West from which potential substrate lexemes can be distilled and assigned an aquatic/marine meaning. Amongst the lexical correspondences he does mention, three deserve mention:⁵²⁵

Lith. *báltas*, Latv. *bal̃ts* ‘white’⁵²⁶ ~ PSlav. **bòlto-* ‘swamp’ ~ Alb. *baltë* ‘mud, swamp’, MoGk. βάλτος ‘swamp’ ~ Rom. *baltă* ‘puddle, pool; swamp’, dial. It. e.g. Lucchese *paltenna* ‘mud’

Lith. *kor̃ys*, Latv. *kāre* ‘honeycomb’ ~ Lat. *cēra* ‘wax’ ~ Gk. κηρός ‘beeswax’, κήρινθος ‘beebread’

Lith. *br̃iedis* ‘elk, deer’, Latv. *briēdis* ‘elk’ ~ Alb. *bri* ‘horn, rack’ ~ Hsch. βρένδον ‘deer’, Gk. βρέντιον (given as Messapic) ‘stag’s head’ ~ PGm. **brinda(n)-* ‘elk’

To these can be added:

PSlav. **grabr̃* ~ Lat. *carpinus* ~ ?Gk. γράβιον ~ ?PBalt. **ske/irp-*

Gk. κυβερνάω ~ Lat. *gubernāre* ~ ?Lith. *kūmbryti* ~ ?OCS *kr̃mīti*

The group with Lith. *báltas* has been suggested to be an Illyrian substrate word (Brüch 1916, Derksen 2014 s.v. *baltas*). But given its late appearance outside Balto-Slavic (i.e. Romance and Modern Greek), the non-Balto-Slavic forms may all be early loans from Slavic (cf. Demiraj 1997: 87-8). As mentioned for Gk. κυβερνάω ~ Lat. *gubernāre*, the Balto-Slavic connections made by e.g. Boisacq (1916: 527-8) and Machek (1955: 61-4) may be the result of chance resemblance. The Baltic and even Greek comparanda for Lat. *carpinus* are not very good, and may simply be unrelated. Instead, it might attest to an Italo-Slavic isogloss and it is thus difficult to determine when this word was borrowed.

That leaves the groups of cervid and apicultural words. For the former group, Kroonen (2013: 77) separates the Germanic word as a coincidental lookalike with an internal Germanic etymology and Demiraj (1997: 110-11) notes formal difficulties with directly comparing the Albanian form. Derksen (2014 s.v. *briedis*) calls the relationship between the Baltic forms and the rest of the comparanda “obscure” but does not reject it. For the

⁵²⁵ Some of his other examples include 1) Lat. *lāma* ‘marsh’ ~ Lith. *lomà* ‘hollow, valley, plot’, Latv. *lāma* ‘hollow in a field or meadow, pool, pit’ ~ PSlav. **lam̃* ‘hollow, bend’ ~ Gk. λάμια ‘chasms’. But the Latin form is probably not related after all (Schrijver 1991: 142, Derksen 2014 s.v. *loma*). Nor can the Greek word be compared with any semantic certainty, removing the Mediterranean comparanda. 2) Latv. *ērms* ‘monkey’ ~ Etr. ἄριμος ‘monkey’. The Etruscan word is given by Strabo. The island today called Ischia was called by Virgil *Inarime* and *Arime* and was called by the Greeks Πιθηκοῦσα, seemingly derived from πίθηκος ‘ape’. No monkeys have lived on the island in recorded history (suggesting some sort of folk etymological interference), and the Etruscan word is too poorly attested to use.

⁵²⁶ The semantic change that must be assumed to connect ‘white’ and ‘swamp’ is variously defended and viewed with suspicion.

latter group, Lat. *cēra* cannot be ruled out as loan from Gk. κηρός, but the suffix of κήρινθος proves its antiquity in Greece. It seems very unlikely that Baltic would ever have been in contact with the Mediterranean, at least in any direct linguistic way. Rather than a Mediterranean stratum in Baltic, an alternative explanation for these would be that they attest to a separate contact situation.⁵²⁷ There are several cases of non-IE words that are attested in Baltic and Greek alongside others: Gk. λύγξ ‘lynx’ ~ PGM. **luhsa-* ‘lynx’ ~ Arm. *lusanun* ‘lynx; hyena; marten’ ~ Lith. *lūšis*, dial. *lūnšis* ‘lynx’; Gk. θρούον ‘reed, rush’ ~ OCS *trbstь* ‘reed, cane’ ~ Lith. *triušis* ‘reed’; Gk. κρόμμυον ‘onion’ ~ PGM. **hramusan-* ‘ramsons’ ~ Mlr. *crem* etc. ‘wild garlic’ ~ Ru. *čeremšá*, etc ‘wild garlic’ ~ Lith. *kermušė*, dial. *kermušà* ‘tip or a drill or flail, wild garlic’. These distributions do not seem convincingly Mediterranean and Alessio’s label seems to be the result of overstretching the term “Mediterranean” as the only substrate language of Europe.

4.2.2.3 Indo-Iranian Connections

Pisani (1938a) had been inspired by recent excavations of the Indus Valley Civilization to look for cultural connections stretching from the Mediterranean to India. He pointed to what he saw as cultural and religious similarities (like a Mother Goddess figure) in all of these areas in the Bronze Age. He only provided two lexical correspondences: Gk. ἐρέβινθος ‘chickpea’ : Skt. *aravinda-* ‘lotus’ and Gk. κάρπασος⁵²⁸ ‘fine flax’ : Skt. *karpāsa-* ‘cotton’.

Gk. ἐρέβινθος has many more irregular comparanda across Europe (Lat. *ervum*, PGM. **arwīt-*, Arm. *a’owoyt*). While the appurtenance of the Sanskrit form Pisani mentions is disputed, there do indeed seem to be comparanda in Iranian languages. This word for an agricultural founder crop attests to an older contact situation with early farmers rather than a Mediterranean substrate language stretching to India. On the opposite side of the spectrum, Gk. κάρπασος is at most a Wanderwort with its origins in the East, having entered Greek after the loss of inherited intervocalic *s and making its way to Latin (at an even later date) via some intermediary language of the Mediterranean. A similar situation, though it may well have occurred earlier, holds for Lat. *rosa* ~ Gk. ῥόδον.

There are a few other irregular Latin-Greek isoglosses (or near isoglosses) that have potential matches in Indo-Iranian however. These include:

Lat. *calix* ~ Gk. κόλιξ ~ ?Skt. *kalása-*

Lat. *frīgō* ~ Gk. φρήγω ~ ?PIr. **bra(i)ǰ-*, Skt. *bhrajj-*

Lat. *pirum* ~ Gk. ἄπιον ~ ?Shina *pisō* ~ ?Burushaski *pheso* ~ ?Khinalug *bzi*

Lat. *racēmus* ~ Gk. ῥάξ, ῥώξ ~ ??Alb. *rrush* ~ ?PIr. **raza-*

In all cases, the appurtenance of the Indo-Iranian material is uncertain. Formally and

⁵²⁷ Contact between populations involved in the trade of Baltic amber would be plausible.

⁵²⁸ He gives κάρβασος, but this form is not actually attested.

semantically, the case for Skt. *kalása-* is strongest, as it points to a reconstruction with *a*-vocalism like Latin. The Indo-Iranian comparanda of *frīgō* and *φρύγω* are semantically compelling, though not all connect them (cf. LIV2 s.v. **b^hreg-*; DV 212, 243). The appurtenance of all the pear words is difficult to evaluate. While Pr. **raza-* can reconstruct to **(H)reg-*, formally aberrant **leg^(h)-* is alternatively possible. In no case is an extended Mediterranean substrate the only conceivable explanation however. As for *ervum*, some distributions might attest to an older contact situation while, like *carbasus*, others may represent Wanderwörter. And others still may be due to chance resemblance.

4.2.2.4 Recurring Features

With some of the confounding variables removed from the dataset, we can search the list of words with a confirmed Mediterranean distribution for recurring irregular correspondences. In addition to the voicing and aspiration alternations mentioned under §4.2.1, three others appear: *l ~ r*, *e ~ i*, and *i ~ u*.

4.2.2.4.1 *L ~ R Alternation*

There are 10 instances of an *l ~ r* alternation. For 6 of these, the lexeme can be argued to have a Mediterranean distribution.⁵²⁹ In all of these 6, the alternation exists between Latin and Greek and there is no secure Semitic comparandum. In the first 5 cases, Lat. *r* corresponds to Gk. *l*:

Lat. *cant(h)ērius* ~ Gk. *κανθήλια*

Lat. *grūmus* ~ Gk. *κλωμαξ*, Hsch. *κρωμαξ* (here the alternation also exists within Greek)

Lat. *hirundō* ~ Gk. *χελιδών* ~ Alb. *dallëndyshe*

Lat. *sirpe* ~ Gk. *σίλφιον*, Hsch. *σέλπον* ~ Berber *azlaf*, *aselbu*, etc.

Lat. *grāmiaie* ~ Gk. *γλάμων* (as detailed above, Slavic comparanda were likely borrowed late)

In one case, the correlation of *r* to *l* is opposite:

Gk. *λείριον* ~ Lat. *līlium*, Hitt. *alēl*, Copt. *hlēli*, etc.

There is one further case where an *l ~ r* alternation exists between Latin and Romance descendants:

PRom. **darbo-* ~ Lat. *talpa*

Lat. *paelex* and Gk. *παλλακή* agree on *l* against OIr. *airech*. But this likely belongs to a different contact phenomenon given the pattern so far and since Celtic does not clearly belong to the Mediterranean region. The same is certainly true for Lat. *filix* ~ Gk.

⁵²⁹For *l(l)/r(r)* as a Mediterranean alternation, cf. Bertoldi (1942: 183-4, fn. 1), Hubschmid (1953: 72), more generally as non-IE in Alessio (1944a: 124).

βλῆχρον, βλῆχρον ~ PGM. *brekna(n)-. In the group of widely distributed pigeon words (including Lat. *columba*), a variant with *r* is present only in Egyptian. Theoretically this could be due to this word's transmission to Egyptian by a language of the Mediterranean, but this is difficult to prove. In any case, 6 or 7 out of 10 cases of an *l ~ r* alternation occur in lemmata with a Mediterranean distribution.

4.2.2.4.2 E ~ I Alternation

There are additionally several cases of an *e ~ i* alternation between comparanda in the dataset. In 6 cases, the alternation occurs between branches, with 5 of these cases being clearly restricted to the Mediterranean.⁵³⁰ In 3 of these 5, Latin has *e* for Gk. *i*:

Lat. *cupressus* ~ Gk. κυπάρισσος (~ Hebr. *gofer*)

Lat. *menta* ~ Gk. μίνθη⁵³¹

Lat. *hedera* ~ Gk. κιστός

In 2 cases, the distribution is reversed:

Lat. *hirundō* ~ Gk. χελιδόν (~ Alb. *dallëndyshe*)⁵³²

Lat. *citrus* ~ Gk. κέδρος

The non-Mediterranean case is Lat. *avēna* ~ PSlav. *ovъsъ ~ PBalt. *(a)viža, whose additional Uralic and potential Germanic comparanda point to a different contact situation.

There is an additional case of an *e ~ i* alternation within Latin (*genesta/genista*) and two within Greek for lexemes that have a Latin relative:

Gk. σίλφιον, Hsch. σέλπιον ~ Lat. *sirpe* ~ Berber *azlaf, aselbu*, etc.

Gk. σικία, Hsch. σεκούα, κύκνον ~ Lat. *cucumis* ~ Arm. *sex* ~ PSlav. *tyky-

If these inner-branch alternations are included in the count, then there are 9 instances of an *e ~ i* alternation in the data, 8 of which occur with a Mediterranean distribution.⁵³³ Some of these cases also include an *l ~ r* alternation and the alternations in voicedness and aspiration detailed above.

4.2.2.4.3 I ~ U Alternation

The set including Gk. κύκνον above also attests to *u*-vocalism. Another case of *i ~ u*

⁵³⁰ *e/i* as a Mediterranean feature with different explanations also in e.g. Bertoldi (1939b: 89), Battisti (1959: 154-7).

⁵³¹ Romance languages also attest to another case like this: PRom. *plenta- 'clod of earth' ~ Gk. πλίνθος 'brick' (cf. Alessio 1944a: 139-41).

⁵³² Its root vocalism, neither **e* nor **i*, perhaps suggests that Albanian was less directly involved.

⁵³³ It seems very unclear whether Lat. *filix/felix* < **b^hel-ik-* ~ Gk. βλῆχρον, βλῆχρον < **b^hl-ēg^hr/n-* ~ PGM. *brekna(n)- < **b^hr-eg-n-* is an example of this same type of alternation.

vocalic alternation is that of Lat. *supparus* ~ Gk. σίφαρος, σείφαρος. The Latin word may be a loan from Oscan, which makes it difficult to analyze the other alternations (geminate *pp* for Gk. φ, unweakened *a*) as original. But even if this is true, the *i* ~ *u* alternation lacks an explanation and makes it likely that the Oscan word was indirectly borrowed. This is reminiscent of two words with a clear *i/ū* alternation: Lat. *fīcus* ~ Gk. τῦκον, σῦκον ~ Arm. *fūz* and Lat. *frīgō* ~ Gk. φρύγω (Indo-Iranian comparanda, as detailed above, may not have much bearing on the words' treatment in the Mediterranean). Such *i* ~ *ū* alternations have been considered characteristic of the Mediterranean substrate (e.g. Bertoldi 1948: 70, Hubschmid 1953: 28, Alessio 1955: 375, 537-40, Battisti 1959: 155) and the examples are indeed restricted to the Mediterranean region. In fact, Lat. *hirundō* ~ Gk. χελιδών attests to a *u* ~ *i* alternation in its second syllable alongside *e* ~ *i* and *l* ~ *r* alternations.

4.2.2.5 A Definitively Mediterranean Substrate

A remarkable pattern emerges wherein the words with a Mediterranean distribution attest to a set of irregular alternations that are also by and large restricted to a Mediterranean distribution. The alternations overlap, such that a single lexeme can attest to more than one. Because they co-occur in the same word, it is highly likely that they are the result of the same language or at least group of closely related languages. The co-occurrence of the Mediterranean *e* ~ *i*, *i* ~ *u*, and *l* ~ *r* alternations in the same words as alternations in voicing and aspiration means that this Mediterranean substrate language was also responsible for at least some cases of this alternation (Figure 4.3).

The feature of voicing and aspiration alternation is more widespread than the Mediterranean. But rather than concluding from this that the Mediterranean substrate language extended far into Europe, it seems most likely that such alternations can be the result of different contact situations.

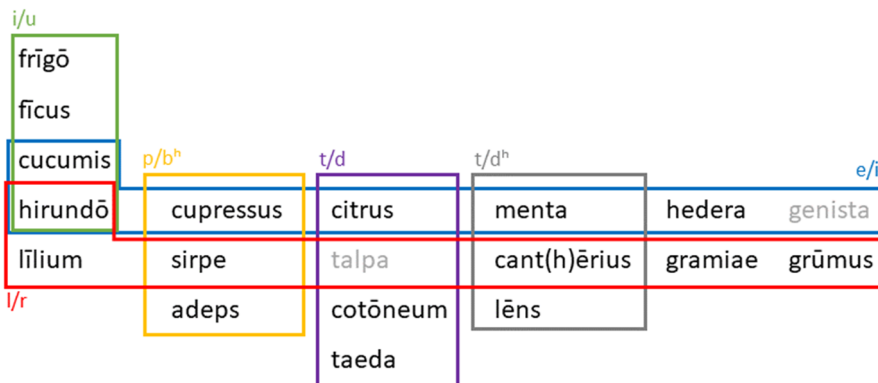


Figure 4.3 Overlapping irregular alternations between words with a Mediterranean distribution. (Words in gray are restricted to Italic)

It is interesting to note that the Mediterranean group contains *ficus*, a word which must have entered at a Proto-Italic date. Thus it is likely that this language was present on the Italian peninsula and represents one of the earliest Mediterranean contact situations in the history of Latin. Though unclear in Figure 4.3, Lat. *talpa* alongside PRom. **darbo-* also attests to a *b ~ p* alternation that is reminiscent of the one represented by Lat. *carbasus* ~ Gk. *κάρπασος*. It is therefore possible the language persisted in the Mediterranean to a recent date. However, it should not be ruled out that there were other languages in the Mediterranean as well. Especially in light of cases that exhibit other alternations than the ones highlighted here, the language responsible for the *l ~ r*, *e ~ i*, and *i ~ u* alternations represents *at least* one distinct Mediterranean substrate language.

That some irregular alternations appearing in words restricted to a Mediterranean distribution are themselves more widespread brings up an important consideration. Claiming that words originated from (a) Mediterranean substrate language(s) based on their distribution alone is risky. What if the reason that the lexemes are restricted to the Mediterranean is due to the fact that their *signifiés* are found only in the Mediterranean? It is plausible that, within one language family, languages in different ecological zones will have portions of mutually exclusive vocabulary based on the presence of something in one language's environment but its absence in that of the other. As concerns the Latin substrate vocabulary with a Mediterranean distribution of attestation, we see this for several plant names. The geographic range of the plants referred to by *alaternus*, *buxus*, *cupressus*, *ficus* (and possibly *tamarix*⁵³⁴) themselves have a Mediterranean distribution (Figures 4.4-7). The species to which *cucumis* refers are likely of Asian origin (Sebastian, Schaefer, Telford & Renner 2010), and early evidence of them in the West comes from Egypt, Mesopotamia, and Greece (Zohary, Hopf, and Weiss 2012: 154-5) meaning their distribution as concerns Europe is Mediterranean.

⁵³⁴ Heywood et al. (2007) give a distribution for the genus *Tamarix* that stretches into central Northern Europe, but most individual species do not seem to range beyond Southeast Europe.



Figure 4.4 Range of *Rhamnus alaternus* (Based on Bolòs & Vigo 1984-2001)



Figure 4.5 Range of *Ficus carica* (Based on Zohary, Hopf & Weiss 2012: 129)



Figure 4.6 Range of *Buxus sempervirens* ([https://commons.wikimedia.org/wiki/File: Buxus_sempervirens_range.svg](https://commons.wikimedia.org/wiki/File:Buxus_sempervirens_range.svg) with refs.)



Figure 4.7 Range of *Cupressus sempervirens* ([https://commons.wikimedia.org/wiki/File:C upressus_sempervirens_range.svg](https://commons.wikimedia.org/wiki/File:Cupressus_sempervirens_range.svg) with refs.)

However, several other plant words with a Mediterranean distribution, including some included in Figure 4.3, indeed refer to plants whose distribution is not limited to the Mediterranean. This includes *hedera*, *malva*, *menta*, and *vaccīnium*. In combination with the overlapping Mediterranean alternations (for *hedera* and *menta*), this suggests that something more than the geographic range of *signifiés* is producing the pattern. Most likely not in all cases, but at least in many, it instead points to a true linguistic boundary.

Given that *cupressus* and *hedera*, Latin words that seem to contain Pre-Greek suffixes, belong to the distinct Mediterranean substrate language isolated above, it seems to have been responsible for the indirect transmission of some Greek words into Latin.

4.2.2.6 Mediterranean Languages with a Wider Distribution

There is a small number of curious cases where non-Indo-European Mediterranean languages attest to comparanda that show a wider distribution. This is the case for Semitic and Sumerian. These must represent older contact situations, quite probably in

the way of Wanderwörter. This is relatively certain for Lat. *ferrum* and its Germanic, Kartvelian, and (probably) NE Caucasian comparanda. Along with several Semitic forms, these words have their ultimate source in an Anatolian language. It is in Anatolia that iron production technology developed, and it spread after the split of the Indo-European languages (cf. Thorsø & Wigman et al. 2023: 120). The group Lat. *ascia* ~ Gk. ἀξίμη ~ PGm. **akwes*(*t*)- has comparanda in Semitic and Sumerian, suggesting a Wanderwort from the East. The Germanic form almost certainly precludes a borrowing from a Mediterranean contact situation.⁵³⁵ But alternations within Akkadian (whence the other Semitic forms seem to be loans) indicate that it is not native there. All forms could be Wanderwörter with an ultimate source in Sumerian, but it cannot necessarily be ruled out that the Sumerian word is also a borrowing. A similar case with strong implications is that of Lat. *raudus* ~ PGm. **arut*- (and possibly PCelt. **rutu*-). They represent a case of non-IE *a*-prefixation, strongly suggesting that they entered from a European substrate source, but their similarity to OSum. *aruda* raises questions that are difficult to answer. Does this hint at a relationship between the language of the *a*-prefixes and Sumerian (cf. Schrijver 2018: 361-3), or is *aruda* a Wanderwort in Sumerian?

The case of Arm. *ƒuz* with its Latin (*ficus*), Greek, and Semitic comparanda indicate that Armenian was in contact with some of the same Mediterranean languages as Latin and Greek. But several other cases of more widespread lexemes with attestation in Armenian show that it participated in some older contact situations involving Latin as well. Such situations as these will now be examined, primarily via their distributions of attestation amongst the IE daughter languages.

4.3 Further Stratification Using Distribution

4.3.1 Potentially Recent Borrowings

4.3.1.1 Words Exclusive to Latin and Romance

There are 5 words restricted to Latin (i.e. without secure comparanda) but which are nevertheless unlikely to be inherited due to their invalid root structure: *faex*, *farcio*, *focus*, *pampinus*, *tabanus*. The antiquity of these words in Latin is difficult to estimate given their lack of comparanda. They may represent words borrowed from a language spoken exclusively on the Italian peninsula, but it cannot be ruled out that their comparanda in other languages—and therefore true distribution—have been lost.

On the other hand, the 7 words restricted to Italic with irregularities recorded by Romance descendants (or alternations within Latin) most likely represent a recent

⁵³⁵ Alessio (1941b: 204, 1946a: 143) had additionally used Ger. *Mohn* ‘poppy’, Gk. μήκων ‘poppy’, etc. to suggest that Germanic indeed has words from a Mediterranean substrate. He entertains part of the proposal by Ribezzo (1934a) that his would have occurred via contacts in the Alps with Rhaetic. Kroonen (2013: 371) however shows that all attestations of this word (which includes Baltic, Slavic, and Albanian) can be reconstructed to **méh_hk-on-*, an IE-looking root albeit only attested in Europe. This seems to be another case of the over-Mediterraneanization of European linguistic features.

stratum of loans, borrowed at least after the dissolution of Proto-Italic, if not after the development of the Latin branch. These include *arbutus*, *cerrus*, *genesta*, *lābrusca*, *lepus*, *sambūcus*, and *talpa*. It is *genesta* and *sambūcus* which show alternations within Latin, making them likely the oldest borrowings of the group. *Genesta* (var. *genista*) attests to an *e ~ i* alternation, quite likely the same one identified for the Mediterranean substrate (§4.2.2.4.2). Similar is likely true for the *l ~ r* alternation between Lat. *talpa* and PRom. **darbo-* (§4.2.2.4.1). They may thus represent examples where comparanda in other Mediterranean languages have been lost.

Sambūcus attests to a *b ~ mb* alternation within Latin. This otherwise appears in the Romance descendants of Lat. *lābrusca*. It is unclear if this is related to the *b ~ m* alternation between the Latin and Romance forms of *arbutus*, but it is interesting that Lat. *plumbum* attests to a nasal that is rather difficult to account for. While lead is present in the Mediterranean from quite an early date, the wide variation in the comparanda of Lat. *plumbum* makes it likely to be a relatively late Wanderwort, perhaps introduced along with the introduction of lead-alloyed copper that reaches Southern Europe around 1000 BCE (Thorsø & Wigman et al. 2023: 119). The irregular appearance of the nasal before *b* in *plumbum*, *sambūcus*, and *lābrusca* may thus be a phenomenon peculiar to the Italian peninsula, representing a feature of a language spoken there.

4.3.1.2 Gemination

Irregular Romance reflexes contain geminates not matched in Latin for *bāca*, *baculum*, *lapis*, and *lepus*. That a substrate language of the Italian peninsula might be responsible for the presence of geminates is suggested by the irregular alternation within Latin between *sappīnus* and *sabīna* (and perhaps the geminates present in Lat. *cerrus* and *vaccīnium*). The presence of both variants suggests, as above, a recent date of entry. However for *bāca*, *baculum*, and *lapis*, where Romance descendants attest to a geminate not recorded in Latin, comparanda exist outside of Italic (Greek, Germanic, and Celtic). Given that a similar pattern occurs between Lat. *lepus* ~ PRom. **lapparo-*, where non-Italic comparanda do not exist, it seems possible that the geminate forms may represent secondary reborrowing after the phonemicization of gemination in Latin and after the expansion of Latin back out of the peninsula (resulting in the Romance reflexes). This chronological differentiation may explain the pair *sappīnus* ~ *sabīnus* (and perhaps *bāca* ~ *vaccīnium*, if they are ultimately related).

4.3.2 Earlier Strata

4.3.2.1 The Oldest Loans in Italy

As mentioned above in §4.2.2, several lexemes from the corpus can be argued to have been borrowed at a time prior to or at the latest during the Proto-Italic developments regarding the voiced aspirates. The case is easiest to demonstrate when the Latin result differs in place of articulation from the reconstruction suggested by the comparanda. While it could be argued that a correspondence like Lat. *f* ~ PGm. **b* (in words with

other indications of being non-native) represents different nativizations of a foreign /β/ or /v/, that Lat. *f* ~ Gk. φ could represent words that entered with /f/ (especially when they are likely to be recent borrowings), and that Lat. *h* ~ Gk. χ could represent nativizations of a foreign /x/, similar explanations are more difficult when Lat. *f* points to the reconstruction of **d^h* or **g^{wh}*.

That loanwords entered such a stage in the history of Latin is not overly surprising for the words with a non-Mediterranean distribution. Such a case can be made for the following words:

In Initial Position

Likely

fracēs < **d^hrak-*: Guaranteed by PGm., PSlav. PBalt. onset **dr-*
fungus < **g^{wh}ong-*: Suggested by Gk. σφ/πόγγος, PGm. **swamb/ppan-*

Plausibly

faba < **b^hab-*: Suggested by PGm. *baunō-*, PBSl. **ba/ob-*
ferrum < **b^herso-*: Suggested by PGm. **brasa-*
fīlix/felix < **b^helik-*: Suggested by Gk. βλήχρον, PGm. **brekna(n)-*
fulica < **b^hulVk-*: Suggested by PGm. **balikōn-*
hasta < **g^hast-*: Suggested by PCelt. **gasdo-*, **gazdo-*, PGm. **gazda-*

In Medial Position

Likely

barba < **bar(s?)d^h*: Suggested by **d* in PGm., PSlav., PBalt.
caulae < **kag^h*: Guaranteed by Lat. *cohūm*, **g* in PCelt., PGm.

Plausibly

columba < **ke/ol-o/umb^h*-. Suggested by PGm. **kulubrōn-*
corbis < **korb^h*-. Suggested by PGm. **krebo-*

In several other cases, the analysis is more complicated. The *b* of Lat. *plumbum* can be reconstructed with a **d^h* like for PCelt. *(*ϕ*)*loudio-*, or it perhaps corresponds to the β of e.g. Gk. μόλιβδος. *Sabulum* could reconstruct to a **b^h* if that is what caused the gemination of Gk. ψάμμος and underlies the *w* of Arm. *awaz*. The analysis of *cucurbita* depends on how the elements align with those of PGm. **hwehwetjō-*. *Sulpur* requires **p* but some Romance forms reconstruct to **b^h*, making its interpretation less than straightforward. And *badius* points to a reconstruction with **d^h* if its suffix is **jō*, rather than **iHo* (**iyō*).

More surprising, or at least informative, is the fact that some Latin loanwords with Mediterranean comparanda likewise entered Italic before or during the Proto-Italic treatment of voiced aspirates. Critically, this means that Proto-Italic did not split up until after it entered the Mediterranean zone. The data is as follows:

In Initial Position

Likely

fīcus < **d^hīk-*: Guaranteed by PGk. **l^(h)ylwāiko-*, Arm. *ī^huz*

faber < **d^hab-*: Guaranteed by Arm. *darbin*

hirundō < **g^hir-*: Guaranteed by Alb. *dallëndyshe*

Plausibly

fascinus < **b^hask-*: Suggested by Gk. βάσκανος if not from /β/ or /v/⁵³⁶

frīgō < **b^hrīg-*: Suggested by Gk. φρύγιω if not from /f/

funda < **b^hend-*: Suggested by Gk. σφενδόνη if not from /f/

hedera < **g^hed^h-*: Suggested by Gk. κηθάρα if not from /x/

In Medial Position

Likely

laurus < **lag^{wh}-ro-*: Suggested by Gk. δάφνη and δαύχνα

ālium < e.g. *aG^h₁-jō-*: Suggested by Gk. ἄλις, PBerb. **agVlum-*

Plausibly

tabānus < **tab^h-*: Suggested by It. *tafano*

As with the non-Mediterranean cases, there are a few forms that resist definitive analysis. *Malva* may have had **g^w* depending on its exact relationship with e.g. Gk. μάλαχη. The *d* of *bolunda* may reconstruct to **d^h* on comparison with Gk. ὄλυνθος depending on how old it is within Latin. *Faex* and *farciō* have no certain comparanda and therefore do not give any indication of the shape of their source forms.

These words (especially *ālium*, *faber*, *fīcus*, *hirundō*, and *laurus*) would thus represent the earliest loanwords taken up into Italic in the Mediterranean contact zone.

4.3.2.2 Early Contact or Widespread Substrate?

As mentioned above, the large number of substrate lexemes attested with comparanda exclusively in geographically Mediterranean languages contrasts highly with a very

⁵³⁶ Though see fn. 121.

small number of lexemes that show a more Europe-wide distribution. 7 cases in the dataset exhibit the most widespread distribution, with comparanda in (at least) Greek, Germanic, and Celtic. These include, with the additional language groups in which they are irregularly attested, the following:

Lat. <i>baculum</i>	Gk.	Gm.	Celt.	PRom.			
Lat. <i>baiulus</i>	?Gk.	Gm.	?Celt.				
Lat. <i>rādīx</i>	Gk.	Gm.	Celt.			Alb.	
Lat. <i>rāpum</i>	Gk.	Gm.	Celt.		Balt.	Slav.	
Lat. <i>turdus</i>	Gk.	Gm.	Celt.		Balt.	Slav.	Arm.
Lat. <i>tilia</i>	Gk.	?Gm.	Celt.		??Balt.	??Slav.	Arm.
Lat. <i>plumbum</i>	??Gk.	??Gm.	Celt.	(Also ?Basque, ?Berb., ??Kartvelian)			

A widespread distribution, especially amongst IE languages, can have two main explanations. The first possible scenario is that a word was borrowed early, shortly after the dissolution of PIE from a (potentially small) source near the PIE homeland into the newly differentiating IE daughter branches. The second possibility is that a word was borrowed late, upon the arrival of the individual PIE daughter languages in Europe (possibly upon reaching their places of eventual attestation, but it cannot be ruled out that words were borrowed somewhere along the way). This scenario has important implications, because it requires the word(s) to have already had a wide distribution in Europe, perhaps due to them belonging to a widespread language (or more probably language family) spoken by Early European Farmers. Genetic studies have demonstrated that agriculture spread through Europe in the Neolithic via demic diffusion, quite possibly from the same starting point in the Aegean (cf. Hofmanová et al. 2016, Shennan 2018: 107). On the other hand, the time between the arrival of agricultural populations in Europe and the arrival of the IE languages amounts to several millennia. Anthony (2007: 80) has postulated on theoretical factors that Neolithic Europe could have been home to up to twenty to forty distinct language *families*.

In the scenario of early loans from a source close to the PIE homeland, we should expect relatively little variation between languages, and especially not within them. The words would have been borrowed into the proto-daughter languages, perhaps even at intermediary subnodes, and should thus mostly follow known sound laws. This does in fact seem to be the case for the group of words represented by Lat. *taurus*, though curiously not for the Latin word itself (which seems to be a loan post-dating the metathesis of **aurV > ar̥V*). Nearly all branches (Celtic, Greek, and Balto-Slavic) attest to a pre-form **tauro-* except for Germanic, which preserves **teuro-*. (Albanian could reflect either.) Etruscan also attests *e*-vocalism in *θevru-*. This in combination with the lexeme's deep entrenchment in Semitic (potentially as deep as Afro-Asiatic) suggests a very early loan into just barely differentiated Indo-European.

On the other hand, in the second situation of later *in situ* borrowing, we should expect the full force of differentiation. The words would have been borrowed into the separate IE daughter languages from probably differentiated substrate daughter languages. And in fact, this is what we seem to find with the list of widely distributed lexemes. Lat. *plumbum* and its comparanda are widely diverse and for archaeological reasons, as mentioned above, it may represent a relatively late Wanderwort. The comparanda of *rādīx* attest to vocalic alternation between $\bar{a} \sim a \sim i$ and what may be the reflex of a schwa. For *rāpum*, the alternations are between $*p \sim *b^h$ as well as $\bar{a} \sim a \sim \bar{e}/oi$ and *a*-prefixation (in Celtic). The comparanda of *tilia* attest to a complex onset cluster (reflecting both $*pt$ and $*tp$) and potentially complex reflexes of the *a*-prefix phenomenon. *Turdus* and its comparanda attest to several alternations, the most prominent of which include the vacillating presence of a sibilant both at the onset and in the interior of the word and $*t \sim *d \sim *d^h$ alternations. *Baculum* and *baiulus* are Italic-Celtic-Greek-Germanic isoglosses whose interpretation is complicated; especially the latter, which seems to have been borrowed multiple times in several branches (Italic: *fascis* \sim *baiulus*, Germanic: $*pakk-$ \sim $*bagg-$, Gk. φάσκωλος \sim φάκελος).

Another indication that Europe may have been home to a large group of related languages is the distribution of the alternating morphological features identified in §3.3. The *a*-prefix especially appears in two lexemes from the list of the most widely distributed (*rāpum* and *tilia*). If the mechanism behind the *a*-prefix is truly as simple as an alternation between a root in the “full-grade” and an *a*-prefixed root in the “zero-grade”, then theoretically all cases could have been borrowed at an early date into the just barely differentiated daughter branches; there would be two variants, as with $*tauro-$ and $*teuro-$. However, there are indications that the mechanism was not this simple. The comparanda of *tilia* (especially if ultimately related to *pōpulus*) demonstrate this for instance. It seems more likely that the IE daughter languages borrowed the variants after their dispersion in Europe. Such is also the case for the widely distributed *n*-suffix, especially given that its vacillating presence is not the only irregular alternation occurring in the lexemes where it is found.⁵³⁷

Interestingly, there is a chance that both morphological features have a distribution that extends into the Mediterranean. Neither of these is represented in the distinct Mediterranean substrate language identified in §4.2.2.5. This may be due to the sample size, but it seems quite likely that this means there were at least two contact languages in the Mediterranean. For the *a*-prefix, Schrijver (1997: 310) cautiously suggests that cases of (mainly) Greek *prothetic* α (cf. Furnée 1972: 368-74 with lit.) involving the vacillating presence of *a*- without vocalic reduction of the root could be a related phenomenon. For the *n*-suffix, cases like Lat. *urna*, *orca* \sim Gk. ὕρχη and Lat. *laurus* \sim Gk. δάφνη, δούχνα are attested with a Mediterranean distribution. For neither of these can it be confirmed that the Mediterranean cases represent the same phenomenon. But it

⁵³⁷ The extent of the *filix-fulica-sōre* type velar suffix is complicated by the inability to confirm its relationship with other cases of velar suffixes.

is difficult to rule out, especially for the *n*-suffix. This makes such cases difficult to stratify. If the *a*-prefix and *n*-suffix were found in a Europe-wide distribution including the Mediterranean, both Italic and Greek could theoretically have received their lexemes with these features upon arriving in the Mediterranean.

Thus an important further question remains whether there exist any indications of lexemes that were borrowed along the way to the areas in which the languages would come to be attested, i.e. traces of their prior homelands. Based on the distribution alone, there is indeed a group of words attested in Latin, Germanic, and Celtic to the exclusion of Greek as well as groups including Latin-Celtic and Latin-Germanic isoglosses. Are these due to coincidental lexical loss, or do they represent different contact situations, such as in Figure 4.8?

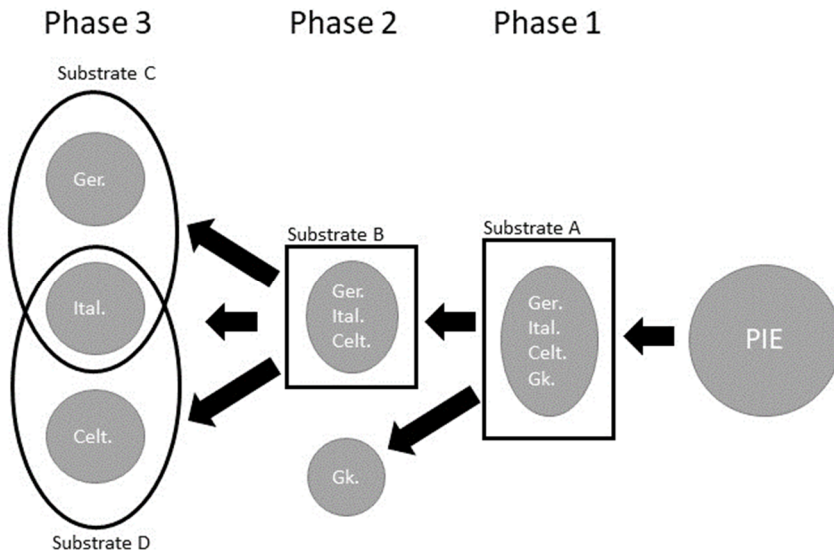


Figure 4.4 Separate contact situations suggested by the distribution of non-inherited lexemes

4.3.2.3 Intermediate Contact Situations

4.3.2.3.1 *Stratum Excluding Greek*

There is a group of 12 words attested beside Latin in Celtic and Germanic to the exclusion of Greek (*caput*, *catulus*, *caulae*, *corbis*, *far*, *fulica*, *hasta*, *merula*, *nux*, *raudus*, *trabs*, *ulmus*). Of these, *caput*, *corbis*, *fulica*, *hasta*, *merula*, and *nux* are relatively securely Latin-Celtic-Germanic isoglosses. *Merula* attests to *a*-prefixation (for which there are examples of Greek participation as well) and the velar of *fulica* corresponds to Germanic in the same way as other lexemes (*filix* and *sōrex*) that have

Greek comparanda, suggesting that these words were borrowed from a language with which Greek was at one point in contact as well. The alternations attested for *corbis* and *hasta* are mainly of voicing and aspiration, types also attested in lexemes for which there are Greek comparanda. It is *caput* and *nux* that furnish the best evidence of a substrate language with which Greek was never in contact. Their comparanda show a dental ~ velar alternation that is further attested in at least Baltic (van Sluis fthc., see §3.2.1.2.7) but not Greek. The only case that might have a Greek comparandum is that of Lat. *pīx* ~ Gk. *πίτρος* and their comparanda.⁵³⁸ Given the semantic relationship between pitch and pine trees, the connection is attractive. It would be more certain if both meanings were attested for both forms. As it stands, all forms meaning pitch can be reconstructed to **k* and for pine to **t*. Thus, the lexemes may truly be unrelated leaving Greek with no examples of a dental ~ velar alternation.

This suggests a situation in which Italic, Celtic, and Germanic (along with at least Baltic) were in contact with a substrate language of Europe with which Greek was not in contact. When exactly this would have occurred depends on how early or late the loanwords in §4.3.2.2 entered Greek. Considerations on Italo-Celtic make this a complicated question to answer.

4.3.2.3.2 *Italo-Celtic Isoglosses and the Italo-Celtic Subnode*

There are three lexemes in the dataset that are attested in Italic and Celtic but not Germanic or Greek. These are *badius*, *sappīnus*, and *bāca*. The first two are Italo-Celtic isoglosses while the latter has only potential Berber comparanda (unless it is ultimately related to Lat. *vaccīnium*). This small number (3% of the data) is probably significant. It may in part be due to lexical loss in the Insular Celtic languages, our best source of Celtic lexical material. But it may well also have to do with the potentially reconstructible Proto-Italic-Celtic subnode. Considerations on the reconstruction of an Proto-Italic-Celtic subnode have important consequences for the stratification of the data. If Proto-Italic-Celtic persisted throughout the period in which loanwords were entering the Indo-European daughter languages, then forms with irregular alternations should be few and due to post-split borrowings. The cases of *badius*, *sappīnus*, and *bāca* are indeed few. The latter two are also plausibly late loans since they attest to irregular alternations even within Italic (though they may have been borrowed at different times, see §4.3.1.2). Some Italo-Celtic isoglosses that reconstruct to the same proto-form (and thus are not included under §2.2.2) may be the result of substrate words borrowed into still unified Italo-Celtic (cf. *caelum*, *hirūdō*; see Stifter fthc. on the latter).

The picture presented by the rest of the Celtic data however seems paradoxical. Of 26 words for which Celtic attests comparanda, only 5 cases (*caput*, *caulae*, *fascis* [as a

⁵³⁸ This also relies on Gk. κεφαλή 'head' < **g^heb^hh₂-l-* being only coincidentally similar to the *caput* group. The presence of a cognate in Tocharian (Toch. A *špāl* 'head', demanding the reconstruction of a laryngeal that the Greek does not otherwise require) is indeed strongly indicative of an unrelated, inherited formation.

relative of *baiulus*], *hasta*, *merula*) can be reconstructed to the same pre-form in Proto-Celtic and Proto-Italic. Most of these are in the Italic-Celtic-Germanic layer excluding Greek which, as argued above, seems likely to be more recent than the layer including Greek. For *tilia* and *rāpum* for instance, Celtic attests to forms with an *a*-prefix against all others. This would imply a contact situation like in Figure 4.9, where Proto-Italo-Celtic and Proto-Germanic had migrated into contact with a language with which Greek did not have contact before reaching their eventual places of attestation and receiving loans from a pan-European substrate language *in situ*. As mentioned above, if the distribution of the *a*-prefix and *n*-suffix includes the Mediterranean, this situation is possible. The Mediterranean loans in Latin would thus have entered Italic in approximately the same region as the other loans, suggesting much linguistic diversity.

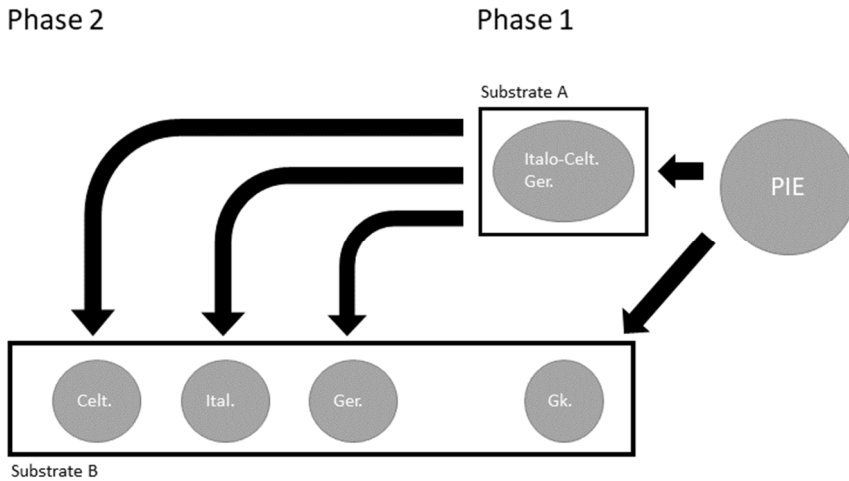


Figure 4.5 One interpretation of contact incorporating Italo-Celtic considerations

Some support of this alternative timeline is given by Stifter (fthc.), who notes that the loss of QPIE **p* in PCelt. **kaϕuto-* (belonging to the stratum excluding Greek) makes it look older than PCelt. **arbīno-* (with widespread comparanda including Greek) if the **b* of the latter indeed represents an attempt at rendering a /p/ after that sound was lost in Celtic. Given the difficulty in being able to confirm this (note the warning of Huld 1990: 394-5 mentioned by Stifter fthc. fn. 24), other potential scenarios are possible.

While an Italo-Celtic subnode is not fully accepted by all, Italic does indeed seem to share more innovative features with Celtic than with other branches (Weiss 2022a: 108). It is not certain how early or late Proto-Italo-Celtic would have split into Proto-Italic and Proto-Celtic. Figure 4.10 illustrates a possibility in which Proto-Italo-Celtic split up early

on. The Italic-Celtic-Germanic lexemes would then have been borrowed in a contact situation that occurred *after* Greek had migrated to an area where this language was not present. Proto-Italic and Proto-Celtic speech communities would theoretically have been in close contact near the Alps at a recent date, and the small number of exclusively Italo-Celtic irregular isoglosses would be the result of 1) close proximity to the same

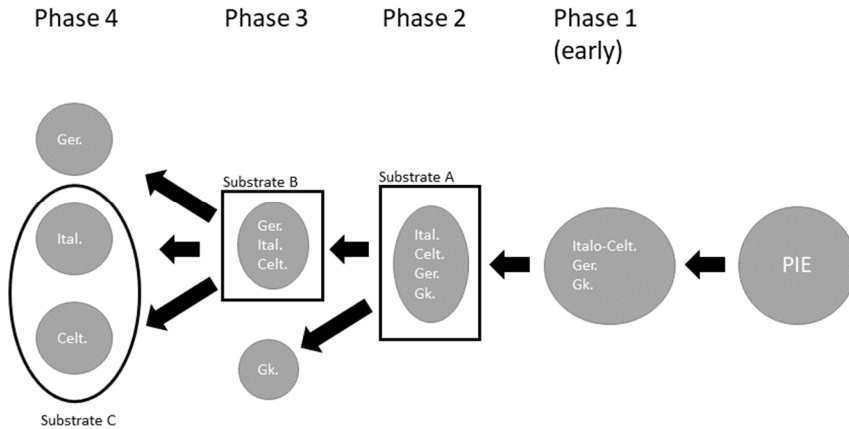


Figure 4.6 An alternative scenario with an early Italo-Celtic split and later close contact

substrate (resulting in many loans entering the languages the same way, mimicking Italo-Celtic unity), 2) lexical loss, and 3) this contact occurring in a part of Europe where there were few speakers of non-IE languages left (cf. Stifter *ftbc.* on the latter point).⁵³⁹ After subsequently entering the Mediterranean region, Italic re-entered a situation where it was in contact with some of the same languages as Greek.

It remains difficult to decide which scenario best represents the contact situations that resulted in the attested data. But the important results remain: very few cases of non-IE loanwords into Italic and Celtic can be reconstructed to the same proto-form and there are very few cases of exclusive Italo-Celtic irregular isoglosses.

4.3.2.3.3 *Italo-Germanic Group*

The comparanda shared between Latin and Germanic to the exclusion of Celtic and Greek (*alnus*, perhaps *avēna*, *barba*, *cucurbita*, *excetra*, *faba*, *ferrum*, *fracēs*, *raia*,

⁵³⁹ Noting cases of alternation within other branches (like Lat. *caulae* ~ Lat. *cohum*; Lat. *trabs* ~ Osc. *trībūm* ~ U *trebeit*; PGM. **habuda-* ~ PGM. **haubuda-*) another alternative is that Proto-Italo-Celtic could have existed in theory up to a recent date, with irregular alternations between comparanda being the result of borrowings from highly regionally diversified dialects of the contact languages. This is the most problematic possibility, as alternations within branches are generally assumed to represent the latest borrowings.

sulpur) often have a strong connection to differentiated Baltic and Slavic (*alnus*, *avēna*, *barba*, *excetra*, *faba*, *fracēs*). That they demonstrate several types of substrate alternations that are found elsewhere (*n*-suffix: *alnus*, *avēna*, *faba*; *s* insertion: *barba*, *fracēs*; *SK* metathesis: *excetra*) makes their analysis difficult. How many such cases are due to lexical loss in other branches that could originally have attested them? How many are due to contact with separate languages that are nonetheless related to other substrate languages of Europe? *Raia*, *sulpur*, and possibly *cucurbita* are Italo-Germanic isoglosses.

