Universiteit
Leiden
The Netherlands

The pre-Roman elements of the Sardinian lexicon

Swanenvleugel, C.

Citation
Swanenvleugel, C. (2025, February 12). *The pre-Roman elements of the Sardinian lexicon*. *LOT dissertation series*. LOT, Amsterdam. Retrieved from https://hdl.handle.net/1887/4180290

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/1887/4180290

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

The pre-Roman elements of the Sardinian lexicon

Published by
LOT
Binnengasthuisstraat 9
1012 ZA Amsterdam
The Netherlands

phone: +31 20 525 2461

e-mail: lot@uva.nl
http://www.lotschool.nl

Cover illustration: view of the Valle di Oddoene, Dorgali, from the Supramonte of Urzulei, near Ghenna ‘e Silana.

ISBN: 978-94-6093-472-8
DOI: https://dx.medra.org/10.48273/LOT0687
NUR: 616

Copyright © 2025 Cid Swanenvleugel. All rights reserved.

The pre-Roman elements of the Sardinian 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 12 februari 2025
klokke 16:00 uur

door

Cid Swanenvleugel

geboren te Delft

in 1998

Promotores:	Prof.dr. G.J. Kroonen
	Prof.dr. M.G. Kossmann
Promotiecommissie:	Prof.dr. G. Mensching (Universität Göttingen)
	Prof.dr. M.P.G.M. Mous
	Dr. G. Murgia (Università di Cagliari)
	Prof.dr. P.C.H. Schrijver (Universiteit Utrecht)

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).

Contents

List of figures xix

List of tables xxv

Acknowledgements xxvii

Abbreviations xxix

I Introduction 1

1 Sardinia and its linguistic background 3

1.1 The languages of Sardinia 3

1.2 The romanization of Sardinia 6

1.3 The Sardinian substrate 7

1.3.1 Features of the Sardinian substrate 8

1.3.1.1 The Sardinian pre-Roman lexicon 8

1.3.1.2 Pre-Roman toponyms in Sardinia 9

1.3.1.3 Substrate-induced features of Sardinian 9

1.3.2 The Mediterranean substrate 10

1.3.2.1 A note on “Ligurian” 14

1.3.2.2 A note on “Iberian” 15

1.3.3 Proposed connections between the Sardinian substrate and attested languages 17

1.3.3.1 Phoenico-Punic 17

1.3.3.2 Berber 18

1.3.3.3 Basque 19

1.3.3.4 Etruscan 20

2 Methodology 23

2.1 Identifying substrate elements 23

2.1.1 Criteria for substrate origin 24

2.1.1.1 Lack of an etymology 24

2.1.1.2 Irregular correspondence 24
2.1.1.3 Unusual phonotactics 25
2.1.1.4 Unusual morphology 26
2.1.1.5 Semantic category 26
2.1.1.6 Limited geographical distribution 27
2.1.2 Form, meaning, and chance similarities 27
2.1.2.1 Substrate origin or irregular inheritance? 28
2.1.2.1.1 A note on folk etymology 29
2.1.2.2 Comparing non-inherited words 29
2.1.2.2.1 Toponyms as evidence 30
2.1.2.2.2 Conclusion: dealing with chance similarities 31
2.2 Analyzing the data 32
2.2.1 Phonological analysis 32
2.2.2 Morphological analysis 35
2.3 Interpreting results 37
2.3.1 Unity of the substrate 38
2.3.2 Nature of language contact 39
2.3.3 Dynamics of language shift 39
2.3.4 Classifying a non-attested language 39
2.3.5 Extralinguistic arguments 40
2.4 Conclusion 40
II Linguistic data 43
3 Flora 45
3.1 Herbs 45
3.1.1 Annual mercury (*Mercurialis annua*); goosefoot (*Chenopodium sp.*) 45
3.1.2 Arum (*Arum italicum*) 49
3.1.3 Betony (*Stachys glutinosa*) 51

3.1.4 Bristly ox-tongue (*Helmintotheca echioides*), wild teasel (*Dipsacus fullonum*) 54

3.1.5 Cat thyme (*Teucrium marum*), cotton lavender (*Santolina chamaecyparissus*) 55

3.1.6 Celery (*Apium graveolens*), fool's watercress (*Apium nodiflorum*) .. 60

3.1.7 False fennel (*Ridolfia segetum*) 62

3.1.8 Fleabane, yellow (*Dittrichia viscosa*) 64

3.1.9 Germander, yellow (*Teucrium flavum*) 66

3.1.10 Hemlock (*Conium maculatum*)$_1$ 68

3.1.11 Hemlock (*Conium maculatum*)$_2$ 71

3.1.12 Mullein, great (*Verbascum thapsus*)$_1$ 72

3.1.13 Mullein, great (*Verbascum thapsus*)$_2$ 75

3.1.14 Mullein (*Verbascum sp.*) 77

3.1.15 Poppy (*Papaver rhoeas*) 79

3.1.16 Mustard$_1$ 83

3.1.17 Mustard$_2$ 84

3.1.18 Mustard, black (*Brassica nigra*) 87

3.1.19 Rue, fringed (*Ruta chalepensis*), cat thyme (*Teucrium marum*) 87

3.1.20 Rush skeletonweed (*Chondrilla juncea*), Sow thistle (*Sonchus oleraceus*) 90

3.1.21 Sea purslane (*Atriplex portulacoides*) 92

3.1.22 Teasel, wild (*Dipsacus fullonum*) 94

3.1.23 Watercress (*Nasturtium officinale*)$_1$; butcher's broom (*Ruscus aculeatus*) 96

3.1.24 Watercress (*Nasturtium officinale*)$_2$, fool's watercress (*Apium nodiflorum*) 98

3.2 Grasses, cereals and weeds 100

3.2.1 Bermuda grass (*Cynodon dactylon*); butcher's broom (*Ruscus aculeatus*), holly (*Ilex aquifolium*) 100

3.2.2 Esparto grass (*Lygeum spartum*) 104

3.2.3 Galingale (*Cyperus longus*) 107
3.2.4 Knotgrass (*Polygonum sp.*); horsetail (*Equisetum palustre*) 110
3.2.5 Ropegrass (*Ampelodesmos mauritanicus*) 110
3.2.6 Weed (unspecified) 113
3.3 Trees and shrubs 114
3.3.1 Bean trefoil (*Anagyris foetida*) 114
3.3.2 Broom, spiny (*Calicotome sp.*, *Genista sp.*) 116
3.3.3 Dog rose (*Rosa canina*) 120
3.3.4 Heather (*Erica scoparia*)$_1$ 122
3.3.5 Heather (*Erica scoparia*)$_2$, tree heather (*Erica arborea*)$_1$ 124
3.3.6 Holly (*Ilex aquifolium*)$_1$ 126
3.3.7 Holly (*Ilex aquifolium*)$_2$ 128
3.3.8 Hop-hornbeam (*Ostrya carpinifolia*) 131
3.3.9 Maple (*Acer monspessulanum*) 133
3.3.10 Osier (*Salix viminalis*) 135
3.3.11 Oleander (*Nerium oleander*) 138
3.3.12 Pine tree (*Pinus sp.*) 140
3.3.13 Rockrose (*Cistus sp.*) 142
3.3.14 Rosemary (*Salvia rosmarinus*) 144
3.3.15 Smilax (*Smilax aspera*) 146
3.3.16 Strawberry tree (*Arbutus unedo*) 148
3.3.17 Strawberry tree flower (*Arbutus unedo*) 150
3.3.18 Strawberry tree fruit (*Arbutus unedo*) 152
3.3.19 Traveller's joy (*Clematis vitalba*), fragrant virgin's bower (*Clematis flammula*) 155
3.3.20 Tree heather (*Erica arborea*)$_2$ 156
3.3.21 Willow (*Salix sp.*) 159
3.3.22 Yew (*Taxus baccata*) 159

3.3.23 Yew (*Taxus baccata*); buckthorn (*Rhamnus alaternus*); spindle (*Euonymus europaeus*) 162
3.4 Plant parts 163
3.4.1 Awn of wild oats (*Avena fatua*) 163
3.4.2 Branch 165
3.4.3 Juniper berry (*Juniperus sp.*) 168
3.4.4 Mushroom 171
3.4.5 Resin 173
3.4.6 Sprout, bud, shoot 173
3.4.7 Tree, shrub 176
3.4.8 Walnut (*Juglans regia*) 176
4 Fauna 179
4.1 Invertebrates 179
4.1.1 Ant 179
4.1.2 Backswimmer (*Notonecta glauca*) 182
4.1.3 Bed bug (*Cimex lectularius*) 182
4.1.4 Caterpillar; earwig (*Forficula auricularia*) 186
4.1.5 Clothes moth (*Tineidae*) 188
4.1.6 Cockroach (*Blatta orientalis*)$_1$; cockchafer (*Melolontha melolontha*); darkling beetle (*Blaps sp.*) 188
4.1.7 Cockroach (*Blatta orientalis*)$_2$; giant darkling beetle (*Blaps gigas*) 190
4.1.8 Cricket 194
4.1.9 Dragonfly; grasshopper 194
4.1.10 Earthworm (*Lumbricus terrestris*) 196
4.1.11 Grasshopper$_1$ 198
4.1.12 Grasshopper$_2$ 201
4.1.13 Rose chafer (*Cetonia aurata*) 202
4.1.14 Spider 204
4.1.15 Slug, snail; cicada 207

4.1.16 Tick (swollen with blood) 209
4.1.17 Worm 212
4.2 Fish 214
4.2.1 European eel (*Anguilla anguilla*) 214
4.2.2 Trout fry (*Salmo cettii*)$_1$ 216
4.2.3 Trout fry (*Salmo cettii*)$_2$ 219
4.3 Reptiles 219
4.3.1 Gecko (*Tarantola mauritanica*) 219
4.3.2 Lizard (*Podarcis tiliguerta*) 221
4.3.3 Lizard; snake 225
4.3.4 Skink (*Chalcides ocellatus tiligugu*); slug 225
4.3.5 Viperine snake (*Natrix maura*)$_1$ 228
4.3.6 Viperine snake (*Natrix maura*)$_2$ 231
4.4 Birds 231
4.4.1 Bee-eater (*Merops apiaster*) 231
4.4.2 Bird of prey 234
4.4.3 Bird species (various) 234
4.4.4 Kestrel (*Falco tinnunculus*)$_1$ 237
4.4.5 Kestrel (*Falco tinnunculus*)$_2$ 237
4.4.6 Kite (*Milvus milvus*) 240
4.4.7 Plover (*Pluvialis apricaria*) 242
4.4.8 Raven (*Corvus corax*) 242
4.4.9 Robin (*Erithacus rubecula*)$_1$; yellow wagtail (*Motacilla flava*); hare (*Lepus sp.*) 244
4.4.10 Robin (*Erithacus rubecula*)$_2$, wren (*Troglodytes troglodytes*)$_1$ 247
4.4.11 Scops owl (*Otus scops*) 248
4.4.12 Wren (*Troglodytes troglodytes*)$_2$ 250
4.5 Mammals 252
4.5.1 Bat$_1$ 252

4.5.2 Bat$_2$ 255
4.5.3 Dog 256
4.5.4 Doggie, puppy 258
4.5.5 Fawn (*Cervus elaphus corsicanus*), lamb 260
4.5.6 Fox (*Vulpes vulpes*)$_1$ 263
4.5.7 Fox, vixen (*Vulpes vulpes*)$_2$ 265
4.5.8 Hare (*Lepus sp.*) 267
4.5.9 Marten (*Martes martes*) 269
4.5.10 Mouflon (*Ovis aries musimon*) 273
5 Geography 275
5.1 Natural features 275
5.1.1 Chasm; pile 275
5.1.2 Clay 277
5.1.3 Den (for animals) 279
5.1.4 Forest 281
5.1.5 Highland 283
5.1.6 Hillock 286
5.1.7 Humid place 288
5.1.8 Marsh 291
5.1.9 Marshy valley 292
5.1.10 Pit 295
5.1.11 Plateau 299
5.1.12 Plateau, characterized by jagged rock outcroppings 299
5.1.13 Plateau, flat terrain 302
5.1.14 Plateau, small 304
5.1.15 Precipice, cliff, ravine 304
5.1.16 Puddle 306
5.1.17 Ravine, gorge 308

5.1.18 Rock, boulder 310
5.1.19 Rock crevice$_1$ 313
5.1.20 Rock crevice$_2$ 315
5.1.21 Rock outcropping$_1$ 317
5.1.22 Rock outcropping$_2$ 317
5.1.23 Rocky plain 320
5.1.24 Secluded place 320
5.1.25 Stream, torrent 322
5.1.26 Summit, in isolation 326
5.1.27 Water basin; source; stream 326
5.1.28 Whirlpool; crevice 328
5.2 Man-made structures 330
5.2.1 Border, limit 330
5.2.2 Nuraghe 332
5.2.3 Pile of rocks 334
5.2.4 Pile of rocks; dry stone wall 337
5.2.5 Source 338
6 Culture and economy 341
6.1 Agriculture 341
6.1.1 Bran 341
6.1.2 Goat or sheep that lost its young 344
6.1.3 Herd 344
6.1.4 Hornless (of goats) 347
6.1.5 Sheep (old, skinny) 347
6.1.6 Sheep (old) 349
6.1.7 Stack of wheat 352
6.1.8 Tilled land 352
6.2 Material culture and food 354

6.2.1 Bread made of acorns and clay ... 354
6.2.2 Clay used to make acorn bread ... 358
6.2.3 Clothes given to servants as a gift on certain holidays ... 358
6.2.4 Mat_1 ... 360
6.2.5 Mat_2 ... 362
7 Miscellaneous ... 367
7.1 Physical ... 367
7.1.1 Blind ... 367
7.1.2 Bristle (of pigs) ... 370
7.1.3 Cheek ... 372
7.1.4 Neck ... 375
7.1.5 Rash, eczema ... 376
7.2 Other ... 378
7.2.1 Evil spirit ... 378
7.2.2 Film, layer ... 381
7.2.3 Forked juniper trunk ... 381
7.2.4 Tangle, bundle of wood ... 384
III Structural features ... 387
8 Phonology ... 389
8.1 Non-inherited phonemes ... 389
8.1.1 Non-inherited *ϑ ... 389
8.1.1.1 Non-inherited sources of *ϑ ... 390
8.1.2 *ϑ in Punic and Berber loans ... 392
8.1.2.1 Sardinian *ϑ and Semitic sibilants ... 393
8.1.2.2 *ϑ and Berber sibilants ... 394
8.1.3 Pre-Roman *ϑ ... 396
8.1.4 Non-inherited *ts* and *č* ... 396
8.1.5 Unexpected final vowels ... 397

8.1.5.1 Final *-i* ... 397
8.1.5.1.1 Paragogic *-i* ... 399
8.1.5.1.2 Replacement of *-i* ... 400
8.1.5.2 Final *-o* ... 401
8.1.5.2.1 Paragogic *-o* ... 402
8.2 Consonant alternations ... 403
8.2.1 Quantity alternations ... 403
8.2.1.1 Quantity alternation between Sardinian and Latin ... 404
8.2.1.2 Quantity alternation within Sardinian ... 406
8.2.1.2.1 **kk ~ *k ~ *g* ... 406
8.2.1.2.2 **ll ~ *l* ... 406
8.2.2 Quality alternations ... 407
8.3 Vowel alternations ... 409
8.4 Observations on pre-Roman phonotaxis ... 411
8.4.1 Word-final consonants ... 411
8.4.1.1 Direct evidence for word-final consonants ... 413
8.4.1.2 Indirect evidence for word-final consonants ... 413
8.4.1.3 Evidence for word-final consonant clusters ... 414
8.4.2 Evidence for pre-Roman vowel harmony ... 414
8.5 Conclusion ... 417
9 Morphology ... 421
9.1 Prefixes ... 421
9.1.1 **kV-* ... 423
9.1.2 **ϑ(i)-* ... 423
9.1.2.1 **ϑVϑV-/*tsVtsV-* ... 426
9.1.2.2 **ϑur-* ... 428
9.1.2.3 **li-* ... 429
9.1.3 Pre-Roman prefixes in Sardinian ... 431

9.1.3.1 The relation between *k(V)- and *ϑ(i)- 431
9.1.3.2 The function of *k(V)- and *ϑ(i)- 431
9.2 Suffixes 432
9.2.1 *-Ví 432
9.2.2 *-Vk- 434
9.2.2.1 *-ák- 434
9.2.2.2 *-ik- 436
9.2.2.3 On the existence of pre-Roman *-Vk- 437
9.2.3 *-l- 437
9.2.3.1 *-al- 437
9.2.3.2 *-el- 438
9.2.3.3 *-il- 438
9.2.3.4 On the existence of pre-Roman *-Vl- 439
9.2.4 *-n 440
9.2.4.1 *-́Vn- 440
9.2.4.2 *-òne 440
9.2.5 *-nk- 442
9.2.6 *-r- 443
9.2.6.1 *-úri 443
9.2.6.2 *-́Vr- 445
9.2.6.2.1 Alternating *-́Vr- 448
9.2.6.2.2 Vocalism of *-́Vr- 448
9.2.6.2.3 Function of *-́Vr- 448
9.2.6.2.4 *-́Vr- outside Sardinia 449
9.2.6.2.5 Secondary productivity of -́Vr- 450
9.2.7 *-Vrr- 450
9.2.8 *-s- 451
9.2.9 *-st(r)- 452

9.3 Conclusion 453
IV External comparisons 457
10 Proposed connections to known languages 459
10.1 Punic 459
10.1.1 Lines of evidence for Punic loans 459
10.1.2 The lexical evidence 460
10.1.3 Discussion 463
10.1.3.1 Distribution of Punic words in Sardinia 465
10.1.3.2 Punic loans with a Semitic etymology 465
10.1.3.3 Punic loans with comparanda in Berber 467
10.1.4 Conclusion 468
10.2 Berber 469
10.2.1 The lexical evidence 469
10.2.2 Discussion 473
10.2.3 Conclusion 474
10.3 Basque 476
10.3.1 The lexical evidence 476
10.3.2 Discussion 482
10.3.3 Conclusion 487
10.4 Etruscan 487
10.4.1 Etruscan lexical comparisons 487
10.4.2 Etruscan structural comparisons 488
10.4.3 Comparing Tuscan 489
10.4.4 The lexical evidence 490
10.4.5 Discussion 495
10.4.5.1 Etruscan and Sardinian 496
10.4.5.2 Latin and Sardinian 496
10.4.5.3 Tuscan and Sardinian 497

10.4.6 Conclusion ... 497
11 Mediterranean comparisons ... 501
11.1 The lexical evidence ... 501
11.2 Discussion ... 513
11.2.1 Geographic distributions ... 515
11.2.1.1 Iberian Peninsula ... 515
11.2.1.2 Southern France ... 517
11.2.1.3 Corsica ... 519
11.2.1.4 Northern Italy ... 520
11.2.1.5 Central Italy ... 521
11.2.1.6 Southern Italy and Sicily ... 522
11.2.1.7 Aegean and the Balkans ... 523
11.3 Interpretation of the observed distributions ... 524
11.3.1 Iberian in Sardinia? ... 525
11.3.2 Etruscan in Sardinia? ... 528
11.3.3 Pre-Greek in Sardinia? ... 530
11.4 Discussion: a pre-Roman language family in the western Mediterranean? ... 530
11.5 Conclusion ... 532
V Conclusion ... 533
12 Conclusion ... 535
12.1 Summary ... 535
12.2 Outlook ... 536
Appendix: Overview of Sardinian developments ... 537
General developments from Latin to Sardinian ... 537
Isoglosses between the Sardinian macro-varieties ... 537
Logudorese and Nuorese ... 537
Logudorese ... 538
Nuorese ... 538

Campidanese 538
Further noteworthy developments 538
Logudorese 538
Nuorese 538
Campidanese 538
Bibliography 541
Abbreviations 541
Bibliography 542
Index 571
Basque 571
Berber 571
Gallo-Romance 572
Greek 573
Ibero-Romance 574
Italo-Romance 575
Latin 577
Sardinian 579
Semitic 593
Other 594
Nederlandse samenvatting 595
Arresùmini sardu 597
Riassunto italiano 599
Curriculum Vitae 601

List of figures

Figure 1.1.1: The languages of Sardinia (Mensching & Remberger 2016: Map 17.1; modified version of Virdis 1988: 905) 5
Figure 3.1.1: Annual mercury (*Mercurialis annua*); goosefoot (*Chenopodium sp.*, *Lipandra polysperma*) 46
Figure 3.1.2: Arum (*Arum italicum*) 48
Figure 3.1.3 Betony (*Stachys glutinosa*) 52
Figure 3.1.4: Bristly ox-tongue (*Helmintotheca echioides*), wild teasel (*Dipsacus fullonum*) 53
Figure 3.1.5: Cat thyme (*Teucrium marum*), cotton lavender (*Santolina chamaecyparissus*) 56
Figure 3.1.6: Celery (*Apium graveolens*), fool's watercress (*Apium nodiflorum*) ... 59
Figure 3.1.7: False fennel (*Ridolfia segetum*) 63
Figure 3.1.8: Yellow fleabane (*Dittrichia viscosa*) 65
Figure 3.1.9: Yellow germander (*Teucrium flavum*) 67
Figure 3.1.10: Hemlock (*Conium maculatum*) 69
Figure 3.1.11: Hemlock (*Conium maculatum*) 70
Figure 3.1.12: Great mullein (*Verbascum thapsus*) 73
Figure 3.1.13: Great mullein (*Verbascum thapsus*) 76
Figure 3.1.14: Mullein (*Verbascum sp.*) 78
Figure 3.1.15: Poppy (*Papaver rhoeas*) 80
Figure 3.1.16: Mustard 82
Figure 3.1.17: Mustard 85
Figure 3.1.18: Black mustard (*Brassica nigra*) 86
Figure 3.1.19: Fringed rue (*Ruta chalepensis*), cat thyme (*Teucrium marum*) 89
Figure 3.1.20: Rush skeletonweed (*Chondrilla juncea*), Sow thistle (*Sonchus oleraceus*) 91
Figure 3.1.21: Sea purslane (*Atriplex portulacoides*) 93
Figure 3.1.22: Teasel, wild (*Dipsacus fullonum*) 95
Figure 3.1.23: Watercress (*Nasturtium officinale*); butcher's broom (*Ruscus aculeatus*) 97
Figure 3.1.24: Watercress (*Nasturtium officinale*), fool's watercress (*Apium nodiflorum*) 99
Figure 3.2.1: Bermuda grass (*Cynodon dactylon*); butcher's broom (*Ruscus aculeatus*), holly (*Ilex aquifolium*) 102
Figure 3.2.2: Esparto grass (*Lygeum spartum*) 105

Figure 3.2.3: Galingale (*Cyperus longus*) 108
Figure 3.2.4: Knotgrass (*Polygonum sp.*); horsetail (*Equisetum palustre*) 109
Figure 3.2.5: Ropegrass (*Ampelodesmos mauritanicus*) 111
Figure 3.2.6: Weed 112
Figure 3.3.1: Bean trefoil (*Anagyris foetida*) 115
Figure 3.3.2: Spiny broom (*Calicotome sp., Genista sp.*) 117
Figure 3.3.3: Dog rose (*Rosa canina*) 121
Figure 3.3.4: Heather (*Erica scoparia*) 123
Figure 3.3.5: Heather (*Erica scoparia*), tree heather (*Erica arborea*) 125
Figure 3.3.6: Holly (*Ilex aquifolium*) 127
Figure 3.3.7: Holly (*Ilex aquifolium*) 129
Figure 3.3.8: Hop-hornbeam (*Ostrya carpinifolia*) 132
Figure 3.3.9: Maple (*Acer monspessulanum*) 134
Figure 3.3.10: Osier (*Salix viminalis*) 136
Figure 3.3.11: Oleander (*Nerium oleander*) 139
Figure 3.3.12: Pine tree (*Pinus sp.*) 141
Figure 3.3.13: Rockrose (*Cistus sp.*) 143
Figure 3.3.14: Rosemary (*Salvia rosmarinus*) 145
Figure 3.3.15: Smilax (*Smilax aspera*) 147
Figure 3.3.16: Strawberry tree (*Arbutus unedo*) 149
Figure 3.3.17: Strawberry tree flower (*Arbutus unedo*) 151
Figure 3.3.18: Strawberry tree fruit (*Arbutus unedo*) 153
Figure 3.3.19: Traveller's joy (*Clematis vitalba*) 154
Figure 3.3.20: Tree heather (*Erica scoparia*) 157
Figure 3.3.21: Willow (*Salix sp.*) 158
Figure 3.3.22: Yew (*Taxus baccata*) 160
Figure 3.3.23: Yew (*Taxus baccata*); buckthorn (*Rhamnus alaternus*); spindle (*Euonymus europaeus*) 161
Figure 3.4.1: Awn of wild oats (*Avena fatua*) 164
Figure 3.4.2: Branch 166
Figure 3.4.3: Juniper berry (*Juniperus sp.*) 169
Figure 3.4.4: Mushroom 170
Figure 3.4.5: Resin 172
Figure 3.4.6: Sprout, bud, shoot 174
Figure 3.4.7: Tree, shrub 175
Figure 3.4.8: Walnut (*Juglans regia*) 177
Figure 4.1.1: Ant 180
Figure 4.1.2: Backswimmer (*Notonecta glauca*) 181

Figure 4.1.3: Bed bug (*Cimex lectularia*)....184
Figure 4.1.4: Caterpillar; earwig (*Forficula auricularia*)....185
Figure 4.1.5: Clothes moth (*Tineidae sp.*)....187
Figure 4.1.6: Cockroach (*Blatta orientalis*); cockchafer (*Melolontha melolontha*); darkling beetle (*Blaps sp.*)....189
Figure 4.1.7: Cockroach (*Blatta orientalis*); giant darkling beetle (*Blaps gigas*)... 191
Figure 4.1.8: Cricket....193
Figure 4.1.9: Dragonfly; grasshopper....195
Figure 4.1.10: Earthworm (*Lumbricus terrestris*)....197
Figure 4.1.11: Grasshopper....199
Figure 4.1.12: Grasshopper....200
Figure 4.1.13: Rose chafer (*Cetonia aurata*)....203
Figure 4.1.14: Spider....205
Figure 4.1.15: Slug; snail; cicada....206
Figure 4.1.16: Tick....211
Figure 4.1.17: Worm....213
Figure 4.2.1: European eel (*Anguilla anguilla*)....215
Figure 4.2.2: Trout fry (*Salmo cettii*)....217
Figure 4.2.3: Trout fry (*Salmo cettii*)....218
Figure 4.3.1: Gecko (*Tarantola mauritanica*)....220
Figure 4.3.2: Lizard (*Podarcis tiliguerta*)....222
Figure 4.3.3: Lizard; snake....224
Figure 4.3.4: Skink (*Chalcides ocellatus tiligugu*); slug....227
Figure 4.3.5: Viperine snake (*Natrix maura*)....229
Figure 4.3.6: Viperine snake (*Natrix maura*)....230
Figure 4.4.1: Bee-eater (*Merops apiaster*)....232
Figure 4.4.2: Bird of prey....233
Figure 4.4.3: Bird species (various)....235
Figure 4.4.4: Kestrel (*Falco tinnunculus*)....236
Figure 4.4.5: Kestrel (*Falco tinnunculus*)....238
Figure 4.4.6: Kite (*Milvus milvus*)....239
Figure 4.4.7: Plover (*Pluvialis apricaria*)....241
Figure 4.4.8: Raven (*Corvus corax*)....243
Figure 4.4.9: Robin (*Erithacus rubecula*); yellow wagtail (*Motacilla flava*); hare (*Lepus sp.*)....245
Figure 4.4.10: Robin (*Erithacus rubecula*), wren (*Troglodytes troglodytes*)....246
Figure 4.4.11: Robin (*Erithacus rubecula*), wren (*Troglodytes troglodytes*)....249
Figure 4.4.12: Wren (*Troglodytes troglodytes*)....251

Figure 4.5.1: Bat 253
Figure 4.5.2: Bat 254
Figure 4.5.3: Dog 257
Figure 4.5.4: Doggie, puppy 259
Figure 4.5.5: Fawn (*Cervus elaphus corsicanus*), lamb 262
Figure 4.5.6: Fox (*Vulpes vulpes*) 264
Figure 4.5.7: Fox, vixen (*Vulpes vulpes*) 266
Figure 4.5.8: Hare (*Lepus sp.*) 268
Figure 4.5.9: Marten (*Martes martes*) 271
Figure 4.5.10: Mouflon (*Ovis aries musimon*) 272
Figure 5.1.1: Chasm; pile 276
Figure 5.1.2: Clay 278
Figure 5.1.3: Animal den 280
Figure 5.1.4: Forest 282
Figure 5.1.5: Highland 284
Figure 5.1.6: Hillock 285
Figure 5.1.7: Humid place 289
Figure 5.1.8: Marsh 290
Figure 5.1.9: Marshy valley 293
Figure 5.1.10: Pit 296
Figure 5.1.11: Plateau 298
Figure 5.1.12: Plateau, characterized by jagged rock outcroppings 300
Figure 5.1.13: Plateau, flat terrain 301
Figure 5.1.14: Small plateau 303
Figure 5.1.15: Precipice, cliff, ravine 305
Figure 5.1.16: Puddle 307
Figure 5.1.17: Ravine, gorge 309
Figure 5.1.18: rock, boulder 311
Figure 5.1.19: Rock crevice 312
Figure 5.1.20: Rock crevice 314
Figure 5.1.21: Rock outcropping 316
Figure 5.1.22: Rock outcropping 318
Figure 5.1.23: Rocky plain 319
Figure 5.1.24: Secluded place 321
Figure 5.1.25: Stream, torrent 323
Figure 5.1.26: Isolated summit 325
Figure 5.1.27: Water basin; source; stream 327
Figure 5.1.28: Whirlpool; crevice 329

Figure 5.2.1: Border, limit 331
Figure 5.2.2: Nuraghe 333
Figure 5.2.3: Pile of rocks 335
Figure 5.2.4: Pile of rocks; dry stone wall 336
Figure 5.2.5: Source 339
Figure 6.1.1: Bran 342
Figure 6.1.2: Goat or sheep that lost its young 343
Figure 6.1.3: Herd 345
Figure 6.1.4: Hornless (of goats) 346
Figure 6.1.5: Old, skinny sheep 348
Figure 6.1.6: Old sheep 350
Figure 6.1.7: Stack of wheat 351
Figure 6.1.8: Tilled land 353
Figure 6.2.1: Bread made of acorns and clay 355
Figure 6.2.2: Clay used to make acorn bread 357
Figure 6.2.3: Clothes given to servants as a gift on certain holidays 359
Figure 6.2.4: Mat 361
Figure 6.2.5: Mat 363
Figure 7.1.1: Blind 368
Figure 7.1.2: Bristle 371
Figure 7.1.3: Cheek 373
Figure 7.1.4: Neck 374
Figure 7.1.5: Rash, eczema 377
Figure 7.2.1: Evil spirit 379
Figure 7.2.2: Film, layer 380
Figure 7.2.3: Forked juniper trunk 382
Figure 7.2.4: Tangle, bundle of wood 383
Figure 9.1.1: Distribution of forms containing **k(V)-* 422
Figure 9.1.2: Distribution of forms containing **ʔi-* showing an alternation 425
Figure 9.1.3: Distribution of forms containing **ʔVʔV-* 427
Figure 9.1.4: Distribution of forms containing **ʔili-* 430
Figure 9.2.1: Distribution of forms with **-úri* 444
Figure 9.2.2: Distribution of forms with **-́Vr-* showing an alternation 447
Figure 10.1.1: Distribution of punicisms in Sardinian 466
Figure 10.2.1: Distribution of Berberisms in Sardinian 475
Figure 10.3.1: Distribution of possible Sardinian-Basque cognates 485
Figure 10.4.1: Distribution of accepted Etruscan and Tuscan comparisons 498

Figure 11.3.1: Distribution of pre-Roman words shared between Sardinia, Iberia, and France 527
Figure 11.3.2: Distribution of pre-Roman words shared between Sardinia, central Italy, and France 529

List of tables

Table 8.1.1: Pre-Roman words containing *ϑ ... 392
Table 8.1.2: Punic loans containing *ϑ ... 393
Table 8.1.3: Irregular occurrences of *ts* and *č*. ... 397
Table 8.1.4: Sardinian words ending in *-*i* ... 399
Table 8.1.5: Words ending in *-*o* ... 402
Table 8.2.1: Words with voicing/length alternations in consonants ... 404
Table 8.2.2: Quantity alternations between Sardinian and Latin ... 404
Table 8.2.3: Plosive quantity alternations within Sardinian ... 406
Table 8.2.4: Quantity alternation involving *l* ... 406
Table 8.2.5: Words showing consonant quality alternations ... 408
Table 8.3.1: words exhibiting vowel alternations ... 410
Table 8.4.1: Words with potential evidence for a word-final consonant ... 413
Table 8.4.2: Pre-Roman words containing three identical consecutive vowels ... 416
Table 8.4.3: Words with two consecutive vowels alternating in unison ... 417
Table 9.1.1: Forms containing *kV- ... 423
Table 9.1.2: Forms containing *ϑ(*i*)- ... 424
Table 9.1.3: Forms containing *ϑVϑV- ... 426
Table 9.1.4: Forms containing *ϑur- ... 428
Table 9.1.5: Forms containing *ϑili- ... 429
Table 9.2.1: Lexical forms containing -V́i̯ ... 433
Table 9.2.2: Forms containing *-*ák*- ... 435
Table 9.2.3: Forms containing *-*ik*- ... 436
Table 9.2.4: Forms containing *-*al*- ... 438
Table 9.2.5: Forms containing *-*el*- ... 438
Table 9.2.6: Forms containing *-*ile* ... 439
Table 9.2.7: Forms containing *-́*Vn*- ... 440
Table 9.2.8: Forms containing *-*òne* ... 441
Table 9.2.9: Forms containing *-*nk*- ... 442
Table 9.2.10: Forms containing *-*úri* ... 443
Table 9.2.11: Forms containing unstressed *-́*Vr*- ... 446
Table 9.2.12: Non-Sardinian comparanda to Sardinian forms with *-́*Vr*- ... 449
Table 9.2.13: Forms containing *-*Vrr*- ... 450
Table 9.2.14: Forms containing *-*Vs*- ... 451
Table 9.2.15: Forms containing *-*st*(*r*)- ... 452
Table 9.3.1: Lexical evidence for pre-Roman morphemes ... 454

Table 10.1.1: Proposed punicisms in Sardinian. 465
Table 10.2.1: Proposed Berberisms in Sardinian 474
Table 10.3.1: Proposed Sardinian-Basque comparisons 483
Table 10.4.1: Proposed lexcial evidence for Tyrrhenian Sardinian contact 496
Table 11.2.1 Distribution of proposed pre-Roman comparanda 515
Table 11.2.2: Pre-Roman Sardinian comparanda in the Iberian Peninsula 517
Table 11.2.3: Pre-Roman Sardinian comparanda in southern France 519
Table 11.2.4: Pre-Roman Sardinian comparanda in Corsica 520
Table 11.2.5: Pre-Roman Sardinian comparanda in northern Italy 521
Table 11.2.6: Pre-Roman Sardinian comparanda in central Italy 522
Table 11.2.7: Pre-Roman Sardinian comparanda in southern Italy and Sicily 523
Table 11.2.8: Pre-Roman Sardinian comparanda in the Aegean and Balkans 524

Acknowledgements

This dissertation is the product of a ten-year journey that began when I moved to Leiden to study Indo-European linguistics at the age of seventeen. While my name appears on the cover, its completion would not have been possible without the support and guidance of many, only a few of whom I am able to acknowledge here.

First and foremost, I wish to express my deep gratitude to my supervisors Guus Kroonen and Maarten Kossmann. Guus, as the supervisor of my bachelor thesis, enabled me to pursue my interest in the Romance languages, even though these are not central to "traditional" Indo-European research. He later provided me with the invaluable opportunity to join the EUROLITHIC research project, first as an intern, then as a student-assistant, and finally as a PhD candidate. His support and confidence in my abilities have been of immeasurable value. Maarten has helped me expand my interests beyond Europe, across the Mediterranean. His indispensable assistance in Berber etymologies and his exceptional talent for demarcating research topics and methodology have been of great help throughout the project.

I also want to thank my teachers at Leiden University especially Alwin Kloekhorst, Michaël Peyrot, Sasha Lubotsky, and Tijmen Pronk. Their outstanding teachings have laid the foundation for my development as a linguist. I extend special appreciation to Lucien van Beek, for giving me the valuable opportunity to acquire teaching experience.

Several people have played important roles, in various ways, throughout my PhD project, and I would like to acknowledge them in roughly chronological order. I thank Laura Dees for her companionship, for keeping me grounded during the pandemic, and for giving me the courage to embark on my academic journey. She, Clemens Mayer, and Jurriaan Wiegertjes were the best classmates I could have wished for, and I look back fondly on the study sessions, dinners, and drinks we shared.

My time at the Leiden Center for Linguistics has been unforgettable, thanks to the many friends I made along the way. My utmost appreciation goes out to my wonderful office mates and fellow project members: Andrew Wigman and Paulus van Sluis at the "western" desk, Anthony Jakob, Axel Palmér and Rasmus Thorsø at the "eastern" desk. I will miss our thought-provoking discussions and

transcendent whiteboard mind mapping sessions. I am equally grateful to my other colleagues, who have become dear friends: Ahmed Sosal, Bente De Graeve, Carmen Spiers, Eleonora Settembrino, Hester Groot, Jesse Wichers Schreur, Jiang Wu, Louise Friis, Matthew Sung, Natasja Delbar, Niels Schoubben, Olga Nozdracheva, Saskia Dunn, Sophia Nauta, and many others. Their presence has made the LUCL feel like a second home to me. My sincere thanks also go to my paranymph Abel Warries, one of my oldest friends in Leiden.

A crucial experience during my project was the semester I spent at the University of Cagliari. I extend my heartfelt thanks to Michele Loporcaro for his invaluable assistance in facilitating this. I am deeply grateful to Giulia Murgia, who so warmly welcomed me to Cagliari and introduced me to the fundamentals of Sardinian philology and linguistics. Gràtzias fintzas a Bruna Sìriu e Càrminu Pintore po m'imparai su pagu sardu chi chistionu. I also wish to thank my classmates Carlo, Giovanna, Michela, and Sandra for their kindness, and to Alessandro and his family for giving me a home during my stay. I am greatly indebted to Rita Porqueddu for her help in accessing essential sources in the University library of Cagliari. My time in Cagliari was greatly shaped by the wonderful friends I met: Asia, Basia, Carolina, Eli, Gabriel, Ibrahim, Johanna, Laura, Maks, Marta, Martin, Mikołaj, Muriel, Nina, Nona, Sanyi, Theodora, Tiago, Tilman, Yana, and many others. Thank you all for exploring Sardinia with me and for making my time there one of the most memorable experiences of my life. To you, I raise an *Uweschluss*.

Many others have contributed to this dissertation, be it by providing constructive criticism, expressing interest, or simply by their friendship. For reasons of space, most of them must remain unnamed, but they include my housemates and friends from Collegium Musicum. However, I would like to extend a special thanks to Tristan van Laar for his affection and for being there for me in moments of doubt. Some old friends have become constants in my life: Bram, Brenda, George, Ellen, Ingo, Jesper, Job, Leopold, Lieke, Maarten, Marla and Ramon. I am deeply grateful for always being able to rely on them. Finally, my most profound gratitude goes to my family, particularly to my grandparents Anneke, Jan, Math and Mieke, my siblings Fee and Gebre, and my parents Bart and Ilja. Zoonder eur oonaafgebroke leefde, sjteun en geluif in mich haw ich nwats zwa wied kinne kome. 'ne Hjelle grwatte merci.

Abbreviations

abl.	ablative	Diosc.	Attested in Dioscorides
Abruzz.	Abruzzese		
acc.	accusative	Du.	Dutch
Aeg.	Aegean	Etr.	Etruscan
Akk.	Akkadian	For.	Forézien
Alb.	Albanian	form.	formal
Algh.	Algherese Catalan	Frpr.	Francoprovençal
alt.	alternation	Gal.	Galician
Ar.	Arabic	Gall.-Rom.	Gallo-Romance
Arag.	Aragonese	Gallur.	Gallurese
Ast.	Asturian	Gasc.	Gascon
Auv.	Auvergnat (Occitan)	Gaul.	Gaulish
Balk.	Balkan Peninsula	Georg.	Georgian
Barb.	Barbagia	Gr.	ancient Greek
Bb.	Berber	Gr. Hes.	Hesychian gloss in Greek
Bearn.	Bearnais		
Bov.	Bovese (Greek dialects around Bova in south Calabria)	Gm.	Germanic
		Ib.	Iberian Peninsula
		Ib.-Rom.	Ibero-Romance
Bq.	Basque	IE	Indo-European
Bret.	Breton	Kb.	Kabylian
Calab.	Calabrese	Lang.	Lengadocian
Camp.	Campidanese	Lad.	Ladin
Campan.	Campanian	Lat.	Latin
Cat.	Catalan	Laz.	Laziale
CIt.	Central Italy	Leon.	Leonese
cogn.	cognacy	Lith.	Lithuanian
comp.	comparanda	Log.	Logudorese
Cors.	Corsican	MGr.	Modern Greek
dep.	département (French administrative unit)	Nap.	Neapolitan
		Navarr.	Navarro-Aragonese
		Nef.	Nefusi
deriv.	derivation	NIt.	Northern Italy
dial.	dialectal	NLog.	North Logudorese
dim.	diminutive	Nuor.	Nuorese

Ogl.	Ogliastran	Salent.	Salentino
OIr.	Old Irish	Sav.	Savoyard
OHG	Old High German	Sem.	Semitic
OProv.	Old Provençal	semant.	semantic
OSrd.	Old Sardinian	SFr.	Southern France
OSp.	Old Spanish	sg.	singular
PC	Proto-Celtic	Sic.	Sicilian, Sicily
PGm.	Proto-Germanic	SIt.	Southern Italy
PIE	Proto-Indo-European	Sl.	Slavic
		Snh.	Senhaja
Pg.	Portuguese	Sp.	Castilian Spanish
Phoen.	Phoenician	Srd.	Sardinian
Pied.	Piemontese	Tmz.	Central Moroccan Berber
pl.	plural		
PN	personal name	Top.	toponym
Prov.	Provençal	Trent.	Trentinan
PSl.	Proto-Slavic	Tsh.	Tashelhiyt
Pugl.	Pugliese	Tuar.	Tuareg
Pun.	Punic	Tun.	Tunisian Arabic
reconst.	reconstruction	Tusc.	Tuscan
Rif.	Tarifit	Val.	Valencian
Rom.	Romanian	Ven.	Venetian

I
Introduction

The aim of this work is to increase our understanding of the linguistic prehistory of pre-Roman Sardinia. Traces of extinct languages spoken in Sardinia before its romanization are preserved in the toponymy and lexicon of the modern Sardinian language varieties, which descend from Latin. This study focuses on the lexical evidence. It consists of five parts, corresponding to the research objectives that will allow us to shed light on the nature and affiliation of the languages of pre-Roman Sardinia.

- Part I provides an overview of the history of research on the pre-Roman languages of Sardinia (§ 1) and establishes a methodology for the analysis and interpretation of the lexical evidence (§ 2).
- Part II presents and discusses the relevant lexical evidence, organized by semantic field (§ 3 – 7).
- Part III examines the lexical evidence for structural properties of pre-Roman Sardinian, specifically in the phonological (§ 8) and morphological (§ 9) domains.
- Part IV aims to uncover the linguistic affiliation of pre-Roman Sardinian, by evaluating a number of previously proposed hypotheses on this topic (§ 10), and by analyzing the geographic distributions of pre-Roman cognates found within and outside Sardinia (§ 11).
- Part V summarizes and synthesizes the findings into an updated view of our understanding of the linguistic prehistory of Sardinia, and proposes possible further avenues of research (§ 12).

1 Sardinia and its linguistic background

Sardinia is the second-largest island of the Mediterranean Sea. Its history has been defined by its relative isolation as an island, as well as by contacts with the three landmasses surrounding it: the Italian Peninsula, the Iberian Peninsula, and North Africa. Currently an autonomous region of Italy, Sardinia has belonged to different spheres of influence at different times. Most of these have left linguistic traces in one way or another. In chronological order, Sardinia was part of, or influenced by, the following foreign political powers: Carthage, Rome, the Vandal Kingdom, the Byzantine Empire, the Pisan and Genovese merchant republics, the Crown of Aragon, the Spanish Empire, the House of Savoy, and finally the Italian state. By far the strongest influence was exerted by the city of Rome, as evidenced by the linguistic romanization of Sardinia. With the possible exception of the Vandal Kingdom, all other aforementioned political powers have left lasting imprints, albeit in varying degrees, on the Sardinian language (see Putzu 2019 for an overview).[1]

Before the Carthaginian and Roman dominance, Sardinia had been inhabited for multiple millennia. The prehistoric Sardinians left a rich archeological record, but we have no direct evidence of the language(s) they spoke. Although Sardinia was eventually completely romanized linguistically, traces of the native pre-Roman language(s) survived in the Sardinian lexicon and toponymy. These linguistic substrate traces have been the subject of many scholarly investigations since the early 20th century (see § 1.3). The present study aims to contribute to this field of research, by synthesizing and evaluating the lexical evidence for the Sardinian substrate, its phonological and morphological features, and its linguistic affiliation.