4.3.2.3.4 Germanic Loans in General

It is interesting to note that most of the Germanic comparanda, where they contain a diagnostic phoneme, can be shown to have entered Germanic before the operation of Grimm's and Verner's Laws. It is a pre-Proto-Germanic stage when their consonantism can be reconstructed to match that of the other comparanda most closely:⁵⁴⁰

Lat.	PItal.	QPIE for PGm.	PGm.	Other Comp.
<i>aper</i>	* <i>apro-</i>	* <i>h₁ep-<i>r</i>-</i>	* <i>ebura-</i>	PGk. * <i>epero-</i>
<i>ardea</i>	* <i>ard-</i>	* <i>h_{2/3}erd-</i>	* <i>artō(n)-</i>	PGk. * <i>erōd-</i>
<i>baculum</i>	* <i>bak-</i>	* <i>bHK-'</i>	* <i>pagjō-</i>	PGk. * <i>bak-</i>
<i>caput</i>	* <i>kaput-</i>	* <i>kHput-</i>	* <i>habuda-</i>	PCelt. * <i>kaϕuto-</i>
<i>caulae</i>	* <i>kaγ-</i>	* <i>kog^h-</i>	* <i>haga(n)-</i>	PCelt. * <i>kagyo-</i>
<i>faba</i>	* <i>b^hab-</i>	* <i>b^hh₂eu-n-</i>	* <i>baunō-</i>	PSlav. * <i>bòbъ</i>
<i>far</i>	* <i>fars-</i>	* <i>b^hHr(V)s-</i>	* <i>bariz-</i>	PCelt. * <i>baragi-</i>
<i>ferrum</i>	* <i>fersom</i>	* <i>b^hros-</i>	* <i>brasa-</i>	
<i>pannus</i>	* <i>panno-</i>	* <i>pHn-</i>	* <i>fanan-</i>	PGk. * <i>pāno-</i>
<i>rādīx</i>	* <i>wrād-</i>	* <i>ureh₂d-</i>	* <i>wrōt-</i>	PCelt. * <i>wrad-</i>
<i>raudus</i>	* <i>raudo-</i>	* <i>h₂erud-</i>	* <i>arut-</i>	
<i>trabs</i>	* <i>trab-</i>	* <i>trb-</i>	* <i>purpa-</i>	PCelt. * <i>treb-</i>
<i>turdus</i>	* <i>to/ur(z)do-</i>	* <i>trosd-</i>	* <i>prastu-</i>	PCelt. * <i>trozdi-</i>

Table 4.1 Consonantism correspondences between Germanic and other branches

Importantly for the nature of the irregular velar suffix (§3.3.3), the correspondence between Lat. *fulica* ~ OHG *belihha* and Lat. *filix*, *felix* ~ PGm. **brekna(n)-* is best reconstructed to initial QPIE **b^h*, requiring a borrowing before the operation of Grimm's Law. Thus the velar suffix that entered Latin and Greek with **k* entered Germanic as voiced **g*.

In a similar vein, PGm. **hadelō-* and **hadnō-* must have entered Germanic pre-Grimm

⁵⁴⁰ A note of caution is that it cannot be ruled out that reconstructions involving the operation of Verner's Law might be misinterpretations. PGm. **pagjō-* can be reconstructed to QPIE **bHK-* to match the Italic and Greek comparanda (as in Table 4.1), but the geminate **kk* of PCelt. **bakko-* corresponds to PGm. **g* < QPIE **g^h* in e.g. PGm. **managa-* 'many' ~ PCelt. **menekki-* 'abundant'. On the other hand, the correspondence is with PGm. **k* < QPIE **g* in PGm. **balikōn-* ~ PCelt. **bo/ula/okk-* 'coot'.

as Italic and Celtic both attest a reflex of PIE **k*. But as to their second consonant, do they go back to **kHd^h-* to match PCelt. **kadVlot-* or to Verner variants of **kHt-* to match PItal. **katVlo-?* In light of PCelt. **arbino-* and Gk. ῥάφους, PGm. **rōbjōn-* may have entered with **b^h* or with **p* (like all other comparanda).

PGm. **kulubrōn-* may attest to QPIE **g* like its Balto-Slavic comparanda, or it may represent a post-Grimm borrowing of **k* like Lat. *columba*. Two cases indeed suggest that a word entered Germanic both before *and* after the operation of Grimm's Law: PGm. **pakka-* (< QPIE **bHG-*) ~ PGm. **bagg-* and PGm. **hrep-* (< QPIE **kreb-*) ~ PGm. **krebō-*.⁵⁴¹ The alternation between PGm. **rugg-* and PGm. **rehhōn-* cannot be explained this way.

That no Germanic comparanda of Latin words securely and exclusively attests to a post-Grimm borrowing suggests that Italic ceased to have contact with the same substrate languages as Germanic before the operation of Grimm's Law. This comes as no surprise if Grimm's Law operated in the mid-first millennium BCE.

4.3.2.3.5 *Other Groups*

There are 4 words attested in Italic, Celtic, and Greek to the exclusion of Germanic (*caballus*, *calpar*, *lapis*, *paelex*). *Caballus* has widespread attestation that points to a more recent Wanderwort. *Calpar* has a Mediterranean distribution save for its Celtic attestation, and van Sluis (fthc.) proposes it might be a Wanderwort into Celtic transmitted by Etruscan. *Paelex* and *lapis* seem also to have been in currency in the Mediterranean (*paelex* perhaps more so given its potential borrowing into Semitic); their presence in Celtic has unclear implications. The Hallstatt and later La Tène cultures indeed maintained trading networks with Greeks and Etruscans (cf. Kristiansen 1998: 292-3), but this seems too late for the Proto-Celtic loss of **p* in all of these words. Otherwise, there seems to be little evidence of Celtic participation in the Mediterranean substrate(s).

Several other words are attested in Italic, Germanic, and Greek to the exclusion of Celtic (*ascia*, *?columba*, *ervum*, *filix*, *fungus*, *sabulum*, *sōrex*, *aper*, *ardea*, *pannus*, *viscum*). It is unclear why Celtic does not have these. If the number of exclusive Italo-Celtic irregular isoglosses is so small due to lexical loss (i.e. replacement), then perhaps this is the case here too: the Celtic languages may well have replaced what would otherwise be loans from the same contact situation that produced widespread loans like *rāpum* and *tīlia*. This is suggested by at least *filix* and *sōrex* (whose irregular correspondence between Latin and Germanic is shared by *fulica*, for which there is a Celtic comparandum) and *ardea* (with an *a*-prefix, a pattern with examples elsewhere in Celtic). The *SK* metathesis of *ascia* and *viscum* also occurs in *mūlus* and possibly in *tamarix* and *excetra*, also

⁵⁴¹ For the latter form, evidence against a later borrowing is the great variation within the Proto-Germanic forms such that they point to an ablauting *n*-stem, making them archaic (Kroonen 2011: 179-82, 2013: 303).

without Celtic comparanda. But PGM. **bahsu-* ~ PCelt. **tazgo-*, **tasko-* ‘badger’ makes it difficult to propose that this is the result of a contact situation in which Celtic did not take part.

4.4 A More Inclusive Visualization

The importance of Greek, Celtic, and Germanic for the major distributional tendencies is clear, but this tripartite representation does not effectively represent the placement of all the other languages for which comparanda exist. This is most easily visualized with the assistance of multidimensional scaling.⁵⁴² Plotted in the following diagrams are the 61 words from §2.2.2.1 *Non-inherited Origin is Probable*.

The plots are Principal Component Analyses (PCAs) using the lexemes as data points (rows) and their languages of attestation as properties (columns). This takes the form of presence (1) and absence (0) data. However, a traditional PCA treats two identical values as the same; thus two zeros would be interpreted in the same way as two ones. This is clearly inaccurate. (The absence of one particular lexeme in two or more languages does *not* mean that they share an alternative lexeme in its place. In reality, each language that does not have the lexeme in question usually has a different alternative lexeme.) In this way, the dataset is similar to datasets on species abundance in the field of Ecology, where Legendre and Gallagher (2001) recommend several transformations of the data before it is put into a PCA. A Hellinger transformation is suited for data like this with low counts and many zeros. Thus the presence-absence tables were imported to R, Hellinger transformed (“vegan” package `decostand(“hellinger”)`), visualized (“ggplot” package `autoplot`), and re-labelled (“ggrepel” package `geom_text_repel`). Because the discussion of the lexemes involved uncertain comparanda, the analysis was performed twice: a loose plot (with the uncertain comparanda included) and a strict plot (with only secure comparanda included).

Figure 4.11 displays the loose analysis, in which Principal Components 1 and 2 account for 50.74% of the variation. The vectors (arrows in red) help visualize in space the importance of the major effect of Greek as opposed to Germanic and Celtic. However it also visualizes the effects of Baltic and Slavic.

⁵⁴² The statistical portion of this section has been greatly informed by discussions with Paulus van Sluis.

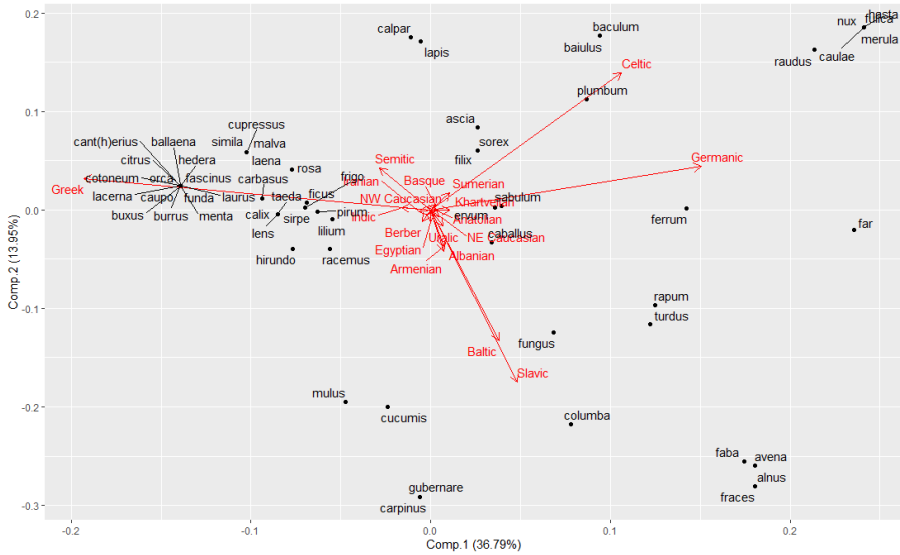


Figure 4.7 PCA of loose analysis

Figure 4.12 contains the strict analysis, in which Principal Components 1 and 2 account for an increased 60.79% of the data. The overall trend remains the same as in the loose plot, with a similar importance played by Greek, Celtic, Germanic, Baltic, and Slavic. Very visible is the Mediterranean cluster, with a major role in placement played by Greek, along with Semitic. The widespread nature of the rest of the data points matches many of the conclusions from above: that there are likely several other contact situations represented by the data. How many different situations are represented as opposed to how many different clusters actually attest to singular large contact situations with subsequent lexical replacement is difficult to confirm.

While the PCAs assist in the visualization of distributional trends, a major caveat is indicated by *plumbum*, *baculum*, and *baulus* sharing the same position in Figure 4.12: Two lexemes in close proximity share similar distributions between languages, but this is not proof they entered Latin as the result of the same contact situation or that they arose in the same source language.

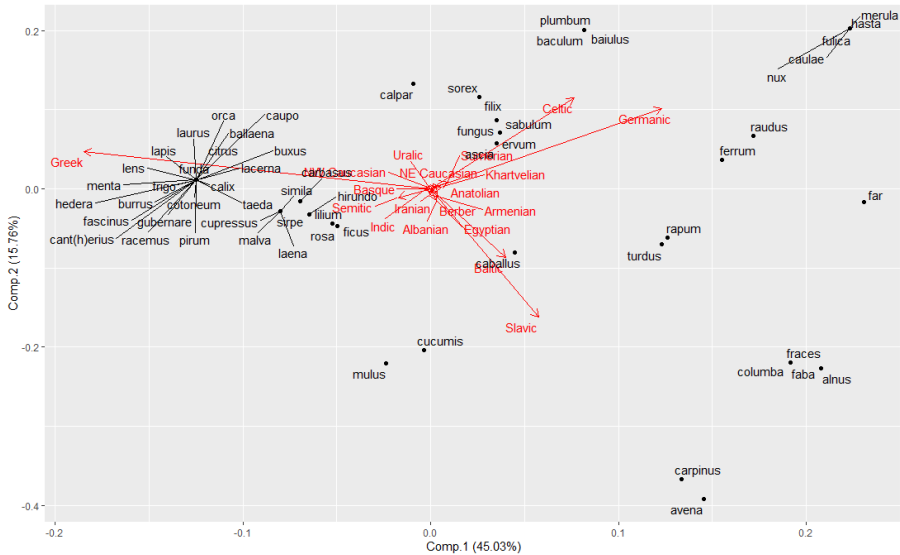


Figure 4.8 PCA of strict analysis

4.5 Summary of Stratigraphy

Good evidence of the very earliest loanwords is partially obscured by the methodology of this thesis, in that loanword status is based on irregular alternations between comparanda. The earliest loanwords would show little variation. *Taurus* is one example, but it happens to be a more recent loan into Latin.

After the split of the Indo-European daughter languages, the stratification of contact situations directly following is difficult and depends in part on whether Proto-Italo-Celtic split very early or late. If early, loanwords with a widespread distribution including Greek are likely some of the earliest. This is followed, as indicated by weak evidence (a velar ~ dental alternation), by a period where Italic, Celtic, and Germanic entered a contact situation with which Greek was not a part. If Proto-Italo-Celtic persisted to a more recent date, then the loanwords with a widespread distribution of attestation may in fact be due to late, *in situ* borrowings from a very widespread, entrenched network of languages. In any case, Italic was also in contact with the same substrate language(s) as pre-Grimm's Law Germanic and separate Baltic and Slavic.

There are very few demonstrable cases of loanwords shared exclusively with Celtic, a fact that must be considered along with the details of the dating of Proto-Italo-Celtic.

Proto-Italic reached the Mediterranean region, at which time it began to borrow lexemes from the same languages as Greek as well as receive loanwords from Greek that were transmitted indirectly through other languages. There is evidence of at least one distinct substrate language of the Mediterranean, along with the possibility of others. Borrowings from these as yet identified sources into Italic languages in the Mediterranean region continued from the period of Proto-Italic until at least the third century BCE. At some point, it seems that some varieties of Latin reborrowed some of the same words at a later date, possibly resulting in Romance forms that attest to geminates.

5 Semantic Analysis

The relevance of semantic category to the study of substrate words has been mentioned several times (cf. Schrijver 1997: 295, who notes that words denoting economically and culturally unimportant animals and plants are possibly more likely to have a substrate origin). As argued in §1.4.2.1, semantic considerations are best kept for after non-inherited material has been identified rather than used to identify it. Now that the non-inherited material has been identified, it will be worthwhile to examine the semantic categories to which it belongs. While it is difficult to know exactly what semantic classifications to assign certain objects and technologies from an ancient culture, I have tried the following:

Animals		21	
	Domestic		4 (<i>caballus, asinus, cant(h)erius, mūlus</i>)
	Birds		6
	Aquatic		2 (<i>ardea, fulica</i>)
	Other		4 (<i>columba, hirundō, merula, turdus</i>)
	Other Wild		10
	Aquatic		3 (<i>ballaena, excetra, raia</i>)
	Insect		3 (<i>arānea, asīlus, tabānus</i>)
	Other		4 (<i>aper, lepus, sōrex, talpa</i>)
	Unspecified		1 (<i>catulus</i>)
Plants		44	
	Domestic		11 (<i>ālium, avēna, cēpa, cucumis, cucurbita, ervum, faba, far, lēns, nāpus, rāpum</i>)
	Other		33
	Tree		17
	Fruit		6 (<i>arbutus, citrus, cotōneum, fīcus, pirum, sambūcus</i>)
	Other		11 (<i>alnus, buxus, carpinus, cerrus, cupressus, laurus, sappīnus, taeda, tamarix, tilia, ulmus</i>)
	Wild		8 (<i>alaternus, filix, genista, hedera, malva, menta, sirpe, viscum</i>)
	Flower		2 (<i>līlium, rosa</i>)
	Berry		2 (<i>bāca, vaccīnium</i>)
	Plant part		2 (<i>rādīx, trabs</i>)
	Nut		1 (<i>nux</i>)
	Fruit		1 (<i>bolunda</i>)
Metallurgy			4 (<i>faber, ferrum, plumbum, raudus</i>)
Textiles			4 (<i>carbасus, lacerna, laena, pannus</i>)

Tool	7 (<i>ascia, baculum, baiulus, caulae, corbis, funda, hasta</i>)
Vessel	3 (<i>calix, calpar, orca</i>)
Viticulture	5 (<i>fracēs, faex, lābrusca, racēmus, pampinus</i>)
Geography	5 (<i>calx, grūmus, lapis, sabulum, sulphur</i>)
Culinary	3 (<i>adeps, frīgō, simila</i>)
Color	2 (<i>badius, burrus</i>)
Body Part	3 (<i>barba, caput, grāmiaie</i>)

Other words are more difficult to classify: *casa* is architectural, *focus* domestic, and *fascinus* magico-religious. *Fungus* is a fungus and could serve as a tool. *Caupō* is an economic profession (cf. *baiulus* classed under tool due to its relationship to *fascis* as well as *faber* classed under metallurgy) and *paelex* a person specified by their role in society. *Farciō* and *gubernō* are actions, the latter nautical.

It is clear that the non-inherited lexemes of Latin are indeed overwhelmingly plants (40%) and animals (19%), though they are certainly not all economically unimportant. Beyond being unable to say with certainty which animals and plants would have been economically unimportant to ancient peoples, several of the words refer to domesticated or otherwise edible species. There are also several words referring to items of material culture, including tools, vessels, and textiles.

One further piece of evidence that may allow us to identify to which sort of contact situations the different identified strata belong is the consideration of the semantic categories present in the different strata. This is partially difficult given the greatly unequal distribution of the lexemes in the different strata. Given that 62% of non-inherited Latin lexemes have Greek comparanda (and 42% of lexemes have Greek comparanda to the exclusion of Germanic and Celtic), it is statistically more likely that strata involving Greek will attest to more semantic categories. Even so, there are some potential patterns in the data.

Firstly, amongst the 15 non-inherited words limited to Latin (and the Romance languages) are found 3 of the 5 viticultural words (*faex, lābrusca, pampinus*). As to the other 2, *racēmus* is (probably) only shared with Greek and *fracēs*, with comparanda in Albanian, Germanic, Baltic, and Slavic, can also refer to other sorts of sediment in liquid. Thus it might not be a strictly viticultural term after all. Interestingly, 2 of the 3 insects (*asīlus, tabānus*) are also restricted to Italic.

Another pattern is found amongst the words attested in Celtic and Germanic to the exclusion of Greek. In §4.3.2.3.1, it was shown that many of the non-inherited words with this distribution likely belonged to languages at least related to those with which Greek had contact. But interestingly, in this group there are no words for domesticated

animals or ceramic vessels. Of 11 non-inherited words for domesticated plants, the only one in this group is *far*. It is unclear how much faith to place in this pattern. There are few confirmed loanwords for domesticated animals in Latin (an interesting fact on its own), thus it is perhaps not surprising that this stratum group does not attest to any. As to the lack of ceramic vessels, all 3 of those words are shared with Greek (*calix* and *orca* exclusively, *calpar* along with Celtic). This might suggest that any borrowed ceramic terms were replaced due to later heavy cultural participation in the Mediterranean sphere. All 3 culinary terms and 3 of 4 textile terms (*carbasus*, *lacerna*, *laena*) likewise show a Mediterranean distribution. The lack of words for domesticated plants seems more likely to be significant and to perhaps indicate the sort of contact situation in which these words were borrowed.

In fact, *far* has further comparanda in Slavic, indicating that the 12 words in the Italic-Celtic-Germanic (non-Greek) group do not all attest to the same contact situation. If only the isoglosses are considered (*fulica*, *merula*, *caput*, *nux*, *hasta*, *corbis*), the pattern of the absence of domesticated semantics is stronger.

The 11 non-inherited words for domesticated plants as a group are important, in that they seem to confirm that a portion of the non-inherited vocabulary in the Indo-European languages was indeed borrowed from a population practicing intensive agriculture (cf. Kallio 2003, Schrijver 2007: 21-2, Kroonen 2012a, Iversen & Kroonen: 516-18, Šorgo 2020: 461-8). As noted in §4.5, the stratification of contact phenomena represented in Latin loanwords prior to the arrival of Proto-Italic in the Mediterranean region is difficult. It is not made any clearer by the distribution of the lexemes for domesticated plants. Some of the words show a Mediterranean distribution (*cucumis*, *ālium*, *cēpa*, *nāpus*, and even the founder crop *lēns*). Others attest to a wider European distribution (always including Germanic, often Baltic and Slavic) more in line with what one might expect from loans originating in a Europe-wide agricultural substrate. That sometimes comparanda do not exist in Greek (*faba*, *cucurbita*) or Celtic (*faba*, *cucurbita*, *ervum*) suggests that these absences might be due to later lexical replacement.

6 Population Genetics of Italy

6.1 Genetics Introduction

The genetic diversity of today's Italian population attests to a history characterized by human migrations. Each event contributed to the complex mosaic of alleles, beginning with the first anatomically modern humans to arrive on the Italian peninsula and continuing to this day. Some historical events are well documented, such as the migration during the Roman Empire and after its fall as well as the settlement of ethnic Albanians and Greeks (both ancient and medieval). Others have only been uncovered due to the work of archaeologists and geneticists.

The term Indo-European is first and foremost a linguistic one, and languages are spoken by people. But given the landmark publications that revolutionized the field of research beginning in 2015 (discussed in detail later), a connection between the Indo-European proto-language and the populations of the Pontic-Caspian steppe is becoming the dominant theory in the research on the Indo-European homeland. No discussion of Indo-European homelands and migrations can be complete without a discussion of these results.

This chapter is an excursion into the background and results of genetic research on Italian populations, serving as a segue between the linguistic patterns identified and the final conclusions. After setting the background for the population genetics of Europe, this chapter details the development of Italian population genetics research with the goal of ultimately identifying the genetic events that can be linked to the arrival of the Indo-European languages in Italy.

6.2 European Genetics from the Origins of Agriculture to the Homeland Debate

6.2.1 Understanding the Origin and Spread of Agriculture

Archaeological inquiry has been making headway into the origins of agriculture for over a century, with the term 'Neolithic Revolution' being coined in 1936 by V. Gordon Childe. We know agriculture developed in Southwest Asia (gradually and originally regionally, cf. Arranz-Otaegui et al. 2016), spreading first to Cyprus, Crete, and mainland Greece (cf. Hofmanová et al. 2016, Douka et al. 2017), then forking into two routes. Via the Cardial Ware Culture, agriculture reached the Mediterranean coast and via the Linearbandkeramik Culture (LBK), it spread through mainland Europe.

Understanding how agriculture spread is not as easy a question to answer. Scholarly opinions differed as to whether farmers from the Near East themselves migrated into Europe (demic diffusion) or whether the technological aspects of farming were adopted

and passed on by the people already living in Europe (cultural diffusion). In 1971 Albert J. Ammerman and Luigi Luca Cavalli-Sforza put forward an argument in favor of the former. They plotted the radiocarbon dates that had been collected for the earliest appearance of agriculture at several archaeological sites across Europe against the distance of those sites from areas in the Near East. The results showed a slow but steady 1 km/year spread of the technology of farming from Southwest to Northeast across Europe. Based on anthropological and archaeological observations along with statistical models, this suggested a “wave of advance” model best explained by demic diffusion rather than cultural diffusion.

Cavalli-Sforza was an influential force in the field of population genetics. In 1964⁵⁴³, with Anthony Edwards, he used 18 alleles of classical markers (mainly blood groups and antigen systems) to examine how 15 widespread population groups in the world diverged from each other. They popularized the then only infrequently used principal components analysis (PCA) and worked on the best ways to generate trees (Edwards and Cavalli-Sforza 1965, 1967). In the 1971 paper that used archaeological material, he and Ammerman had already postulated how genetic information could be used to inform their work in the future. In 1975, in a report with Alberto Piazza, the results of a tree analysis using 58 alleles and a larger number of populations basically corroborated the results of the 1964 study. The distribution of classical markers amongst modern populations seemed to be powerful enough, when used in combination with each other, to perform large-scale studies of the history of human migrations. Interestingly, up until this point, the main focus of the articles was on the development and presentation of the statistical methods. The genetic analyses were used as examples to show how the methods worked.

In 1978, Menozzi, Piazza, and Cavalli-Sforza did what the 1971 article had proposed, applying genetic methodology to the question of the spread of agriculture. The statistical approach they employed had been used for clustering analysis in genetics before, but this time they would display the results in the form of a map. The distribution of individual classical markers across Europe was beginning to be understood by this time. The first principal component of 38 genetic alleles analyzed produced a pattern that matched the Southeast to Northwest cline of the spread of farming known from archaeological research. This first principal component explained a full third of the genetic variation. Assuming that the hunter-gatherers of Europe and the agriculturalists of the Near East had been in isolation from each other long enough to be genetically quite different, this cline across the geography most likely represented the result of one population beginning in the Near East, and spreading to the Northwest, intermarrying with local populations along the way such that it came to be genetically more and more similar to those populations. In other words, this is precisely what one would expect if farming had spread via demic rather than cultural diffusion.

⁵⁴³ Edwards and Cavalli-Sforza 1964; after two unpublished conference talks in 1963.

The second principal component revealed a general East to West cline and the third principal component revealed a Northeast (centering in Ukraine) to Southwest cline, with all three together explaining 50% of the variation. It was clear that more investigation was needed to fully understand what had occurred, but the stage had been set for the use of genetics to understand the population history of Europe. Ammerman and Cavalli-Sforza augmented their wave of advance model with statistics and examples from archaeological investigation in 1979 and combined this augmented approach with the genetic data and a detailed explanation of how it was analyzed in a capstone book in 1984.

6.2.2 The Indo-European Homeland Debate

Ever since the realization that the relatedness of the Indo-European languages meant that they had all descended from a parent language, hypotheses about the homeland of this language came into circulation. This homeland or *Urheimat* was the geographic location where the speakers of the reconstructed proto-language would have lived. In the century or so of research, several locations have been proposed as the potential homeland. The two that have widely been considered the most compelling are the “kurgan hypothesis” (particularly its later revision as the “steppe hypothesis”) and the “Anatolian hypothesis”.

The earliest concrete codification of the kurgan hypothesis came in the form of archaeologist Marija Gimbutas’ 1956 book *The Prehistory of Eastern Europe, Part I*. In it and its further developments, she proposed the umbrella term ‘Kurgan Culture’ to denote several archaeological cultures in Eastern Europe in the 4th and 3rd millennia BCE. They all seemed to share features (including the construction of burial mounds called kurgans) and overlap in terms of movement and development through time. The changes in settlement patterns of the cultures of the kurgan horizon including, at some stages, drastic expansion, led Gimbutas to propose a series of waves of expansion, some of which were responsible for bringing the Indo-European languages along with them. Schrader (1883) had already argued on the basis of linguistic methodology, which would come to be called linguistic paleontology, for a similar Pontic-Caspian homeland for Indo-European.

In 1987, a strong contender to the kurgan hypothesis was put forward by archaeologist Colin Renfrew in his book *Archaeology and Language: the Puzzle of Indo-European Origins*. The book was very much in line with the genetic discoveries that had been made up until that point. If there had indeed been a large-scale migration of people from the Near East during the Neolithic, we can assume that they would have brought their language with them. They brought farming technology with them that resulted in geographically enormous cultural horizons that spread across Western Europe. If we are looking for a group of people spreading in unison with closely related languages across the European continent, we need look no further than this. In proposing that the Indo-European languages spread with Neolithic farmers from Anatolia, this position came to be called the Anatolian hypothesis.

Despite seeming like a smoking gun both archaeologically and genetically, there were major problems with the Anatolian hypothesis from the linguistic side. In a line of research that has culminated most recently in Kroonen et al. (2022), the society for which Proto-Indo-European is reconstructed could not have been fully agricultural.

The kurgan hypothesis as laid out by Marija Gimbutas, while generally in agreement with linguistic data, received criticism in the field. Older explanations for culture change that frequently involved the assumption of migration had given way to a swing of the pendulum in which such explanations were viewed with suspicion (cf. Anthony 2007: 214).⁵⁴⁴ But given the linguistic problems with the Anatolian hypothesis, the field of Indo-European studies never fully rejected Gimbutas' hypothesis. Jim Mallory's 1989 *In Search of the Indo-Europeans: Language, Archaeology and Myth* followed the explanation she presented, and David Anthony's *The Horse, the Wheel and Language* presents a revised steppe hypothesis.

6.2.3 Refining and Overturning the Understandings

While archaeologists and linguists were refining and expanding homeland theories, geneticists continued, with the aid of improving technology, to sample and sequence human genomes.

6.2.3.1 Building Genetic Databases

With the hypothesis that farmers had spread into Europe from the Near East via demic diffusion, i.e. in person, during the Neolithic, it was initially unclear how much of a genetic footprint they would have left on the population. Richards et al. (1996) examined a sequence of mitochondrial DNA in living individuals from Europe and the Middle East. The divergence times that they calculated for the lineages led them to conclude that most mitochondrial alleles spread to Europe during the Upper Paleolithic, with only around 10% of the alleles dating to the spread of farming. Perhaps farming had been more of a local development after all. Cavalli-Sforza and Minch (1997) disagreed with the methodology behind this conclusion. The segment of mitochondrial DNA used was subject to higher rates of mutation, and the mitochondrial patterns might actually be revealing a sex-bias in terms of who was migrating. Staying true to the earlier conclusions about demic diffusion, Cavalli-Sforza and Minch estimated the genetic influence of farmers at the more significant but still quite small amount of ca. 27%. Barbujani, Bertorelle, and Chikhi (1998) sided with the latter on this, emphasizing the archaeological data on the spread of farming and the arguments generated by linguists on the post-Neolithic age of many of the language families in question. Additionally, they argued that the age of a haplotype does not equate to the age of the population in which that haplotype is found, especially when the haplotype is not found exclusively within

⁵⁴⁴ There was less ideological criticism of her work too. Anthony (2007: 306-7) for example points out the problematic way that she grouped several different burial mound-building archaeological cultures into her one "Kurgan culture".

that population. This would become an important argument in understanding the data, repeated for instance by Chikhi et al. (1998) when they found using modern autosomal DNA that all population separation times were less than 10 kya and thus unlikely to be pre-Neolithic.

Torroni et al. (1998) examined modern European mitochondrial DNA and proposed that it demonstrated a large genetic component from a Mesolithic (post-glacial) Iberian population. Mitochondrial haplogroup V was found in its highest concentrations among Berbers, Basque, and Saami, and so was most likely to have spread from an Iberian refugium after the ice sheets melted. Its sister haplogroup H, though it originated in the Near East, seems to have spread to Europe during the Paleolithic, perhaps along with V from Iberia. This was countered by Simoni et al. (2000a; criticized by Torroni et al. 2000 and defended in Simoni et al. 2000b). They found that the Saami were mitochondrially quite unique, and their inclusion in analysis was skewing the data. When removed, European populations were mitochondrially extremely similar to one another, with little geographic structuring. However, when Southern Europe was examined separately, there was a clinal distribution. While the alleles and their frequencies did not differ between North and South, along the Mediterranean, they were distributed in a very structured way. In combination with the autosomal cline from the Near East to Northwest Europe, the results in theory only supported a large-scale directional expansion. If all human populations had been forced into glacial refugia and re-expanded during the Mesolithic, this would have erased the original Paleolithic distribution patterns. As the patterns persist to the present day, they could most likely only have resulted from a directional expansion during the Neolithic.

The debate continued. Richards et al. (2000) took back-migration into the Near East from Europe into account, suggesting that the Neolithic migration contributed at most 20% of the mitochondrial genepool of Europe—substantial but in the minority. Most of the mitochondrial haplogroups arrived in Europe during the Paleolithic and their distribution dates back to founder effects during the post-glacial re-expansion. Using Y-chromosome data, Semino et al. (2000) agreed. They found that ten lineages account for more than 95% of the Y-chromosomes in their study. Clustering analysis seemed to identify two Paleolithic post-glacial migration events and one Neolithic migration event, with most of the genepool nevertheless being of Paleolithic origin. Chikhi et al. (2002) concluded that it was instead the *Paleolithic* component of the genepool that was less than 30%. They used the Y-chromosome dataset from Semino et al. (2000) but instead of looking for clines, they directly calculated admixture with Near Eastern populations. With an average of 50% across samples, they found European populations closest to the Near East had a Neolithic component of between 85 and 100%, decreasing by France, Germany, and Catalonia to 15-30%.⁵⁴⁵ They also identified differences between

⁵⁴⁵ These results should however be taken with caution, as the descendant population used to represent the Paleolithic genome was the Basques, which was a big assumption. When Sardinia was used as the Paleolithic descendant population, the admixture with Neolithic genes increased to 65%.

Mediterranean and non-Mediterranean populations.

6.2.3.2 A Paradigm Shift Waiting to Happen: the Sequencing of Ancient DNA

It is crucial to bear in mind that all results up until this point had been reached through analysis of modern populations. The problem with this is that the modern population is the result of every migration and all the gene flow that has occurred up until the present moment. The only way to actually determine which genes were Paleolithic, which were Neolithic, and what percent of each made up the European population through time was to actually test ancient samples. Developing genome sequencing technology eventually made this possible.⁵⁴⁶ Initially, only ancient mitochondrial DNA could be sequenced, as each cell contains hundreds of copies of the mitochondrial genome as opposed to only one copy of the full nuclear genome.

The very first results were fascinating. Haak et al. (2005) sequenced the mitochondrial genomes of 24 early farmers associated with geographically widespread LBK sites. 18 of the 24 individuals belonged to mtDNA haplogroups that are common in modern Europe, the Near East, and Central Asia. The remaining 6 however belonged to haplogroup N1a, today found in only 0.2% of people in the same area as the individuals sampled. Given how geographically widespread the samples were, it was determined that this represented a 150-fold decrease in the frequency of this haplogroup between the Neolithic and the present day and that this could not have been due to genetic drift. Thus, either agriculture spread in way fitting with the cultural diffusion hypothesis, or European farmers were replaced again later. Given the archaeological evidence, the latter scenario seemed unlikely. Who could the farmers have been replaced with? Based on the arguments proposed thus far, Haak et al. postulated that modern Europeans might genetically be mainly of Paleolithic origin, with the incoming farmers having intermarried with the local hunter gatherer populations to the point of being genetically assimilated into them.

Then however, Malmström et al. (2009) found that, at least in Scandinavia, the hypothesis of genetic continuity between hunter gatherers (in this case the Pitted Ware Culture) and modern Swedes, Norwegians, and even Saami could not be upheld.⁵⁴⁷ They were not able to collect enough data on farmers (in this region, the Funnel Beaker Culture), and thus concluded that a population replacement must have occurred during the Neolithic or post-Neolithic periods. The former seemed to be ruled out by the Haak et al. (2005) study, but could this have been due to the regional nature of Malmström's data? Bramanti et al. (2009) were able to confirm the genetic discontinuity between European hunter-gatherers and LBK farmers. The hunter-gatherers belonged mainly to mitochondrial haplotype U, especially type U5. LBK farmers brought haplotype N1a along with some others, but never types U4 or U5. They discovered that hunter-gatherers

⁵⁴⁶ For a thrilling overview of the developments in the field, in narrative form, see Pääbo 2014.

⁵⁴⁷ It could *not* however be rejected between hunter gatherers and modern Baltic populations, suggesting that the eastern Baltic area remained a genetic refugium for some hunter-gatherer populations.

and farmers, although living side-by-side, did not intermarry, at least for a few thousand years. More importantly, modern European mtDNA diversity cannot be explained by admixture between hunter-gatherers and farmers. There were indeed other, post-Neolithic population turnovers.

Haak et al. (2010) sequenced mitochondrial DNA from 43 LBK farmers, finding amongst them 25 haplotypes. Of those, 11 are still found in high frequency among present-day populations, 10 are found with limited distribution (purportedly near the core of the LBK as well as the Near East), and 4 no longer exist. When compared with some other datasets, previous patterns in haplogroup distribution were confirmed. N1a was present in some 14% of the LBK samples, but none of the hunter-gatherers. Conversely, hunter-gatherer mitochondrial DNA consisted of ca. 80% haplogroups U4 and U5, which were almost non-existent in the LBK individuals. Contrary to Torroni et al. (1998), haplogroups H and V were only common in LBK individuals. Haak et al. (2010) were also able to sequence three LBK Y-chromosomal genomes. Rather than the most common haplogroups today (R1a, R1b, I, and E1b1), they found the males belonged to haplogroups G2a and F, rare today but with slightly higher frequencies in the Near East. They concluded that the modern European population was shaped by *both* Mesolithic re-peopling *and* the LBK expansion, with the later having contributed much more. However, it was also shaped by post-Neolithic events, demonstrated by the 4 unique mitochondrial haplogroups and the 3 rare Y-chromosome haplogroups, perhaps having to do with the movement of Y-chromosome haplogroup R1a.

Haak et al. (2008) had identified Y-chromosome haplogroup R1a1 in Late Neolithic Corded Ware individuals along with diverse mitochondrial haplogroups (K1, U5, I, H, and X). Keyser et al. (2009) found R1a1 at very high frequencies in southern Siberia. Out of a group of 26 sequenceable genomes dating from the Middle Bronze Age Andronovo Culture through the Karasuk, Tagar, and Iron Age Tashtyk cultures, all males save for one (of haplogroup C) were of haplogroup R1a1. As with other analyses, the mitochondrial haplogroups were more diverse: the Bronze Age sites were dominated by western Eurasian haplogroups (HV, H, T, I, U, and K) while the eastern Eurasian haplogroups (Z, G2a, C, F1b, and N9a) increased in frequency in the Iron Age. Working with some data and interpretations by Semino et al. (2000), Keyser et al. (2009) posited that the distribution of Y-chromosomal haplogroup R1a1, with its peak in Poland and a decreasing cline to the Northwest and to the East towards Anatolia and the Caucasus, might represent the re-peopling of Europe from an Eastern European glacial refugium, magnified later by the spread of Kurgan people bearing Indo-European languages. Despite being more diverse, the distribution of mitochondrial DNA haplogroups could not rule out the idea that whole populations, rather than just men, had been migrating. Could this have been a part of the post-Neolithic population changes?

Skoglund et al. (2012) used partial whole genome analysis, and though they only included four individuals (three Pitted Ware hunter-gatherers and one Funnel Beaker farmer) the samples were all from the same geographic region and roughly

contemporaneous. They found that, for all European populations, allele-sharing with hunter-gatherers is negatively correlated with allele-sharing with Neolithic farmers. The trend is mainly along a North-South axis with hunter-gatherers sharing most alleles with northern Europeans and the farmers with southeastern Europeans. All in all, the fraction of farmer DNA decreases from a high of ca. 95% in Sardinians through ca. 52% in individuals of northwestern European descent to ca. 11% in Russians. The non-farmer percentage is not fully hunter-gather however. On the contrary, despite allele-sharing, Skoglund et al. (2012) found that the genetic profile of the Scandinavian hunter-gatherers they sequenced is not fully represented in any modern population. By this point, it had become clear that farming had spread, at least in some parts of Europe, via a largescale demic diffusion.

Gamba et al. (2014) provided further evidence for these results, sequencing autosomal genomes from a period of about 5,000 years on the Hungarian plain. This longitudinal analysis showed two clear shifts in genomic affinities on either side of a period of about 2,800 years with little change (during the Neolithic). The first change was due to the incoming farmers, but already the two earliest farmers tested showed intermarriage with hunter-gatherers. In fact, the combination of Mesolithic hunter-gatherer Y-chromosomal haplogroups I2 and C6 with Neolithic farmer mitochondrial haplogroup N1a suggested that hunter-gatherer males were being incorporated into farming societies in Central Europe. This stood in potential contrast to the Mediterranean route, where Lacan et al. (2011) had found Y-chromosomal haplogroups G2a and Elb1b1a1b in Spain. Once widespread, as attested by Haak et al. (2010) finding it in Germany, G2a is now very rare. Lacan et al. (2011) suggested that its high frequency at this southern farming site meant that men had played a greater role in the spread of farming along the Mediterranean route. Sampietro et al. (2007) had also noted a difference between the two farming routes, going as far as to suggest that, since Iberian Neolithic remains were still similar to modern Iberians, farming had perhaps spread via demic diffusion along the Mediterranean route but more via acculturation in Central Europe. In any case, Gamba et al. (2014) found that the Neolithic genomes cluster with affinity to southern Europeans, especially Sardinians. After the period of Neolithic genetic stasis, Bronze Age individuals cluster with modern Central Europeans. This suggested some genetic influence from the North. Finally, the Iron Age individuals clustered between modern Eastern Europeans and individuals from the Caucasus. Now that the archaeologically well-attested migration of Neolithic farmers throughout Europe was being genetically fine-tuned, more attention was being granted to these mysterious post-Neolithic population turn-overs.

Patterson et al. (2012) had found that a portion of European ancestry was shared with Amerindians. The source and time-depth of this added to the mystery. Lazaridis et al. (2014) were able to further contextualize those results. They added seven individuals to the ancient whole genomes that had been sequenced to date and then performed an admixture analysis on modern European and Near Eastern populations to understand

how the modern and ancient genomes were related. In a PCA plot, they found that modern European and Near Eastern populations formed two separate North-South clines, bridged by a few Mediterranean populations. The ancient genomes formed several clusters in relation to the modern genomes. Western European hunter-gatherers formed two clusters (Western and Scandinavian hunter-gatherers), beyond present-day Europeans in the direction of European differentiation from the Near East. Early European farmers, including the Tyrolean Ice Man, clustered together. Finally, at the far northern end of the Europeans on the plot clustered a group of Siberian Upper Paleolithic hunter-gatherers, including Mal'ta Boy. Lazaridis et al. (2014) called this cluster Ancient North Eurasian (ANE). Admixture analysis shows that this ANE component is a better representation of the mysterious component of European ancestry than Amerindian ancestry (of which it is also a part, leading to the similarity noticed). The admixture analysis furthermore revealed that nearly all Europeans were genetically a mix of three ancestral populations: 1) Western European hunter-gatherers (WHG) who contributed ancestry to Europeans but not Near Easterners, 2) early European farmers (EEF) mainly of Near Eastern descent but with small amounts of hunter-gatherer ancestry as well, and 3) this ANE Siberian hunter-gatherer component, which contributed to both Europeans and Near Easterners. Haak et al. (2010), mentioned above, had confirmed the role of both hunter-gatherer and farmer ancestry in making up the modern European genome, along with post-Neolithic factors. Now Lazaridis et al. (2014) showed how important these post-Neolithic factors really were. Early European farmer ancestry ranges from 30% in the Baltic region to 90% in the Mediterranean. But the ANE ancestry, absent from both hunter-gatherers and farmers and therefore post-Neolithic, is found in all Europeans at at least 20%. Nearly a quarter of the European genome comes from a *third* ancestral population. This conclusion, resulting from the ability to sequence ancient whole genomes, set the stage for a major paradigm shift. What role did ANE ancestry play in post-Neolithic Europe?

6.2.3.3 The 2015 Paradigm Shift

2015 was an enormously important year for the field of ancient genomics as well as for Indo-European studies. Two major papers, Haak et al. (2015) and Allentoft et al. (2015) appeared in the same issue of *Nature* and provided convincing answers to many of the questions that had surfaced up to that time.

Haak et al. (2015) sequenced whole genomes for 69 ancient individuals. This large sample allowed many trends to become visible. The earliest European farmers from across Europe cluster with present-day Sardinians, suggesting that both the farming expansion across the Mediterranean and the one into Central Europe had similar ancestral origins. Middle Neolithic Europeans are intermediate between the earlier farmers and Western hunter-gatherers, meaning that hunter-gatherer ancestry began to increase again after the farmers first arrived.

In European Russia, Y-chromosome haplogroups R1a and R1b are found in 100% of

samples from all periods. Before the Late Neolithic period, outside of Russia, these haplogroups were virtually absent. Then, in late Neolithic and Bronze Age sites, these haplogroups are found in 60% of the male populations, suggesting that they spread from the East in ca. 3000 BCE. Furthermore, Late Neolithic/Bronze Age Europeans plot between the Middle Neolithic farmers and the Yamnaya steppe herders from ca. 3000 BCE, indicating that they are the result of admixture between these two populations.⁵⁴⁸ By 2500 BCE, genetic signatures consisting of up to 79% Yamnaya-like ancestry appear in the Corded Ware Culture. As individuals of the Corded Ware Culture are some of the earliest to differentiate from the Middle Neolithic population of Europe and because they also do so most strongly, it appears that Yamnaya people migrated into Europe quite suddenly. This represented, as they titled it, a “massive migration from the steppe”. The once-common mitochondrial haplogroup N1a and Y-chromosomal haplogroup G2a virtually disappear at this time and are replaced by mitochondrial haplogroups I, T1, U2, U4, U5a, W, and subgroups of H as well as Y-chromosomal haplogroups R1a and R1b. In fact, the Yamnaya are likely the source of the ANE ancestry component found throughout all Europeans.⁵⁴⁹

Research had focused on the Neolithic expansion since it was archaeologically so clearly a major force of change across Europe. Now though, a second expansion, the full extent of which had been invisible, was brought into the light of day. Current European populations were all the result of a three-way admixture of Western hunter-gatherers, Early Neolithic farmers, and Yamnaya steppe populations. The proposals of Gimbutas, Mallory, Anthony and the linguists since Schrader as to the origin and spread of the Indo-European language family had to be reckoned with in a huge way. Further support continued to shift the paradigm.

Haak et al. (2015) further found that Bell Beaker and Únětice populations had reduced Yamnaya ancestry compared with earlier Corded Ware Culture individuals, suggesting that Early European farmer ancestry began to rise again after the initial incursion of steppe peoples (in the same way that hunter-gatherer ancestry rose again after the initial appearance of farmers). Allentoft et al. (2015) confirmed and expanded these. Individuals from the Afanasievo Culture are genetically indistinguishable from the Yamnaya and demonstrate that the expansion went East as well as West. The later Sintashta, although in geographic proximity to the Afanasievo, are not genetically descended from them. Instead, their mix of Yamnaya and Neolithic farmer ancestry matches the Corded Ware, and requires that a population from Western Europe moved

⁵⁴⁸ This makes Keyser et al. (2009) seem extremely prescient.