1.1 The languages of Sardinia

The native language of most of Sardinia is Sardinian, which is the direct continuation of Latin spoken in Sardinia. As such, it is considered a separate language or language branch in the Romance family tree (Bossong 2016: 65; Mensching & Remberger 2016: 270; Lupinu 2023: 27). Modern Sardinian is characterized by considerable internal variation. Its varieties are usually classified into a number of dialect groups. The main distinction is between Logudorese in the north and Campidanese in the south. Nuorese in the center-east is either considered a

[1] For an overview of the history of Sardinia, see Guidetti (1987).

sub-variety of the former, or a variety of its own. In the west, the Arborense dialects are transitional between Logudorese and Campidanese (Virdis 1988; Mensching & Remberger 2016: 270; Molinu & Floricic 2017). Overviews of the Sardinian language, its linguistics, and its philology, can be found in Wagner (1997 [1950]), Atzori (1982), Contini & Tuttle (1982), Blasco Ferrer (1984a), Jones (1997), Bossong (2008: 225–245), Mensching & Remberger (2016), Blasco Ferrer, Koch & Marzo (2017), and Lupinu (2023). The appendix to this volume contains an overview of some of the major historical phonological developments of Sardinian as a whole, and of its individual dialect groups.

Besides Sardinian, other Romance languages are spoken in smaller areas: the Sardo-Corsican dialects of Sassarese and Gallurese in the north (Maxia 2017), Catalan in and around the city of Alghero (Dessì Schmid 2017), and the Tabarchino variety of Ligurian in the Sulcis archipelago in the southwest (Toso 2017)(see Figure 1.1.1). These language varieties have gained a presence in Sardinia in post-classical times and are a testimony to its diverse linguistic contacts throughout history.

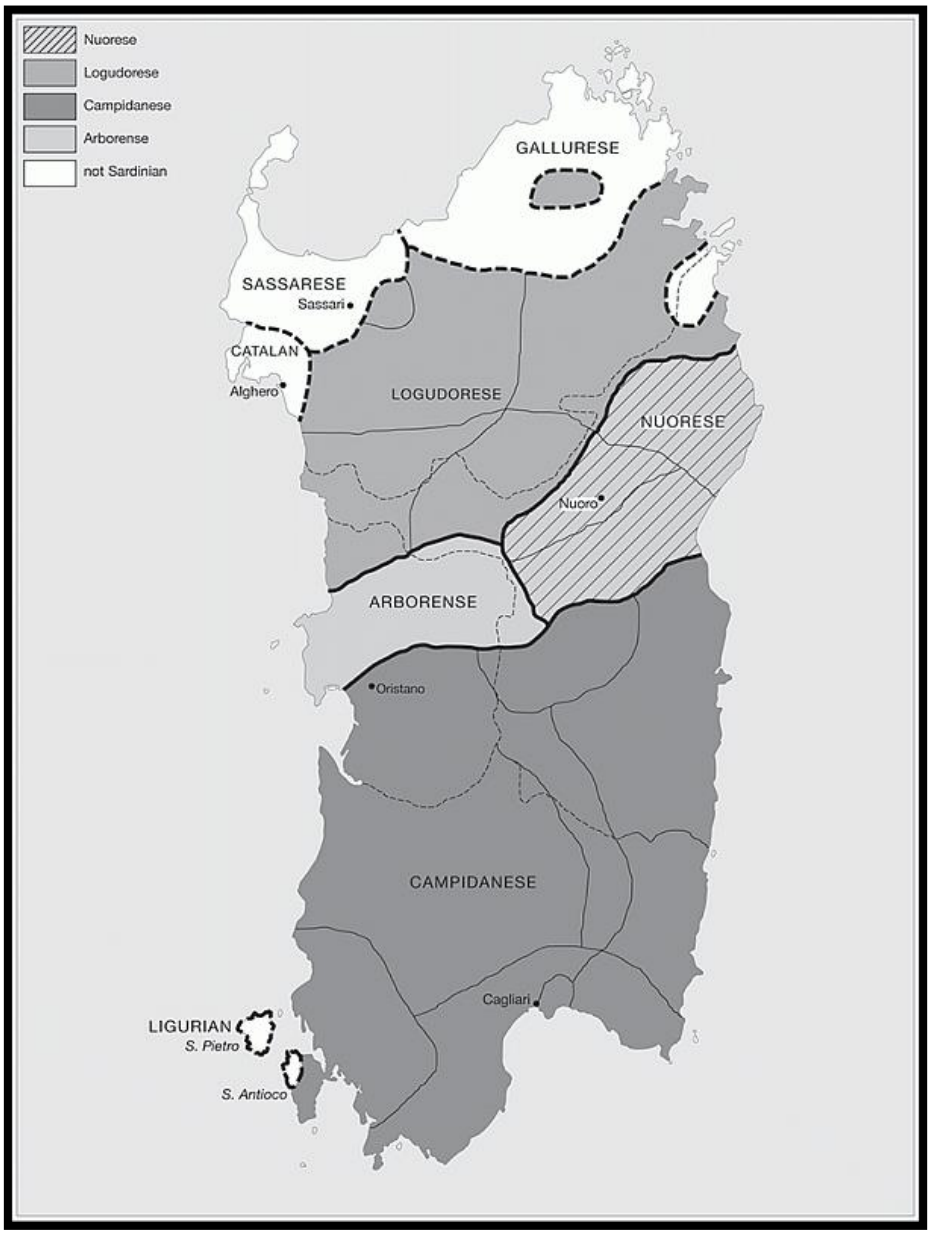

Figure 1.1.1: The languages of Sardinia (Mensching & Remberger 2016: Map 17.1; modified version of Virdis 1988: 905)

1.2 The romanization of Sardinia

In 238 BCE, Sardinia was annexed by the Roman republic, after which it entered a process of administrative, cultural, and linguistic romanization. This was a gradual process. It is clear that the Romans did not abolish the institutions and infrastructure left behind by the Carthaginians, but rather integrated them into the new political order (Meloni 1987). Roman authority was established relatively quickly in the more densely populated fertile plains in the south and west of the island. The more central mountainous regions have been argued to have resisted significantly longer (Dyson 1985; Rowland 1985).[2]

Various hypotheses have been posited regarding the linguistic romanization of Sardinia. Wagner (1928; 1997: 105–120) argues for an early romanization of Sardinia on the basis of certain lexical archaisms. He further notes lexical variation between northern and southern Sardinia (e.g. *éliɣe* < Lat. *ēlex* next to *íliži* < Lat. *īlex* 'holm-oak' etc.), which he interprets as resulting from two separate waves of romanization. The first of these would have reached all of Sardinia, while the second was more limited to the south (Wagner 1928: 8). However, this does not explain why some of the more archaic formations are found in the south, and the innovative ones in the north: e.g. Camp., Barb. *fórru* < Lat. *fornus* vs. Log., Nuor. *fúrru* < Lat. *furnus* 'oven' (Wagner 1928: 8–14). Blasco Ferrer (1989: 26) disagrees with Wagner's (1928) two-step romanization, but does agree that an initial wave of romanization reached all the way into the mountainous center of the island (Blasco Ferrer 1989: 16–23). He dates the Latin of this first wave to the time of Plautus and Terence (i.e. 3rd–2nd century BCE), and assumes it to be closely related to the Latin of Africa (Blasco Ferrer 1989: 34–40). Later waves of Latin would have continued to influence Sardinia's more accessible regions, with some particular local sources of innovation at various points in time (Blasco Ferrer 1989: 74–76). Thus, Wagner and Blasco-Ferrer agree that the presumed conservative nature of the dialects of Sardinia's mountainous center is the result of the first wave of romanization, after which it developed in relative isolation. This stance is also taken by Lupinu (2005), who essentially follows Wagner.[3]

A different model is proposed by Virdis (2018: 12–25; 2020), who argues that the Campidanese dialects do in fact preserve some more archaic features than the

[2] But recently more nuanced views have been put forward (Trudu 2012; Farre 2016; Canu 2016).

[3] On the question of the archaicity of Sardinian as a whole, see Mensching (2004).

northern Sardinian dialects.[4] Virdis argues that the Campidanese dialects therefore continue a more archaic stage of Latin, indicative of an earlier time of romanization, even if the modern Logudorese-Nuorese dialects are more conservative.[5] An important realization in Virdis' hypothesis is that greater conservativism cannot be equated with greater archaicity of language varieties, or vice versa. This crucial distinction is absent from Wagner's (1928) and Blasco Ferrer's (1989) accounts. It is in any case clear that the linguistic romanization of Sardinia was not a single process, as it was initiated and completed at different times in different places.[6]

1.3 The Sardinian substrate

The romanization of Sardinia was at the same time its Indo-Europeanization. While the starting date of 238 BCE is early in terms of the former, it is comparatively late with regard to the latter. It has long been noted that the Sardinian toponymy and lexicon contain numerous elements that cannot satisfactorily be etymologized within Latin (cf. Wagner 1932; 1997: 254–280 for an overview). It is probable that at least a part of these stem from the language or languages spoken on Sardinia at the time of the Roman conquest. These non-inherited elements have been extensively studied by various authors, often in the context of substrates in the wider geographical region. For a detailed history of the study of linguistic substrates in Sardinia and elsewhere in the Mediterranean, see Devoto (1955), Serra (1956), Paulis (2008), and especially Craddock's (1969: 18–47) critical review of almost two centuries of literature on the topic. In what follows I will present an overview of the most important literature regarding the Sardinian substrate specifically. Previous work on the Sardinian substrate can roughly be divided into three categories, all of which have contributed to the ongoing debate regarding Sardinia's linguistic prehistory in different ways:

1. Studies that focus on the specificities of the non-inherited linguistic material within Sardinia, without necessarily comparing it to phenomena found outside Sardinia. These are discussed in § 1.3.1.
2. Studies that observe and compile similarities in non-inherited linguistic material between Sardinia and various other areas. This category com-

[4] The examples include some of the lexical variants discussed by Wagner (1928), as well as phonological and morphological features of Campidanese.

[5] Cf. Mastino (1993), who shows that literacy, and probably fluency in Latin, was lower in the center than in the coastal areas of Roman Sardinia.

[6] The attestations of Punic inscriptions well into the imperial period is one indication that the definitive romanization must have been a lengthy process (Adams 2003: 209).

prises the bulk of the studies in the earliest phases of substrate research. An overview of these studies is given in § 1.3.2, and the relevant lexical evidence is discussed in § 11.

3. Studies that argue for the identification of the pre-Roman language of Sardinia as a relative of some other, attested, language. These present a clearly defined linguistic scenario that can be tested by examination of the data. The various proposals in this category are discussed in § 1.3.2.1 and in § 10.

Various terms have been used to refer to the pre-Roman languages of Sardinia. Most common is the term "Paleo-Sardinian". The term "Nuragic", based on the eponymous archeological culture, is used specifically by Pittau (1984; 1995). Both of these terms may give the impression that we are dealing with a single, homogenous language. While this is possible (if we exclude Punic), it is by no means certain (cf. § 2.3.1). As such, I prefer to use the more neutral, be it less compact, term "pre-Roman Sardinian" to refer to the non-inherited features of the modern Sardinian varieties of putative substrate origin.

1.3.1 Features of the Sardinian substrate

The language-internal analysis of possibly non-inherited features does not immediately shed light on the linguistic affiliations of any pre-Roman language in Sardinia, but it does allow us to differentiate between what should and what should not be compared, i.e. which features may indeed be of pre-Roman origin, and which are due to more recent developments. Moreover, a detailed understanding of the distribution of substrate features is essential in recovering the pre-Roman linguistic dynamics within Sardinia.

1.3.1.1 The Sardinian pre-Roman lexicon

With the exception of Punic loans (§ 1.3.2.1), only a single word from a pre-Roman language is directly attested and identifiable; i.e. the form *NURAC* (§ 5.2.2), inscribed on the nuraghe *Aidu Entos* in Bortigali (Gasperini 1992: 303–306). As a result, our main source of evidence for language contact between Sardinian Latin and other, local languages, is the lexicon of the Romance varieties of Sardinia. Collections of Sardinian lexical material of proposed pre-Roman origin can be found in Wagner (1932; 1997: 254–280) and Hubschmid (1953). Important etymological sources regarding the Sardinian lexicon are Wagner's *Dizionario etimologico sardo* (DES), and Paulis' (1992) work on Sardinian plant names. The phonological and morphological features of Sardinian lexical items of pre-Roman origin, have usually been discussed per individual etymon, across

many studies (cf. § 8 and § 9). No complete overview and evaluation of all proposed pre-Roman phenomena in the lexicon has as of yet been given. In chapter § 8, I discuss phonological features found in words of presumed pre-Roman origin, and evaluate which of these can be projected back to a pre-Roman language. In chapter § 9, the evidence for proposed pre-Roman morphological features in the Sardinian lexicon is reviewed.

1.3.1.2 Pre-Roman toponyms in Sardinia

Another important line of evidence are Sardinian place names, many of which have been argued to be of pre-Roman origin. Overviews can be found in Serra (1956: 181–198; 1960) Paulis (1987; 2008), Wolf (1998a), and Pittau (2013). From Meyer-Lübke's (1925) discussion of pre-Roman toponyms of the Iberian Peninsula onwards, various scholars have relied on toponymic evidence to posit phonological or morphological properties of pre-Roman Sardinian languages as well (e.g. Terracini 1927; Bertoldi 1928; Serra 1960; Hubschmid 1963a; Wolf 1998a; Paulis 2008), or to argue for certain linguistic affiliations of the substrate (Bertoldi 1928; Serra 1954a; Hubschmid 1963a; Pittau 1995; Blasco Ferrer 2017). For methodological reasons outlined in § 2.1.2.2.1, this study will focus primarily on the lexical evidence, rather than toponymic evidence.

1.3.1.3 Substrate-induced features of Sardinian

Besides lexical and toponymic elements of potential pre-Roman origin, it has been proposed that some of the phonological characteristics of Sardinian too might have arisen by influence from one or more local substrate languages. Most commonly discussed in this regard are the following phenomena:

- Loss of inherited **f* in the Barbaricine dialects (e.g. Bertoldi 1937a: 169; Wagner 1997: 285; Iribarren Argaiz 1997: 338; Paulis 2008: 34; Blasco Ferrer 2010a: 85).
- Development of inherited **k* as *ʔ* in the Barbaricine dialects (e.g. Wagner 1997: 285).
- Vowel prothesis before inherited initial **r-* in Campidanese (e.g. Bertoldi 1937a: 169; Wagner 1997: 285; Iribarren Argaiz 1997: 338; Paulis 2008: 34; Virdis 2020: 35).

Both the loss of **f-* and the rise of a prothetic vowel before initial **r-* have been seen as evidence for the presence of a Basque or Iberian substrate in Sardinian (Bertoldi 1937a: 169; Blasco Ferrer 2010a: 85; Virdis 2020: 35). Wagner (1997: 285) compares the presence of *ʔ* (as the outcome of **k* in various Sardinian dialects) to the phonological systems of Afro-Asiatic languages.

However, both the loss of **f* and the development **k* > *ʔ* have been shown be of more recent nature in the Barbaricine dialects, and are thus not informative on the phonological characteristics of any substrate language (cf. Paulis 1984a: lxxxi–lxxxv, xc–xcvii; Wolf 2011: 613; Loporcaro 2013: 57). The idea of vowel prothesis before initial **r-* due to some Iberian substrate influence is more resilient (e.g. recently Virdis 2020: 35). Basque does not allow word-initial *r-*, adding a prothetic vowel in Latin and Romance loans (e.g. Bq. *errege* 'king' << Lat. *regem*), a feature that is probably shared with ancient Aquitanian and Iberian (Michelena 1990: 332–333). Proving that the Sardinian phenomenon must be caused by substrate influence from a pre-Roman language of the Iberian Peninsula is difficult. Paulis (1984a: xxx fn. 45) notes that similar phenomena are found elsewhere in Romance too, and restrictions against initial /r/ are cross-linguistically common (Labrune 2021). It is therefore possible that the development of a prothetic vowel before **r-* in Campidanese is simply the result of Sardinian-internal phonological developments.[7] A pre-Roman Iberian origin of this phenomenon remains speculative.

1.3.2 The Mediterranean substrate

The study of the pre-Roman linguistic situation in Sardinia cannot be seen separately from the study of such linguistic substrates in the wider geographical region. From the early 20th century onwards, numerous studies have been conducted on the pre-Roman elements in Romance languages. Especially in the early stages of substrate research, many proposals revolved around the concept of a Mediterranean unity, in cultural as well as linguistic terms, before the Indo-Europeanization of Europe (cf. Ribezzo 1919: 104; 1920: 83; Bertoldi 1928: 222; Terracini 1936: 59). Initially, the hypothesized prehistorical linguistic connections between various regions were to a large extent based on the similarities between certain toponyms, as well as on traditional views on the origins of peoples recorded by Greek and Roman authors (e.g. Philipp 1920: 2486; Ribezzo 1920: 87–97; Terracini 1927; Bertoldi 1928).[8] In most studies on this "Mediterranean" substrate, Sardinia has had a prominent role (cf. Devoto 1955; Craddock 1969: 18–47; Wigman 2023: 18–20). As the details of the Mediterranean hypothesis have always remained somewhat vague with regard to the concrete historical and linguistic scenarios that are to be inferred, it is no surprise that diverging proposals have been put forward to account for the Sardinian attested data. The

[7] Cf. Paulis (1984a: xxxi–xxxii) for a possible scenario.

[8] On the use of toponyms as evidence for substrate connections, cf. § 2.1.2.2.

proposals usually involve "layers" or "currents" of linguistic influence, connecting different regions in the Mediterranean in varying chronological orders.

One of the first hypotheses that were firmly grounded in linguistic analysis is Terracini's (1927: 145–146). On the basis of an analysis of the endings found in Sardinian toponyms, he argues that Sardinia was "intersected by two currents in particular: one that unites it with Africa and Iberia, and another that seems to have its center in the northern Tyrrhenian region",[9] the first of which would have been stronger and more durable.[10] Terracini (1927: 146) believes that these "Afro-Iberian" toponyms in Sardinia, which would possibly be related to Punic presence in Sardinia, should be able to be analyzed with the help of Berber and Libyan. Bertoldi (1928) discusses pre-Roman suffixes and roots which he identifies in various Sardinian toponyms. Largely in agreement with Terracini, he posits a "Libyco-Sardo-Iberian" current (Bertoldi 1928: 241). Later, he connects this to the Libyan and Iberian migrants and mercenaries in Sardinia described by Pausanias (X: 17; Bertoldi 1937a: 148, 151).[11] He also interprets Terracini's "northern Tyrrhenian current" as an east-to-west "Anatolian-Aegean-Iberian" current (Bertoldi 1928: 250; 1939: 100). The idea behind this "Anatolian-Aegean-Iberian current" is based on a combination of "Ibero-Sardinian", "Pyrenaean-Alpine", and "Pyrenaean-Caucasian" comparisons, all of which are compared to Sardinian toponyms as well (Bertoldi 1928: 249–250). With regard to the pre-Roman substrate of Italy, Bertoldi (1936: 296; 1939: 100) makes a distinction between an "Aegean-Anatolian" element introduced with the (putative) Etruscan colonization, and a native "Italo-Rasennic" element that was somehow assimilated to Etruscan due to their mutual linguistic affinity.[12] This is contradictory to his idea that the original pre-Roman language of Sicily (i.e. Sicanian) was related to languages in Iberia and Sardinia, suggesting a connection to his "Anatolian-Aegean-Iberian current", but that it was unrelated to Etruscan (Bertoldi 1937a: 169). Unfortunately, it is not made clear how the idea of all these independent linguistic "currents" relates to the idea of a single Mediterranean-wide substrate, to which Bertoldi (1928: 222) subscribes. In later articles, Bertoldi takes the idea of a Mediterranean

[9] "[…] intersecata particolarmente da due correnti: una che l'unisce all'Africa e all'Iberia, l'altra che pare aver il suo centro nel Tirreno settentrionale."

[10] Cf. Bottiglioni (1929: 101), who identifies toponyms of Iberian, Ligurian, and Etruscan origin in Corsica, agreeing well with Terracini's (1927: 145–146) "north Tyrrhenian" idea.

[11] Bertoldi (1937a: 151–152) relates this to Wagner's (1932: 223–224) proposed correspondences between the Sardinian substrate and Berber languages.

[12] This is especially curious because Bertoldi's "Rasennic" is derived from Etr. *Rasenna*, the Etruscan endonym denoting Etruria (Bonfante & Bonfante 1983: 218).

substrate as a given (e.g. Bertoldi 1929; 1930), without specifying the extent of this phenomenon, or the historical reality that would have been behind it. In yet other articles (Bertoldi 1931; 1932; 1933), numerous lexical comparisons are made, without a clear conclusion as to the "current" in which the posited wide-spread word families would have originated (cf. Craddock 1969: 37–39 for a discussion of Bertoldi's various strata). In contrast to Bertoldi's intricate scenarios, Alessio returns to a more clearly defined view of the Mediterranean substrate as a genealogical linguistic unit (e.g. Alessio 1935: 141–143; 1944: 105), although he allows for internal dialectal variation (Alessio 1944: 107–108). However, this does not mean that Alessio's hypotheses are more concrete or easier to verify, and he does in fact operate with an additional, more recent substrate layer in Sardinia that would have its provenance in North-Africa (Alessio 1957: 263). In general, Alessio frequently revisits Bertoldi's comparisons and adds less convincing comparanda to the equation. I agree with Craddock (1969: 42) that "Alessio leaves even the impartial reader with a far less favorable impression than in the case of Bertoldi".

Most specific, but also most complex, in terms of stratification of the substrate lexicon, are the studies by Hubschmid, who also dedicated considerable attention to the Sardinian substrate specifically (Hubschmid 1953). His comparisons of putative pre-Roman forms are, if not convincing, supported by an impressive amount of lexical material from and by detailed formal, semantic, and etymological analysis of said material. More than Bertoldi and Alessio, Hubschmid attempts to explain the observed linguistic facts with historical scenarios — although often of speculative nature (e.g. Hubschmid 1950a: 225–226; 1950b: 133). Hubschmid does not depart from the *a priori* assumption of a unified Mediterranean substratum, but posits various pre-Indo-European language families that would at various times have expanded and receded in the Mediterranean region. Among his hypotheses are that of a "Eurafrican" substrate, and a more recent and wide-spread "Hispano-Caucasian" substrate, both of which would have had some relation to Basque (Hubschmid 1953: 91; 1960a: 24, 40; 1965: 156). The Eurafrican substrate is based mainly on comparisons between Sardinian, Berber and Basque, as well as on comparanda from Gallo-Romance and Insular Celtic. The Hispano-Caucasian substrate includes comparisons between words from Sardinian, Basque, Occitan, Alpine Romance varieties, Greek, Caucasian languages, even reaching Burushaski (Hubschmid 1953: 109). These two language families (as Hubschmid sees them) would have caused intricate contact situations with each other and with yet other languages at various times in the past. This becomes clearest in Hubschmid (1963a: 147–148):

"In Sardinia we can assume at least six Paleo-Sardinian linguistic strata:

1) an oldest Eurafrican substrate, which in the south encompassed Northwest-Africa, and that reached until Lusitania in the west, to Gaul in the north, to Italy in the east;
2) an Iberian substrate, related to Eurafrican, but stemming rather from Hispania;
3) a Hispano-Caucasian substrate stemming from the east, which reached until Lusitania in the west, Gaul in the north, the Caucasus in the east, and which in the south probably encompassed parts of North-Africa too;
4) a Tyrrhenian or Etruscan substrate, stemming from the east as well;
5) a Libyan "Hamitic" substrate related to Berber, or an early Semitic substrate;
6) perhaps another, unknown, very old Paleo-Sardinian substrate."[13]

On the whole, no amount of linguistic data adduced by the substratists discussed has been able to lend wide-spread acceptance to any configuration of the Mediterranean hypothesis. The main problem is that the scenarios proposed are so all-encompassing and abstract, and at the same time so vaguely defined, that they are almost impossible to prove or disprove — even if some of their proposals may someday prove to be close to the truth. This is exacerbated by the lack of a clearly delineated methodology as to how formally and semantically different two comparanda are allowed to be until the comparison is rejected. Another weakness is the heavy reliance on toponyms as evidence.[14]

Even if the resulting hypotheses fail to convince, the large amounts of lexical data gathered from a wealth of disparate sources especially by Bertoldi, Alessio and Hubschmid, are impressive in their own right, and I have made grateful use of these scholars' efforts in this study. The many cross-Mediterranean comparisons with relevance to the Sardinian substrate that have been proposed are discussed in § 11. On the basis of the geographic distribution of the accepted comparisons, I discuss conclusions that can potentially be drawn on any prehistoric linguistic connections in § 11.4 – 11.5.

---

[13] (Translation mine).

[14] See § 2.1.2 for a discussion of these methodological issues.

1.3.2.1 A note on "Ligurian"

The notion of an ancient "Ligurian" language often recurs in the context of Mediterranean substrate elements. Classical sources mention a people called *Liguri* in Latin, Λίγυες in Greek. Even in antiquity, ideas on the ethnological, political, and linguistic identity of the *Liguri*, varied and sometimes contradicted each other. Some sources place *Liguri* in the Iberian Peninsula, others between the Iberian Peninsula and the Rhône, and yet others east of the Rhône into northern Italy (see Bats 2019 on the use of Gr. Λίγυες). In addition, there are mentions of a Ligurian presence in Corsica and in Sicily (Cuno 1878: 90–150; Müllenhoff 1892: 171–193; Kretschmer 1905: 108–109; Mullen 2013: 27). In the 19th century, the *Liguri* were widely held to be a pre-Indo-European people (Kretschmer 1905: 109–110).[15]

There are no inscriptions that can be attributed with certainty to a language by the name of Ligurian. The sole evidence of a language spoken by the *Liguri* (provided that the term *Liguri* indeed refers to the speakers of a single language) is limited to personal names and toponyms (Mullen 2013: 26). This very limited record has paradoxically not stopped authors from indulging in various types of far-reaching speculations. Classifications of the Ligurian language have ranged from pre-Indo-European to a separate branch of Indo-European to a variety of Celtic. Cuno (1878: 91) is the first to identify the *Liguri* as a Celtic people. Based on inscriptional evidence, Kretschmer (1905: 99, 127) favored classification as an independent Indo-European branch, close to Celtic.[16] The idea of Ligurian being an unattested Indo-European language is continued in more recent scholarship (e.g. Hubschmid 1953: 116–117; Paulis 1992: 305; Dworkin 2012: 20; Mullen 2013: 29). Most frequently however, the term "Ligurian" features as a more general label referring to substrate phenomena roughly in the area of the southwestern Alps, i.e. modern-day Liguria and surroundings (e.g. Battisti 1928; 1960; Bertoldi 1928; 1937b; Alessio 1944; 1949; Ambrosini 1983). It is not always clear whether Ligurian is used purely as a geographic indicator (as in Alessio 1957: 238), or as a specific hypothetical pre-Indo-European language (as in Bertoldi 1937b: 140). It is sometimes proposed that there was a connection between Ligurian and Iberian (Bertoldi 1928: 229; Alessio 1935: 141–142; Hubschmid 1953: 116), or that Ligurian was also found in Sicily (e.g. Alessio 1944: 96 fn. 15). Notably, Ligurian has also

[15] This idea was partly rooted in contemporary racial theories (cf. Kretschmer 1905: 109).

[16] However, the inscription on the basis of which Kretschmer (1905: 99, 127) classifies Ligurian as an Indo-European language is now considered to be Lepontic (cf. Stifter et al. 2022: VB·3.1).

been mentioned as one of the linguistic substrates of Sardinia (e.g. Alessio 1944: 96 fn. 15; Devoto 1955).

In short, both in ancient and in modern times there has been considerable disagreement on the geographic localization of the *Liguri* or Λίγυες, and on the linguistic classification of these people's unattested language. Mullen (2013: 26–27) rightly calls for caution in attributing any specific linguistic or historical reality to the term "Ligurian". As such, the term "Ligurian" has lost much of its usefulness as a label for substrate phenomena, regardless of whether these substrates are of Indo-European or pre-Indo-European nature.

1.3.2.2 A note on "Iberian"

Like in the case of "Ligurian", the term "Iberian" is frequently used as a label for certain substrate phenomena by proponents of the Mediterranean substrate hypothesis. "Iberian-Sardinian" substrate features have been posited time and time again (Forsyth Major 1893: 154; Wagner 1907a: 408; 1932: 238; 1997: 259; Bertoldi 1928: 229, 238–239; Terracini 1936: 67; Bertoldi 1937b: 144, 151, 160, 169; Alessio 1944: 263; Devoto 1955: 219; Battisti 1960: 377–378; Hubschmid 1960a: 7; Paulis 1992: 21). These have usually been explained as testifying to a migration from the Iberian Peninsula to Sardinia in prehistoric times (e.g. already Wagner 1907a: 408). In this context, frequent reference is made to Pausanias' (X: 17: 5) remark that there was at some point an Iberian migration to Sardinia led by Norax (Frazer 2012a: 523 [1898]). It is difficult to verify if this semi-mythical migration could explain the linguistic facts.

Unlike Ligurian, however, Iberian refers to an actually attested language. Iberian was a non-Indo-European language, attested in ca. 2300 inscriptions, found from the southeastern coast of Spain to the Languedoc region in France (Moncunill & Velaza 2020: 593). De Hoz (2009; 2011: 56–57) argues that Iberian was spoken natively in the southeastern part of the Iberian Peninsula, and that it expanded to the northeast as a commercial lingua franca. Velaza (2019: 162) on the other hand argues that Iberian was in fact spoken as a native language across the area where it is attested, even if it would have expanded just prior to the earliest inscriptional evidence. Whatever the case, there is onomastic evidence documenting the existence of yet another, non-Indo-European language spoken in the northeastern part of the area where Iberian is attested (Untermann 1969). Untermann notes structural similarities of non-Iberian personal names in Iberian inscriptions in southwestern France to personal names recorded more eastward along the modern French Mediterranean coast. Despite being aware of the issues surrounding that term (cf. § 1.3.2.1), he labels this onomastic phenomenon "Ligu-

rian" (Untermann 1969: 8–9). On the basis of this non-Iberian onomastic material, Gorrochategui and Vallejo (2019: 360) posit the existence of an "autochthonous language in Roussillon and Catalonia, different from Iberian". Beyond the onomastic evidence, lexical substrate connections between Iberian and Ligurian (and Sardinian) have been suggested (Kretschmer 1905: 109–110; Battisti 1928: 672; Alessio 1935: 141–142; Terracini 1936: 67).

Iberian was long thought to be related to Basque (Orduña 2019: 219–221). However, with the decipherment of the Iberian script came the realization that Basque is irrelevant to understanding Iberian texts (Orduña 2019: 223). Basque and Iberian do not appear to be closely related, or in fact related at all. Any structural similarities may be explained as resulting from language contact (Gorrochategui 1995; Trask 1995: 79).[17] The conflation of Iberian and Basque as constituting a single "substrate layer" is present in many publications by proponents of the Mediterranean substrate hypothesis (e.g. Bertoldi 1937b; Alessio 1944: 109 fn. 114; Hubschmid 1953: 117–118; Battisti 1960: 352). It persists even in more recent scholarship (e.g. Blasco Ferrer 2017: 81). However, the fact that Iberian and Basque were not (closely) related implies that comparisons of pre-Roman linguistic features of Sardinian to Basque (§ 10.3) cannot be used as direct evidence in favor of an Iberian migration to Sardinia. Conversely, pre-Roman correspondences between Sardinian and languages in the Iberian Peninsula other than Basque (§ 11.2.1.1) are no direct evidence in favor of the presence of a Vasconic language in pre-Roman Sardinia.

Paulis (2008) notes some similarities between features of the Sardinian substrate and of the Iberian language. The phonological constraints that appear to operate in the pre-Roman toponyms of the Barbagia di Ollolai, as described by Wolf (1998a: 28–35), are similar to those observed in Iberian, as well as to those reconstructed by Gorrochategui and Lakarra (2001) for Proto-Basque (Paulis 2008: 32–33). Additionally, various Iberian personal names are reminiscent of medieval Sardinian personal names of probable pre-Roman origin (Paulis 2008: 39). Paulis (2008: 48) interprets these structural similarities, as well as the handful of Sardinian-Basque lexical correspondences, as evidence for an Iberian origin of the pre-Roman language of Sardinia. In his view, the linguistic ancestors of Basque, Iberian, and the Sardinian substrate would have coexisted in the Iberian Peninsula ca. 10 000 – 9 000 years ago. This time depth is roughly twice that of the

[17] In general, Vasconists seem to be less optimistic about a genetic link, at whatever time-depth, between Basque and Iberian than Iberianists (e.g. Trask 1995: 77–80; cf. Orduña 2019: 221).

Indo-European language family, however, making it unlikely that superficial linguistic features as personal names would be recognizably preserved into the attested descendent languages. Moreover, this date appears to primarily be based on genetic insights (such as Vona & Calò 2006; Caramelli et al. 2007) that have since been surpassed due to the rapid developments in the study of ancient DNA (cf. Calò et al. 2021).

1.3.3 Proposed connections between the Sardinian substrate and attested languages

Apart from the unattested substrata of unknown affiliation (cf. § 1.3.2), various known languages or relatives thereof have been proposed to be spoken in Sardinia prior to the Roman conquest. These include Punic (§ 10.1), Berber (§ 10.2), Basque (§ 10.3), and Etruscan (§ 10.4). With the exception of Punic, for which the historical context is well-understood, each of these languages figures prominently in Mediterranean substrate comparisons (§ 1.3.2), as part of larger "currents" of linguistic influence. In more recent times, attempts have been made to explain all or most of the Sardinian substrate lexicon as coming from a language related to Berber (Argiolas 2012; 2020), Basque (Blasco Ferrer 2010a; 2010b; 2011; 2017), or Etruscan (Pittau 1995). These hypotheses are mutually exclusive. In § 10, the relevant evidence for each of these hypotheses is reviewed.

1.3.3.1 Phoenico-Punic

Of all proposed connections between the Sardinian non-inherited lexicon and some non-Indo-European language, Punic is the most well-established. The first Phoenician settlements emerged in Sardinia in the eighth century BCE (Bondì 1987a; 1987b). These eventually came under Carthaginian influence from about 500 BCE (Bondì 1987c), until Sardinia was conquered by the Roman Republic in 238 BCE. Phoenician and Punic inscriptions in Sardinia are attested from the ninth century BCE until the second (or early third) century CE (Bondì 1987d; Adams 2003: 209):[18] a presence of more than a millennium with centuries of continuity into the Roman period. Given these facts, the presence of Phoenician or Punic elements in non-inherited Sardinian is expected. The first attempts at identifying these were made by Spano (1872). Unfortunately, Spano sees a Canaanite or Phoenician etymology behind almost any Sardinian word. His etymologies can almost invariably be dismissed (cf. Wagner 1957: 37). Wagner (1932: 212; 1997: 158–161; 1957: 104–109) instead proposes a Phoenico-Punic origin for a

[18] There may in fact be a Punic graffiti, in the Roman alphabet, on a hypogeum in San Salvatore di Cabras dated to the era of the Constantinian imperial dynasty — i.e. the 4th century CE (Meloni 1987: 290).

small number of lexical items,[19] hypothesizing that Phoenician and Punic linguistic influence was restricted to the plains of Campidano and Planargia (1997: 158).[20] The corpus of Phoenico-Punic etymologies in the Sardinian lexicon has been expanded by Paulis (1990). While Paulis accepts more evidence, he also rejects some of Wagner's earlier proposals at the same time. A difficulty in analyzing the material is that not all accepted punicisms in Sardinia have a clear Semitic etymology. In fact, Argiolas (2010: 23–24) has suggested some to be of Berber origin. The evidence for Punic loans in Sardinian is discussed in § 10.1.

1.3.3.2 Berber

Multiple ancient sources mention a migration of "Libyans" or "Africans" from North-Africa to Sardinia; e.g. Pausanias X: 17: 2 (Frazer 2012a: 523), Cicero XIX: 42, 45 (Cicero 1984: 19).[21] In linguistics too, various scholars have operated with an "African" or "Libyan" component within the Sardinian substrate, typically informed by lexical comparisons to Berber (e.g. Terracini 1927; 1936; Bertoldi 1928: 241; 1937b: 148, 151; Wagner 1932: 224; 1997: 263–263, 280; Serra 1954a; 1960; Hubschmid 1953: 91; 1960a: 24, 40; 1965: 156; Alessio 1957: 263). Recently, Argiolas (2012; 2020) has argued for a Berber substrate in Sardinia. Given the possibility that some Berber linguistic material entered Sardinia in a Punic colonial context (Paulis 2008: 42; Argiolas 2010: 23–24), a key question is whether it is necessary to posit an independent North-African substrate, predating Punic presence in Sardinia. This, and the proposed Berber-Sardinian comparisons, are discussed in § 10.2.

On a practical note, there have been different conventions for the phonological representation of forms from the Berber languages. Wherever relevant, I have attempted to represent these in a uniform fashion. Wherever the notation used here differs from that used by the source cited, the original notation is given in a footnote. Forms found in sources that do not employ a linguistic transcription method, are cited between angle brackets (i.e. <…>). For the reconstruction of Proto-Berber, I adhere to the system used by Múrcia Sànchez (2011), which is the main source of the reconstructed Berber forms cited.

---

[19] See also Friedrich's (1957) mildly positive evaluation of Wagner's proposals.

[20] Wagner also discusses possible Phoenician origins of a range of toponyms (Wagner 1932: 208–212; 1997: 152–157; 1957: 38–43, 78–104), which are outside the scope of this study. Cf. also Paulis (1990: 625–639).

[21] Cf. Frazer (2012b: 320–321 [1898]) for a discussion.

1.3.3.3 Basque

Being the sole surviving pre-Indo-European language in Europe, it is no surprise that Basque is a major suspect when it comes to the identification of non-inherited linguistic material in the languages of Europe. While it is in itself reasonable to assume that the Basque language at some point in prehistory had relatives that were spoken in the wider region, it does not follow that all pre-Indo-European languages were related to Basque. In fact, even the Basque language's closest neighbor in antiquity, Iberian, appears to be genetically unrelated to it, despite being a pre-Indo-European language too (cf. Gorrochategui 1995; Trask 1995: 79).[22] For the Sardinian-Basque connections posited in the framework of the "Mediterranean substrate hypothesis", see § 1.3.2.

During the last decades, a number of authors have discussed apparent connections between Basque and Sardinian. Santo Moreno (2000: 174) interprets the putative Sardinian-Basque lexical correspondences as pre-Roman *Wanderwörter*, rather than remnants of extinct language families. Bengtson (2017: 86–88) sees the pre-Roman Sardinian substrate as part of a larger Basque-Caucasian macrofamily. Paulis (2008: 20) notes some possible similarities of medieval Sardinian personal names to Aquitanian and Iberian personal names.

In a recent account, the relation between Basque and the pre-Roman language(s) of Sardinia is explored highly concretely by Eduardo Blasco Ferrer. In a series of publications, he argues that the pre-Roman toponymic evidence on Sardinia demonstrates an ancient migration of Vasconic speakers to Sardinia (Blasco Ferrer 2010a; 2010b; 2011; 2017). Partly based on Wolf's (1998a) observations on recurring patterns in Barbaricine toponyms of pre-Roman origin, he divides synchronically opaque Sardinian toponyms into segments supposedly related to Basque lexical roots. For instance, the Ogliastran toponyms *Dolai*, *Dodoliai* and *Iridolai*, analyzed as **dol-ai*, **do-dol-iai* and **iri-dol-ai*, in his view, all three contain a cognate of Bq. *odol* 'blood' (Blasco Ferrer 2011: 88).

Blasco Ferrer's methodology, like his conclusions, are deeply problematic. First and foremost, it is impossible to establish the *Benennungsmotive* of the hundreds of pre-Roman Sardinian toponyms. As a consequence, the meanings attributed to them are exclusively based on the proposed Basque etymologies and therefore speculative to the extreme. Additionally, our understanding of the derivational processes behind the formation of these toponyms is limited to the observation

[22] See Trask (1995) for an overview of the many languages that Basque has been proposed to be related to.

that some phonological sequences seem to reoccur. This essentially does not inform us on anything as long as there is no way to securely identify them as morphemes. As a result, Blasco Ferrer's (2011: 88) analysis of a toponym like *iri-dolai* as < **iri-dol-ai* is entirely arbitrary. As such, the forms and meanings of the "morphemes" that are identified in Sardinian toponyms and compared to Basque are the outcome of speculative assumptions that are neither provable nor falsifiable on the basis of the evidence.

What is worse, the Proto-Basque phonological system as reconstructed by Michelena (1990) is dated to around two thousand years ago (Trask 1995: 67). This post-dates both the Phoenician and Roman arrivals on Sardinia, and also post-dates the arrival of "paleo-Basque" speakers on Sardinia between the 9th and 3rd millennium BCE, if we follow Blasco Ferrer (2010a: 162; 2017: 80–81). Even if we depart from Lakarra's (2009; 2016) internally reconstructed earlier phase of Proto-Basque, there are multiple millennia in which virtually no language change must have occurred, neither in the direct ancestor of Basque, nor in its hypothetical sister language on Sardinia.

In short, Blasco Ferrer's view of pre-Roman Sardinia as linguistically Vasconic can be considered unsubstantiated, due to its reliance on morphologically and semantically opaque toponyms.[23] This does not mean that a prehistoric migration from Iberia to Sardinia, or any other type of connection between pre-Roman Sardinian and Basque is ruled out; it simply does not follow from the material adduced by Blasco Ferrer. Any claim of prehistoric ties between Sardinia and Iberia must be supported by lexical evidence, including both form and meaning. Semantic scrutiny is the only tool we have to minimize the risk of chance correspondence between two or more phonologically similar forms (cf. § 2.1.2).[24] The proposed lexical correspondences between Sardinian and Basque are discussed in § 10.3.

1.3.3.4 Etruscan

Like Basque, Etruscan was a non-Indo-European language spoken amidst Indo-European languages. Unlike Basque, Etruscan died out around the 1st century CE (Wallace 2008: 1). As for most of the apparently pre-Indo-European languages

[23] Cf. Wolf (2011) and Loporcaro (2013) for a more in-depth discussion of Blasco Ferrer's methodology.

[24] Although Blasco Ferrer's views do not seem to have gained widespread popularity in linguistic circles, they have unfortunately found their way into some genetic studies (e.g. Chiang et al. 2018: 1431).

that survived long enough to be attested, including Basque, Minoan, Tartessian etc., the linguistic affiliation of Etruscan has long puzzled scholars (Wallace 2008: 215–218). At this point, the only languages that are commonly accepted to be related to Etruscan are the Rhaetic language in the Alps, and the Lemnian language on the Greek island of Lemnos (Rix 1998: 60; Wallace 2008: 218–225). Regarding the eventual provenance of this "Tyrrhenian" language family, there are two main hypotheses. One is that the Etruscans as well as the Rhaetians originated in Anatolia, of which Lemnian would be a local vestige (Schachermeyr 1929; Beekes 1993; 2003; Van der Meer 2004; Kloekhorst 2022a). The other is that Etruscan and Rhaetic were native to Italy, and Lemnian a colonial offshoot of Etruscan (De Simone 1998: 409–410; Wallace 2008: 222; Oettinger 2010; Simon 2021: 245).[25] With regard to pre-Roman Sardinia, the idea of (one of) its languages being related to Etruscan is by no means absurd. Sardinia and Etruria are geographically in close proximity, and an Etruscan presence has been demonstrated in Corsica (Wallace 2008: 3). Still, any hypothesis of a Tyrrhenian language being spoken in pre-Roman Sardinia must be supported by compelling linguistic evidence.

The scholar that has argued in favor of a connection between the pre-Roman language(s) of Sardinia and Etruscan most fervently, is Pittau. Pittau (1984: 18–20; 1995; 1997: 29) argues that the pre-Roman language of Sardinia ("Nuragic", in his definition), was a sister language to Etruscan, and as such part of the Tyrrhenian language family. Nuragic and Etruscan would have been introduced to Sardinia and Etruria respectively ca. 1 000 BCE from Anatolia (Pittau 1995: 147, 171–172). More specifically, the putative introduction of Tyrrhenian in these places, as well as in Corsica, the Balearic islands and the Iberian Peninsula (Pittau 1995: 148–157), is connected to a so-called "thalassocracy" and ensuing colonization by the Lydian people of Anatolia. Pittau's hypothesis rests on a combination of historiographical, archeological and linguistic arguments. Here, I will focus on Pittau's linguistic arguments.