⁵⁴⁹ The Yamnaya ancestry signature is composed of admixture between ANE ancestry (from Eastern hunter-gatherers [EHG]) and what Haak et al. (2015) identified as Near Eastern ancestry (cf. Mathieson et al. [2015] who call it Armenian-like Near Eastern but without additional EEF). The latter has been demonstrated by Jones et al. (2015) to most likely be Caucasus hunter-gatherer (CHG) ancestry. Yamnaya, Afanasievo, and Poltavka individuals are homogenous in comprising 48-58% of this component (Mathieson et al. 2015). Lazaridis et al. (2016) find that this component is related to Chalcolithic farmers from Iran.

back across Europe toward the steppe and provided the ancestry of the Sintashta. The Andronovo Culture then represents a temporal and geographic expansion of the Sintashta gene pool. Bronze Age cultures that came after the Andronovo began to introduce East Asian ancestry. At the same time, Mathieson et al. (2015) found that male individuals of the Srubnaya Culture to the West of Andronovo were all of Y-chromosome haplogroup R1a, but that some were of haplotype R1a-Z93. This haplotype is not found among ancient central Europeans, but is found today in Central/South Asians. This suggested a potentially more eastern source for the Early European Farmer ancestry component on the steppe than the Corded Ware.

Further developments saw the refining of our understanding of past populations movements in Eurasia. Each newly sequenced ancient whole genome provided greater resolution. Fu et al. (2016) took a closer look at Paleolithic and Early Neolithic genomics and concluded that population change was a common occurrence in Eurasian history. Surprisingly, they found an R1b male from Villabruna in Italy 14,000 years ago, which, when taken with an R1b farmer from 7,000 years ago in Iberia (cf. Haak 2015), shows that the distribution of Y-haplogroup R1b, as is the case with R1a above, has yet to be fully understood. Their analysis supported the conclusions of Haak et al. (2015) and Allentoft et al. (2015) that ANE ancestry spread with steppe migrations.

Also around this time, the genetic results being published in journals like *Science* and *Nature* were causing concern in archaeological circles, where it was feared that the interpretations were too simplistic, too generalized, and ignored the understanding of human interactions that had been gathered from decades of archaeological inquiry. Vander Linden (2016) writes that “the narrative set forth in several high-profile publications seems as first sight to rise from the darkest depths of culture history.” Other issues include the patchy sampling, particularly in the Mediterranean region and the supra-regional conclusions being made on this potentially unrepresentative data. Individuals are listed as belonging to the Corded Ware Culture or the Bell Beaker Culture, both of which are characterized by their variability over wide-spread geographic areas. He does not reject that the ancient DNA analyses offer a wealth of new information but instead argues for more nuance, more collaboration, and the recognition of the important role of regional differences.

As if in response to these critiques, some subsequent studies addressed regional phenomena and targeted regions with a lack of data. Lipson et al. (2017) sampled Neolithic and Chalcolithic individuals from Hungary, Germany, and Spain in order to better understand the interactions between hunter-gatherers and incoming farmers. They found that each region showed a distinct pattern of interaction between EEFs and hunter-gatherers. WHG ancestry increased over time in all regions, but more slowly in Hungary. The genetic distribution that resulted showed that the admixture occurred locally, rather than suggesting that farmers were quick to admix and then carry that admixture across Europe. After the migrating, settling, and admixing, populations did not

seem to be mobile enough to recreate the original homogeneity of the farmers.⁵⁵⁰

Mathieson et al. (2018) helped fill in the picture for Southeastern Europe. The Balkan peninsula was an important genetic transition zone throughout prehistory. Hunter-gatherers from the Iron Gates region show ancestry from both WHG and EHG populations, but admixture analyses show that it is not a perfect fit, suggesting other influences. Amongst them are also some early individuals of entirely Northwestern Anatolian ancestry, which means that region was a zone of interaction. Individuals from the Peloponnese and some associated with the Minoans show more CHG-oriented than WHG-oriented ancestry. Once farming spread to the area, Neolithic populations in modern Bulgaria, Croatia, North Macedonia, Serbia, and Romania are of up to 98% Near Eastern ancestry. The Copper Age sees a rise in hunter-gatherer ancestry, as it does elsewhere in Europe. Mathieson et al. (2018) analyzed this resurgence and concluded that a bias toward male hunter-gatherer ancestry is strong in Iberia and Central Europe at this time, but weak in the Balkans. Some Balkan individuals, including one from a burial at the site of Varna, show steppe ancestry components two millennia before the advent of Corded Ware, but steppe ancestry otherwise remains rare until the Late Bronze Age. Although they suggested it might turn up later, Bronze Age Anatolians have the CHG component but not the EHG component of the steppe signature. This has implications for the time and manner of the spread of the Anatolian Indo-European languages.

Olalde et al. (2018) made an important contribution to the understanding of the Iberian peninsula, where it had been noted that populations did not show steppe ancestry at the time when the rest of Europe had begun to (Mathieson et al. 2015). Outside of Iberia, 84 of 90 Bell Beaker males analyzed belonged to Y-chromosome haplotype R1b-M269. Within Iberia, this haplotype was found in four males with genome-wide steppe ancestry. But otherwise there was a higher proportion of haplogroups like I and G2, known to have been common across Europe during the earlier Neolithic. Steppe ancestry was only present in 8 of the 32 Iberian Bell Beaker individuals sequenced, though it is with the Bell Beaker that steppe ancestry components first appear in Iberia. Across Europe, Bell Beaker-associated individuals are heterogenous, even sometimes within sites, across a cline with one extreme being Yamnaya individuals and the other extreme being Middle Neolithic and Copper Age Europeans. But no Iberia-related ancestry could be discerned in Beaker-complex-associated individuals outside of Iberia. Thus this early spread of the Bell Beaker complex seems to have involved cultural diffusion rather than migrations. This was not the case in the British Isles, where steppe ancestry arrives with the Bell Beaker complex from the mainland, with genetic affinities to Central European individuals. By ca. 2450 BCE, an approximately 90% population turnover had occurred in Britain in autosomal DNA, previously absent Y-chromosomal haplogroup R1b, and previously absent mitochondrial haplogroups I, R1a, and U4. By 2000 BCE, as happened

⁵⁵⁰ Both the Danubian (LBK) and Mediterranean (Cardial Ware) agricultural waves seem to have originated from one single population of farmers in the Balkans (Haak et al. 2015, Mathieson et al. 2015, Mathieson et al. 2018).

elsewhere, Neolithic-related ancestry modestly increased again. Olalde et al. (2019) further demonstrated the uniqueness of the hunter-gatherer population of Iberia, with a substructuring between northwestern and southeastern hunter-gatherers having existed. It was ca. 2000 BCE that steppe ancestry became entrenched in Iberia through male-biased admixture. Steppe ancestry replaced 40% of the previous ancestry components but nearly 100% of the Y-chromosomes (with R1b-M269). By the Iron Age, steppe ancestry had even spread to the non-Indo-European-speaking Basque.

6.2.4 Summary

The results of ancient DNA analysis provide strong support for the steppe hypothesis of Indo-European origins. Many of the details will continue to be fine-tuned in the future, for example the precise relationship of the R1a and R1b Y-chromosome haplogroups. But it has justified a number of assumptions that relate to the substrate lexicon of Latin:

The Italic language family spread through Europe from the steppe. Its ancestors arrived in a Europe that had been settled for centuries by farmers from Southwest Asia. But they had also admixed to a significant extent with the original inhabitants of Europe: the hunter-gatherers. Thus there is the potential for great substrate language diversity but also the possibility that some aspects could be widely distributed.

6.3 The Italian Peninsula

Genetic results pertaining to the Italian peninsula have been sequestered to a second section of this chapter for two reasons. The first, most obvious reason, is that Italy is the subject of this thesis and it is therefore worth discussing them separately in detail. The second reason is that, in comparison to the rest of the Europe and as lamented by Vander Linden (2016), not nearly as much work has been done on ancient Mediterranean genomes until extremely recently. Population genetic studies with the aim of adding to medical understanding abound in the earliest literature, and a research tradition in line with that of Cavalli-Sforza evolved from that. It serves best to discuss the Italian data after the general stage has been set for Europe, so the particularities of what the data suggest can be most visible. Ancient DNA studies on remains found in Italy are very recent, having been performed even later than for other regions. But more are forthcoming.

6.3.1 Studies on Modern Populations

6.3.1.1 The Earliest Studies

As with population genetics in general, the earliest studies of this sort on the Italian peninsula had begun already in the 1970s, investigating “classical” genetic markers. The purpose was often principally medical and involved understanding the distributions of blood group and protein markers (cf. Bargagna, Domenici, and Morali 1975). Soon however, recurring patterns began to appear. Menozzi, Piazza, and Cavalli-Sforza (1978)

had already noticed that the distribution of traits in Sardinian populations was closer to that of Middle Eastern populations than anywhere else in the Mediterranean. They attributed this to Phoenician and Punic immigration, assuming that the first farmers would have come from southern Italy. Southern Italy grouped together with southern France, Greece, and Turkey, which they took to be the influence of Greek colonizations.

Zappacosta et al. (1980), when they compared HLA frequencies in Bergamo and Campania, found that frequencies of certain antigens in Campania were similar to the Middle East whereas in Bergamo they parallel the European distribution. Giari, Domenici, and Bargagna (1986) found that the frequency of red cell esterase D (EsD*5) in Tuscany is similar to that expected in Northern and Central Europe, while the frequency of EsD*1 is lower in Italy as a whole than in Europe. Olivetti et al. (1986) found from HLA-A, -B, and -C allelic frequencies throughout Italy that “genetic distances of Sardinia from the rest of Italy are one order of magnitude higher than all the other distances between Italian regions” and that “the Southern regions (with the possible exception of the island of Sicily) are genetically more distant from the Northern and Central regions than from each other.” Some alleles showed a North-South gradient.

Piazza et al. (1988) performed an analysis in which they also factored in linguistic and archaeological evidence. They produced a PCA from several blood group gene frequencies, and saw in the first principal component that Northern Italy clusters with Central and northern European countries, while Southern Italy clusters with Greece. The second principal component separated a Northern European cluster from a Mediterranean one (Italy, Greece, France, Spain, Portugal) plus Hungary and Czechoslovakia. As the other studies had done, this multiple allele analysis showed a North-South gradient. Piazza et al. interpreted the genetic similarity of southern Italy to Greece in the same way as Menozzi et al. (1978): possible genetic input from Magna Graecia. Proportional population estimates of the Greeks in Italy led them to believe they should almost certainly see a Greek genetic signature. The rest of their conclusions are based on contemporary understandings of the archaeological and linguistic landscape of Italy (cf. §7.2). They suggested that an area of genetic distinctness in the West could be the influence of Etruscans. Furthermore, they found the population of Liguria to be distinct, and found potential affinities between Oscans and Picenes. They concluded from all of this that the Italic languages were most likely brought by small groups from the area of the Danube, either over the Alps or from the Adriatic, in two waves of first cremating and then inhuming Italici. Citing the lack of any Proto-Messapic archaeological traces along the peninsula, they proposed that Messapic could have been brought to Apulia from Illyria across the Adriatic. This analysis was the first to incorporate genetic lines of evidence into the arguments that had been built using archaeology and linguistics.

6.3.1.2 More Modern Methods

Fascinatingly, the major trends that had been observed from classical/protein markers

were *not* overturned, but rather confirmed by DNA testing. Barbujani et al. (1995) found that around one quarter of the mitochondrial haplogroups they identified (12 of 42) across Italy and Sardinia showed substantial variation in frequency among populations. Some of these took the form of North-South gradients, with Sardinia being quite distinct. They took these patterns to be the result of a rapid demographic expansion, but they calculated a date of 8,200 - 20,525 years ago for the expansion. They could not decide if this was from the Neolithic spread of farming or from population growth after the Last Glacial Maximum. In any case however, Sardinia produced other results, suggesting slow if any growth and then isolation. Semino et al. (2000) found that Sardinian Y-chromosomes cluster with Greece and Albania, but not Basques. Chikhi et al. (2002) concluded that Sardinia is an outlier with a significantly higher proportion of Paleolithic genes, and that its pattern represented genetic drift from the Paleolithic rather than a Neolithic immigration. These conclusions fit into the debate at the time that was centered around the true contribution of the Neolithic migrants.

Di Giacomo et al. (2003) examined the distributions of 9 Y-chromosome haplogroups from 30 populations in Italy and Greece, finding major differences between them. In Italy, only 3 haplogroups have frequencies above 10% while in Greece, the three most common haplogroups account for only 55% of chromosomes. The only haplogroup to show a significant distribution was P*(xR1a)⁵⁵¹ in Italy, decreasing from Northwest to Southeast. Because this haplogroup appeared in Basque and Celtic populations, they thought it was a Paleolithic group. But recall that Olalde et al. (2018) would show that this spread to Basque groups relatively late, by the time of the Iron Age. Capelli et al. (2006) sampled more markers with a larger sample size and found that in fact 70% of the Y-chromosome diversity of Italy is structured along a North-South axis. 80% of chromosomes were in haplogroups R1*(xR1a1)(itself 40%), J2 (20%), G (11%), and E3b1 (10%). R1 frequencies decrease from North to South, while J2 and E3b1 frequencies increase in from North to South. Like Di Giacomo et al. (2003), Capelli et al. (2006) assumed Iberian populations were a good proxy for Paleolithic ancestry. Thus when they tested admixture using Iberians to represent Paleolithic Europeans and Anatolians to represent Neolithic farmers, they interpreted the 70-90% Anatolian contribution in the South and 50-70% Anatolian contribution in the North as well as the decreasing frequencies of R1 from North to South to represent Mesolithic-Neolithic contact effects. What they were truly seeing however was steppe influence descending in decreasing frequency from the North.

Di Gaetano et al. (2012) were able to replicate the patterns of 1) Northern and Central Italy clustering separately from Southern Italy, the former with the rest of Europe and the latter with the Middle East and 2) Sardinia clustering separately from all groups using genome-wide data. Keller et al. (2012) found that the Neolithic Ice Man of the Ötztal Alps clusters with modern Sardinian populations, suggesting that Sardinia's signature

⁵⁵¹ Haplogroup R (to which R1a and R1b belong) is a subclade of P. The x indicates that, though the exact subclade was not determined, it is not R1a.

was not Paleolithic, but rather Neolithic. Boattini et al. (2013) looked again at uniparental markers, finding that 62% of modern Italian Y-chromosomes fit into 8 haplogroups, including 4 R1b lineages. Three R1b haplogroups and G-P15 (an EEF haplogroup) were mainly responsible for the Northwest-Southeast cline, with the Southeast showing more intra-population variation and the Northwest showing more inter-population variation. Sardinia differed in having the lowest Y-chromosome haplogroup diversity, and had I-M26, which was virtually absent from the peninsula. In contrast to the Y-chromosome picture and suggesting sex-biased population movements, mitochondrial DNA haplogroups are much more homogeneously distributed. De Fanti et al. (2015a) determined that the Italian mitochondrial HV* lineages split very early, and some are unique to Italy, suggesting an ancient local presence perhaps from the time of a glacial refugium in Southern Italy. Sarno et al. (2014) agreed with these results, emphasizing that NW Italy clusters Iberia and Central Europe while SE Italy (including Sicily) clusters with the Balkans and the Levant, and that males seem to have been the mobile ones. The NW to SE gradient was again replicated by De Fanti (2015b), this time with the genes for lactase persistence (North Italy was intermediate between high frequencies of lactose intolerance in Southern Italy to the highest level of lactase persistence in Finland, showing that European populations appear more homogenous than those of the Italian peninsula), and by Fiorito et al. (2016) again with whole genome analysis. Grugni et al. (2018) refined the Y-chromosome picture, finding that in Northern Italy, 69% of haplogroups were R1b (specifically R1b-U152) with no other haplogroup reaching 10%. In the South, only 27.5% were R1b, representing a demographic expansion from North-West and Central-North Europe that greatly affected Northern Italy.⁵⁵² This was the traces of steppe ancestry having entered Italy.

6.3.2 Ancient DNA

Ancient DNA analyses with relevance to Italy took considerable time to appear, and there is still a great need for more results in order to understand how and when the patterns found in the modern population data came to be. Mathieson et al. (2018) confirmed the presence of Western hunter-gatherers in Sicily between 12000 and 6100 BCE. Emery (2017) and Emery et al. (2018) confirmed that mitochondrial genomes from Iron Age and Roman period South Italy are comparable to elsewhere in Europe.⁵⁵³

⁵⁵² They also noted, following a paper by Günther et al. 2015, that Basques may not be the “living fossils” from the Paleolithic they have long been considered to be. Instead they might represent a long-lasting isolated population that originated from admixture of local HGs and early farmers. Olalde et al. (2018) would come out in that same year, showing the importance of ancient DNA in understanding population history. The presence of R1b in Basque populations does not date to the Neolithic.

⁵⁵³ In fact, in Emery (2017) when pre-Late Glacial Maximum and pre-Neolithic mitochondrial genomes were removed from comparison, the Iapygian samples shared similar haplogroup compositions with Armenian Iron Age populations. Emery concluded that, if the results are not due to Late Glacial Maximum gene flow, the Iron Age southern Italians descended from Neolithic farmers (from Anatolia and possibly as far East as the Caucasus) and from migrants arriving from eastern Europe in the Late Neolithic or Early Bronze Age (consistent with the ancestors of the Iapygians arriving in Southern Italy from the Balkans, or at least sharing a source population with them). Emery et al. (2018) show that

Serventi et al. (2018) got the same result for Novilara, the area where North Picene is found, showing that Iron Age Picene populations show similar mitochondrial profiles to modern populations from the same region. Antonio et al. (2019) contextualized the results by performing an analysis on whole genomes from Rome, spanning from the Mesolithic through the Iron Age. Mesolithic individuals from Sicily and the Italian mainland indeed clustered with Western hunter-gatherers. Between 7000 and 6000 BCE, agriculture arrived with the appearance of individuals clustering with Anatolian farmers. Interestingly, they carried a small amount of a component lacking in central European and Iberian individuals and found in high quantities in Iranian farmers and Caucasus hunter-gatherers. This suggests a different or at least additional source population from that involved central and western Europe was involved in Italy's Neolithic transition. Then, just like elsewhere in Europe, hunter-gatherer ancestry began to rise again starting in the Late Neolithic.

Allentoft et al. (2015) had determined the Copper Age Remedello Culture in Northern Italy (ca. 3400-2400 BCE, cf. De Marinis & Pedrotti 1997: 298) did not show steppe ancestry, suggesting a potential *terminus post quem* for the arrival of steppe-derived populations in Italy (though they also note that it may have in some way been unaffected by the Yamnaya expansion). Antonio et al. (2019) were able to push this date to ca. 2900 BCE. A gap in their samples due to cremation meant that they had no data between 2900 BCE and 900 BCE, at which point steppe ancestry is already present (with individuals exhibiting highly variable ancestries, suggesting several sources of migration). Olalde et al. (2018) had sampled two individuals in a Bell Beaker context in Northern Italy (Parma) that showed components of steppe ancestry, dating to between 2200 and 1930 BCE. Saupe et al. (2021) were able to fill in the gaps, with samples dating between 3200 and 1500 BCE. They found the first traces of steppe ancestry in individuals from North Italy around 2000 BCE (a Bell Beaker individual from 2195-1940 calBCE as opposed to two without steppe ancestry, an individual from Broion from 1952-1752 calBCE, and even one of the Remedello individuals described by Allentoft et al. [2015], from 2134-1773 calBCE), with the amount increasing with time. The first traces of steppe ancestry in Central Italy were found four centuries later.⁵⁵⁴

Sarno et al. (2017) tested modern Italian populations, but they viewed the results as a PCA overlain on other modern and ancient sequences, allowing them to provide one of the most-up-to-date interpretations of the trends that had been repeatedly found. They found that modern southern Italian and southern Balkan populations plotted in “an almost uninterrupted bridge between two parallel clines of distribution where most of the other modern populations are found, one stretching along the East-West axis of Europe

pre-Roman Iron Age individuals from Vagnari cluster slightly differently than Roman-period individuals from Vagnari, possibly representing the result of Roman subjugation.

⁵⁵⁴ Steppe ancestry appears in Sicily ca. 2200 BCE, but seems to arrive from Iberia (Fernandes et al. 2020).

and the other from the Near East to the Caucasus.” Thus, Sicily, southern Italy, and Mediterranean regions stretching East to the Anatolian Greek islands formed a Mediterranean genetic continuum, with a Sardinian (Neolithic-like) ancestry component making up half of their ancestry. Near Eastern-like ancestry is more frequent in Sicily, southern Italy, and the Greek islands, whereas a European-like component appears in increasing frequencies in the populations of Albania, mainland Greece, and the rest of the Balkan peninsula. In contrast to continental Europe, Caucasus-related admixture via Yamnaya is present to a lower degree in the Mediterranean groups, detected mainly in Balkan-related groups. They concluded that “any significant Steppe/northern component may have arrived in the south Balkan mainland and southern Italy only later, by which time Indo-European languages of the Italic, Greek and various Balkan branches had already established themselves there.” This suggests that the spread of Italic languages into Southern Italy would have involved a certain amount of indirect diffusion and a large amount of contact with pre-IE languages.⁵⁵⁵

6.3.3 Other Questions

6.3.3.1 Greek Colonization

Sarno et al. (2015) had suggested that remnants of the Greek settlement of Southern Italy might be represented by Y-chromosome haplogroup E-V13. Tofanelli et al. (2016) noted that no Italian populations seem to show a closer affinity with Greek and Greek-related sources from a Y-chromosome perspective, but designed an approach that took mutational processes into account. In this way, they purported to find evidence of Greek colonization from Euboea in East Sicily in Y-chromosome and mitochondrial DNA from modern populations. They found a sex bias, and evidence for low numbers: probably thousands of men and a few hundred women. Interestingly, they found that E-V13 was actually a very poor indicator of Greek origin. Later, Sarno et al. (2017) found that the modern Greek-speaking communities of Apulia and Calabria show no clear signs of a recent continental Greek origin, instead clustering within the Mediterranean continuum they found. This suggests that either a large degree of admixture has occurred, that these Greek-speaking communities date back to antiquity (as opposed to the Middle Ages), or perhaps both. Grugni et al. (2018) also found potential genetic traces of the Greek presence in Italy in form of Y-chromosome haplogroup R1b-M412*, which is found in Turkey, Iran, Cyprus, and Greece along with all of their Southern Italian samples but which is only sporadically found in North Italian samples. They suggested this might date to the colonization of Magna Graecia (especially from Greek islands) or perhaps to the Byzantine era. They also found R1a-M17, with the highest frequencies in Italy being in the South and otherwise found in significant amounts in mainland Greece and Thrace. This might represent traces of Balkan migrations into Southeast Italy.

⁵⁵⁵ Fernandes et al. (2020) found in Sicily Iranian-related ancestry consistent with a shift towards Minoans and Mycenaeans by the Middle Bronze Age. Raveane et al. (2019) had found this signature in modern southern Italian populations as well.

6.3.3.2 Etruscans

The question of Etruscan origins, whether they are from Anatolia or autochthonous to Italy, has also received attention from geneticists. At a remarkably early date, Vernesi et al. (2004) sequenced the mitochondrial genomes of 30 Etruscans, finding little heterogeneity between sites or time periods sampled, confirming that the Etruscans were a unitary population. But they were unsure how to further interpret the data, because most of the haplotypes they found no longer existed in modern populations. Some evidence pointed to their genomes being typical for Europe or West Asia, but they also found similarities with Turkish gene pools. Achilli et al. (2007) found that the mitochondrial haplogroups of the modern population of Murlo, a town of Etruscan origin, are composed of 17.5% of Near Eastern haplogroups. No other Tuscan populations had this high of an amount, but overall 5% of mitochondrial haplotypes in Tuscany were found to be shared exclusively with Near Easterners, supporting a direct and recent input from the Near East.

Ghirotto et al. (2013) found from mitochondrial DNA that the Etruscan individuals they sequenced were the ancestors of modern populations in Casentino and Volterra (but, interestingly, not of the more general population of the formerly Etruscan territory) and that the genetic links between Tuscany and Anatolia dated back to at least 5,000 years ago. Thus the Etruscan civilization likely developed locally. Tassi et al. (2013) repeated the conclusion that the genetic links between Tuscany and Anatolia were too old to have brought the Etruscans, instead possibly dating to the spread of Neolithic farmers. But the analysis of results would continue to go back and forth.

Gómez-Carballa et al. (2015) concluded that the Near Eastern components in Tuscan mitochondrial genomes was 8% and in autosomal DNA was 21%. As the Near Eastern mitochondrial haplogroups do not show local or regional variation, it suggests a recent arrival from the Near East with no subsequent founder effects or genetic bottlenecks. Thus the Etruscans were probably from Anatolia. Fiorito et al. (2016) used an admixture analysis on whole genome data that suggested an event ca. 3000 years ago between populations from the Caucasus and Middle East and populations from Central Italy. Pardo-Seco et al. (2014) had calculated a similarly recent date for an admixture event. This might be the arrival of the Etruscans from Anatolia. Grugni et al. (2018) found that the Y-chromosome haplogroup J2a-M67* had high microsatellite variation specifically in Volterra. Given that its origin might be close to the Middle East, they took this to support the Asia Minor origin of the Etruscans.

More recent ancient DNA studies call into question the possibility of this 3000-year-old admixture event. Antonio et al. (2019) in their small sample size found no significant difference between Latins and Etruscans in their allele-sharing with other populations. Most recently, Posth et al. (2021) sequenced individuals dating back to 800 BCE. They confirmed that the Etruscan genetic profile was the same as that of its Latin neighbors (including a large proportion steppe ancestry), without any recent population influx from

the Near East. The Etruscan language seems to have been preserved despite large amounts of admixture, similar to the situation of Basque. It is unclear if an elite dominance situation like that responsible for the spread of Hungarian could instead be responsible.

6.4 Conclusions

There are several important conclusions that we can draw from the picture of Italian population genetics. Firstly, the genetic history of the Italian peninsula is one of a melting pot. Steppe-derived populations begin to appear in Northern Italy around 2000 BCE, their presence increasing gradually. Steppe ancestry reaches Central Italy by around 1600 BCE. Modern genetic profiles show that the spread of steppe-derived populations into the South of Italy involved large amounts of admixture with previous populations. Complete admixture seems to have occurred with the ancestors of the Etruscans as well. All signs point to heavy contact with speakers of pre-Indo-European languages. The possibility of non-IE language communities, especially in the South, existing up until a late date is certainly not ruled out.

Genetic studies on the population history of the Mediterranean continue to come out, and a refined understanding of population movements will certainly provide clues about the linguistic diversity of Italy. How different would the language of the Cardial Ware farmers of the Mediterranean coast be from that of those that established Linearbandkeramik in the North? And does the Mediterranean continuum help explain the existence of a Mediterranean substrate? What about the traces of Iranian-related ancestry (consistent with Minoans and Mycenaean) in modern southern Italians? We may someday be able to postulate what other languages, perhaps attested ones, might have played a role in contact situations.

7 Archaeological Theories on the Italicization of Italy

The history of the field of prehistoric archaeology in Italy is a nuanced one, presented by Guidi (2010) as a discipline that competed with the much more well-funded historical (Classical) archaeological field and which was shaped by methodological rivalries and political ideologies. Discoveries beginning in the mid-1800s led to the recognition of many of the major prehistoric cultures of the Italian peninsula and its environs (though their contemporary recognition as prehistoric was not always immediate). One of the first to begin to using the archaeological evidence to theorize on the arrival of Italic-speaking populations in Italy was Luigi Pigorini.

Having had his start in the field of prehistoric archaeology under the mentorship of Pellegrino Strobel (the first to recognize Terramare sites prehistoric and with whom he published a seminal work on the Terramare: Strobel and Pigorini 1864), Pigorini became a juggernaut that dominated the field for decades. Throughout numerous articles in the *Bullettino di Paleontologia Italiana* (starting with e.g. Pigorini 1875), he developed what would come to be called by some his *teoria pigoriniana*. The narrative that he supported was that Neolithic Italy had been inhabited by a homogenous population of autochthonous Mediterranean origin that practiced inhumation burial. The Copper Age saw the arrival of peoples who built pile-dwellings over lakes (the *palafitticoli* of the Polada Culture) like those across the Alps in Austria and Switzerland. Then in the Bronze Age, the eastern part of these regions saw the arrival of a people who built their pile-dwellings over dry land and practiced cremation and intensive metallurgy. This was the Terramare Culture, responsible for disseminating bronze metallurgy throughout the Italian peninsula. They left their settlements in the Po Valley, either willingly or due to another wave of invaders, and marched/treked southwards and over the Apennines whereupon they became the Iron Age Villanova and Latial cultures and eventually founded the city of Rome using the city plan of a Terramare settlement (summarized in e.g. Randall-MacIver 1939, Guidi 2010).

Though highly influential, the sacrosanctity of Pigorini's *teoria* began to come into question in the 1900s. The Terramare origin of Rome and the Roman military *castra* was revealed to have been based on precious little evidence (Säflund 1939, further e.g. Barocelli 1942). Several alternative accounts reduced the role of foreign invasions, with varying degrees of ideological bias. Patroni (e.g. 1939: 215), on the basis of racial phenotypical analysis, argued that Indo-European languages could only have come to Italy "by exchanges and imitations and by slow but very extended infiltrations of individuals."⁵⁵⁶ Ugo Rellini's work on the Apennine Culture (originally called the *extra-terramaricoli*), highlighted the continuity in many areas with the earlier Neolithic and Copper Age cultural materials and disproved the existence of Terramare sites outside of the Po Valley (cf. Rellini 1929, 1933). Thus several scholars gave up the idea that

⁵⁵⁶ "Per scambi e imitazione e per lente ma lunghissime infiltrazioni di individui".

Terramare was “portato bell’e fatto” by a “popolo speciale” (Barocelli 1942: 138) and some suspected instead that it was but a local development instigated by geographical and environmental conditions (cf. Barocelli 1942: 136) of “una diretta emanazione” of the Apennine Culture (Laviosa-Zambotti 1937: 54): rather than Terramare being a great influencer, it would instead have been a local development in the North of a culture further South.

After the era of Fascism, archaeological research on the Italian peninsula continued to develop, gradually at first (especially with the work of Massimo Pallottino beginning in the 1940s, then Salvatore Puglisi and Renato Peroni beginning in the 1950s), with methodological and ideological innovations throughout the 1970s and into the 1980s when Processualist methodology began to be taken up (Guidi 2010: 17-18). Several additional hypotheses as to the dating of the arrival and archaeological cultural affiliation of the first Italic speakers in Italy have been proposed, summarized in archaeological chronological order.

7.1 Single Origin Theories

7.1.1 The Copper Age Cultures

The transition from the Neolithic to the Copper Age (Chalcolithic/Eneolithic) in Italy (ca. 3600 to 2200 calBCE) is marked by the intensification of the use of copper and the appearance of three archaeological cultures: Remedello (North), Rinaldone (central), and Gaudio (Southwest)(Baldi 2002: 98, Dolfini 2014: 477).⁵⁵⁷ Metalwork was theorized to have been brought to these areas by a nomadic warrior elite (e.g. Laviosa Zambotti 1939: 58, Puglisi 1959: 89-90, Trump 1966: 69), possibly Indo-European speaking due to the appearance of e.g. horses and battle axes at this time (e.g. Laviosa Zambotti 1949, discussion in Mallory 1989: 93). However, even if the technology of copper metallurgy may have been introduced from abroad (cf. Dolfini 2014), others are not convinced that there is strong enough evidence for its introduction being accompanied by a migration (Barfield 1971: 59, Barker 1981: 81-9, Mallory & Adams 1997: 318). More definitively as regards a possible Indo-European connection, genetic sequencing of individuals from Remedello (Allentoft et al. 2015, Mathieson et al. 2015) and Rinaldone/Gaudio (Antonio et al. 2019) contexts do not show steppe ancestry, indicating that they predate the arrival of Indo-European speakers.

7.1.2 Side Note: Bell Beaker and Polada Cultures

Bell Beaker cultural material (potsherds) appears in over 80 North Italian Copper Age sites (Trump 1966: 70, Nicolis 2001: 208, Dal Santo et al. 2014), and is found in more limited amounts at least as far South as Campania (Aurino & De Falco 2022).⁵⁵⁸ Three

⁵⁵⁷ Dolfini also includes a fourth culture, Laterza (Southeast).

⁵⁵⁸ It is also found in Sicily, but here its development is separate from the Italian mainland (Aurino & De Falco 2022: 211).

Bell Beaker burials have been found near to the cemetery at Remedello, culturally distinct from the cemetery burials (Barfield 1971: 62), along with perhaps a few other sites (Nicolis 2001: 209). As noted above (§6.3.2), it is in one of these three burials that some of the earliest steppe ancestry in Italy is found (Saupe et al. 2021). This nicely parallels the conclusion of e.g. Barfield (2001: 516), on the basis of a combination of new (Bell Beaker) and persistent (Copper Age) lithic forms, that the appearance of Bell Beaker material in Italy involved some population movement but also intensive interaction with already present populations. Thus the Bell Beaker culture seems to be the earliest possible candidate for a bearer of Proto-Italic into Italy. But compared to later possibilities, there seems to be simply too little Bell Beaker material to represent the arrival of a language that would dominate the peninsula. Thus, even if the bringers of Bell Beaker materials were Indo-European speaking (cf. Gallay 2001: 54-6, Waldman & Mason 2006: 453-4), and even if their languages persisted, it seems quite likely that “more evolved forms of Indo-European, including Italic, may have spread across Italy at a later stage” (Posth et al. 2021, cf. Stifter fthc. on a similar situation Celtic regions).

Nor has the Polada Culture (ca. 2200-1500 BCE) been proposed as a vector of the Italic languages, at least not in isolation. It does feature in some of the multiple-origin theories below (§7.2). The Polada sites are characterized by pile-dwellings (*palafitte*) at the edges of lakes and watercourses. The culture is often connected to pile-dwelling cultures in Alpine Switzerland and Austria/Southern Germany (though confirmation of the connection is still elusive, cf. Marzatico 2004), from which it would represent a “numerically weak infiltration” (wording from Pulgram 1958: 108-9, cf. also Watmough 1937: 210-11, Bietti Sestieri 2010: 21). On the basis of this trickle, like the case for Bell Beaker, some exclude that the Polada Culture individuals could have spoken Indo-European languages (Palmer 1954: 34, Kaschnitz-Weinberg 1954: 343-6, Devoto 1962: 382-5). A clear continuation with the previous Neolithic Lagozza culture of the area has been claimed (e.g. Barfield 1971: 68, 70), though it is no longer supported by e.g. Fasani (2002: 108), who emphasizes that the style of pile-dwelling and the material culture is not comparable to Neolithic geographic predecessors. Instead, the sudden appearance of the Polada over a relatively large area contemporaneously, interrupting the development of Bell Beaker sites where it appears (though Bell Beaker elements remain visible in the grave goods related to archery [Barfield 1971: 77] and ceramics [Gallay 2001: 47, Nicolis 2001: 212, 218]), suggests the arrival of a new people; whence they originated remains unknown (Fasani 2002: 108, Dal Santo et al. 2014: 225). Their technologically advanced bronze work attests to contact with cultures north of the Alps (Barfield 1971: 77 mentions specifically Únětice) and the *Brotlaibidole/tavolette enigmatiche* attest to contact with the Carpathian basin (Barfield 1971: 74-5, 77; Fasani 2002: 109, Cardarelli 2009: 458, Cavazzuti et al. 2022: 46).

7.1.3 Terramare Culture

The Terramare Culture existed from ca. 1650-1150 BCE (Middle to Late Bronze Age) in

the Po Valley, principally in the modern provinces of Cremona, Mantua, and Verona as well as south of the Po River in the historical region of Emilia. Terramare settlements were quadrangular, surrounded by an embankment and ditch into which a natural watercourse was diverted. The houses were raised on piles, even when the settlement was on fully dry land, in a grid-like network of perpendicular streets. In the beginning of the Terramare period, the settlements rarely exceeded 2 hectares, but by the end of their existence a few reached sizes of up to 20 hectares. The name of the culture comes from the *terra marna*, thick deposits of fertile soil produced by piled up refuse (Cardarelli 2009: 449-51).

The idea that the Terramare Culture represents the first Italic speakers goes back at least to Gaetano Chierici⁵⁵⁹ (Chierici 1871, 1881: 69), who conceptualized similarities between the wall and ditch of the Terramare sites and the ditch and *agger* in Roman city founding as well as between the grid structure of Terramare roads and the *cardō* and *decumānus* of Roman military camps. As mentioned above, Pigorini was inspired to develop the idea further (e.g. Pigorini 1903). Another early proponent of the idea was Wolfgang Helbig (e.g. Helbig 1879). Many would come to disagree with the idea, and a crucial part of the debate centered on whether a foreign origin could actually be established for the Terramare. Some claimed that similar settlement types and ceramics found in the Danube valley, especially Hungary, meant that this was where the Terramare Culture began (e.g. Peet 1909: 505-7, Wilke 1919: 177, Kaschnitz-Weinberg 1954: 344, Trump 1966: 128). Others disagreed that the Terramare materials were related or even similar to those discovered at the Hungarian site of Tószeg (Leopold 1929: 26-7, Rellini 1933: 93-4). V. Gordon Childe agreed that the identification of Tószeg as a *terramare* was baseless, but nevertheless argued that the Terramare Culture represented an invasion from across the Alps, descended from “an as yet undiscovered prototype developing perhaps in Carinthia or Western Hungary” (Childe 1925: 269), with it being “not therefore impossible that Tószeg, or some more westerly stations of the same type, may contain the germs from which the *Terramar[e]* culture of Italy sprang” (Childe 1929: 265). Modern research indeed points in this direction. Around 1500 BCE, the population of the Terramare region suddenly drastically increased (as did human activity like deforestation), too quickly to have been due to demographic growth alone and instead suggesting something akin to a colonization (Cardarelli 2009: 450, 458-9). This corresponds to the same time at which the Tell cultures of the Hungarian plain were collapsing. Noting a similarity in the organization of the Tell settlements and the Terramare, Kristiansen (2018: 118) suggests that inhabitants of the Tells migrated to the Po Valley and became incorporated into the Terramare Culture (cf. earlier Barfield 1971: 95).

Around 1150 BCE, the region of the Terramare settlements was abandoned, perhaps due to a combination of overpopulation, climatic stress, and the political turmoil that resulted

⁵⁵⁹ The archaeologist and priest who lived from 1819-1886, not the painter who lived from 1838-1920.

in the collapse of the Bronze Age elsewhere around the Mediterranean (Cardarelli 2009: 465-72).

7.1.4 Apennine Culture

The Apennine Culture is *sensu stricto* the technological complex, especially the ceramic assemblage, of Central and Southern Italy in the 15th and 14th centuries BCE. *Sensu lato* it is sometimes extended to include the prior 16th c. Proto-Apennine and Grotta Nuova as well as the subsequent 13th c. Sub-Apennine facies. Earlier studies of material and settlements concluded that the Apennine Culture represented a nomadic pastoralist economy, with seasonal grazing lands connected through a series of seasonal camps (Puglisi 1959, Trump 1966: 109-113, Barker 1975). The perceived unity and in part a nationalistic desire to reduce the importance of external influences led Rellini (1933: 94) to write, “I have come to believe that in the Apennine peoples, having reached the advanced stage of their civilization, we can recognize the *italici* or, if you will, the *proto-italici*.”⁵⁶⁰ He was not the only one to consider that the pastoralist Apennine peoples represented the origins of the Italic *koine* (cf. Puglisi 1959: 96, Barker 1975: 157-8). But his position of an autochthonous Apennine Culture was difficult to reconcile with the fact of the Italic languages’ Indo-European pedigree (cf. Pallottino 1975: 40).

More recent research has shown the Apennine Bronze Age to have been home to a mixed economy of nomadic pastoralism, stock-breeding, and agriculture (Östenberg 1967, Barker 1981: 90-5, Lewthwaite 1981, Albarella 1999: 326-7, Skeates et al. 2021). That some of the Apennine Bronze Age population may have been Indo-European speaking is in fact not impossible. Saupe et al. (2021) confirmed the presence of steppe ancestry in Central Italy by 1600 BCE, and at least one of the sites where it was present (Grotta Regina Margherita) has Middle Bronze Age potsherds assignable to Grotta Nuova and Proto-Apennine facies (Skeates et al. 2021). An additional point of interest is contact between Apennine sites and the Mycenaean civilization. Trade networks have long existed in the Mediterranean, and Mycenaeans came to play an important role in transmitting influences between East and West (Kristiansen 1998: 360), attested in part by the Mycenaean pottery in Apennine sites (e.g. Puglisi 1959: 92-3, Trump 1966: 124-7) and in the Po Valley in the 12th c. BCE (Smith 1996: 25).

7.1.5 Urnfield Horizon

The Urnfield Culture represents the adoption of a series of burial practices (deposition of cremated remains into an urn/container, placement of the urn into a pit, frequent exclusion or reduction of grave goods, esp. weapons) that first arose in central Hungary ca. 2000 BCE. From around 1300 BCE, the Urnfield burial package spread over wide swaths of Central Europe (Cavazzuti et al. 2021), where it gave rise to or at least strongly influenced, among others, the Hallstatt Culture of Western and Central Europe (ca. 1200

⁵⁶⁰ “Ho creduto che nelle genti apenniniche, pervenute alla fase progredita della loro civiltà, si possano riconoscere gli ‘italici’ o se si vuole i ‘protoitalici’.”

to 500 BCE), the Canegrate Culture of Northwest Italy (ca. 1300 to 1200 BCE), and the Proto-Villanova Culture in the whole rest of Italy (ca. 1200 to 1000 BCE).⁵⁶¹ The Proto-Villanova culture underwent regionalization at the beginning of the Iron Age (e.g. Pallottino 1975: 45) into the Este Culture in Veneto (ca. 1000 BCE to 1st c. BCE, famous for its *situlae*), the Villanova Culture of (principally) Tuscany (ca. 900 to 700 BCE and regarded as the earliest phase of the Etruscans),⁵⁶² and the Latial Culture in historic *Latium* (ca. 900 to 700 BCE, famous for its hut urns). Early Latial Culture necropoleis are very similar to those of late the Proto-Villanova in Etruria. While it would continue to be influenced by the Villanova Culture, it would also develop in its own independent ways (Poucet 1985: 21-2).

Because the Latial Culture extends into the historical period and the area of Rome, many have postulated that it was it was the (Proto-)Villanovans, offshoots of the Urnfield tradition, that represented the appearance of the Italic languages in Italy. Gimbutas (1965: 340) proposed that Proto-Italic was spoken in one of the populations of the Urnfield Culture. Von Mehrtart (1942: 65-66) considered it possible that all Italic groups had entered Italy as Proto-Villanovans, perhaps Indo-Europeanizing the Terramare Culture along the way (cf. similarly Freu 1989: 28, who proposed that the Terramare Culture may have spoken an older dialect of Indo-European). Kaschnitz-Weinberg (1954: 354-6) suggested that the regionalization of the Proto-Villanova Culture was due to the influence of a autochthonous Bronze Age Mediterranean substrate, perhaps in some places strong enough to outcompete non-Italic Indo-European dialects. That the Italic languages were brought into Italy with Proto-Villanova is further supported by e.g. Sergent (1995: 418), Gamkrelidze and Ivanov (1995 I: 845), and Anthony (2007: 367).

It was with the arrival of the Urnfield tradition (already in some Terramare sites) that an interesting pattern appeared on the peninsula, one which has led to a series of multiple origin theories for the Italic languages. Cremation burial as a rite was introduced to the Italian peninsula. But it was not ubiquitous. Instead, it was restricted to the North, Latium, Tuscany, and part of Umbria. In the rest of Italy southwards, inhumation was the exclusive funerary practice (Pallottino 1975: 45, cf. von Duhn 1924 on the *verbrennenden* and *bestattenden Italiker*). Several scholars have interpreted these as representing different migrations of peoples, leading to theories involving multiple origins for the Italic languages in Italy.

⁵⁶¹ The latter term was coined by Patroni (1937).

⁵⁶² This is of course potentially problematic for theories that have Italic languages arrive with Proto-Villanova: the descendant populations spoke both Indo-European and non-Indo-European languages (i.e. Etruscan, cf. Mallory 1989: 92-3, Mallory & Adams 1997: 318 on this being a reason to doubt a Villanova origin of Italic). Several possible explanations present themselves: 1) If the Proto-Villanova Culture arrived with a migration, it may have been a multi-ethnic confederation. 2) The Proto-Villanova Culture may have spread into an Italy where Etruscan was already spoken. 3) The Etruscan language may have arrived sometime after the spread of Proto-Villanova.

7.2 Multiple Origin Theories and the Question of Proto-Italic Unity

The geographical position of the Italian peninsula is such that it can be entered relatively directly from the North over the Alps by land or from the East over the Adriatic by sea. There are archaeological connections in both directions (cf. Mallory 1989: 91-2, Recchia 2020 on trans-Adriatic links). Thus Italic languages have been envisioned as arriving from across the Alps (Meyer 1893: 499; von Duhn 1924: 116-17, 439; Ebert 1927: Tafel 103; Watmough 1937: 118), across the Adriatic (Meyer 1909: 792, Patroni 1939: 214-15, Pallottino 1975: 56), or, especially given the possibility of a non-monolithic Italic, both (Palmer 1954: 36-9; Devoto 1962, 1974).⁵⁶³

Early works on Indo-European (e.g. Aufrecht & Kirchhoff 1849: 11, Mommsen 1850: 101, Brugmann 1886: 8-9) took for granted that the Latin (or Latino-Faliscan) and Sabellic (or Osco-Umbrian) language families descended from a Proto-Italic node intermediate to Proto-Indo-European. Walde (1917) challenged this assumption by arguing that a more accurate grouping comprised Proto-Gaelo⁵⁶⁴-Latin and Proto-Sabellic (alongside Proto-Brittonic). A combination of archaeological (the two different burial practices) and linguistic arguments led the idea to develop, rather rapidly, that Latin and Sabellic were independent Indo-European daughter languages. Most proponents of this idea argued that they had converged after migrating into a geographic Italic *Sprachbund* (e.g. Kretschmer 1923a: 105; Devoto 1929: esp. 239-40; Devoto 1931: 51-2; Pisani 1932: 88;⁵⁶⁵ Devoto 1940: 54, 59-69;⁵⁶⁶ Kretschmer 1943: 136-7). Muller (1926: v) had suggested that, if there ever had been a Proto-Italic period, it had been well before the arrival of Latin and Sabellic in Italy. And Ribezzo (1932) proposed that Latin and Sabellic had earlier belonged to a dialect continuum, and that Sabellic diverged due to contact with other languages.

Despite this, the Italic branch continued to be accepted without any mention of question in several handbooks (e.g. Hirt 1927: 20, Buck 1928:2-3, Buck 1933: 23-5, Bloomfield 1933: 61⁵⁶⁷). A brief recognition of the controversy was made by Meillet (1948: 48,

⁵⁶³ Similarly, there are those who are explicitly undecided and consider both options possible (Devoto 1940: 5, Pulgram 1958: 157 [though he prefers the Alpine direction on pg. 136], Mallory 1989: 91-2).

⁵⁶⁴ i.e. Goidelic.

⁵⁶⁵ “Besides the fact that I do not believe in a common Italic at all, it would be time to frankly declare that there has never existed an *Urgriechisch* from which the Greek dialects are derived, an *Urbaltisch-slavisch*, father of the various Baltic and Slavic languages, an *Urgermanisch*, *Urkeltisch*, *Urarisch* in similar relationship with the Germanic, Celtic, and Aryan languages. There have existed, and exist, and will always exist territorial zones...” (“Astraendo dal fatto, che io non credo affatto ad un italice commune, sarebbe ora di dichiarare francamente che non è mai esistito un *Urgriechisch* donde siano derivati i dialetti greci, un *Urbaltisch-slavisch* padre delle varie lingue baltiche e slave, un *Urgermanisch*, un *Urkeltisch*, un *Urarisch* in simile realzione colle lingue germaniche, celtiche, arie. Sono esistite, ed esistono, e sempre esisteranno zone territoriali...”).