First and foremost, Pittau's (1995: 172; 1997: 28–33) connection between Etruscan and Lydian cannot be accepted. Even if one does not agree with the common view that Etruscan is not an Indo-European language (e.g. Wallace 2008: 1, 216–217), it is most definitely not an Anatolian language, which Lydian clearly is (Simon 2021; Kloekhorst 2022b). As a logical consequence, the "Nuragic" language posited by Pittau for pre-Roman Sardinia, cannot have been related to both

[25] The genetic findings of Posth et al. (2021) are more consistent with an autochthonous Italian origin of the Etruscans.

Etruscan and Lydian. To my knowledge, no Indo-European Anatolian presence has ever been proposed for Sardinia. The "Lydian" hypothesis will therefore not be pursued any further here.

A hypothetical connection between pre-Roman Sardinia and the Tyrrhenian language family is still on the table however, regardless of the ultimate origins of the latter. The evidence for this, adduced by Pittau (1995: 197–233), consists of four categories. These are: 1) comparisons of Sardinian forms of suspected pre-Roman origin with Etruscan words, 2) comparisons with Latin words of probable Etruscan origin, 3) comparisons with Tuscan words of obscure etymology for which an Etruscan origin is plausible, and 4) comparisons with dialectal Italian words of obscure etymology for which an Etruscan origin is plausible. For these categories, we must address a number of caveats. Firstly, even though Pittau (1995: 196) objects against the use of person names and toponyms in linguistic comparisons, due to their semantic opacity, he does make use of both of these himself in categories 1 and 2. He argues that onomastic and toponymic material can be used if the compared forms consist of a significant string of phonemes, diminishing the risk of "fortuitous homophony". Yet, the issue of semantic opacity remains. We could find an eight-syllable-long place name in Sardinia and in Anatolia and it would be an enormous coincidence, but a coincidence nonetheless as long as we cannot verify that their meanings are comparable. These comparisons can therefore not be used to build a strong case for Tyrrhenian linguistic presence in Sardinia. Secondly, without direct evidence, an Etruscan origin of the Tuscan and dialectal Italian forms in categories 3 and 4 remains rather speculative, due to our scarce knowledge of the Etruscan lexicon and our limited knowledge of any other pre-Roman languages in the region that could have been the source of etymologically obscure words. The lexical and structural lines of evidence proposed to point to a connection between Etruscan and pre-Roman Sardinian are discussed in § 10.4.

2 Methodology

The study of prehistoric substrate languages comes with several issues. By definition, these languages are often not directly attested or they did not leave any attested descendants. As a result, the existence of any non-attested substrate language, including its properties, can only be inferred from the traces it left through contact with the languages that survived it. Expectedly, this leaves ample room for speculation concerning phonological and grammatical features of the substrate, not to mention its origins and affiliation. The main problem is that no direct linguistic evidence exists to either falsify or verify any claims made. All evidence is transmitted through the filter of contact with at least one other language, rendering the comparative method, a historical linguist's most trusted tool, only partially applicable (cf. § 2.1.1.2). Moreover, the indirect attestation of a hypothetical substrate language makes claims regarding its lexicon, grammar, and genetic affiliation difficult to evaluate. However, suggestions on how to address these methodological difficulties have been made at least since the late nineteenth century. For in-depth overviews of the field of linguistic substrate research, cf. Craddock (1969: 18–47; with a focus on Romance) and Kroonen (2024a; with a focus on Indo European).

In the present chapter, the most prominent challenges in linguistic substrate research are discussed, along with the methodological approaches adopted to address these challenges. In § 2.1, approaches to the identification and collection of the relevant data are discussed. Next, § 2.2 treats the analysis of the gathered data in terms of phonological and morphological features. Finally, § 2.3 explores the potential of what the interpretation of the linguistic data reveals about language relationships, and which processes and events they imply.

2.1 Identifying substrate elements

The identification of substrate elements in attested languages, as opposed to elements inherited from an ancestor language, is notoriously challenging. Especially when the ancestor of the language under research is not attested, discerning inherited from non-inherited formations is not always straightforward. The reliability with which a non-inherited element in a language can be identified inevitably depends on the degree of understanding we have of the ancestor of that language. Even for a well-researched language family like Indo-European, this has proven difficult (cf. Kroonen 2024a).

2.1.1 Criteria for substrate origin

The concept of linguistic substrates has featured in historical linguistic research from its early stages onwards (e.g. Hervás 1979: 18 [1800]; Pott 1853: 451–452; Ascoli 1882; Kretschmer 1896: 121; see Craddock 1969: 18–47 for an overview of the field). The first steps toward formalizing the methodology for the identification of substrate words can already be seen in Meillet (1908) and Cuny (1910). On the basis of semantic domain, unusual phonotactics and morphology, and irregular sound correspondences, they identify a number of pre-Indo-European words in Latin and Greek. However, clear criteria to study words of substrate origin in a more structured and objective manner were formulated only in the nineties of the previous century. None of the criteria is decisive on its own, but some constitute stronger evidence in favor of substrate origin than others. Roughly in descending order of strength, the proposed criteria are:

- lack of an etymology
- irregular correspondence
- unusual phonotactics
- unusual morphology
- specific semantic categories
- limited geographical distribution

Each of these will be discussed below.

2.1.1.1 Lack of an etymology

A straightforward reason to suspect non-inheritance, and potentially a substrate origin, is the absence of a clear etymology. This criterion has been employed in various Indo-European branches (Meillet 1908: 161; Polomé 1986: 661; 1989: 54; 1990: 276–277). In the case of Sardinian, its ancestor Latin is well attested and much of the Sardinian lexicon is therefore etymologically transparent. If a Sardinian word does not have a clear Latin pre-form, this can suggest a non-inherited origin. Naturally, there is always a possibility that a word was not recorded by chance. Given our extensive knowledge of Latin however, these chances are slim. Conversely, if a plausible Latin etymology is available, it should always be preferred over a hypothetical substrate origin.

2.1.1.2 Irregular correspondence

There are cases in which a certain word is found in several related languages and could thus, based on its linguistic distribution, be inherited. Still, a non-inherited origin may be posited when the attested cognates deviate from the known sound laws and sound correspondences. In such cases, one would assume that the word

was borrowed from a third language, or that it spread through borrowing between the related languages.

This criterion derives from a principle central to the comparative method: the regularity of sound change. It is used already by Meillet (1908) and Cuny (1910: 158–160), and remains important in substrate methodology today (Polomé 1989: 55; 1990: 278; Hamp 1990: 296, 300; Huld 1990: 392–393; Schrijver 1997: 294; Lubotsky 2001: 303). The criterion can be used on different levels: either between two languages that are "vertically" related, such as Sardinian and Latin, or between two languages that share a common ancestor, like Latin and Greek. Irregular correspondence between dialectal varieties of a single language can indicate a non-inherited origin of words too: in this case it is often referred to as "irregular alternation" (e.g. Cuny 1910: 156). With sufficient material, the nature of the observed irregularities can potentially offer clues about the phonology of the substrate source language (cf. § 2.2). The Sardinian evidence is discussed in § 8.

A special case are seemingly corresponding forms in unrelated languages (e.g. Cuny 1910: 161–164), such as Sardinian and Basque (§ 10.2.2). Since these languages do not share a common ancestor, the detection of regular correspondence is not a possibility, nor is the detection of phonological irregularity. When vocabulary items in unrelated languages share a common origin, this is likely the result of lexical borrowing. If the cognates cannot be etymologized in either language, they must have been borrowed from a third source.

There is an important caveat. The comparative method can reveal when two comparanda do not correspond regularly to each other, but it can only prove a common etymological origin when the comparanda do correspond regularly to each other. This is a fundamental limitation that causes any comparison of two or more irregular comparanda to run a risk of chance correspondence. Without the possibility to systematically verify or falsify the validity of a comparison, we must be watchful not to invalidate our investigations of linguistic prehistory by lumping any two vaguely similar words together. The methodological strategies applied in this study to mitigate this risk are discussed in § 2.1.2.2. Another problem for detecting potential substrate words is the fact that, at times, inherited formations exhibit irregular developments too. The phenomena responsible for this, and how they should be taken into consideration, are discussed in § 2.1.2.1.

2.1.1.3 Unusual phonotactics

The criterion of unusual phonotactics is closely related to the previous criterion and applies to phonological sequences that are unusual in the relevant language

or its ancestor, whether reconstructed or not, (Polomé 1989: 55; Hamp 1990: 298; Huld 1990: 393; Salmons 1991: 267; Beekes 1996: 218; Schrijver 1997: 294; Lubotsky 2001: 303–304). This could for instance be a phoneme that does not have a regular source in the historical phonology of the language, such as e.g. intervocalic *s* in Lat. *rosa* 'rose' (Weiss 2009: 151 fn. 12; Wigman 2023: 121–123) and IE **a* and **b* in Lat. *faba* 'bean' < **b*^{h}*ab-* (Kuiper 1995: 65–67; Wigman 2023: 80). Phonemes that do have a regular source, but that occur in a phonological context that did not occur in the ancestor language, are equally suspicious. An example of this is Srd. *ɸ* before *i* (supposing Lat. **c̯i̯-* or **t̯i̯-*) (§ 8.1.1). Here too however, irregular processes can falsely suggest non-inheritance (§ 2.1.2.1). In Sardinian, for instance, *s-* was irregularly replaced by *ɸ-* in *ɸilímba* (Baunei) 'bean pod' < Lat. *siliqua* (DES II: 417). For the interpretation of unusual phonotactics, cf. § 2.2.1. The Sardinian evidence is presented in § 8.

2.1.1.4 Unusual morphology

Another indication of borrowing is when lexemes exhibit morphological features that cannot be explained by inheritance (Polomé 1989: 55; Huld 1990: 390; Salmons 1991: 267; Schrijver 1997: 294–295; Lubotsky 2001: 304). This concerns affixes without an etymology, such as Gr. -ι/υνθ- (Pott 1853: 451; Meillet 1908: 162; Cuny 1910: 154–156; Katičić 1976: 41–43). It may also include other types of morphology, like irregular ablaut (cf. Schrijver 1997: 307–312). In the context of this work, a concrete example are Sardinian nouns ending in *-i* that suggest a Latin accusative in *-im*, which is only plausible for a small, closed class of inherited nouns (§ 8.1.5.1). The interpretation of morphological features is discussed in § 2.2.2. The Sardinian evidence is discussed in § 9.

2.1.1.5 Semantic category

It has been observed that lexical material of hypothesized substrate origin often clusters in specific semantic fields. Among these, prominent categories are those referring to the natural world, such as flora, fauna, geography, as well as terminology referring to cultural practices and technologies specific to the area (Meillet 1908: 161; Cuny 1910: 154; Bertoldi 1931: 94; Polomé 1989: 54; 1990: 276; Hamp 1990: 300; Schrijver 1997: 295; Lubotsky 2001: 304–305). This does not mean that any word in such semantic categories must be of substrate origin. As long as a concept existed in the predecessor of a language, the corresponding word could have been inherited. As this criterion much depends on the socio-economic context of the language contact, it is weaker than the purely linguistic criteria discussed above. However, it can inform on the nature of language contact once the material has been identified as being non-inherited (e.g. Polomé 1986:

669)(cf. § 2.3.2). In this study, the lexical material discussed is organized according to semantic category. The categories are as follows: flora (§ 3), fauna (§ 4), geography (§ 5), culture and economy (§ 6), and miscellaneous (§ 7).

2.1.1.6 Limited geographical distribution

A limited geographical distribution is the last criterion introduced here. It has been invoked with regard to the branches of Indo-European. When a word occurs only in Indo-European languages of northern or western Europe, for instance, projecting that word back to Proto-Indo-European is considered dubious (Polomé 1986: 663–665; Huld 1990: 393; Schrijver 1997: 294): even when the comparanda are regular, there is a risk of borrowing (Schrader 1883: 201–202). Within the context of the present study, a word occurring exclusively in Sardinia could quite possibly be a loan from an extinct language once spoken on the island. If, however, a word is found in all of Romance, chances are that we are dealing with a formation that was inherited from (vulgar) Latin. In isolation, this criterion is not conclusive. Inherited material may have been preserved only in a single branch of a language family due to chance. For Sardinian, examples of this are Srd. *dòmo*, *dòmu* 'house' < Lat. *domō* (abl.) 'id.', and Srd. *yánna*, *yènna* etc. 'door' < Lat. *ianua* 'id.' (REW 4575; DES I: 707), which have been replaced in most other Romance varieties by reflexes of Lat. *casa* 'house, hut' and Lat. *porta* 'gate' respectively (REW 1728, 6671). Nevertheless, the geographical distribution of putative substrate words is important for interpreting prehistoric linguistic connections between regions (Polomé 1986: 669; 1990: 282) (§ 11).

2.1.2 Form, meaning, and chance similarities

In isolation, none of the criteria discussed in § 2.1.1 can conclusively demonstrate a word's substrate origin. In practice, it is often the cumulative evidence that decides the issue. The case for a substrate origin becomes stronger if a word answers to two or more of the criteria in § 2.1.1 (Schrijver 1997: 294–296). And even then, rejecting an inherited origin may be impossible. For instance, Srd. *iṡbírru*, *skírru* 'marten' < **squiriolu* (DES I: 646–647) satisfies multiple criteria suggesting a non-inherited origin. It belongs to the semantic category of animal names, is found in the meaning 'marten' only in Sardinia, and corresponds irregularly to Lat. *sciurus* 'squirrel'. Yet, it was likely inherited from Lat. *sciurus* 'squirrel', with a shift in meaning and an irregular metathesis from **skʲur-* to **skʷir-* (DES I: 646-647).

On the other hand, there are instances where accepting an inherited origin is problematic, even if the word question does not satisfy any of the criteria dis-

cussed. Such is the case for Srd. *kála* 'cove, inlet (for ships)', which corresponds to Sp., Cat., OProv., Cors., It., Sic. *cala* 'id.'. It is thus attested in a wider region, in a semantic category different from the ones mentioned, and shows no trace of phonological irregularity. However, it lacks a reliable Latin etymology and has therefore been argued to be of pre-Roman origin (cf. § 5.1.3).

Since no criterion by itself is decisive, the identification of lexical substrate elements depends on careful consideration of form, meaning, etymological status, and geographic distribution. A Romance language like Sardinian has the advantage that its Latin ancestor is exceedingly well attested, allowing us to follow its diachronic development for over two millennia. Because of this, any Romance word lacking a Latin correspondent, is suspect of borrowing from the outset, even if it is otherwise unremarkable in form, meaning, or distribution. Although by no means a conclusive criterion either, it is relatively strong compared to cases in which the linguistic ancestor is not attested. This is the case for many ancient Indo-European subgroups, such as Latin, Greek, Sanskrit, and Hittite, but also for ones that were attested relatively late, such as Baltic and Slavic. Still, even though the study of Sardinian linguistic prehistory is greatly aided by our knowledge of Latin, there are complicating factors that deserve attention. Specifically, inherited material too sometimes exhibits irregularities (§ 2.1.2.1). This raises a methodological question as to when two irregularly corresponding forms should be seen as independently borrowed cognates, or as the result of local irregularizing processes (§ 2.1.2.2).

2.1.2.1 Substrate origin or irregular inheritance?

The exceptionlessness of sound change is the cornerstone of the comparative method (cf. § 2.1.1.2). However, this exceptionlessness is frequently obscured by analogy, folk etymology, onomatopoetic processes, sound-symbolism, and basically any type of "expressivity". This obstacle is all the more evident on the dialectal level. Local varieties of one and the same language can exhibit slightly irregular manifestations of an etymologically identical formation, resulting from local processes or inter-dialectal borrowing (Schuchardt 1885: 11). These irregularizing processes are crucial to take into account, as they may falsely suggest a non-inherited origin for inherited lexical items.

Within the context of this study, it must be noted that many allegedly pre-Roman words in Sardinian have received competing Latin etymologies or etymologies rooted in the historical contact languages of Sardinian. Not infrequently, the choice between a substrate origin and an inherited or recent loan origin

depends on the interpretation of irregular phonological developments or opaque semantic shifts. To an extent, the acceptance of either an inherited or a substrate origin for these cases is thus a matter of intuition. Some scholars are willing to accept a certain degree of formal or semantic deviation to sustain an etymology. Others prefer to take formal and semantic incongruencies at face value, leading them to posit an unattested substrate. Since both approaches entail uncertainties, systematically prioritizing one over the other is counter-productive: we must instead face the reality that both irregular developments of inherited words and borrowings from unattested languages do occur. In order to a achieve a maximally robust methodology, we must acknowledge these uncertainties and carefully weigh the arguments in favor an irregular inherited origin against those in favor of a substrate origin, for each individual word. This is the approach that I adopt in the discussion of the lexical material in § 3 – 7.

2.1.2.1.1 A note on folk etymology

Folk etymology is the process by which the form of a word is altered on the basis of an association with the form of another word. This process is often aided by semantic association. Various linguists have stressed the role of folk etymology in concepts related to animals, plants, and belief systems (Alinei 1984; Beccaria 1995). Paulis (1992), in his work on Sardinian plant names, often invokes folk etymology to account for irregular interdialectal comparanda found across different dialects. The plausibility of a folk-etymological origin depends on the similarity between the irregular form and the putative model, and on the possibility of a semantic connection between the two forms. Needless to say, such evaluation demands a good understanding of the cultural and practical applications and associations of the concept in question (e.g. Paulis 1992: 11–12). As a general rule, I consider the semantic motivation of a supposed folk etymology convincing as long as it does not depend on hypothetical unverifiable cultural contexts; cf. § 2.1.2.2, where a similar attitude towards non-inherited comparanda is adopted.

2.1.2.2 Comparing non-inherited words

Traditionally, a large part of the research into substrate languages has centered on identifying lexical correspondences between the substrate lexicons of different languages or areas. The motivation for this is typically to recover the prehistoric language relationships and contacts between peoples. However, as mentioned in § 2.1.1.2, the comparison of irregularly corresponding cognates, or similar words in unrelated languages, always involves a risk of chance similarity. In order to render the corpus of potential substrate connections as reliable as pos-

sible, it is pertinent to be strict regarding their formal and semantic similarity. A clean corpus is important for a reliable analysis of the linguistic data, on which hypotheses of prehistoric language relationship are necessarily based.

Ideally, a set of non-inherited comparanda are formally and semantically (near-)identical. An example of this is the comparison between Srd. *golóstḭu*, *golóstri* etc. 'holly' and Bq. *gorosti* 'id.', both of which can go back to **golosti* (§ 3.3.7)(Bertoldi 1929: 261 fn. 3). The majority of comparisons, however, require varying degrees of liberty on the formal or semantic side, or both. Statistically, longer corresponding phonological sequences are less likely to be the result of chance similarity. For instance, the shared sequence **al-* in Srd. *aláše* 'butcher's broom' and Prov. *árẹ* 'broom' (§ 3.2.1) stands a higher chance of being due to coincidence than the much longer sequence of sound correspondences between aforementioned Srd. *golóstḭu* and Bq. *gorosti* 'holly'. It is however difficult to draw a line, and great coincidences can of course happen sometimes.

The semantic evaluation of lexical comparisons is even more intuitive (cf. § 2.1.2.1). The approach of this study is to accept only those comparisons whose semantic shifts are trivial. As indicated above, this includes only those shifts that do not require hypothetical cultural contexts (cf. Polomé 1990: 282). Thus, the shift 'maple' > 'ash tree' needed to connect Srd. *kòsti*, OProv. *agast* 'maple' to Tusc. *còstolo* 'ash tree' is accepted under this criterion (§ 3.3.9). The change 'poppy' > 'lady' needed for Lang. *ãndèr* etc. 'poppy' and Bq. *andere* 'lady' (§ 3.1.15), on the other hand, is not. Needless to say, the approach described here is but a crude way of drawing a line. Inevitably, it will exclude comparisons that are correct despite their opaque semantics. For instance, the Romance words for 'liver' (Nuor. *fíkatu*, Camp. *fɨɣáu*, Sp. *hígado*, Fr. *foie*, It. *fegato* etc.), which are known to be derived from Lat. *ficus* 'fig', would not be considered convincing for reasons of stringency.[26] However, it is preferable to infer prehistoric language relationships on a smaller corpus of strong comparisons than on a larger corpus of weak ones.

2.1.2.2.1 Toponyms as evidence

Placenames, or toponyms, take up a prominent place in substrate research in general (e.g. already Kretschmer 1896: 400). Especially in the earliest days of the "Mediterranean hypothesis", many of the proposed pre-Roman connections between different areas in the Mediterranean were based in large part on toponyms (§ 1.3.2). Toponyms are among the parts of a language that are most resilient even in the event of a wholesale language-shift (Bertoldi 1931: 95). In situa-

[26] Cf. FEW III: 490–493 and DES I: 518–520 for the specificities of this etymology.

tions where speakers shifted from a language prior to its attestation, the toponymical evidence potentially contains valuable traces of the preexisting linguistic landscape. However, even though toponyms can be preserved over long periods of time and across multiple language shifts, they suffer from a significant drawback: they are usually semantically opaque. As discussed in § 2.1.2.2, scrutiny of the semantics of comparisons is an important step to minimize the risk of chance correspondence. To quote Bertoldi (1933: 260): "abandoning the lexicon […] means letting go of one of our most precious research instruments: the possibility of semantic verification".[27] Since the original meaning of substrate toponyms cannot typically be retrieved, such comparisons are limited to formal similarity.[28]

Toponymic evidence can on the other hand be mobilized as an additional check for verifying the characteristics of a substrate, once it has been identified in the lexicon. Provided that the source language of substrate toponyms is the same as the source language of the substrate lexicon, we should expect features exhibited by the lexicon to reappear in local toponymy (Aikio 2004: 9). In addition, toponyms can support the identification of recurring sequences, i.e. potential morphemes of the source language. These too can then be compared to the lexical evidence. In Sardinia, such an approach has been applied to the toponymy of the Barbagia-region by Wolf (1998a) and integrated with the findings from lexical comparisons by Paulis (2008). Thus, toponyms can aid the reconstruction of the phonology and morphology of a hypothetical substrate source language (§ 8 – 9).

2.1.2.2.2 Conclusion: dealing with chance similarities

As discussed, the main methodological issue in comparing potential substrate words is that chance correspondence can never fully be ruled out. This does not mean that the matter is to be abandoned altogether. In order to isolate compelling evidence, etymological proposals should be evaluated along the following two lines:

---

[27] "Abbandonare il lessico significa però lasciarsi sfuggire uno degli strumenti di ricerca più preziosi: la possibilità del controllo semantico." (Translation mine).

[28] A quick glance at some similar European city names whose etymologies are known, reveals the danger of postulating cognacy without semantic checks: cf. *Lyon* (France) < Lat. *Lugdūnum* << Gaul. **Lugudūnon* 'stronghold of Lug' vs. *León* (Spain) < Lat. *legiō, -ōnem* 'legion'; *Almería* (Spain) << Arab. *al-Mariyya* vs. *Almere* (The Netherlands) < *Aelmere* 'big lake'. For additional examples, see Bottiglioni (1929: 15–16) and Wolf (2011: 607).

- The forms compared must formally be sufficiently similar or identical. This is difficult to quantify, but the longer the corresponding sequence, the better. In doing this, it is crucial that the internal phonological developments of the compared languages be taken into account.
- The meanings of the comparanda should be evident. They should as a minimum not involve any semantic shifts that are only understandable with additional, unverifiable assumptions (about *Benennungsmotive* or cultural practices etc.).

Positing these guidelines is easier than putting them into practice. There is no final answer as to how divergent the forms and meanings of two or more comparanda can be, before their comparison is to be rejected. Each comparison should be weighed carefully on the basis of the relevant data. The semantic guideline eliminates toponyms as reliable evidence.

2.2 Analyzing the data

In the above, we have explored the methodological strategies for identifying and comparing non-inherited material between different languages to detect prehistoric linguistic relationships between regions. It is equally informative to investigate the phonological and morphological features of the non-inherited lexicon. This enables us to exploit the body of non-inherited material found in Sardinian to gain insights into the properties of the substrate language(s) of Sardinia (cf. Devoto 1940: 40–43), as well as the dynamics of its contact with Latin.

2.2.1 Phonological analysis

The notion that phonological properties of suspected substrate words reflect phonological properties of the donor language has been present from an early age. As early as 1800, Don Lorenzo Hervás (1979: 18) notes:

"Had these things [i.e. the presence of other languages in France and Spain in Roman times] not been known to us through history or tradition, we could infer them by observation and comparison of the Latin dialects, which the French and the Spanish speak presently, with Celtic and Cantabrian [i.e. Basque], which are still spoken among them, with ancient Latin, and with the pure dialects of the latter, which in Italy are spoken by the Romans' descendants. This observation and comparison would make us notice and discover many features of Celtic in French, and many of Cantabrian in Spanish. If by chance the latter and the Celtic language had perished, and we would thus consequently not have been able to compare to them the languages that the French and Spanish actually speak, nevertheless, because of the differences of the idioms and of other things in

those languages, we would conjecture that in ancient times the French and Spanish spoke different languages."[29]

This view was taken up by Ascoli (1882), who argues that certain typical Gallo-Romance sound shifts resulted from a Celtic substrate.[30]

Later, the attention was directed to phonological irregularities and alternations found in specific lexical items. Wagner (1907a: 408–410) already considers ancient Iberian as a source for unexplained lexical similarities between Sardinia and the Iberian Peninsula (cf. § 11.3.1). Meillet (1908) sees a pre-Indo-European language impacting Greek, Latin, Armenian and Iranian. Both scholars express a degree of methodological hesitation. Wagner (1907a: 408) asks: "if the Sardinian word [i.e. *vega*] is old, as we must believe, the etymon **vica* [proposed for Sp. *vega*] is not acceptable; but is it therefore certain that it is Iberian, or do we need to look for another explanation?".[31] Meillet (1908: 164) warns that "it would be unwise to want to specify too much detail in such matters [i.e. which populations the people speaking Indo-European dialects encountered]"[32] — words of caution that are valid to this day. Nevertheless, the idea that irregularly corresponding words could represent parallel borrowings from a substrate language was well-established by the beginning of the 20th century.

What followed was the interpretation of the irregularities and alternations in these putative substrate words. Meillet (1908: 163) and Cuny (1910) account for various irregular correspondences between Greek and Latin words by postulat-

[29] "Si todas estas cosas no nos constaran por la historia ó tradicion, podriamos inferirlas de la observacion y cotejo de los dialectos latinos que al presente hablan los francesses y españoles, con el céltico y cántabro, que aun se hablan entre ellos, con el latin antiquo, y con los dialectos puros de éste, que en Italia hablan los descendientes de los romanos. Esta observacion y cotejo nos harian advertir y descubrir en la lengua francesa muchas cosas de la céltica, y en la española muchas de la cántabra. Si por ventura ésta y la lengua céltica hubieran perecido, y consiguientemente no pudiéramos hacer entónces cotejo con ellas y los lenguages que en la actualidad hablan los franceses y los españoles, no obstante, por razon de la diferencia de los idiotismos y de otras cosas en dichos lenguages, conjeturariamos que antiguamente los franceses y los españoles hablaban diversos idiomas." (Translation mine).

[30] Cf. Kretschmer (1896: 121) for some more early examples.

[31] "Se il vocabolo è anche del sardo antico, come bisogna credere, l'etimo **vica* non è ammissibile per questo; ma è per questo sicuro che è iberico o bisogna cercare un'altra spiegazione?" (Translation mine).

[32] "Il serait imprudent de vouloir trop préciser le détail en pareil matière". (Translation mine).

ing independent borrowings from a third language. They attribute the irregularities to a phonological mismatch between the respective source and target languages. Several alternations in the Greek lexicon have since been attributed to "imperfect" borrowing of non-native phonemes from a substrate language (e.g. Kretschmer 1940: 269; Heubeck 1949: 201–202; 1959: 56–57; Kuiper 1956: 219–220; Schachermeyr 1964: 260–261; Ruijgh 1967: 53 fn. 35 etc.). For examples outside Greek, cf. Martinet (1955: 387) and Kronasser (1966: 41).

The notion that formal irregularities in substrate words could be attributed to phonological properties of the source language, has become fundamental to linguistic substrate research. Attempts to find recurring patterns in irregular alternations are made by Oštir (e.g. 1921; 1930) and later by Furnée (1972; 1979). Both try to interpret irregular alternations between words in various languages as a function of the phonology of substrate languages. Neither scholar's efforts have been met with great enthusiasm. In Oštir's case, this is likely due to the density and idiosyncrasy of the data presentation and of his complicated writing style. Furnée (1972: 92–83) interprets the phonological alternations observed in Greek as reflecting independently borrowed "expressive" and/or "euphonic" variants in the source language, thus pushing the problem from Greek to the unattested source language. This approach is continued and improved by Hamp (1979; 1990), Huld (1990), Salmons (1991), Beekes (1996), Schrijver (1997) and Boutkan (1998). More so than Oštir and Furnée, these scholars identify well-defined irregular correspondences recurring in words found across the (northern) European branches of Indo-European. On the basis of these, they posit clearly formulated ideas on some of the morphophonological properties of a hypothetical pre-Indo-European language that could be the source of these words. Furnée's (1972) extensive material is used by Beekes (2014) to interpret "pre-Greek" alternations as the result of phonological mismatches between source and target language.

This is the methodology that will also be adopted in this study, specifically in § 8. Detailed analysis of the dialectal variation found in the Sardinian lexicon (in § 3 – 7) allows us to identify irregular correspondences that cannot plausibly be explained by means of processes such as folk etymology (cf. § 2.1.2.1). If any of these irregularities recur in multiple words, this can be indicative of features of the phonology of their pre-Roman Sardinian source language(s) (cf. Salmons 1991: 267; Schrijver 1997: 312).[33] Irregular correspondences with single occurrenc-

[33] The phonology of pre-Roman Sardinian toponyms has been studied by Serra (1960) and Wolf (1998a: 77–81), with a focus on syllable structure.

es only cannot be used to make any reliable inferences about the phonology of a substrate language and are therefore left out of consideration.

2.2.2 Morphological analysis

The methodology for identifying and interpreting morphological features of substrate words has gone hand in hand with that of phonological features (§ 2.2.1). Pott (1853: 451–452) considers the Greek elements -νθ- and -σσος/-σσα, also found in toponyms, to "maybe [be] a remainder of from the language of a people that preceded the Hellenes".[34, 35] For Sardinian, Forsyth Major (1893: 154) and Guarnerio (1904b: 259) analyze the element *činči-*, *tsintsi-*, found in various animal names, as a diminutive prefix of Iberian origin. Kretschmer's (1896: 402–405) discussion of Greek forms containing -νθ- allows us to distill three criteria for establishing a substrate origin for a segment:

1. The sequence is not an isolated phenomenon but occurs in various words of non-inherited origin.
2. The "morpheme-hood" of the sequence is secured through alternations in identifiable cognates.
3. The sequence cannot be of inherited origin because of formal or semantic reasons.

Kretschmer (1896: 404) additionally notes that a pre-Greek origin of -νθ- is suggested by the fact that it is not productive and only occurs in words of obscure origin. However, this criterion only works in one direction. While its occurrence on etymologically obscure words may indeed point to a substrate origin for -νθ-, a word containing -νθ- cannot mechanically be considered a substrate word, as secondary productivity of the suffix cannot easily be ruled out. This applies to all hypothetical substrate morphemes.[36] Over the years, numerous substrate morphemes have been posited on the basis of both toponyms and lexical nouns, especially in the context of the Mediterranean hypothesis (see § 1.3.2). The foundations for the identification of pre-Roman Sardinian morphemes were laid by Terracini (1927), who identifies various pre-Roman suffixes found in Sardinian toponyms (*-V́ɻ́*, *-án*, *-́in*, *-énnVr*, *-́Vr* etc.; cf. § 9.2).

A risk concerning the study of substrate morphology is to give a sequence morphemic status when there is in fact no compelling reason to do so. An example of

[34] "… vielleicht Ueberrest aus der Sprache eines den Hellenen vorausgegangenen Geschlechts, …" (Translation mine).

[35] Cf. Katičić (1976: 39–55) and Kroonen (2024b) for an overview.

[36] Cf. the complex case of athematic velar suffixes in Greek (Chantraine 1933: 376–383).

this is Bertoldi's (1928: 231) suggestion of a "Mediterranean" derivational suffix *-*st*(*r*)-, represented in Srd. *golóstḭu* etc. 'holly' (§ 3.3.7) and *giḍḍòstre* etc. 'tree heather' (§ 3.3.20)(cf. also Bertoldi 1929: 261 fn. 3; 1930). Although the first and third criteria listed above are fulfilled, with two non-inherited words containing an etymologically unexplained sequence, neither of these words have corresponding comparanda without *-*st*(*r*)-.[37] Even if such an alternation can be found elsewhere (e.g. Prov. *árę* 'broom' besides Lang. *orousto* 'broom branch'; § 3.2.1), the Sardinian data do not provide evidence that *-*st*(*r*)- was an independent morpheme, rather than a part of the lexical stem, in the Sardinian substrate. This example illustrates how one's evaluation of the second criterion depends on the cognates that one accepts, which in the case of non-inherited lexicon suffers from the additional uncertainties described in § 2.1.2.2. Thus, FEW's (II: 490) and Hubschmid's (1953: 30) comparison between Srd. *kaðúmbu*(*lu*) 'great mullein' < **katúmbu*- and Campan. *kwátenə*, *kwātanə* 'id.' is only possible if both Srd. *-úmbu* and Campan. *-enə* are analyzed as suffixes to a root **k*(*w*)*at*-. However, identifying these morphemes is only possible in the first place if we accept the cognacy between these two words.[38] This approach evidently runs the risk of becoming circular, and for that reason I exclusively consider potential substrate morphemes if their alternations are attested within Sardinian, in indubitably cognate forms (in § 9). This will inevitably exclude many of the comparisons proposed in previous scholarship (cf. § 1.3), but in turn greatly increases the reliability of the accepted cases.

Once a substrate morpheme has been identified, the logical next step is to determine its function. This is usually challenging, however. If we exclusively accept a sequence as a morpheme when it alternates in forms recognizable as cognates due to their (near-)identical meaning (e.g. *-(*i*)*k*- in *kòstike* besides *kòsti* 'maple'; § 3.3.9, § 9.2.2.2), we inevitably lose any derivational morphemes whose very purpose it was to alter the original meaning of a word. What remains are those morphemes that are semantically bland, including but not limited to case markers, gender/noun class markers, number markers etc. It does not come as a surprise that so many alleged substrate affixes have been interpreted as having a plural or "collective" function: cf. § 9.2.1, § 9.2.2.1, § 9.2.6.2.1 for the Sardinian cases alone. It is implausible that all of these proposals are correct.

---

[37] The comparison between Srd. *giḍḍòstre* and Bq. *gillar* etc. by Hubschmid (1953: 29) is uncertain (cf. § 3.3.20).

[38] For many more such comparisons to Srd. *kaðúmbu* 'mullein', cf. Hubschmid (1953: 29–33).

The two pre-Roman morphemes identified in the Sardinian lexicon that have received most attention with regard to the question of their original function, are **-ar/-Vr* and **ϑi-* (and variants). The former is found in lexical items as well as toponyms. On the basis of early attested toponyms like *Gennor* and *Mandara*, whose endings were over time replaced with the inherited plural suffix *-s*, resulting in *Gennos* and *Mandas* (§ 9.2.6.2), Terracini (1927: 139) has convincingly analyzed *-r* as a plural marker. The prefix **ϑi-* is found in lexical items and has been interpreted as a diminutive marker (Forsyth Major 1893: 154; Guarnerio 1904a: 57–58; 1904a: 260; Wagner 1932: 223–224; 1997: 263) as well as a morphosyntactic element (Pittau 1995: 197; Swanenvleugel 2024)(see § 9.1.2). It goes to show that, unless there is a fortuitous attestation of relevant forms (as was the case for **-Vr-*), establishing the original function of a non-productive morpheme can remain challenging despite more than a century of debate.

2.3 Interpreting results

Once the evidence has been examined in the greatest possible detail, a logical next step is to offer a linguistic scenario that is consistent with the findings. If sufficient evidence for a substrate language has been identified, possible follow-up questions are:

1. Stratification: Was there a single substrate language or were there multiple?
2. Sociolinguistics: What was the nature of contact between the substrate language and the attested language leading up to the eventual language shift?
3. Chronology: For how long did the substrate and superstrate languages coexist?
4. Origin: What was the linguistic affiliation of the substrate language(s)?
5. Interdisciplinarity: Is there any extralinguistic (e.g. historical, archeological or genetic) evidence corroborating the linguistically inferred scenario?

It is unlikely that all of these questions can be satisfactorily answered for every single linguistic substrate. Even in a region like southern Europe, where relatively many non-Indo-European languages have survived to the brink of history (e.g. Tartessian, Iberian, Etruscan, Rhaetian, Minoan, Lemnian, Eteocypriot) and where the existence of other, non-attested languages is apparent from classical sources, it is still difficult to reconstruct the specifics of linguistic affiliation and the degree of contact between these languages.

Situations of language contact are inherently complex. A survey of modern-day Sardinia may serve as an illustration. In most of the island Sardinian is spoken, but with considerable variation between the various dialects. In addition, there are areas with Catalan, Ligurian and Corsican dialects. To add to this complexity, Italian has imposed itself as a superstrate to all these speech varieties. If someone in the future, when the shift to standard Italian has been completed,[39] were to attempt to reconstruct the island's former linguistic diversity, chances are low that they would succeed to fully capture it. There is *a priori* no reason to think that Sardinia in pre-Roman and Roman times was linguistically less diverse. It is thus important to acknowledge the fact that many aspects of the original linguistic landscape are irrecoverable. Any hypothesis claiming the opposite is inescapably wrong. This does not, however, mean that all attempts to elucidate linguistic prehistory are necessarily futile. Even extracting the tiniest bit of information on the past linguistic configuration of a certain region, is potentially of great value to our understanding of general (pre-)history.

2.3.1 Unity of the substrate

A good starting point for the investigation of any prehistoric linguistic situation, is to assume that it was to some extent linguistically diverse, unless there are strong reasons to assume homogeneity. When analyzing the corpus of non-inherited material in a given language, we should therefore ask ourselves the question whether the assumption of a single extinct language can account for all the evidence.[40] One way to demonstrate such linguistic homogeneity, is to analyze the dialectal distribution of the features of the linguistic substrate: lexemes, phonology, morphemes etc. If a set of features is consistently found in a particular area, this may point to the existence of a distinct language limited to that area. Conversely, if many different features are found across all of the speech varieties in the researched region, it seems more likely that a single language was spoken throughout that region. Naturally, the latter scenario does not preclude that other languages were present, in smaller pockets, like in present-day Sardinia.

With regard to the linguistic situation on Sardinia at the time of the Roman conquest, at least two languages were spoken on the island. One was Phoenician, the language of Punic settlers from Carthage, who left ample inscriptional evidence

[39] This is, and will hopefully remain, a purely hypothetical scenario.

[40] For a case in point, cf. Beekes' (2010: xli–xlii; 2014: 45) assumption of the unity of pre-Greek, and the critique by Verhasselt (2011: 279), Meissner (2013: 6–7), De Decker (2015: 5), and Meester (2024).

(Bondì 1987d; Adams 2003: 209). In addition, one or more non-colonial languages native to Sardinia were spoken, but these left no written record. In order to be able to stratify the pre-Roman lexicon, it is necessary to analyze and interpret the distributions of non-inherited lexemes, and of their phonological (§ 8) and morphological features (§ 9).

2.3.2 Nature of language contact

Thomason and Kaufman (1988: 121) argue that it is impossible to predict the outcome of linguistic interference in case of a language shift. The opposite, interpreting the nature of language contact on the basis of linguistic interference, is not much easier. As discussed in § 1.3.1.3, it is difficult to prove that a certain feature of Sardinian reflects influence from a substrate language. However, on the basis of the lexical data, discussed in this study, we may attempt to employ linguistic paleontology to make inferences about the nature of economic and cultural contacts between speakers of Latin, Phoenician, and the other pre-Roman language(s) spoken on Sardinia.

2.3.3 Dynamics of language shift

The question of the nature of the language contact between Latin and the preexisting languages in Sardinia is intrinsically linked to the romanization process (§ 1.2). Multiple scenarios can be considered. Did native Sardinians shift to Phoenician prior to shifting to Latin? If not, did speakers of Phoenician and other pre-Roman language(s) shift to Latin at the same time, or did one language outlive the other(s)? The patterns of distribution of non-inherited linguistic elements can be employed to shed more light on the various scenarios that can be proposed for the relative chronology and direction of the romanization process.

2.3.4 Classifying a non-attested language

A question that has proven particularly popular in the study of linguistic substrates is to which languages they are related. The Sardinian substrate is no exception. It has been proposed to be related to Basque, Etruscan, and Berber, as well as to other unattested hypothetical languages constituting substrates elsewhere in the Mediterranean (cf. § 1.3.2, § 1.3.2.1). Like in the case of living languages, the classification of substrate elements as belonging to a certain language family should ideally be based on morphological as well as lexical evidence. The difference with living languages is however that the available material for substrate languages is usually much scarcer, especially in the morphological domain, and intrinsically transmitted indirectly through the language that replaced it.

In order to confidently posit a genetic link between a presumed substrate language and another language, the hypothesis must be able to account for a significant portion of the identified substrate elements. In Sardinia, this is the case for Punic, to which a number of non-inherited words can be ascribed (§ 10.1) (Paulis 1990).[41] A hypothetical genetic link between two unattested substrate languages is harder to evaluate, due to the greater amount of uncertainties in these cases. It is difficult to distinguish "native" substrate words from *Wanderwörter* (e.g. into 1st century BCE Sardinian Latin), or from words borrowed into the substrate language (e.g. a loan from Greek into 4th century BCE Sardinian Punic). We must thus rely on the number of shared forms between two putative substrate languages when evaluating the likelihood of their genetic relatedness. An important caveat is that it cannot be expected that the geographical extent of extinct languages coincides exactly with that of modern languages. In order to reliably interpret the potentially putatively shared between two substrate languages, close attention should be paid to the geographical distribution of these features on both sides of the comparison. This will be discussed in § 11.

2.3.5 Extralinguistic arguments

Once we have formulated hypotheses about the pre-Roman linguistic configuration, and subsequent romanization of Sardinia on the basis of linguistic evidence, a possible next step is to establish whether these can be corroborated by other lines of evidence. Such evidence may come from the fields of archeology, genetics, or from historiographical sources. For overviews of Sardinian archeology, see Lilliu (1963; 2002), Webster (1996), and Dyson and Rowland (2007). In recent times, Sardinia has been the subject of numerous genetic studies. Both ancient and modern Sardinian populations display a remarkably low degree of steppe-derived ancestry, whose prominence in the rest of Europe has been connected to the spread of the Indo-European languages (e.g. Chiang et al. 2018; Fernandes et al. 2020; Calò et al. 2021). These non-linguistic lines of evidence are best considered by experts in the respective fields, and are therefore left outside the scope of this study.

2.4 Conclusion

When it comes to methodology of substrate research, an important realization is that the evidence should be treated with great caution, due to the many uncertainties surrounding extinct, unattested, indirectly transmitted languages. We

[41] Cf. § 10.3.3, where it is argued that Basque–Sardinian comparisons do not live up to this expectation, thus making Basque an unlikely relative to the Sardinian substrate.

must accept that it will never be possible to recover all of the intricacies of lost prehistoric languages. To avoid ending up with grandiose theories (cf. § 1.3.2), only the strongest possible evidence should be used when formulating hypotheses about the pre-Roman languages of Sardinia. Although this means that less of the non-inherited material is exploited to reach conclusions, the conclusions that are reached should be relatively more reliable. In practice, I argue for deploying a maximal degree of strictness in the various avenues of substrate research discussed. Some points that are observed throughout this study are the following:

- Etymological obscurity does not necessarily imply a substrate origin. The latter should only be posited in the presence of compelling positive evidence, e.g. non-native phonemes, irregular alternations, irregular cognates. Isolated words of unknown origin generally do not provide reliable evidence. Cf. § 2.1.1.
- On the other hand, an inherited etymology should not involve too many unmotivated irregular sound changes aor non-trivial semantic shifts. Cf. § 2.1.2.2.
- Comparisons between non-inherited comparanda can only be plausibly made when they display sufficient similarity and when semantic differences are trivial. Cf. § 2.1.2.2.
- The previous point generally rules out toponyms as evidence for substrate language relatedness, as their semantic opacity means that the risk of formal chance correspondence cannot be mitigated by semantic comparison. Cf. § 2.1.2.2.1.
- Phonological alternations are only potentially indicative of phonological features of the substrate language if they occur sufficiently frequently throughout the non-inherited lexicon. Cf. § 2.2.1.
- Morphological features can only be posited for a substrate language if there is sufficient positive evidence, i.e. an alternation between non-inherited endings. “Morpheme-hood” should only be posited based on language-internal evidence, as positing suffix alternation on the basis of forms in different languages renders both the comparison of the forms and the identification of the suffix circular. Cf. § 2.2.2.