⁵⁶⁶ Often cited is his quip, “Le affinità fra latino e osco-umbro sono recenti, le diversità sono antiche” (pg. 67).

⁵⁶⁷ Pagination from the 1984 reprint, (perhaps) not the 1933 first edition.

“L’unité «italique»...est évidente, bien qu’elle ait été récemment contestée”), followed by a full defense of Proto-Italic (pp. 53-72). Jones (1950: 61-2) used Meillet’s own logic against him: after showing in *Dialectes indo-européens* that Indic and Iranian were so similar that passages of Avestan could be transformed into valid Vedic through the application of sound laws, he (Meillet) nevertheless concluded that Indic and Iranian were separate families that fused due to contact. “It is then clear,” Jones (1950: 62) wrote,

that the hypothesis of a common Italic language, parent of Latin and of Osco-Umbrian is not immediately proved by the existence of a few innovations common and peculiar to Latin and Osco-Umbrian, nor on the other hand is it wrecked if some isoglosses of Indo-European date be discovered separating them. An attempt must be made to weigh the *prima facie* evidence for and against unity and to strike a balance.

Touching on the infuriating difficulty of differentiating between archaism and common innovation, he further wrote, “With the exception of obvious borrowings of late date, almost any common feature peculiar to Latin and Osco-Umbrian may, taken in isolation, be regarded as evidence for their earlier unity or as a dialect phenomenon of Indo-European date or again as a product of the period of contact in Italy” (pg. 66). Like those before him and Beeler (1952) after him, Jones (1950) rejected Proto-Italic because he did not see the shared innovations of Latin and Sabellic as numerous, non-trivial, and exclusive enough to reject positions like that of Ribezzo: separate dialects in close proximity whose closeness to each other waxed and waned. Those who supported Proto-Italic (cf. at this time Martinet 1950: 188, Watmough 1951: 82⁵⁶⁸) on the other hand did.

The back-and-forth would continue for decades. Those who rejected Italic unity factored this into their ideas on the Indo-Europeanization of Italy. Devoto (1940: 17) argued that Latin was more archaic than Sabellic and proposed that it had occupied a place on the Western margin of the expanding Indo-European languages. Thus it had arrived in the Italian peninsula first and was subsequently pushed aside and partially overlain by Sabellic (pp. 59-61). He considered an interim homeland to be in Central Germany, but was at first not sure of the details, writing “The two endpoints Thuringia and Rome were joined by a line, how twisted and in what ways we cannot say”⁵⁶⁹ (pg. 5). Pallottino (1940: 28-30) expanded on the idea by pointing out that the trench graves (*Fossakultur*, coined by Säflund 1938: 23, in Italian *cultura della tombe a fossa*) of Italy’s South and Southwest potentially matched the exact area in which Proto-Latin would have been spoken.

⁵⁶⁸ His credibility on linguistic topics is perhaps bolstered by his willingness to admit a less than perfect competence in biology: “The whole discussion reminds me most vividly of a spider, spinning threads out of its belly—or wherever spiders spin threads from.”

⁵⁶⁹ “I due punti estremi Turingia e Roma sono stati congiunti da una linea, quanto e come totuosa non sappiamo dire.”

Devoto (1962: 383-4) accepted the suggestion and revised his hypothesis: Proto-Latin had spread from across the Adriatic to Apulia and fanned out westwards. Other inhuming cultures representing Sabellic, with links to the Danube, spread from the area around Novilara southwards, pushing Proto-Latin further into the West. A third wave of Indo-Europeanization entered Northern Italy as the Terramare Culture, continuing on through the Urnfield horizon, resulting in the languages Lepontic and Venetic and a third language, spreading with the Proto-Villanova culture, which left no trace. Devoto (1974: 51) refined the theory even further by conceiving of a Proto-Latin-Venetic language. It had split up North of the Alps, with Proto-Latin travelling through the Balkans to enter Italy from the Southeast. He again saw Sabellic (this time explicitly travelling across the Adriatic) pushing Proto-Latin to the West and Venetic coming in from the Northeast. His ideological bent culminated in his discussion of the mechanisms behind the language shifts. “The Mediterranean world, including the Italian one, was superior in civilization; an Indo-European cultural conquest is unthinkable,”⁵⁷⁰ he wrote.

This force could only have been social. The nuclei of Indo-European linguistic tradition...maintained their compactness [and] constituted a force of attraction and comparison for the previous inhabitants: first a source of attraction and curiosity, then models of a psychologically urban life, then a solid, fixed point of reference in the changing of daily life, something comparable to a “market”. Only in this way is it possible to effect such a powerful and lasting accomplishment, and at the same time an invisible one”⁵⁷¹ (Devoto 1974: 46).

Palmer (1954: 11) criticized Devoto’s conceptualization of the convergence of Latin and Sabellic as “too remote from the realities of actual speech” but in general agreed with the rejection of Italic unity. He saw Latin and Sabellic as both ultimately originating in Central Europe, with Latin having entered Italy from the North as cremators and being cut off by Sabellic entering from across the Adriatic as inhumers. Venetic, as a separate branch originally closely related to Latin, entered into the North (Palmer 1954: 36-43). The same year, Pisani (1954: 56) quite patronizingly lamented that there were still “hervorragende Anhänger des alten Glaubens” who saw language as an organism that can only change via sound laws and that, “wenn sie sich daher vor zwei ähnlichen Sprachen befinden, so können diese Gelehrten nur an eine Muttersprache denken, davon sich jene unabhängig entwickelt haben.”⁵⁷²

⁵⁷⁰ “Il mondo mediterraneo, ivi compreso quello italiano, era superiore per civiltà; una conquista culturale indeuropea è impensabile.”

⁵⁷¹ “Questa forza non poteva essere che sociale. I nuclei di tradizione linguistica indeuropea...mantenevano la loro compattezza [e] costituivano una forza di attrazione e confronto per gli indigeni: prima, fonte di attrazione e curiosità, poi modelli di vita psicologicamente urbana, poi solido, fisso punto di riferimento nel mutare della vita quotidiana, qualcosa di paragonabile a un “mercato”. Solo in questo modo è possibile rendersi conto di una affermazione così potente e durevole, e nel tempo stesso invisibile.”

⁵⁷² Astoundingly, he continued, “Sie [diese Gelehrten] sind immer bereit zu schwören, dass das Englische eine germanische Sprache ist, ohne darum zu kümmern, wie viel französisches—und nicht nur

Pulgram (1958: 217-18) did not make his position on the existence of Proto-Italic entirely clear, but he envisioned the Indo-European languages of Italy trickling into the peninsula from a source in the Danube Valley as the Polada, Terramare, and Villanova Cultures (pp. 115-15, 120-21). He criticized the idea of the cremating vs. inhuming Italic, suggesting that only cremation was introduced by Indo-European speakers (pp. 220-21).⁵⁷³

Polomé (1972: 59-64) supported a separate Latin and Sabellic as well as a multi-wave migration to Italy based on shared innovations in the political vocabulary between Sabellic and Germanic (e.g. lack of the inherited lexeme *rēx*), suggesting that Latin left before Sabellic.

Pallottino (1975) marks a relatively drastic shift from the variations on a theme so far, undoubtedly because he was informed by a much more up-to-date understanding of the archaeological situation, specifically the fact that Proto-Villanova cultural materials had in fact spread over the whole of the Italian peninsula, with regional differences—including cremation vs. inhumation—developing in response to this (pg. 45). He maintained that the linguistic situation of ancient Italy “is a far cry from the over-simple ‘Italic’ unity conceived by scholars in the past” (pg. 53), and still supported the idea that Latin was pushed into the West by incoming Sabellic (pp. 54-5). But he now argued that 1) all Indo-European languages had entered Italy from across the Adriatic, given that the position of the non-IE languages of Italy cuts them off from the North and 2) that Indo-Europeanization had started well before cremation, already with the Neolithic Square-Mouthed Vases Culture (pp. 58-9). Venetic looks close to Latin (cf. Devoto’s 1974: 51 Proto-Latin-Venetic) simply because it was in contact with the same non-IE languages as it (pg. 55).⁵⁷⁴ Nevertheless, Piazza et al. (1988: 210-11) still seemed to prefer a multi-Italic trickling of tribes into Italy as cremators and later inhumers.

Though Rix (1983: 95, 104) subtly yet negatively assessed the ability to reconstruct a Proto-Italic that was significantly different from Proto-Indo-European, he would change his mind, emphasizing e.g. that the existence of an Italic cultural *koiné* could explain *how* linguistic borrowings occurred, but it is not proof *that* they occurred; each potential case of borrowing must have its own valid, linguistic proof (Rix 1994). Szemerényi (1991:

französisches—darin zusammen geflossen ist...” For someone who is meant to be arguing that Latin and Sabellic were always two separate branches that influenced each other, one would think he might cherish the case of French influence on English as an apt parallel. Instead he uses it to reject the Neogrammarian model. No amount of French influence on English can erase its historically documented development from clearly Germanic Old English to heavily Romance-influenced Present Day English.

⁵⁷³ He further proposed that iron technology was introduced to Italy both from the North, with these cremating Indo-European speakers, as well as from the South, with non-Indo-European Mediterranean peoples who practiced inhumation like the already-present inhabitants of Italy (pp. 121-3).

⁵⁷⁴ He also suggested (pp. 59-60) that Indo-Europeanization may actually have spread from the culturally advanced centers in the South to the more backwater North, but “we have no reliable evidence, and speculation is therefore pointless.”

682-5), based on an analysis of the treatment of the inherited voiced aspirates, had proposed that all the Italic languages had spent time in the region of the Danube, in contact with Greek. He then had a Sichel-Ausonian (cf. §1.2.1.2.3.1) group travel down the West of Italy to reach Sicily, with Proto-Latin moving in behind it. Sabellic would have been displaced and sent into Italy by the Illyrian migrations. Rix (1994: 24-5), though he admitted it was based on little more than linguistic possibility, proposed a similar idea (cf. also the much earlier position of Muller 1926). He saw Proto-Italic (freshly defended by him) as having been spoken in the Sava-Drava watersheds. Venetic left first, then Latin, and finally Sabellic.

Despite some assessments of the Proto-Italic school as outdated, despite e.g. Silvestri (1998) still rejecting Proto-Italic (using “Italic” to refer exclusively to Sabellic) and Clackson and Horrocks (2007: 65-74) considering the evidence cautiously “vague”, the model of a reconstructible Proto-Italic is still held by the majority of scholars today (cf. Poccetti 2017, definitively Weiss 2022b: 116-22). The question is not if Latin and Sabellic share innovations, but rather if they share *enough* to warrant classification as a subgroup (cf. similarly for Italo-Celtic, §4.3.2.3.2). The debate is thus sustained by the fact that an answer to such a question is necessarily subjective. I side with those who consider the body of evidence to be large enough (the gerundive, the treatment of the voiced aspirates, the imperfect subjunctive in **-sē-*, the imperfect indicative in **-fū-*,⁵⁷⁵ etc.). Furthermore, the archaeological details are even better understood today. Thus many of the archaeological scenarios proposed in the literature are impossible, and others can be reevaluated from the better-informed perspective of today’s scholarship (see §8.2).

⁵⁷⁵ Following the Proto-Italic notation I have been using in the linguistic section.

8 Conclusion

8.1 Summary

On the basis of positive evidence—irregular phonological alternations and to a certain extent non-IE morphophonology—I have compiled a list of Latin lexemes that are likely not inherited but have instead originated from or have been mediated by non-IE languages with which Italic (and its changing neighbors) came into contact on its migrations from the steppe to the Italian peninsula.

The phonological alternations on their own, while indicators of non-native origin, are poorly diagnostic of stratigraphy. They become more informative when considered alongside distributional details. In cases of alternation between what can be artificially reconstructed to voicedness and/or aspiration of plosives, Latin tends to attest to the plain, unvoiced variant. Some of the substrate languages with which Italic was in contact left traces of palatalization and labialization in the IE branches that borrowed from them. Other alternations include those between labial plosives and labial nasals as well as between labial plosives and labial approximants. A case can be made for an intrusive nasal before labial plosives being particular to Latin reflexes of some loanwords. Another class of alternation involves the vacillating presence of a sibilant at the beginning of words, word-internally, and especially in interaction with **k*. Other alternations include those between *l* and *r*, a dental and a velar, and several vocalic alternations. Gemination sometimes appears in alternation between Classical Latin and Romance descendants, suggesting that it represents a recent stratum. Morphological features of substrate origin include the phenomenon of *a*-prefixation and concomitant root vocalic reduction and an *n*-suffix; both features are potentially attested across Europe and into the Mediterranean. In a small group of Latin words ending in *-ix/-ex*, the suffix can be shown to be of non-native origin. In other cases it is from different sources, some of which may be inherited. That any of the substrate languages with which Latin was in contact had reduplication as a feature is difficult to confirm.

Distributionally, the non-inherited lexicon of Latin attests to a large group of loanwords acquired in the Mediterranean region as well as a smaller group that must represent earlier contact phenomena. The latter are much more difficult to stratify. There is possible evidence of a velar ~ dental alternation in a contact situation in which Greek did not participate. Considerations on a stage of Italo-Celtic unity lead to different stratigraphic interpretations. Otherwise, comparanda for Latin lexemes of non-IE origin are found (in limited numbers) in languages as far away as the Caucasus, Uralic, Sumerian, and Indo-Iranian. The Uralic case (comparandum to *avēna* ‘oats’) may attest to a Wanderwort or an Eastern substrate language. Certain Indo-Iranian comparanda likely attest to Wanderwörter, although the Iranian comparanda of Lat. *ervum* ‘bitter vetch’ are relevant to the discussion surrounding the descent of individuals of the Sintashta Culture from Corded Ware (cf. Allentoft et al. 2015). As to Caucasian and

Sumerian, both the latter (Schrijver 2017: 362) and the former (Schrijver fthc.) have been suggested as possible points of origin for the *a*-prefix phenomenon.

In any case, I have found very few certain cases of Caucasian comparanda for Latin words of non-IE origin (*ferrum* ‘iron’, *fungus* ‘mushroom, sponge’, and *plumbum* ‘lead’ are probably in Kartvelian and NE Caucasian; *sabulum* ‘sand’ perhaps in NW Caucasian). The metallurgical words are Wanderwörter, though *fungus* may be the result of a very old, widespread word. Similarly, Basque plays a very small role in the data. Both aspects are surprising, given some proposals of a Euskaro-Caucasian substrate in Europe (cf. early on Trombetti 1927: 220, recently Bengtson & Leschber 2021; also Bossong 2017: 859 on Vasconic) for which I find no evidence as concerns what Latin preserves.

Of the non-inherited lexemes in Latin presented here, 40% refer to plants and 20% to animals. All 3 words for vessels, all 3 culinary terms, and 3 of 4 textile terms are shared with Greek and/or attest to a Mediterranean distribution. This is undoubtedly the result of large cultural changes that occurred upon and after settling in the Mediterranean region. On the other hand, the 6 Italo-Celto-Germanic isoglosses lack domesticated species and include *corbis* ‘basket’ and *hasta* ‘spear’, hinting at the much earlier cultural contexts in which they were borrowed. Several (11) non-native words for domesticated plants seem to suggest that at least a portion of the borrowed vocabulary originated in an agricultural substrate. The possibility of a single agricultural substrate language family distributed throughout Europe, in both Mediterranean and non-Mediterranean regions, provides a layer of difficulty in determining where and when agricultural terminology was acquired and therefore where the ancestors of the Italic languages were at various periods of prehistory.

Ancient DNA analyses provide strong support for the steppe hypothesis of Indo-European origins. Steppe ancestry first appears in Northern Italy ca. 2000 BCE, increasing gradually and spreading southwards. Large amounts of admixture with non-steppe-derived populations occurred, resulting in a genetic cline from North to South. The spread of Indo-European languages into Southern Italy seems to have involved much less population replacement than in the North. Archaeologically, there are several possibilities of identifying the arrival of the Italic language branch in the Italian peninsula. With the following combination of the most recent archaeological, genetic, and linguistic data, we can make the most accurate assessment so far.

8.2 Discussion: Triangulating Italic Prehistory

Given that some of the earliest human remains with steppe ancestry were found in Bell Beaker contexts dating to ca. 2000 BCE in Northern Italy (Saupe et al. 2021), the appearance of the Bell Beaker material in Italy marks the earliest point at which Italic languages could have entered Italy. The technology of halberds and their ritual elite use spread to El Argar in Spain ca. 2100-2075 BCE, in some cases in Bell Beaker contexts,

from where they seem to have directly influenced the development of halberd technology in western Central Italy by ca. 2050 BCE (Schuhmacher 2002: 282-4). Given that steppe ancestry appeared on the Iberian peninsula by 2500 BCE and had begun to spread by 2000 BCE (Olalde et al. 2019), the clear communication of these areas in association with their attestation of Bell Beaker materials could be further evidence of the Indo-Europeanization of Italy through the Bell Beaker Culture. However the influence seems to have come from Southeast Spain (Schuhmacher 2002: 282), and it is in the South that steppe ancestry had less of an impact (Olalde et al. 2019). Furthermore, the influence is on the Italian Rinaldone Culture which, despite a very small sample size having been sequenced, has not yet provided evidence of individuals with steppe ancestry (Antonio et al. 2019); in fact steppe ancestry does not seem to appear in Central Italy until ca. 1600 BCE (Saupe et al. 2021). Thus it seems clear that not all Bell Beaker material in Italy marks the arrival of individuals with steppe ancestry. While the Bell Beaker grave that does attest to an individual with steppe ancestry in Northern Italy is found close to a Remedello Culture cemetery (a Copper Age culture roughly contemporary with Rinaldone), it seems possible that any migrations of individuals associated with the diffusion of Bell Beaker Culture may have been too small to have introduced major language shift. Harrison and Heyd (2007: 206-7) note that only a few individuals are needed to “proselytize” the solar cult and specific outwards displays that the Bell Beaker package represented.

A more likely scenario for the Italicization of Italy hearkens in part back to Pigorini. Firstly, the analysis of the linguistic data has shown that loanwords with a Mediterranean distribution (plausibly *ālium*, *faber*, *ficus*, *hirundō*, and *laurus*, cf. §4.3.2.1) entered still-intact Proto-Italic. This means, in other words, that Proto-Italic did not split up until after it entered the Mediterranean zone. On this alone, we cannot rule out a situation whereby, for instance, Proto-Italic was spoken across the Adriatic, still in Mediterranean territory, and multiple waves of migration brought the separate Italic daughter languages to the peninsula. But we also have the evidence of the North-to-South genetic cline as well as the position of the Italic languages themselves. Besides Latin and Sabellic, Venetic, which was spoken mainly in what is now Veneto and Friuli, is most likely to be classified as an Italic language (Weiss 2022b, or at least intermediate to Italic and Celtic, cf. Schrijver 2016). Ancient personal names attested in the area around the Slovenian town of Ig show some similarities with those of Venetic (cautiously, Stifter 2012: 255, 260). All of this points to a(n interim) Proto-Italic homeland in Northeast Italy, placing it within the realm of the Terramare Culture.

The Terramare civilization collapsed around 1150 BCE, perhaps over the course of a few decades, after which site numbers fell to less than a quarter North of the Po River and most settlements were abandoned completely South of it (Moloy, Bruyère, & Jovanović 2023: 148), leading to a diaspora of its inhabitants (Cardarelli 2009, Iacono et al. 2021: 384). The widespread presence of Terramare style pottery in northern Tuscany, Romagna, Umbria, and the Marches (as opposed to its infrequency in southern Etruria

and Latium) and Terramare-style elements in the pottery of some Campanian settlements suggests that some of these areas received significant numbers of Terramare groups (Cardarelli 2009: 507). Additionally, cremation and subsequent interment in Urnfields had begun to be widely adopted at Terramare sites by ca. 1450 BCE (Iacono et al. 2021: 385-6; Moloy, Bruyère, & Jovanović 2023: 149). The diffusion of this rite throughout the Italian peninsula coincides with the diaspora (Cavazzuti et al. 2022: 74), ushering in the era of the Urnfield horizon in Italy. Cremation was not adopted with equal acceptance everywhere (Iacono et al. 2021: 386, Cavazzuti et al. 2022), implying an important ideological component to its spread (cf. Cavazzuti et al. 2021). Note that, instead of a subsequent invasion of ‘inhuming Italici’ (cf. §7.2), the inhuming areas were instead the inhabitants of the peninsula that did not adopt the new burial ritual.⁵⁷⁶ In any case, the collapse of the Terramare settlement region seems to provide evidence of a sizable population movement that could certainly have had a linguistic impact on the Italian peninsula. But the story of course does not begin with Terramare, and events before its collapse provide further support of its role as a potential bearer of the Italic languages.

The first Urnfields actually appeared ca. 2000 BCE amongst the Tell cultures in Hungary. It is from the Pannonian/Hungarian Plain that they would spread into the Po Valley (Trump 1966: 137, Cavazzuti et al. 2022). Contact between these two regions from ca. 1600 BCE onwards was extensive. Both areas’ metalwork evolved in tandem and there are similarities in their development timelines across all levels of society and, as mentioned, all aspects of their mortuary practice (Moloy, Bruyère, & Jovanović 2023: 150-60). It is to such an extent that Moloy, Bruyère, and Jovanović (2023: 158) suggest that “people moving between these regions was predictable and normal in this linked-up social world and it included people settling in communities that were distant from those into which they were born.” Around 1550 BCE, Terramare sites began to see a substantial population increase, difficult to explain in terms of natural population growth, leading Cardarelli (2009: 450) to use the term “colonization” of the Po Plain. While it is likely that an amount of this increase was due to inward migration from surrounding regions (Cardarelli 2018: 362), it is precisely around this time (ca. 1600-1450 BCE) that the Kozsider Period on the Pannonian Plain saw a drastic change in settlement patterns concomitant with contact with the Tumulus Culture from the West. While it was formerly thought that the flourishing Tell civilization in Hungary collapsed under an onslaught of pastoralist warriors, the absence of drastic depopulation suggests more modern interpretations involving demographic decline and the simplification of social structures (Fischl et al. 2013: 355, 360-4), perhaps the toppling of elite systems upon the

⁵⁷⁶ What had been seen as evidence of the trans-Adriatic movement of Sabellic is the Cetina phenomenon, in which ca. 2500 BCE Cetina pottery from Dalmatia appears in Italy in the Northern and Western Adriatic, and crossing the Apennines into Campania (at three sites). This seems to have been brought to Italy by the movement of small groups with an interest in occupying inland areas (Recchia 2020). I consider the possibility of these migrants having been speakers of an Italic language highly unlikely in light of the argumentation here for Proto-Italic from the North.

establishment of new trade connections (Moloy, Bruyère, & Jovanović 2023: 160). It seems likely that some amount of the population, perhaps those disenfranchised by the societal changes (Moloy, Bruyère, & Jovanović 2023: 153) migrated along the previously established corridors into “homophilous” communities, especially the Po Valley (Kristiansen 2018: 118). It is almost certain that the ensuing period of similarity between the Hungarian Plain, especially the Tisza Site Group of the South and the Po Plain, especially the Terramare, between 1500 and 1200 BCE (Moloy, Bruyère, & Jovanović 2023: 157) established an amount of homophily that allowed Urnfield burial customs to spread so rapidly into Italy (Cavazzuti 2022: 74). Moloy, Bruyère, and Jovanović (2023: 161-2) further suggest, given that the megasites of the Tisza Site Group entered a period of crisis and depopulation ca. 1200 BCE, that a migration of inhabitants southwards could have stressed the already ecologically strained Terramare society (cf. Cardarelli 2009: 459, 468), precipitating its ca. 1150 BCE collapse.

The concept of homophily also seems applicable to the Terramare diaspora southwards into central Italian regions. Beside the cultural technology of the halberd, the production, use, and circulation of copper and bronze daggers became an important symbolic phenomenon (van Rosenberg 2013). Of the solid-hilted dagger (*Vollgriffdolch*) there were several regional types. The Italian peninsula was home to two significant categories. One, the “Baltisch-Padanisch” type, is interesting because it is found from the Po Valley over the Bohemian Basin (including in the Únětice Culture) up to the Baltic. But even more interesting are the “Italian types” from Central and Southern Italy (beginning ca. 2000 BCE), because these may in fact be the oldest in Europe, with their technology spreading North in part via their influence on the Baltisch-Padanisch type (Schwenzer 2004: 240-3).

Amongst these Italian types, van Rosenberg (2013) focuses on unique super-sized full-hilted daggers, the earliest of which are from central East Italy on the Adriatic. Amongst their circulation, he located two separate cross-Appennine trading networks with a cultural boundary between them: one to the South amongst the southern Proto-Appennine facies and one to the North amongst the Grotta Nuova facies (both precursors of the Appennine Culture, cf. §7.1.4). The sites on the northern margins of the Grotta Nuova group were integrated into the Terramare. From the other side of things, Cardarelli (2009: 487) notes that, beginning around 1350 BCE, during the formation of the Sub-Appennine facies in northern Central Italy, “a more than marginal role was played by the adoption and reworking of styles and tastes deriving from the Terramare.” Thus, communication between the Terramare Culture and the parts of Italy into which its inhabitants would migrate was relatively long-standing. Perhaps the Italicization of those regions had begun even before the collapse of the Terramare.

Up to now, it seems that a case can be made that the Terramare Culture represents at least in part a Proto-Italic speaking population, and that the Proto-Italic element reached Northern Italy via intensive contact with and possible migrations from the area of the Hungarian Plain. A further consideration is the fact that the Urnfield burial rite spread

from Hungary not only to the Po Valley but from ca. the 14th c. BCE onwards also westward (e.g. Cavazzuti et al. 2022: 74), becoming the Hallstatt Culture in areas that would come to be Celtic-speaking. It seems quite possible then that the homeland of Proto-Italo-Celtic could have been in this region. How can this area be connected to the steppe however? Part of the Yamnaya expansion saw it reach up to the Carpathian Basin (cf. Harrison & Heyd 2007: 194). Anthony (2007: 305) suggests that this could have “incubated the ancestral dialects for several western Indo-European language branches, including Pre-Italic and Pre-Celtic.”

Unless Proto-Germanic was also incubated in that wave of Yamnaya, then the wave of steppe migrants into the Carpathian Basin could not have been speakers of Proto-Italo-Celtic. As the linguistic data in this work shows, the non-native vocabulary of Latin requires it to have been in contact with a substrate language shared with Germanic. Interestingly, this has also been proposed from the perspective of inherited linguistic data (Polomé 1972, 1974, 1981; Bossong 2017: 859). Ideally, such an area would also exclude Greek (cf. already Devoto 1936: 535). Steppe ancestry entered North/Central Europe as the Corded Ware Culture (cf. Haak et al. 2015) beginning around 2900 BCE (Papac et al. 2021). On the other hand, it reached Northern Greece between 2600 and 2000 BCE (Clemente et al. 2021) appearing in reduced proportion in Mycenaean individuals (cf. also Lazaridis et al. 2022). It thus seems likely that the ancestors of Greek- and Armenian-speakers were not a part of the Corded Ware horizon.⁵⁷⁷ The ancestors of Italic-speakers may well have been.

The subsistence strategies of the Corded Ware Culture were diverse, differing between sites and sexes and correlating to mobility. A dietary shift from earlier Neolithic periods is visible in the isotopic data; while it can have several causes, one likely reason is an increased reliance on milk products, suggesting an increasing importance of pastoralism. At the same time, it is clear from especially non-local women that some groups were practicing intensive agriculture (Sjögren, Price & Kristiansen 2016). This has a good parallel in the genetic data, where it is clear that during the Corded Ware period, intrusive males with steppe ancestry were marrying non-local females without steppe ancestry, leading to a period of assimilation whereby Neolithic ancestry components increased into the Bell Beaker period (Papac et al. 2021). It seems plausible that, given such a mosaic of economies in this time and place, the non-domesticated vocabulary shared by Italic, Celtic, and Germanic not present in Greek (and words attesting to the dental ~ velar alternation lacking attestation in Greek) may have entered at this period alongside vocabulary from the more widespread agricultural substrate with which Greek did have contact. Being one of the oldest strata, it makes sense that so few lexemes can be traced back to it with certainty.

⁵⁷⁷ Papac et al. 2021 note that the Y-chromosomal haplogroups of Yamnaya, Corded Ware, and Bell Beaker are different, suggesting that the Yamnaya were not the direct source of the steppe ancestry in Corded Ware or Bell Beaker. Lazaridis et al. 2022 do not agree. In any case, improved resolution of sequences from Italy might allow similar conclusions to be made.

If Italic speakers were indeed descended from Corded Ware individuals, we would need a way to link Corded Ware (ca. 3000-2350 BCE) to the second millennium Hungarian Plain. Gamkrelidze and Ivanov (1995 I: 845, cf. also Gimbutas 1963: 828) for instance suspected that the Únětice Culture (2300-1600 BCE) might represent a time and place in which at least Italic, Celtic, and Germanic would have (still) been in contact.⁵⁷⁸ Furthermore, Únětice developed out of a region in the Bell Beaker horizon upon the addition of population influx from a more Northeastern source, while there is some evidence that the preceding Bell Beaker horizon may have developed locally with some genetic continuity from earlier Corded Ware (Papac et al. 2021). The Tumulus Culture (1600-1300 BCE) subsequently developed in Únětice areas. It spread into the Hungarian Plain after (perhaps helping to precipitate) the collapse of the Tell cultures, remaining during the period of extensive contact between the Hungarian and Po Plains (e.g. Fischl et al. 2013; Moloy, Bruyère, & Jovanović 2023). It was also present in some areas in Italy North of the Po as indicated by e.g. prestige goods in inhumation graves at Olmo di Nogara (Iacono et al. 2021: 385-6). Like for the spread of the Urnfield horizon into Italy (e.g. Cavazzuti 2022 et al.), an aspect of homophily may have been involved in the spread of the Urnfield horizon into areas previously home to the Tumulus Culture. A small amount of cremation had begun to be practiced in the late phases of the Tumulus Culture, gradually growing to put inhumation burials into the minority; and this already before the Urnfield style of cremation ritual rapidly spread westwards, superseding the previous style of cremation (Falkenstein 2012).

Additional support for such a scenario comes in the form of the word for iron. Lat. *ferrum* is plausibly a Wanderwort with its source in Luw. **parza-*. Compelling comparanda include PGM. **brasa-* ‘brass’, Svan *berež* ‘iron’, and Ingush/Chechen *borza* ‘bronze’ (cf. Thorsø et al. 2023: 111-12). Especially the Germanic form suggests that the *f* of *ferrum* could be the result of a borrowing into the chain of developments from PIE **b^h*, i.e. into Proto-Italic. But despite the lexeme’s non-Mediterranean distribution, it is unlikely to have entered Proto-Italic outside of the Mediterranean zone. The earliest certain evidence for iron smelting dates to ca. 1800 BCE in Central Anatolia (Thorsø et al. 2023: 120 with lit.), and iron objects begin to appear in mainland Italy during the Late and Final Bronze Age, including Proto-Villanova contexts (Giardino 2005). But the technology did not reach Western Europe until the Late Hallstatt and La Tène periods in the early first millennium BCE (Thorsø et al. 2023: 120 with lit.). Thus, in order for Italic to have acquired its word for iron from contact with the same substrate as Germanic,⁵⁷⁹ it probably could not have done so before the first century BCE. And there is no plausible archaeological proxy for the arrival of Proto-Italic speakers in Italy at

⁵⁷⁸ Additionally, the Únětice and Nordic Bronze Age cultures were also in extensive contact with the Carpathian Basin (Kristiansen and Larsson 2005: 128-86, Vandkilde 2014).

⁵⁷⁹ This also assumes that PGM. **brasa-* ‘brass’ is not the original meaning of the source lexeme from which Italic and Germanic could have borrowed. But this is a justifiable assumption if Luwian is the closest to the origin of the word. It is uneconomical to assume that Latin later, independently, changed the meaning “back” to iron.

such a late date. Instead, *ferrum* must have reached the Italian peninsula independently of Germanic, as a Wanderwort. The possibility remains that it was borrowed into the chain of Proto-Italic developments from PIE **bʰ*, and it can thus be placed alongside e.g. *ficus* as one of the earliest loanwords taken up on the Italian peninsula. This would further point to the presence of Italic in Italy at the boundary of the Bronze and Iron Ages.

Thus an argument can be made for the origin of Proto-Italic ultimately amongst the inhabitants of the Corded Ware Culture, through the Únětice and Tumulus Cultures, into the Hungarian Plain and its extensive population contact with, most proximally, the Terramare Culture. The actual spread of the Proto-Italic and Proto-Celtic languages seems like it can be associated with the dispersion of the Urnfield Horizon (due at least in part to human migrations, cf. Kristiansen 1998: 385-6). Given the possibility of the transmission of new regional developments back into homophilous adjacent areas, it is not easy to rule out that perhaps an earlier stage, like the Tumulus Culture itself, could have been home to Proto-Italo-Celtic. Perhaps this is more likely, given that the expansion of the Urnfield horizon was not limited to Italic and Celtic areas (continuing into Northern Europe and the Balkans). As the resolution of ancient genome analysis increases, an alternative scenario may present itself: The Tell Cultures of the Hungarian Plain seem to have developed from the contact of a Balkan and a Central/Northwestern European Bell Beaker network ca. 2500/2400-2000/1900 BCE (Fischl et al. 256). If future analyses exclude the descent of Proto-Italic speakers from individuals of the Corded Ware Culture, then perhaps the Hungarian Plain served as more ultimate rather than proximal intermediate homeland of Proto-Italic.

The Mediterranean stratum visible in the Latin words of non-inherited origin is almost certainly composed of material that was present on the peninsula before the arrival of (Proto-)Italic speakers. Apulia, the Eastern Adriatic, and Sicily had been pulled into the periphery of the Mesopotamian Uruk Civilization between ca. 2600-2200 BCE, which at that point already included Anatolia, the Levant, and the Aegean (Harrison & Heyd 2007: 193), probably spurring e.g. the development of the symbology and technology of the full-hilted dagger (Schwenzer 2004: 245). But there were plenty of interactions with Mediterranean societies after the arrival of Italic yet still before the dawn of the historic era. Mycenaean would eventually play an important role in transmitting influences between East and West (Kristiansen 1998: 360). The Adriatic coast of Italy was importing Greek pottery by the Middle Bronze Age and imitating it locally by the Late Bronze Age (Iacono et al. 2021: 386-7). (Recall how in two cases of Mediterranean words, *cupressus* and *hedera*, it seems that the lexemes originated in Pre-Greek and were transmitted to Latin.) Later, contact with Phoenicians and Greeks starting around the 9th c. BCE seems to have had a crucial influence on the economic and political development of central and northern Italy (Kristiansen 1998: 136). In the middle of the first millennium BCE, emporia were established to assist the import of products from across the Mediterranean, from Greece to Egypt. Cultural hybridizations began to occur, visible in e.g. the Greek and local kitchenware of the emporium of Spina (Zamboni 2021: 394).

This relatively late period still seems early enough for Mediterranean terminology to have been available to enter the Italic languages before the widespread occurrence of written texts.

The role of trade networks on the transmission of vocabulary is an important consideration in general. The sort of intensive and shifting trade relationships of the Bronze Age means in theory that words can have been exchanged over long distances without the borrowing languages having ever been spoken near each other. Both Únětice and Nordic Bronze Age Cultures were already in extensive contact with the Carpathian Basin, for example (Kristiansen and Larsson 2005: 128-86, Vandkilde 2014). While trade networks between Northern Europe and the Mediterranean may have led to the exchange of some words for cultural items (*corbis*, *hasta*), it is difficult to imagine that this could be the explanation for e.g. wild animals (*fulica*, *merula*).

In §1.2.2.3, I noted the difficulty of confirming the Etruscan origin of Latin lexemes. This is an unfortunate result of the paucity of well understood Etruscan lexical items. Roman sources attest to the political importance of Etruscan, and Etruscan prestige goods were traded deep into Central Europe (cf. Kristiansen 1998: 322-3). Etruscan even seems to have contributed personal names and morphological elements to (at least Cisalpine) Celtic (cf. McCone 2005: 396). Thus the lack of ample Etruscan influence on Latin is conspicuous. A new perspective on research of this sort, including other poorly understood *Kleinkorpussprachen* of Italy like Rhaetic and North Picene and perhaps with wider implications for languages of the Mediterranean (like Linear A), could include a search for the substrate features attested in the Latin and Greek lexicons (cf. the discussion under §3.3.4 on the potential relationship of the Etruscan and European substrate *n*-suffixes).

I have argued that the **Italicization** of Italy must have occurred from a single Proto-Italic kernel in Northeast Italy. But archaeological and especially genetic studies have made it very likely that the **Indo-Europeanization** of Italy may have occurred in several waves. Could there have been other IE languages spoken in Italy that disappeared? There is no reason why not (cf. Stifter's proposal of a "Bell Beakerish" on the British Isles before Celtic). But so far, it seems they have disappeared without a trace, at least given the current study's methodology. Forthcoming genetic studies on Mediterranean populations will surely help further contextualize the linguistic results of this dissertation (and vice versa). Thus we are growing nearer and nearer to an answer to the question: Latin, *unde vēnistī?*

Bibliography

Electronic Sources

[EDIL] eDIL. 2019. An Electronic Dictionary of the Irish Language, based on the Contributions to a Dictionary of the Irish Language (Dublin: Royal Irish Academy, 1913-1976) (www.dil.ie).

[eDiAna] Digital Philological-Etymological Dictionary of the Minor Ancient Anatolian Corpus Languages. (<https://www.ediana.gwi.uni-muenchen.de/index.php>).

[GPC] *Geiriadur Prifysgol Cymru*. (<https://welsh-dictionary.ac.uk/gpc/gpc.html>).

[LKŽ] *Lietuvių Kalbos Žodynas*. (<http://www.lkz.lt/>).

[NavigAIS] Digital Database of Jaberg, K., Jakob Jud, and Paul Scheuermeier. 1928. *Sprach- und Sachatlas Italiens und der Südschweiz*. (<https://navigais-web.pd.istc.cnr.it/>)

[OED] Oxford English Dictionary Online. Oxford University Press. (<https://www.oed.com/>).

[PSD] The Pennsylvania Sumerian Dictionary. Completed 2006. (<http://psd.museum.upenn.edu/nepsd-frame.html>).

[TLL] *Thesaurus Linguae Latinae*. (<https://tll.degruyter.com/>).

Note on the journal KZ: This is the journal titled from 1852-1874 *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete des Deutschen, Griechischen und Lateinischen*, from 1877-1967 as *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprachen*, from 1968-1987 as *Zeitschrift für vergleichende Sprachforschung*, and from 1988 until now as *Historische Sprachforschung*.

Print Sources

Abaev, Vasily I. 1958-99. *Istoriko-etimologičeskij slovar' osetinskogo jazyka* [Historical-Etymological Dictionary of the Ossetian language]. Moscow/Leningrad: Russian Academy of Science.

Ačařyan, Hrač'ya Hakobi. 1971-79. *Hayerēn armatakan bařaran* [Armenian Root Dictionary]. Second edition. Yerevan: University Press.

Achilli, Alessandro, Anna Olivieri, Maria Pala, Ene Metspalu, Simona Fornarino, Vincenza Battaglia, Matteo Accetturo, Ildus Kutuev, Elsa Khusnutdinova, Erwan Pennarun, Nicoletta Cerutti, Cornelia Di Gaetano, Francesca Crobu, Domeico Palli, Giuseppe Matullo, A. Silvana Santachiara-Benerecetti, Luigi Luca Cavalli-Sforza,

Ornella Semino, Richard Villems, Hans-Jürgen Bandelt, Alberto Piazza, and Antonio Torroni. 2007. Mitochondrial DNA Variation of Modern Tuscans Supports the Near Eastern Origin of Etruscans. *American Journal of Human Genetics* 80(4): 759-68.

Adams, Douglas Q. 1985. Latin *mas* and *masturbari*. *Glotta* 63(3/4): 241-7.

Adams, Douglas Q. 2013. *A Dictionary of Tocharian B*. Revised and Greatly Enlarged. Amsterdam/New York: Rodopi.

Adams, James Noel. 2007. *The Regional Diversification of Latin 200 BCE - AD 600*. Cambridge: Cambridge University Press.

Agostiniani, Luciano. 1998. Sull'origine del nome del vino in etrusco e nelle altre lingue dell'Italia antica. In Agostiniani, Luciano, Maria Giovanna Arcamone, Onofrio Carruba, Fiorella Imparati, and Riccardo Rizza (eds.). *do-ra-qe pe-re. Studi in memoria di Adriana Quattordio Moreschini*, pp. 1-13. Rome: Istituti editoriali e poligrafici internazionali.

[Ahd. Wb.] 1968-present. *Althochdeutsches Wörterbuch*. Leipzig: Sächsische Akademie der Wissenschaften.

Aikio, Ante. 2014. On the reconstruction of Proto-Mari vocalism. *Journal of Language Relationship* 11: 125-158.

Albarella, Umberto. 1999. The Animal Economy after the Eruption of Avellino Pumice: the Case of La Starza (Avellino, Southern Italy). In Livadie, Claude Albore (ed.). *L'Eruzione vesuviana delle "pomici di Avellino" e la facies di Palma Campania (Bronzo Antico)*, pp. 317-30. Bari: Edipuglia.

Alessio, Giovanni. 1935. La base preindoeuropea *Kar(r)a/Gar(r)a: pietra. *Studi Etruschi* 9: 133-51.

Alessio, Giovanni. 1936. La base preindoeuropea *Kar(r)a/Gar(r)a: pietra (Continuazione). *Studi Etruschi* 10: 165-89.

Alessio, Giovanni. 1937. Una voce toscana di origine etrusca. *Studi Etruschi* 11: 253-62.

Alessio, Giovanni. 1939. Un nome di pianta ligure preindoeuropeo *verbascum*. *Studi Etruschi* 13: 317-30.

Alessio, Giovanni. 1941a. L'etrusco e due problemi etimologici latini. *Aevum* 15(4): 545-58.

Alessio, Giovanni. 1941b. Fitonimi mediterranei. *Studi Etruschi* 15: 178-224.

Alessio, Giovanni. 1941c. I dialetti romanzi e il problema del sostrato mediterraneo. *Archivum Romanicum* 25: 140-83.

Alessio, Giovanni. 1943. Raddoppiamento normale in presunte voci del sostrato mediterraneo. *Studi Etruschi* 17: 227-35.

Alessio, Giovanni. 1944a. Suggestimenti e nuove indagini sul problema del sostrato "mediterraneo". *Studi Etruschi* 18: 93-157.

Alessio, Giovanni. 1944b. Il problema dell'individualità dell' Etrusco in un articolo di C. Battisti (Review). *Studi Etruschi* 18: 407-421.

Alessio, Giovanni. 1946a. Un'oasi linguistica preindoeuropea nella regione baltica? *Studi Etruschi* 19: 141-76.

Alessio, Giovanni. 1946b. Relitti mediterranei nel lessico botanico greco e latino. *Annali della Scuola Normale Superiore di Pisa. Lettere, Storia e Filosofia*. Serie II, Vol. 13: 24-51.

Alessio, Giovanni. 1948-9. Vestigia etrusco-mediterranee nella flora toscana. *Studi Etruschi* 20: 109-49.

Alessio, Giovanni. 1950. Problemi di etimologia romanza. *Revue de Linguistique Romane* 17: 28-75.

Alessio, Giovanni. 1955. *Le Lingue indoeuropee nell'ambiente mediterraneo*. Bari: Adriatic editrice.

Alessio, Giovanni. 1957. Stratificazione dei nomi del 'tasso (*Taxus baccata* L.)' in Europa. *Studi Etruschi* 25: 219-64.

Alessio, Giovanni. 1976. *Lexicon Etymologicum*. Naples: Arte Tipografica.

Ališan, Lewond. 1985. *Haybusak kam haykakan busabarut' iwn* [Armenian Botanical Dictionary]. Venice: S. Lazarus Armenian Academy.

Allen, James. 2020. *Ancient Egyptian Phonology*. Cambridge: Cambridge University Press.

Allen, W. Sidney. 1987. *Vox Graeca: A Guide to the Pronunciation of Classical Greek*. 3rd edition. Cambridge: Cambridge University Press.

Allentoft, Morten E., Martin Sikora, Karl-Göran Sjögren, Simon Rasmussen, Morten Rasmussen, Jesper Stenderup, Peter B. Damgaard, Hannes Schroeder, Torbjörn Ahlström, Lasse Vinner, Anna-Sapfo Malaspinas, Ashot Margaryan, Tom Higham, David Chivall, Niels Lynnerup, Lise Harvig, Justyna Baron, Philippe Della Casa, Paweł Dąbrowski, Paul R. Duffy, ...Eske Willerslev. 2015. Population Genomics of Bronze Age Eurasia. *Nature* 522: 167-72.

Alves, Mark. 2022. Lexical Evidence of the Vietic Household Before and After Language Contact with Sinitic. In Phan, Trang, John Phan, and Mark Alves (eds.). *Vietnamese Linguistics: State of the Field*, pp. 15-58. (Journal of the Southeast Asian Linguistics Society Special Publication No. 9). Honolulu: University of Hawai'i Press.

Ammerman, Albert J. and Luigi Luca Cavalli-Sforza. 1971. Measuring the Rate and

Spread of Early Farming in Europe. *Man* 6(4): 674-88.

Ammerman, Albert J. and Luigi Luca Cavalli-Sforza. 1979. The Wave of Advance Model for the Spread of Agriculture in Europe. In Renfrew, C. and K.L. Cooke (eds.). *Transformations. Mathematical Approaches to Culture Change*, pp. 275-93. New York, NY: Academic Press.

Ammerman, Albert J. and Luigi Luca Cavalli-Sforza. 1984. *The Neolithic Transition and the Genetics of Populations in Europe*. Princeton, NJ: Princeton University Press.

Andersen, Henning. 1996. *Reconstructing Prehistorical Dialects: Initial Vowels in Slavic and Baltic*. Berlin: Mouton de Gruyter.

André, Jacques. 1956. Recherches étymologiques sur certains noms de plantes latins. *Latomus* 15(3): 290-307.

André, Jacques. 1959. Remarques sur l'apophonie de -ā- latin en syllabe ouverte devant -l- vélaire. *Bulletin de la Société de Linguistique de Paris* 54(1) : 83-9.

André, Jacques. 1967. *Les noms d'oiseaux en latin*. Paris: Klincksieck.

André, Jacques. 1978. *Les mots a redoublement en latin*. Paris: Klincksieck.

Anthony, David W. *The Horse, the Wheel, and Language*. Princeton/Oxford: Princeton University Press.

Antonio, Margaret L., Ziyue Gao, Hannah M. Moots, Michaela Lucci, Francesca Candilio, Susanna Sawyer, Victoria Oberreiter, Diego Calderon, Katharina Devitofranceschi, Rachael C. Aikens, Serena Aneli, Fulvio Bartoli, Alessandro Bendini, Olivia Cheronet, Daniel J. Cotter, Daniel M. Fernandes, Gabriella Gasperetti, Renata Grifoni, Alessandro Guidi, Francesco La Pastina, ...Jonathan K. Pritchard. 2019. Ancient Rome: A genetic crossroads of Europe and the Mediterranean. *Science* 366(6466): 708-14.