It goes without saying that all of these points are subject to a researcher’s personal preferences and intuitions. Still, they provide guidelines, followed throughout the rest of this study, that allow us to distinguish plausibilities from uncertainties, and to formulate hypotheses based on the former.

II
Linguistic data

The collection of Sardinian lexical data presented and discussed in § 3 – 7, on which the discussion in other chapters is based, uses as its point of departure Wagner's *Dizionario etimologico sardo* (DES), the most complete etymological resource on the Sardinian lexicon. The material from the DES is complemented with material from other sources, most notably AIS, Paulis (1992), and Puddu (2023). Included are those Sardinian words that have been proposed to be of pre-Roman origin, or that lack a convincing inherited origin. For each lemma, the likelihood of a substrate origin will be discussed along the criteria discussed in § 2.1.1, as well as on the basis of its proposed cognates outside Sardinia (cf. § 2.1.2.2). Detailed attention will be given to dialectal variation, which can provide us with valuable information about the original form, meaning, and distribution of a word.

As a living language, Sardinian exhibits a large degree of dialectal variation, which is well-documented. The forms documented for specific localities in AIS, DES, Paulis (1992) provide a wealth of information about a word family's distribution, as well as interdialectal correspondences that allow us to reconstruct its original form. However, this dialectal variation does not manifest itself as simple regular phonological correspondences between forms with identical meanings attested in different places. Rather, we can observe slight differences in meaning from area to area, or even from town to town, as well as various local and irregular deformations of words as a result of their semantic connotations. Since the study into Sardinian linguistic (pre)history at hand is based predominantly on lexical data, it is crucial to present these data in an accurate and insightful manner. Every lemma in chapters § 3 – 7 is therefore provided with a map showing the locations of all relevant forms attested across the Sardinian local dialects.

Especially when dealing with irregular correspondences between dialects, evaluating possible etymologies requires a significant amount of knowledge about the speakers' lived reality; i.e. a *Wörter und Sachen* approach. For instance, the semantic connection between some species of plant and its use in traditional medicine may be fully transparent to a native speaker, but harder to understand from the perspective of a linguist's armchair. Such context-specific connotations have been explored in detail by Wagner (1921) on daily life in rural Sardinia, and by Paulis (1992) for Sardinian plant names. I have followed the Paulis' (1992) and Puddu's (2023) example to include the scientific name of plants and animals wherever possible, to avoid any confusion.

3 Flora

Flora is the semantic field to which the largest number of Sardinian words of proposed pre-Roman origin belong. The lemmata discussed in this chapter are divided into herbaceous plants (§ 3.1), grasses, cereals and weeds (§ 3.2), trees and shrubs (§ 3.3), and plant parts and other plant-related terms (§ 3.4).[42] In order to minimize the risk of semantic uncertainty, the scientific names of the plant species discussed has been provided wherever possible.[43]

Various plant species discussed are endemic to Sardinia, and it is therefore not surprising that the majority of the forms do not have any demonstrable cognates outside Sardinian. Still, the phonological and morphological features exhibited by these forms may inform us about the specificities of their source language.

3.1 Herbs

3.1.1 Annual mercury (*Mercurialis annua*); goosefoot (*Chenopodium sp.*)

Reconstruction	**Forms**
***katòne**	DES (I: 259): ***kaðòne***, ***koðòne*** (Log.), ***kaðòni*** (Camp.) ‘annual mercury (*Mercurialis annua* L.), goosefoot (*Chenopodium album* L.)’
	Paulis (1992: 374–375): ***kaðòne***, ***koðòne*** (Log.), ***kaðòne*** (Camp.) ‘goosefoot (*Chenopodium album*), ‘nettle-leaved goosefoot (*Chenopodium murale* L.)’, ‘many-seed goosefoot (*Chenopodum polyspermum* L.)’, ‘stinking goosefoot (*Chenopodium vulvaria* L.)’, ***kaðòni*** [***burdu***] (1: Escalaplano) ‘stinking goosefoot (*Chenopodium vulvaria*)’

[42] The words for ‘mushroom’ have been included in this latter section, by lack of a more suitable place.

[43] All botanical binomial names follow the current classifications and naming conventions of the WFO Plant List (2024).

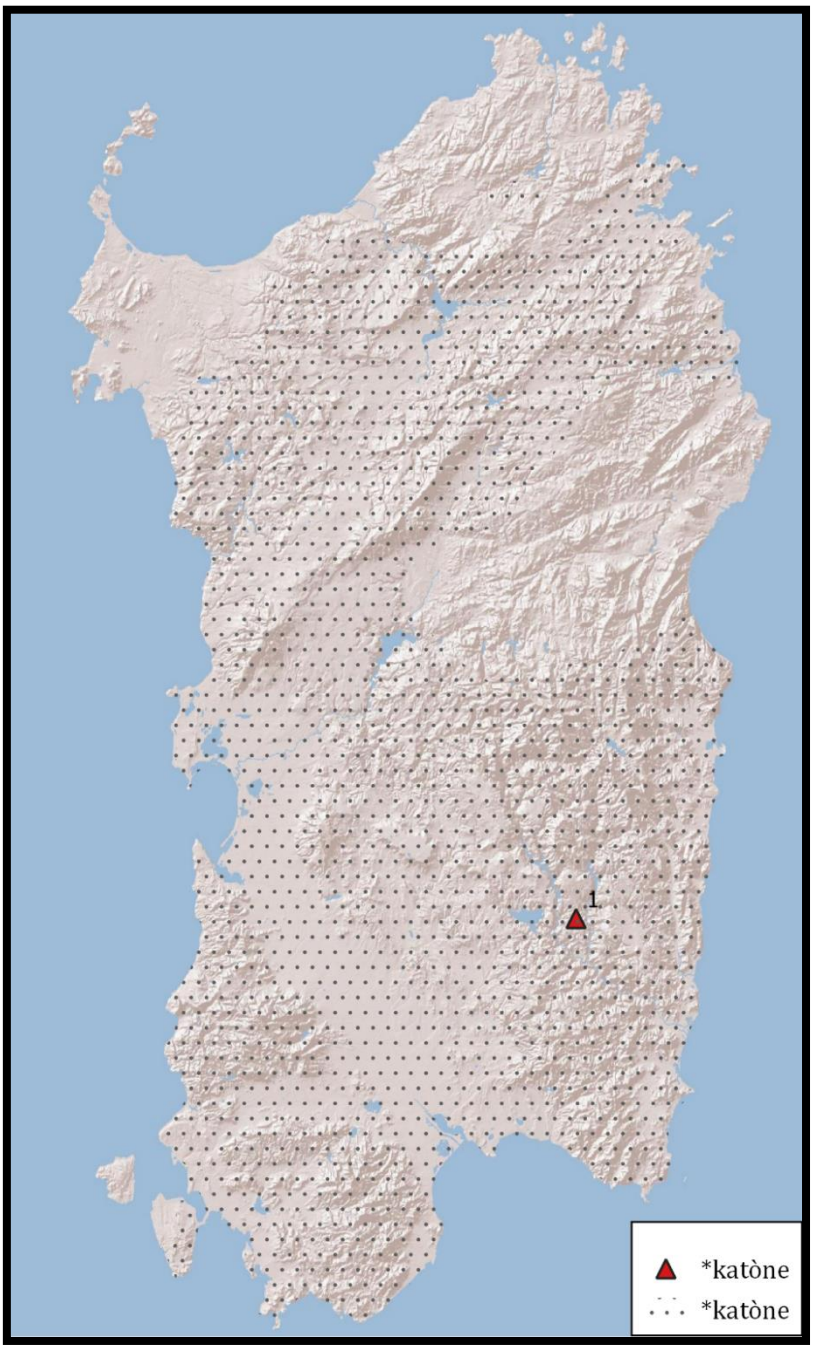

Figure 3.1.1: Annual mercury (*Mercurialis annua*); goosefoot (*Chenopodium sp., Lipandra polysperma*)

The leaves of all above-mentioned species of plants are known to be edible, and most species (except annual mercury) also have tiny edible seeds. This etymon, reconstructible to **katòne*, is widespread in Sardinia. Wagner does not discuss its etymology (simply noting "Et.?"). Pittau (1990: 43) compares it to Etr. *kathuni*, without giving a meaning or source for this word, making it impossible to judge its plausibility. Paulis (1992: 375) connects Srd. *kaðòne* etc. to Lat. *catone* (abl.) 'orache (*Atriplex sp.* L.)' from Dioscorides' *De Materia Medica* (Stadler 1899: 186, 219). Paulis follows André (1956: 77; 1985: 53) in deriving Lat. *catone* from *catus* 'sharp', on account of the pointy leaves of this plant. However, although *catus*, originally a Sabine word (Paulis 1992: 375; de Vaan 2008: 99), is generally translated as 'sharp', it seems to refer to sharpness exclusively in the cognitive sense; i.e. 'sly, cunning etc.' (Lewis & Short 1879: s.v. *catus*). This makes the derivation to 'plant with pointy leaves' semantically rather difficult. Moreover, the plant name *catone* 'orache' is attested only once, in a work that was originally composed in Greek and that is known to contain plant names from other languages. The single attestation of *catone* in a relatively late work renders its attribution to the inherited vocabulary of Latin rather tenuous.

If Adams (2007: 533–540) is correct in proposing an African origin for the Latin translation of *De Materia Medica* (Stadler 1899, 1901), an African origin is imaginable for both Lat. *catone* and Srd. **kaðòne* too. And in fact, we do find Kb. *waktun*, *aktun* 'goosefoot (*Chenopodium sp.*)',[44] in which *wa-* and *a-* respectively are prefixes (cf. Brugnatelli 1998). It is likely that this Berber base *-ktun* 'goosefoot' is somehow related to Lat. *catone* and Srd. *kaðòne* etc. 'id.', but the question is exactly how. Possibly Kb. (*w*)*aktun* 'goosefoot' < **(ā)k(V̆)tūn* is a descendant of an earlier Berber (or Libyco-Berber) word that was borrowed into African Latin as *catone*. Alternatively, both the Berber and the Latin word were borrowed from an unattested Punic word for 'goosefoot'. For Srd. *kaðòne* it is in any case plausible that it entered Sardinian Latin by contact with the local Punic population, regardless whether they spoke Phoenician or Berber (cf. § 10.1, §10.2).

[44] Massinissa Garaoun, p.c., who also informs me that *aktun* is found in Djidjelli Arabic, which contains a significant Berber substrate.

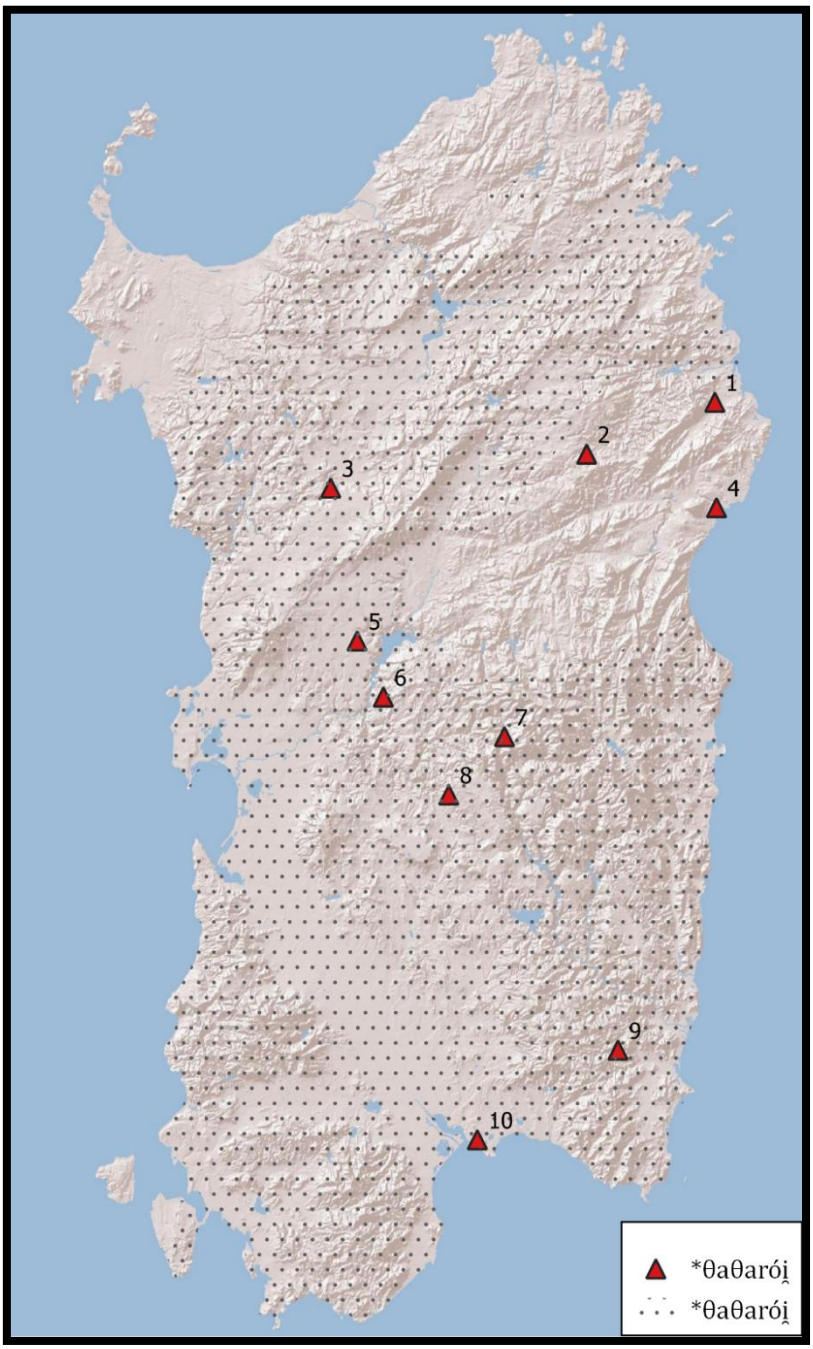

Figure 3.1.2: Arum (*Arum italicum*)

3.1.2 Arum (*Arum italicum*)

Reconstruction	Forms
ϑaϑarói̯	DES (II: 552): ***tattaróyu***, ***tottoróyu*** (Log.), ***sattsarói̯*** (Camp.), ***ϑoϑϑóri̯u*** (1: Siniscola), ***ϑoϑoróyu*** (2: Bitti, 4: Orosei), ***tattaróǧu*** (3: Bonorva), ***tsottsorói*** (5: Norbello, 8: Laconi), ***sottsoróyu*** (6: Busachi), ***tsottsoróyu*** (7: Belvì) 'arum (*Arum italicum* Mill.)'
	Alessio (1949: 111): ***sazzaroi*** (10: Cagliari), ***tattaroiu*** 'id.'
	Böhne (1950: 120): ***sattsarói̯*** (Sarrabus reg.) 'id.'

These Sardinian forms, which are found throughout the island, have generally been attributed to a pre-Roman language, due to their peculiar formation and the presence of **ϑ*. Within Sardinian, they can go back to **ϑaϑarói̯*; the pretonic o-vocalism found in some forms are likely secondary, as well as the word-final *-u*. They have been connected by Alessio (1949: 110–111) to Sic. *anzara/-u* 'arum (*Arum italicum*)',[45] Sic. *azzaru*, *azareddu*, *anzareddu* 'friar's cowl (*Arisarum vulgare* O.Targ.Tozz.)', Calab. *zara*, *anzara* (Cetraro, Verbicaro), *anzana* 'arum; cyclamen' (cf. Rohlfs 1932: 99), Pugl. *azara* 'arum', to which Hubschmid (1953: 27) has added Calab. *źanźarra* 'arum' (cf. Rohlfs 1932: 409; 1977: 84b, 802b). Alessio (1949: 110–111), Hubschmid (1953: 27), Wagner (DES II: 552) and Paulis (1992: 341) agree on deriving the Sardinian and South-Italian forms from a common pre-Roman substrate. This requires an analysis of Srd. **ϑaϑarói̯* as < **ϑ-aϑar-ói*, with prefixation of **ϑ(i)-* (§ 9.1.2) and a suffix *-Ví* (§ 9.2.1). The Pre-Roman Sardinian base **-aϑar-* would then correspond to **azar-* in the South-Italian forms. However, Pittau (1995: 107) proposes the Italian forms to be from Lat. *lancearius* 'lancer', due to its upright shape, with reinterpretation and detachment of initial *l-* as the article. If this is correct, the South Italian forms are of no relevance to Srd. **ϑaϑarói̯*.

The situation is complicated by the attestation of the form *asaru* in the Latin version of Dioscorides' *De Materia Medica* (Stadler 1899: 238). In this work, the plant named *asaru* is described as similar to but smaller than the arum. According to Stadler (1899: 238 fn. 1) this name should have been *arisarum* 'friar's cowl (*Arisarum vulgare*), which is what we find in the Greek version (ἀρίσαρον; cf. Dioscorides Anazarbei 1907: 234). Moreover, there is another chapter titled "*De asaro*" in the first book of *De Materia Medica* (Hofmann & Auracher 1883: 60), which corresponds to ἄσαρον in the Greek version (Dioscorides Anazarbei 1907: 14), which has the meaning 'hazelwort (*Asarum europaeum* L.)'. In short, where the Greek original has ἄσαρον 'hazelwort (*Asarum europaeum*), ἄρον 'arum (*Arum sp.*)', and ἀρίσαρον 'friar's cowl (*Arisarum vulgare*)' (Dioscorides Anazarbei

[45] Domenico Muscianisi informs me that in Messina (Sicily), *zara* is also used for 'arum'.

1907: 14, 233, 234), the Latin translation has [*De*] *asaro* 'hazelwort', [*De*] *aron* 'arum', and [*De*] *asaru* 'friar's cowl' (Hofmann & Auracher 1883: 60; Stadler 1899: 237, 238). While a confusion between *asaru* and expected *arisarum* on the part of the translator of the Latin text is possible, the form *asaru*, which refers to a species of plant that is closely related to the arum and is indeed quite similar to it, is very reminiscent of the South-Italian names for the arum mentioned above. This raises the question whether *asaru* in *De Materia Medica* is in fact not the result of a mistake, but a Vulgar Latin word for a species of arum, dispersed at least in Southern Italy and Africa.[46]

If *asaru* is indeed related to the southern Italian forms, this precludes Pittau's (1995: 107) derivation from Lat. *lancearius*, as the needed phonological developments are not expected to have been completed by Dioscorides' time already. It is unclear whether *asaru* 'species of arum' is etymologically related to Gr. ἄσαρον 'hazelwort (*Asarum europaeum*)', whence Lat. *asarum* 'id.'. Hazelwort and arum are not closely related, but could be perceived as superficially similar because of their heart-shaped, glossy leaves and tube or jug-shaped flowers. Gr. ἄσαρον has been argued to be of Semitic origin by Lewy (1895: 47), who compares it to Aramaic *zərīr* 'sneeze'. This is formally and semantically unconvincing.

If the Sardinian forms are to be connected to (Late) Lat. *asaru*, Sic. *anzara/u*, Calab. *zara*, *anzara*, *anzana*, *źanźarra* and Pugl. *azara*, we need to assume the addition of the Pre-Roman Sardinian affixes **ϑ(i)-* and **-V̊ḭ̊* to the base **aϑar-*, as proposed by Alessio (1949: 111), Hubschmid (1953: 27), and Wagner (DES II: 552). In light of the necessary assumptions, and the etymological uncertainty regarding the southern Italian forms, it is safest to treat the relation between Srd. **ϑaϑaróḭ̊* and the southern Italian forms, as well as Lat. *asaru*, as no more than a possibility. The connection of Sic. *anzara/-u* 'arum (*Arum italicum*) etc. to Lat. *asaru* 'friar's cowl (*Arisarum vulgare*) is in any case plausible.

[46] That is, if Adams (2007: 533–540) is right in attributing this version of De Materia Medica to an African translator (cf. § 3.1.1).

3.1.3 Betony (*Stachys glutinosa*)

Reconstruction	Forms
1: *lokáṡ-	DES (II: 611): ***lokáṡi*** (2: Sassari) 'hyssop (*Hyssopus officinalis* L.)', ***lokáṡu*** (6: Nuoro) '*Stachys glutinosa* L.'
	Paulis (1992: 115): ***lokáṡu*** (5: Orune, 8: Sarule), ***lokáu*** (9: Urzulei), ***lukéṡu*** (10: Aritzo) '*Stachys glutinosa*'
2	DES (II: 611): ***lukréžu*** (10: Aritzo) '*Stachys glutinosa*'
3	Paulis (1992: 115): ***kalekáṡu*** (1: Berchidda) 'Italian strawflower (*Helichrysum italicum* (Roth) G.Don.)', ***kalakáṡu*** (5: Orune) '*Stachys glutinosa*'
4	Paulis (1992: 115): ***vrattakáṡu*** (3: Siniscola) '*Stachys glutinosa*'
5	Paulis (1992: 115): ***olokáṡu*** (4: Lula), ***alakáṡu*** (7: Orani) '*Stachys glutinosa*'
6	Paulis (1992: 115): ***lukkíttu*** (11: Meana Sardo) '*Stachys glutinosa*'

Stachys glutinosa is a species of betony endemic to Sardinia, Corsica, and smaller surrounding islands. Its names in the Sardinian dialects are diverse. The forms listed here roughly cover the Barbagia and Baronie areas, with outliers in Berchidda (1) and Sassari (2). The DES follows Hubschmid (1960b: 172) who considers *lokáṡu* to be pre-Roman (without mentioning the other forms). Paulis (1992: 115) adds various forms that complicate the picture. He argues that any pre-Roman features of these words could have been introduced from other Sardinian phytonyms. He argues that *lukkíttu* (type 6; 11: Meana Sardo) is a loan from Cat. *lluquet* 'bundle of dry and sulfated twigs or straw that easily catches fire', and gives Cat. *lluqueta* 'heart-leaved globe daisy (*Globularia cordifolia* L.)' as a parallel. Similarly *lukréžu* (type 2; 9: Aritzo) would reflect It. *elicriso* 'everlasting (*Helichrysum sp.* Mill.)', with influence of Srd. *lúži* 'light'. Paulis considers the remaining types to be derived from *lukéṡu* (type 1; 10: Aritzo). Forms like *lokáṡu* (type 1), *kalekáṡu* (type 2), *vrattakáṡu* (type 4) and *olokáṡu* (type 5) would have been influenced by Srd. *káṡu* 'cheese', because this plant's use to remove chicken lice supposedly led to the comparison of falling lice to grated cheese (Paulis 1992: 112). It is clear that many secondary folk-etymological influences have affected the various shapes of these forms, although I am not quite convinced that Paulis' proposals succeed to explain all forms satisfactorily. Due to the complexity of the dialectal variation and of this word and the resulting unclarity as to its original shape however, it cannot be used to provide us with any reliable information about the pre-Roman language(s) of Sardinia.

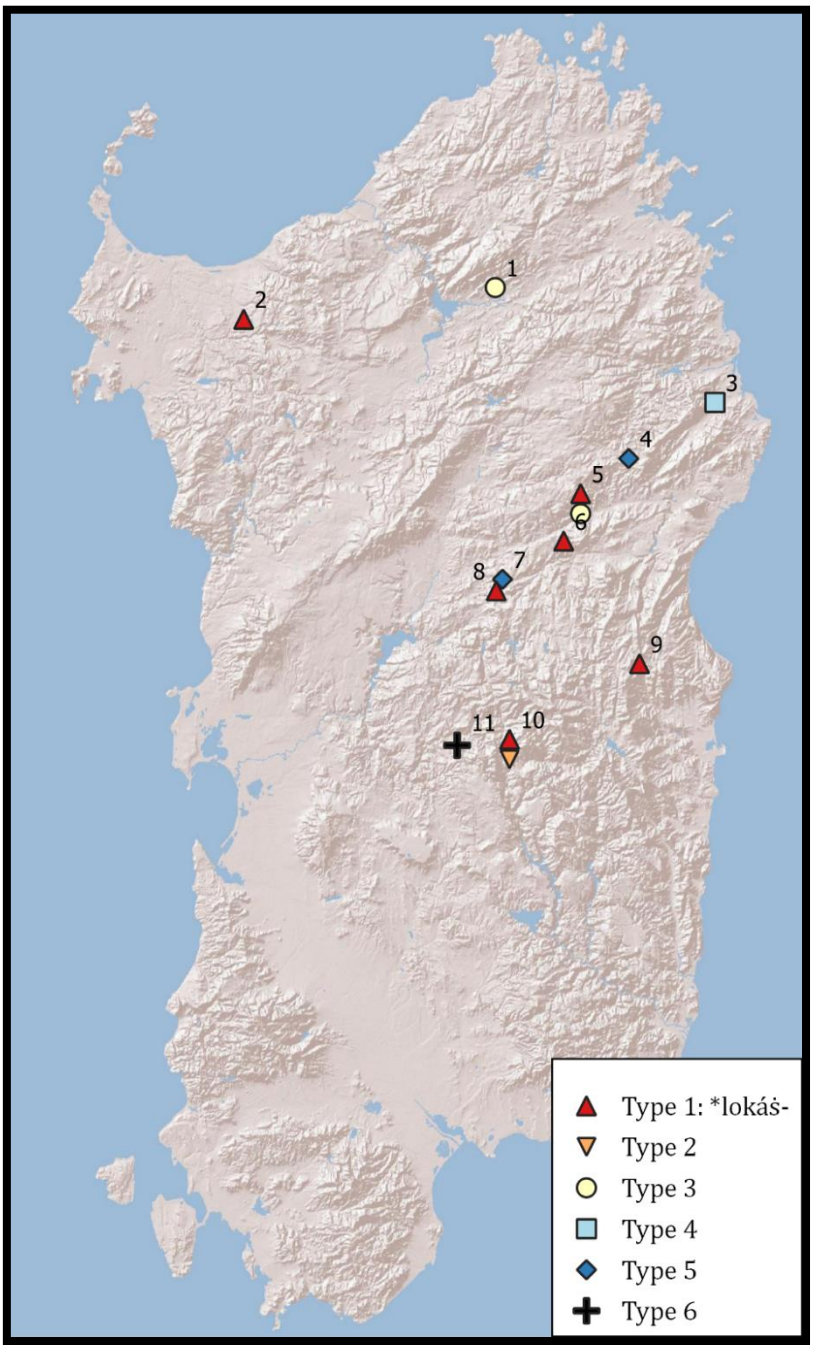

Figure 3.1.3 Betony (*Stachys glutinosa*)

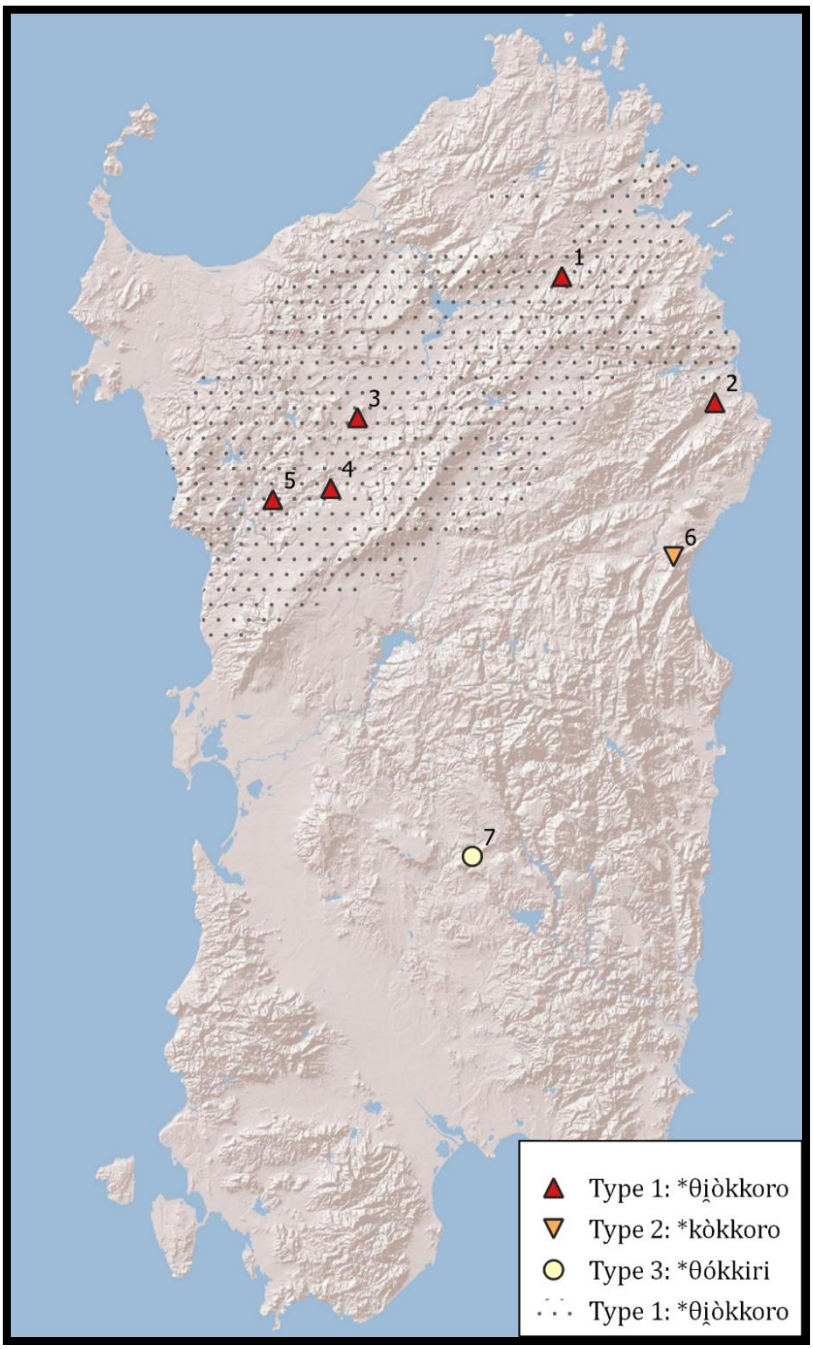

Figure 3.1.4: Bristly ox-tongue (*Helmintotheca echioides*), wild teasel (*Dipsacus fullonum*)

3.1.4 Bristly ox-tongue (*Helmintotheca echioides*), wild teasel (*Dipsacus fullonum*)

Reconstruction	Forms
1: *ϑi̯òkkoro	DES (II: 548): ***artiòkkoro*** (Log.), ***išòkkoro*** (NLog., 1: Monti, 3: Mores) ***ϑiòkkoro*** (Nuor., 2: Siniscola), ***istiòkkoro*** (4: Bonorva, 5: Padria) 'bristly ox-tongue (*Helminthotheca echioides* (L.) Holub); wild teasel (*Dipsacus fullonum* L.)'
	Hubschmid (1953: 36): ***čòkkoro*** (Log.) 'bristly ox-tongue (*Helminthotheca echioides*)'
2: *kòkkoro	Paulis (1992: 192): ***kòkkoro*** (6: Dorgali) 'bristly ox-tongue (*Helminthotheca echioides*)'
3: *ϑókkiri	DES (II: 548): ***sóčči̯ri*** (Camp.) 'id.'
	Paulis (1992: 190, 192): ***čóčči̯ri*** (7: Isili) 'bristly ox-tongue (*Helminthotheca echioides*); wild teasel (*Dipsacus fullonum*)'

Paulis (1992: 190) suggests that these words, which are generally taken to be of pre-Roman origin (cf. Hubschmid 1953: 36; DES II: 548; Paulis 1992: 190), may originally have referred specifically to 'bristly ox-tongue', as it is the only Sardinian word denoting this species. Its extension to 'wild teasel' would be due to the fact that both plants are covered in small spines or bristles, and to the similarity of their leaves. Paulis (1992: 192) proposes the forms of type 2 and 3 to be influenced by the Sardinian reflexes of Lat. *cicer* 'chickpea' (> Camp. *čížiri*), as *cicer rusticum* is attested in several classical sources as one of the Latin names for 'wild teasel' (e.g. Dioscorides Anazarbei 1906: 18; cf. further Paulis 1992: 192). This is likely at least for Camp. *sóčči̯ri, čóčči̯ri* (7: Isili). However, Lat. *cicer* was not preserved in Logudorese and Nuorese (Paulis 1992: 281), so assuming influence from unattested **kíkere* 'chickpea' to explain *kòkkoro* (6: Dorgali) < **ϑiòkkor*(*o*) is slightly problematic. Perhaps the Dorgalese variant is the result of consonant assimilation.[47] A pre-Roman origin of this word is suggested by the sequence **ϑio-* that must be reconstructed. This cannot go back to any Latin sequence, as it would require Lat. **ci̯io-*/**ti̯io-*, which did not exist (§ 8.1.1.1). Outside Sardinian, Hubschmid (1953: 36) compares Bq. *txokorro* 'milk thistle (*Silybum marianum* (L.) Gaertn.)' and various other Basque and Romance forms denoting 'trunk',

[47] Dorgalese *kòkkoro* 'bristly ox-tongue' is formally identical to Barb. *kòkkoro* 'nut', from which *kokkoròni* (Dorgali) 'bump' is derived (DES I: 356–57). Establishing a semantic link between 'bristly ox-tongue' and 'nut/bump' is difficult however.

'branch' or 'knot'. I share Wagner's (DES II: 548) and Paulis' (2008: 17) hesitation in accepting these comparisons.[48]

3.1.5 Cat thyme (*Teucrium marum*), cotton lavender (*Santolina chamaecyparissus*)

Reconstruction	Forms
**1: ****múkra*	DES (II: 139): ***murga*** **[*de bovis*]** (5: Aritzo) 'cotton lavender (*Santolina chamaecyparissus* L.)'
2: ****múkru véru* **(?)	DES (II: 139): ***murgulèu*** (Ogl., Camp.) 'cat thyme (*Teucrium marum* L.); cotton lavender (*Santolina chamaecyparissus*)'; ***murguèu*** (11: Cagliari) 'wine'
	Paulis (1992: 118): ***muyuòes*** (6: Meana Sardo) 'Massilian germander (*Teucrium massiliense* L.)', ***murguéu*** (8: Seulo) 'cotton lavender (*Santolina chamaecyparissus*)', ***murguèus*** (9: Laconi) 'everlasting *Helicrysum sp.*'
3: ****mámmu véru* **(?)	Paulis (1992: 118): ***mummuléu*** (1: Lula) 'cat thyme (*Teucrium marum*)', ***murmuréu*** (2: Dorgali, 3: Oliena, 4: Villagrande Strisaili, 7: Arzana) 'cotton lavender (*Santolina chamaecyparissus*)', ***mammuléu*** (6: Meana Sardo, 9: Laconi) '*Micromeria graeca* (L.) Benth. ex Rchb.',[49] ***murmuréu*** (10: Escalaplano) 'species of betony (*Stachys glutinosa*)'[50]

The various plant species referred to by the forms listed above are not all closely related, but are superficially somewhat similar in their size, shape, and the colour of their leaves. DES (II: 139) regards these words as "probably pre-Roman" because of a toponym *Murgueliai*. Paulis (1992: 118) points out that there is no evidence that this toponym is connected to the phytonym, and that the presence of a pre-Roman lexical item in a toponym does not prove its substrate origin. For the forms of type 3, he instead proposes a derivation of Srd. *mammuléu* (6: Meana Sardo, 9: Laconi) '*Micromeria graeca*' from **maruléu* << Lat. **marum verum* 'real cat thyme', by assimilation of the **-r-* in **marum* to the initial *m-* (Paulis 1992: 119). This formation would be parallel to Srd. *karduréu* 'thistle' < Lat. *carduum verum* 'real thistle' (cf. DES I: 300).

[48] I have not been able to verify Hubschmid's (1953) mention of Bq. *txokorro* 'milk thistle'.
[49] Paulis gives the scientific name *Satureja graeca* L., which is a synonym of *Micromeria graeca* (L.) Benth. ex Rchb. (WFO 2024c: s.v. *Micromeria graeca*).
[50] Paulis gives the binomial *Betonica glutinosa*. This does not exist as a scientific name in WFO (2024a) however, and should probably be interpreted as *Stachys glutinosa* instead — the common Italian name for *Stachys* species is *betonica*.

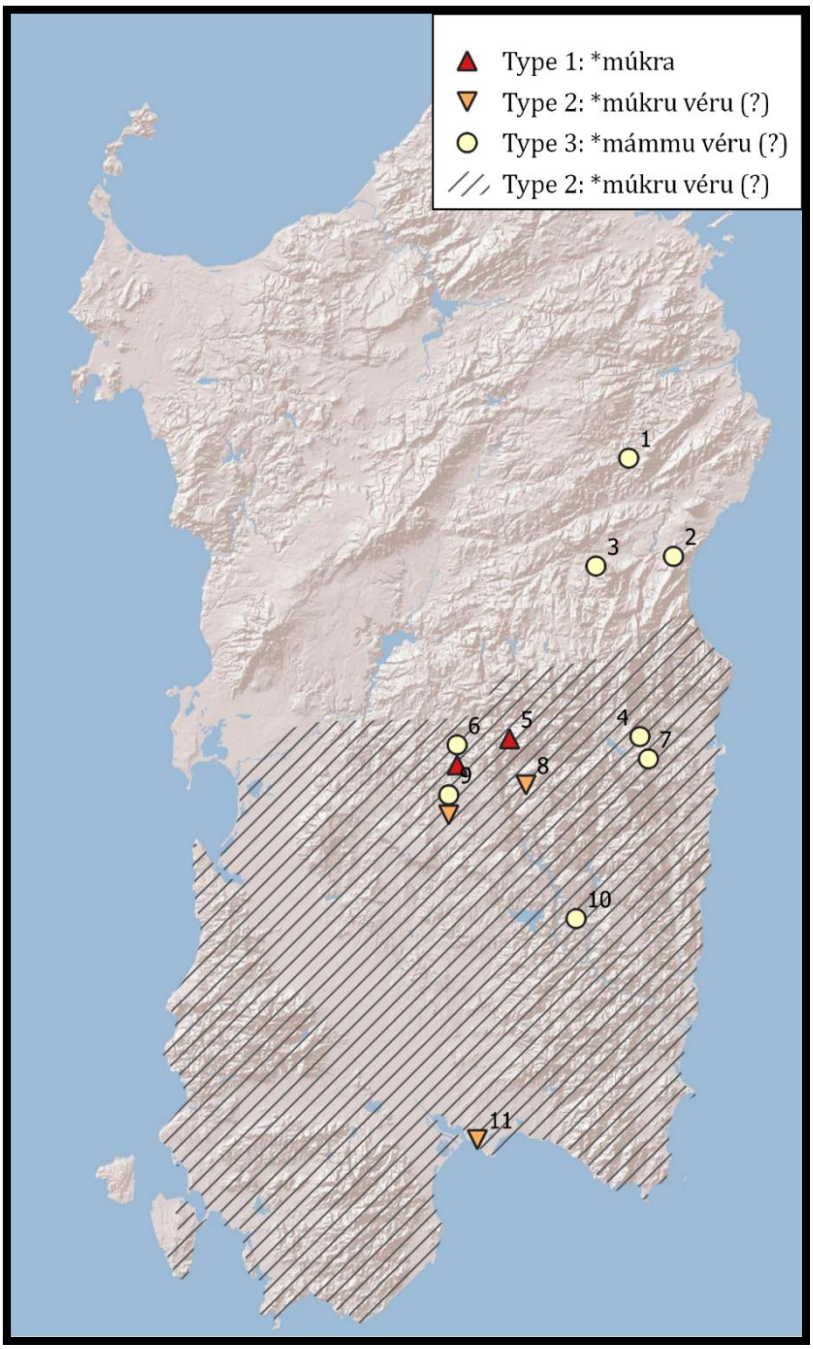

Figure 3.1.5: Cat thyme (*Teucrium marum*), cotton lavender (*Santolina chamaecyparissus*)

Concerning the development of *verum* > -(*u*)*léu*/-(*u*)*réu*/-(*u*)*éu* in types 2 and 3, there are parallels for *-r-* dissimilating to *-l-* (Wagner 1941a: 232) and *-r-* disappearing altogether (*arišéu* < *erišéru* 'yesterday evening'; Wagner 1941a: 229 fn. 1). It is peculiar, however, that this development would occur in *mammuléu* < **mammuréu*, where there is no other *-r-* to dissimilate from. Especially since a similar dissimilation and dissolution of *-r-* does not occur in *karduréu* etc. 'thistle', whose base *kardu-* does contain an *-r-*.[51] Paulis accounts for this discrepancy by assuming original **maruléu* << *marum verum*, which would create the conditions for a dissimilation. Note however that, if the *-l-* is to be explained as the dissimilatory reflex of *-r-* in presence of an *-r-* in the stem, it is all the more striking that we still find *murmuréu*. If we adhere to Paulis' explanation of the various forms, we would need to assume the following steps to account for all attested forms:

- *marum verum* > **maruréu* > **maruléu* > *mammuléu* (> *mummuléu*)
- *marum verum* > **maruréu* > **maruréu* > **mammuréu* > **mummuréu* > *murmuréu*

While theoretically possible, it is suspicious that none of the necessary intermediate steps is attested across the dialectal variation found in this word. Moreover, the only certain example of progressive dissimilation of *r – r* > *r – l* in Wagner (1941a: 232) is *ferréli* < *ferréri* 'smith' in Perdasdefogu, which is outside the attestation area of the forms under discussion. These cumulative difficulties cast doubt on the possibility of the sequence *-uléu/-uréu/-uéu* in these words going back to Lat. *verum* 'true', although I cannot offer a better alternative explanation. The similarity between *-uréu* < *verum* in *karduréu* 'thistle' etc. and *-uléu/-uréu/-uéu* in the words for 'cat thyme, cotton lavender etc.' is remarkable. For now, the best option is therefore to assume *verum* as the origin of the ending *-uléu/-uréu/-uéu*, even though the details of the phonological development are not quite clear.

This leaves us with the first half of the word, which shows considerable variation. As mentioned before, Paulis (1992: 119) wants to derive *mammuléu* (6: Meana Sardo, 9: Laconi) '*Micromeria graeca*' from Lat. *marum verum* 'true cat thyme'. The other forms of type 3 would be due to subsequent alterations of *mammuléu* to *mummuléu* (1: Lula) by assimilation of the *-a-* to the next syllable, and hence to *murmuréu* (2: Dorgali, 3: Oliena, 4: Villagrande Strisaili, 7: Arzana, 10: Escalaplano) by *-r-*insertion, commonly found in Sardinian (Paulis 1992: 119). I have not

[51] The only example of dissimilation of *-r-* to *-l-* in the word for 'thistle' occurs in *garduléu*, which is restricted to Siniscola (Paulis 1992: 67).

been able to find any parallel for the assimilation of intervocalic *-r-* to *-m-* that is needed to derive *mammuléu* from **maruléu* < *marum verum* (cf. Wagner 1941a: 230–232). While the derivation of *mummuléu* and *murmuréu* from *mammuléu* do not pose any serious problems, an etymological relation to Lat. *marum* 'cat thyme' is problematic.

For the forms of type 1 and 2, viz. respectively *murga* [*de bovis*] (5: Aritzo) 'cotton lavender' and *murguèus* (9: Laconi) 'everlasting' etc., a derivation from Lat. *marum* is even more difficult. The forms in these types are found in Campidanese, and all contain a base **mukr-*, which can in turn be from original **mukr-*, **mukl-* or **mutl-*; cf. Camp. *sórgu* < Lat. *socru-* 'father in law', Camp. *márga* < Lat. *macula* 'stain' and Camp. *múrgu* 'tree trunk' < Lat. *mutilus* 'maimed' (cf. Wagner 1941a: 157, 167; DES II: 132). There are some inherited Sardinian words that are nearly or completely homophonous with this base *murg-* found in types 1 and 2. One is Camp. *múrgu* 'tree trunk' < Lat. *mutilus*, which Paulis (1992: 120) thinks could have influenced original **maruléu/mammuléu* < *marum verum*, yielding *murguléu* etc. The other is Camp. *múrga* 'dregs' < Lat. *amurga, amurca* 'id.' (DES II: 138). The latter form cannot have yielded *muɣuòes* (6: Meana Sardo) regularly, because the *-r-* in original **-rg-* is expected to persist. Semantically the link between 'dregs' and the plant species discussed is difficult to understand too. Paulis further (1992: 120) proposes that Cat. *múrgola* 'morel (*Morchella sp.* Dill. ex Pers.)' could have exerted influence, by being applied first to dry fungi used to start fires and hence to plants used for the same purpose. Neither 'tree trunk', nor 'morel' bears a clear semantic connection to the plant species referred to by *múrga*, *murguléu* etc., and various assumptions are needed to uphold these etymologies.