Arranz-Otaegui, Amaia, Sue Colledge, Lydia Zapata, and Juan José Ibáñez. 2016. Regional Diversity on the Timing for the Initial Appearance of Cereal Cultivation and Domestication in Southwest Asia. *PNAS* 113(49): 14001-6.

Aufrecht, S.T. and A. Kirchoff. 1849. *Die Umbrischen Sprachdenkmäler*. Berlin: Ferdinand Dümmler.

Aurino, Paola and Maria De Falco. 2022. Mind the Gap: Bell Beaker Presence in Copper Age Campania, Southern Italy. In Abegg, Claudine, Delia Carloni, Florian Cousseau, Eve Derenne, and Jessica Ryan-Despraz (eds.). *The Bell Beaker Culture in All its Forms*, pp. 211-24. Summertown: Archaeopress.

von Bahder, Karl. 1903. Zur hochdeutschen Lautlehre (hd. *f* = wgerm. *b*). *Indogermanische Forschungen* 14: 258-65.

- Bailey, Harold W. 1979. *Dictionary of Khotan Saka*. Cambridge: Cambridge University Press.
- Bailey, T. Grahame. 1924. *Grammar of the Shina (Ṣiṇā) Language*. London: The Royal Asiatic Society.
- Baldi, Philip. 2002. *The Foundations of Latin*. Berlin/New York: Mouton de Gruyter.
- Bakkum, Gabriël C.L.M. 2009. *The Latin Dialect of the Ager Faliscus. 150 Years of Scholarship*. Amsterdam: University of Amsterdam Press.
- Bammesberger, Alfred. 1990. Lateinisch *fīnis* und urgermanisch **baina-*. *KZ* 103(2): 264-8.
- Barbuĵani, Guido, Giorgio Bertorelle, G. Capitani, and Rosaria Scozzari. 1995. Geographical Structuring in the mtDNA of Italians. *PNAS* 92(20): 9171-5.
- Barbuĵani, Guido, Giorgio Bertorelle, and Lounès Chikhi. 1998. Evidence for Paleolithic and Neolithic gene flow in Europe. *American Journal of Human Genetics* 62(2): 488-92.
- Barfield, Lawrence. 1971. *Northern Italy before Rome*. London: Thames and Hudson.
- Barfield, Lawrence. 2001. Beaker Lithics in Northern Italy. In Nicolis, Franco (ed.). *Bell Beakers Today. Pottery, People, Culture, Symbols in Prehistoric Europe*. Vol. 2, pp. 507-18. Trento: Ufficio Beni Archeologici.
- Bargagna, Mario, Ranieri Domenici, and Agorasti Morali. 1975. Red cell esterase-D polymorphism in the population of Tuscany. *Human Genetics* 29(3): 251-3.
- Barker, Graeme. 1975. Prehistoric Territories and Economies in Central Italy. In Higgs, E.S. (ed.). *Palaeoeconomy*, pp. 111-75. Cambridge: Cambridge University Press.
- Barker, Graeme. 1981. *Landscape and Society. Prehistoric Central Italy*. London/New York/Toronto/Sydney/Tokyo/San Francisco: Academic Press.
- Barocelli, Piero. 1942. Rassegne. Terremare, palatino, orientazione dei *castra* e delle città romane. *Bullettino della Commissione Archeologica del Governatorato di Roma*. 70: 131-44.
- Barrios-Garcia, M. Noelia and Sebastian A. Ballari. 2012. Impact of Wild Boar (*Sus scrofa*) in its Introduced and Native Range: a Review. *Biological Invasions* 14: 2283-2300.
- Bartholomae, Christian. 1961. *Altiranisches Worterbuch*. 2. unver. Aufl. Berlin: de Gruyter.
- Bathe, Max. 1955. Der Vielbaum (Schwarzpappel). *Zeitschrift für Mundartforschung* 23(1): 1-13.
- Batisti, Roberto. 2021. Birds of a Feather? *Indogermanische Forschungen* 126: 205-228.

Battisti, Carlo. 1929. Per lo studio dell'elemento etrusco nella toponomastica italiana. *Studi Etruschi* 1: 327-349.

Battisti, Carlo. 1931. Recensioni e repertorio bibliografico. B. Sez. II – Lingua, Epigrafica. Rivista linguistica etrusca 1929. IV: Rapporti etrusco-latini. *Studi Etruschi* 5: 647-662.

Battisti, Carlo. 1943. Voci mediterranee contestate (*carra, sala, bova, napa, marra, toba, mala, pala, cala* e derivati). *Studi Etruschi* 17: 243-85.

Battisti, Carlo. 1959. *Sostrati e parastrati nell'Italia preistorica*. Florence: F. Le Monier.

Battisti, Carlo. 1960. Il sostrato mediterraneo nella fitonimia greco-latina. *Studi Etruschi* 28: 349-84.

Battisti, Carlo and Giovanni Alessio. 1950-1957. *Dizionario etimologico italiano*. Florence: G. Barbèra.

Beeler, Madison S. 1949. *The Venetic Language*. Berkeley/Los Angeles: University of California Press.

Beeler, Madison S. 1952. The Relation of Latin and Osco-Umbrian. *Language* 28(4): 435-43.

Bedrossian, Matthias. 1875-9. *New Dictionary Armenian-English*. Venice: S. Lazarus Armenian Academy.

van Beek, Lucien. 2022. *The Reflexes of Syllabic Liquids in Ancient Greek*. Leiden: Brill.

Beekes, Robert S.P. 1969. *The Development of the Proto-Indo-European Laryngeals in Greek*. Berlin/Boston: De Gruyter Mouton.

Beekes, Robert S.P. 1971. A European Substratum Word. *Orbis* 20: 132-137.

Beekes, Robert S.P. 1975. E.J. Furnée, Die wichtigsten konsonantischen Erscheinungen des Vorgriechischen (Review). *Lingua* 36: 69-82.

Beekes, Robert S.P. 1988. Laryngeal Developments: A Survey. In Bammesberger, A. (ed.). *Die Laryngaltheorie und die Rekonstruktion des indogermanischen Laut- und Formensystems*, pp. 59-105. Heidelberg: C. Winter.

Beekes, Robert S.P. 1992. Widow. *KZ* 105(2): 171-88.

Beekes, Robert S.P. 1996. Ancient European Loanwords. *KZ* 109: 215-236.

Beekes, Robert S.P. 1998. The Origin of Lat. *aqua*, and of **teutā* 'People'. *Journal of Indo-European Studies* 26(3/4): 459-466.

Beekes, Robert S.P. 1999. The Greek Word for 'Lead'. *Münchener Studien zur Sprachwissenschaft* 59: 7-14.

- Beekes, Robert S.P. 2000. European Substratum Words in Greek. In Ofitsch, Michael and Christian Zinko (eds.). *125 Jahre Indogermanistik in Graz*, pp. 21-31. Graz: Leykam.
- Beekes, Robert S.P. 2003. *The Origin of the Etruscans*. Amsterdam: Koninklijke Nederlandse Academie van Wetenschappen.
- Beekes, Robert S.P. 2009. Pre-Greek Names. *Journal of Indo-European Studies* 37(1/2): 191-7.
- Beekes, Robert S.P. 2011. *Comparative Indo-European Linguistics: An Introduction*. Second edition. Amsterdam: John Benjamins.
- Beekes, Robert S.P. 2014. *Pre-Greek: Phonology, Morphology, Lexicon*. Leiden: Brill.
- Belardi, Walter. 1950. Linguistica etimologie greche. *DOXA* 3: 191-224.
- Bellwood, Peter. 2005. *First Farmers: The Origins of Agricultural Societies*. Malden, MA: Blackwell.
- Bellwood, Peter and Colin Renfrew (eds.). 2002. *Examining the farming/language dispersal hypothesis*. Cambridge: McDonald Institute for Archaeological Research.
- Bengtson, John D. and Corinna Leschber. 2021. Notes on Some Pre-Greek Words in Relation to Euskaro-Caucasian (North Caucasian + Basque). *Journal of Language Relationship* 19(2): 71-98.
- Benveniste, Émile. 1954. Études hittites et indo-européennes. *Bulletin de la Société de Linguistique de Paris* 50(1): 29-43.
- Berger, Hermann. 1956. Mittelmeerische Kulturpflanzen aus dem Burushaski. *Münchener Studien zur Sprachwissenschaft* 9: 4-33.
- Berger, Hermann. 1966. Remarks on Shina Loans in Burushaski. In Dil, Anwar S. (ed.). *Shahidullah Presentation Volume*, pp. 79-88. Lahore: Linguistic Research Group of Pakistan.
- Bernabé, Alberto. 2021. The Noun for 'Horse' in Mycenaean and Some Related Terms. In Giannakis, Georgios, Luz Conti, Jesús de la Villa, and Raquel Fornieles (eds.). *Synchrony and Diachrony of Ancient Greek*, pp. 115-124. Berlin/Boston: De Gruyter.
- Berneker, Erich. 1908-14. *Slavisches Etymologisches Wörterbuch*. Heidelberg: Carl Winter's Universitätsbuchhandlung.
- Bertoldi, Vittorio. 1928. Antichi filoni nella toponomastica mediterranea incrociatisi nella Sardegna. *Revue de Linguistique Romane* 4: 222-50.
- Bertoldi, Vittorio. 1931. Problèmes de substrat: Eassi de méthodologie dans le domaine préhistorique de la toponymie et du vocabulaire. *Bulletin de la Société Linguistique de Paris* 32: 93-184.

- Bertoldi, Vittorio. 1933a. Relitti etrusco-campani. *Studi Etruschi* 7: 279-93.
- Bertoldi, Vittorio. 1933b. “Preellenico” βάτος, μαντία “cespuglio, rovo” e “preromano” *matta, * mantia “cespuglio, rovo”. *Glotta* 21(3/4): 258-67.
- Bertoldi, Vittorio. 1936. “Nomina tusca” in Dioscoride. *Studi Etruschi* 10: 295-320.
- Bertoldi, Vittorio. 1937a. Contatti e conflitti di lingue nell’antico Mediterraneo. *Zeitschrift für romanische Philologie* 57(2): 139-169.
- Bertoldi, Vittorio. 1937b. Plurale Mediterraneo in Residui Fossili. In *Mélanges de linguistique et de philologie offerts à Jacq. van Ginneken à l’occasion du soixantième anniversaire de sa naissance (21 avril 1937)*, pp. 157-69. Paris: Klincksieck.
- Bertoldi, Vittorio. 1939a. *Questioni di metodo nella linguistica storica*. Naples: Stabilimento etnografico editoriale.
- Bertoldi, Vittorio. 1939b. Prerogative culturali mediterranee nel lessico Greco e Latino. *Archivio Glottologico Italiano* 31: 85-101.
- Bertoldi, Vittorio. 1942. *Linguistica storica. Questioni di metodo*. Seconda edizione interamente riveduta e rinnovata. Naples: Dante Alighieri.
- Bertoldi, Vittorio. 1948. Storia d’una tradizione mediterranea di lingua e di cultura. Aspetti di cultura mediterranea espressi dalle voci gemelle *finis* e *funis* e da altri nomi inidcanti “giunco” e “fune”. *Museum Helveticum* 5(2): 69-82.
- Bertoni, Giulio. 1910. Ancora di *-f-* italico e *-b-* latino e dei loro continuatori romanzi. *Rivista di Filologia e di Istruzione Classica* 38: 25-37.
- Bertoni, Giulio. 1925. Franc. “sapin”. *Archivum Romanicum* 9: 422-3.
- Bezenberger, Adalbert. 1911. Lat. *faex*. *KZ* 44(1/2): 22.
- Bezenberger, Adalbert and August Fick. 1881. Nachträge zum indogermanische Wörterbuch. *Beiträge zur Kunde der indogermanischen Sprachen* 6: 235-40.
- Bichlmeier, Harald. 2016. Zur Frage des nichtindogermanischen Substrats im Germanischen: ein kurzer Überblick. *Slovo a slovesnost* 77: 316-36.
- Bietti Sestieri, Anna Maria. 2010. *L’Italia nell’età del bronzo e del ferro*. Rome: Carocci.
- Billerbeck, Julius. 1824. *Flora Classica*. Leipzig: J.C. Heinrichssche Buchhandlung.
- [Biville] Biville, Frédérique. 1990-5. *Les emprunts du latin au grec: approche phonétique*. Louvain/Paris: Éditions Peeters.
- Blažek, Václav. 2003. The Prefix **sm-* / **som-* in the Indo-European Tree- & Plant-Names. *Sborník prací filozofické fakulty brněnské university* 52(A51): 5-11.

- Blažek, Václav. 2005. Lexica Anatolica. Hieroglyphic Luwian contribution to the Indo-European lexicon. *Balkansko Ezikoznanie [Linguistique Balkanique]* 43(1): 3-11.
- Blažek, Václav. 2010. *Indo-European “Smith” and his Divine Colleagues* (JIES Monograph No. 58). Washington D.C.: Institute for the Study of Man.
- Blažek, Václav. 2013. Indo-European Zoonyms in Afroasiatic Perspective. *Journal of Language Relationship* 9: 37-54.
- Blažek, Václav. 2014. Etymologizing ‘Unetymologizable’ Greek Dendronyms. *Graeco-Latina Brunensia* 19(1): 41-50.
- Bloomfield, Leonard. 1933. *Language*. New York: Henry Holt.
- Boisacq, Émile. 1911-12. Grec κράνος “cornouiller”, κέρασος “cerisier”, et congénères. *Mémoires de la Société de Linguistique* 17: 56-59.
- Boisacq, Émile. 1916. *Dictionnaire étymologique de la langue Grecque*. Heidelberg: Carl Winter/ Paris: Klincksieck.
- de Bolòs, Oriol and Josep Vigo. 1984-2001. *Flora dels Països Catalans*. 4 vols. Barcelona: Editorial Barcino.
- Bonfante, Giuliano. 1985. Etruscan Words in Latin. *Word* 36(3): 203-10.
- Bosson, Georg. 2017. The evolution of Italic. In Klein, Jared, Brian Joseph, and Matthias Fritz (eds.). *Handbook of Comparative and Historical Indo-European Linguistics*. Vol. 2, pp. 858-74. Berlin/Boston: De Gruyter Mouton.
- Bottiglioni, Gino. 1943. Di alcuni presunti sabinismi del lessico latino. *Studi Etruschi* 17: 315-26.
- Boutkan, Dirk. 1995. Altgermanisch ‘Held’ und die Entwicklung von unbetontem *e im Altnord- und Westgermanisch. *Amsterdamer Beiträge zur Älteren Germanistik* 41: 1-7.
- Boutkan, Dirk. 1998. On the form of North European substrate words in Germanic. *KZ* 111(1): 102-33.
- Boutkan, Dirk and Marten Kossmann. 1999. Some Beber Parallels of European Substratum Words. *Journal of Indo-European Studies* 27(1/2): 87-100.
- Boutkan, Dirk and Sjoerd Michiel Siebinga. 2005. *Old Frisian Etymological Dictionary*. Leiden/Boston: Brill.
- Bramanti, Barbara, Mark G. Thomas, Wolfgang Haak, Martina Unterländer, Pia Jores, Kristiina Tambets, I. Antanaitis-Jacobs, Miriam N. Haidle, Rimantas Jankauskas, C.J. Kind, Friedrich Lüth, Thomas Terberger, J. Hiller, Shuichi Matsumura, Peter Forster, and Joachim Burger. 2009. Genetic Discontinuity Between Local Hunter-Gatherers and Central Europe’s First Farmers. *Science* 326(5949): 137-40.

Braune, Wilhelm. 2018. *Althochdeutsche Grammatik I*. 16th edition by Frank Heidermanns. Berlin/Boston: Walter de Gruyter.

Bréal, Michel. 1908. Étymologies latines et grecques. *Mémoires de la Société linguistique de Paris* 15: 137-51.

Breyer, Gertraud. 1993. *Etruskisches Sprachgut im Lateinischen unter Ausschluss des spezifisch onomastischen Bereiches*. Leuven: Uitgeverij Peeters.

Briand, Michel. 1997. Grec κάπρος: de (porc) vorace au sanglier. In Mellet, Sylvie (ed.). *Les Zoonymes: actes du colloque international tenu à Nice les 23, 24 et 25 janvier 1997*, pp. 91-115. Nice: Centre de recherches comparatives sur les langues de la Méditerranée ancienne.

[Brown-Driver-Briggs] Brown, Francis, Samuel R. Driver, and Charles A. Briggs. 1906. *A Hebrew and English Lexicon of the Old Testament*. Oxford: Clarendon Press.

Brüch, Josef. 1914. Zwei ligurische Wörter im Lateinisch-Romanischen. *KZ* 46(4): 351-73.

Brüch, Josef. 1916. Lat. *blatea*, *balatro*, und genues. *brata*. *Glotta* 8(1/2): 83-5.

Brüch, Josef. 1919a. Lat. *ballaena*. *Glotta* 10(3): 198-9.

Brüch, Joseph. 1919b. Sabinisches *alpus*. *Glotta* 19(3): 193-8.

Brüch, Josef. 1921. Etymologien. *Zeitschrift für Romanische Philologie* 41(3): 582-6.

Brüch, Josef. 1922. Lateinische Etymologien. *Indogermanische Forschungen* 40(1): 198-247.

Brüch, Josef. 1923. Lat. *larix* "Lärche". *Indogermanische Forschungen* 41(1): 377-9.

Brüch, Josef. 1933. Lat. *sulphur* und nhd. Schwefel. *Glotta* 21: 72-6.

Brugmann, Karl. 1894. Μοῦσα; τρίαῖνα, θρῖναξ, Θρῖνακίη; ἦνεικα. *Indogermanische Forschungen* 3: 253-64.

Brugmann, Karl. 1886. *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*. Vol. 1, 1st edition. Strasbourg: Karl. J. Trübner.

Brugmann, Karl. 1906. *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*. Vol. 2, Part 1. 2nd edition. Strassburg: Karl J. Trübner.

Brugmann, Karl. 1908. Griechisch ἵνυος und ὄνος. *Indogermanische Forschungen* 22: 197-202.

Brust, Manfred. 2008. *Die indischen und iranischen Lehnwörter im Griechischen*. 2nd improved and expanded edition. Innsbruck: IBS.

Buck, Carl Darling. 1904. *A Grammar of Oscan and Umbrian*. Boston: Ginn &

Company.

Buck, Carl Darling. 1933. *Comparative Grammar of Greek and Latin*. Chicago: University of Chicago Press.

Budenz, József. 1859. Lateinische etymologien. *KZ* 8(4): 287-93.

Buecheler, Franz. 1927. *Kleine Schriften*. Vol. 2. Leipzig/Berlin: B.G. Teubner.

Bugge, Sophus. 1893. Beiträge zur etymologischen erklärung der armenischen sprache. *KZ* 32(1): 1-87.

Çabej, Eqrem. 1976. *Studime Gjuhësore*. Vol 1. Prishtinë: Rilindja.

[CAD] 1956-2010. *The Assyrian Dictionary of the Oriental Institute of the University of Chicago*. Chicago: The Oriental Institute.

Campanile, Enrico. 1974. *Profilo Etimologico del Cornico Antico*. Pisa: Pacini.

Campbell, Lyle. 1990. Indo-European and Uralic Tree Names. *Diachronica* 7(2): 149-80.

Capelli, Cristian, Francesca Brisighelli, Francesca Scarnicci, Barbara Arredi, Alessandra Caglià, Giuseppe Vetrugno, Sergio Tofanelli, Valerio Onofri, Adriano Tagliabracci, Giorgio Paoli, Vincenzo L. Pascali. 2006. Y Chromosome Genetic Variation in the Italian Peninsula is Clinal and Supports an Admixture Model for the Mesolithic-Neolithic Encounter. *Molecular Phylogenetics and Evolution* 44: 228-239.

Cardarelli, Andrea. 2009. The Collapse of the Terramare Culture and Growth of New Economic and Social Systems during the Late Bronze Age in Italy. *Scienze dell'Antichità* 15: 449-520.

Cardarelli, Andrea. 2018. Before the city: the last villages and proto-urban centres between the Po and Tiber Rivers', in Frangipane, M. and L. Manzanilla (eds.). *Rethinking Urbanization and its Living Landscapes from the Inspiring Perspective of a Great 'Maestro'*, pp. 359-82. Rome: Gangemi Editore.

Carnoy, Albert. 1955a. Etyma Pelasgica. *L'Antiquité Classique* 24(1): 5-28.

Carnoy, Albert. 1955b. Les noms du lièvre et du lapin. *Revue belge de Philologie et d'Histoire* 33(3): 597-601.

Cavalli-Sforza, Luigi Luca and Eric Minch. 1997. (Letters to the Editor) Paleolithic and Neolithic Lineages in the European Mitochondrial Gene Pool. *American Journal of Human Genetics* 61(1): 247-51.

Cavalli-Sforza, Luigi Luca and Albert Piazza. 1975. Analysis of Evolution: Evolutionary Rates, Independence and Treeness. *Theoretical Population Biology* 8: 127-65.

Cavazzuti, Claudio, Tamás Hajdu, Federico Lugli, Alessandra Sperduti, Magdolna

Vicze, Aniko Horváth, István Major, Mihály Molnár, László Palcsu, and Viktória Kiss. 2021. Human mobility in a Bronze Age Vátya ‘urnfield’ and the life history of a high-status woman. *PLoS ONE* 16(7): e0254360.

Cavazzuti, Claudio, Alberta Arena, Andrea Cardarelli, Michaela Firtzl, Mario Gavranović, Tamás Hajdu, Viktória Kiss, Kitti Köhler, Gabriella Kulcsár, Eszter Melis, Katharina Rebay-Salisbury, Géza Szabó, and Vajk Szeverényi. 2022. The First ‘Urnfields’ in the Plains of the Danube and the Po. *Journal of World Prehistory* 35: 45-86.

Chadwick, John and Lydia Baumbach. 1963. The Mycenaean Greek Vocabulary. *Glotta* 41(3/4): 157-271.

Chantraine, Pierre. 1933. *La formation des noms en grec ancien*. Paris: Librairie ancienne Honoré Champion.

Chantraine, Pierre. 1968-80. *Dictionnaire étymologique de la langue grecque*. Paris: Klincksieck.

Charpentier, Jarl. 1917. Zur Italischen Wortkunde. *Glotta* 9(1): 33-69.

Cheung, Johnny. 2007. *Etymological Dictionary of the Iranian Verb*. Leiden: Brill.

Chierici, Gaetano. 1871. *Le antichità preromane della provincia di Reggio nell’Emilia*. Reggio Emilia: Tip. di S. Calderini.

Chierici, Gaetano. 1881. La terramara di Bellanda nel Mantovano. *Bullettino di Paleontologia Italiana* 7(5/6): 68-86.

Chikhi, Lounès, Giovanni Destro-Bisol, Giorgio Bertorelle, Vincenzo Pascali, and Guido Barbujani. 1998. Clines of nuclear DNA markers suggest a largely Neolithic ancestry of the European gene pool. *PNAS* 95(15) : 9053-8.

Chikhi, Lounès, Richard A. Nichols, Guido Barbujani, and Mark A. Beaumont. 2002. Y genetic data support the Neolithic demic diffusion model. *PNAS* 99(17): 11008-13.

Childe, Vere Gordon. 1926. *The Aryans: A study of Indo-European origins*. London: Kegan Paul, Trench, Trübner & Co.

Childe, Vere Gordon. 1925. *The Dawn of European Civilization*. New York: Alfred A. Knopf.

Childe, Vere Gordon. 1929. *The Danube in Prehistory*. Oxford: Clarendon Press.

Childe, Vere Gordon. 1936. *Man Makes Himself*. London: Watts & Company.

Christol, Alain. 1996. Le rhotacisme: anomalies phonétiques, anomalies morphologiques. *Latomus* 55(4): 806-14.

Ciorănescu, Alexandru. 1958-1966. *Dicționarul etimologic al limbii române*. Tenerife:

Universidad de la Laguna.

Clackson, James. 1994. *The Linguistic Relationship between Armenian and Greek*. Oxford/Cambridge: Blackwell.

Clackson, James. 2017. The Lexicon of Armenian. In Klein, Jared, Brian Joseph, and Mathias Fritz (eds.). *Handbook of Comparative and Historical Indo-European Linguistics*, Vol. 2, pp. 1115–1132. Berlin/Boston: De Gruyter Mouton.

Clackson, James and Geoffrey Horrocks. 2007. *The Blackwell History of the Latin Language*. Malden, MA: Blackwell.

Clemente, Florian, Martina Unterländer, Olga Dolgova, Carlos E.G. Amorim, Francisco Coroado-Santos, Samuel Neuenschwander, Elissavet Ganiatsou, Diana I. Cruz Dávalos, Lucas Anchieri, Frédéric Michaud, Laura Winkelbach, Jens Blöcher, Yami Ommar Arizmendi Cárdenas, Bárbara Sousa de Mota, Eleni Kalliga, Angelos Souleles, Ioannis Kontopoulos, Georgia Karamitrou-Mentessidi, Olga Philaniotou, Adamantios Sampson, ...Christina Papageorgopoulou. 2021. The Genomic History of the Aegean Palatial Civilizations. *Cell* 184(10): 2565-2586.e21

Cohen, Marcel. 1931. Quelques mots périméditerranéens: lis, scorpion, sabre. *Bulletin de la Société de linguistique de Paris* 31: 37-41.

Collitz, Hermann. 1912. *Das schwache Präteritum und seine Vorgeschichte*. Göttingen: Vandenhoeck & Ruprecht.

Conway, R. Seymour. 1893. On the Change of *d* to *l* in Italic. *Indogermanische Forschungen* 2: 157-67.

Conway, R. Seymour. 1897. *The Italic Dialects*. Cambridge: Cambridge University Press.

Cooper, Brian. 2005. The Russian Names for Some Grasses of the Triticeae Tribe. *Russian Linguistics* 2: 227-42.

Čop, Bojan. 1969. Beiträge zur indogermanischen Wortforschung VII. *Linguistica* 9(2): 187-96.

Corominas, Joan and José A. Pascual. 1984-91. *Diccionario crítico etimológico castellano e hispánico*. Madrid: Gredos.

Corssen, Wilhelm. 1863. *Kritische Beiträge zur lateinischen Formenlehre*. Leipzig: B.G. Teubner.

Craddock, Jerry Russell. 1969. *Latin Legacy versus Substratum Residue*. Berkeley/Los Angeles: University of California Press.

Cunningham, Ian C. (ed.). 2018-20. *Hesychii Alexandrini Lexicon (recensuit et emendavit Kurt Latt e editionem alteram curavit Ian C. Cunningham)*. Berlin/Boston: De

Gruyter.

Cuny, Albert. 1909. Grec αἰγίλωψ ‘espèce de chêne’, latin *īlex*. *Indogermanische Forschungen* 26(2): 21-6.

Cuny, Albert. 1910. Les mots du fonds préhellénique en grec, latin et sémitique occidental. *Revue des Études Anciennes* 12(2): 154-64.

Curtius, Georg. 1858. *Grundzüge der griechischen Etymologie*. Vol. I. Teubner: Leipzig.

Curtius, Georg. 1894. Miscellen. *KZ* 13(5): 395-400.

Dal Santo, Nicola, Alessandro Ferrari, Gabriella Morico, and Giuliana Steffè. 2014. Bell Beaker in Eastern Emilia (Northern Italy). In Besse, Marie (ed.). *Around the Petit-Chausseur Site in Sion (Valais, Switzerland) and New Approaches to the Bell Beaker Culture*, pp. 205-36. Oxford: Archaeopress.

De Fanti, Sara, Chiara Barbieri, Stefania Sarno, Federica Sevini, Dario Vianello, Erika Tamm, Ene Metspalu, Mannis van Oven, Alexander Hübner, Marco Sazzini, Claudio Franceschi, Davide Pettener, and Donata Luiselli. 2015a. Fine Dissection of Human Mitochondrial DNA Haplogroup HV Lineages Reveals Paleolithic Signatures from European Glacial Refugia. *PLoS ONE* 10(12): e0144391. <https://doi.org/10.1371/journal.pone.0144391>

De Fanti, Sara, Marco Sazzini, Cristina Giuliani, Federica Frazzoni, Stefania Sarno, Alessio Boattini, Elena Marasco, Vilma Mantovani, Claudio Franceschi, Pedro Moral, Paolo Garagnani, and Donata Luiselli. 2015b. Inferring the genetic history of lactase persistence along the Italian peninsula from a large genomic interval surrounding the *LCT* gene. *American Journal of Physical Anthropology* 158(4): 708-18.

Defrémery, Charles. 1862. *Journal asiatique* 1: 89.

Delamarre, Xavier. 2003. *Dictionnaire de la langue gauloise: une approche linguistique du vieux-celtique continental*. 2^e édition. Paris: Editions Errance.

De Marinis, Raffaele C. and Annaluisa Pedrotti. 1997. L’età del rame nel versante italiano delle alpi centro-occidentali. In *Atti della XXXI Riunione Scientifica: la valle d’Aosta nel quadro della preistoria e protostoria dell’arco alpino centro-occidentale*, pp. 247-300. Florence: Istituto Italiano di Preistoria e Protostoria.

Demiraj, Bardhyl. 1997. Albanische Etymologien. Amsterdam/Atlanta : Rodopi.

Deroy, Louis. 1956a. La valeur du suffixe préhellénique *-nth-* d’après quelques noms grecs [*sic*] en *-θος*. *Glotta* 35 (3/4) : 171-95.

Derksen, Rick. 1999. North European **KulP-* ‘beak, snout’. *Baltistica* 34(1): 69-76.

Derksen, Rick. 2002. “Rozwadowski’s change“ in Baltic. *Baltu Filologija* 11(1): 5-12.

Derksen, Rick. 2007. *Etymological Dictionary of the Slavic Inherited Lexicon*. Leiden:

Brill.

Derksen, Rick. 2014. *Etymological Dictionary of the Baltic Inherited Lexicon*. Leiden: Brill.

Deroy, Louis. 1956b. A propos du nom étrusque de l'haruspice. *Latomus* 15(1): 206-14.

Deshayes, Albert. 2003. *Dictionnaire étymologique du Breton*. Douarnenez: Le Chasse-Marée.

Devoto, Giacomo. 1928. L'Etrusco come intermediario di parole greche in latino. *Studi Etruschi* 2: 307-41.

Devoto, Giacomo. 1929. Italo-greco e italo-celtico. *Archivio Glottologico Italiano* 22-3: 200-40 (*Silloge Linguistica. Dedicata alla memoria di Graziadio Isaia Ascoli nel primo centenario della nascita*).

Devoto, Giacomo. 1931. *Gli antichi italici*. Florence: Vallecchi.

Devoto, Giacomo. 1936. Germanisch-lateinisch und Germanisch-oskisch-umbrisch. In Arntz, Helmut (ed.). *Germanen und Indogermanen. Volkstum, Sprache, Heimat, Kultur. Festschrift für Herman Hirt*, Vol. 2, pp. 533-47. Heidelberg: Carl Winter.

Devoto, Giacomo. 1940. *Storia della lingua die Roma*. Bologna: Licinio Cappelli.

Devoto, Giacomo. 1943. Pelasgo e peri-indoeuropeo. *Studi Etruschi* 17: 359-67.

Devoto, Giacomo. 1962. *Origini Indoeuropee*. Florence: Sansoni.

Devoto, Giacomo. 1974. *Il linguaggio d'Italia*. Milan: Rizzoli.

Diamond, Jared and Peter Bellwood. 2003. Farmers and Their Languages: The First Expansions. *Science* 300(5619): 597-603.

Dickson, James H., Klaus Oeggl, and Linda L. Handley. 2003. The Iceman Reconsidered. *Scientific American* 288: 70-9.

Diebold, A. Richard. 1985. *The Evolution of Indo-European Nomenclature for Salmonid Fish. The Case of Huchen (Hucho spp.)*. Washington, D.C.: Institute for the Study of Man.

Di Gaetano, Cornelia, Floriana Voglino, Simonetta Guarrera, Giovanni Fiorito, Fabio Rosa, Anna Maria Di Blasio, Paola Manzini, Irma Dianzani, Marta Betti, Daniele Cusi, Francesca Frau, Cristina Barlassina, Dario Mirabelli, Corrado Magnani, Nicola Glorioso, Stefano Bonassi, Alberto Piazza, and Giuseppe Matullo. 2012. An Overview of the Genetic Structure within the Italian Population from Genome-Wide Data. *PLoS ONE* 7(9): e43759.

Di Giacomo, Francesco, Francesca Luca, Nicholas Anganou, G. Ciavarella, Rosa Maria Corbo, Marie Cresta, F. Cucci, L. Di Stasi, V. Agostiano, M. Giparaki, A. Loutradis,

Corrado Mammi, Emmanuel N. Michalodimitrakis, Franco Papola, G. Pedicini, Erving Plata, Luciano Terrenato, Sergio Tofanelli, Patrizia Malaspina, and Andrea Novelletto. 2003. Clinal Patterns of Human Y Chromosomal Diversity in Continental Italy and Greece are Dominated by Drift and Founder Effects. *Molecular Phylogenetics and Evolution* 28(3): 387-95.

Dolfini, Andrea. 2014. Early Metallurgy in the Central Mediterranean. In Roberts, B.W. and C.P. Thornton (eds.) *Archaeometallurgy in Global Perspective*, pp. 473-506. New York: Springer.

Douka, Katerina, Stefano Grimaldi, Giovanni Boschian, Angiolo del Lucchese, and Thomas F.G. Higham. 2012. A new chronostratigraphic framework for the Upper Paleolithic of Riparo Mochi (Italy). *Journal of Human Evolution* 62(2): 286-99.

Douka, Katerina, Nikos Efstratiou, Mette Marie Hald, Peter Steen Henriksen, and Alexandra Karetsou. 2017. Dating Knossos and the arrival of the earliest Neolithic in the southern Aegean. *Antiquity* 91(375): 304-21.

Driessem, C. Michiel. 2005. On the Etymology of Lat. *fulvus*. In Meiser, Gerhard and Olav Hackstein (eds.). *Sprachkontakt und Sprachwandel: Akten der XI. Fachtagung der Indogermanischen Gesellschaft, 17.-23. September 2000, Halle an der Saale*, pp. 39-64. Wiesbaden: Dr. Ludwig Reichert Verlag.

von Duhn, Friedrich. 1924. *Italische Gräberkunde*. Vol. 1. Heidelberg: Carl Winter.

[DV] de Vaan, Michiel. 2008. *Etymological Dictionary of Latin and the Other Italic Languages*. Leiden: Brill.

Ebert, Max. 1927. *Reallexikon der Vorgeschichte*. Vol. 9. Berlin: De Gruyter.

[EDG] Beekes, Robert S.P. and Lucien van Beek. 2010. *Etymological Dictionary of Greek*. Leiden: Brill.

Edwards, Anthony W.F. and Luigi Luca Cavalli-Sforza. 1964. Reconstruction of Evolutionary Trees. In Heywood, V.H. and J. McNeil (eds.). *Phenetic and Phylogenetic Classification*, pp. 67-76. London: The Systematics Association.

Edwards, Anthony W.F. and Luigi Luca Cavalli-Sforza. 1965. A Method for Cluster Analysis. *Biometrics* 21(2): 362-75.

Edwards, A.W.F. and L.L. Cavalli-Sforza. 1967. Phylogenetic Analysis: Models and Estimation Procedures. *Evolution* 21(3): 550-70.

Egetmeyer, Markus. 2010. *La dialecte grec ancien de Chypre*. Berlin/New York: De Gruyter.

[EIEC] Mallory, James P. and Douglas Q. Adams. 1997. *Encyclopedia of Indo-European Culture*. London/Chicago: Fitzroy Dearborn.

El Arifi, Samir. 2014. *Tachelhit de l'Atlas blidéen*. 4th edition. Algiers: Haut Commissariat à l'Amazighité.

Emery, Matthew. 2017. *Assessing Migration and Demographic Change in pre-Roman and Roman Period Southern Italy Using Whole-Mitochondrial DNA and Stable Isotope Analysis*. (Doctoral dissertation, McMaster University) https://macsphere.mcmaster.ca/bitstream/11375/22810/2/PhD_Dissertation_Emery_Final.pdf

Emery, Matthew, Ana T. Duggan, Tyler J. Murchie, Robert J. Stark, Jennifer Klunk, Jessica Hider, Katherine Eaton, Emil Karpinski, Henry P. Schwarcz, Hendrik N. Poinar, and Tracy L. Prowse. Ancient Roman mitochondrial genomes and isotopes reveal relationships and geographic origins at the local and pan-Mediterranean scales. *Journal of Archaeological Science: Reports*. 20: 200-9.

Emmerick, Ronald E. 1981. The Consonant Phonemes of Khotanese. In *Monumentum Georg Morgenstierne* (Acta Iranica 21-22), 185-209. Leiden: Brill.

[EM] Ernout, Alfred and Antoine Meillet. 2001. *Dictionnaire étymologique de la langue Latin*. 4th edition reprint. Paris: Klincksieck.

Ernault, Emile. 1895. *Glossaire Moyen-Breton*. Première Partie (A-G). Second Edition. Paris: Librairie E. Bouillon.

Erichsen, Wolja. 1954. *Demotisches Glossar*. Copenhagen: Ejnar Munksgaard.

Erman, Adolf and Hermann Grapow. 1971. *Wörterbuch der Aegyptischen Sprache*. Berlin: Akademi-Verlag.

Ernout, Alfred. 1909. *Les éléments dialectaux du vocabulaire latin*. Paris: H. Champion.

Ernout, Alfred. 1925. Remarques sur l'étymologie latine. *Revue des études latines* 3: 101-29.

Ernout, Alfred. 1946. *Philologica*. Paris: Klincksieck.⁵⁸⁰

Ernout, Alfred. 1954. *Aspects du vocabulaire latin*. Paris: Klincksieck.

Ernout, Alfred. 1965. *Philologica III*. Paris: Klincksieck.

[ESSJa] Trubachev, Oleg. 1974-present. *Etimologicheskij slovar' slavyanskikh yazykov*. Moscow: Nauka.

Ettmayer, Karl. 1926. Der Ortsname "Luzerne". *Indogermanische Forschungen* 43: 10-39.

[EWA] 1988-. *Etymologisches Wörterbuch des Althochdeutschen*. Göttingen:

⁵⁸⁰ Pages 21-51 "Les éléments étrusques du vocabulaire latin" are a reprint of the article from 1930, *Bulletin de la société de linguistique de Paris* 30: 82-124.

Vandenhoeck & Ruprecht.⁵⁸¹

[EWAia] Mayrhofer, Manfred. 1992-2001. *Etymologisches Wörterbuch des Altindoarischen*. Heidelberg: Carl Winter.

Fähnrich, Heinz. 2007. *Kartwelisches Etymologisches Wörterbuch*. Leiden/Boston: Brill.

Falk, Hjalmar and Alf Torp. 1960. *Norwegisch-dänisches etymologisches Wörterbuch*. 2. Aufl. Oslo/Bergen: Universitetsforlaget.

Falkenstein, Frank. 2012. The Development of Burial Rites from the Tumulus to the Urnfield Culture in Southern Central Europe. In Borgna, Elisabetta and Sylvie Müller Celka (eds.). *Ancestral Landscape. Burial mounds in the Copper and Bronze Ages (Central and Eastern Europe – Balkans – Adriatic – Aegean, 4th-2nd millennium B.C.)*, pp. 329-340. Lyon: Maison de l'Orient et de la Méditerranée.

Fasani, Leone. 2002. Età del Bronzo. In Aspes, Alessandra (ed.). *Preistoria Veronese: Contributi e Aggiornamenti*, pp. 107-153. Verona: Museo Civico di Storia Naturale.

Fay, Edwin. 1917. Dreams, the Swelling Moon, the Sun. *The Classical Quarterly* 11(4): 212-17.

Feist, Sigmund. 1910. Die germanische und die hochdeutsche Lautverschiebung sprachlich und ethnographisch betrachtet. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 36: 307-54.

Feist, Sigmund. 1932. The Origin of the Germanic Languages and the Europeanization of North Europe. *Language* 8: 245-54.

Fenwick, Rhona S.H. 2016. Descendants and Ancestry of a Proto-Indo-European Phytonym **meh₂l-*. *Journal of Indo-European Studies* 44(3/4): 441-65.

Fernandez, Daniel M., Alissa Mittnik, Iñigo Olalde, Iosif Lazaridis, Olivia Cheronet, Nadin Rohland, Swapan Mallick, Rebecca Bernardos, Nasreen Broomandkhoshbacht, Jens Carlsson, Brendan J. Culleton, Matthew Ferry, Beatriz Gamarra, Martina Lari, Matthew Mah, Megan Michel, Alessandra Modi, Mario Novak, Jonas Oppenheimer, Kendra A. Sirak, ...David Reich. 2020. The Spread of Steppe and Iranian-Related Ancestry in the Islands of the Western Mediterranean. *Nature Ecology & Evolution* 4: 334-45.

[FEW] Wartburg, Walther von. 1928- present. *Französisches Etymologisches Wörterbuch*. Bonn: Klopp.

Fick, August. 1890-1909. *Vergleichendes Wörterbuch der indogermanischen Sprachen*. 4. Aufl. Göttingen: Vandenhoeck & Ruprecht.

Fiesel, Eva and Paul-Max Groth. 1932. Etruskisch *tupi* und Lateinisch *tofus*. *Studi*

⁵⁸¹ In place of page numbers, column numbers are given.

Etruschi 6: 261-72.

Fiorito, Giovanni, Cornelia Di Gaetano, Simonetta Guarrera, Fabio Rosa, Marcus W. Feldman, Alberto Piazza, and Giuseppe Matullo. 2016. The Italian Genome Reflects the History of Europe and the Mediterranean Basin. *European Journal of Human Genetics* 24: 1056-62.

Finkelberg, Margalit. 2006. *Greeks and Pre-Greeks. Aegean Prehistory and Greek Heroic Tradition*. Cambridge: Cambridge University Press.

Fischl, Klára P., Viktória Kiss, Gabriella Kulcsár, and Vajk Szeverényi. 2013. Transformations in the Carpathian Basin around 1600 B.C. In von Harald, Meller, François Berteme, Hans-Rudolf Bork, and Roberto Risch (eds.) *1600 - Kultureller Umbruch im Schatten des Thera-Ausbruchs? / 1600 – Cultural Change in the shadow of the Thera-Eruption?*, pp. 255-371. Halle (Saale): Landesamt für Denkmalpflege und Archäologie Sachsen-Anhalt.

Flemestad, Peder and Birgit Anette Olsen. 2017. Sabellic Textile Terminology. In Gaspar, Salvatore, Cécile Michel, and Marie-Louise Nosch (eds.). *Textile Terminologies from the Orient to the Mediterranean and Europe 1000 BC – 1000 AD*, pp. 210-27. Lincoln, NE: University of Nebraska.

Fohalle, René. 1925. A propos de κυβερνᾶν *gubernāre*. In *Mélanges linguistiques offerts a M.J. Vendreyes par ses amis et ses élèves*, 157-78. Paris: Librairie Ancienne Édouard Champion.

Foy, W. 1896. Altindisch *kápr̥t* : griech. κάπρος usw. *Indogermanische Forschungen* 8: 295-297.

Fraenkel, Siegmund. 1886. *Die Aramäischen Fremdwörter im Arabischen*. Leiden: Brill.

Fraenkel, Ernst. 1910-12. *Geschichte der griechischen Nomina agentis auf -τήρ, -τωρ, -της (-τ-)*. Strassburg: Karl J. Trübner.

Fraenkel, Ernst. 1962-5. *Lithauisches etymologisches Wörterbuch*. Heidelberg: Winter.

Freu, Jacques. 1989. L'arrivée des Indo-Européens en Europe. *Bulletin de l'Association Guillaume Budé* 1: 3-41.

Friedrich, Paul. 1970. *Proto-Indo-European Trees*. Chicago/London: University of Chicago Press.

Frisk, Hjalmar. 1960-72. *Griechisches etymologisches Wörterbuch*. Heidelberg: Carl Winter.

Froehde, F. 1889. Zur lateinischen Lautlehre. *Beiträge zur Kunde der indogermanischen Sprachen* 4: 80-116.

Fruyt, Michèle. 1986. *Problèmes méthodologiques de dérivation à des suffixes latins en*

...cus. Paris: Klincksieck.

Fu, Qiamei, Cosimo Posth, Mateja Hajdinjak, Martin Petr, Swapan Mallick, Daniel Fernandes, Anja Furtwängler, Wolfgang Haak, Matthias Meyer, Alissa Mittnik, Birgit Nickel, Alexander Peltzer, Nadin Rohland, Viviane Slon, Sahra Talamo, Iosif Lazaridis, Mark Lipson, Iain Mathieson, Stephan Schiffels, Pontus Skoglund, ...David Reich 2016. The Genetic History of Ice Age Europe. *Nature* 534: 200-205.

Furnée, Edzard J. 1972. *Die wichtigsten konsonantischen Erscheinungen des Vorgriechischen*. The Hague/Paris: Mouton.

Furnée, Edzard J. 1979. *Vorgriechisch-Kartvelisches*. Leuven: Editions Peeters.

Gallay, Alain. 2001. L'énigme campaniforme. In Nicolis, Franco (ed.). *Bell Beakers Today. Pottery, People, Culture, Symbols in Prehistoric Europe*. Vol. 1, pp. 41-57. Trento: Ufficio Beni Archeologici.

Gamba, Cristina, Eppie R. Jones, Matthew D. Teasdale, Russell L. McLaughlin, Gloria Gonzalez-Fortes, Valeria Mattiangeli, László Domboróczki, Ivett Kóvári, Ildikó Pap, Alexandra Anders, Alasdair Whittle, János Dani, Pál Raczky, Thomas F.G. Higham, Michael Hofreiter, Daniel G. Bradley, and Ron Pinhasi. 2014. Genome Flux and Stasis in a Five Millennium Transect of European Prehistory. *Nature Communications* 5, 5257.

Gamillscheg, Ernst. 1920. Französische Etymologien. *Zeitschrift für Romanische Philologie* 40(5): 513-42.

Gamillscheg, Ernst. 1922. Zu Zs. 41, S. 583. *Zeitschrift für Romanische Philologie* 42(1): 86-9.

Gamkrelidze, Thomas V. and Vjačeslav V. Ivanov. 1973. Sprachtypologie und die Rekonstruktion der gemeinindogermanischen Verschlüsse. *Phonetica* 27(3): 150-6.

Gamkrelidze, Thomas V. and Vjačeslav V. Ivanov. 1995. *Indo-European and the Indo-Europeans*. (Transl. by Johanna Nichols). Berlin/New York: Mouton de Gruyter.

Ganieva, Faida A. 2002. *Xinalugsko-russkij slovar'* [Khinalug-Russian Dictionary]. Maxačkala: Rossijskaya Akademiya Nauk, Dagestanskij Nauchnyj Tsent.

Garnier, Romain. 2006. Nouvelles réflexions autour de gr. ψάμαθος. *Die Sprache* 46(1): 81-93.

Garnier, Romain. 2017a. *La dérivation inverse en latin*. Innsbruck: IBS.

Garnier, Romain. 2017b. Une loi phonétique méconnue en latin : la *lex hircus*. *Pallas* 103: 53-8.

Garnier, Romain. 2017c. Étymologie du gr. χλαῖνα: un mot louvite en grec? *Nouvelles Assyriologiques Brèves et Utilitaires* 2017(2): 90-2.

Garnier, Romain and Benoît Sagot. 2017. A shared substrate between Greek and Italic.

Indogermanische Forschungen 122(1): 29-60.

Garnier, Romain and Benoît Sagot. 2020. New results on a centum substratum in Greek: the Lydian connection. In Garnier, Romain (ed.). *Loanwords and Substrata: Proceedings of the Colloquium held in Limoges (5th-7th June, 2018)*, pp. 169-200. Innsbruck: IBS.

Georgiev, Vladimir. 1941. Vorgriechische Sprachwissenschaft. Erste Lieferung. *Annuaire de l'Université de Sofia. I. Faculté Historico-Philologique* 37(6): 1-162.

Georgiev, Vladimir. 1971 [1973]. Furnée, Edzard J.: Die wichtigsten konsonantischen Erscheinungen des Vorgriechischen. (Review). *Kratylos* 16: 164-67.

Georgiev, Vladimir. 1974. Drei Griechische Etymologien. *Philologus* 118(1): 272-3.

Gerola, B. 1942. Substrato mediterraneo e latino. *Studi Etruschi* 16: 345-68.

Gershevitch, Ilya. 1969. *Amber at Persepolis*. Rome: Bardi.

Ghirotto, Silvia, Francesca Tassi, Erica Fumagalli, Vincenza Colonna, Anna Sandionigi, Martina Lari, Stefania Vai, Emmanuele Petiti, Giorgio Corti, Ermanno Rizzi, Gianluca De Bellis, David Caramelli, and Guido Barbujani. 2013. Origins and Evolution of the Etruscans' mtDNA. *PLoS ONE* 8(2): e55519.

Giacomelli, Roberto. 1994. Lingua quotidiana e gecismo nel lessico culinario latino. *ACME* 47(2)⁵⁸²: 29-49.

Giardino, Claudio. 2005. Metallurgy in Italy between the Late Bronze Age and the Early Iron Age: The Coming of Iron. In Attema, Peter, Albert Nijboer, and Andrea Zifferero (eds.). *Papers in Italian archaeology VI: Communities and settlements from the Neolithic to the Early Medieval period*, pp. 491-505. Oxford: Archaeopress.

Giari, A., Ranieri Domenici, and Marta Bargagna. 1986. EsD*5 gene frequency in Tuscany (Italy). *Zeitschrift für Rechtsmedizin* 97: 1-6.

Gil Fernández, Luis. 1959. *Nombres de insectos en griego antiguo*. Madrid: Antonio de Nebrija.

Gimbutas, Marija. 1956. *The Prehistory of Eastern Europe, Part 1: Mesolithic, Neolithic and Copper Age Cultures in Russia and the Baltic Area*. Cambridge, MA: Peabody Museum.

Gimbutas, Marija. 1963. The Indo-Europeans: Archaeological Problems. *American Anthropologist* 65(4): 815-36.

Gimbutas, Marija. 1965. *Bronze Age Cultures in Central and Eastern Europe*. The Hague: Mouton & Co.

Gimbutas, Marija. 1989. *The Language of the Goddess: Unearthing the Hidden Symbols*

⁵⁸² Mislabeled as Volume 48.

of Western Civilization. San Francisco: Harper & Row.

Gliwa, Bernd. 2008. Ist Lit. *túopa* 'Pappel' wirklich ein Erbwort? *Baltistica* 43(2): 239-44.

Goldmann, . 1938. B. Sez. II - Lingua, Epigrafia. Recensioni. In Recensioni e repertorio bibliografico. *Studi Etruschi* 12: 393-419.

Gómez-Carballa, Alberto, Jacobo Pardo-Seco, Jorge Amigo, Federico Martín-Torres, and Antonio Salas. 2015. Mitogenomes from The 1000 Genomes Project Reveal New Near Eastern Features in Present-Day Tuscans. *PLoS ONE* 10(3): e0119242.

Gorton, Luke. 2017. Revisiting Indo-European "Wine". *Journal of Indo-European Studies* 45(1/2): 1-26.

Grassmann, Hermann. 1863. Ueber die Aspiraten und ihr gleichzeitiges Vorhandensein im An- und Auslaute der Wurzeln. *KZ* 12(2): 81-110.

Gray, Louis H. 1906. Contributions to Old Italic Etymology. *Beiträge zur Kunde der indogermanischen Sprachen* 27: 297-310.

Green, Dennis Howard. 1998. *Language and History in the Early Germanic World*. Cambridge: Cambridge University Press.

Greppin, John A. C. 1978. *Classical and Middle Armenian Bird Names*. New York: Caravan Books.

Greppin, John A. C. 1981. Reduplicated Patterns in Classical Armenian. *Annual of Armenian Linguistics* 2: 1-14.

Greppin, John A. C. 1998. Arm. գինի *gini*, Grg. ვინო *vinno* 'wine'. *Annual of Armenian Linguistics* 19: 65-9.

Grugni, Viola, Alessandro Raveane, Francesca Mattioli, Vincenza Battaglia, Cinza Sala, Daniela Toniolo, Luca Ferretti, Rita Gardella, Alessandro Achilli, Anna Olivieri, Antonio Torroni, Giuseppe Passarino, and Ornella Semino. 2018. Reconstructing the Genetic History of Italians: New Insights from a Male (Y-chromosome) perspective. *Annals of Human Biology* 45(1): 44-56.

Guggenheimer, Heinrich. 1998. *The Scholar's Haggadah*. London/Northvale: Jason Aronson Inc.

Guidi, Alessandro. 2010. The Historical Development of Italian Prehistoric Archaeology: A Brief Outline. *Bulletin of the History of Archaeology* 20(2): 13-21.

Güntert, Hermann. 1914. *Über Reimwortbildungen im Arischen und Altgriechischen*. Heidelberg: Carl Winter.

Güntert, Hermann. 1933. Labyrinth. Eine sprachwissenschaftliche Untersuchung. *Sitzungsberichte der Heidelberger Akademie der Wissenschaften*.

Philosophisch-historische Klasse 23: 3-49.

Güntert, Hermann. 1934. *Der Ursprung der Germanen*. Heidelberg: Carl Winters.

Günther, Torsten, Cristina Valdiosera, Helena Malmström, Irene Ureña, Ricardo Rodriguez-Varela, Óddny Osk Sverrisdóttir, Evangelia A. Daskalaki, Pontus Skoglund, Thijessen Naidoo, Emma M. Svensson, José M. Bermúdez de Castro, Eudald Carbonell, Michael Dunn, Jan Storå, Eneko Iriarte, Juan Luis Arsuaga, José-Miguel Carretero, Anders Götherström, and Mattias Jakobsson. 2015. Ancient Genomes Link Early Farmers from Atapuerca in Spain to Modern-Day Basques. *PNAS* 112(38): 11917-22.

Gysseling, Maurits. 1975. Enkele Belgische leenwoorden in de toponymie. *Naamkunde* 7: 1-6.

Haak, Wolfgang, Peter Forester, Barbara Bramanti, Shuichi Matsumura, Guido Brandt, Marc Tänzler, Richard Villems, Colin Renfrew, Detlef Gronenborn, Kurt W. Alt, and Joachim Burger. 2005. Ancient DNA from the First European Farmers in 7500-Year-Old Neolithic Sites. *Science* 310(5750): 1016-18.

Haak, Wolfgang, Guido Brandt, Hylke N. de Jong, Christian Meyer, Robert Ganslmeier, Volker Heyd, Chris Hawkesworth, Alistair W.G. Pike, Harald Meller, and Kurt W. Alt. 2008. Ancient DNA, Strontium Isotopes, and Osteological Analyses Shed Light on Social and Kinship Organization of the Later Stone Age. *PNAS* 105(47): 18226-31.

Haak, Wolfgang, Oleg Balanovsky, Juan J. Sanchez, Sergey Koshel, Valery Zaporozhchenko, Christina J. Adler, Clio S.I. Der Sarkissian, Guido Brandt, Carolin Schwarz, Nicole Nicklisch, Veit Dresely, Barbara Fritsch, Elena Balanovska, Richard Villems, Harald Meller, Kurt W. Alt, Alan Cooper, and the Genographic Consortium. 2010. Ancient DNA from European Early Neolithic Farmers Reveals their Near Eastern Affinities. *PLoS Biology* 8(11): e1000536.

Haak, Wolfgang, Iosif Lazaridis, Nick Patterson, Nadin Rohland, Swapan Mallick, Bastien Llamas, Guido Brandt, Susanne Nordenfelt, Eadaoin Harney, Kristin Stewardson, Qiomei Fu, Alissa Mittnik, Eszter Bánffy, Christos Economou, Michael Francken, Susanne Friederich, Rafael Garrido Pena, Fredrik Hallgren, Valery Khartanovich, Aleksandr Khokhlov, ...David Reich. 2015. Massive Migration from the Steppe was a Source for Indo-European Languages in Europe. *Nature* 522: 207-11.

Haas, Otto. 1959. Die Lehre von den indogermanischen Substraten in Griechenland. *Balkansko Ezikoznanie [Linguistique Balkanique]* 1: 29-56.

Haas, Otto. 1960. *Das Frühitalische Element: Versuch über die Sprache der ersten Indogermanen Italiens*. Vienna: Notring.

Haas, Otto. 1962. *Messapische Studien*. Heidelberg: Carl Winter.

Hamp, Eric. 1957. Albanian and Messapic. In Pulgram, Ernst (ed.). *Studies Presented to Joshua Whatmough on his Sixtieth Birthday*, pp. 73-89. The Hague: Mouton & Co.

- Hamp, Eric. 1973. Mōrum. *The American Journal of Philology* 94(2): 167-9.
- Hamp, Eric. 1974. 'Ivy' in Italic and Celtic. *Journal of Indo-European Studies* 2(1): 86-93.
- Hamp, Eric. 1975. 'Cut' and 'Meat' in Germanic. *Acta Philologica Scandinavica* 30: 49-51.
- Hamp, Eric. 1978. Sound change and the etymological lexicon. In Farkas, Donka, Wesley Jacobsen, and Karol Todrys (ed.). *Papers from the Parasession on the Lexicon, Chicago Linguistic Society, April 14-15, 1978*, pp. 184-95. Chicago: Chicago Linguistic Society.
- Hamp, Eric. 1979. The North European word for 'apple'. *Zeitschrift für Celtische Philologie* 37: 158-66.
- Hamp, Eric. 1981. Refining Indo-European Lexical Entries. *KZ* 95(1): 81-3.
- Hamp, Eric. 1983. A revised amendment to Fortunatov's Law. *Indo-Iranian Journal* 25(4): 275-6.
- Hamp, Eric. 1984. Varia I: I. Irish *fert, fertae*. *Zeitschrift für celtische Philologie* 40(1): 275-9.
- Hamp, Eric. 1985. Latin *callum, callis*, Indic *kiṇa-*. *KZ* 98(1): 59.
- Hamp, Eric. 1990. The Pre-Indo-European Language of Northern (Central) Europe. In Markey, Thomas L. and John A. C. Greppin (eds.). *When Worlds Collide: Indo-European and the Pre-Indo-Europeans. The Bellagio Papers*, pp. 291-310. Ann Arbor: Karoma.
- Hamp, Eric. 1992. focus. *Glotta* 70(1/2): 82.
- Hamp, Eric. 2000. Greek *ῥῶξ ῥῥῶγός*, Albanian *rrush*, and *Ragusa*. *Ponto-Baltica* 10: 7-11.
- Harrison, Richard, and Volker Heyd. 2007. The Transformation of Europe in the Third Millennium BC: the example of 'Le Petit-Chasseur I + III' (Sion, Valais, Switzerland). *Praehistorische Zeitschrift* 82(2): 129-214.
- Haupt, Paul. 1915. Lat. *asinus* und semit. *atân*, Eselin. *Orientalistische Literaturzeitung* 7: 203-5.
- Havet, Louis. 1911. *Manuel de critique verbale appliquée aux textes latins*. Paris: Librairie Hachette et C^{ie}.
- Hehn, Victor and Otto Schrader (ed.). 1911. *Kulturpflanzen und Haustiere in ihrem Übergang aus Asien nach Griechenland und Italien sowie in das übrige Europa*. 8th edition. Berlin: Gebrüder Borntraeger.

- Helbig, Wolfgang. 1879. *Die Italiker in der Po-Ebene*. Leipzig: Breitkopf.
- Hellquist, Elof. 1922. *Svensk etymologisk Ordbok*. Lund: C. W. K. Gleerups.
- Hemmerdinger, Bertrand. 1968. Noms communs grecs d'origine égyptienne. *Glotta* 46(3/4): 238-47.
- Henning, Walter Bruno. 1951. *Zoroaster : Politician or Witch-Doctor?* London: Oxford University Press.
- Henry, Victor. 1900. *Lexique étymologique des termes les plus usuels de Breton moderne*. Rennes: J. Plihon & L. Hervé.
- Herbig, Gustav. 1917. Etruskisches Latein. *Indogermanische Forschungen* 37: 163-87.
- Hester, D.A. 1965. "Pelasgian" - A New Indo-European Language? *Lingua* 13: 335-84.
- Heywood, Vernon H., Richard K. Brummit, Alastair Culham, and Ole Seberg. 2007. *Flowering Plant Families of the World*. Buffalo, NY: Firefly Books Ltd.
- Hiersche, Rolf. 1964. *Untersuchungen zur Frage der Tenues Aspiratae im Indogermanischen*. Wiesbaden: Otto Harrassowitz.
- Hill, Eugen. 2003. *Untersuchungen zum inneren Sandhi des Indogermanischen* (Münchener Forschungen zur historischen Sprachwissenschaft 1). Bremen: Hempen Verlag.
- Hinz, Walther. 1973. *Neue Wege im Altpersischen*. Wiesbaden: Otto Harrassowitz.
- Hinz, Walther and Heidemarie Koch. 1987. *Elamisches Wörterbuch*. Berlin: Dietrich Reimer.
- Hirt, Hermann. 1892. Die Verwandtschaftsverhältnisse der Indogermanen. *Indogermanische Forschungen* 4: 36-45.
- Hirt, Hermann. 1907. Miscellen. *Indogermanische Forschungen* 21: 162-74.
- Hirt, Hermann. 1907/8. Untersuchungen zur indogermanischen Altertumskunde. *Indogermanische Forschungen* 22: 55-95.
- Hirt, Herman. 1909. *Etymologie der neuhochdeutschen Sprache: Darstellung des deutschen Wortschatzes in seiner geschichtlichen Entwicklung*. Munich: C.H. Beck.
- Hirt, Hermann. 1927. *Indogermanische Grammatik*. Teil I. Heidelberg: Carl Winter.
- Hoffler, Jr., Harry A. 1967. An English-Hittite Glossary. *Revue Hittite et Asiatique* Tome 25 (Fascicule 80). Paris: Klincksieck.
- Höfler, Stefan. 2015. Dark Matter. The Root **k_uel* 'Dark, Black'. *Indo-European Linguistics* 3: 24-41.
- Höfler, Stefan. 2017. Observations on the *plama* rule. *PALLAS* 103: 15-23.

Höfler, Stefan and Johan Ulrik Nielsen. 2022. A Proto-Indo-European word for ‘spider’? Un-weaving the prehistory of the Greek ἀράχνη and the Latin *arāneus*. *Greco-Latina Brunensia* 27: 69-89.

Hofmanová, Zuzana, Susanne Kreutzer, Garrett Hellenthal, Christian Sell, Yoan Diekmann, David Díez-Del-Molino, Lucy van Dorp, Saioa Lópe, Athanasios Kousathana, Vivian Link, Karola Kirsanow, Lara M. Cassidy, Rui Martiniano, Melanie Strobel, Amelie Scheu, Kostas Kotsakis, Paul Halstead, Sevi Triantaphyllou, Nina Kyparissi-Apostolika, Dushka Urem-Kotsou, ...Joachim Burger. 2016. Early Farmers from across Europe Directly Descended from Neolithic Aegeans. *PNAS* 113(25): 8668-91.

Holder, Alfred. 1896-1904. *Alt-Celtischer Sprachschatz*. Leipzig: Teubner.

Holmes Jr., Urban T. 1947. The Suffix -“erna” in Latin and Romance. *Romance Philology* 1(2): 105-11.

Holopainen, Sampsa. 2019. *Indo-Iranian Borrowings in Uralic*. (Doctoral dissertation, University of Helsinki.) <https://helda.helsinki.fi/handle/10138/307582>

Holton, David, Geoffrey Horrocks, Marjolijne Janssen, Tina Lendari, Io Manolessou, and Notis Toufexis. 2020. *The Cambridge Grammar of Medieval and Early Modern Greek*. Cambridge: Cambridge University Press.

Holton Pierce, Richard. 1971. Egyptian loan-words in Ancient Greek. *Symbolae Osloenses* 46(1): 69-107.

Holzer, Georg. 1989. *Entlehnungen au seiner bisher unbekanntem indogermanischen Sprache im Urslavischen und Urbaltischen*. Vienna: Verlag der österreichischen Akademie der Wissenschaften.

Hommel, Fritz. 1881. *Allgemeine Zeitung* 231 (August 19, 1881).

Hopper, Paul. 1973. Glottalized and murmured occlusives in Indo-European. *Glossa* 7(2): 141-66.

Hubschmid, Johannes.⁵⁸³ 1943. Bezeichnungen für ‘Kaninchen’ - ‘Höhle’ - ‘Steinplatte’. In *Sache Ort und Wort. Jakob Jud zum sechzigsten Geburtstag 12. Januar 1942*, pp. 246-80. Zürich/Erlenbach: Eugen Rentsch.

Hubschmid, Johannes. 1950a. Vorindogermanische und jüngere Wortschichten in den romanischen Mundarten der Ostalpen. *Zeitschrift für romanische Philologie* 66(1/2): 1-94.

Hubschmid, Johannes. 1950b. Oscar Bloch et W.v. Warburg, *Dictionnaire étymologique de la langue française* (review). *Vox Romanica* 11: 288-292.

⁵⁸³ Given as “J. Hubschmied jun.”, though he would later come to spell his surname “Hubschmid”

Hubschmid, Johannes. 1953. Sardische Studien. *Romanica Helvetica*, Vol. XLI. Bern: A. Francke.

Hubschmid, Johannes. 1954-5. Praeindogermanica. 3. Altprovenzalisch *cadarauc* 'ruisseau des rues' und das vorindogermanische Suffix *-ouko-*. *Romance Philology* 8: 12-26.

Hubschmid, Johannes. 1955. Hispano-ägäische Pflanzennamen. *Zeitschrift für Romanische Philologie* 71(3/4): 236-48.

Hubschmid, Johannes. 1958. Kritische Bemerkungen zum mediterranen Substrat Sardiniens und Hispaniens. *Zeitschrift für Romanische Philologie* 74(3/4): 212-33.

Hubschmid, Johannes. 1960a. *Mediterrane Substrate*. Bern: A. Francke.

Hubschmid, Johannes. 1960b. Substratprobleme. Eine neue iberoromanisch-alpinlombardische Wortgleichung vorindogermanischen Ursprungs und die vorindogermanischen Suffixe *-ano-* und *-s(s)-*. *Vox Romanica* 19: 124-179.

Hubschmid, Johannes. 1963-5. *Thesaurus Praeromanicus*. Bern: Francke.

Hubschmid, Johann Ulrich. 1942. Rätoromanisch. *Zeitschrift für Romanische Philologie* 62: 107-28.

Huld, Martin E. 1981. Albanian corrigenda to Friedrich's 'Proto-Indo-European Trees'. *KZ* 95(2): 302-308.

Huld, Martin E. 1983/4. *Basic Albanian Etymologies*. Los Angeles: Slavica Publishers.

Huld, Martin E. 1990. The Linguistic Typology of the Old European Substrata in North Central Europe. *JIES* 18(3/4): 389-423.

Huld, Martin E. 2012. Some observations on the development of Indo-European metallurgy. In: M.E. Huld, K. Jones-Bley & D. Miller (eds.), *Archaeology and language: Indo-European studies presented to James P. Mallory* (JIES Monograph No. 60), pp. 281–356. Washington D.C.: Institute for the Study of Man.

Hyllested, Adam and Brian D. Joseph. 2022. Albanian. In Olander, Thomas (ed.). *The Indo-European Language Family. A Phylogenetic Perspective*, pp. 223-45. Cambridge: Cambridge University Press.

Iacono, Francesco, Elisabetta Borgna, Maurizio Cattani, Claudio Cavazzuti, Helen Dawson, Yannis Galanakis, Maja Gori, Cristiano Iaia, Nicola Ialongo, Thibault Lachenal, Alberto Lorrio, Rafael Micó, Barry Molloy, Argyro Nafplioti, Kewin Peche-Quilichini, Christina Rihuete Herrada, and Roberto Risch. 2021. Establishing the Middle Sea: The Late Bronze Age of Mediterranean Europe (1700-900 BC). *Journal of Archaeological Research* 30: 371-445.

Imberciadori, Giulio. The Etymology of Toch. B *tapre*, A *tpär* 'high'. Paper presented at

the Symposium *Hansakansan kanssa*, 16th - 18th May 2022, University of Greifswald, Germany.

Iversen, Rune and Guus Kroonen. 2017. Talking Neolithic: Linguistic and Archaeological Perspectives on how Indo-European was Implemented in Southern Scandinavia. *American Journal of Archaeology* 121(4): 511-25.

Jagersma, Abraham H. 2010. *A descriptive grammar of Sumerian*. (Doctoral dissertation, University of Leiden.)
<https://scholarlypublications.universiteitleiden.nl/handle/1887/16107>

Jahowkyan, Gevorg B. 1987. *Hayoc' lezvi patmowt'yown: naxagrayin žamanakašrjan* [History of the Armenian Language: the Period Preceding Writing]. Yerevan: Haykakan SSH GA Hratarakč'owt'yown.

Jahowkyan, Gevorg B. 2010. *Hayeren stugabanakan bařaran* [Armenian Etymological Dictionary]. Yerevan: Asoghik.

Jakob, Anthony. fthc. Three pre-Balto-Slavic bird names, or A more austere take on Oštir. In Kroonen, Guus (ed.). *Sub-Indo-European Europe: Continuity and discontinuity in the linguistic landscape of prehistoric Europe*. Berlin/Boston: De Gruyter Mouton.

Janda, Michael. 1999. Ἐβοσίχθων „Erdschütterer“. In Heiner Eichner and Hans Christian Luschützky (eds.). *Compositiones Indogermanicae in Memoriam Jochem Schindler*, pp. 183-203. Prague: Enigma.

Janick, Jules, Harry S. Paris, and David C. Parrish. 2007. The Cucurbits of Mediterranean Antiquity: Identification of Taxa from Ancient Images and Descriptions. *Annals of Botany* 100(7): 1441-57.

Johansson, Karl Ferdinand. 1890a. Etymologische beiträge. *KZ* 30(1/3): 346-51.

Johansson, Karl Ferdinand. 1890b. Etymologische beiträge. *KZ* 30(4): 428-52.

Johansson, Karl Ferdinand. 1902. Arische Beiträge. *Indogermanische Forschungen* 14: 265-339.

Johansson, Karl Ferdinand. 1906. Arische Beiträge. *Indogermanische Forschungen* 19: 112-39.

Johnston, Richard F. and Marian Janiga. 1995. *Feral Pigeons*. Oxford: Oxford University Press.

Jones, D.M. 1950. The Relation of Latin to Osco-Umbrian. *Transactions of the Philological Society*: 60-87.

Jones, Eppie R., Gloria Gonzalez-Fortes, Sarah Connell, Veronika Siska, Anders Eriksson, Rui Martiniano, Russell L. McLaughlin, Marcos Gallego Llorente, Lara M. Cassidy, Cristina Gamba, Tengiz Meshveliani, Ofer Bar-Yosef, Werner Müller, Anna

Belfer-Cohen, Zinovi Matskevich, Nino Jakeli, Thomas F.G. Higham, Mathias Currat, David Lordkipanidze, Michael Hofreiter, Andrea Manica, Ron Pinhasi, and Daniel G. Bradley. 2015. Upper Palaeolithic Genomes Reveal Deep Roots of Modern Eurasians. *Nature Communications* 6, 8912.

Juret, A.-C. 1918. Questions de phonétique du latine II. *Mémoires de la Société de Linguistique de Paris* 20: 190-209.

Kahane, Henry and Renée Kahane. 1960. Animalia Pyrricha. *Glotta* 39(1/2) : 133-45.

Kallio, Petri. 2003. Languages in the Prehistoric Baltic Sea Region. In Bammesberger, Alfred and Theo Vennemann (eds.). *Languages in Prehistoric Europe*, pp. 227-244. Heidelberg: Universitätsverlag Winter.

Karulis, Konstantīns. 1992. *Latviešu etimologijas vārdnīca*. Vol. II. Riga: Avots.

Kaschnitz-Weinberg, Guido. 1954. Italien mit Sardinien, Sizilien und Malta. In *Handbuch der Archäologie*. Vol. 2, pp. 396-402. Munich: C.H. Beck.

Katičić, Radoslav. 1976. *Ancient Languages of the Balkans*. Part 1. The Hague/Paris: Mouton.

Katz, Joshua T. 1998. How to be a Dragon in Indo-European: Hittite *illuyankaš* and its Linguistic and Cultural Congeners in Latin, Geek, and Germanic. In Jasanoff, Jay, H. Craig Melchert, and Lisi Oliver (eds.). *Mír Curad: Studies in Honor of Calvert Watkins*, pp. 317-34. Innsbruck: IBS.

Keller, Otto. 1891. *Lateinische Volksetymologie und Verwandtes*. Leipzig: Teubner.

Keller, Andreas, Angela Graefen, Markus Ball, Mark Matzas, Valesca Boisguerin, Frank Maixner, Petra Leidinger, Christina Backes, Rabab Khairat, Michael Forster, Björn Stade, Andre Franke, Jens Mayer, Jessica Spangler, Stephen McLaughlin, Minita Shah, Clarence Lee, Timothy T. Harkins, Alexander Sartori, Andres Moreno-Estrada, ...Albert Zink. 2012. New Insights into the Tyrolean Iceman's Origin and Phenotype as Inferred by Whole-Genome Sequencing. *Nature Communications* 3, 698.

[Keil GL] Keil, Heinrich. 1855-80. *Grammatici Latini*. Leipzig: Teubner.

[KEWA] Mayrhofer, Manfred. 1956-1980. *Kurzgefaßtes etymologisches Wörterbuch des Altindischen*. Heidelberg: Carl Winter.

Keyser, Christine, Caroline Bouakaze, Eric Crubézy, Valery G. Nikolaev, Daniel Montagnon, Tatiana Reis, and Bertrand Ludes. 2009. Ancient DNA Provides New Insights into the History of South Siberian Kurgan People. *Human Genetics* 126: 395-410.

Kibrik, Aleksandr E. and Sandro V. Kodzasov. 1990. *Sopostavitel'noe izučenie dagestanskix jazykov. Imja. Fonetika* [Comparative Study of Dagestani Languages: The Noun. Phonetics]. Moscow: Izdatel'stvo moskovskogo universiteta.

Klein, Ernest. 1987. *A Comprehensive Etymological Dictionary of the Hebrew Language for Readers of English*. Jerusalem: Carta.

Klimov, Georgiy A. 1998. *Etymological Dictionary of the Kartvelian Languages*. Berlin/New York: Mouton de Gruyter.

Klingenschmidt, Gert. 1982. *Das altarmenische Verbum*. Wiesbaden: Reichert.

Kloekhorst, Alwin. 2007. *Etymological Dictionary of the Hittite Inherited Lexicon*. Leiden: Brill.

Kloekhorst, Alwin. 2022. Luwians, Lydians, Etruscans, and Troy. The Linguistic Landscape of Northwestern Anatolia in the Pre-Classical Period. In Hajnal, Ivo, Eberhard Zangger, and Jorrit Kelder (eds.). *The Political Geography of Western Anatolia in the Late Bronze Age. Proceedings of the EAA Conference Bern, 7 September 2019*, pp. 201-27. Budapest: Archaeolingua.

Kluge, Friedrich. 1885. Sprachhistorische Miscellen. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 10(2): 439-45.

Kluge, Friedrich and Elmar Seebold. 1989. *Etymologisches Wörterbuch der deutschen Sprache*. Berlin/New York: Walter de Gruyter.

Knobloch, Johann. 1955. Holunder im Hethitischen? *PHMA* 1: 5-10.

Kobyliński, Zbigniew. 2005. The Slavs. In Fouracre, Paul (ed.). *The New Cambridge Medieval History*, Vol. 1, pp. 524-44. Cambridge: Cambridge University Press.

Koch, John T. 2020. *Celto-Germanic: later prehistory and Post-Proto-Indo-European vocabulary in the north and west*. Aberystwyth: Canolfan Uwchefrydiau Cymreig a Cheltaidd Prifysgol Cymru.

Kock, Axel. 1909. *Svensk Ljudhistoria*. Part 2. Lund/Leipzig: C.W.K. Gleerund/Harrassowitz.

Kogan, Leonid. 2011. Proto-Semitic Lexicon. In Weninger, Stefan (ed.). *The Semitic Languages. An International Handbook*, pp. 179-258. Berlin/Boston: de Gruyter Mouton.

Kölligan, Daniel. 2017. Gr. *σάλαξ, σκούραξ* und die griechischen Nomina auf *-ak-*. In Hajnal, Ivo, Daniel Kölligan, and Katharina Zipser (eds.). *Miscellanea Indogermanica. Festschrift für José Luis García Ramón zum 65. Geburtstag*, pp. 369-82. Innsbruck: IBS.

Körting, Gustav. 1908. *Etymologisches Wörterbuch der Französischen Sprache*. Paderborn: Ferdinand Schöningh.

Kortlandt, Frederik. 1981. More Evidence for Italo-Celtic. *Ériu* 32: 1-22.

Kortlandt, Frederik. 1987. PIE. **s* in Albanian. In Barentsen, A., B. Groen, and R. Sprenger (eds.). *Dutch studies in South Slavic and Balkan linguistics*, pp. 219-26.

Amsterdam: Rodopi.

Kortlandt, Frederik. 1989. Lachmann's Law. In Vennemann, Theo (ed.). *The New Sound of Indo-European*, pp. 103-5. Berlin/New York: De Gruyter Mouton.

Kortlandt, Frederik. 1992. The Germanic Fifth Class of Strong Verbs. *North-Western European Language Evolution* 19(1): 101-7.

Kortlandt, Frederik. 2007. *Italo-Celtic Origins and Prehistoric Development of the Irish Language*. Amsterdam: Brill.

Kortlandt, Frederik. 2010. *Studies in Germanic, Indo-European and Indo-Uralic*. Amsterdam/New York: Rodopi.

Krahe, Hans. 1937. Zu CIL. I² 379. *Indogermanische Forschungen* 55: 121-2.

Krahe, Hans. 1955. *Die Sprache der Illyrer. Erster Teil: die Quellen*. Wiesbaden: Otto Harrassowitz.

Krahe, Hans. 1964. *Unsere ältesten Flussnamen*. Wiesbaden: Otto Harrassowitz.

Kretschmer, Paul. 1896. *Einleitung in die Geschichte der griechischen Sprache*. Göttingen: Vandenhoeck und Ruprecht.

Kretschmer, Paul. 1905. Die inschriften von Ornavasso und die ligurische sprache. *KZ* 38(1): 97-128.

Kretschmer, Paul. 1912. Literaturbericht für das Jahr 1909. *Glotta* 3(4): 296-343.

Kretschmer, Paul. 1921. Pelasger und Etrusker. *Glotta* 11(3/4): 276-85.

Kretschmer, Paul. 1923a. *Einleitung in die Altertumswissenschaft*. Band 1, Heft 6 Sprache. Leipzig/Berlin: B.G. Teubner.

Kretschmer, Paul. 1923b. Messapische Göttinnen. *Glotta* 12(3/4): 278-283.

Kretschmer, Paul. 1925a. Das *nt*-Suffix. *Glotta* 14(1/2): 84-106.

Kretschmer, Paul. 1925b. Die protindogermanische Schicht. *Glotta* 14(3/4): 300-19.

Kretschmer, Paul. 1928. [Ernst Maass, Eunochos und Verwandtes, Rhein. Mus. 74, 432-476] Literaturbericht für das Jahr 1925. *Glotta* 16(3/4): 190-2.

Kretschmer, Paul. 1931. Fiesel, Eva: Namen des griechischen Mythos im Etruskischen. In Literaturbericht für das Jahr 1928. *Glotta* 19(3/4): 216-17.

Kretschmer, Paul. 1932a. [A. Walde, Vergleichendes Wörterbuch der indogermanischen Sprachen I. 3-4 Lief., Berlin, 1929, 339-642] Literaturbericht für das Jahr 1929. *Glotta* 20(3/4): 247-8.

Kretschmer, Paul. 1932b. Die Herkunft der Umbrier. *Glotta* 21(1/2): 112-25.

Kretschmer, Paul. 1933a. Zur ältesten Sprachgeschichte Kleinasiens. *Glotta* 21(1/2): 76-100.

Kretschmer, Paul. 1933b. Literaturbericht für das Jahr 1930. *Glotta* 21(3/4): 151-213.

Kretschmer, Paul. 1934. Nordische Lehnwörter im Altgriechischen. *Glotta* 22(1/2): 100-22.

Kretschmer, Paul. 1940. Die vorgriechischen Sprach- und Volksschichten. *Glotta* 28 (3/4): 231-78.

Kretschmer, Paul. 1943. Die vorgriechischen Sprach- und Volksschichten (Fortsetzung von Gl. XXVIII 231–278). *Glotta* 30: 84-218.

Kristiansen, Kristian. 1998. *Europe Before History*. Cambridge: Cambridge University Press.

Kristiansen, Kristian. 2018. The Rise of Bronze Age Peripheries and the Expansion of International Trade 1950–1100 BC. In Kristiansen, Kristian, Thomas Lindkvist, and Janken Myrdal (eds.). *Trade and Civilisation: Economic Networks and Cultural Ties, from Prehistory to the Early Modern Era*, pp. 87-112. Cambridge: Cambridge University Press.

Kristiansen, Kristian and Thomas B. Larsson. 2005. *The Rise of Bronze Age Society. Travels, Transmission and Transformations*. Cambridge: Cambridge University Press.

Krogmann, Willy. 1937. Lat. *ferrum*. *KZ* 64(3/4): 267–269.

Krogmann, Willy. 1938. Der Name der ewigen Stadt (I.). *Wörter und Sachen* 19: 115-34.

Kroonen, Guus. 2009. *Consonant and vowel gradation in the Proto-Germanic n-stems*. (Doctoral dissertation, University of Leiden.) <https://scholarlypublications.universiteitleiden.nl/handle/1887/14513>

Kroonen, Guus. 2011. *The Proto-Germanic n-stems*. Amsterdam/New York: Rodopi.

Kroonen, Guus. 2012a. Non-Indo-European root nouns in Germanic: evidence in support of the Agricultural Substrate Hypothesis. In Grünthal, R. and P Kallio (eds.). *A linguistic map of prehistoric Northern Europe*, 239-60. Helsinki: Suomalais-Ugrilainen Seura.

Kroonen, Guus. 2012b. On the etymology of Greek ἄλγλις and γέλιγλις ‘garlic’: an Akkadian loanword in Pre-Greek. *JIES* 40(3/4): 289-99.

Kroonen, Guus. 2013. *Etymological Dictionary of Proto-Germanic*. Leiden: Brill.

Kroonen, Guus. 2017. The development of the Proto-Indo-European instrumental suffix in Germanic. *Indogermanische Forschungen* 122: 105-10.

Kroonen, Guus. fthc. For the nth time: the Pre-Greek vθ-suffix Revisited. In Kroonen, Guus (ed.). *Sub-Indo-European Europe: Continuity and discontinuity in the linguistic*

landscape of prehistoric Europe. Berlin/Boston: De Gruyter Mouton.

Kroonen, Guus, Andrew Wigman, and Rasmus Thorsø. 2021. Proto-Indo-European **sneig*^{wh}- ‘to fall down; to snow’. *KZ* 134(1): 214-24.

Kroonen, Guus, Anthony Jakob, Axel Palmér, Paulus van Sluis, and Andrew Wigman. 2022. Indo-European Cereal Terminology suggests a Northwest Pontic homeland for the core Indo-European languages. *PLoS ONE* 17(10): e0275744.

Kuhn, Hans. 1959. Vor- und frühgermanische Ortsnamen in Norddeutschland und den Niederlanden. *Westfälische Forschungen* 12: 5-44.

Kuhn, Hans, Rolf Hachmann, and Georg Kossack. 1962. *Völker zwischen Germanen und Kelten. Schriftquellen, Bodenfunde und Namengute zur Geschichte des nördlichen Westdeutschlands um Christi Geburt*. Neumünster: Karl Wachholz.

Kuiper, Franciscus B.J. 1948. *Proto-Munda Words in Sanskrit*. Amsterdam: Noord-Hollandsche Uitg. Mij.

Kuiper, Franciscus B.J. 1956. The Etymology of ἄνθρωπος. *Μνήμη χαρίτιν* 1: 211-26.

Kuiper, Franciscus B.J. 1991. *Aryans in the Rigveda*. Amsterdam/Atlanta: Rodopi.

Kuiper, Franciscus B.J. 1995. Gothic *bagms* and Old Icelandic *ylgr**. *North-Western Language Evolution* 25: 63-88.

Kurdadze, Ramaz, Dodo Shonia, Lile Tandilava, and Lela Nizharadze. 2015. *Georgian-Megrelian-Laz-Svan-English Dictionary*. Tbilisi: Petiti.

Kuryłowicz, Jerzy. 1956. *L'apophonie en indo-européen*. Wrocław: Polska Akademia Nauk.

Lacan, Marie, Christine Keyser, François-Xavier Ricaut, Nicolas Brucato, Francis Duranthon, Jean Guilaine, Eric Crubézy, and Bertrand Ludes. 2011. Ancient DNA Reveals Male Diffusion through the Neolithic Mediterranean Route. *PNAS* 108(24): 9788-91.

Ladefoged, Peter and Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford/Cambridge, MA: Blackwell.

Lafon, René. 1934. Mots “méditerranéens” en géorgien et dans quelques autres langues caucasiques. *Revue des Études Anciennes* 36(1): 32-46.

de Lagarde, Paul. 1877. *Armenische Studien*. Göttingen: Dieterich'sche Verlags-Butchhandlung.

Latte, Kurt. 1955. Zur griechischen Wortforschung II. *Glotta* 34(3/4): 190-202.

Laviosa Zambotti, Pia. 1937. La civiltà enea della valle padana studiata specialmente nella ceramica. *Studi Etruschi* 11: 9-56.

Laviosa Zambotti, Pia. 1939. Sulla costituzione dell'Eneolitico italiano e le relazioni eneolitiche intermediterranee. *Studi Etruschi* 13: 11-83.

Laviosa Zambotti, Pia. 1949. Movimenti culturali ed etnici in Italia durante l'età del rame e del bronzo (Protolatini ed Osco-Umbri). *Rivista di Antropologia* 37: 3-33.

Lazaridis, Iosif, Nick Patterson, Alissa Mittnik, Gabriel Renaud, Swapan Mallick, Karola Kirisanow, Peter H. Sudmant, Joshua G. Schraiber, Sergi Castellano, Mark Lipson, Bonnie Berger, Christos Economou, Ruth Bollongino, Qiaomei Fu, Kirsten I. Bos, Susanne Nordenfelt, Heng Li, Cesare de Filippo, Kay Prüfer, Susanna Sawyer, ...Johannes Krause. 2014. Ancient Human Genomes Suggest Three Ancestral Populations for Present-Day Europeans. *Nature* 513: 409-13.

Lazaridis, Iosif, Dani Nadel, Gary Rollefson, Deborah C. Merrett, Nadin Rohland, Swapan Mallick, Daniel Fernandes, Mario Novak, Beatriz Gamarra, Kendra Sirak, Sarah Connell, Kristin Stewardson, Eadaoin Harney, Qiaomei Fu, Gloria Gonzalez-Fortes, Eppie R. Jones, Songül Alpaslan Roodenberg, György Lengyel, Fanny Bocquentin, Boris Gasparian, ...David Reich. 2016. Genomic Insights into the Origin of Farming in the Ancient Near East. *Nature* 536: 419-24.

Lazaridis, Iosif, Songül Alpaslan-Roodenberg, Ayşe Acar, Ayşen Açıkkol, Anagnostis Agelarakis, Levon Aghikyan, Uğur Akyüz, Desislava Andreeva, Gojko Andrijašević, Dragana Antonović, Ian Armit, Alper Atmaca, Pavel Avetisyan, Ahmet İhsan Aytek, Krum Bacvarov, Ruben Badalyan, Stefan Bakardzhiev, Jacqueline Balen, Lorenc Bejko, Rebecca Bernardos, ...David Reich. 2022. The Genetic History of the Southern Arc: A Bridge between West Asia and Europe. *Science* 377: eabm4247.

Łazaryan, Ruben S. 1981. *Bowsanownneri hayeren-latineren-řowseren-angleren-germaneren bařaran* [Armenian-Latin-Russian-English-French-German Dictionary of Plant Names]. Yerevan: Erevani Hamalsaran.

Legendre, Pierre and Eugene D. Gallagher. 2001. Ecologically meaningful transformations for ordination of species data. *Oecologia* 129: 271-80.

Lehmann, Wilhelm. 1907. Etymologisches. *KZ* 41(4): 390-4.

Lehmann, Winfred. 1986. *A Gothic Etymological Dictionary*. Leiden: Brill.

[LEIA] Vendryes, Joseph, Édouard Bachallery & Pierre-Yves Lambert. 1959–1996. *Lexique étymologique de l'irlandais ancien*. Dublin: Dublin Institute for Advanced Studies.

Lejeune, Michel. 1972. *Phonétique historique de mycénien et du grec ancien*. Paris: Klincksieck.

Lejeune, Michel. 1993. Le nom de mesure λίτρα: essai lexical. *Revue des Études Grecques* 106: 1-11.

- Leopold, H.R.M. 1929. La sede originaria dei Terramaricoli. *Bullettino di Paleontologia Italiana* 49: 19-31.
- Leumann, Manu. 1977. *Lateinische Grammatik: Erster Band, Lateinische Laut- und Formenlehre*. München: C.H. Bech.
- Levin, Saul. 1983. Hebrew {pi(y)lēḡeš}, Greek παλλακή, Latin *paelex*: The origin of intermarriage among the early Indo-Europeans and Semites. *General Linguistics* 23(3): 191-7.
- Lévy, Isidore. 1900. Sur quelques noms sémitiques de plantes en grèce et en égypte. *Revue Archéologique* 36: 334-344.
- Lewthwaite, James. 1981. Plains Tails from the Hills: Transhumance in Mediterranean Archaeology. In Sheridan, Alison and Geoff Bailey (eds.). *Economic Archaeology*, pp. 57-66. Oxford: British Archaeology Reports (B.A.R.).
- Lewy, Ernst. 1934. Sprachgeographische Probleme des Mittelmeerrandes. *Studi Etruschi* 8: 171-8.
- Lewy, Heinrich. 1895. *Die semitischen Fremdwörter im Griechischen*. Berlin: Gaertner.
- Lewy, Heinrich. 1930. Etymologien. *KZ* 58(1/2); 16-35.
- Lidén, Evald. 1897. *Studien zur Altindischen und vergleichenden Sprachgeschichte*. Upsala: Almqvist & Wiksell.
- Lidén, Evald. 1905. Baumnamen und Verwandtes. *Indogermanische Forschungen* 18(1): 485-509.
- Lidén, Evald. 1906. *Armenische Studien*. Gothenburg: Wald. Zachrissons.
- Lipp, Reiner. 2020. The Word for *Wine* in Anatolian, Greek, Armenian, Italic, Etruscan, Semitic and Its Indo-European Origin. In Kim, Ronald I., Jana Mynářová, and Peter Pavúk (eds.). *Hrozný and Hittite. The First Hundred Years*, pp. 195-229. Leiden/Boston: Brill.
- Lipson, Mark, Anna Szécsényi-Nagy, Swapan Mallick, Annamária Pósa, Balázs Stégmár, Victoria Keerl, Nadin Rohland, Kristin Stewardson, Matthew Ferry, Megan Michel, Jonas Oppenheimer, Nasreen Broomandkoshbacht, Eadaoin Harney, Susanne Nordenfelt, Bastien Llamas, Balázs G. Mende, Kitti Köhler, Krisztián Oross, Mária Bondár, Tibor Marton, ...David Reich. 2017. Parallel Palaeogenomic Transects Reveal Complex Genetic History of Early European Farmers. *Nature* 551: 368-72.
- [LIV2] Rix, Helmut, Martin Kümmel et al. 2001. *Lexicon der indogermanischen Verben*. Second edition. Wiesbaden: Reichert.
- Lockwood, William B. 1990. Latin *columba*, *plaumbēs*, Greek κόλυμβος. *KZ* 103(2): 261-263.

Lockwood, William B. 2001. On the origin of Lat. *hirundo* and Gr. *χελιδών*. *Glotta* 77(3/4): 217-18.

Loewe, Gustav. 1884. *Glossae Nominum*. Leipzig: B.G. Teubner.

Loma, Aleksandar. 2012. Achaism vs. Innovation, Composition vs. Derivation. In Furlan, Metka and Alenka Šivic-Dular (eds.). *Praslovanska dialektizacija v luči etimoloških raziskav*, pp. 151-61. Ljubljana: Založba ZRC.

Loprieno, Antonio. 1995. *Ancient Egyptian. A linguistic introduction*. Cambridge: Cambridge University Press.

[LS] Lewis, Charlton T. and Charles Short. 1879. *A Latin Dictionary*. Oxford: Oxford University Press.

Lubotsky, Alexander. 1989. Against a Proto-Indo-European phoneme **a*. In Vennemann, Theo (ed.). *The New Sound of Indo-European: Essays in Phonological Reconstruction*, pp. 53-66. Berlin/New York: Mouton de Gruyter.

Lubotsky, Alexander. 1994. Avestan *θβōrāštar-* and the Indo-European root **turk-*. *Die Sprache* 36(1): 94-102.

Lubotsky, Alexander. 1998. Vedic roots of the type **TERD^h*. *Studia indogermanica Lodziensia* 2: 75-81.

Lubotsky, Alexander. 2001. The Indo-Iranian substratum. In Carpelan, C., A. Parpola, and P. Koskikallio (eds.). *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations. Papers presented at an international symposium held at the Tvärminne Research Station of the University of Helsinki 8-10 January 1999*, pp. 301-17. Helsinki: Suomalais-ugrilainen seura.

Lubotsky, Alexander. 2004. Avestan *siiazd-*, Sanskrit *sedh-*, Latin *cedere*. In: Hyllested, A., A.R. Jørgensen, J.H. Larsson, T. Olander (eds.). *Per aspera ad asteriscos: Studia Indogermanica in honorem Jens Elmegard Rasmussen sexagenarii Idibus Martiis anno MMIV*, pp. 322-332. Innsbruck: IBS.

Lühr, Rosemarie. 1985. Fälle von Doppelkonsonanz im Keltischen: Zur Frage ihrer Genese. *Sprachwissenschaft* 10(3/4): 274-346.

Lühr, Rosemarie. 1988. *Expressivität und Lautgesetz im Germanischen*. Heidelberg: Universitätsverlag Winter.

Maass, Ernst. 1925. Eunuchos und Verwandtes. *Rheinisches Museum für Philologie* 74(4): 432-76.