In conclusion, Wagner's (DES II: 139) assumption of a pre-Roman origin on the basis of the structure of these words is unnecessary. The ending *-uléu/-uréu/-uéu* might ultimately be the same as the *-uréu* in *karduréu* 'thistle' (< Lat. *carduum verum*), even though the details surrounding the development of *-r-* remain unclear. The base of the forms in types 1 and 2 (in *murg-*) may have arisen through influence from Camp. *múrgu* 'tree trunk' and Cat. *múrgola* 'morel' (cf. Paulis 1992: 120). For the forms in type 3, I do not think that Lat. *marum* 'cat thyme' can be upheld as an etymology, since it is based more on semantic proximity than on regular correspondence. Whether this element, which alternates between *mamm-*, *mumm-* and *murm-*, is of pre-Roman origin after all, cannot be established on the basis of the formal appearance of the forms, even though a convincing etymology is still lacking.

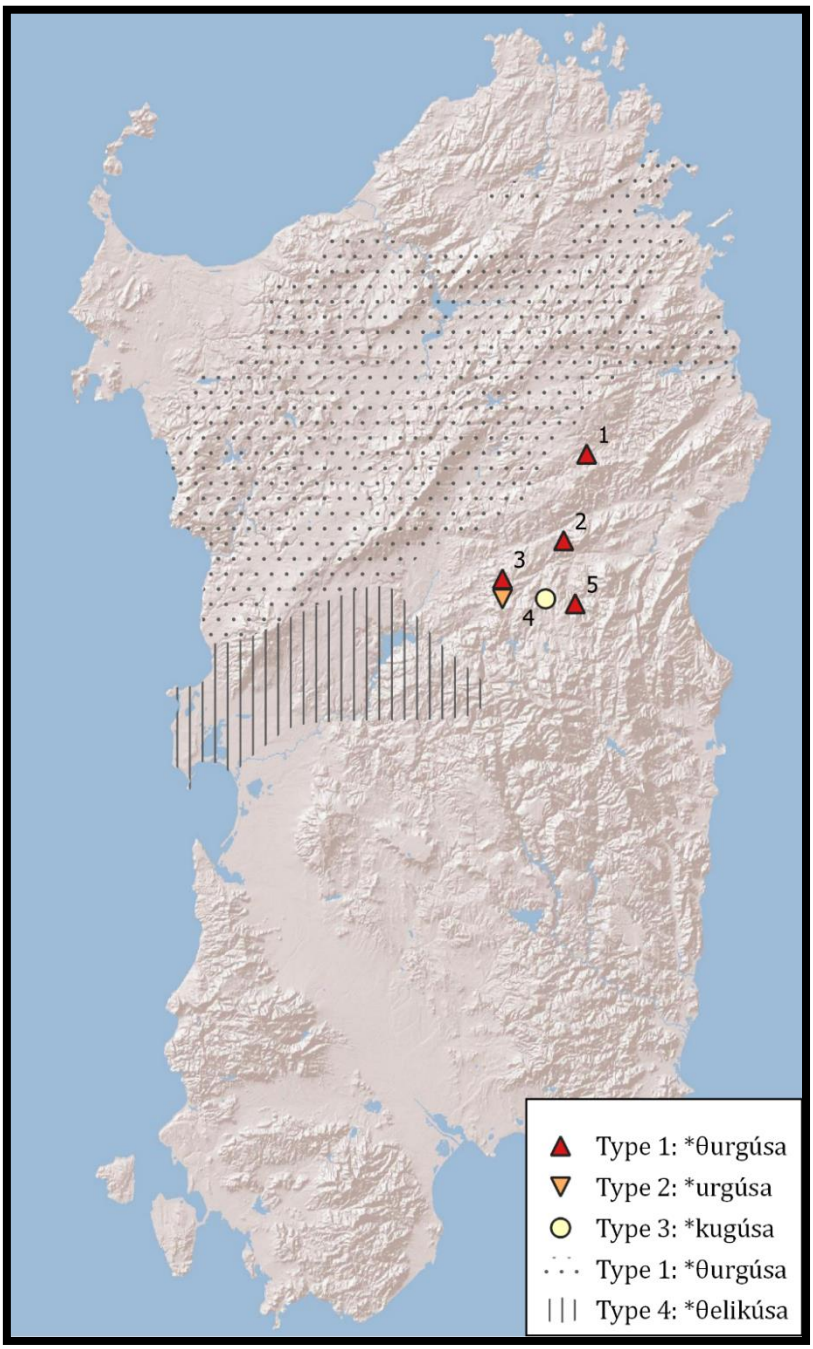

Figure 3.1.6: Celery (*Apium graveolens*), fool's watercress (*Apium nodiflorum*)

3.1.6 Celery (*Apium graveolens*), fool's watercress (*Apium nodiflorum*)[52]

Reconstruction	Forms
1: *ʤurgúsa	DES (II: 555): ***ʤurgúsa*** (1: Bitti), ***ʤrugúsa*** (2: Nuoro) 'wild celery (*Apium graveolens* L.)', ***ʤurgúsa*** (5: Orgosolo) 'water hemlock (*Cicuta sp.* L.)'
	Paulis (1992: 141): ***turgusònе*** (Log.), ***ʤurgúsa*** (3: Orani) 'wild celery (*Apium graveolens*)'
	Puddu (2023: s.v. *tirgúsa*, *tzurgúsa*): ***tirgúsa***, ***trugúsa***, ***turgúsa***, ***tzurgúsa***, ***tzurvúsa*** 'fool's watercress (*Apium nodiflorum* (L.) Lag.)'; ***turgusòne***, ***tzurgúsa*** 'celery (*Apium graveolens*)'
2: *urgúsa	Blasco Ferrer (1988: 175): ***uryúsa*** (3: Orani) 'fool's watercress (*Apium nodiflorum*)'
3: *kugúsa	DES (I: 422): ***kugúṡa*** (4: Mamoiada) 'fool's watercress (*Apium nodiflorum*),[53] wild celery (*Apium graveolens*)
	Puddu (2023: s.v. *tirgúsa*, *tzurgúsa*): ***cugúsa*** 'fool's watercress (*Apium nodiflorum*), celery (*Apium graveolens*)'
4: *ʤelikúsa	Paulis (1992: 141): ***tseliɣúṡa*** (Arborense) 'wild celery (*Apium graveolens*)'
	Puddu (2023: s.v. *tirgúsa*, *tzurgúsa*): ***tzeligúsa*** 'id.'
5: *ʤikúsa	Puddu (2023: s.v. *tirgúsa*, *tzurgúsa*): ***atigúsa*** 'celery (*Apium graveolens*)'

The cited forms refer to a number of plant species. The forms from Bitti (1), Nuoro (2), and Mamoiada (4) all refer to species of the genus *Apium*. For the variant from Orgosolo (5), Wagner gives the meaning 'water hemlock (*Cicuta sp.*)'. This meaning is not represented by Puddu, who only has 'fool's watercress' and 'celery'. The lexical forms themselves have in common that they end in *-gúsa*. If they are etymologically related, we can observe an alternation between initial *ku-* vs. *ʤur-* (*ʤru-* is due to metathesis).

Besides the forms presented by Wagner, Blasco Ferrer (1988: 175) records *uryúsa* 'fool's watercress', found only in Orani (3). Its relation to *ʤurgúsa*, also attested in Orani, is not quite clear; Paulis (1992: 141) proposes that the *ʤ-* may have been detached by confusion with the article. The alternating presence of **ʤ-* suggests the presence of the pre-Roman element **ʤ(i)-* (§ 9.1.2). An interpretation **ʤ-urgúsa* would preclude a direct comparison to **kugúsa*, **ʤelikúsa* etc., unless

[52] These data are discussed in a similar fashion in Swanenvleugel (2024: 226–228).

[53] Older sources often use the binomial *Helosciadium nodiflorum* (L.) W.D.J.Koch, which is a synonym of *Apium nodiflorum* (cf. WFO 2021: s.v. *Helosciadium nodiflorum*).

**urgúsa* is assumed to be some kind of compound with **k/gúsa* in other forms. This is of course purely speculative. Paulis (1992: 141) lists *turguṡòne* for Logudorese and *tseliɣúṡa* for South Logudorese, which corresponds to Virdis' Arborense variety (Virdis 1988: 912). *Tseliɣúṡa* 'celery' also has *-gúsa* but starts in *tseli-*, which could be the element *ϑili-* that is found in various non-inherited words (cf. § 9.1.2). Puddu (2023) has many forms, most of which can be compared to the forms attested in Wagner and Paulis. An exception to this is *atigúsa* 'celery'. Its initial sequence *ati-* likely represents pre-Roman **ϑi-* (cf. Macomer *attiliɣèrta* < **ϑilikèrta*).

Paulis (1992: 141–142) treats *ϑurgúṡa* and *kuɣúṡa* as unrelated words. *Kuɣúṡa* would be identical to *kukúsa* 'flake'. The problem is that this is expected to yield ***kuʔúṡa* in Mamoiada. Paulis proposes a connection between *ϑurgúṡa* 'wild celery' and *orgòsa*, *urgúsa* 'source of water' (§ 5.1.7). These words are themselves of obscure origin, so this connection is speculative. It is rather unattractive to separate *ϑurgúṡa* and *kuɣúṡa*, which both contain **-gúsa* and which both mean 'wild celery'. If all of the presented forms are taken to be etymologically related, we are dealing with an alternation between initial *ku-*, *ϑur-*, *ur-*, **ϑili-* and **ϑi-*. Except for the *ϑur-*, this coincides with alternations observed in some other lexical items (cf. § 9.1). Whether *ϑur-* is some variant of **ϑili-*, or a different morpheme altogether in the source language, is difficult to establish on the basis of one word. The Logudorese and Arborense forms listed by Paulis and Puddu (*turguṡòne* and *tseliɣúṡa*) show that the distribution of this family of words for plants of the genus *Apium* must be rather larger than Wagner's attestations would suggest. Pittau (1995: 209) proposes a shared pre-Roman origin for the Sardinian forms and Lat. *cicūta* 'poison hemlock'. This is formally problematic.[54]

One difficulty in the comparison of Wagner's forms to those listed by Paulis and Puddu is the reconstruction of the velar obstruent. Types 1, 2 and 3 all need an original **-g-*, since **-k-* would either have been preserved, or have yielded ***-ʔ-* in Mamoiada and Orgosolo (Wagner 1941a: 71). The same goes for the *-g-* in Logudorese *turguṡòne*, where **-k-* is also expected to be preserved post-consonantally (Wagner 1941a: 175). In Arborense *tseliɣúṡa* and probably also *atigusa* (if it is Logudorese), the *-ɣ-/-g-* should go back to **-k-*, since intervocalic **-g-* would have disappeared in these dialects (Wagner 1941a: 79). Perhaps the preservation of *-g-* in the Logudorese and Arborense dialects is due to influence

[54] Pittau (1995: 209) also includes Srd. *aússa* 'poison hemlock' (*Conium maculatum* L.) in this comparison (see § 3.1.10). Lat. *cicūta* has been compared to Celtic forms (Pedersen 1909: 209; Wigman 2023: 193).

from forms like Camp. *kuɣúttsula* 'wild artichoke' < Lat. *cucutium* 'hood, cap', as Paulis (1992: 142) suggests. It is in any case plausible that the sequence *-*gusa* (or *-*kusa*) in the discussed forms is etymologically one and the same pre-Roman element. The alternations of the initial sequence are not isolated, and may indicate pre-Roman morphological patterns (§ 9.1).

3.1.7 False fennel (*Ridolfia segetum*)

Reconstruction	**Forms**
***ʒikkiría**	DES (II: 589): ***tsikkiría*** (Camp.), ***tsirikkía*** (1: Villacidro) 'a species of dill/false fennel (*Ridolfia segetum* Moris)'[55]

This word has been connected in DES (II: 589) to the word σικκιριά 'dill (*Anethum graveolens*)' cited in *De materia medica* (Dioscorides Anazarbei 1906: III: 58 RV) as a Punic word (i.e. "Αφροι").[56] The close formal and semantic match between the Sardinian and the Punic forms make it very likely that the Sardinian forms go back directly to Punic spoken on Sardinia. Blau (1873: 527) does not, as Paulis (1990: 601; 1992: 167) suggests, discuss the Sardinian forms, but he does propose a connection between the Punic word and Hebr. שֵׁכָר (*šēkār*) >> Gr. σίκερα 'fermented liquor, strong drink' (cf. Liddell et al. 1978: s.v. σίκερα). The descendants of Proto-Semitic *šikar- 'alcoholic drink' and the related verbal root *škr- 'to become inebriated' are widely attested in the Semitic languages (Kogan 2012a: 241). An etymological connection between Pun. σικκιριά 'dill' and Hebr. שֵׁכָר (*šēkār*) therefore requires a semantic shift from 'alcoholic drink' to 'dill' on the Punic side. Whether this is accepted or not, does not alter the likelihood of a Punic origin for the Sardinian forms.

[55] DES defines this word as "a species of dill" ("*una specie di aneto*"), but gives the scientific name *Ridolfia segetum*, which is in fact 'false fennel' rather than 'dill (*Anethum graveolens* L.)'. The plants are rather similar in appearance and odor though.

[56] Wellmann (Dioscorides Anazarbei 1906: III: 58 RV) has σικκιριά, with the accent on the final syllable. Blau (1873: 527) cites σικκίρια however, with proparoxytone accent and Wagner's (DES II: 589) form is yet different, with paroxytone σικκιρία. Wagner's form may have been influenced by Sardinian paroxytone *tsikkiría*. It is not clear to me whether to follow Blau or Wellmann's edition of Dioscorides Anazarbei, but it is of little consequence to the etymological discussion of this word. There is also Pun. σιχκιριάμ 'parsnip' (Dioscorides Anazarbei 1906: III: 52 RV), which may or may not be related (cf. Múrcia Sànchez 2011: 443).

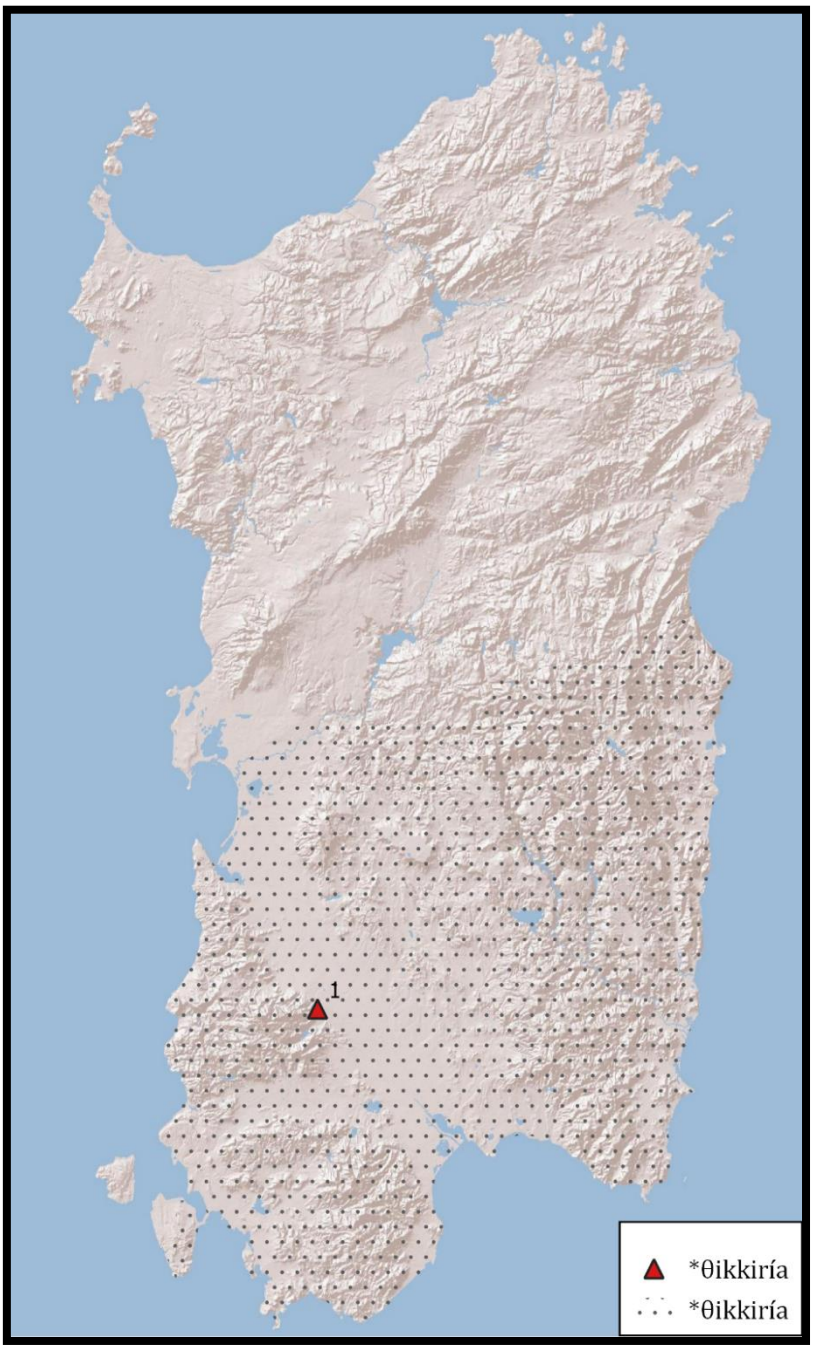

Figure 3.1.7: False fennel (*Ridolfia segetum*)

A peculiar feature of the Sardinian forms is their lack of expected palatalization, as **tsikkiría* would be expected to yield ***tsiččiría*. This does not discredit its etymology in Punic σικκιριά though. In Berber, Múrcia Sànchez (2011: 432–433) compares Tsh. *taskra* 'thistle with small and very prickly leaves; *Echinops spinosissimus* Turra' (cf. van den Boogert 1998: 178, 184),[57] <*tāsakra*, *tāsikrā*, *sakrā*> 'black chamoeleon (*Cardopatium corymbosum* Pers.)' (*'Umdat aṭ-ṭabīb*; cf. Tilmatine & Bustamante 2002: 433), <*taskrat*, *teskera*> '*Echinops spinosissimus*' (Trabut 1935: 96). He reconstructs these forms as **tā-sĭkrā* < **sĭkkĭryā*,[58] which would correspond to Punic σικκιριά. However, as Múrcia Sànchez (2011: 433) himself notes, the fact that the Berber forms refer to prickly, "thistle-like" plants, while the Punic and Sardinian forms refer to 'dill' and 'false fennel', renders the comparison semantically rather difficult.

It should be noted, however, that in two anonymous medieval Arabic-Berber word lists, Tsh. <*taskrā*> is defined as Ar. شِبِتّ *šibitt* 'dill' (van den Boogert 1998: 178, 184; cf. Múrcia Sànchez 2011: 432 fn. 2985). This identification is deemed "incorrect" by Van den Boogert (1998: 31, 178, 184). The fact that Pun. σικκιρία and Tsh. <*taskrā*> are formally compatible (Múrcia Sànchez 2011: 433), and are attested in the meaning 'dill', are evidence in favor of an etymological relationship between these forms. How exactly the meaning 'dill/false fennel' would have developed into that of 'thistle', or vice versa, is not clear to me.

3.1.8 Fleabane, yellow (*Dittrichia viscosa*)

Reconstruction	**Forms**
****frís(s)a***	DES (I: 547): ***fríša*** (Camp.) 'yellow fleabane (*Dittrichia viscosa* (L.) Greuter)';[59] ***frísa***, ***fríssa*** *'Artemisia* L.'; ***vríssa*** (1: Dorgali) 'great mullein (*Verbascum thapsus* L.)'[60]

This word occurs already in the Condaghe of San Pietro di Silki as [*erua*] *frisa* (Bonazzi 1900: sec. 309). Wagner (DES I: 547) remains agnostic about the etymology of these forms, but does consider a pre-Roman origin.

---

[57] Van den Boogert's *Echinops spinosus* L. is a synonym of *Echinops spinosissimus* Tura (WFO 2022a: s.v. *Echinops spinosissimus* Turra).

[58] If one reconstructs not two but just one short high vowel for Proto-Berber, this reconstruction would be **tā-səkrā* < **səkkəryā* (cf. Kossmann 2020a: 35–36).

[59] Wagner's *Inula viscosa* (L.) Aiton is a synonym of *Dittrichia viscosa* (cf. WFO 2022b: s.v. *Dittrichia viscosa* susbsp. *viscosa*).

[60] Wagner marks the meaning of the Dorgalese form with "(?)", presumably indicating that this meaning is not completely certain.

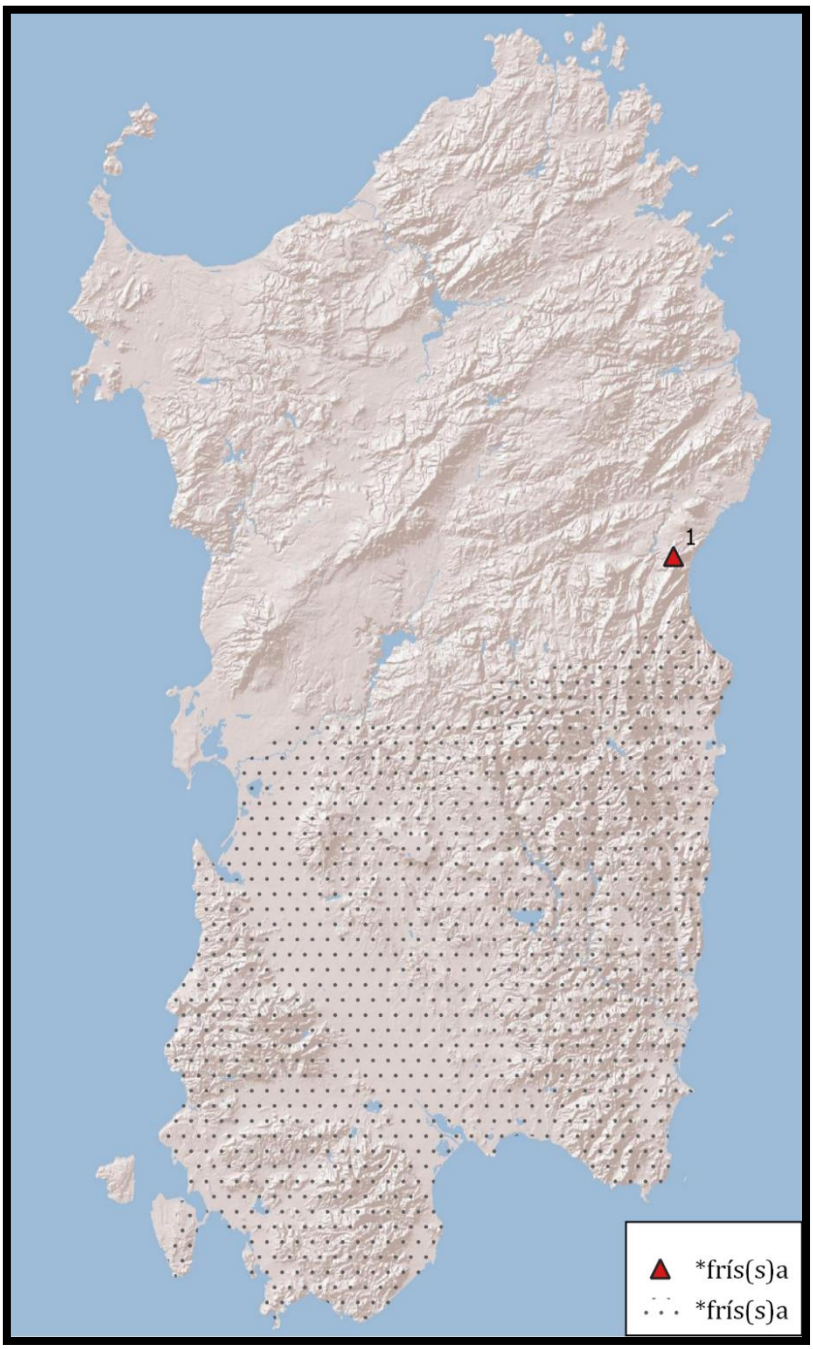

Figure 3.1.8: Yellow fleabane (*Dittrichia viscosa*)

Paulis (1992: 89) on the other hand derives the Sardinian forms from the Latin phytonym *afrissa* 'dragon lily (*Dracunculus vulgaris* Schott), attested once in Pseudo-Apuleius' Herbarium (cf. André 1985: 7).[61] Although dragon lily and yellow fleabane are not similar in appearance, Paulis attributes the connection to the fact that both plants are known for their foul smell, as well as to some shared medical uses. The meaning 'mullein (*Verbascum thapsus*)' may have come about due to the superficial similarity between mullein and yellow fleabane, which are both rather tall herbaceous plants with yellow flowers. The meaning '*Artemisia*', is found in Spano (1851: 221), but is not confirmed anywhere else (Paulis 1992: 89). Perhaps it is somehow the result of confusion of various scientific plant names containing *dracunculus* (cf. Paulis 1992: 89–90).

It is likely that the Sardinian forms are related to Lat. *afrissa* 'dragon lily'. The question regarding its etymology is all but resolved however. It is not quite clear where Pseudo-Apuleius' herbarium was composed. If it is North-African, as has been proposed (Pradel-Baquerre 2013: 20–21), this might be yet another instance of a Sardinian-African Romance isogloss, possibly due to Punic influence (cf. § 10.1). However, without direct evidence for a Punic provenance of these words and/or for the geographic origin of Pseudo-Apuleius, no reliable conclusions can be drawn.

3.1.9 Germander, yellow (*Teucrium flavum*)

Reconstruction	**Forms**
***bunnánaru**	DES (I: 240): ***bunnánneru*** (1: Bonorva) 'yellow germander (*Teucrium flavum* L.)'
	Paulis (1992: 121): **[*èrβa*] *bunnánaru*** (1: Bonorva) 'id.'

DES does not propose an etymology for this word, but merely notes its similarity to the toponym *Bunnánnaru* (standard Italian *Bonnanaro*). Paulis (1992: 122) suggests that this similarity may be due to chance or folk etymology, which is plausible. He proposes an etymology **kunn-án-aru*, from Lat. *cunnus* 'vagina' because of this plant's apparent emmenagogue (i.e. menstruation-stimulating) properties. This would require a derivation by means of the inherited substantivizing suffix *-ánu/a* and the "phytonymic suffix" *-́aru/a* (Paulis 1992: 122). Even though the semantic development needed for this etymology is acceptable, the evidence for nominalizing derivations with *-ánu/a* is scarce (Wagner 1952: 51–52), and the combination of *-ánu/a* and *-́aru/a* is found nowhere else.

[61] Paulis' (1992: 89) *Arum dracunculus* L. is a synonym of *Dracunculus vulgaris* Schott (cf. WFO 2024b: s.v. *Dracunculus vulgaris*).

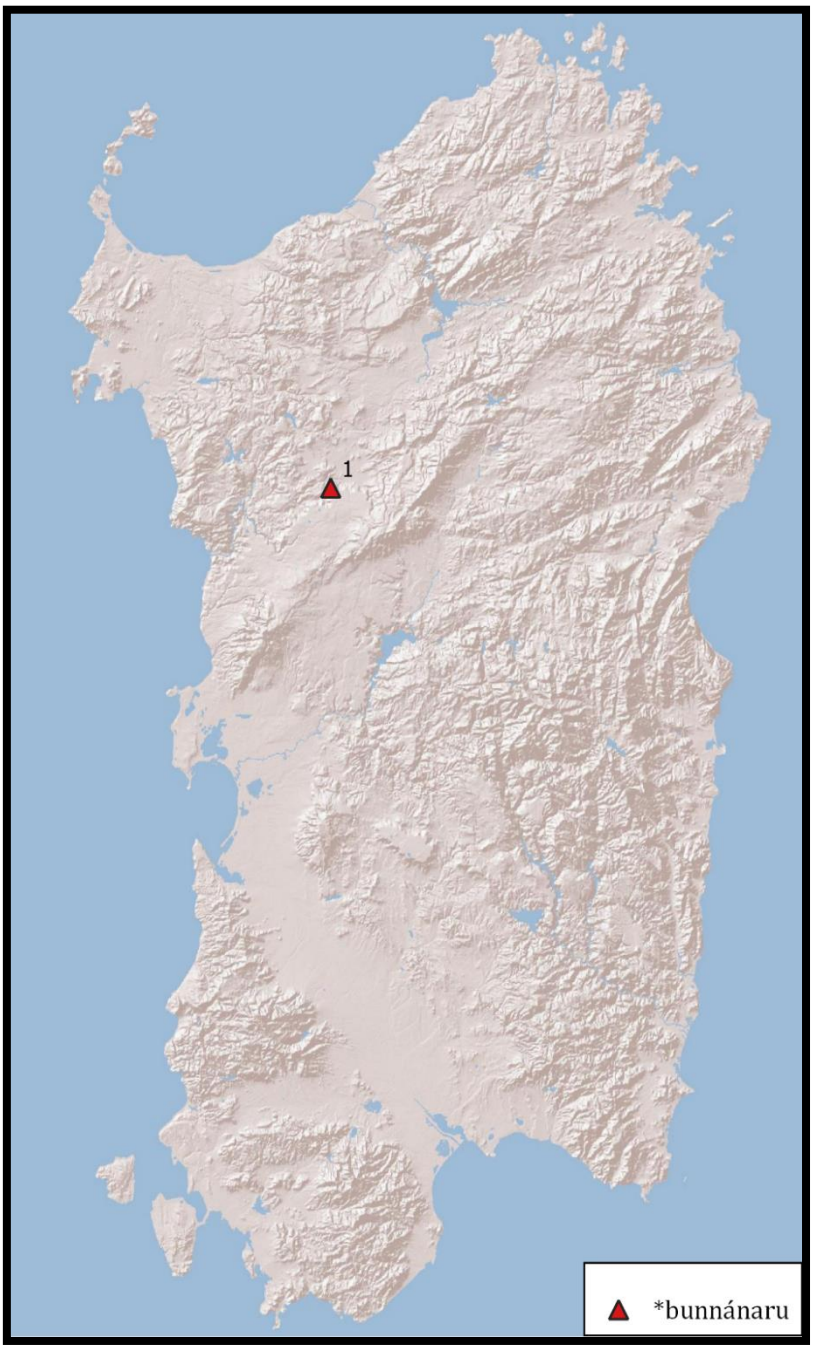

Figure 3.1.9: Yellow germander (*Teucrium flavum*)

Moreover, whereas we would need to assume reanalysis of the initial **k-* of **kunn-án-aru* through **g-* (regular in post-vocalic position) to attested *b-*, this is not paralleled in the regular reflex of Lat. *cunnus* itself, which is *kúnnu* everywhere (DES I: 433). Each of these developments is theoretically possible, but the combined number of assumptions needed to explain *bunnánneru* 'yellow germander' from *kúnnu* 'vagina' makes this etymology all but certain. The fact that this word has no apparent cognates outside of Bonorva, makes it difficult to comparatively evaluate Paulis' proposed etymology, let alone to propose an alternative one.

3.1.10 Hemlock (*Conium maculatum*)$_1$

Reconstruction	**Forms**
***aCússa**	DES (I: 153): ***aússa*** (1: Villaputzu) 'poison hemlock (*Conium maculatum* L.)'

This word is only recorded once, and the informant allegedly had to think before giving this name for 'poison hemlock' (DES I: 153). Wagner considers it to be "probably pre-Roman", but does not discuss it any further. The sequence *-aú-* must originally have contained a consonant (i.e. **-aCú-*), as inherited *-au-* regularly yields *-a-*. It is formally similar to Camp. *alaússa*, which refers to various kinds of mustard and other similar plants (§ 3.1.16). Since Villaputzu is in the Sarrabus region, where intervocalic *-l-* > *-ʔ-* (Wagner 1941a: 123), *aússa* could perhaps have developed from **aʔaússa*. From a semantic point of view this shift from 'mustard etc.' to 'poison hemlock' is more difficult to understand, although both plants have a superficially similar shape. Pittau (1995: 209) connects *aússa* to *kugúsa*, *ɔurgúsa* etc. 'fool's watercress; celery' < pre-Rom. **-gusa* (§ 3.1.6). This is formally and semantically imaginable, although the geminate *-ss-* in *aússa* is problematic. On the basis of a single attestation it is difficult to draw any firm conclusions.

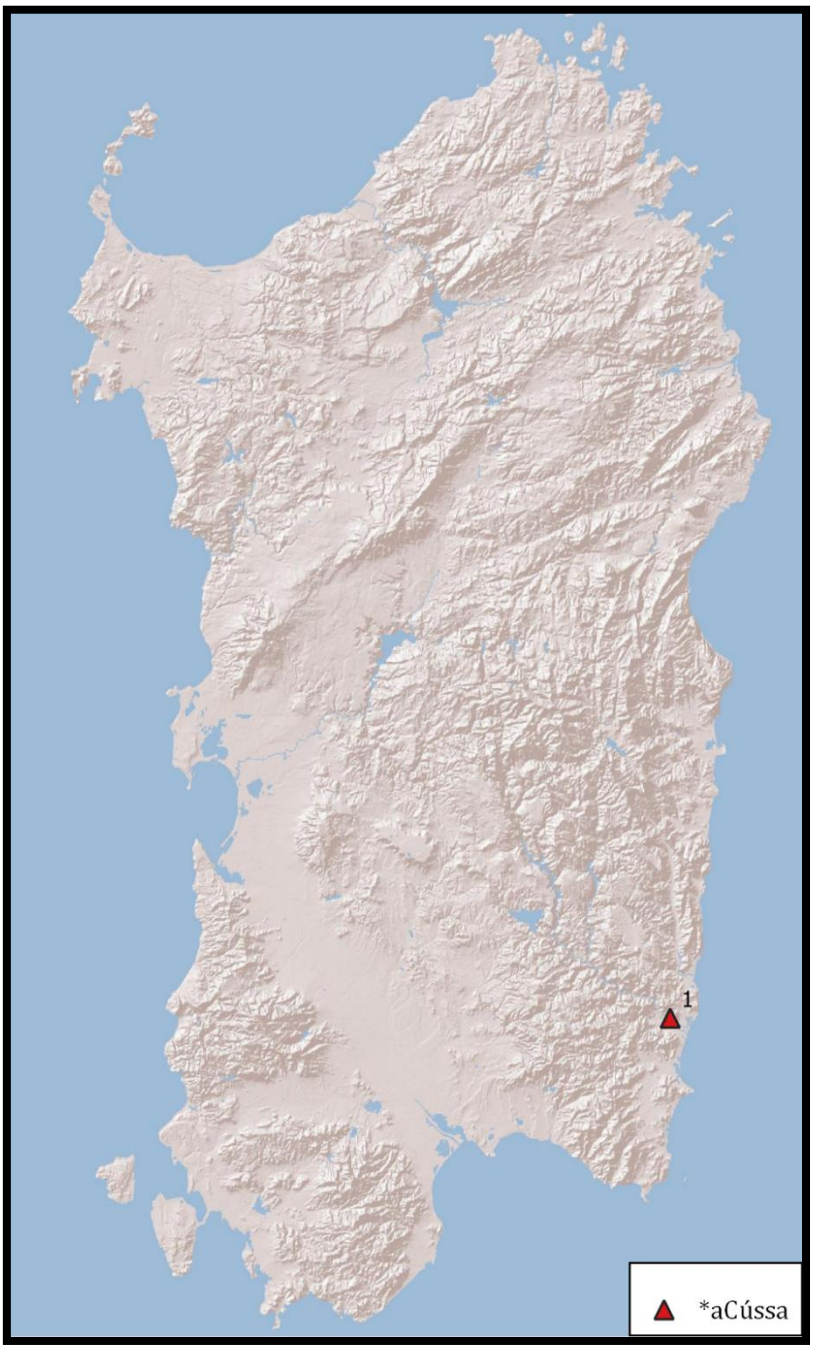

Figure 3.1.10: Hemlock (*Conium maculatum*)

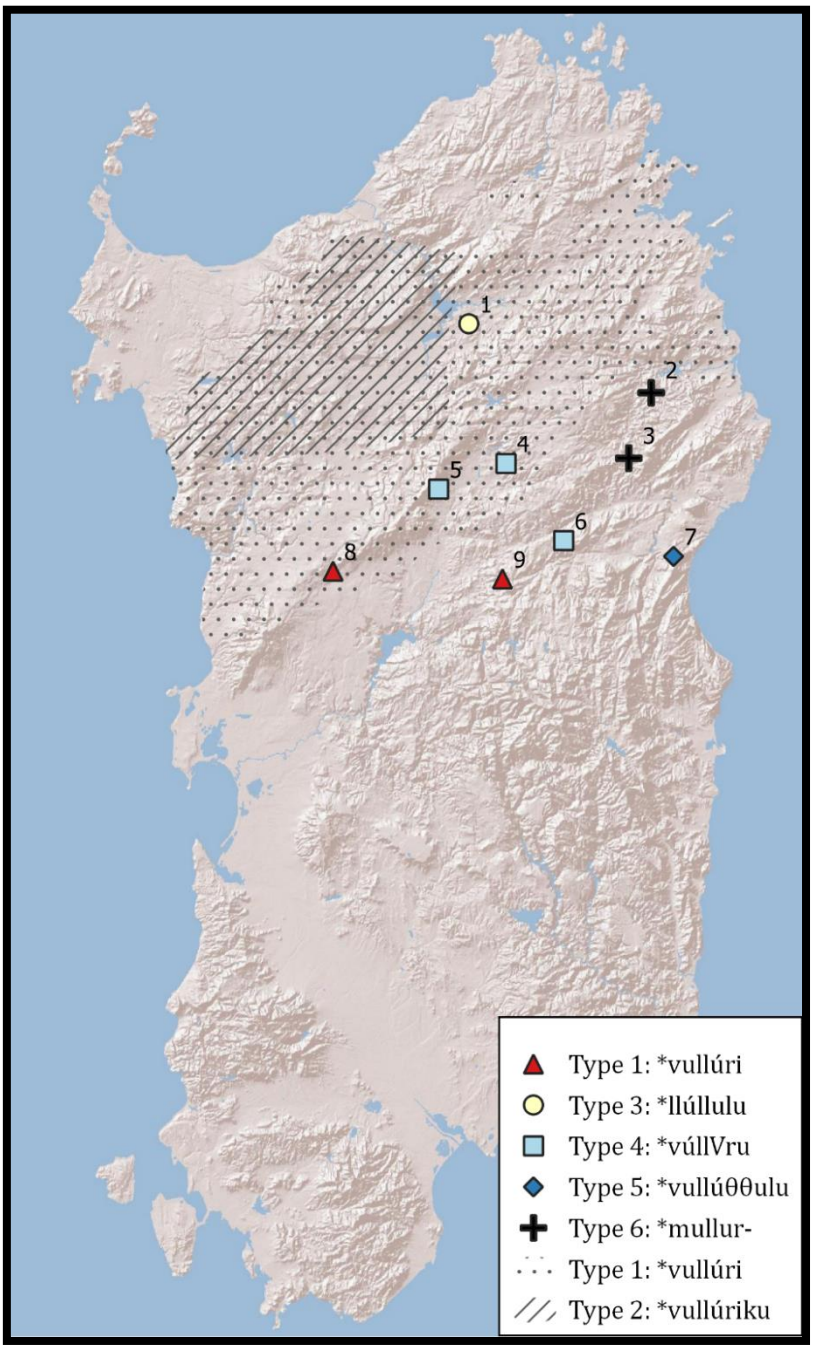

Figure 3.1.11: Hemlock (*Conium maculatum*)

3.1.11 Hemlock (*Conium maculatum*)$_2$

Reconstruction	Forms
1: *vullúri	DES (I: 204): ***biḍḍúri*** (Log.), ***buḍḍúri,*** [***su***] ***uḍḍúri*** (8: Macomer) 'poison hemlock (*Conium maculatum* L.)'
	Paulis (1992: 151): ***biḍḍúri, buḍḍúri*** (Log.), ***vuḍḍúri*** (9: Orani) 'id.'
2: *vullúriku	DES (I: 204): ***uḍḍúriɣu*** (NLog.) 'id.'
3: *llúllulu	DES (I: 204): [***su***] ***ḍḍúḍḍulu*** (1: Oschiri) 'id.'
4: *víllVru	DES (I: 204): ***uḍḍuréḍḍu*** (4: Nule), ***úḍḍuru*** (5: Bono) 'id.'
	Paulis (1992: 151): ***búḍḍaru*** (6: Nuoro) 'id.'
5: *vullúʒʒulu	Paulis (1992: 151): ***guḍḍútulu*** (7: Dorgali) 'id.'
6: *mullur-	Paulis (1992: 151): ***meḍḍure*** (2: Lodè), ***muḍḍuru*** (3: Lula) 'id.'

There is quite some variation in these words for 'poison hemlock', which lack a clear Latin etymology. Most widespread is type 1, from which type 2 is derived by means of the suffix **-ˊ(i)ku* (cf. Wagner 1952: 16). Types 3 and 4 are accented on the first syllable and have different suffixes than *-úri* found in type 1. Type 3 ([*su*] *ḍḍúḍḍulu*; 3: Oschiri) contains the inherited diminutive suffix *-ˊulu* (Wagner 1952: 35). The forms of type 4 contain *-ˊaru* and *-ˊuru*, which are not inherited from Latin. Unaccented *-ˊaru/a* occurs with some frequency in words of suspected pre-Roman origin (§ 9.2.6.2); *-ˊuru* may be due to assimilation of *-ˊaru* to preceding accented *-ú-*. Type 5 (*guḍḍútulu*; 7: Dorgali) has *-út(t)ulu* < Lat. *-úceu* + *-ˊulu* (Wagner 1952: 35). In the forms of type 6, the place of the accent is not indicated, which makes it impossible to establish which suffix we are dealing with. As for the initial consonant, we find *b-* in types 1 and 4. The forms without an initial consonant (e.g. *uḍḍuréḍḍu*; 4: Nule) as well as with initial *g-* (in *guḍḍútulu*; 7: Dorgali) can be explained from original **b-* by reanalysis of the postvocalic form, in which the onset was regularly lost, as the base form (cf. Wagner 1941a: 206–209). The onset *ḍḍ-* in type 3 is likely due to assimilation to the following *-ḍḍ-*. The *m-* of the forms in type 6 is slightly more difficult to explain morpho-phonologically, but cf. Wagner (1941a: 102). All in all, most forms are most easily explained from **buḍḍ-* < **bull-/vull-*. The variant *vuḍḍúri* (9: Orani) provides evidence that *v-* may be original. Although all Sardinian dialects, except for the dialect of Bitti, generally merge inherited *-v-* and *-b-* in all positions (Wagner 1941a: 97–98), there are some traces of preservation of original *-v-* as distinct from *-b-*. For Orani, cf. *vervèke* 'sheep' vs. Log. *berβèɣe* 'id.' < Lat. *vervēx* 'wether' (Wagner 1941a: 101; DES II: 573); *vinnínna* 'grape harvest' vs. Log., Camp. *binnènna* 'id.' < Lat. *vindēmia* 'id.' (DES II: 578). Wagner (DES I: 204) argues that Log. *biḍḍúri* is due to vowel dissimilation from older *buḍḍúri*, whose vocalism is more widely

attested across the dialects. Against this, Paulis (1992: 152) proposes that the vocalism of *buḍḍúri* etc. may have arisen through association with *búḍḍa* 'gut, rectal intestine', because of the plant's unpleasant smell and its frequent presence in dunghills.[62] Positing original **billúri* would allow for a comparison with Tusc. *bìlleri* (pl.) 'mayflowers (*Cardamine pratensis* L.)', as proposed by Pittau (*apud* Paulis 1992: 151; Pittau 1995: 203; cf. also Alessio 1949: 132). However, poison hemlock and mayflowers are rather different plants, especially given the fact that the latter is edible while the former is extremely poisonous. Paulis' proposal that *buḍḍúri* etc. was influenced by *búḍḍa* 'gut, rectal intestine', does not quite explain forms such as *ḍḍúḍḍulu*, *guḍḍútulu* and *muḍḍuru*. The fact that these all have *-u*-vocalism, suggests that it may be original after all.

In conclusion, the two types that cannot be explained in terms of inherited material, are type 1 *buḍḍúri* etc., and type 4 *búḍḍaru* etc. The question is whether the difference in suffixation results from the existence of two differently suffixed pre-Roman forms that were borrowed into Sardinian independently, or from secondary formations within Sardinian. In light of the occurrence of both *-úri* and *-́ara/u* in other Sardinian words of pre-Roman origin (§ 9.2.6), we are likely dealing with a pre-Roman base **vull-* referring to 'poison hemlock'.

3.1.12 Mullein, great (*Verbascum thapsus*)$_1$[63]

Reconstruction	**Forms**
1: *túmbar-	DES (I: 261-262): ***túmbaru*** (1: Monti) 'sow thistle (*Sonchus oleraceus* L.)', ***túmbara*** (4: Urzulei) 'great mullein (*Verbascum thapsus* L.)'
2: *tutúmbar-	Paulis (1992: 61): ***tutúmbaru*** (2: Lodè, 3: Siniscola) 'spurge (*Euphorbia sp.* L.)'
3: *katúmb(ul)u	DES (I: 261-262): ***kaðúmbu*** (5: Milis, 9: Oristano, 13: Usellus, 14: Mogoro), ***kardúmmulu*** (6: Tonara), ***kaðúmbulu*** (7: Meana Sardo, 9: Laconi), ***karúmbulu*** (8: Tortolì), ***karćúmbulu*** (11: Gairo), ***kardúmbulu*** (12: Seui), ***kaðrúmbulu*** (15: Escalaplano), ***kaðúmbuʔu*** (16: Siurgus), ***kadúmbuʁu*** (17: San Nicolò Gerrei, 18: Dolianova) 'great mullein (*Verbascum thapsus*)'

[62] I am however not convinced by Paulis' (1992: 152) remark that poison hemlock's hollow stems would necessarily remind speakers of rectal intestines.

[63] These data are discussed in a similar fashion in Swanenvleugel (2024: 236–239).

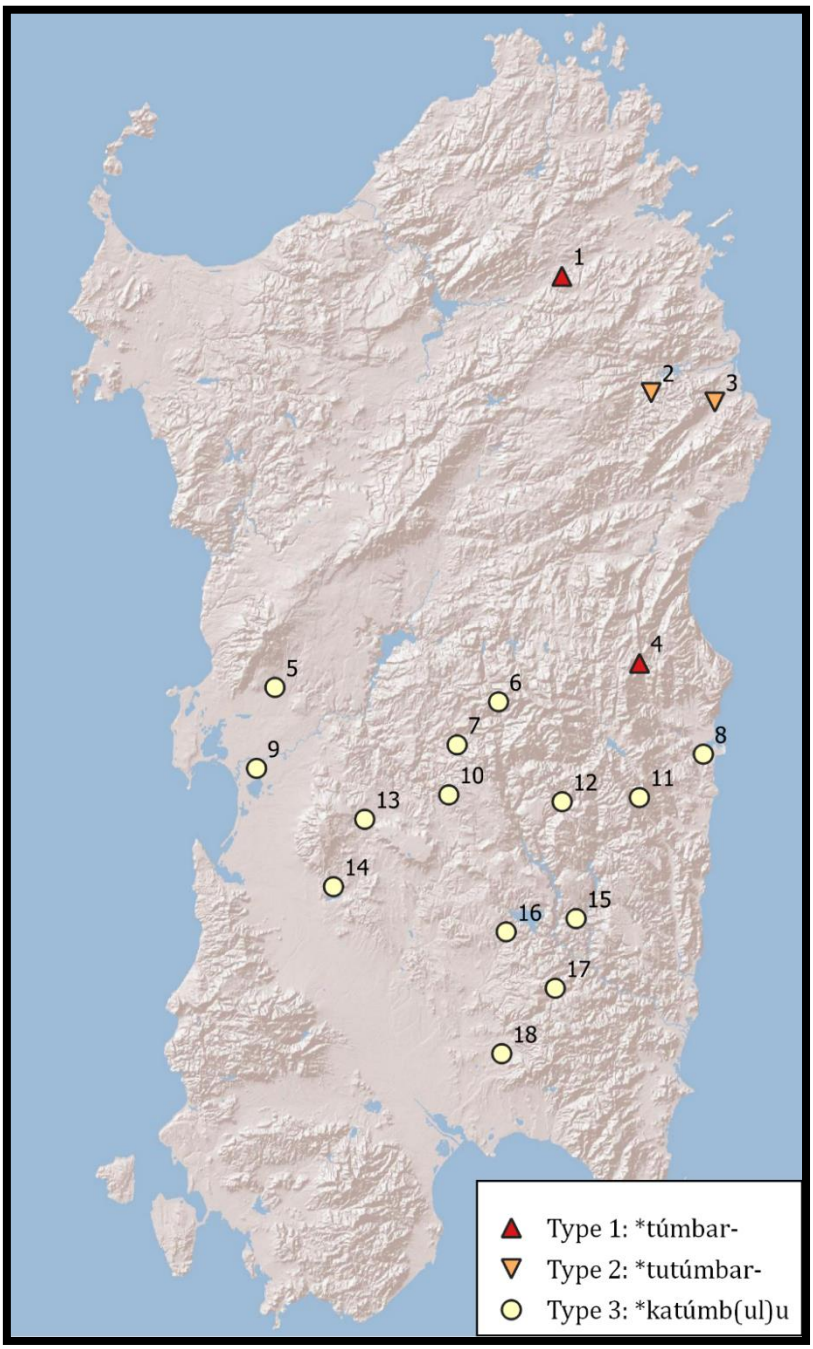

Figure 3.1.12: Great mullein (*Verbascum thapsus*)

In the meaning 'great mullein', we find forms of type 1 next to type 3. The latter of these invariably refer to 'great mullein'. Type 1 *túmbara* (2) 'great mullein' however has a near-homonym *túmbaru* (1) that denotes 'sow thistle'. Mullein and sow thistle are two rather different, unrelated plants. Both do have yellow flowers however, and both are used in traditional medicine. The question is whether these superficial similarities are enough to motivate a semantic shift from 'mullein' to 'sow thistle'.

Paulis (1992: 61, 354–355) argues that *tutúmbaru* 'spurge (*Euphorbia sp.*)' (type 2) was inherited from Lat. *tithymalus* (which in turn is from Gr. τιθύμαλ(λ)ος 'milkweed (*Euphorbia peplus* L.)'; Beekes 2010: 1483–1484). This *tutúmbaru* would have been extended semantically, on the one hand, to 'sow thistle (*Sonchus oleraceus*)', due to both plants' "milk-like" sap,[64] and on the other hand to 'mullein (*Verbascum thapsus*)', because of both plants' use as a piscicide (cf. Heizer 1953: 263). There are several issues with this scenario. Firstly, Lat. *tithymalus* 'spurge' is continued in Sardinian as Log. *titímbalu*, Nuor. *ϑiϑímbalu* (Paulis 1992: 170). This does not correspond regularly to *tutúmbaru* found in Lodè (2) and Siniscola (3), where we would also expect ***ϑiϑímbalu*. The same goes for *túmbaru* 'sow thistle', which Paulis (1992: 355) explains as the result of influence from Log. *túmbaru* 'empthy beehive', on account of the hollow stem of the sow thistle, citing Wagner's (DES II: 532) hypothesis that the word for 'beehive' is a loan from It. *túbolo* 'tube'. Regardless of the likelihood of the phonetic developments needed for *túbolo* 'tube' >> *túmbaru* 'empty beehive', it is important to note that *túmbaru* itself does not mean 'tube'. Moreover, spurges (*Euphorbia sp.*) do not have hollow stems. It is thus unlikely that this word for 'empty beehive' has had anything to do with the botanical terms discussed here, and the forms *tutúmbaru* 'spurge' and *túmbaru* 'sow thistle' remain unresolved. Paulis is probably right in attributing the similarity between *tutúmbaru* 'spurge', *túmbaru* 'sow thistle' and *túmbara* 'mullein' to popular association of these plants because of similar uses or botanical traits, but it seems more likely that it was *túmbara* 'mullein' that influenced the outcome of Lat. *tithymalus* 'spurge' in Lodè (2) and Siniscola (3).

For the forms of type 3, Paulis (1992: 355) considers two different etymologies. On the one hand he compares it to the Hesychian gloss καράμβας 'shepherd's staff' (Hesychius Alexandrinus 2020: 523), and to Skt. *kaḍamba-* 'end, point; stalk' (Monier-Williams 1872: 245a). On the other hand, he proposes *kaðúmbu* to be a contraction of **kánn'e ðúmbu* 'reed of the "*túmbu*" (one of the three pipes of the traditional Sardinian musical instrument *launèḍḍas*). For the Sanskrit (and Hes-

[64] Cf. the Dutch names *wolfsmelk* 'spurge' and *melkdistel* 'sow thistle'.

ychian) comparison, the geographical as well as the semantic distance are simply too great for it to be of any use. The second comparison is unlikely because *launèḍḍas* are made of reed, which is not at all similar to mullein. The best option is still to treat this word as being of possible pre-Roman origin, and to analyze it within Sardinian before jumping to far-flung comparisons.

Within type 3, we find a large variety of forms. Most of these can be explained by secondary insertion of *-r-*. In the case of this specific word, this may have been due to influence from *kardu* 'thistle', but non-etymological *-r-*epenthesis is a common phenomenon in Sardinian in general (Wagner 1941a: 225–227). As such, *kaðúmbu* and its diminutive *kaðúmbulu* are the base form of this type.[65] The intervocalic *-ð-* in this word must go back to an original *-t-* (Wagner 1941a: 69), leading to a reconstruction **katúmbu*. This leaves us with **katúmbu* next to *túmbaru/a*. An apparent suffix *-ára/u* occurs with some frequency in the non-inherited Sardinian lexicon (§ 9.2.6.2). This allows us to identify a stem *-tumb-* in both **ka-túmb-* and *túmb-ara*. Prefixed *ka-* occurs in several words of suspected pre-Roman origin (cf. § 9.1.1).

3.1.13 Mullein, great (*Verbascum thapsus*)$_2$

Reconstruction	**Forms**
***trofòlla**	DES (II: 525): ***truèḍḍa*** (Gallur.), ***travòḍḍa*** (1: Sennori), ***trovòḍḍa*** (2: Bonorva, 3: Padria, 4: Nuoro, 8: Ghilarza), ***trivòḍḍa*** (5: Macomer, 6: Scano di Montiferro, 7: Santu Lussurgiu, 10: Busachi, 11: Baunei), ***tròḍḍa*** (9: Fonni) 'great mullein (*Verbascum thapsus* L.)'

Besides the forms listed above, Wagner (DES II: 525) mentions a non-localized variant *truvèḍḍa*, which he believes must be somewhere from the north because of its similarity to Gallurese *truèḍḍa*. Wagner (DES II: 525) and Paulis (1992: 355–356) both follow Schuchardt (1899: 183), who derives the Sardinian forms from Lat. *turbāre* 'to disturb', because of the practice of stunning fish by putting mullein in the water (cf. Heizer 1953). Although this etymology would work well semantically, cf. Srd. *truvísku* 'flax-leaved daphne (*Daphne gnidium* L.)' < Lat. *turbiscus* (Schuchardt 1899: 182–183; DES II:530), it is phonologically and morphologically problematic.

[65] Some of the other differences are due to local dialectal developments, such as the development of **-l-* in *kaðúmbuɾu* (San Nicolò Gerrei, Dolianova), *kaðúmbuʔu* (Siurgus) < *kaðúmbulu*.

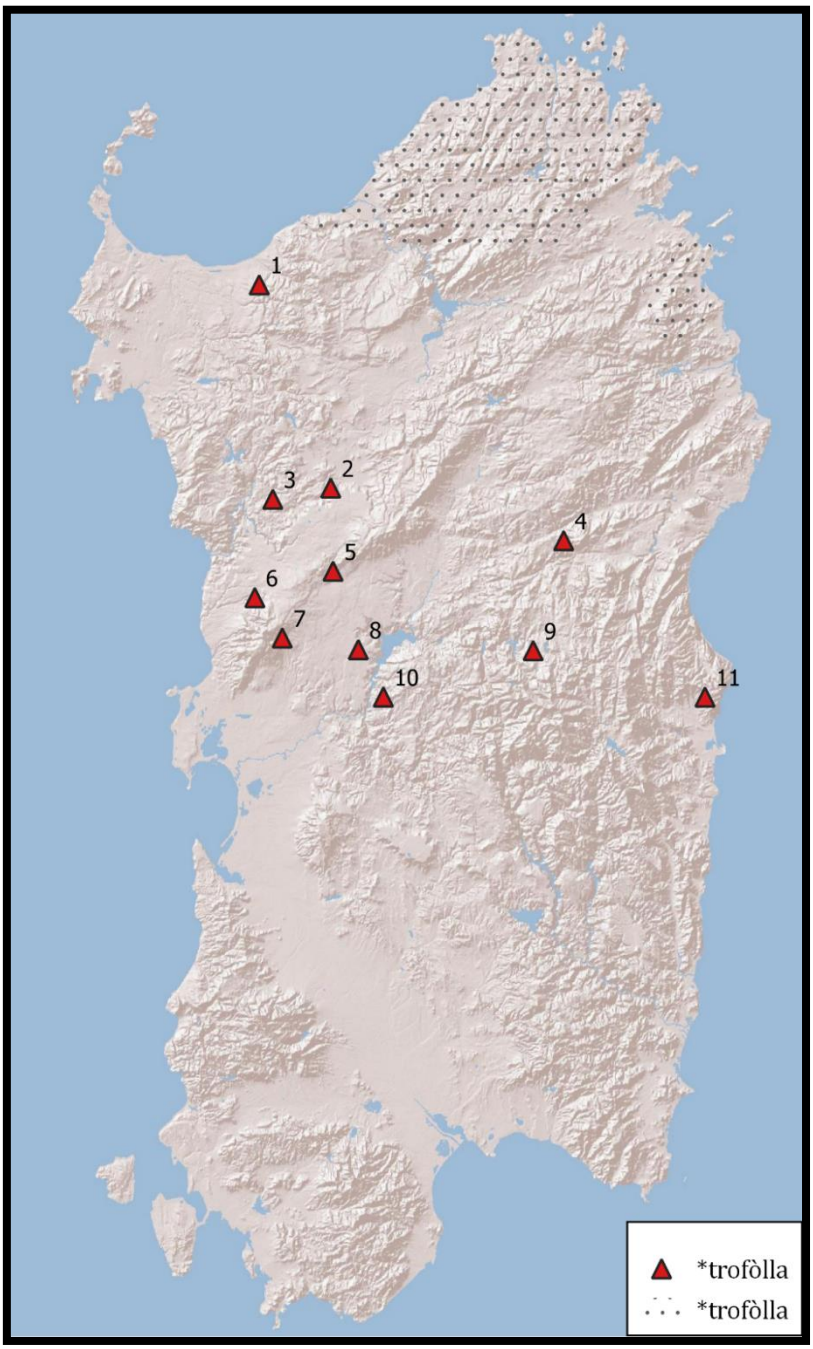

Figure 3.1.13: Great mullein (*Verbascum thapsus*)

Wagner (DES II: 525) notes that the ending *-òḍḍa*, which should go back to **-olla*, is peculiar and not found elsewhere in Sardinian. On the phonological side, there are two issues in deriving *trovòḍḍa* etc. from Lat. *turbare*. The first is that the expected vocalism would be ***truvòḍḍa*, which is attested only in the Sardo-Corsican dialect of the Gallura. Although sporadic pretonic vowel changes are common (Wagner 1941a: 24–33), we would still expect pretonic *-u-* to be preserved somewhere; cf. also Log. *truβare* 'to flush, to incite (animals)' < Lat. *turbare*, Log. *truβísku* 'flax-leaved daphne' < Lat. *turbiscus* (DES II: 529–530), which preserve pretonic *-u-*. The actually attested vocalism for this word rather points to original *-o-*. The second issue concerns its consonantism. If this word were from *turb-*, we would expect *-b-* to yield *-β-* in most dialects (Wagner 1941a: 98–100, 176). Even if we take into account sporadic outcomes of **-b-* as *-v-*, such as in *truvísku* 'flax-leaved daphne' (Nuoro) < Lat. *turbiscus* (DES II: 530), this does not explain the complete absence of this consonant in *tròḍḍa* (9: Fonni). For these reasons, it is more consistent with the data to assume an original **-f-* in this word. In the Barbaricine dialects, to which Fonni belongs, *-f-* is lost, while it becomes *-v-* intervocalically elsewhere (Wagner 1941a: 90–91).

At face value, these Sardinian words for 'great mullein' thus point to original **trofòlla*. The morphological and phonological difficulties described above cause Schuchardt's (1899: 183) etymology to lose much of its explanatory power, and it is best abandoned, although there may still have been influence from *truváre* < Lat. *turbare* in the forms *truvèḍḍa* and Gallur. *truèḍḍa*. It is in any case clear that in these forms the ending *-òḍḍa* was remade into *-èḍḍa*, which is the productive diminutive suffix < Lat. *-ella/us*. The reconstruction of *trovòḍḍa* etc. 'great mullein' as **trofòlla* and the rejection of its etymology mean that it is of obscure origin, just like several other words for 'mullein'. It may be of pre-Roman origin, although there is no direct evidence beside its lack of an etymology.

3.1.14 Mullein (*Verbascum sp.*)

Reconstruction	**Forms**
***lókri**	DES (II: 34): ***lòkri*** (1: Busachi) 'species of mullein (*Verbascum sp.* L.) used to make mats'
	Puddu (2023: s.v. *lócri*): ***lócri*** 'species of mullein (*Verbascum sp.*)'

This word from Busachi is attested with an open *-ò-* in Wagner, and with a closed *-ó-* in Puddu. One wonders whether the open *-ò-* in Wagner could be due to an error. We would expect **lokri* to yield *lókri* with closed *-ó-*, due to the following *-i*.

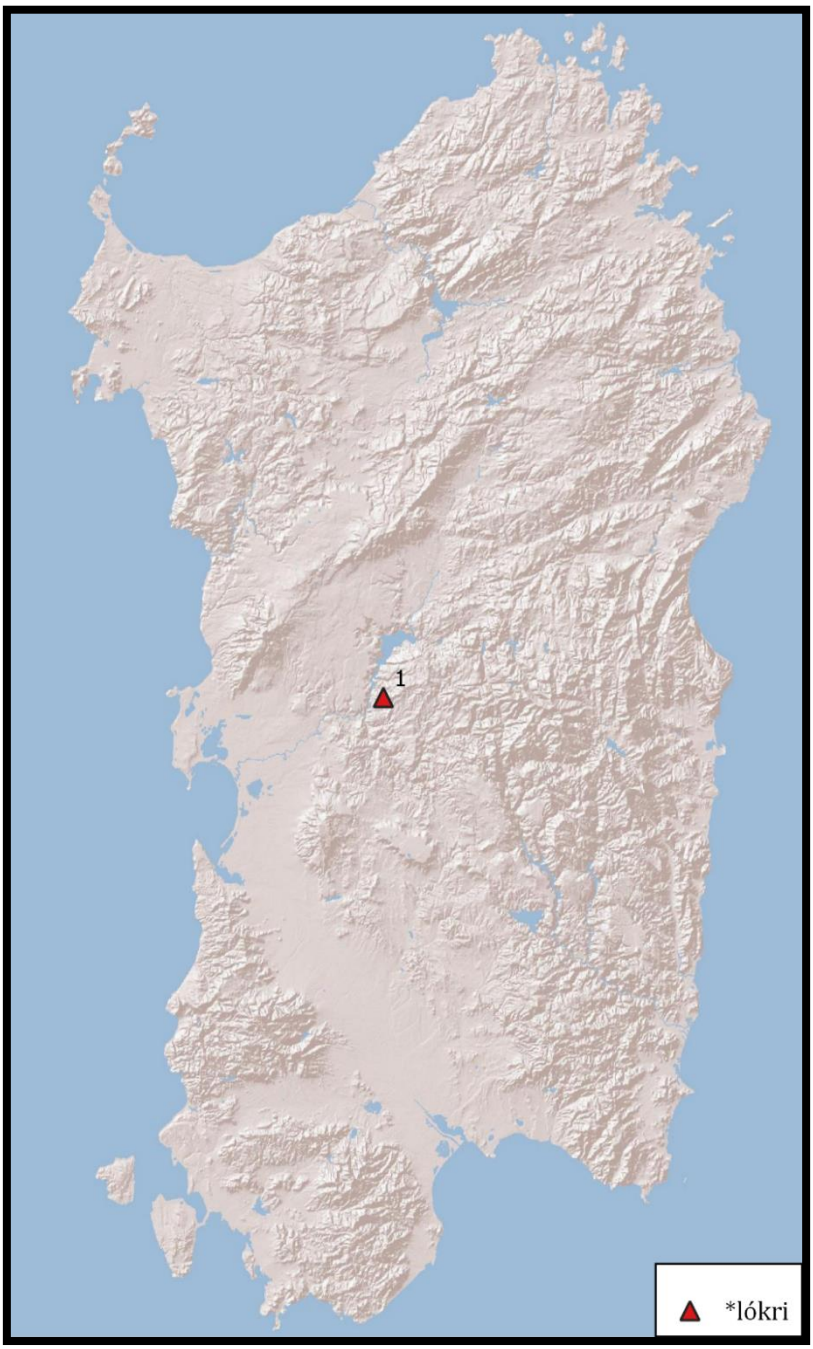

Figure 3.1.14: Mullein (*Verbascum sp.*)

Note furthermore that Busachi does not belong to the dialects that raise the unaccented mid vowels *e*, *o* to *i*, *u* (cf. Wagner 1907b: Map 1), so an original **lòkre* is impossible too. Wagner (DES II: 34) does not express any thoughts on the etymology of this isolated form. Paulis (1992: 356) proposes a back-formation from *lukréžu* '*Stachys glutinosa*'. This etymology is based on his assumption that *lukréžu*, which is attested only in Aritzo, goes back to Lat. *helicrysum* (Paulis 1992: 115)(cf. § 3.1.3). It is not clear to me what the motivation for a back-formation of *lukréžu* '*Stachys glutinosa*' >> *lò/ókri* 'mullein' could have been, and by which mechanism *-ézu* was detached. In my view Srd. *lòkri*, *lócri* remains obscure; it could be of pre-Roman origin, but this is difficult to establish on the basis of a single form.

3.1.15 Poppy (*Papaver rhoeas*)

Reconstruction	**Forms**
1: *ánda	Wagner (1928: Map 9): **ánna** [**e lóɣu**] (1: Torpè), **ánna** [**e lóku**] (2: Siniscola, 3: Bitti) 'common poppy (*Papaver rhoeas* L.)'
2: *(a)ɵánda	Wagner (1928: Map 9):[66] **aɵáṇḍa** (Nuor., 4: Orune, 5: Lollove, 6: Nuoro, 7: Dorgali, 8: Oliena, 10: Orgosolo, 14: Urzulei), **ɵáṇḍa** (12: Ollolai, 13: Gavoi) 'id.'
	DES (I: 151): **aɵáṇḍa** (13: Gavoi) 'id.'; [**fròre e**] **tɵáṇḍa** (14: Triei) 'id.'
3: *ɵránda	Wagner (1928: Map 9): **ɵráṇḍa** (9: Orani) 'id.'
4: *ɵánɵa	Wagner (1928: Map 9): **tsántsa** (11: Olzai) 'id.'
5: *ɵándara	Puddu (2023: s.v. *tàndha*, *tzàndha*): **tzàndhara** 'id.'

This word, which is restricted to the Nuorese and some Ogliastran dialects, is generally agreed to be of pre-Roman origin (Wagner 1922: 251; Bertoldi 1937a: 151 fn. 1; DES I: 151; Paulis 1992: 319). It shows quite some formal variation. Most common is type 2 (**ɵánda*), from which type 3 may have arisen by *-r-*insertion. In type 4 the sequence *-nd-* was replaced with *-nts-* due to assimilation to initial *ts-* < **ɵ-*. The forms of type 1 (*ánna*) lack the onset **ɵ-* altogether.[67] Puddu (2023: s.v. *tzàndha*) also notes *tzándhara* (without location; type 5), ostensibly with the suffix *-ára* (§ 9.2.6.2).

[66] These forms are already partially listed in Wagner (1922: 251), but all instances of *-ɵ-* have mistakenly been printed as *-p-*.

[67] The realization of *-nd-* as *-nn-* is not uncommon in the Nuorese dialects (Wagner 1952: 179). According to Wagner (DES I: 151) this pronunciation was in this case reinforced by association with the female personal name *Anna*.

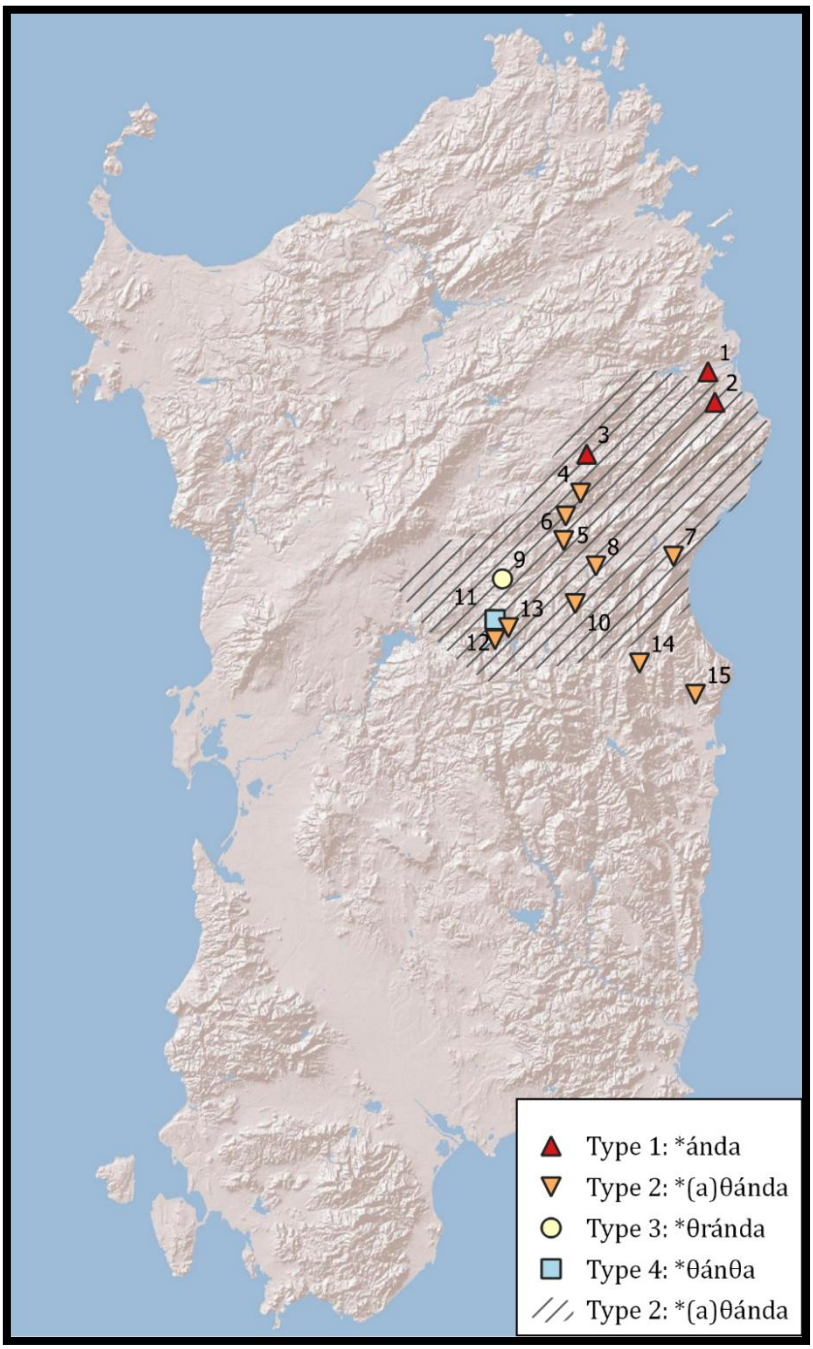

Figure 3.1.15: Poppy (*Papaver rhoeas*)

Outside Sardinia, this word has been compared to Egyptian ναντί 'poppy' (Bertoldi 1937a: 151 fn. 1; cf. Dioscorides Anazarbei 1906: 218), Lang. *ãndèr*, *õndèr* (dep. Lot) etc. 'id.' (cf. FEW XXIV: 550 fn. 21), Bq. *andere* 'lady' (Bertoldi apud Wagner 1932: 237; cf. FEW I: 93), and Gr. ἄνθος 'flower' (Pittau apud Paulis 1992: 319; Pittau 1995: 108). All of these comparisons require the assumption that the Sardinian forms of types 2 – 5 contain the pre-Roman prefix **ϑ-* (§ 9.1.2), which would in that case not be part of the lexical stem (Wagner 1932: 237; Pittau 1995: 108). This is confirmed by the existence of type 1 **ánda*. However, Gr. ἄνθος 'flower' is inherited from an Indo-European root **h₂end^{h}-* (Beekes 2010: 104–105). As Wagner (DES I: 151) notes, the Egyptian comparandum is rather difficult to unite to the Sardinian forms phonetically, especially in absence of more precise information on the provenance of this form. The comparison of Bq. *ander* 'woman' seems to be based more on superficial similarity than anything else; it is to be rejected on semantic grounds.

The Gallo-Romance forms for 'poppy' are formally similar to local words for 'ulcer', e.g. MFr. *enderce*, Lang. *endèrs* (dep. Aveyron) etc. < **derbice*. This word would be related to Gaul. *derbita* 'ulcer', which is attested in glosses and has a cognate in Bret. *dervoed* (FEW III: 46). FEW (XXIV: 550 fn. 21) therefore regards the meaning 'poppy' as a semantic extension of 'ulcer'. However, although the forms for 'poppy' and 'ulcer' in these dialects are similar, they are not homonymous everywhere: e.g. Lang. *ender* 'ulcer' vs. *onder* 'poppy' (Cahors; FEW XXIV: 550 fn. 21), *dervese* 'ulcer' vs. *ẽdẹ́r* 'poppy' (dep. Cantal; FEW III: 46, XXIV: 550 fn. 21). Moreover, a shift in meaning from 'ulcer' to 'poppy' is not easily understood. It is in my view preferable to keep the Gallo-Romance words for 'poppy' and 'ulcer' etymologically separated, although they may have exerted some influence on each other due to formal similarities. This leaves Lang. *ãndèr*, *õndèr* etc. 'poppy' as a possible comparandum to the Sardinian forms. It is in any case plausible that this set of Sardinian words for 'poppy' is of pre-Roman origin, because alternating **ϑ-* (and perhaps *-ára*).

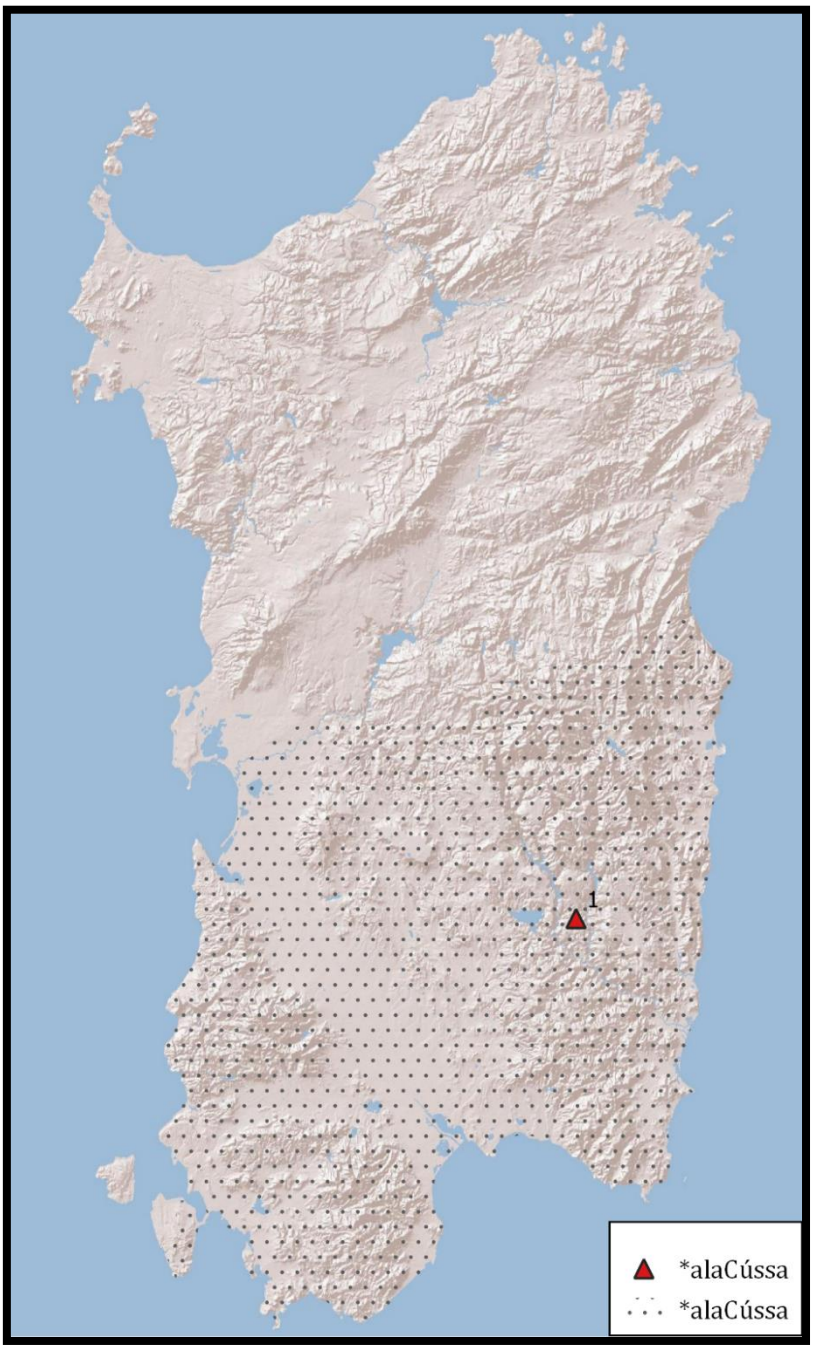

Figure 3.1.16: Mustard

3.1.16 Mustard$_1$

Reconstruction	Forms
***alaCússa**	DES (I: 68): ***alaússa*** (Camp.) 'white mustard (*Sinapis alba* L.), shortpod mustard (*Hirschfeldia incana* (L.) Lagr.-Foss.)', ***alaússa*** (1: Escalaplano) 'species of yellow ryegrass (*Lolium sp.* L.)'
	Paulis (1992: 236–237): ***alaússa*** (Camp.) 'white mustard (*Sinapis alba*), shortpod mustard (*Hirschfeldia incana*),[68] London rocket (*Sisymbrium irio* L.), hedge mustard (*Sisymbrium officinale* (L.) Scop.), corn rocket (*Bunias erucago* L.), common giant mustard (*Rapistrum rugosum* (L.) All.)'

All species referred to by *alaússa* bear similarity to each other by having small yellow flowers and edible seeds or leaves with a sharp taste. DES (I: 68) is uncertain as to its etymology. Paulis (1992: 237) derives *alaússa* from a hypothetical **saleús* < Lat. **sale iūs* 'salt-sauce', with detachment of initial *s-* by confusion with the definite article *sa*. The semantic motivation would be that mustard, with pine kernels, was commonly used to make a sauce for a specific turnip dish, for which he adduces a passage from Columella (12: 57; cf. Columella 1982: 430, 433). Even if such a specific use of the plant in question could have led to its name being derived from a name of a sauce — Srd. *fíkatu*, It. *fegato* etc. 'liver' < Lat. *fīcātus* 'fattened with figs' proves that culinary culture can bring about major semantic shifts — this etymology cannot be maintained on morphological grounds. The required form **sale-i̯us* would be a head-final compound, whereas virtually all other compounds in Sardinian are head-initial (Wagner 1952: 137–138), in line with Romance tendencies (Forza & Scalise 2016). It is also problematic that Lat. *iūs* 'broth, sauce' is not continued in Sardinian; the only direct reflexes of this word are found in Gallo-Romance (FEW V: 85). The existence of such a compound is therefore unlikely. Additionally, the development of word-final **-s* into geminate *-ss-* is unexpected. It thus remains etymologically obscure, and a pre-Roman origin cannot be excluded. Any pre-form would probably require the presence of a consonant between *-a-* and *-ú-* (i.e. **alaCússa*), as inherited *-au-* would have yielded ***-a-* (cf. Wagner 1941a: 13–14). The form *aússa* (Villaputzu) 'poison hemlock' could perhaps be the same word (cf. § 3.1.10).

[68] Wagner's and Paulis' *Sinapis incana* L., and Paulis' *Brassica adpressa* Boiss. are synonyms of *Hirschfeldia incana* (L.) Lagr.-Foss. (cf. WFO 2022c: s.v. *Hirschfeldia incana*).

3.1.17 Mustard$_2$

Reconstruction	**Forms**
i̯uspinu	DES (I: 609): ***ǧúspinu***, ***ǧuspínu*** (Log.) 'shortpod mustard (*Hirschfeldia incana* (L.) Lagr.-Foss.)'[69]
	Paulis (1992: 235): ***ǧúspinu, ǧuspínu*** (Log.) 'white mustard (*Sinapis alba* L.)'

DES (I: 609) does not offer an etymology for this word. Like for Camp. *alaússa* 'white mustard' in § 3.1.16, Paulis (1992: 236) proposes Log. *ǧúspinu*, *ǧuspínu* 'id.' to be derived from the name of a specific sauce, on the basis of a recipe found in Columella (12: 57; cf. Columella 1982: 430, 433). In this case, the mustard seeds would, together with pine nuts and vinegar, be ground into a kind of sauce. The forms *ǧúspinu*, *ǧuspínu* would thus go back to Lat. *iūs pīnūs* 'sauce of the pine'. While it is still uncertain whether Lat. *iūs* 'sauce, broth' was continued in Sardinian, this derivation is morphologically more plausible than the one proposed for *alaússa* (§ 3.1.16). As for the semantics, we would have to assume that mustard was so strongly associated with this sauce that it was referred to as "[herb of the] pine sauce"; cf. again Romance **ficatu* 'liver' < Lat. *ficatus* '[liver] fattened with figs'. If this etymology is correct, Srd. *ǧuspínu* probably continues the original Latin stress pattern, while proparoxytone *ǧúspinu* must be an innovation.

[69] Wagner gives the meaning: "senape bianca (*Sinapis incana*)". However, *Sinapis incana* L. is a synonym of *Hirschfeldia incana* (L.) Lagr.-Foss., which refers to 'shortpod mustard' rather than 'white mustard'. Cf. fn. 68.

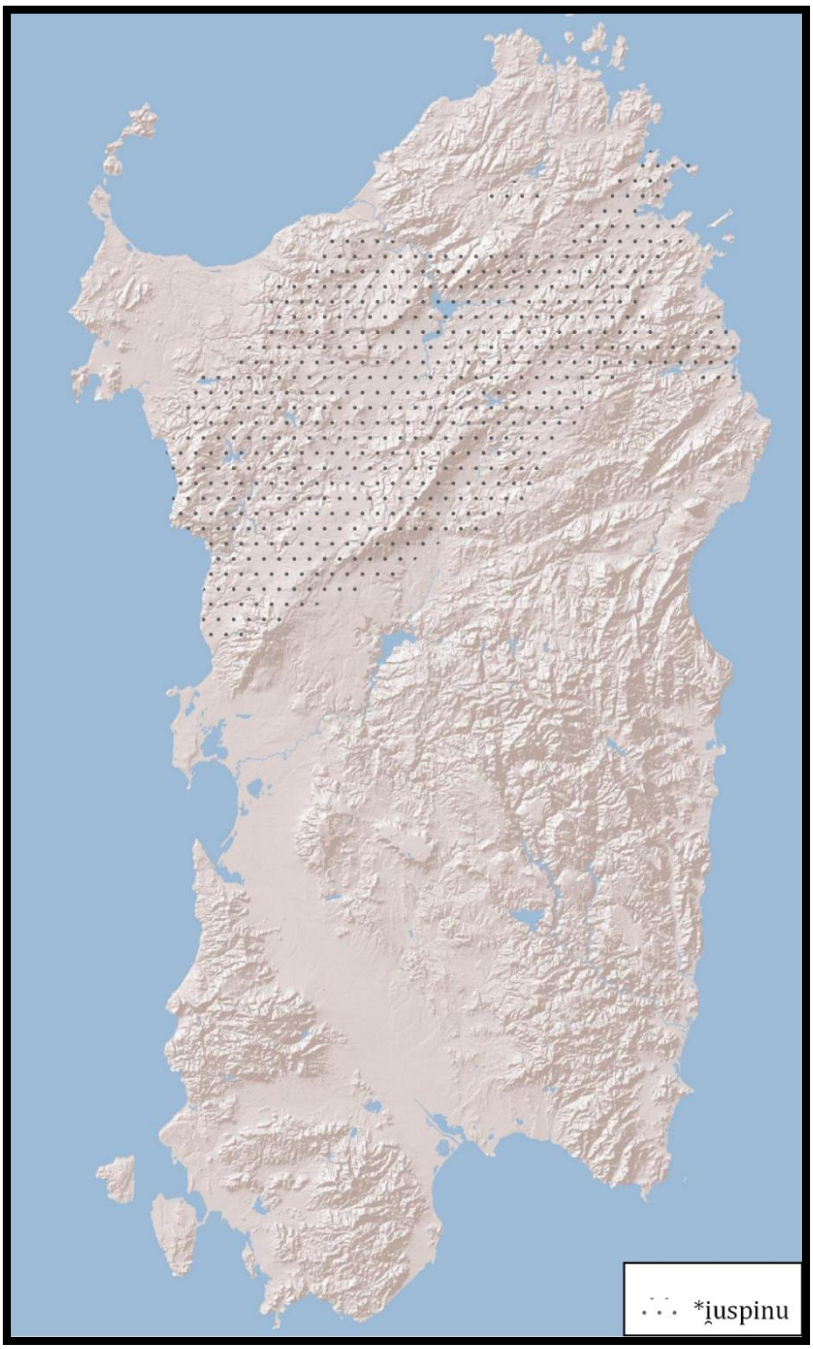

Figure 3.1.17: Mustard

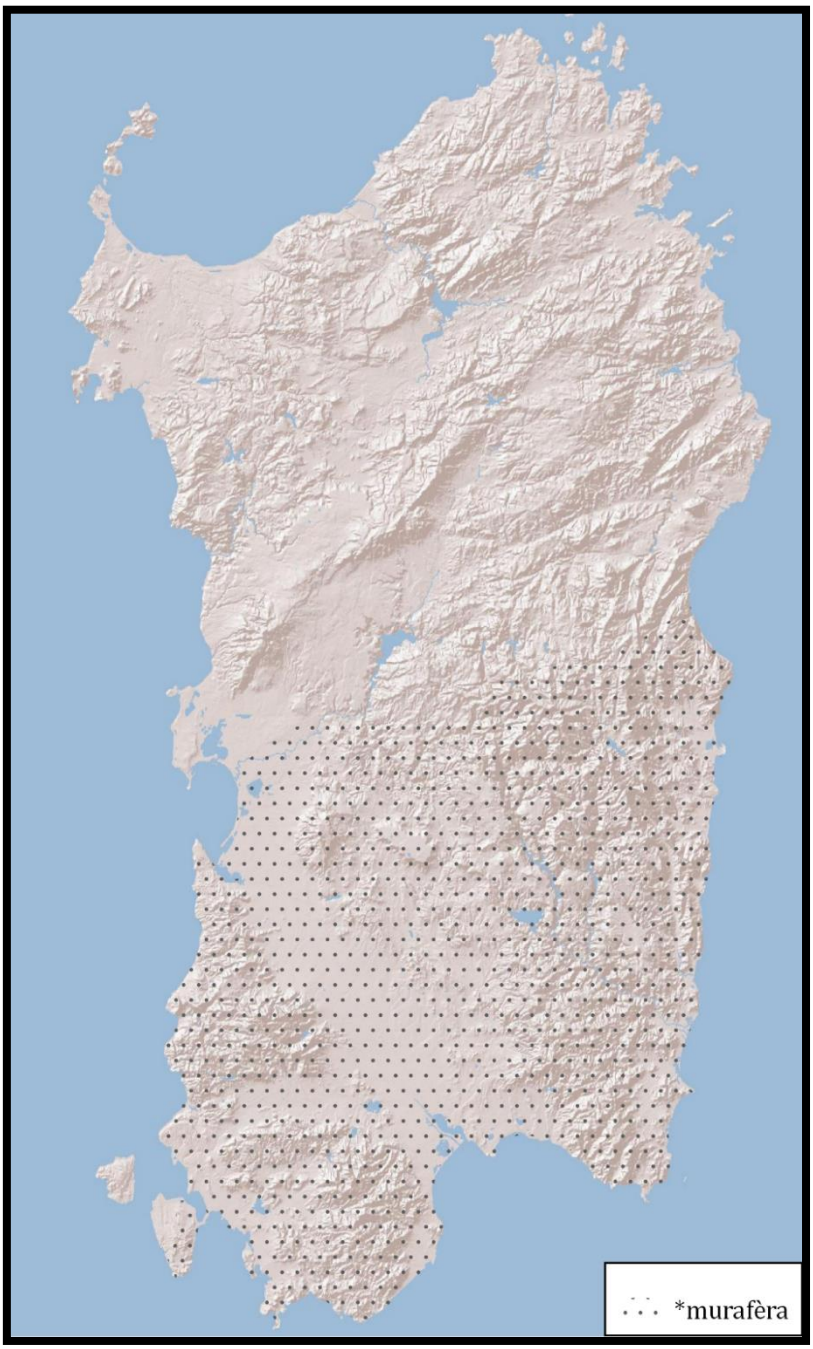

Figure 3.1.18: Black mustard (*Brassica nigra*)

3.1.18 Mustard, black (*Brassica nigra*)

Reconstruction	Forms
***murafèra**	DES (II: 138): ***muravèra*** (Camp.) 'mustard plant'
	Paulis (1992: 235): ***muravèra*** (Camp.) 'black mustard (*Brassica nigra* W.D.J.Koch)'

Paulis (1992: 235) derives Camp. *muravèra* 'black mustard' from Srd. *múra* 'mulberry' + *bèra* 'true', i.e. "true mulberry". Semantically this is quite a stretch. Paulis argues that this *Benennungsmotiv* for 'black mustard' stems from the color of mustard seeds, which would be perceived to be similar to that of mulberries. If Srd. *múra* 'mulberry' is at all involved, the second part must surely go back to Lat. *ferus* 'wild, uncultivated' rather than *verus* 'true'. This has two advantages. The first is that intervocalic **-f-* regularly develops into *-v-* (Wagner 1941a: 93), making Paulis' assumption that **-v-* in this form was retained under Italian influence unnecessary. The second is that explaining the meaning 'mustard' from 'wild mulberry' is perhaps slightly less unlikely than from 'true mulberry'. The doubts about the likelihood of a semantic shift from 'mulberry' to 'black mustard' remain however. There is no positive evidence for a pre-Roman origin.