MacBain, Alexander. 1911. *An Etymological Dictionary of the Gaelic Language*. Stirling: Eneas MacKay.

Machek, Václav. 1955. Graeco-slavica. *Lingua Posnaniensis* 5: 59-71.

- Machek, Václav. 1971. *Etymologický Slovník Jazyka Českého*. Prague: Československé Akademie Věd.
- Mackenzie, D.N. 1971. *A Concise Pahlavi Dictionary*. London/New York: Routledge.
- Magni, Elisabetta. 2017. Suffix Borrowing and Conflict through Latin-Greek Hybrid Formations. *Pallas* 103: 283-91.
- Magnússon, Ásgeir Blöndal. 1989. *Íslensk orðsifjabók*. Reykjavik: Orðabók Háskólans.
- Malkiel, Yakov. 1984. Six Categories of Nasal Epenthesis: Their Place in the Evolution from Latin into Romance. In Brugman, Cladia and Monica Macaulay (eds.). *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society February 17-20, 1984*, pp. 27-46. Berkeley, CA: Berkeley Linguistics Society.
- Mallory, James. 1989. *In Search of the Indo-Europeans: Language, Archaeology, and Myth*. London: Thames & Hudson.
- Mallory, James and Douglas Q. Adams. 2006. *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*. Oxford: Oxford University Press.
- Malmström, Helena, M. Thomas P. Gilbert, Mark G. Thomas, Mikael Brandström, Jan Storå, Petra Molnar, Pernille K. Andersen, Christian Bendixen, Gunilla Holmlund, Anders Götherström, and Eske Willerslev. 2009. Ancient DNA Reveals Lack of Continuity between Neolithic Hunter-Gatherers and Contemporary Scandinavians. *Current Biology* 19(20): 1758-62.
- Malone, Caroline. 2003. The Italian Neolithic: A Synthesis of Research. *Journal of World Prehistory* 17(3): 235-312.
- Mann, Stuart E. 1963. *Armenian and Indo-European (Historical Phonology)*. London: Luzac & Co.
- Marchesini, Simona. 2012. The Elymian Language. In Tribulato, Olga (ed.). *Language and Linguistic Contact in Ancient Sicily*, pp. 95-114. Cambridge: Cambridge University Press.
- Marchesini, Simona. 2013. Considerazioni storico-linguistiche. In de Simone, Carlo and Simona Marchesini (eds.). *La Lamina di Demlfeld*, pp. 73-89. Pisa/Rome: Fabrizio Serra.
- Martinet, André. 1950. *Manual de lingüística indoeuropea*, dirigido por Antonio Tovar (Review). *Word* 6(2): 187-9.
- Martinet, André. 1955. Le couple *senex-senatvs* et le "suffixe" *-k-*. *Bulletin de la Société de Linguistique de Paris* 51(1): 42-56.
- Martirosyan, Hrach. 2009. *Etymological Dictionary of the Armenian Inherited Lexicon*. Leiden: Brill.

Martzloff, Vincent. 2019. La malédiction osque de Capoue à l'encontre de PAKIS KLUVATIIS (*ST* Cp 37, Vetter 6, Audolent 193). Considérations phraséologiques, morpho-syntaxiques et phonologiques. *Wékʷos* 5: 263-340.

Marzatico, Franco. 2004. 150 Years of Lake-Dwelling Research in Northern Italy. In Menotti, Francesco (ed.). *Living on the Lake in Prehistoric Europe. 150 Years of Lake-Dwelling Research*, pp. 83-97. London/New York: Routledge.

Masson, Émilie. 1967. *Recherches sur les plus anciens emprunts sémitiques en grec*. Paris: Klincksieck.

Mastrelli, Carlo Alberto. 2002. A Proposito di gr. *λάγυρος, λεκάνη/λακάνη, λέκος, λήκυθος* e di itt. *lahanni-*. In de Martino, Stefano and Franca Pecchioli Daddi (eds.). *Anatolia Antica. Studi in memoria di Fiorella Imparati*, Vol. 2, pp. 563-70. Vicchio: LoGisma editore.

Mata Oroval, Xavier. 2017. Zoònims Llatins Motivats: asilus, colubra, columba, hirundo, scrofa. *Liburna* 10: 47-78.

Matasović, Ranko. 2009. *Etymological Dictionary of Proto-Celtic*. Leiden: Brill.

Matasović, Ranko. 2013. Substratum words in Balto-Slavic. *FILOLOGIJA* 60: 75-102.

Matasović, Ranko. 2016. Latin tree and plant names with the suffix *-ix, -ex*. In Kazanskij, Nikolaj N. (ed.). *Indo-evropejskoe jazykoznanie i klassičeskaja filologija* Vol. 20, pp. 700-9. St. Petersburg: Nauka.

Matasović, Ranko. 2020. Language of the bird names and the pre-Indo-European substratum. In Garnier, Romain (ed.). *Loanwords and substrata: proceedings of the colloquium held in Limoges (5th-7th June, 2018)*, pp. 329-44. Innsbruck: IBS.

Matasović, Ranko. fthc. Proto-Slavic Forest Tree Names: Substratum or PIE Origin? In Kroonen, Guus (ed.). *Sub-Indo-European Europe: Continuity and discontinuity in the linguistic landscape of prehistoric Europe*. Berlin/Boston: De Gruyter Mouton.

Mathieson, Iain, Iosif Lazaridis, Nadin Rohland, Swapan Mallick, Nick Patterson, Songül Alpaslan Roodenberg, Eadaoin Harney, Kristin Stewardson, Daniel Fernandes, Mario Novak, Kendra Sirak, Crostoma Gamba, Eppie R. Jones, Bastien Llamas, Stanislav Dryomov, Joseph Pickrell, Juan Luís Arsuaga, José María Bermúdez de Castro, Eudald Carbonell, Fokke Gerritsen, ...David Reich. 2015. Genome-Wide Patterns of Selection in 230 Ancient Eurasians. *Nature* 528: 499-503.

Mathieson, Iain, Songül Alpaslan-Roodenberg, Cosimo Posth, Anna Szécsényi-Nagy, Nadin Rohland, Swapan Mallick, Iñigo Olalde, Nasreen Broomandkhoshbacht, Francesca Candilio, Olivia Cheronet, Daniel Fernandes, Matthew Ferry, Beatriz Gamarra, Gloria González Fortes, Wolfgang Haak, Eadaoin Harney, Eppie Jones, Denise Keating, Ben Krause-Kyora, Isil Kucukkalipci, ...David Reich. 2018. The Genomic History of Southeastern Europe. *Nature* 555: 197-203.

- Matzinger, Joachim. 2005. Messapisch und Albanisch. *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 2: 29-54.
- Matzinger, Joachim. 2019. *Messapisch*. Wiesbaden: Reichert.
- Mayrhofer, Manfred. 1950. Etymologische Miszellen. *Archiv Orientální* 18(4): 68-77.
- Mayrhofer, Manfred. 1961. Indogermanistische Randglossen zu „Kluge-Mitzka“. *Die Sprache* 7: 177-189.
- Mawet, Francine. 2008. Les suffixes complexes en -s- de l'indo-européen et l'élargissement *-s-. *Studi e Saggi Linguistici* 45: 19-70.
- McCone, Kim. 2005. Mögliche nicht-indogermanische Elemente in den keltischen Sprachen. In Meiser, Gerhard and Olaf Hackstein (eds.). *Sprachkontakt und Sprachwandel*, pp. 395-436. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Mees, Bernard. 2003. Stratum and shadow: a Genealogy of Stratigraphy Theories from the Indo-European West. In Andersen, H. (ed.). *Language Contacts in Prehistory: Studies in Stratigraphy*, pp. 11-44. Amsterdam/Philadelphia: John Benjamins.
- Meester, Lotte. fthc. Substrate Stratification: an Argument Against the Unity of Pre-Greek. In Kroonen, Guus (ed.). *Sub-Indo-European Europe: Continuity and discontinuity in the linguistic landscape of prehistoric Europe*. Berlin/Boston: De Gruyter Mouton.
- von Mehrhart, Gero. 1942. *Donauländische Beziehungen der früheisenzeitlichen Kulturen Mittelitaliens*. Darmstadt: Bonner Jahrbücher.
- Meid, Wolfgang. 1986. Hans Kuhns "Nordwestblock"-Hypothese. Zur Problematik der "Völker zwischen Germanen und Kelten". In Beck, Heinrich (ed.). *Germanenprobleme in heutiger Sicht*, pp. 183-212. Berlin/New York: Walter de Gruyter.
- Meillet, Antoine. 1894. Notes arméniennes. *Mémoires de la Société de Linguistique de Paris* 8: 153-65.
- Meillet, Antoine. 1908. De quelques emprunts probables en grec et en latin. *Mémoires de la Société de linguistique de Paris* 15: 161-4.
- Meillet, Antoine. 1912. À propos de Avestique *zrazdā*-. *Mémoires de la Société de linguistique de Paris* 18: 60-4.
- Meillet, Antoine. 1922a. Wageningen, Jan. *Latijnsch woordenboek* (Review). *Bulletin de la Société de Linguistique de Paris* 23(2): 72-4.
- Meillet, Antoine. 1922b. K. Oštir. *Beiträge zur alarodischen Sprachwissenschaft* (Review). *Bulletin de la Société de Linguistique de Paris* 23(1): 128-30.
- Meillet, Antoine. 1925. Les origines du vocabulaire slave: I. – le problème de l'unité balto-slave. *Revue des études slaves* 5(1/2): 5-13.

Meillet, Antoine. 1948. *Esquisse d'une histoire de la langue latine*. 5th ed. Paris: Klincksieck.

Meillet, Antoine and Joseph Vendryes. 1979. *Traité de grammaire comparée des langues classiques*. 5th ed. Paris: Champion.

Meiser, Gerhard. 1986. *Lautgeschichte der Umbrischen Sprache*. Innsbruck: Institut für Sprachwissenschaft.

Meiser, Gerhard. 2010. *Historische Laut- und Formenlehre der lateinischen Sprache*. 3rd edition (unchanged reprint of the 2nd edition). Darmstadt: WBG.

Meister, Richard. 1889. *Die griechischen Dialekte auf Grundlage von Ahrens' Werk: „De Graecae linguae dialectis“*. 2. Vol. Göttingen: Vandenhoeck & Ruprecht.

Melchert, H. Craig. 2003. Language. In Melchert, H. Craig (ed.). *The Luwians*, pp. 170-210. Boston/New York: Brill.

Melchert, H. Craig. 2008. Greek *mólybdos* as a loanword from Lydian. In: B.J. Collins, M.R. Bachvarova & I.C. Rutherford (eds.). *Anatolian interfaces: Hittites, Greeks and their Neighbours*, pp. 153–158. Oxford: Oxbow.

Melchert, H. Craig. 2022. Lengthened grades of PIE *(C)aC roots. In Malzahn, Melanie, Hannes A. Fellner, and Theresa-Susanna Illés (eds.). *Zurück zur Wurzel: Struktur, Funktion und Semantik der Wurzel im Indogermanischen. Akten der 15. Fachtagung der Indogermanischen Gesellschaft vom 13. bis 16. September 2016 in Wien*, pp. 197-204. Wiesbaden: Reichert.

Méndez Dosuna, Julián. 2012. Ancient Macedonian as a Greek dialect: A critical study on recent work. In Giannakis, G.K. (ed.). *Ancient Macedonia: Language, History, Culture*, pp. 133-45. Thessaloniki: Kéntro Ellinikís Glóssas.

Menozi, Paolo, Alberto Piazza, and Luigi Luca Cavalli-Sforza. 1978. Synthetic Maps of Human Gene Frequencies in Europeans. *Science* 201(4358): 786-92.

Merrit, Andrew. 2021. κέλϝφοϝ and κάλϝπτϝ. *Indogermanische Forschungen* 126: 305-24.

Meyer, Eduard. 1893. *Geschichte des Altertums*. Vol. 1. Stuttgart: J.G. Cotta.

Meyer, Eduard. 1909. *Geschichte des Altertums*. Vol. 1, Part 2. 2nd edition. Stuttgart: J.G. Cotta.

Meyer, Gustav. 1883. Albanesische Studien. I. Die Pluralbildungen der albanesischen Nomina. *Sitzungsberichte der philosophisch-historischen Classe der kaiserlichen Akademie der Wissenschaften* 104: 257-362.

Meyer, Gustav. 1892. Etymologisches. *Indogermanische Forschungen* 1: 319-329.

Meyer, W. 1887. Kleine Beiträge zur lateinischen Grammatik. *KZ* 28(1/2): 162-76.

- Meyer-Lübke, Wilhelm. 1903. Zu den lateinischen Glossen. *Wiener Studien* 25: 90-109.
- Meyer-Lübke, Wilhelm. 1920. *Einführung in das Studium der romanischen Sprachwissenschaft*. 3rd edition. Heidelberg: Carl Winter.
- Mikić, Aleksandar and Margarita Vishnyakova. 2012. A Short Survey on the Lexicological Continuum Related to the Most Ancient Eurasian Pulses in the North Caucasian Languages. In Asatrian, Garnik (ed.). *Iran and the Caucasus*, pp. 217-23. Leiden: Brill.
- Mikkola, Joos J. Baltische etymologien II. *Beiträge zur Kunde der indogermanischen Sprachen* 25(1/2): 73-76.
- Milevski, Ianir and Liora K. Horwitz. 2019. Domestication of the Donkey (*Equus asinus*) in the Southern Levant: Archaeozoology, Iconography and Economy. In Kowner, Rotem, Guy Bar-Oz, Michal Biran, Meir Shahar, and Gideon Shelach-Lavi (eds.). *Animals and Human Society in Asia*, pp. 93-148. London: Palgrave Macmillan.
- Militarev, Alexander and Leonid Kogan. 2005. *Semitic Etymological Dictionary*. Vol. II. Animal Names. Münster: Ugarit-Verlag.
- Möller, Hermann. 1911. *Vergleichendes indogermanisch-semitisches Wörterbuch*. Göttingen: Vandenhoeck & Ruprecht.
- Molloy, Barry, Caroline Bruyère, and Dragan Jovanović. 2023. Rethinking Material Culture Markers for Mobility and Migration in the Globalising European Later Bronze Age: A Comparative View from the Po Valley and Pannonian Plain. *Proceedings of the British Academy* 254: 142-69.
- Mommsen, Theodor. 1850. *Die Unteritalischen Dialekte*. Leipzig: Georg Wigand.
- Morpurgo Davies, Anna. 1986. The Linguistic Evidence. Is There Any? In Cadogan, Gerald (ed.). *The End of the Early Bronze Age in the Aegean*, pp. 93-124. Leiden: Brill.
- Much, Rudolf. 1898. Etymologisches. *Zeitschrift für deutsches Altertum und deutsche Literatur* 42(2): 163-72.
- Muller Jzn., Frederik. 1918. Etymologiae Graecae. *Mnemosyne* 46(2): 135-155.
- Muller Jzn., Frederik. 1926. *Altitalisches Wörterbuch*. Göttingen: Vandenhoeck and Ruprecht.
- Mussi, Margherita, Patrizia Gioia, and Fabio Negrino. 2006. Ten small sites: the diversity of the Italian Aurignacian. In O. Bar-Yosef and J. Zilhão (eds.). *Towards a Definition of the Aurignacian*, pp. 189-290. Lisbon: American School of Prehistoric Research/Instituto Português de Arqueologia.
- National Research Council. 1994. *An Assessment of Atlantic Bluefin Tuna*. Washington, DC: The National Academies Press.

- Nehring, Alfons. 1923. Zum Namen der Quitte. *Glotta* 13(1/2): 11-16.
- Nehring, Alfons. 1925. Griech. τίτος, τῆνη und ein vorgriechisches *k*-Suffix. *Glotta* 14(1/2): 153-92.
- Nehring, Alfons. 1927. Lateinische Grammatik. In: Literaturbericht für das Jahr 1924. *Glotta* 15(3/4): 245-279.
- Nehring, Alfons. 1928. Pererga zur lateinischen Wortforschung. *Glotta* 17(1/2): 117-34.
- Nehring, Alfons. 1949. Die Wortsippe von gr. καθάλλης. *Die Sprache* 1: 164-70.
- Nehring, Alfons. 1959. Idg. **mari*, **mori*. In Rasch, Wolfdietrich (ed.). *Festschrift für Franz Rolf Schröder zu seinem 65. Geburtstag. September 1958*, pp. 122-38. Heidelberg: Carl Winter.
- Neri, Sergio. 2016. Kroonen, Guus: Etymological Dictionary of Proto-Germanic [Review]. *Kratylos* 61: 1-51.
- Neri, Sergio. 2017. Lat. *umbra* und Verwandtes. In Hajnal, Ivo, Daniel Kölligan, and Katharina Zisper (eds.). *Miscellanea Indogermanica: Festschrift für José Luis García Ramón zum 65. Geburtstag*, pp. 561-74. Innsbruck: IBS.
- Neumann, Günter. 1964. Meriggi Piero. Hieroglyphisch-Hethitisches Glossar (Review). *Indogermanische Forschungen* 69: 56-62.
- Neumann, Günter. 1971a. Studien zu den Boğazköy-Texten 7-13 (Review). *Indogermanische Forschungen* 76: 260-74.
- Neumann, Günter. 1971b. Substrate im Germanischen? *Nachrichten der Akademie der Wissenschaften in Göttingen. Philologisch-historische Klasse*. 4: 75-99.
- Neumann, Günter. 1987. Griechisch κβεγάω. *KZ* 100(1): 64-9.
- Neumann, Günter. 1992. Zu κβεγάω. *KZ* 105(2): 188.
- Newmark, Leonard. 2005. *Albanian-English Dictionary*. <http://www.seelrc.org:8080/albdict/>
- Nicolis, Franco. 2001. Some Observations on the Cultural Setting of the Bell Beakers of Northern Italy. In Nicolis, Franco (ed.). *Bell Beakers Today. Pottery, People, Culture, Symbols in Prehistoric Europe*. Vol. 1, pp. 41-57. Trento: Ufficio Beni Archeologici.
- Niedermann, Max. 1903. Etymologische Forschungen. *Indogermanische Forschungen* 15: 104-21.
- Niedermann, Max. 1905-6. Walde A., Lateinisches etymologisches Wörterbuch. (Review). *Anzeiger für indogermanische Sprach- und Altertumskunde. Beiblatt zu den Indogermanischen Forschungen* 18: 72-81.
- Niedermann, Max. 1909. Zur griechischen und lateinischen Wortkunde.

Indogermanische Forschungen 26: 43-59.

Niedermann, Max. 1916. Zur indogermanischen Wortforschung. *Indogermanische Forschungen* 37: 145-55.

Niedermann, Max. 1918. *Essais d'étymologie et de critique verbale latines*. Neuchâtel: Attinger Frères.

Niedermann, Max. 1919. Die Names des Storches im Litauischen. In *Festgabe Adolf Kaegi*, pp. 66-92. Frauenfeld: Huber & Co.

Niedermann, Max. 1927. De quelques noms indo-européens du millet. In Taszycki, W. (ed.). *Symbolae grammaticae in honorem Ioannis Rozwadowski*, pp. 109-17. Kraków: Z zasilku Wydziału Nauki Ministerstwa W.R. i O.P.

Niedermann, Max. 1930. Zur lateinischen und griechischen Wortgeschichte. *Glotta* 19(1/2): 1-15.

Nielsen, Benedicte. 2004. On Latin instrument-nouns in */-lo-/. In Clackson, James and Birgit Annette Olsen (eds.). *Indo-European word-formation: proceedings of the conference held at the University of Copenhagen October 20th-22nd 2000*, pp. 189-213. Copenhagen: Museum Tusulanum.

Nieminen, E. 1956. Die Benennungen der Hirse in den baltischen Sprachen. *KZ* 74(3/4): 162-76.

Nussbaum, Alan. 1986. *Head and Horn in Indo-European*. Berlin/New York: Walter de Gruyter.

Nussbaum, Alan. 1997. The "Saussure Effect in Latin and Italic". In Lubotsky, Alexander (ed.). *Sound Law and Analogy*, pp. 181-203. Amsterdam: Rodopi.

Oettinger, Norbert. 2010. Seevölker und Etrusker. In Cohen, Yoram, Amir Gilan, and Jared L. Miller (eds.). *Pax Hethitica*, pp. 233-46. Wiesbaden: Harrassowitz.

Oettinger, Norbert. 2021. Language Contact between Lydian and Greek or the Origin of Lydian *k*. In Bianconi, Michele (ed.). *Linguistic and Cultural Interactions between Greece and Anatolia*, pp. 116-30. Leiden/Boston: Brill.

Ogle, Marbury B. 1945. As a Tale That Is Told. *Monatshefte für Deutschen Unterricht* 37(4/5): 130-4.

Olalde, Iñigo, Selina Brace, Morten E. Allentoft, Ian Armit, Kristian Kristiansen, Thomas Booth, Nadin Rohland, Swapan Mallick, Anna Szécsényi-Nagy, Alissa Mittnik, Eveline Altena, Mark Lipson, Iosif Lazaridis, Thomas K. Harper, Nick Patterson, Nasreen Broomandkoshbacht, Yoan Diekmann, Zuzana Faltyskova, Daniel Fernandes, Matthew Ferry,...David Reich. 2018. The Beaker Phenomenon and the Genomic Transformation of Northwest Europe. *Nature* 555: 190-6.

Olalde, Iñigo, Swapan Mallick, Nick Patterson, Nadin Rohland, Vanessa Villalba-Mouco, Marina Silva, Katharina Dulias, Ceiridwen J. Edwards, Francesca Gandini, Maria Pala, Pedro Soares, Manuel Ferrando-Bernal, Nicole Adamski, Nasreen Broomandkhoshbacht, Olivia Cheronet, Brendan J. Culleton, Daniel Fernandes, Ann Marie Lawson, Matthew Mah, Jonas Oppenheimer, ...David Reich. 2019. The Genomic History of the Iberian Peninsula over the Past 8000 Years. *Science* 363(6432): 1230-4. DOI: 10.1126/science.aav4040

Olivetti, E., Sabina Rendine, Nazario Cappello, E.S. Curtoni, and Alberto Piazza. 1986. The HLA System in Italy. *Human Heredity* 36: 357-72.

Olsen, Birgit Anette. 1999. *The Noun in Biblical Armenian*. Berlin/New York: De Gruyter Mouton.

Olsen, Birgit Anette. 2009. Martinet's rule of laryngeal hardening - a reappraisal. In Jamieson, Stephanie W., H. Craig Melchert, and Brent Vine (eds.). *Proceedings of the 21st Annual UCLA Indo-European Conference, October 30th and 31st, 2009*, pp. 207-22. Bremen: Hempen.

Orel, Vladimir E. 1987. Der indogermanische Akzent im Albanischen. *Zeitschrift für Balkanologie* 23(2): 140-50.

Orel, Vladimir E. 1998. *Albanian Etymological Dictionary*. Leiden: Brill.

Östenberg, Carl Eric. 1967. *Luni sul Mignone e problem della preistoria d'Italia*. Lund: C.W.K. Gleerup.

Osthoff, Hermann. 1890. Anlautend indog. *sr-* im lateinischen. *Morphologische Untersuchungen auf dem Gebiete der indogermanischen Sprachen* 5: 62-84.

Osthoff, Hermann. 1892. Eine bis jetzt nicht erkannte Präsensstamm bildung im Indogermanischen. In *Verhandlungen der einundvierzigsten Versammlung deutscher Philologen und Schulmänner in München vom 20. bis 23. Mai 1891*, pp. 300-304. Leipzig: B.G. Teubner.

Osthoff, Hermann. 1893. Lat. *gero*. *Beiträge zur Kunde der indogermanischen Sprachen* 19: 320-2.

Osthoff, Hermann. 1901. *Etymologische Parerga. Erster Teil*. Leipzig: S. Hirzel.

Oštir, Karel. 1921. *Beiträge zur alarodischen Sprachwissenschaft*. Vienna: Beyer.

Oštir, Karel. 1930. *Drei voroslavisch-etruskische Vogelnamen*. Ljubljana: Učiteljska Tiskarna.

Otrębski, Jan. 1939. *Studja Indoeuropeistyczne/Indogermanische Forschungen*. Vilnius: Towarzystwo Przyjaciół Nauk w Wilnie.

Pääbo, Svante. 2014. *Neanderthal Man: In Search of Lost Genomes*. New York, NY:

Basic Books.

- Pallottino, Massimo. 1940. Appunti di protostoria latina ed etrusca. *Studi Etruschi* 14: 27-32.
- Pallottino, Massimo. 1975. *The Etruscans*. Revised and Enlarged, Translated by J. Cremona. Bloomington/London: Indiana University Press.
- Palmer, Leonard Robert. 1954. *The Latin Language*. London: Faber and Faber.
- Palmer, Leonard Robert. 1963. *The Interpretation of Mycenaean Greek Texts*. Oxford: The Clarendon Press.
- Papac, Luka, Michal Ernée, Miroslav Dobeš, Michaela Langová, Adam B. Rohrlach, Franziska Aron, Gunnar U. Neumann, Maria A. Spyrou, Nadin Rohland, Petr Velemínský, Martin Kuna, Hana Brzobohatá, Brendan Culleton, David Daněček, Alžběta Danielisová, Miluše Dobisíková, Josef Hložek, Douglas J. Kennett, Jana Klementová, Michal Kostka, ... Wolfgang Haak. 2021. Dynamic Changes in Genomic and Social Structures in Third Millennium BCE Central Europe. *Science Advances* 7(35): eabi6941.
- Pardo-Seco, Jacobo, Alberto Gómez-Carballa, Jorge Amigo, Federico Martínón-Torres, and Antonio Salas. 2014. A Genome-Wide Study of Modern-Day Tuscans: Revisiting Herodotus's Theory on the Origin of the Etruscans. *PLoS ONE* 9(9): e105920.
- Parejko, Ken. 2003. Pliny the Elder's Silphium: First Recorded Species Extinction. *Conservation Biology* 17(3): 925-7.
- Parker, Holt N. 1988. Latin *sisō > serō and Related Rules. *Glotta* 66: 221-41.
- Patroni, Giovanni. 1937. *La Preistoria*. Milan: Francesco Vallardi.
- Patroni, Giovanni. 1939. L'indoeuropeizzazione d'italia. *Athenaeum* 17(3-4): 213-26.
- Patterson, Nick, Priya Moorjani, Yontao Luo, Swapan Mallick, Nadin Rohland, Yiping Zhan, Teri Genschoreck, Teresa Webster, and David Reich. 2012. Ancient Admixture in Human History. *Genetics* 192(3): 1065-93.
- Patterson, Nick, Michael Isakov, Thomas Booth, Lindsey Büster, Claire-Elise Fischer, Iñigo Olalde, Harald Ringbauer, Ali Akbari, Olivia Cheronet, Madeleine Bleasdale, Nicole Adamski, Eveline Altena, Rebecca Bernardos, Selina Brace, Nasreen Broomandkhoshbacht, Kimberly Callan, Francesca Candilio, Brendan Culleton, Elizabeth Curtis, Lea Demetz, ... David Reich. 2022. Large-Scale Migration into Britain during the Middle to Late Bronze Age. *Nature* 601: 588-594.
- Paulis, Giulio. 1992. *I nomi popolari delle piante in Sardegna*. Sassari: Carlo Delfino.
- Pedersen, Holger. 1893. *r-n-Stämme*. Studien über den Stammwechsel in der Declination der idg. Nomina. *KZ* 32(2): 240-73.

Pedersen, Holger. 1895. Das indogermanische *s* im Slavischen. *Indogermanische Forschungen* 5: 33-87.

Pedersen, Holger. 1904. Przyczynki do gramatyki porównawczej języków słowiańskich. *Materyaly i prace Komisji Językowej Akademii Umiejętności w Krakowie* 1(2): 165-76.

Pedersen, Holger. 1906. Armenisch und die Nachbarsprachen. *KZ* 39(3): 334-485.

Pedersen, Holger. 1909-1913. *Vergleichende Grammatik der Keltischen Sprachen*. Göttingen: Vandenhoeck und Ruprecht.

Peet, Thomas Eric. 1909. *The Stone and Bronze Ages in Italy and Sicily*. Oxford: Clarendon Press.

Persson, Per. 1909. Zur lateinischen Grammatik und Wortkunde. *Indogermanische Forschungen* 26: 60-8.

Petersson, Herbert. 1914. Einige Tier- und Pflanzennamen aus idg. Sprachen. *KZ* 46(1/2): 128-50.

Petr, V.J. 1896. Slavische etymologien. *Beiträge zur Kunde der indogermanischen Sprachen* 21: 207-17.

Petr, V.J. 1899. Über den wechsel der laute *d* und *l* im Lateinischen. *Beiträge zur Kunde der indogermanischen Sprachen* 25: 127-58.

Peust, Carsten, 1999. *Egyptian Phonology: an Introduction to the Phonology of a Dead Language*. Göttingen: Peust & Gutschmidt.

Pfiffig, Amos. J. 1969. *Die Etruskische Sprache*. Graz: Akademische Druck- und Verlagsanstalt.

Philippa, Marlies, Frans Debrabandere, Arend Quak, Tanneke Schoonheim, and Nicoline van der Sijs. 2003-09. *Etymologisch Woordenboek van het Nederlands*. Amsterdam: Amsterdam University Press.

Piazza, Alberto, Nazario Cappello, E. Olivetti, and Sabina Rendine. 1988. A Genetic History of Italy. *Annals of Human Genetics* 52(3): 203-13.

Pigorini, Luigi. 1875. Scavi nella Terramara di Castione parmense. *Bullettino di Palemologia Italiana* 1(4/5): 53-7.

Pigorini, Luigi. 1903. Le più antiche civiltà dell'Italia. *Bullettino di Palemologia Italiana* 29(10/11): 189-211.

Pinault, Georges-Jean. 2000. Nouveautés dans un commentaire de la Discipline bouddhique. *Tocharian and Indo-European Studies* 9: 77-120.

Pisani, Vittore. 1930. Miscellanea etimologica. *Atti della Reale Accademia dei Lincei. Rendiconti*. Ser. 6, Vol. 6: 171-188.

Pisani, Vittore. 1932. Ernst Kieckers, *Die Sprachstämme der Erde* (Review). *Rivista Indo-Greco-Italica* 19(1/2): 86-90.

Pisani, Vittore. 1938a. L'unità culturale indo-mediterranea anteriore all'avvento di Semiti e Indoeuropei. In *Scritti in onore di Alfredo Trombetti*, pp. 199-213. Milan: U. Hoepli.

Pisani, Vittore. 1938b. Neupersisch *sarv*. (Ein Beitrag zur Geschichte der idg. Palatale.). *Indogermanische Forschungen* 56: 285-9.

Pisani, Vittore. 1954. Zur Sprachgeschichte des alten Italiens. *Rheinisches Museum für Philologie* 97(1): 47-68.

von Planta, Robert. 1892-7. *Grammatik der Oskisch-Umbrischen Dialekte*. Strassbourg: Karl. J. Trübner.

Pleiner, Radomír. 1996. Das frühe Eisen: Von den Kleinwaagemengen zu der ältesten Industrie. *Ethnographisch-Archäologische Zeitschrift* 37(3): 283-91.

Pocetti, Paolo. 2017. The Documentation of Italic. In Klein, Jared, Brian Joseph, and Matthias Fritz (eds.). *Handbook of Comparative and Historical Indo-European Linguistics*. Vol. 2, pp. 733-42. Berlin/Boston: De Gruyter Mouton.

[Pokorny] Pokorny, Julius. 1959. *Indogermanisches etymologisches Wörterbuch*. Bern/Munich: Francke.

Pokorny, Julius. 1923. Die Stellung des Tocharischen im Kreise der indogermanischen Sprachen. *Berichte des Forschungs-Institutes für Osten un Orient* 3: 24-57.

Pokorny, Julius. 1933. Air. capall 'Pferd, Packpferd'. *Zeitschrift für celtische Philologie* 19: 160.

Pokorny, Julius. 1936. Substrattheorie und Urheimat der Indogermanen. *Mitteilungen der anthropologischen Gesellschaft in Wien* 66: 69-91.

Polomé, Edgar. 1952. On the source of Hittite *ḫ*. *Language* 28(4): 444-56.

Polomé, Edgar. 1986. The Non-Indo-European Component of the Germanic Lexicon. In Etter, Annermarie (ed.). *O-o-pe-ro-si: Festschrift für Ernst Risch zum 75. Geburtstag*, pp. 661-672. Berlin: de Gruyter.

Polomé, Edgar. 1989. Substrate Lexicon in Germanic. *NOWELE* 14: 53-73.

Posth, Cosimo, Valentina Zaro, Maria A. Spyrou, Stefania Vai, Guido A. Gencchi-Ruscone, Alessandra Modi, Alexander Peltzer, Angela Mötsch, Katharin Nägele, Åshild J. Vågane, Elizabeth A. Nelson, Rita Radzevičūtė, Cäcilia Freund, Lorenzo M. Bondioli, Luca Cappuccini, Hanna Frenzel, Elsa Pacciani, Francesco Boschin, Giulia Capecchi, Ivan Martini, ...Johannes Krause. 2021. The Origin and Legacy of the Etruscans through a 2000-Year Archaeogenomic Time Transect. *Science*

Advances 7(39): abi7673.

Porzig, Walter. 1927. Kleinasiatisch-indische Beziehungen. *Zeitschrift für Indologie und Iranistik* 5: 265-280.

Pott, August Friedrich. 1853. *Die Personennamen insbesondere die Familiennamen und ihre Entstehungsarten: auch unter Berücksichtigung der Ortsnamen*. Leipzig: F.A. Brockhaus.

Poucet, Jacques. 1985. *Les origines de Rome: Tradition et histoire*. Brussels: Facultés universitaires Saint-Louis.

Prasse, Karl-G. 2003. *Dictionnaire Touareg-Français (Niger)*. Copenhagen: Museum of Tusculanum Press.

Prellwitz, Walther. 1897. Etymologische miscellen. *Beiträge zur Kunde der indogermanischen Sprachen* 22: 118-24.

Prellwitz, Walther. 1905. *Etymologisches Wörterbuch der griechischen Sprache*. Göttingen: Vandenhoeck und Ruprecht.

Prokosch, E. 1939. *A Comparative Germanic Grammar*. Baltimore: The Linguistic Society of America.

Pronk, Tijmen. 2011. The “Saussure Effect” in Indo-European Languages Other Than Greek. *Journal of Indo-European Studies* 39(1/2): 176-93.

Pronk, Tijmen. 2019. Proto-Indo-European *a. *Indo-European Linguistics* 7: 122-63.

Pronk, Tijmen and Saskia Pronk-Tiethoff. 2018. Balto-Slavic Agricultural Terminology. In Kroonen, G., J.P. Mallory, and B. Comrie (eds.). *Talking Neolithic. Proceedings of the workshop on Indo-European origins held at the Max Planck Institute for Evolutionary Anthropology, Leipzig, December 2-3, 2013*, pp. 278-314. Washington: Institute for the Study of Man.

Pronk-Tiethoff, Saskia. 2012. The Germanic loanwords in Proto-Slavic: Origin and Accentuation. (Doctoral dissertation, University of Leiden.) <https://scholarlypublications.universiteitleiden.nl/access/item%3A2930460/view>

Prósper, Blanca María. 2019. What became of “Sabine l’”? An Overlooked Proto-Italic Sound Law. *Journal of Indo-European Studies* 47(3/4): 457-506.

Puglisi, Salvatore. 1959. *La civiltà appenninica: origine della comunità pastorali in Italia*. Florence: Sansoni.

[Puhvel] Puhvel, Jaan. 1984-2021. *Hittite Etymological Dictionary*. Berlin/New York: Mouton de Gruyter.

Pulgram, Ernst. 1958. *The Tongues of Italy*. Cambridge, MA: Harvard University Press.

- Randall-MacIver, David. 1939. Modern Views on the Italian Terremare. *Antiquity* 13(51): 320-3.
- Rata, Cristian G. 2008. Observations on the Language of the Book of Job. *Scripture and Interpretation* 2(1): 5-24.
- Raveane, Alessandro, Serena Aneli, Francesco Montinaro, G. Athanasiadis, Simona Barlera, Giovanni Birolo, G. Boncoraglio, A.M. Di Blasio, Cornelia Di Gaetano, Luca Pagani, Silvia Parolo, P. Paschou, Alberto Piazza, George Stamatoyannopoulos, Andrea Angius, Nicolas Brucato, Francesco Cucca, Garrett Hellenthal, A. Mulas, Marine Peyret-Guzzon, ...Cristian Capelli. 2019. Population Structure of Modern-Day Italians Reveals Patterns of Ancient and Archaic Ancestries in Southern Europe. *Science Advances* 5(9): eaaw3492.
- Recchia, Giulia. 2020. Reaching across the Adriatic: Northern and Western Interactions of the Cetina Phenomenon (25th-20th centuries BC). *Godišnjak Centra Za balkanološka Ispitivanja* 49: 5-28.
- Rédei, Károly. 1986. *Zu den indogermanisch-uralischen Sprachkontakt*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Rédei, Károly. 1988. *Uralisches etymologisches Wörterbuch*. Vol. 1. Budapest: Verlag der Ungarischen Akademie der Wissenschaften.
- Reichert, Hans. 1906. Der sekundäre ablaut. *KZ* 39(1): 1-80.
- Reichert, Hans. 1914. Studien zur lat. Laut- und Wortgeschichte. *KZ* 46(4): 309-50.
- Rellini, Ugo. 1929. *Le origini della civiltà italiana*. Rome: Libreria di Scienze e Lettere del Dott. G. Bardi.
- Rellini, Ugo. 1933. La civiltà enea in italia. *Bullettino di Paletnologia Italiana* 53: 63-94.
- [REW] Meyer-Lübke, Wilhelm. 1935. *Romanisches etymologisches Wörterbuch*. 3rd edition. Heidelberg: C. Winter.
- Ribezzo, Francesco. 1920a. La originaria unità tirrena dell'Italia nella toponomastica. *Rivista Indo-Greco-Italica* 4(1/2): 83-97.
- Ribezzo, Francesco. 1920b. Carattere mediterraneo della più antica toponomastica italiana. *Rivista Indo-Greco-Italica* 4(1/2): 221-36.
- Ribezzo, Francesco. 1928. Le origini mediterranee dell'accento iniziale italo-etrusco. *Rivista Indo-Greco-Italica* 12: 183-204.
- Ribezzo, Francesco. 1929. I testi etruschi CIE 5237 e 4538. *Rivista Indo-Greco-Italica* 13: 59-104.
- Ribezzo, Francesco. 1930. Roma delle origini, Sabini e Sabelli. *Rivista Indo-Greco-Italica* 14: 59-99.

Ribezzo, Francesco. 1932. Unità italica ed unità italo-celtica. *Rivista Indo-Greco-Italica* 16(1/2): 27-40.

Ribezzo, Francesco. 1934a. Sostrato mediterraneo e *Lautverschiebung* germanica. *Rivista Indo-Greco-Italica* 18: 61-101.

Ribezzo, Francesco. 1934b. P. Chantraine La formation de noms en grec ancien (Review). *Rivista Indo-Greco-Italica* 18: 121-6.

Richards, Martin, H. Côrte-Real, P. Forster, V. Macaulay, H. Wilkinson-Herbots, A. Demaine, S. Papiha, R. Hedges, Hans-Jürgen Bandelt, and B. Sykes. 1996. Paleolithic and Neolithic Lineages in the European Mitochondrial Gene Pool. *American Journal of Human Genetics* 59(1): 185-203.

Richards, Martin, V. Macaulay, E. Hickey, E. Vega, B. Sykes, V. Guida, C. Rengo, D. Sellitto, E. Cruciani, T. Kivisild, Richard Villems, M. Thomas, S. Rychkov, O. Rychkov, Y. Rychkov, M. Gölge, D. Dimitrov, E. Hill, Daniel Bradley, V. Romano, ...Hans-Jürgen Bandelt. 2000. Tracing European Founder Lineages in the Near Eastern mtDNA Pool. *American Journal of Human Genetics* 67(5): 1251-76.

Rieken, Elizabeth. 1999. *Untersuchungen zur nominalen Stammbildung des Hethitischen*. Wiesbaden: Harrassowitz.

Rifkin, Matthew J. 2007. A Spatial Analysis of Neolithic Cultures throughout Eastern, Central, and Northern Europe in Relation to Proto-Germanic. *Journal of Indo-European Studies* 35(1/2): 53-81.

Ringe, Don. 2006. *From Proto-Indo-European to Proto-Germanic*. Oxford: Oxford University Press.

Ritter, Hans. 2009. *Wörterbuch zur Sprache und Kultur der Twareg*. Wiesbaden: Harrassowitz.

Rix, Helmut. 1983. Umbro e Proto-Osco-Umbro. In Vineis, Edoardo (ed.). *Le lingue indoeuropee di frammentaria attestazione = Die indogermanische Restsprachen: atti del convegno della Società Italiana di Glottologia e della Indogermanische Gesellschaft: Udine, 22-24 settembre 1981*, pp. 91-107. Pisa: Giardini.

Rix, Helmut. 1985. Schrift und Sprache. In Cristofani, Mauro (ed.). *Die Etrusker*, pp. 210-38. Stuttgart/Zürich: Belsen.

Rix, Helmut. 1994. Latein und Sabellisch: Stammbaum und/oder Sprachbund? *Incontri Linguistici* 17: 13-29.

Rix, Helmut. 1997. *Autumnus* "Herbst" und andere lateinische Vertreter der Wurzel **temh₁*- "schneiden". In Ambrosini, Riccardo, Maria Patrizia Bologna, Filippo Motta, and Chatia Orlandi (eds.). *Scríbthair a ainm n-ogaim. Scritti in Memoria di Enrico Campanile*. Vol. 2, pp. 871-89. Pisa: Pacini.

- Rix, Helmut. 2005. Lehnbeziehungen zwischen den Sprachen Altitaliens. In Meiser, Gerhard and Olav Hackstein (eds.). *Sprachkontakt und Sprachwandel: Akten der XI. Fachtagung der Indogermanischen Gesellschaft, 17.-23. September 2000, Halle an der Saale*, pp. 559-72. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Rix, Helmut. 2008. Etruscan. In Woodard, Roger (ed.). *The Ancient Languages of Europe*, pp. 141-64. Cambridge: Cambridge University Press.
- [RLA] Edzard, Dietz O. and Michael P. Streck. 2003-5. *Reallexikon der Assyriologie und Vorderasiatischen Archäologie*. Vol. 10. Berlin/New York: de Gruyter.
- Rohlf, Gerhard. 1972. *Studi e ricerche su lingua e dialetti d'Italia*. Florence: Sansoni.
- Rönsch, Hermann. 1886. Das Subst. bolunda. *Romanische Forschungen* 2: 317-18.
- Rosén, Haiim. 1994. Lat. *rete*. *Indogermanische Forschungen* 100: 210-2.
- Rosén, Hannah. 1994. Lat. *sacēna*, Heb. *šakkān* and the Mediterranean Substrate. *Journal of Indo-European Studies* 22(1/2): 165-75.
- Rosół, Rafał. 2013. *Frühe semitische Lehnwörter im Griechischen*. Frankfurt am Main: Peter Lang.
- van Rosenberg, Erik. 2013. Forging relationships with Vollgriffdolche: tournaments of value, boundary work and Early Bronze Age network changes in Central Italy. Paper presented at the CinBA Conference *Creativity: An Exploration Through the Bronze Age and Contemporary Responses to the Bronze Age*, 10th - 11th April 2013, Magdalene College, University of Cambridge, UK.
- Rozwadowski, Jan. 1907. Parerga. *Materyały i prace Komisji Językowej Akademii Umiejętności w Krakowie* 2: 341-356.
- Ruijgh, C.J. 1986. R. Giacomelli. *Graeca Italica* (Review). *Mnemosyne* 39(3/4): 507-9.
- Ruijgh, C.J. 1997. La date de la création de l'alphabet grec et celle de l'épopée homérique. *Bibliotheca Orientalis* 54(5-6): 533-603.
- Säflund, Gösta. 1938. Bemerkungen zur Vorgeschichte Etruriens *Studi Etruschi* 12: 17-55.
- Säflund, Gösta. 1939. *Le terremare delle province di Reggio Emilia, Modena, Parma e Piacenza*. Lund: Gleerup.
- Salmons, Joe. 1992. Northwest Indo-European vocabulary and substrate phonology. In Pearson, Roger (ed.). *Perspectives on Indo-European language, culture and religion: studies in Honor of Edgar C. Polomé*. Vol. 2, pp. 265-279. Washington, DC: Institute for the Study of Man.
- Salmons, Joe. 1993. *The glottalic theory: survey and synthesis*. McLean, VA: Institute for the Study of Man.

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*, pp. 478–554. Leiden: Brill.

Sampietro, Maria Lourdes, Oscar Lao, David Caramelli, Martina Lari, R. Pou, M. Martí, Jaume Bertranpetit, and Carles Lalueza-Fox. 2007. Palaeogenetic Evidence Supports a Dual Model of Neolithic Spreading into Europe. *Proceedings of the Royal Society B* 274: 2161-7.

Sarno, Stefania, Alessio Boattini, Marilisa Carta, Gianmarco Ferri, Milena Alù, Daniele Yang Yao, Graziella Ciani, Davide Pettener, and Donata Luiselli. 2014. An Ancient Mediterranean Melting Pot: Investigating the Uniparental Genetic Structure and Population History of Sicily and Southern Italy. *PLoS ONE* 9(4): e96074.

Sarno, Stefania, Alessio Boattini, Luca Pagani, Marco Sazzini, Sara De Fanti, Andrea Quagliariello, Guido A. Gneccchi Ruscone, Etienne Guichard, Graziella Ciani, Eugenio Bortolini, Chiara Barbieri, Elisabetta Cilli, Rosalba Petrilli, Ilia Mikerezi, Luca Sineo, Miguel Vilar, Spencer Wells, Donata Luiselli, and Davide Pettener. 2017. Ancient and Recent Admixture Layers in Sicily and Southern Italy Trace Multiple Migration Routes along the Mediterranean. *Scientific Reports* 7, 1984.

Saupe, Tina, Francesco Montinaro, Cinzia Scaggion, Nicola Carrara, Toomas Kivisild, Eugenia D’Atanasio, Ruoyun Hui, Anu Solnik, Ophélie Lebrasseur, Greger Larson, Lica Alessandri, Ilenia Arienzo, Flavio De Angelis, Mario F. Rolfo, Robin Skeates, Letizia Silvestri, Jessica Beckett, Sahra Talamo, Andrea Dolfini, Monica Miari, ...Christiana L. Scheib. 2021. Ancient Genomes Reveal Structural Shifts after the Arrival of Steppe-Related Ancestry in the Italian Peninsula. *Current Biology* 31(12): 2576-91.

Schaffner, Stefan. 2014. Zur Wortbildung und Etymologie von lateinisch *autumnus* “Herbst”. *Graecolatina et Orientalia* 35-6: 67-102.

Scheftelowitz, J. 1904. Die begriffe für „schädel“ im Indogermanischen. *Beiträge zur kunde der indogermanischen sprachen* 28: 143-158.

Schindler, Jochem. 1994. Alte und neue Fragen zum indogermanischen Nomen. In Rasmussen, Jens E. (ed.). *In memoriam Holger Pedersen. Kolloquium der Indogermanischen Gesellschaft vom 26. bis 28. März 1993 in Kopenhagen*, pp. 397-400. Wiesbaden: Reichert.

Schmalstieg, William R. 1976. *Studies in Old Prussian*. University Park/London: Pennsylvania State University Press.