3.1.19 Rue, fringed (*Ruta chalepensis*), cat thyme (*Teucrium marum*)

Reconstruction	Forms
***kúrma**	Paulis (1990: 608–612; 1992: 229–231): ***kúrma***, ***kúruma*** (1: Lodè, 2: Siniscola) 'fringed rue (*Ruta chalepensis* L.)'
	Farina (1987: 363; 1989: 620): ***crúma*** (Baronie reg.), ***úrma*** (3: Oliena) 'cat thyme (*Teucrium marum* L.)'

The forms listed above can all go back to **kúrma*. The variant *úrma* (3: Oliena) must probably be interpreted as /ʔúrma/, with the regular Barbaricine shift *k* > *ʔ* (cf. Wagner 1941a: 71). The variant *kúruma* is due to epenthesis, which is common (Wagner 1941a: 50–52; Paulis 1992: 231). Paulis (1990: 608–612; 1992: 231) convincingly connects this word to Pun. χουρμά 'rue', attested in Dioscorides' *Materia Medica* (Dioscorides Anazarbei 1906: 57).[70] Possible other links are Akk. *ḫurmu* 'a plant species' (CAD VI: 266) and Syrian (i.e. Dioscorides' Σύροι) ἁρμαλά 'rue (*Ruta chalepensis*)' (Dioscorides Anazarbei 1906: 57).[71] Múrcia Sànchez (2011: 440) ar-

[70] Dioscorides' use of Ἄφροι 'Africans' usually indicates the Punic name of a plant (Vattioni 1976: 519). See Vattioni (1976: 527–528) for all relevant glosses.

[71] Paulis (1992: 230) also connects supposed Tun. *kharma* 'St. John's wort (*Hypericum perforatum* L.)', for which he cites Rolland (1903: 8). This citation is probably mistaken however. In volume 4, p. 8, Rolland discusses 'rue', for which the Arabic name is not given. St. John's wort is instead discussed in volume 3 (Rolland 1900: 178), but the

gues that a Punic form **ḥurmal* was borrowed into Berber as **(h)urmay* > **urme*, represented by Bb. <*āwurmī*> (*ʿUmdaṭ aṭ ṭabīb*; Tilmatine & Bustamante 2002: 428), Tsh. *iwrmi* 'rue' (van den Boogert 1998: 84).[72] The existence of formally similar words in other Semitic languages makes it likely that Pun. χουρμά 'rue' is of Phoenician, rather than of local North African origin.

Because of the great semantic and formal correspondence between Pun. χουρμά 'rue' and Srd. *kúrma* etc. 'rue, cat thyme', it is attractive to see the Sardinian word as a loan from Punic. It is surprising that this word was only preserved in the Nuorese dialects, which is much farther north than one would perhaps expect for a Punic loan. Paulis (1992: 231) notes that this word must at one time have had a wider distribution in Sardinia, but that it was lost in many places due to competition with inherited Lat. *rūta*. Whether its preservation in the Baronie is indeed due to its association with magic (Paulis 1992: 231), is difficult to prove. Paulis (1992: 231) further observes that the fact that this Punic word is attested in the northern part of the island, does not necessarily imply that Punic was actually spoken there. If we assume that **kúrma* vel sim. was at some point a general word for 'rue' in Punic-influenced Latin spoken on Sardinia, it may have expanded up to its present-day location during the romanization of Sardinia, after it was lost everywhere else.

Tunisian Arabic form that is given, is *hhamra*. Paulis' *kharma* is thus probably a ghost word, and should not be used in the discussion of Srd. *kúrma*. The actual form *hhamra* is phonologically further removed from *kúrma* and χουρμά and may be unrelated.

[72] This form is attested both in modern Tashelhiyt and in the 17th c. lexicon of *al-Hīlālī* as <*iwrmi*> (van den Boogert 1998: 84; Múrcia Sànchez 2011: 440).

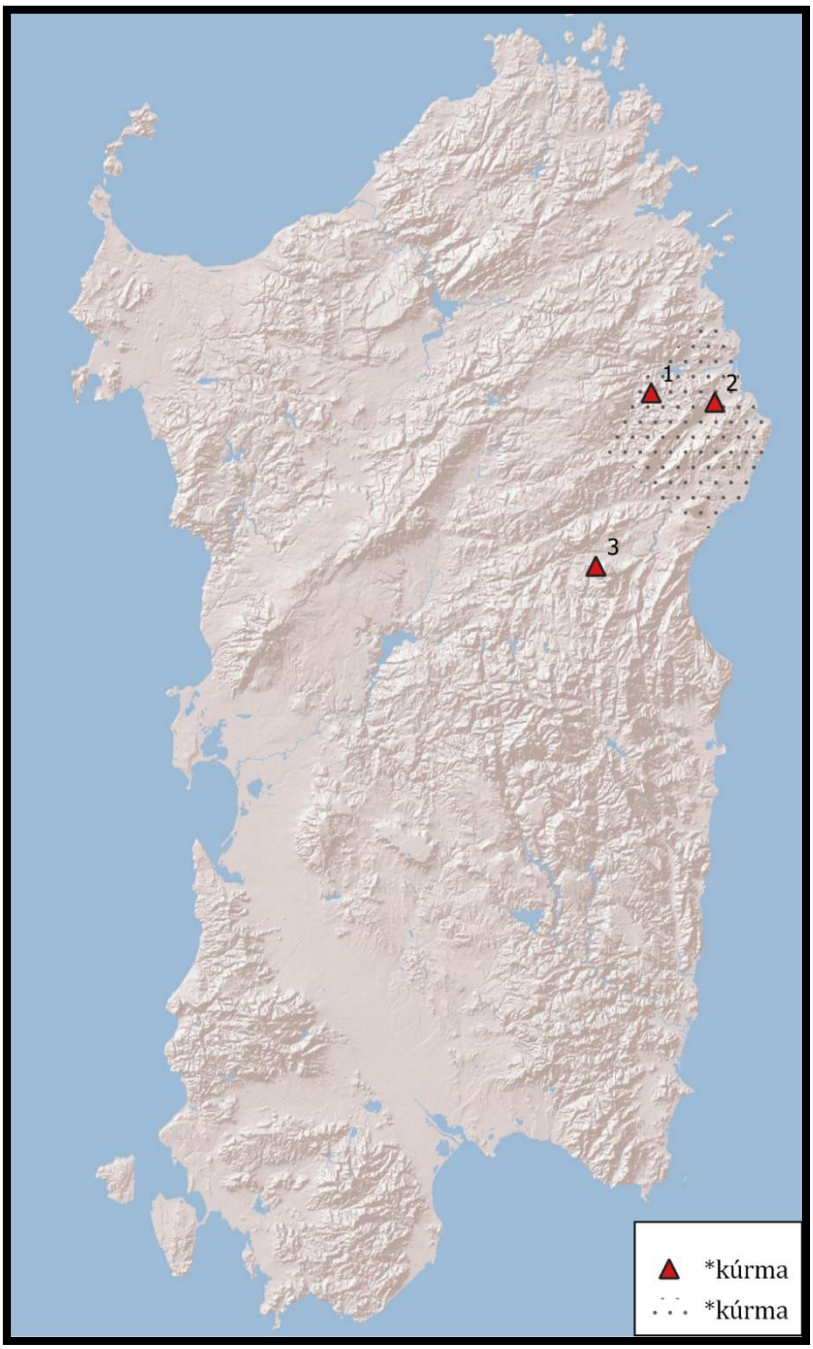

Figure 3.1.19: Fringed rue (*Ruta chalepensis*), cat thyme (*Teucrium marum*)

3.1.20 Rush skeletonweed (*Chondrilla juncea*), Sow thistle (*Sonchus oleraceus*)

Reconstruction	**Forms**
1: *lippòrra	DES (II: 31): ***lipòrra*** (Log.) 'rush skeletonweed (*Chondrilla juncea* L.), sow thistle (*Sonchus oleraceus* L.)'
2: *limpòrra	DES (II: 31): ***limpòrra, limpòra*** (Camp.) 'id.'
3: *lispòrra	DES (II: 31): ***lispòrra*** (Camp.) 'id.'

Wagner does not express any thoughts on the etymology of these words for 'rush skeletonweed' and 'sow thistle', which are rather similar plants, other than his remark that "the suffix *-orra* leads one to presuppose a pre-Roman origin". Paulis (1992: 62–63) proposes a derivation from the words for 'hare' — Nuor. *lèppore*, Log. *lèppere*, Camp. *léppuri*, *lèppuri*, *lèppiri* (§ 4.5.8) — by means of the same hypothetical pre-Roman suffix *-òrra*, which has supposedly become productive as a derogatory suffix in surnames. The association of these plants to hares has parallels in other languages (Paulis 1992: 62–63). For this, Paulis needs to assume haplology from **leppor-òrra* to *lipòrra* etc., as well as a raising of pretonic **-e-* to *-i-* and influence from other words to account for the vocalism of *limpòrra* and *lispòrra*. Although a derivation of the name of these plants from a word for 'hare' does not appear problematic semantically, the amount of assumptions needed for this etymology — i.e. derivation of a plant name from a word for 'hare' by means of an originally pre-Roman suffix usually present in surnames, haplology, vowel assimilation and influence from other words explaining the consonantism — significantly weaken its explanatory value. It must moreover be noted that the Sardinian lexicon does not in fact present evidence for the identification of *-Vrr-* as a pre-Roman suffix (§ 9.2.7). More convincing is Wolf's (1998b: 345–346) identification of this word as a continuant of **lepórna* < **lepórĭna* << Lat. *leporīnus* 'of hares, hare-'. The irregular presence of geminate **-pp-* in the Sardinian forms is also found in the word for 'hare' itself (§ 4.5.8).

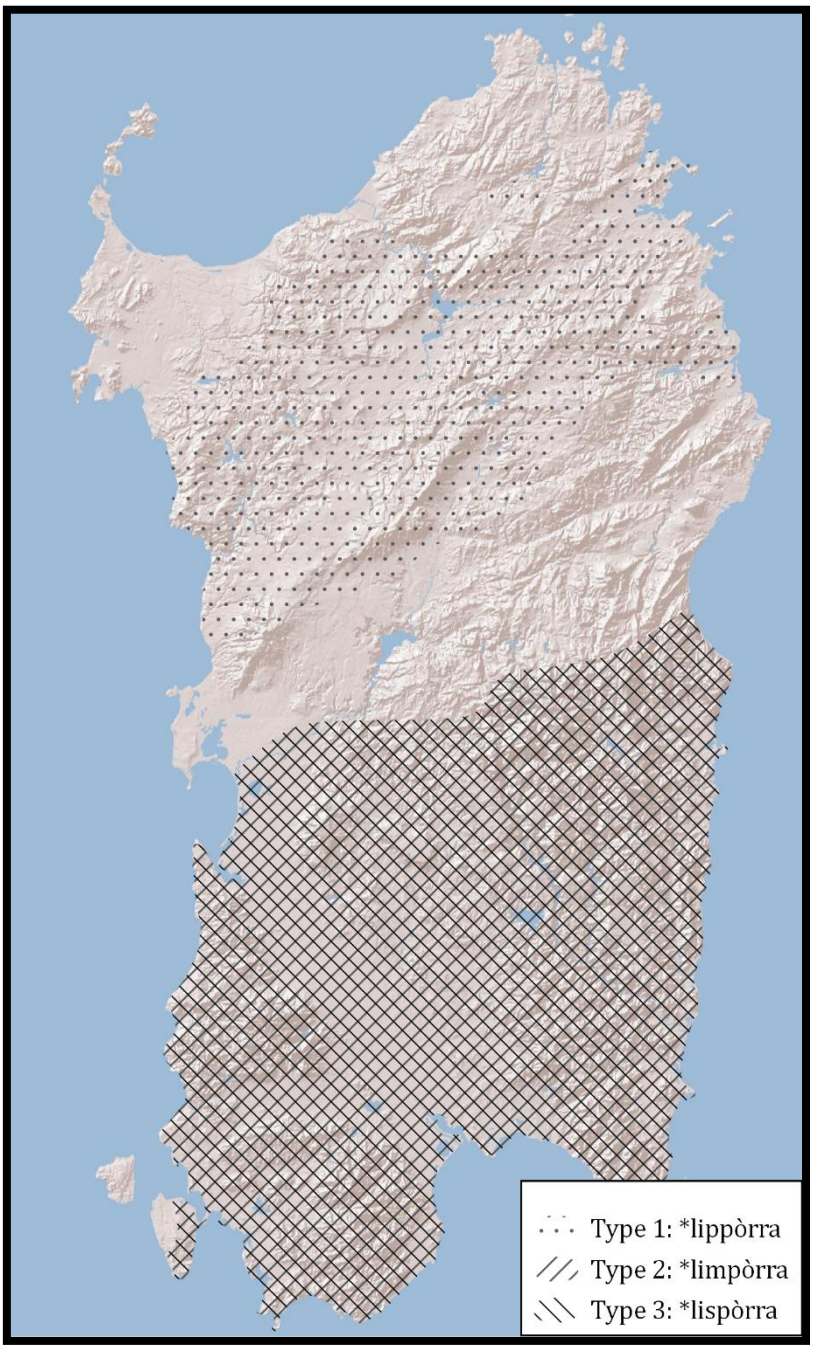

Figure 3.1.20: Rush skeletonweed (*Chondrilla juncea*), Sow thistle (*Sonchus oleraceus*)

3.1.21 Sea purslane (*Atriplex portulacoides*)

Reconstruction	Forms
***tsíbba**	DES (II: 589): ***tsíbba*** (1: Cabras) 'glasswort (*Salicornia europaea* L.)'
	Paulis (1992: 381): ***tsíbba*** (1: Cabras), ***attsíbba*** (2: Oristano) 'sea purslane (*Atriplex portulacoides* L.)'

After Paulis (1992: 381), the Sardinian forms listed above probably do not refer to 'glasswort', as Wagner (DES II: 589) suggests, but to 'sea purslane'. Both plants are halophytes and the leaves of both plants are eaten as a vegetable. Both Wagner and Paulis note that the plant in question is used in Cabras (1) to wrap boiled fish. Whereas Wagner does not propose any etymology, Paulis (1992: 381–382) compares it to Sic. *scebba*, *scerba* 'opposite-leaved saltbush (*Salsola soda* L.)',[73] Sp. *jebe*, OSp. *axebe*, *enxebe* 'alum' << Ar. شَبّ (*šabb*) 'alum, vitriol'. It is clear that the Sardinian forms ultimately go back to this Arabic word too, albeit not directly. Paulis (1992: 382) thinks it most likely that the Sardinian forms were borrowed from Spanish, with replacement of *š-* (= OSp. *-x-*) by *ts-* and *-é-* by *-í-*. One problem is that we would expect Sp. *-b-* to be borrowed into Sardinian as **-*β*- rather than *-bb-*; cf. Camp. *kaβeli̯èra* << Sp. *cabellera* 'hair', Camp. *skaβéčču* << Sp. *escabeche* or Cat. *escabetx* 'marinated fish' etc. (Wagner 1997: 207, 210). Geminate *-bb-* is however found in Sic. *scebba* 'saltbush', which is also closer in meaning to the Sardinian forms than Sp. *jebe* 'alum'. It is therefore more attractive to view *tsíbba* and *attsíbba* as a Sicilian loan. This is not surprising, as a good deal of Sardinian maritime terminology was borrowed from Sicilian (Wagner 1997: 251). It is in any case evident that this word is a relatively recent loan, and definitely not of pre-Roman origin.

[73] Paulis (1992: 381): "pianta che bruciata serve a fare il vetro, a sgrassare i panni e a fabbricare sapone: *Soda salicornia* 'salsola'" (= "plant that, when burnt, is used to make glass, to degrease clothes and to produce soap: *Soda salicornia* 'saltbush'"). Paulis cites Traina (1877; Paulis has 1888 however), but the page numbers cited do not exist. The word does in fact occur in Traina (1877: 384), but without the description quoted by Paulis. Note that *Soda salicornia* is not the currently accepted scientific name of this plant, which is *Salsola soda* L.

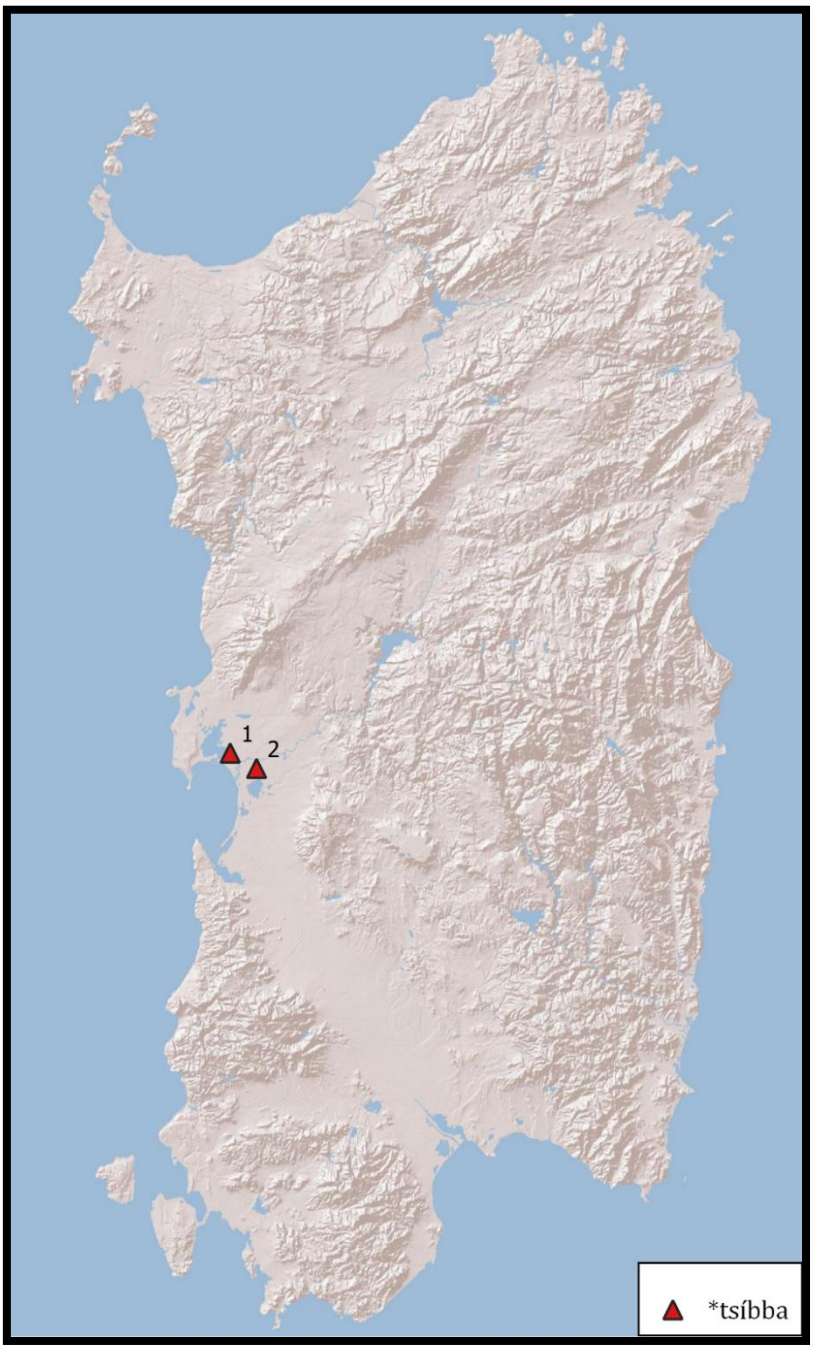

Figure 3.1.21: Sea purslane (*Atriplex portulacoides*)

3.1.22 Teasel, wild (*Dipsacus fullonum*)

Reconstruction	**Forms**
1: *bil̰ióttiri	DES (I: 206): ***billóttiri*** (1: Usellus) 'wild teasel (*Dipsacus fullonum* L.)'
	Paulis (1992: 190): ***billóttiri*** (5: Orroli) 'id.'
2: *bil̰ótti	Paulis (1992: 190): 2: ***billótti*** (2: Nuragus), [***kanna de***] ***billótti*** (3: Isili, 4: Serri) 'id.'

Paulis (1992: 190) rejects Wagner's (DES I: 206) suspicion that this word is "probably pre-Roman". He prefers to derive these forms from Camp. *baballótti, -u* 'insect', because this species of plants has leaves growing from the stem that form cuplike structures collecting rainwater, in which insects get trapped and dissolved for their nutrients. The phonetics would be guaranteed by the surnames *bibillótti* and *billótti* that are locally interpreted as meaning 'insect'.

While not impossible, there are some issues regarding this etymology. Paulis cites Camp. *baballótti, -u* 'insect' from Wagner, but Wagner (DES I: 161) actually has *babbalótti*. It is important to observe that *babbalótti* has a singleton *-l-*, which would be expected to yield *-ʙ-* in Orroli, Nuragus and Isili, and *-β/w-* in Serri (Wagner 1941a: 121–122). Geminate *-ll-* is instead the regular outcome of **-l̰-*.[74] Because of this formal problem, and the fact that the forms *bibillótti* and *billótti* are not actually attested in the meaning 'insect', it is possible that the similarity between these words is no more than a coincidence. The interpretation of the surnames as meaning 'insect' may be due to formal similarity to *baballótti, -u*. Moreover, a semantic shift from 'insect' to 'plant in which insects get trapped' without any derivational morphology would be hard to understand.

For these reasons, I prefer to assume a pre-form **bil̰ótti/-iri* 'wild teasel'. If this word is not to be derived from forms like *babbalótti* etc. 'insect', which is likely at least partly of sound-symbolic nature (cf. DES I: 162), a pre-Roman origin becomes attractive again. Paulis (1992: 190) explains the suffix *-́iri* as introduced from Camp. *sóččíri* 'wild teasel' (cf. § 3.1.22). This could be the case, but it may just as well have been originally present on this word as well. We do in any case observe an alternation between *billótti* and *billóttiri*, suggesting that *-́iri* was originally some morpheme rather than belonging to the stem (cf. § 9.2.6.2).

[74] It is theoretically possible that the geminate *-ll-* is due to Cagliaritan influence (cf. Wagner 1941a: 125), but since this specific word does not exist in the other Campidanese dialects, I do not think this is very likely.

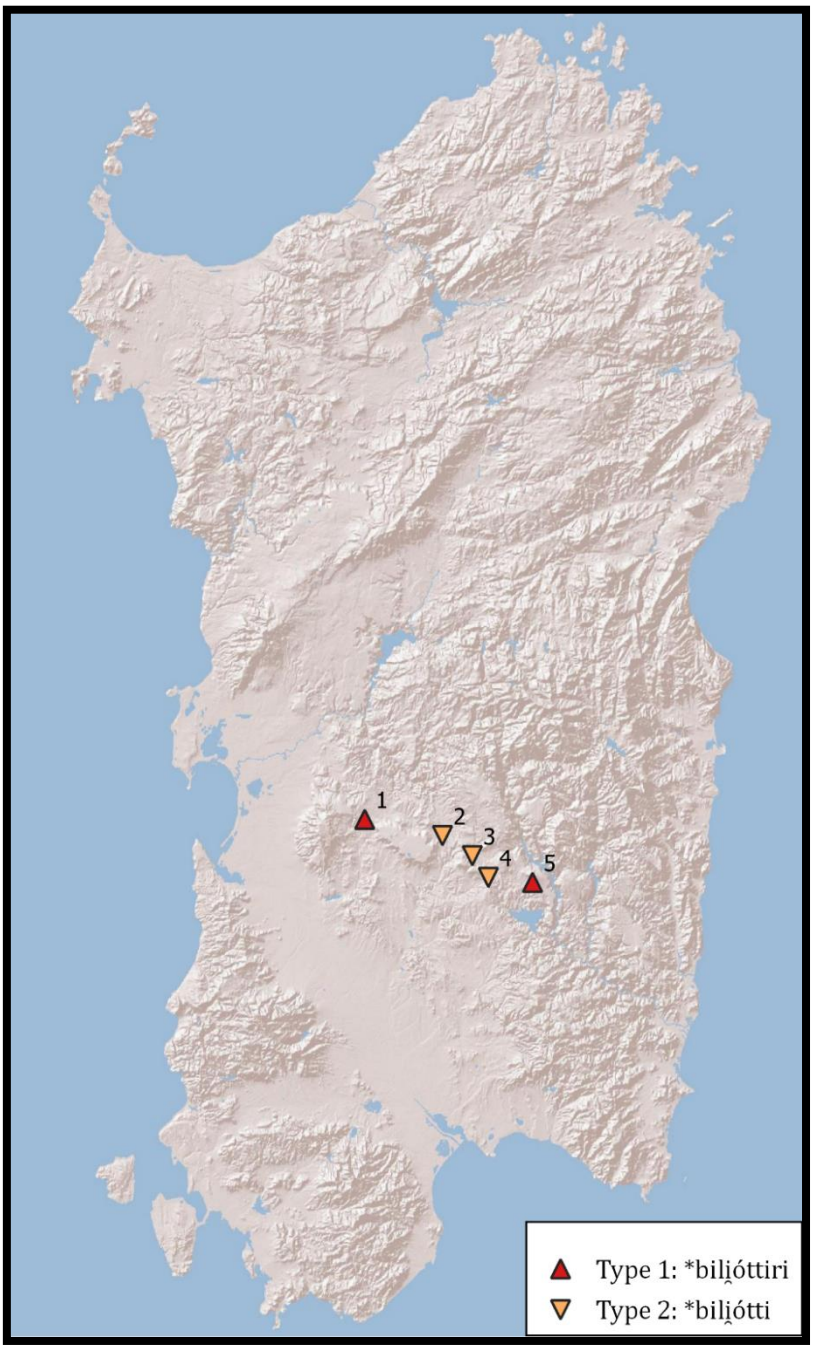

Figure 3.1.22: Teasel, wild (*Dipsacus fullonum*)

3.1.23 Watercress (*Nasturtium officinale*)ı; butcher's broom (*Ruscus aculeatus*)

Reconstruction	Forms
1: *gúspinu	DES (I: 595): ***grúspinu*** (1: Bitti), ***gúspinu*** (5: Orgosolo) 'watercress (*Nasturtium officinale* R.Br.)'
	Paulis (1992: 241): ***óspinu*** (3: Bolotana) 'watercress (*Nasturtium officinale*)'
2: *grúspi(n)u	DES (I: 595): ***grúspi̯u*** (2: Orosei), ***grúspinu*** (4: Dorgali) 'butcher's broom (*Ruscus aculeatus* L.)'

Wagner (DES I: 595) and Paulis (1992: 241) agree that the forms denoting 'butcher's broom' must be secondary. Wagner attributes this to the phonetic similarity between *grúspinu* etc. 'watercress' and *frúskru, frúski̯u* etc. 'butcher's broom (cf. DES I: 553). Paulis instead proposes a semantic cause for the shift from 'watercress' to 'butcher's broom'. Because of its use to treat abscesses, which in Sardinian are called *fruskèḍḍa* etc., watercress would have come to be associated to phonetically similar *frúskru, frúski̯u* 'butcher's broom' and hence this latter plant would have adopted watercress' original name. I am not fully convinced by this complex scenario; perhaps it was the combination of phonetic similarity between *grúspinu* 'watercress' and *frúskru, frúski̯u* 'butcher's broom', and the potential semantic association that caused this peculiar homonymy.

In the meaning 'watercress', Paulis (1992: 242) proposes this word to be borrowed from Pun. *cusmin* 'watercress', attested in Pseudo-Apuleius (Howald & Sigerist 1927: 59). This requires voicing of the initial consonant and adaptation of the cluster *-sm-* to *-sp-*, neither of which is particularly problematic.[75] In light of the semantic and phonetic correspondence of these forms, Paulis is likely correct. It is interesting to see that a word of Punic origin has such a northern distribution (cf. § 3.1.18). The dialects more to the south all have reflexes of Lat. *nasturtium* 'watercress' (cf. AIS 635).

[75] Paulis (1992: 242) suggests the change *-sm-* > *-sp-* to have taken place under influence of forms like Camp. *spinaðòppis* 'butcher's broom', again due to the same semantic association between both plant species. I rather believe it may simply be an adaptation of the cluster *-sm-* to the Latin phonological system.

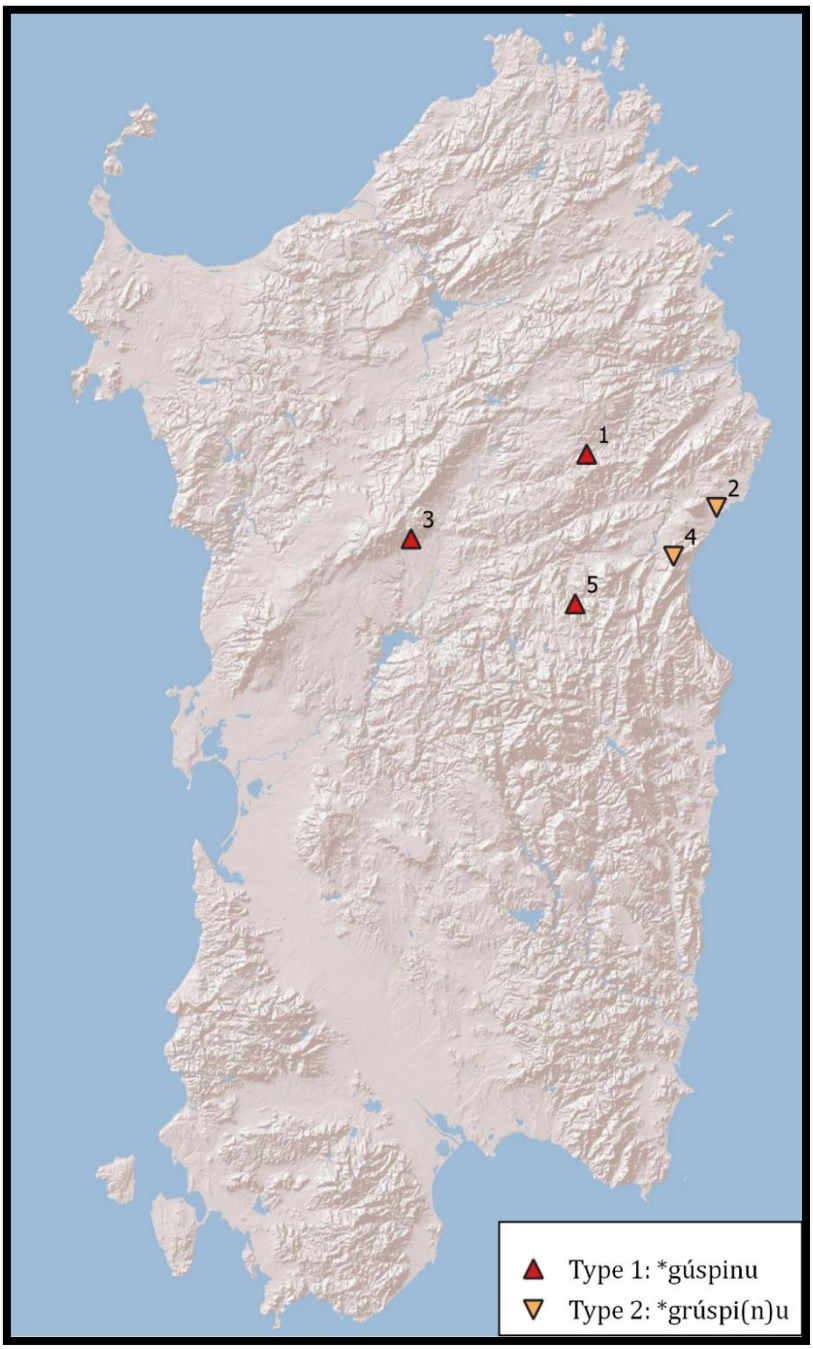

Figure 3.1.23: Watercress (*Nasturtium officinale*); butcher's broom (*Ruscus aculeatus*)

3.1.24 Watercress (*Nasturtium officinale*)$_2$, fool's watercress (*Apium nodiflorum*)

Reconstruction	Forms
1: *i̯úguru	DES (I: 601): ***ǵúru*** (1: Olbia) 'watercress (*Nasturtium officinale* R.Br.)', ***ǧúru*** (Log., 2: Sennori, 7: Mores, 9: Padria), ***ǧúɣuru*** (11: Fonni) 'fool's watercress (*Apium nodiflorum* (L.) Lag.)'
	Ferraro (1891: 90): ***giuru*** (5: Siligo) 'fool's watercress (*Apium nodiflorum*)'
2: *i̯úru-i̯úru	DES (I: 601): ***ǧuruyúru***, [***lu***] ***yuruyúru*** (3: Sassari) 'watercress (*Nasturtium officinale*)'
3: *i̯úi̯uru	DES (I: 601): ***ǵúǵǵuru*** (8: Bitti) 'fool's watercress (*Apium nodiflorum*)', ***ǧúǧǧuru*** (10: Nuoro) 'watercress (*Nasturtium officinale*)'
	AIS (635): ***ǵúǵuru*** (8: Bitti) 'some other species of plant growing on the banks of streams'[76]
	Paulis (1992: 241): ***ǧúǧaru*** (4: Lodè), ***ǧuǧaréḍḍu*** (6: Siniscola) 'fool's watercress (*Apium nodiflorum*)'

Two meanings are found for the forms listed above: 'watercress' and 'fool's watercress'. The former is found only in Olbia (1), Sassari (3) and Nuoro (10). Everywhere else this word refers to 'fool's watercress', which is therefore likely the original meaning. The confusion between both plant species, which are not closely related, is probably due to the fact that they often occur together in the same habitat, i.e. on the banks of streams. Both plants are edible, although fool's watercress is very similar to cutleaf waterparsnip (*Berula erecta* (Huds.) Coville), which is toxic.

Formally, different types can be distinguished. Type 2 *ǧuruyúru* is restricted to Sassari, and appears to be a reduplicated form of Log. *ǧúru* (type 1). Log. *ǧúru* and *ǧúɣuru* (11: Fonni) in type 1 go back to **i̯úguru*, with regular loss of intervocalic voiced stops in Logudorese (Wagner 1941a: 79). Type 3 requires **i̯úi̯uru/*i̯úi̯aru*. I agree with Paulis (1992: 246) that the forms of type 3 are probably due to assimilation of the second consonant to the onset of the word, i.e. **i̯úguru* >> **i̯úi̯uru*.

[76] Probably 'fool's watercress', since this form is included in the map dedicated to 'watercress'.

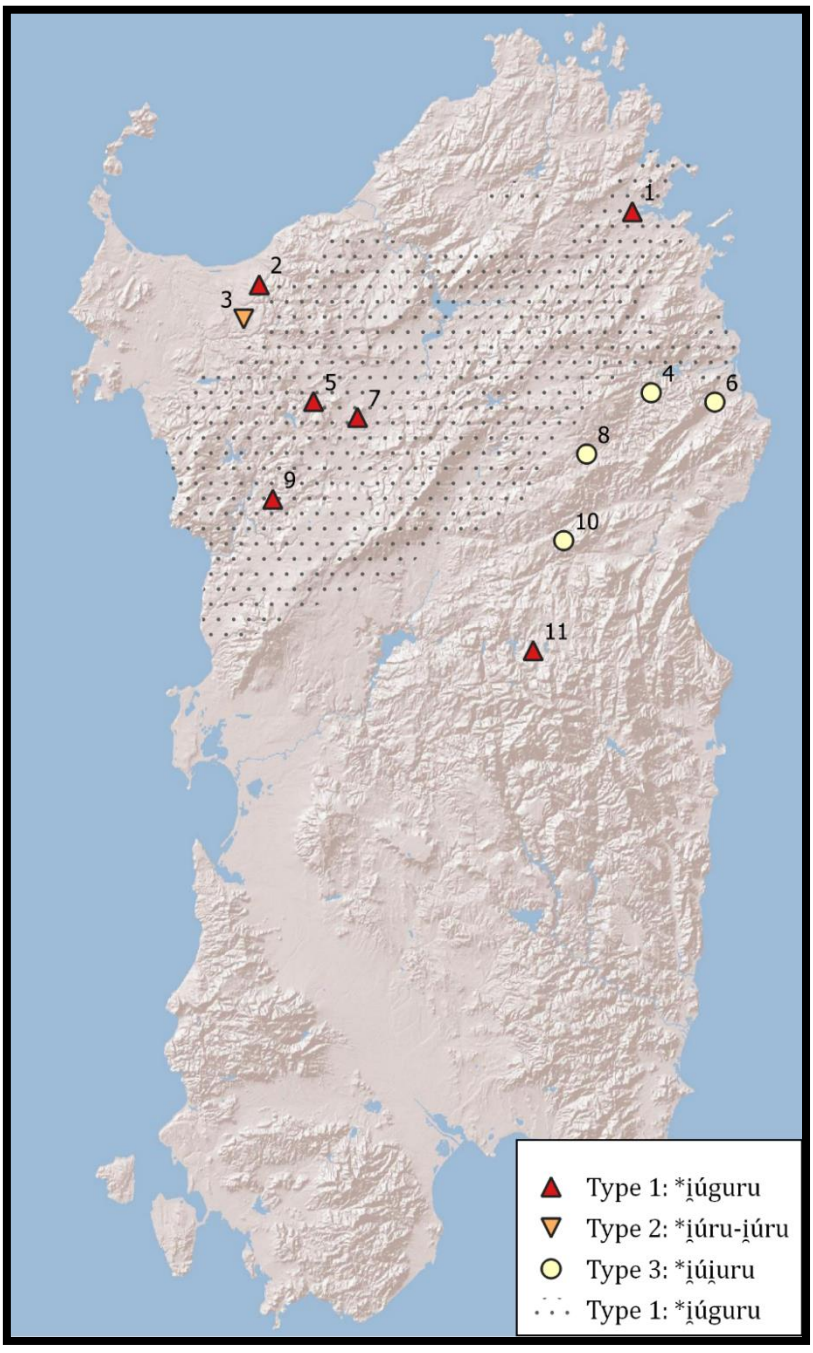

Figure 3.1.24: Watercress (*Nasturtium officinale*), fool's watercress (*Apium nodiflorum*)

Paulis (1992: 245–246) proposes these forms to be derived from Lat. *iūgis* 'continuous', which often occurs in the construction *aqua iugis* 'flowing water' (cf. Lewis & Short 1879: s.v. *jūgis*). The reason for this would be that both watercress and fool's watercress grow on the banks of streams. However, as Paulis (1992: 246) notes, the suffix *-́uru/-́aru* cannot be explained on the basis of inherited material and should most likely be connected to the element *-́Vr-* that is often found in words of suspected pre-Roman origin (§ 9.2.6.2). This etymology requires a semantic shift from 'continuous' or 'flowing [water]' to '(fool's) watercress'. Despite the habitat of these plants, such a semantic shift is not convincing. Watercress and fool's watercress are by no means the only plants growing on riverbanks and it is not clear to me how an adjective referring to a property of water could come to refer to a specific species of plant. Since we already need to assume a pre-Roman origin for the ending of this word, it is more attractive posit a pre-Roman origin for the entire form, as Wagner (DES I: 601) already suspected.

3.2 Grasses, cereals and weeds

3.2.1 Bermuda grass (*Cynodon dactylon*); butcher's broom (*Ruscus aculeatus*), holly (*Ilex aquifolium*)

Reconstruction	**Forms**
1: *(*nodu*)-*ollásu*	Paulis (1992: 417): ***oroḍḍáṡile*** (1: Olzai), ***noroḍḍáṡile*** (2: Gavoi), ***oroḍḍáṡu*** (3: Ovodda) 'Bermuda grass (*Cynodon dactylon* (L.) Pers.)'
2: *alásu	DES (I: 67-68): ***aláṡu*** (4: Urzulei, 6: Triei, 9: Villagrande Strisaili, 12: Tortolì), ***olláṡu*** (13: Perdasdefogu, 14: Villaputzu) 'Bermuda grass (*Cynodon dactylon*)'
3: *alase/-i	DES (I: 67-68): ***aláṡe*** (6: Désulo), ***aláṡ̱u*** (8: Belvì), ***aláṡi*** (10: Aritzo), ***oláṡi*** (11: Meana Sardo) 'butcher's broom (*Ruscus aculeatus* L.); holly (*Ilex aquifolium* L.)'
	Blasco Ferrer (2019: 40): ***aláṡe*** (5: Austis) 'butcher's broom (*Ruscus aculeatus*); holly (*Ilex aquifolium*)'

In the material listed above, we find several similar formations for very different plants. On the one hand, there are the forms of types 1 and 2 referring to 'Bermuda-grass (*Cynodon dactylon*)'; on the other hand, the forms of type 3 are formally nearly identical with those of type 2, but refer to 'holly' or 'butcher's broom'. Type 2 and 3 seem to contain a base **alas-*. Type 1 has been identified by Paulis

(1992: 417) as *nóðu* 'knot'[77] + **ollas-*(*ile*), which would be the same etymon as type 2 *ollásu*.

There are several issues that require attention in these forms. Most obvious is the striking difference in meaning between 'Bermuda grass' on the one hand and 'holly/butcher's broom' on the other. Type 2 'Bermuda grass' and type 3 'holly/butcher's broom', though very similar, do differ in their final vowel, with *-u* in type 2 and *-e*, *-i*, *-i̯u* in type 3.[78] This may suggest that these words reflect different original formations and that their similarity is the result of a coincidence rather than of a drastic semantic change.

The single form in type 1 that does not carry the suffix *-ile* (i.e. *oroḍḍáśu*), shares its final *-u* with type 2, speaking in favor of a shared origin of these types. It is difficult, however, to explain the difference between *-l-/-ll-* in type 1 and *-ḍḍ-* in type 1. While *-l-/-ll-* must go back to singleton **-l-*, the *-ḍḍ-* found in type 1 can only reflect original geminate **-ll-*. We are thus confronted with three reconstructions: **ollásu* 'Bermuda grass' (type 1), **alásu* 'Bermuda grass' (type 2), and **aláse/-i* 'holly/butcher's broom' (type 3). Due to the perfect semantic match and phonological similarity, it is unattractive to separate types 1 and 2, even if the phonetic details cannot readily be understood. To account for the semantics of the apparent link between 'Bermuda grass' and 'holly/butcher's broom', Paulis (1992: 418) proposes that hypothetical **alas-* may originally have referred to some other plant with spines. On the one hand this would have developed into 'Bermuda grass', on account of the spiny appearance of this grass species (as reflected by its scientific name *Cynodon* 'dog's tooth'), and on the other hand into holly, which has spiny leaves. Though imaginable, it would still be a rather striking semantic divergence of a word that occurs only in a rather limited area in Sardinia.

[77] The presence of the word for 'knot' would be due to the tendency of this type of grass to propagate itself through its stalks and roots, forming a very dense and tangly mat. The initial *n-* was supposedly lost due to its misinterpretation as belonging to the indefinite article (Paulis 1992: 418).

[78] The *-i̯u* must be secondary from **-i*, since **-si̯-* regularly produces voiced *-ṡ-* in Sardinian (Wagner 1941a: 150).