Schmid, Stephan. 2016. Segmental Phonology. In Ledgeway, Adam and Martin Maiden (eds.). *The Oxford Guide to the Romance Languages*, pp. 471-83. Oxford: Oxford University Press.

Schmidt, Johannes. 1895. *Kritik der Sonantentheorie*. Weimar: Hermann Böhlhaus

Nachfolger.

Schmidt, Karl-Horst. 1966. Keltisches Wortgut im Lateinischen. *Glotta* 44: 151-74.

Schrader, Otto. 1883. *Sprachvergleichung und Urgeschichte: Linguistisch-historische Beiträge zur Erforschung des indogermanischen Altertums*. Jena: Hermann Costenoble.

Schrader, Otto and Alfons Nehring. 1917-23. *Reallexikon der indogermanischen Altertumskunde*. Vol 1. Zweite, vermehrte und umgearbeitete Auflage. Berlin: Walter de Gruyter & Co.

Schrijnen, Joseph. 1914. Das "sabinische l" im Lateinischen. *KZ* 46(4): 376-80.

Schrijver, Peter. 1991. *The Reflexes of the Proto-Indo-European Laryngeals in Latin*. Amsterdam: Rodopi.

Schrijver, Peter. 1995. *Studies in British Celtic Historical Phonology*. Amsterdam: Rodopi.

Schrijver, Peter. 1997. Animal, vegetable and mineral: some Western European substratum words. In Lubotsky, Alexander (ed.). *Sound Law and Analogy*, pp. 293-316. Amsterdam: Rodopi.

Schrijver, Peter. 1999. On Henbane and Early European Narcotics. *Zeitschrift für celtische Philologie* 51(1): 17-45.

Schrijver, Peter. 2001. Lost Languages in Northern Europe. In Carpelan, Christian, Asko Parpola, and Petteri Koskikallio (eds.). *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations*, pp. 419-25. Helsinki: Suomalais-Ugrilainen Seura.

Schrijver, Peter. 2003. Athematic *i*-Presents: the Italic and Celtic Evidence. *Incontri Linguistici* 26: 59-86.

Schrijver, Peter. 2007. Keltisch en de buren: 9000 jaar taalcontact. Inaugural speech at the University of Utrecht, 7 March 2007.

Schrijver, Peter. 2015. The Meaning of Celtic **eburos*. In Guillaume Oudaer, Gael Hily, Herve Le Bihan (eds.), *Mélanges en l'honneur de Pierre-Yves Lambert*, pp. 65 - 76. Utrecht : TIR.

Schrijver, Peter. 2016. Sound Change, the Italo-Celtic Linguistic Unity, and the Italian Homeland of Celtic. In Koch, John T. and Barry Cunliffe (eds.). *Celtic from the West* 3, pp. 489-502. Oxford: Oxbow.

Schrijver, Peter. 2018. Talking Neolithic: The Case for Hatto-Minoan and its relation to Sumerian. In Kroonen, Guus, James P. Mallory, and Bernard Comrie (eds.). *Talking Neolithic: Proceedings of the workshop on Indo-European origins held at the Max Planck Institute for Evolutionary Anthropology, Leipzig, December 2-3, 2013*, pp.

336-374. Washington, DC: Institute for the Study of Man.

Schrijver, Peter. fthc. East Caucasian Perspectives on the Origin of the Word 'Camel' and Some Notes on European Substrate Lexemes. In Kroonen, Guus (ed.). *Sub-Indo-European Europe: Continuity and discontinuity in the linguistic landscape of prehistoric Europe*. Berlin/Boston: De Gruyter Mouton.

Schroeder, Otto. 1930-31. Assyrische Gefäßnamen. *Archiv für Orientalforschung* 6: 111-12.

Schuchardt, Hugo. 1918. *Die romanischen Lehnwörter im Berberischen*. (Kaiserliche Akademie der Wissenschaften in Wien, Philosophisch-historische Klasse 188 [4]). Vienna: k.u.k. Hof- und Universitäts-Buchhändler.

Schuchardt, Hugo. 1922. Heimisches und fremdes Sprachgut. *Revista Internacional de los Estudios Vascos* 13(1): 69-82.

Schulze, Wilhelm. 1895. Ἀραξάρης, λίτρα. *KZ* 33(2): 214-24.

Schulze, Wilhelm. 1933. *Kleine Schriften von Wilhelm Schulze zum 70. Geburtstag am 15. Dezember 1933*. Göttingen: Vandenhoeck & Ruprecht.

Schumacher, Stefan. 2004. *Die Keltischen Primärverben*. Innsbruck: IBS.

Schumacher, Stefan and Joachim Matzinger. 2013. *Die Verben des Altalbanischen*. Wiesbaden: Harrassowitz.

Schwarz, Martin. 2008. Iranian *L, and Some Perisan and Zaza Etymologies. *Iran and the Caucasus* 12: 281-7.

Schwenzer, Stefan. 2004. *Frühbronzezeitliche Vollgriffdolche. Typologische, Chronologische und Technische Studien auf der Grundlage der Materialaufnahme von Hans-Jürgen Hundt*. Mainz: Verlag des Römisch-Germanischen Zentralmuseums.

Schwyzer, Eduard. 1913. Zur griechischen Laut- und Wortbildungslehre. *Glotta* 5(3): 193-7.

Schwyzer, Eduard. 1930. Die Bezeichnung des Zahnfleisches in den indogermanischen Sprachen. *KZ* 57(3/4): 256-75.

Schwyzer, Eduard. 1931. Griechische Interjektionen und griechische Buchstabennamen auf -α. Mit Exkursen über die Geschichte der Buchstabennamen und des Wortes Alphabet. *KZ* 58(3/4): 170-205.

Schwyzer, Eduard. 1934. Dissimilatorische Geminatenauflösung als Folge von Übersteigerung, zunächst im Neugriechischen und im Spätgriechischen. *KZ* 61(3/4): 222-52.

Schwyzer, Eduard. 1939-50. *Griechische Grammatik*. Vol. I-II. [Vol. I 5th ed. 1977, Vol. II 4th ed. 1975]. Munich: C.H. Beck.

Sebastian, Patrizia, Hanno Schaefer, Ian R.H. Telford, and Susanne S. Renner. 2010. Cucumber (*Cucumis sativus*) and melon (*C. melo*) have numerous wild relatives in Asia and Australia, and the sister species of melon is from Australia. *PNAS* 107(32): 14269-14273.

Seidl, Christian. 2003. Les variétés du latin. In Ernst, Gerhard, Martin-Dietrich Gleßgen, Christian Schmitt, and Wolfgang Schweickard (eds.). *Romanische Sprachgeschichte / Histoire linguistique de la Romania*. Vol. I, pp. 515-530. Berlin/New York: Walter de Gruyter.

Semino, Ornella, Giuseppe Passarino, Peter J. Oefner, Alice A. Lin, Svetlana Arbuzova, Lars E. Beckham, Giovanna De Benedictis, Paolo Francalacci, Anastasia Kouvatsi, Svetlana Limborska, Mladen Marcikić, Anna Mika, Barbara Mika, Dragan Primorac, A. Silvana Santachiara-Benerecetti, Luigi Luca Cavalli-Sforza, and Peter A. Underhill. 2000. The Genetic Legacy of Paleolithic *Homo sapiens sapiens* in Extant Europeans: A Y Chromosome Perspective. *Science* 290(5494): 1155-9.

Sergent, Bernard. 1995. *Les Indo-Européens: Histoire, langues, mythes*. Paris: Payot.

Serventi, Patrizia, Chiara Panicucci, Roertha Bodega, Sara De Fanti, Stefania Sarno, Manuel Fondevila Alvarez, Francesca Brisighelli, Beniamino Trombetta, Paolo Anagnostou, Gianmarco Ferri, Antonio Vazzana, Chiara Delpino, Giorgio Gruppioni, Donata Luiselli, and Elisabetta Cilli. 2018. Iron Age Italic Population Genetics: the Piceni from Novilara (8th-7th century BC). *Annals of Human Biology* 45(1): 34-43.

Shennan, Stephen. 2018. *The First Farmers of Europe. An Evolutionary Perspective*. Cambridge: Cambridge University Press.

Sherratt, Andrew and Susan Sherratt. 1988. The Archaeology of Indo-European: An Alternative View. *Antiquity* 62: 584-95.

Shorto, Harry L. 2006. *A Mon-Khmer Comparative Dictionary*. Canberra: The Australian National University.

Sihler, Andrew L. 1995. *New Comparative Grammar of Greek and Latin*. New York/Oxford: Oxford University Press.

Silvestri, Domenico. 1997. The Italic Languages. In Giacalone Ramat, Anna, and Paolo Ramat. *The Indo-European Languages*, pp. 322-344. London: Routledge.

Simon, Zsolt. 2005. Die Etymologie von *caballus*. *Journal of Latin Linguistics* 9(1): 405-16.

Simon, Zsolt. 2021. The alleged Anatolian loanwords in Etruscan: A reconsideration. In Giusfredi, Federico and Zsolt Simon (eds.). *Studies in the languages and language contact in Pre-Hellenistic Anatolia*, pp. 227-54. Barcelona: Universitat de Barcelona Edicions.

Simon, Zsolt. fthc. Indo-European Substrates: the Problem of the Anatolian Substrate. In Kroonen, Guus (ed.). *Sub-Indo-European Europe: Continuity and discontinuity in the linguistic landscape of prehistoric Europe*. Berlin/Boston: De Gruyter Mouton.

de Simone, Carlo. 1963. VII Congresso Internazionale di Scienze Onomastiche (Review). *Indogermanische Forschungen* 68: 194-7.

de Simone, Carlo. 1968-70. *Die griechischen Entlehnungen im Etruskischen*. Wiesbaden: Harrassowitz.

de Simone, Carlo. 2018. Messapic. In Klein, Jared, Brian Joseph, Matthias Fritz, and Mark Wenthe (eds.). *Handbook of Comparative and Historical Indo-European Linguistics*. Vol. 3, pp. 1839-50. Berlin/Boston: Walter de Gruyter.

Simonet, Francisco Javier. 1888. *Glosario de voces ibéricas y Latinas usadas entre los mozárabes*. Vol. 1. Madrid: Fortanet.

Simoni, Lucia, Francesc Calafell, Davide Pettener, Jaume Bertranpetit, and Guido Barbujani. 2000a. Geographic Patterns of mtDNA Diversity in Europe. *American Journal of Human Genetics* 66(1): 262-78.

Simoni, Lucia, Francesc Calafell, Davide Pettener, Jaume Bertranpetit, and Guido Barbujani. 2000b. (Letters to the Editor) Reconstruction of Prehistory on the Basis of Genetic Data. *American Journal of Human Genetics* 66(3): 1177-9.

Sims-Williams, Nicholas. 2016. *A Dictionary: Christian Sogdian, Syriac and English*. Wiesbaden: Reichert Verlag.

Sjögren, Karl-Göran, T. Douglas Price, and Kristian Kristiansen. 2016. Diet and Mobility in the Corded Ware of Central Europe. *PLoS ONE* 11(5): e0155083.

Skeat, Walter William. 1888. *An Etymological Dictionary of the English Language*. Oxford: Oxford Clarendon.

Skeates, Robin, Jessica Beckett, Daniela Mancini, Claudio Cavazzuti, Letizia Silvestri, W. Derek Hamilton, Kerry L. Sayle, Kayla D. Crowder, Mario Rolfo, and Micaela Angle. 2021. Rethinking Collective Burial in Mediterranean Caves: Middle Bronze Age Grotta Regina Margherita, Central Italy. *Journal of Field Archaeology* 46(6): 382-98.

Skoglund, Pontus, Helena Malmström, Maanasa Raghavan, Jan Storå, Per Hall, Eske Willerslev, M. Thomas P. Gilbert, Anders Götherström, and Mattias Jakobsson. 2012. Origins and Genetic Legacy of Neolithic Farmers and Hunter-Gatherers in Europe. *Science* 336(6080): 466-9.

Skutsch, Franz. 1901-3. Altitalische Sprachen und allgemeine lateinische Grammatik. In Vollmöller, Karl (ed.). *Kritischer Jahresbericht über die Fortschritte der Romanischen Philologie* Vol. 5 (1897-8), Part 1: 52-74.

van Sluis, Paulus. 2022. Beekeeping in Celtic and Indo-European. *Studia Celtica* 56:

1-28.

van Sluis, Paulus. fthc. Substrate alternations in Celtic. In Kroonen, Guus (ed.). *Sub-Indo-European Europe: Continuity and discontinuity in the linguistic landscape of prehistoric Europe*. Berlin/Boston: De Gruyter Mouton.

van Sluis, Paulus, Anders Jørgensen, and Guus Kroonen. 2023. European Prehistory between Celtic and Germanic: the Celto-Germanic Isoglosses Revisited. In Kristiansen, Kristian, Guus Kroonen, and Eske Willerslev (eds.). *The Indo-European Puzzle Revisited*, pp. 193- 243. Cambridge: Cambridge University Press.

Smith, William. 1854. *Dictionary of Greek and Roman Geography*. Vol. 1. London: Walton and Maberly.

Smoczyński, Wojciech. 2018. *Lithuanian Etymological Dictionary*. Berlin: Peter Lang.

von Soden, Wolfram. 1965-81. *Akkadisches Handwörterbuch*. Wiesbaden: Otto Harrassowitz.

Solmsen, Felix. 1888. Sigma in Verbindung mit nasalen und liquiden im griechischen. *KZ* 29(1/2): 59-124.

Solmsen, Felix. 1894. *Studien zur lateinischen Lautgeschichte*. Strassburg: Trübner.

Solmsen, Felix. 1897. Beiträge zur geschichte der lateinischen sprache. *KZ* 34: 1-36.

Solmsen, Felix. 1898. Etymologien. *KZ* 35(4): 463-84.

Solmsen, Felix. 1901. Untersuchungen zur griechischen Laut- und Verslehre. Strassburg: Karl J. Trübner.

Solmsen, Felix. 1904. Beiträge zur geschichte der lateinischen sprache. *KZ* 37: 1-26.

Solmsen, Felix. 1911. Zur Geschichte des Namens der Quitte. *Glotta* 3(3): 241-5.

Sommer, Ferdinand. 1900. Lateinischer Vokallumlaut in haupttonigen Silben. *Indogermanische Forschungen* 11: 325-341.

Sommer, Ferdinand. 1914. *Kritische Erläuterungen zur lateinischen Laut- und Formenlehre*. Heidelberg: C. Winter.

Sommer, Ferdinand and Raimund Pfister. 1977. *Handbuch der lateinischen Laut- und Formenlehre: eine Einführung in das sprachwissenschaftliche Studium des Lateins*. Vierte, neubearbeitete Auflage. Heidelberg: Carl Winter.

Šorgo, Aljoša. 2020. Characteristics of lexemes of a substratum origin in Proto-Germanic. In Garnier, Romain (ed.). *Loanwords and substrata: proceedings of the colloquium held in Limoges (5th-7th June, 2018)*, 427-72. Innsbruck: IBS.

Specht, Franz. 1947. *Der Ursprung der indogermanischen Deklination*. Neudruck. Göttingen: Vandenhoeck & Ruprecht.

- Sperber, Hans. 1917. Rom. “alapa”. *Zeitschrift für Romanische Philologie* 38(5): 537-43.
- Stang, Christian S. 1972. *Lexikalische Sonderübereinstimmungen zwischen dem Slavischen, Baltischen und Germanischen*. Oslo/Bergen/Tromsø: Universitetsforlaget.
- Stanley, Phillip V. 1982. KN Uc 160 and Mycenaean Wines. *American Journal of Archaeology* 86(4): 577-8.
- Steinbauer, Dieter Hubertus. 1989. *Etymologische Untersuchungen zu den bei Plautus belegten Verben der lateinischen ersten Konjugation. Unter besonderer Berücksichtigung der Denominative*. (Doctoral dissertation, University of Regensburg.) Altendorf b. Bamberg: Druckerei Gräbner.
- Steinbauer, Dieter Hubertus. 1999. *Neues Handbuch des Etruskischen*. St. Katharinen: Scripta Mercaturae.
- Stifter, David. 1998. Study in Red. *Die Sprache* 40: 202–223.
- Stifter, David. 2010. The Invisible Third. The Basque and Celtic Words for ‘Swallow’. *Ériu* 60: 145-57.
- Stifter, David. 2012. On the Linguistic Situation of Roman-period Ig. In Meißner, Torsten (ed.). *Personal Names in the Western Roman World*, pp. 247-65. Hagen: Curach Bhán.
- Stifter, David. 2017. The Phonology of Celtic. In Klein, Jared, Brian Joseph, Matthias Fritz, and Mark Wenthe (eds.). *Handbook of Comparative and Historical Indo-European Linguistics*. Vol. 2, pp. 1188-1202. Boston/Berlin: De Gruyter Mouton.
- Stifter, David. 2020. Old Irish etymology through the ages. *Language & History* 63(1): 24-46.
- Stifter, David. 2023. The Rise of Gemination in Celtic. *Open Research Europe* 3(24).
- Stifter, David. fthc. Prehistoric Layers of Loanwords in Old Irish. In Kroonen, Guus (ed.). *Sub-Indo-European Europe: Continuity and discontinuity in the linguistic landscape of prehistoric Europe*. Berlin/Boston: De Gruyter Mouton.
- Stokes, Whitley. 1888. Keltic Etymologies. *Beiträge zur Kunde der indogermanischen Sprachen* 9:86-92.
- Stolz, Friedrich. 1902. Zur lateinische Sprachgeschichte. *Indogermanische Forschungen* 13: 95-117.
- Straka, Georges. 1979. *Les sons et les mots. Choix d'études de phonétique et de linguistique*. Paris: C. Klincksieck.⁵⁸⁴

⁵⁸⁴ Includes reprint of Straka, George. 1968. Contribution à la description et à l'histoire des consonnes L. *Travaux de Linguistique et de Littérature* 6(1): 267-326.

- Streitberg, W. and V. Michels. 1927. *Geschichte der indogermanischen Sprachwissenschaft seit ihrer Begründung durch Franz Bopp, 2: Die Erforschung der indogermanischen Sprachen*. Band 2, Lieferung 1. Berlin/Leipzig: Walter de Gruyter.
- Strobel, Pellegrino and Luigi Pigorini. 1864. *Le terremare e le palafitte del Parmense*. Milan: Tipi di Giuseppe Bernadoni di Gio.
- Strodach, George Kleppinger. 1933. Latin Diminutives in *-ello/a-* and *-illo/a-*: *A Study in Diminutive Formation*. (Doctoral dissertation, University of Pennsylvania.) *Language* 9(1) = *Language Dissertation No. 14*: 7-98.
- Strömberg, Reinhold. 1937. *Theophrasta: Studien zur botanischen Begriffsbildungen*. Gothenburg: Elanders Boktryckeri Aktiebolag.
- Strömberg, Reinhold. 1940. Griechische Pflanzennamen. *Göteborgs Högscolas Årsskrift* 46. Gothenburg: Elanders boktryckeri aktiebolag.
- Strong, James. 1890. *The Exhaustive Concordance of the Bible*. Cincinnati: Jennings & Graham.
- Strunk, Klaus. 1993. Lateinisch *naucum* (-i, -o) und Fragen einer Alternanz *-au/-u-*. In Meiser, Gerhard (ed.). *Indogermanica et Italica: Festschrift für Helmut Rix zum 65. Geburtstag*, 424-35. Innsbruck: IBS.
- Szemerényi, Oswald. 1952. The development of the Indo-European *mediae aspiratae* in Latin and Italic. *Archivum Linguisticum* 4: 27-53.
- Szemerényi, Oswald. 1960. Etyma Latina I. (1-6). *Glotta* 38(3/4): 216-51.
- Szemerényi, Oswald. 1974. The Origins of the Greek Lexicon: *ex oriente lux*. *The Journal of Hellenic Studies* 94: 144-57.
- Szemerényi, Oswald. 1991. The Development of the Indo-European Mediae Aspiratae in Latin and Italic. (3 parts). In Considine, P. and J.T. Hooker (eds.). *Scripta Minora: Selected Essays in Indo-European, Greek, and Latin*, pp. 628-93. Second printing. Innsbruck: IBS.
- Szemerényi, Oswald. 1995. Etyma Latina VII, 38-44. In Smoczyński, Wojciech (ed.). *Kuryłowicz Memorial Volume Part One*, pp. 409-16. Cracow: Universitas.
- Tafrahi, Orest. 1925. La cité pontique de Callatis recherches et fouilles. *Revue Archéologique* 21 : 238-292.
- Tassi, Francesca, Silvia Ghirotto, David Caramelli, and Guido Barbujani. 2013. *American Journal of Physical Anthropology* 152(1): 11-18.
- Tedesco, P. 1943. Slavic **lozã* and Persian *rãz*, both 'vine'. *Journal of the American Oriental Society* 63(2): 149-55.
- Terracini, Benvenuto. 1921. Gallico e Latino. A proposito di un recente libro sulla lingua

gallica. *Rivista di Filologica e di Istruzione Classica* 49: 401-30.

Terracini, Benvenuto. 1929. Su alcune congruenze fonetiche fra Etrusco e Italico. *Studi Etruschi* 3: 209-48.

Thomason, Sarah Grey and Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley/Los Angeles/Oxford: University of California Press.

Thorsø, Rasmus. 2020. Two Balkan Indo-European loanwords. In Serangeli, M. and T. Olander (eds). *Dispersals and diversification: linguistic and archaeological perspectives on the early stages of Indo-European*, pp. 251-62. Leiden/Boston: Brill.

Thorsø, Rasmus. fthc. Alternation of Diphthong and Monophthong in Armenian Words of Substrate Origin. In Kroonen, Guus (ed.). *Sub-Indo-European Europe: Continuity and discontinuity in the linguistic landscape of prehistoric Europe*. Berlin/Boston: De Gruyter Mouton.

Thorsø, Rasmus, Andrew Wigman, Anthony Jakob, Axel Palmér, Paulus van Sluis, and Guus Kroonen. 2023. Word Mining: Metal Names and the Indo-European Dispersal. In Kristiansen, Kristian, Guus Kroonen, and Eske Willerslev (eds). *The Indo-European Puzzle Revisited: Integrating Archaeology, Genetics, and Linguistics*, pp. 105-26. Cambridge: Cambridge University Press.

Thurneysen, Rudolf. 1882. Urspr. dt tn cn im lateinischen. *KZ* 26(2/3): 301-14.

Thurneysen, Rudolph. 1887. Lateinischer Lautwandel. *KZ* 28(1/2): 154-62.

Thurneysen, Rudolph. 1888. Zur Inschrift von Corfinium. *Rheinisches Museum für Philologie* 43(Neue Folge): 347-54.

Thurneysen, Rudolph. 1890. Vokalisches z im indogermanischen. *KZ* 30(1/3): 351-3.

Thurneysen, Rudolph. 1907. A. Walde, Lateinisches etymologisches Wörterbuch (review). *Göttingische gelehrte Anzeigen* 169(1): 795-807.

Thurneysen, Rudolph. 1921. Grammatisches und Etymologisches. *Zeitschrift für celtische Philologie* 13(1): 101-108.

Thurneysen, Rudolph. 1924. Keltisches. *Indogermanische Forschungen* 42: 143-8.

Thurneysen, Rudolph. 1946. *A Grammar of Old Irish*. Dublin: Dublin Institute for Advanced Studies.

Tischler, Johann. 1983. *Hethitisches etymologisches Glossar*. Vol. 1. Innsbruck: IBS.

Todd, Evelyn T., Laure Tonasso-Calvière, Lorelei Chauvey, Stéphanie Schiavinato, Antoine Fages, Andaine Seguin-Orlando, Pierre Clavel, Naveed Khan, Lucía Pérez Pardal, Laura Patterson Rosa, Pablo Librado, Harald Ringbauer, Marta Verdugo, John Southon, Jean-Marc Aury, Aude Perdereau, Emmanuelle Vila, Matilde Marzullo, Ornella Prato, Umberto Tecchiati,...Ludovic Orlando. 2022. The Genomic History and

Global Expansion of Domestic Donkeys. *Science* 377: 1172-80.

Tofanelli, Sergio, Francesca Brisighelli, Paolo Anagnostou, George B.J. Busby, Gianmarco Ferri, Mark G. Thomas, Luca Taglioli, Igor Rudan, Tatijana Zemunik, Caroline Hayward, Deborah Bolnick, Valentino Romano, Francesco Cali, Donata Luiselli, Gillian B. Shepherd, Sebastiano Tusa, Antonio Facella, and Cristian Capelli. 2016. The Greeks in the West: Genetic Signatures of the Hellenic Colonisation in Southern Italy and Sicily. *European Journal of Human Genetics* 24: 429-36.

Tomaschek, Wilhelm. 1880. *Centralasiatische Studien II: Die Pamir Dialekte*. Vienna: Carl Gerold's Sohn.

Topalli, Kolec. 2017. *Fjalor etimologjik i gjuhës shqipe*. Durrës: Jozef.

Torroni, Antonio, Hans-Jürgen Bandelt, L. D'Urbano, P. Lahermo, P. Moral, D. Sellitto, C. Rengo., P. Forster, M.L. Savontaus, B. Bonn -Tamir, and R. Scozzari. 1998. mtDNA Analysis Reveals a Major Late Paleolithic Population Expansion from Southwestern to Northeastern Europe. *American Journal of Human Genetics* 62(5): 1137-52.

Torroni, Antonio, Martin Richards, V. Macaulay, P. Forster, Richard Villems, S. N rby, M.L. Savontaus, K. Huoponen, R. Scozzari, and Hans-J rgen Bandelt. 2000. (Letters to the Editor) mtDNA Haplogroups and Frequency Patterns in Europe. *American Journal of Human Genetics* 66(3): 1173-7.

T tfalusi, Istv n. 2001. *Magyar etimol giai nagysz t r*. Budapest: Arcanum Adatb zis. [Database].

Trask, R.L. 2008. *Etymological Dictionary of Basque*. Edited for web publication by Max W. Wheeler. University of Sussex.

Trombetti, Alfredo. 1927. La lingua etrusca e le lingue preindoeuropee del Mediterraneo. *Studi Etruschi* 1: 213-38.

Trombetti, Alfredo. 1928. *La lingua etrusca*. Florence: Rinascimento del Libro.

Trump, David. 1966. *Central and Southern Italy before Rome*. London: Thames and Hudson.

Turner, Ralph Lilley. 1966-9. *A Comparative Dictionary of the Indo-Aryan Languages*. London: Oxford University Press.

Untermann, J rgen. 2000. *W rterbuch des Oskisch-Umbrischen*. Heidelberg: Carl Winter.

de Vaan, Michiel. 1999. The PIE root structure **Te(R)D^h*. *KZ* 112: 1-25.

de Vaan, Michiel. 2014. West Germanic **þþ* and **þm* in Dutch. *Amsterdamer Beitr ge zur  lteren Germanistik* 72: 1-34.

de Vaan, Michiel. 2018. The Phonology of Albanian. In Klein, Jared, Brian Joseph, and

Matthias Fritz (eds.). *Handbook of Comparative and Historical Indo-European Linguistics*. Vol. 3, 1732-49. Berlin/Boston: De Gruyter Mouton.

Väänänen, Veikko. 1981. *Introduction au Latin Vulgaire*. 3rd edition. Paris: Klincksieck.

Vander Kloet, S.P. 1992. On the Etymology of *vaccinium* L. *Rhodora* 94(880): 371-3.

Vander Linden, Marc. 2016. Population History in Third-Millennium-BC Europe: Assessing the Contribution of Genetics. *World Archaeology* 48(5): 714-28.

Vandkilde, Helle. 2014. Breakthrough of the Nordic Bronze Age: Transcultural Warriorhood and a Carpathian Crossroad in the Sixteenth Century BC. *European Journal of Archaeology* 17(4): 602-33.

Vaniček, Alois. 1881. *Etymologisches Wörterbuch der lateinischen Sprache*. 2. umgearb. Aufl. Leipzig: Teubner.

Valério, Miguel and Ilya Yakubovich. 2010. Semitic Word for ‘Iron’ as Anatolian Loanword. In T.M. Nikolaev (ed.). *Issledovanija po lingvistike i semiotike: Sbornik statej k jubileju Vyač. Vs. Ivanova* [Studies in Linguistics and Semiotics: A Collection of Articles for the Anniversary of Vyacheslav V. Ivanov], pp. 108–116. Moscow: Jazyki slavjanskix kul’tur.

Vasmer, Max. 1908. Etymologisches und Grammatikalisches. *Byzantinische Zeitschrift* 17: 108-120.

Vasmer, Max. 1959-61. *Ètimologičeskij slovar’ russkogo jazyka* [Etymological Dictionary of the Russian Language], translated from German and supplemented by Oleg Trubachyov. Moscow: Progress.

Vennemann, Theo. 1998. Germanica Semitica: *Biene* und *Imme*. Mit einem Anhang zu lat. *apis*. *Sprachwissenschaft* 23(4): 473-87.

Vennemann, Theo. 2003. *Europa Vasconica – Europa Semitica*. Berlin/New York: Mouton de Gruyter.

Vernesi, Cristiano, David Caramelli, Isabelle Dupanloup, Giorgio Bertorelle, Martina Lari, Enrico Cappellini, Jacopo Moggi-Cecchi, Brunetto Chiarelli, Loredana Castrì, Antonella Casoli, Francesco Mallegni, Carles Lalueza-Fox, and Guido Barbujani. 2004. The Etruscans: A Population-Genetic Study. *American Journal of Human Genetics* 74: 694-704.

Vine, Brent. 1999a. On “Cowgill’s Law” in Greek. In Heiner Eichner and Hans Christian Luschützky (eds.). *Compositiones Indogermanicae in Memoriam Jochem Schindler*, pp. 555-600. Prague: Enigma.

Vine, Brent. 1999b. Greek *πίζα* ‘root’ and ‘schwa secundum’. In Ivanov, Vyacheslav and Brent Vine (eds.). *UCLA Indo-European Studies* (vol. 1), pp. 5-30. Los Angeles: Program in Indo-European Studies.

- Vine, Brent. 2006. On 'Thurneysen-Havet's Law' in Latin and Italic. *KZ* 119: 211-49.
- Vine, Brent. 2016. Latin *bēs/bessis* 'two-thirds of an *as*'. In Byrd, Andrew Miles, Jessica DeLisi, and Mark Wenthe (eds.). *Tavet Tat Satyam. Studies in Honor of Jared S. Klein on the Occasion of His Seventieth Birthday*, pp. 324-330. Ann Arbor/New York: Beech Stave Press.
- de Vries, Jan. 1962. *Altnordisches etymologisches Wörterbuch*. 2nd edition. Leiden/Boston/Cologne: Brill.
- Vycichl, Werner. 1963. Der sumerische Ursprung von lateinisch *porrus, porrum* „Lauch“. *Die Sprache* 9: 21-2.
- Vycichl, Werner. 1983. *Dictionnaire étymologique de la langue copte*. Leuven: Peeters.
- Vycichl, Werner. 1990. *La vocalisation de la langue égyptienne*. Vol. I. Cairo: Institut Français d'Archéologie Orientale.
- Wachter, Rudolf. 2006. Rez. Lexikon des frühgriechischen Epos, Lief. 19 (Πασιδίκη - πλεῦμων). Lief. 20 (πλευραί - πῶν) (Göttingen 1999-2000). *Kratylos* 51: 136-44.
- Wagner, Heinrich. 1957. Indogermanisch-Vorderasiatisch-Mediterranes. *KZ* 75(1/2): 58-75.
- Wagner, Max Leopold. 1941. *Historische Lautlehre des Sardischen*. Halle (Saale): Max Niemeyer Verlag.
- Wagner, Max Leopold. 1953. Etymologische Randbemerkungen zu neueren iberoromanischen Dialektarbeiten und Wörterbüchern. *Zeitschrift für romanische Philologie* 69(5/6): 347-91.
- Wagner, Max Leopold. 1960-4. *Dizionario etimologico Sardo*. Heidelberg: Carl Winter.
- Walde, Alois. 1906. Aspiratendissimilation im Latein. *Indogermanische Forschungen* 19: 98-111.
- Walde, Alois. 1910. *Lateinisches etymologisches Wörterbuch*. Heidelberg: Carl Winter.
- Walde, Alois. 1917. *Über älteste sprachliche Beziehungen zwischen Kelten und Italikern*. Innsbruck: Wagner.
- Walde, Alois. 1921. Lateinische Etymologien. *Indogermanische Forschungen* 39: 74-93.
- Waldman, Carl and Catherine Mason. 2006. *Encyclopedia of European Peoples*. New York: Facts on File.
- Waldman, Nahum M. 1972. *Akkadian Loanwords and Parallels in Mishnaic Hebrew*. (Doctoral dissertation, Dropsie College.) <https://repository.upenn.edu/dropsietheses/16/>
- Wallace, Rex. 2008. *Zikh Rasna: A Manual of the Etruscan Language and Inscriptions*.

Ann Arbor/New York: Beech Stave Press.

Watmough, Joshua. 1937. *The Foundations of Roman Italy*. London: Methuen & Co.

Watmough, Joshua. 1951. *La posición del latín en el dominio indoeuropeo*. By Michel Lejeune (Review). *Language* 27(1): 81-2.

Watmough, Margaret M.T. 1997. *Studies in the Etruscan Loanwords in Latin*. Florence: Leo. S. Olschki Editore.

Weise, Fr. Oscar. 1881. Etymologische Beiträge. *Beiträge zur Kunde der indogermanischen Sprachen* 6: 233-5.

Weise, Fr. Oscar. 1882. *Die Griechischen Wörter im Latein*. Leipzig: S. Hirzel.

Weiss, Michael. 1993. *Studies in Italic Nominal Morphology*. (Doctoral dissertation, Cornell University.)

Weiss, Michael. 2010a. *Language and Ritual in Sabellic Italy*. Leiden: Brill.

Weiss, Michael. 2010b. Observations on the *Littera* Rule. Handout presented at ECIEC XXIX, Cornell University, June 19, 2010.

Weiss, Michael. 2016. Cosmogonia italica. In Ancillotti, Augusto, Alberto Calderini, and Riccardo Massarelli (eds.). *Forme e strutture della religione nell'Italia mediana antica. (Forms and Structures of Religion in Ancient Central Italy)*, pp. 661-7. Rome: "L'Erma" di Bretschneider.

Weiss, Michael. 2018. Limited Latin Grassmann's Law: Do We Need It? In Gunkel, Dieter, Stephanie W. Jamison, Angelo Mercado, and Kazuhiko Yoshida (eds.). *Vine Diem Celebrent: Studies in Linguistics and Philology in Honor of Brent Vine*, pp. 438-47. Ann Arbor: Beech Stave Press.

Weiss, Michael. 2020. *Outline of the Historical and Comparative Grammar of Latin*. 2nd edition. Ann Arbor/New York: Beech Stave Press.

Weiss, Michael. (2021, January 5). The Etymology of Latin *libra*. Society for Classical Studies Greek and Latin Linguistics Panel, Chicago, IL.

Weiss, Michael. 2022a. Italo-Celtic. In Olander, Thomas (ed.). *The Indo-European Language Family. A Phylogenetic Perspective*, pp. 102-13. Cambridge: Cambridge University Press.

Weiss, Michael. 2022b. Italic. In Olander, Thomas (ed.). *The Indo European Language Family. A Phylogenetic Perspective*, pp. 114-34. Cambridge: Cambridge University Press.

West, Martin L. 2007. *Indo-European Poetry and Myth*. Oxford: Oxford University Press.

[WH] Walde, Alois and Johann Baptist Hofmann. 1938-54. *Lateinisches etymologisches Wörterbuch*. Heidelberg: Winter.

Wharton, Edward Ross. 1890. *Etyma Latina*. London: Rivingtons.

Wigman, Andrew. fthc. A European Substrate Velar “Suffix”. In Kroonen, Guus (ed.). *Sub-Indo-European Europe: Continuity and discontinuity in the linguistic landscape of prehistoric Europe*. Berlin/Boston: De Gruyter Mouton.

Wilamowitz-Moellendorff, Ulrich von. 1931-2. *Der Glaube der Hellenen*. Berlin: Weidmannsche Buchhandlung.

Wilke, Georg. 1919. Die Herkunft der Italiker. Neue Beiträge zur Indogermanenfrage. *Archiv für Anthropologie* 45 (Neue Folge 17): 162-80.

Willi, Andreas. 2008. *Sikelismos. Sprache, Literatur und Gesellschaft im Griechischen Sizilien (8.-5. Jh. v. Chr.)*. Basel: Schwabe.

Willi, Andreas. 2012. Lateinische *scīre* und Verwandtes. *Glotta* 88: 253-72.

Willms, Lothar. 2013. *Klassische Philologie und Sprachwissenschaft*. Göttingen: Vandenhoeck & Ruprecht.

van Windekens, Albert Joris. 1952. *Le Pélasgique: Essai sur une langue indo-européenne préhellénique*. Louvain: Publications Universitaires.

van Windekens, Albert Joris. 1959. Die Herkunft von gr. *καβάλλης*, lat. *caballus* usw. *KZ* 76(1/2): 78-80.

Witczak, Krzysztof Tomasz. 1996. The Pre-Germanic Substrata and Germanic Maritime Vocabulary. In Jones-Bley, Karlene and Martin E. Huld (eds.). *The Indo-Europeanization of Northern Europe*, pp. 166-80. Washington D.C.: Institute for the Study of Man.

Witczak, Krzysztof Tomasz. 1999. Studies in Armenian Etymology (1-30). *Studia Etymologica Cracoviensia* 4: 172-190.

Witczak, Krzysztof Tomasz. 2003. *Indoeuropejskie nazwy zbóż*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

Wood, Francis A. 1910. Greek and Latin Etymologies. *Classical Philology* 5(3): 303-8.

Worrell, William H. 1934. *Coptic Sounds*. Ann Arbor: University of Michigan Press.

Zair, Nicholas. 2012. *The Reflexes of the Proto-Indo-European Laryngeals in Celtic*. Leiden/Boston: Brill.

Zair, Nicholas. 2013. Latin *glārea* ‘gravel’. *KZ* 126: 280-6.

Zair, Nicholas. 2016. Vowel weakening in the Sabellic languages as language contact. *Indogermanische Forschungen* 121: 295-315.

Zair, Nicholas. 2017. The origins of *-urC-* for expected *-orC-* in Latin. *Glotta* 93: 255-89.

Zair, Nicholas. 2018. Latin *bardus* and *gurdus*. *Glotta* 94(1): 311-18.

Zamboni, Lorenzo. 2021. The Urbanization of Northern Italy: Contextualizing Early Settlement Nucleation in the Po Valley. *Journal of Archaeological Research* 29: 387-430.

Zappacosta, S., M. De Felice, M. Fiore, and G.B. Ferrara. 1980. The HLA System in the Campania Region: A Genetic Study. *Tissue Antigens* 16(4): 286-93.

Zimmern, Heinrich. 1915. *Akkadische Fremdwörter als Beweis für babylonischen Kultureinfluss*. Leipzig: A. Edelmann.

Zinkevičius, Zigmās. 1966. *Lietuvių dialektologija* [Lithuanian dialectology]. Vilnius: Mintis.

Zohary, Daniel, Maria Hopf, and Ehud Weiss. 2012. *Domestication of Plants in the Old World*. Fourth edition. Oxford: Oxford University Press.

Nederlandse Samenvatting

Het Latijn behoort tot de Italische taalgroep, onderdeel de Indo-Europese taalfamilie. Alle Indo-Europese talen kunnen worden herleid tot een oertaal, het Proto-Indo-Europees. Deze oertaal werd vermoedelijk gesproken op de Pontisch-Kaspische steppe door een bevolking die zich rond 3000 v.Chr. hiervandaan begon te verspreiden. De voorouders van de sprekers van het Latijn verlieten de steppe dus aan het begin van het derde millennium v.Chr. en verspreidden zich door Europa, waarbij ze uiteindelijk het Apennijns schiereiland bereikten, lang voor de vroegste overlevering van het Latijn rond 700 v.Chr. In de lange periode tussen het verlaten van de steppe en de vroegste geschreven bronnen kwamen ze in contact met de bevolkingen die al in Europa woonden en talen spraken die sindsdien zijn verdwenen. Dit proefschrift onderzoekt het deel van de Latijnse woordenschat dat is ontleend aan zulke verdwenen talen en gebruikt de bevindingen om te bepalen hoe en wanneer de Italische talen Italië hebben bereikt.

Met behulp van de vergelijkende methode weten we hoe Indo-Europese talen onderling verwant zijn met regelmatige klankveranderingen, en daardoor zijn onregelmatige klankcorrespondenties en tot op zekere hoogte niet-Indo-Europese morf fonologie de betrouwbaarste indicatoren van ontlening. Ik heb deze methodologie toegepast om een lijst samen te stellen van Latijnse lexemen die waarschijnlijk niet zijn overgeërfd, maar hun oorsprong vinden in, of ontleend zijn aan de verdwenen talen waarmee het Latijn en verwante Indo-Europese buurtalen in aanraking kwamen na het uiteenvallen van het Proto-Indo-Europees.

In de onregelmatige woordenschat van het Latijn komen verschillende soorten alternanties voor in de fonologie, fonotaxis, en morfologie. Een analyse van de geografische distributie van deze alternanties levert nieuwe inzichten op. De niet-overgeërfde woordenschat van het Latijn bevat 1) een grote groep woorden die zijn ontleend in het Middellandse Zeegebied, bijna geheel gedeeld met het Grieks, en daarnaast 2) een kleinere groep leenwoorden, deels zonder Grieks vergelijkingsmateriaal, die uit een eerdere contactsituatie is voortgekomen. Sommige leenwoorden hebben wijdverspreide parallellen en lijken terug te gaan op onderling verwante niet-Indo-Europese talen, mogelijk behorend tot een taalgroep die zich door Europa heeft verspreid met de komst van de landbouw. Andere woorden zijn zwerfwoorden van onduidelijke oorsprong.

De twee grootste semantische groepen zijn plantennamen (ongeveer veertig procent van de leenwoordenschat) en diernamen (ongeveer twintig procent). Alle woorden voor vaatwerk en kookgerei hebben parallellen in het Grieks of vertonen een mediterrane distributie, net als de meeste woorden voor geweven stoffen. Dit weerspiegelt ongetwijfeld de grote culturele veranderingen die plaatsvonden tijdens en na de vestiging van de sprekers van de voorouder van het Italisch in het Middellandse Zeegebied. Tot

een groep woorden gedeeld met alleen het Keltisch en het Germaans behoren geen gedomesticeerde soorten, maar wel namen voor gebruiksvoorwerpen en wapens. Dit zou erop kunnen wijzen dat ze in een vroege culturele context zijn ontleend. Het feit dat meerdere Latijnse lexemen van niet-Indo-Europese oorsprong daarentegen naar gedomesticeerde planten verwijzen, duidt erop dat een deel van de geleende woordenschat zijn oorsprong had in een landbouwsubstraattaal. Echter, de mogelijkheid dat de taalfamilie waartoe deze taal behoorde in heel Europa aanwezig was, zowel in het Middellandse Zeegebied als daarbuiten, maakt het lastig om te bepalen waar en wanneer landbouwtermen zijn ontleend en zodoende ook waar de voorouder van de Italische talen zich op verschillende momenten in de prehistorie bevond.

Onderzoek naar prehistorisch DNA ondersteunt de steppehypothese over de oorsprong van de Indo-Europese talen. Steppe-DNA verschijnt rond 2000 v.Chr. in Noord-Italië, waarna het zich verspreidde in zuidelijke richting. Er vond een grote vermenging plaats met populaties die niet van de steppe kwamen, waardoor er een genetische gradiënt van noord naar zuid ontstond. Er zijn verschillende manieren om de aanwezigheid van de Italische taalgroep op het Apennijnse schiereiland af te leiden: 1) aan de hand van de historische verspreidingsgebieden van de verschillende Italische talen het Latijn, het Sabellisch en het Venetisch; 2) aan de hand van de verspreiding van steppe-DNA; en 3) aan de hand van de woordenschat die door het Proto-Italisch in het Middellandse Zeegebied is ontleend. Op basis hiervan valt af te leiden dat het Proto-Italisch waarschijnlijk vanuit het noordoosten het Apennijns schiereiland is binnengekomen. Een plausibel scenario is dat de voorloper van het Italisch met de touwbekercultuur geassocieerd was, welke vervolgens (door vermenging met de klokbekercultuur) overging in de Úněticecultuur en daarna in de grafheuvelcultuur, die zich verplaatste naar de Pannonische Vlakte. Deze regio stond vervolgens meerdere eeuwen in contact met de Terramarecultuur, waarvan tenminste een deel van de bevolking mogelijk al Italische talen sprak. De verdere verspreiding naar het zuiden kan wellicht verbonden worden met de verspreiding van de urnenveldencultuur in de vorm van de Proto-Villanovacultuur.

Curriculum Vitae

Andrew Michael Wigman was born on 7 June 1991 in Voorhees, New Jersey, the United States of America. He finished his secondary education at Downingtown High School East Campus in 2009. Andrew subsequently attended the Pennsylvania State University, entering the Schreyer Honors College, and in 2014 received a BSc in Archaeological Science, a BSc in World Languages Education, and a BA in Classics and Ancient Mediterranean Studies, with highest distinction. The programs included a study trip to Rome, an archaeological dig at Tel es-Safi in Israel, and a semester abroad in Vienna. He then moved to Munich on a DAAD Graduate Study Scholarship, earning his MA in Comparative Indo-European Linguistics in 2017 from the Ludwig Maximilian University. While a Master's student, Andrew spent a semester as a teaching assistant for introductory level Sanskrit as well as a year as a student assistant at the Informatics Department of the Technical University of Munich. After graduation, Andrew worked for a year at the daycare and after-school programming of the Jules Verne Campus in Munich before moving to Leiden in 2018 to start as a PhD candidate at the Leiden University Centre for Linguistics (LUCL). Under the supervision of Guus Kroonen and Kristian Kristiansen, he completed the research for the present dissertation on the European Research Council (ERC)-funded project "The Linguistic Roots of Europe's Agricultural Transition". At Leiden, Andrew was a member of the LUCL PhD council and served briefly as its chairperson. He spent a semester at Cornell University just as the world entered the Covid-19 lockdown.

Latin is one of the most important Indo-European languages in European history. Between the dissolution of Proto-Indo-European on the Pontic-Caspian steppe and the first attestation of written Latin on the Italian Peninsula, the ancestors of Latin-speakers had more than two millennia to migrate across Europe. The Europe that they entered was not empty however. It had been populated by farmers for three thousand years, and by hunter-gatherers for nearly ten thousand years before that. This dissertation investigates the lexemes in Latin that may have been borrowed from the languages that these populations spoke and combines the insights gained with lines of evidence from genetics and archaeology to hypothesize on the route that brought the ancestors of Latin-speakers into Italy.

A large portion of the prehistoric loanwords in Latin attests to its presence in the Mediterranean. But older strata from its ancestors' travels north of the Alps are also preserved, albeit in smaller quantities. They attest to contact with a language or group of languages with which the Germanic and Celtic language subgroups also had contact, suggesting they were taken up during the time of the early steppe migrations to the Northwest. Other loanwords that are present beside Latin in several other languages of Europe point to their having been borrowed from a widespread language family, perhaps that of the European farmers. Several lines of evidence point to the spread of the Italic language family into the Italian Peninsula from the Northeast, around the end of the transition of the Bronze Age into the Iron Age.