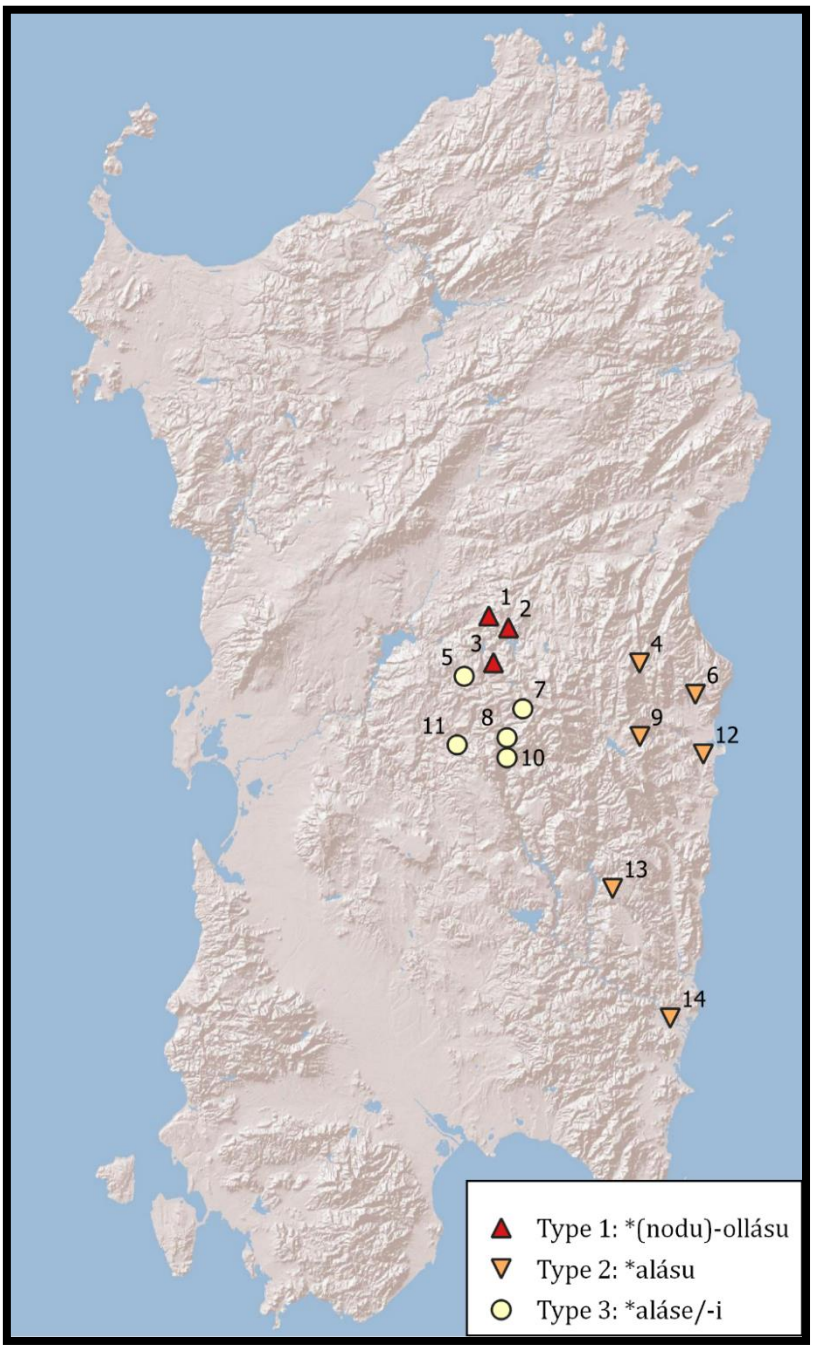

Figure 3.2.1: Bermuda grass (*Cynodon dactylon*); butcher's broom (*Ruscus aculeatus*), holly (*Ilex aquifolium*)

Outside Sardinian, these words have been compared to Sic. *alastra* 'spiny broom (*Calicotome spinosa* (L.) Link)', Lig. *la lastra*, *lástroga* (Bertoldi 1928: 233 fn. 3; Battisti 1960: 373), Lig. *arástra* 'spiny broom (*Calicotome spinosa*), spiny shrub' (Battisti 1960: 373; Paulis 1992: 418); Prov. *ářę* (dep. Alpes-de-Haute-Provence) 'broom', (dep. Var) 'broom, branch of broom', Bearn. *aroû* 'broom' < **ala*; Auv. *alète* (dep. Allier) 'common broom' < **alista*; Lang. *orousto* (dep. Aveyron) 'branch of broom' < **alusta* (FEW XXI: 107; Paulis 1992: 418); Lat. *alaternus* 'buckthorn (*Rhamnus alaternus* L.)' (Battisti 1960: 374).[79] The Gallo-Romance and Italo-Romance forms suggest a base **ala* that was extended with an **-st*-suffix in some cases, with the meaning 'broom'. In order to connect the Sardinian forms, we would have to assume a different suffix **-ás-* (cf. § 9.2.8), and a semantic change that could explain the coexistence of the meanings 'broom', 'holly' and 'Bermuda grass'.

Although a connection between Srd. *aláṡe* etc. 'butcher's broom/holly' and **ala*, **alastra*, **ali/usta* 'broom' seems plausible both semantically and phonologically, the existence of Srd. *aláṡu* etc. and *noroḍḍáṡile* etc. << **ollás-* 'Bermuda grass', which in fact has a wider distribution than the near-homophonous word for 'butcher's broom/holly', poses some serious semantic and phonological challenges. Perhaps this problem can be resolved by assuming two phonologically similar, but not quite identical forms referring to 'holly' and 'Bermuda grass' respectively. An indication for this may be the fact that the words for 'Bermuda grass' consistently end in *-u*, unlike the words for 'holly, butcher's broom', as well as the geminate *-ḍḍ-* < **-ll-* in *noroḍḍáṡile* etc. 'Bermuda grass' which cannot correspond regularly to the *-l/ll-* found in *aláṡu*, *olláṡu* 'Bermuda grass'. Nevertheless, if the Sardinian forms (in either meaning, but most likely in the meaning 'butcher's broom/holly') are to be connected to the other Romance forms mentioned above, this provides us with potentially interesting information about pre-Roman linguistic connections in the Mediterranean (§ 11). Finally, Wagner (DES I: 67) also compares Srd. *alaϑúli* (§ 3.3.6), which probably means either 'butcher's broom' (DES I: 68), or 'holly' (Paulis 1992: 417).

[79] Battisti also mentions Gr. ἄλιξ in the meaning 'holly'. According to Beekes (2010: 69), however, it means 'groats of rice-wheat'.

3.2.2 Esparto grass (*Lygeum spartum*)

Reconstruction	**Forms**
1: *ϑinníga/ *ϑinníka	DES (II: 547-548): ***tinnía*** (Log.), ***zinnía*** (Sass.), ***tsinníɣa*** (Camp.), ***ϑinnía*** (1: Orosei, 3: Triei), ***ϑinníɣa*** (2: Nuoro) 'esparto grass (*Lygeum spartum* Loefl. ex L.)'
2: *ϑónni	DES (II: 547-548): ***tsónni, tsónn̰ia, sónn̰ia*** (Camp. "in the villages") 'id.'

The Sardinian forms for this word come in two types. Type 1 must go back to **ϑinníga*, attested in the Condaghe di San Pietro di Silki as *thinnigas* (Bonazzi 1900: sec. 425) and in the Condaghe di San Michele di Salvennor as *tinigas* (Maninchedda & Murtas 2003: sec. 164). The apparent preservation of intervocalic **-g-* as *-ɣ-* in Camp. *tsinníɣa* is unexplained; we would expect ***tsinnía*. The forms of type 2 require original **ϑónni*, to which a secondary *-a* was added in *tsónn̰ia, sónn̰ia*, since inherited **-nn̰-* would have yielded ***-nǧ-* in Campidanese (cf. Wagner 1941a: 144).

These Sardinian forms have been related by Bertoldi (1950: 16) to Nef. *tsənnít* 'esparto grass (*Lygeum spartum*)', and *tsunít* 'basket woven from rushes'. If we compare more Berber material however — e.g. Nef. *tsənnít* 'esparto grass, of inferior quality for ropes, mats etc.' (Beguinot 1942: 237), Nef. *tsunít* 'basket' (Beguinot 1942: 223), Tuar. *tesănit* 'big basket with two handles' (Ritter 2009: 434), Kb. *isni* 'big basket made from esparto grass' (Dallet 1982: 783), Tazerwalt *tisnit, tiznit* 'basket used for work in the garden' (cf. Múrcia Sànchez 2011: 468; Argiolas 2020: 41; Laoust 1920: 37 fn. 4) — it seems evident that the original Berber meaning is 'basket (made from rushes/esparto grass)' and that Nef. *tsənnít* 'esparto grass' represents a semantic extension. The Berber forms can go back to **sănV̌y/*sūnV̌y/*sănī/*sūnī.*[80]

As far as the Sardinian forms are concerned, especially Camp. *tsónni* etc. may be comparable to Bb. **sūnī*. The other Sardinian forms, from **ϑinníga*, contain a *-g-* that is not present in Berber and that cannot easily be explained as an inherited Romance suffix either. Thus, while the comparison of Camp. *tsónni* 'esparto grass' to Bb. **sūnī* 'basket made from rushes/esparto grass' may look promising if we allow for either a semantic shift from 'basket' to 'esparto grass' independently in Sardinian and Nefusi Berber or a similar polysemy at the time of borrowing, a direct relation between Srd. **ϑinníga* and the Berber forms is much less convincing.

[80] Maarten Kossmann, p.c.

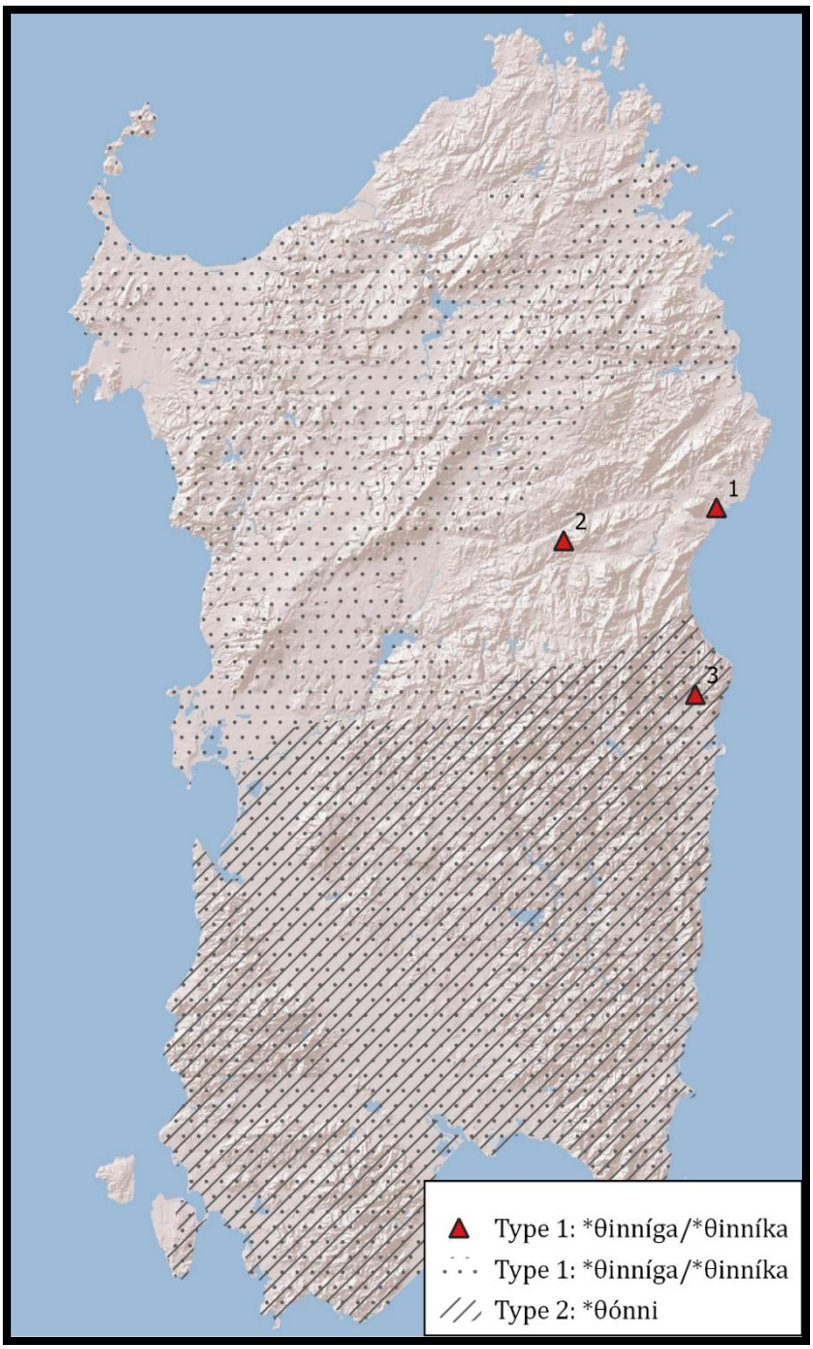

Figure 3.2.2: Esparto grass (*Lygeum spartum*)

Paulis (1992: 271) proposes a scenario in which both Sardinian types of this word entered Sardinia during Carthaginian presence in Sardinia, but that the more restricted and southerly distribution of type 2 **ϑónni* is due to it being introduced by a later wave of "Lybian" settlers in this same period. Argiolas (2020: 41) instead argues for a double introduction, at different times. The form **ϑinníga*, which has a wider distribution, would have belonged to a pre-Punic "Libyco-Berber" substrate in Sardinia. The etymologically related form **ϑónni*, restricted to Campidanese, would have been reintroduced in Sardinia during the Punic colonization, presumably through Berber-Punic language contact.

Yet, due to the phonetic problems of comparing **ϑinníga* to the Berber forms, either scenario is rather speculative. If Camp. *tsónni* 'esparto grass' is indeed related to a Bb. **sūnī* 'basket', we either need to assume that there was some unattested North-African variant containing a *-g-* that entered Sardinian as **ϑinníga*, which is unattractive because of a complete lack of evidence, or that Camp. *tsónni* and pan-Srd. **ϑinníga* are not related to each other. Alternatively, if one prefers to regard *tsónni* and **ϑinníga* as cognates, the Berber link must be rejected. As both scenarios still have their problems — the semantic development from 'basket' to 'esparto grass' needed to connect Camp. *tsónni* and Bb. **sănV̆y*/**sūnV̆y*/**sănī*/**sūnī*, although probably attested in Berber, is far from trivial, but the phonetic and morphological developments explaining Srd. **ϑinníga* vs. **ϑónni* are not easily explained either. Deciding which one to accept is difficult. There is in any case no compelling reason to assume a double introduction of the same North-African word.

Decidedly less convincing are Pittau's (apud Paulis 1992: 270) connection of both the Sardinian and Berber forms to Lat. *fūnis* 'rope', which is of Indo-European origin (de Vaan 2008: 250–251), and Argiolas' (2020: 41) connection of all these forms to Gr. θῶμιγξ 'cord, string'. While the latter is of probable pre-Greek origin (Beekes 2010: 569), its similarity to Srd. **ϑinníga*, **ϑónni* 'esparto grass' is restricted to initial <ϑ->. The Vulgar Latin form **seniculum* that Múrcia Sànchez (2011: 468) posits on the basis of Cat. *senill* 'type of reed', is formally incompatible with the Sardinian forms.

3.2.3 Galingale (*Cyperus longus*)

Reconstruction	**Forms**
***sèssene**	DES (II: 412): ***sèssini*** (Camp.), ***sèssene*** (1: Desulo, 2: Laconi), ***sèssiri*** (4: Villacidro) 'galingale (*Cyperus longus* L.)'; deriv.: ***fuiṡèssini*** (3: Ballao) 'Italian three-toed skink (*Chalcides chalcides* Linnaeus, 1758)'
	Paulis (1992: 449): [***mamma de***] ***séssini***, ***sessinárğu*** (Camp.) 'purple nut-sedge (*Cyperus rotundus* L.)'
	Puddu (2023: s.v. sènsene): ***sènsene***, ***sèntzini***, ***sèssi***, ***sèssini***, ***sèssiri*** 'galingale (*Cyperus longus*)'; [***su***] ***sessinédhu*** 'kind of rope'; [***mamma de***] ***sèssini*** (etc.), ***sessinàrgiu*** 'great bulrush (*Schoenoplectus tabernaemontani* (C.C.Gmel.) Palla)'

Galingale, also known as 'water rush', is a species of sedge that grows in wet places such as the banks of ponds etc. and has been used to make mats (DES II: 412). The Sardinian forms can be united in an original **sèssen(e)*. This word is generally considered to be of pre-Roman origin and has been compared to various Berber forms: e.g. Kb. *azəzzu* etc. 'broom (plant)' (Hubschmid 1953: 27; cf. Dallet 1982: 927); Beni Snous *tiyzzi* 'dry esparto grass' (DES II: 412; cf. Destaing 1914: 10),[81] Beni Snous *ẓẓiw* 'strand' (DES II: 412; cf. Destaing 1914: 50).[82] The connection to the words for 'broom' and 'strand' can readily be rejected on semantic grounds, and I fail to see the similarity between Srd. **sèssen(e)* and Beni Snous *tiyzzi*. Argiolas (2020: 40–41), who has her doubts about these earlier comparisons, proposes a connection to Kb. *ẓẓəmẓəm* 'to be slim', *aẓəmẓum* 'twig' (cf. Dallet 1982: 945), Tmz. *azmu* 'very fine rush used to make mats' (cf. Taïfi 1991: 807) etc. The only formally acceptable comparanda to Srd. *sèssene* 'galingale' are reduplicated forms like Kb. *aẓəmẓum* 'twig', which are etymologically unrelated to Tmz. *azmu* 'very fine rush used to make mats' and its cognates.[83] The semantic distance between words meaning 'twig' or 'slim' and Srd. *sèssene* 'galingale' is too great to confidently posit an etymological connection.

[81] In Destaing (1914) and DES as *ϑiizzi*.

[82] In Destaing (1914) as *ẓẓẹu*, and in DES as *ž'zeu*.

[83] Maarten Kossmann (p.c.).

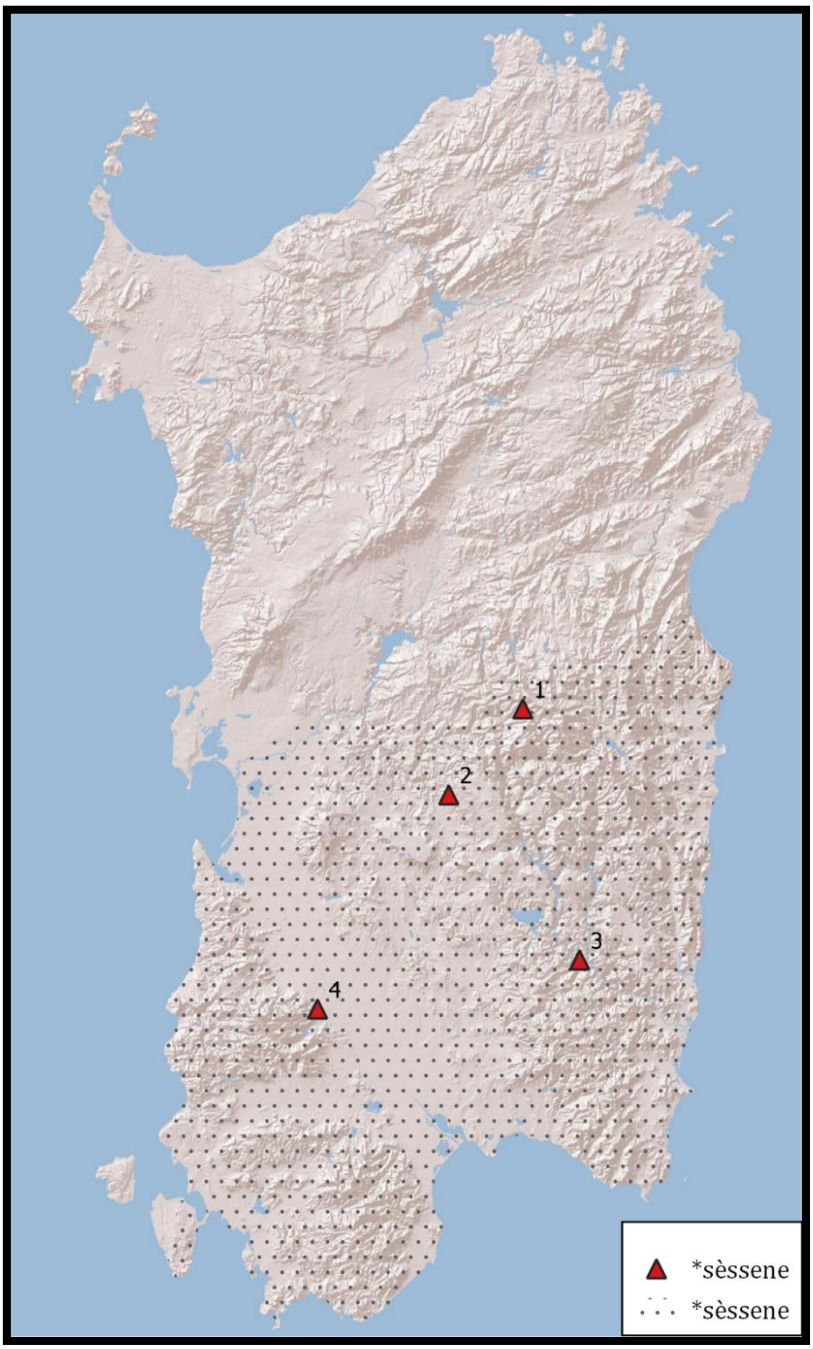

Figure 3.2.3: Galingale (*Cyperus longus*)

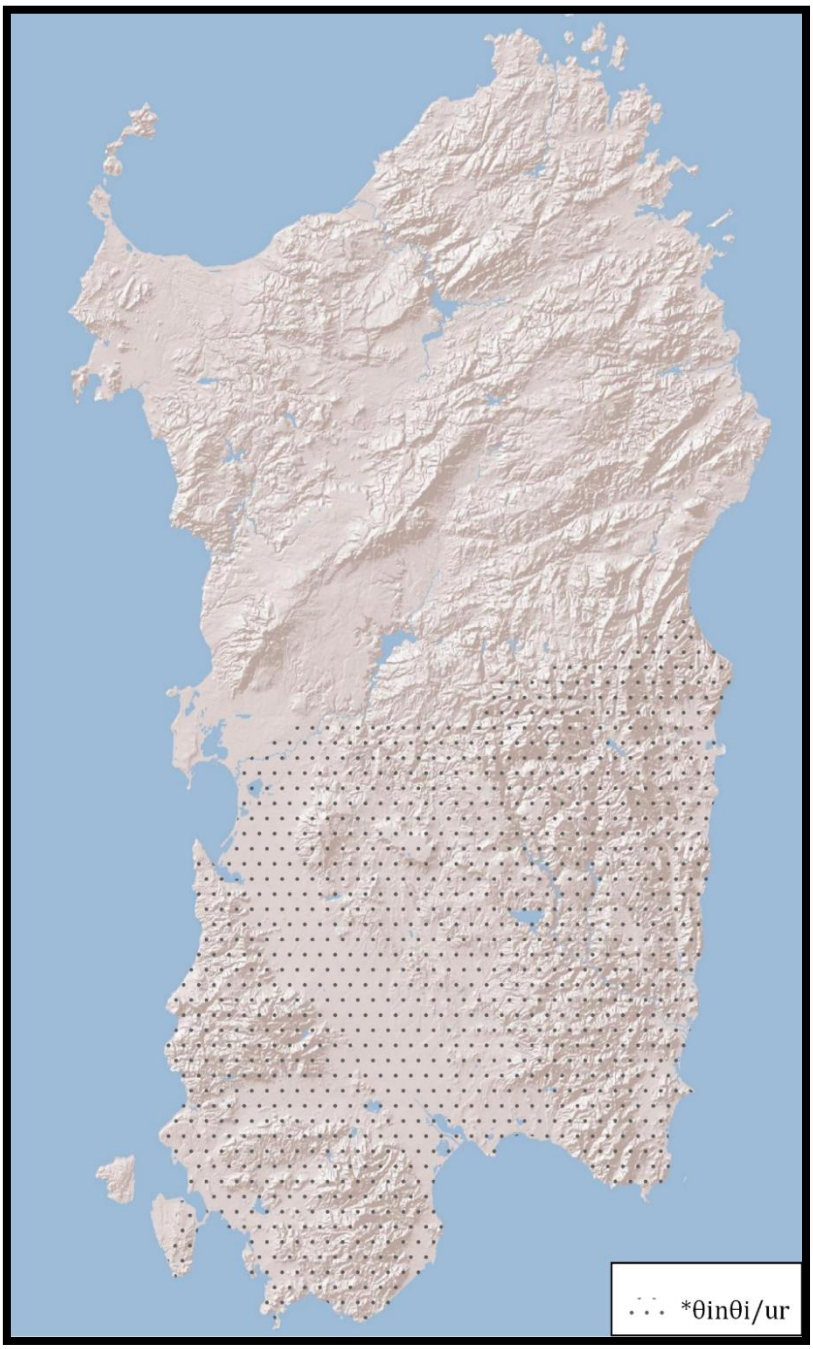

Figure 3.2.4: Knotgrass (*Polygonum sp.*); horsetail (*Equisetum palustre*)

3.2.4 Knotgrass (*Polygonum sp.*); horsetail (*Equisetum palustre*)

Reconstruction	Forms
***ϑinϑi/ur**	Paulis (1992: 386–387): ***síntsiri, sintsurru***[84] (Camp.) 'knotgrass (*Polygonum aviculare* L., *P. equisetiforme* Sm.); horsetail (*Equisetum palustre* L.)'

These Campidanese forms have convincingly been connected by Paulis (1992: 386–387) to Pun. *sensur*, *zenzur*, *zunzur, sumsur* 'knotgrass' (Howald & Sigerist 1927: 55; cf. André 1985: 279). As in other instances of possible Punic loans, sibilants are borrowed into Sardinian as *ϑ- (§ 8.1.2). This is yet another interesting example of a word of Punic origin preserved in the Sardinian dialects.

3.2.5 Ropegrass (*Ampelodesmos mauritanicus*)

Reconstruction	Forms
***kalkúri**/ ***karkúri**	DES (I: 298): ***karkúri, krakkúri, krukkúri; krukkúr̥u*** (1: Mogoro) 'ropegrass (*Ampelodesmos mauritanicus* (Poir.) T.Durand & Schinz)'

Except for *krukkúr̥u*, Wagner does not provide any information on the place of attestation of these words for 'ropegrass'. Nor does he provide any etymological discussion apart from noting that they are "probably pre-Roman". The pretonic *-u-* in *krukkúri*, *krukkúr̥u* is likely due to assimilation to the following stressed vowel. Paulis (1992: 269) suggests that these forms may somehow be related to Log. *kárku*, Camp. *krákku* 'thick, dense, packed' and Camp. *krákkili*, *krákkiri* 'thick bush, dense thickets', be it etymologically or folk-etymologically. This is likely, as ropegrass forms large, thick tufts. However, the ending *-úri* is not an inherited morpheme, but is occasionally found in words of supposed pre-Roman origin. This makes one wonder if the formation of *karkúri* etc. is the result of the addition of a pre-Roman suffix to an inherited stem. Alternatively, an already existing word of pre-Roman origin for 'ropegrass' may have been reformed into *karkúri* etc. by influence of *kárku*, *krákku* 'thick, dense, packed' << *karkáre* 'to tread, to press' < Lat. *calcāre* 'id.'.

[84] This form has been recorded without information about its accentuation.

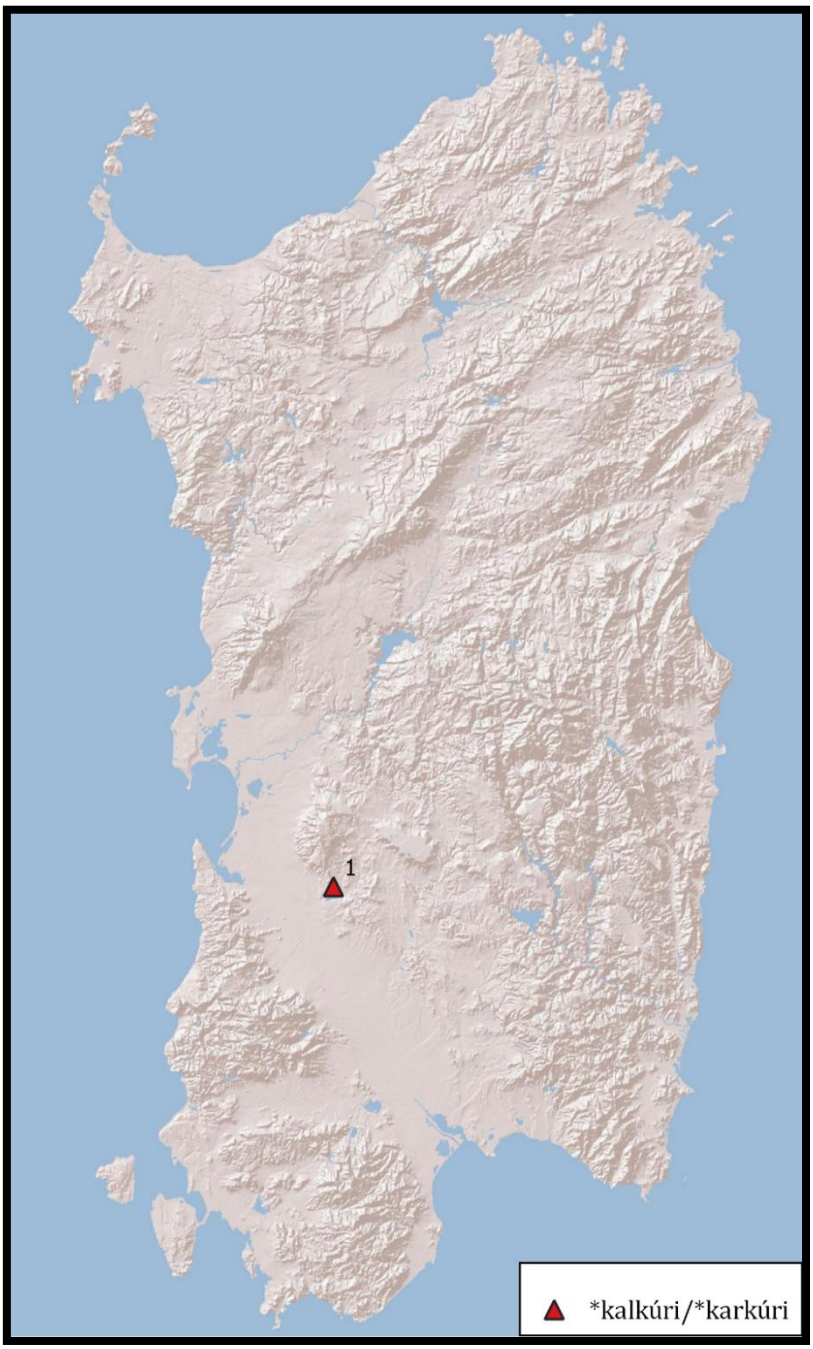

Figure 3.2.5: Ropegrass (*Ampelodesmos mauritanicus*)

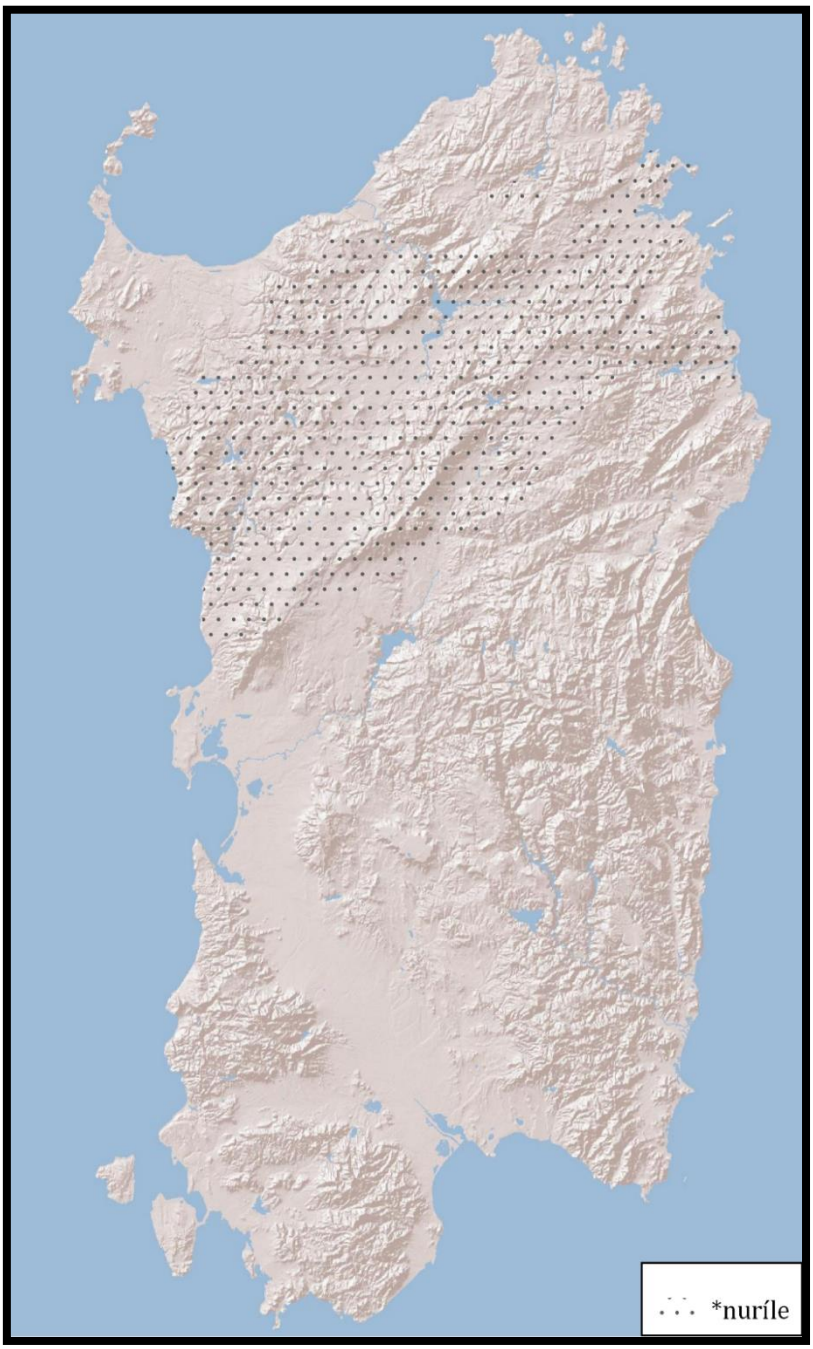

Figure 3.2.6: Weed

3.2.6 Weed (unspecified)

Reconstruction	**Forms**
***nuríle**	DES (II: 179): ***nuríle*** (Log.) 'weed that attaches to clods of earth that have not been cultivated for a long time'; [85] deriv.: ***annurilare*** 'hardening of the terrain, which develops crusts by lack of weeding'[86]

Wagner believes that these forms, recorded by Spano,[87] may be of pre-Roman origin. Paulis (1992: 386), however, suggests that the rather vague definition of the forms may refer to 'knotgrass (*Polygonum aviculare*)', and proposes this word to be a derivation of Lat. *nōdus* 'knot' with the suffix *-íle* < Lat. *-īlis* (cf. Wagner 1952: 40–41). While such a derivation would semantically fit 'knotgrass' very well, we cannot be certain that this word actually refers to 'knotgrass'. Moreover, it is difficult to explain the phonetic development of *-r-* < *-d-* that is needed to get *nuríle*. Paulis (1992: 386) provides parallels for *-r-* < *-d-*; e.g. *gurínu* 'basket, wooden bucket' < *kaðínu*; Camp. *pirai* 'to fart' < *piðai*. However, these examples show a different development, because their *-r-* < *-d-* is the reflex of *-t-* in Lat. *catinus* and *peditare* respectively (cf. DES I: 259; II: 264). In both Logudorese and Campidanese, Latin intervocalic voiced stops are regularly lost (Wagner 1941a: 79). This means that *-d-* in Lat. *nōdus* could not have developed into *-r-*.[88] This leaves *nuríle* without etymology, but in absence of a more specific definition it is impossible to provide an alternative etymology or to confirm its pre-Roman origin.

[85] "[E]rba parassita che si attacca alle zolle di terreno da molto tempo non coltivato".

[86] "[I]ndurirsi del terreno che mette le croste per mancanza di sarchiatura".

[87] In Spano's handwritten notes (*non vidi*) on his own dictionary (i.e. Spano 1851).

[88] Besides expected Log. *nóu* 'knot', Log. *nóðu* 'knot' is attested (DES II: 169). Its *-ð-* is attributed to influence from Italian however (DES II: 169), and there is no evidence for a secondary development of *-ð-* > *-r-* in this word.

3.3 Trees and shrubs

3.3.1 Bean trefoil (*Anagyris foetida*)

Reconstruction	Forms
***i̯òlva*/*i̯òlba*/ *i̯òrva*/*i̯òrba**	DES (I: 608): ***ǧòlva*, *ǧòrva*** (Log.) 'bean trefoil (*Anagyris foetida* L.)', ***ǵòlva*** (1: Bonorva) 'a herb'
	Hubschmid (1953: 85): ***giolva*** (Sassari)[89] 'bean trefoil' (*Anagyris foetida*)'
	Paulis (1992: 293): ***ǧòlva*, [*fáβa*] *ǧòlva*, *ǧòrva*** (Log.), ***fadzòrβa*** (2: Bolotana) 'bean trefoil (*Anagyris foetida*)'

The bean trefoil is a small tree native to the Mediterranean region, with bright yellow flowers and seeds that grow in long pods. Its leaves emit an unpleasant odor when crushed (Christian 2022). Its Logudorese names presuppose a reconstruction **i̯olva* (or **i̯olba/*i̯orva/*i̯orba*) (Hubschmid 1953: 85), which cannot be connected to any known Latin word. Hubschmid (1953: 85) notes that its formation looks similar to that of Lat. *malva* 'mallow' and other plant names they ascribe to the "Mediterranean substrate".[90] Paulis (1992: 293) instead proposes that the Sardinian forms go back to dialectal Modern Greek forms like ἀζώγυρος, ἀζώϋρο, ἀζώγουρας, ἀζώουρος etc. 'bean trefoil' (ΛΕΓ I: 295). As Paulis points out, the Sardinian forms would require a Greek form **azóuras*, with subsequent replacement of Gr. /z/ by *-ǧ-* and with metathesis of **-óura-* to **-óru̯a-* > *-ólva-*. Paulis' (1992: 293) etymology is supported by the existence of similar forms for 'bean trefoil' in southern Italy and Sicily going back to Gr. ἀζώγυρος: Bov. *źojero* etc., Calab. *źòjaru*, *źòjiru* etc., Sic. *azòiru*, *azòiuru*, *żżòira*, *żżòiru* etc. 'bean trefoil' (Rohlfs 1964: 15; Piccitto & Tropea 1977: 344; 2002: 1302).

[89] It is difficult to interpret this form. In the dialect of Sassari, we would expect ***ǵóybbu* (cf. *naybbúttsa* << Lat. *malva* 'mallow', *áybburu* < Lat. *arbor* 'tree'; AIS 642, 1251). It is possible that Hubschmid (1953) is referring to the province of Sassari, which contains the region where North Logudorese dialects are spoken. Because of this uncertainty, the form is not included in the map.

[90] See also Wigman (2023: 106–108).

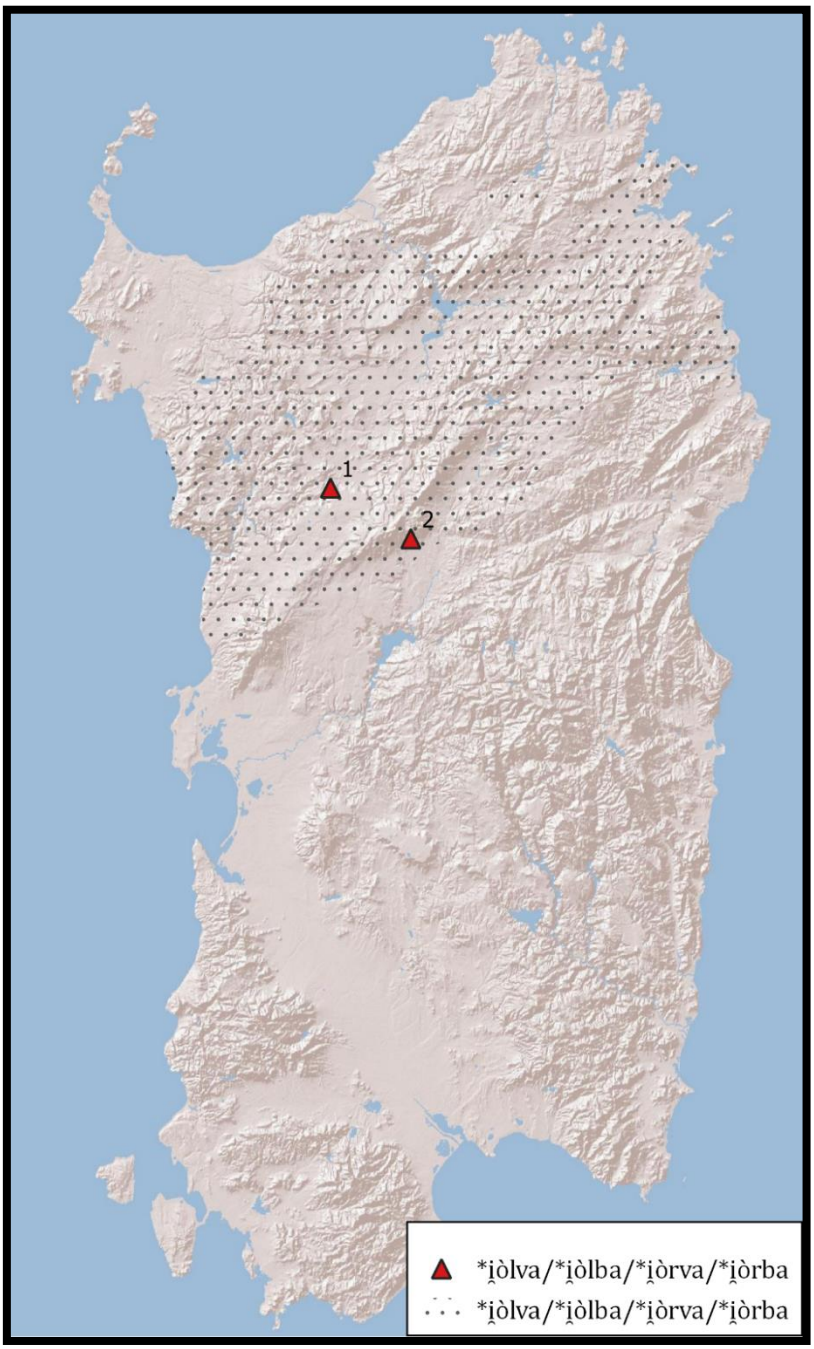

Figure 3.3.1: Bean trefoil (*Anagyris foetida*)

There are, however, some problems on the Greek side. It is unclear why classical Greek -υ- was borrowed into Sardinian as /u/ rather than /y/ or /i/. While in early Greek loans -υ- is sometimes represented as *-u-* (e.g. Srd. *túm*(*b*)*u* << Gr. θύμον 'thyme'), -υ- is usually borrowed as *-i-* in Sardinian (Wagner 1941a: 13). This is probably why Paulis points to ἀζώουρος from Kos, ἀζώγουρας from Egina and ἀζώγ̣ουρας from Megara as phonetically possible sources for the Sardinian forms. However, the Greek varieties of Egina and Megara (both of which are extinct) belong to a very small group of archaic dialects that regularly continue ancient Greek -υ- as /u/ rather than /i/ (Trudgill 2003: 54). It is not clear that the form of Byzantine Greek present in medieval Sardinia preserved /u/ too. Thus, although Paulis' (1992: 293) derivation of Log. *ǧòlva* from Greek is plausible, the exact source form remains to be clarified. A pre-Roman origin can in any case be discarded.

3.3.2 Broom, spiny (*Calicotome sp., Genista sp.*)

Reconstruction	**Forms**
1: *tería	DES (II: 548): ***tiría*** [***másu***] (1: Bonorva) 'Corsican broom (*Genista corsica* (Loisel.) DC.)', ***tería***, [***sa***] ***ðería*** (4: Baunei, 5: Jerzu, 6: Perdasdefogu) 'broom species (*Calicotome sp.* Link, *Cytisus sp.* Desf., *Genista sp.* L.)'
	Paulis (1992: 297–300): ***tiría*** (2: Bolotana), ***tiría*** [***búrda***] (7: Escalaplano) 'Corsican broom (*Genista corsica*)'
2: *ϑería	DES (II: 548): ***ϑiría*** (Nuor.), ***ϑería*** (3: Talana) 'broom species (*Calicotome sp., Cytisus sp., Genista sp.*)'
	Paulis (1992: 297–300): ***ϑiría*** [***ayrèste***] (Nuor.), 'bladder-senna (*Colutea arborescens* L.)'

All sources agree that the plant referred to as *ϑiría*, *tiría* or *tería* is particularly spiny. The DES (II: 548) further specifies that *ϑiría* etc. refers to a spiny species of broom, whereas *maϑrikúṩa* refers to species of broom without spines. I agree with Paulis (1992: 299) that the *ϑ-* in the Nuorese dialects must be secondary from **t-*, as we would expect Campidanese ***ts-* from original **ϑ-*.[91]

[91] Secondary introduction of *-ϑ-* instead of etymological *-t-* is rather common in the Nuorese dialects, potentially because of the correspondence between etymological Nuor. *-ϑ-* and Log. *-t-* (Wagner 1941a: 119).

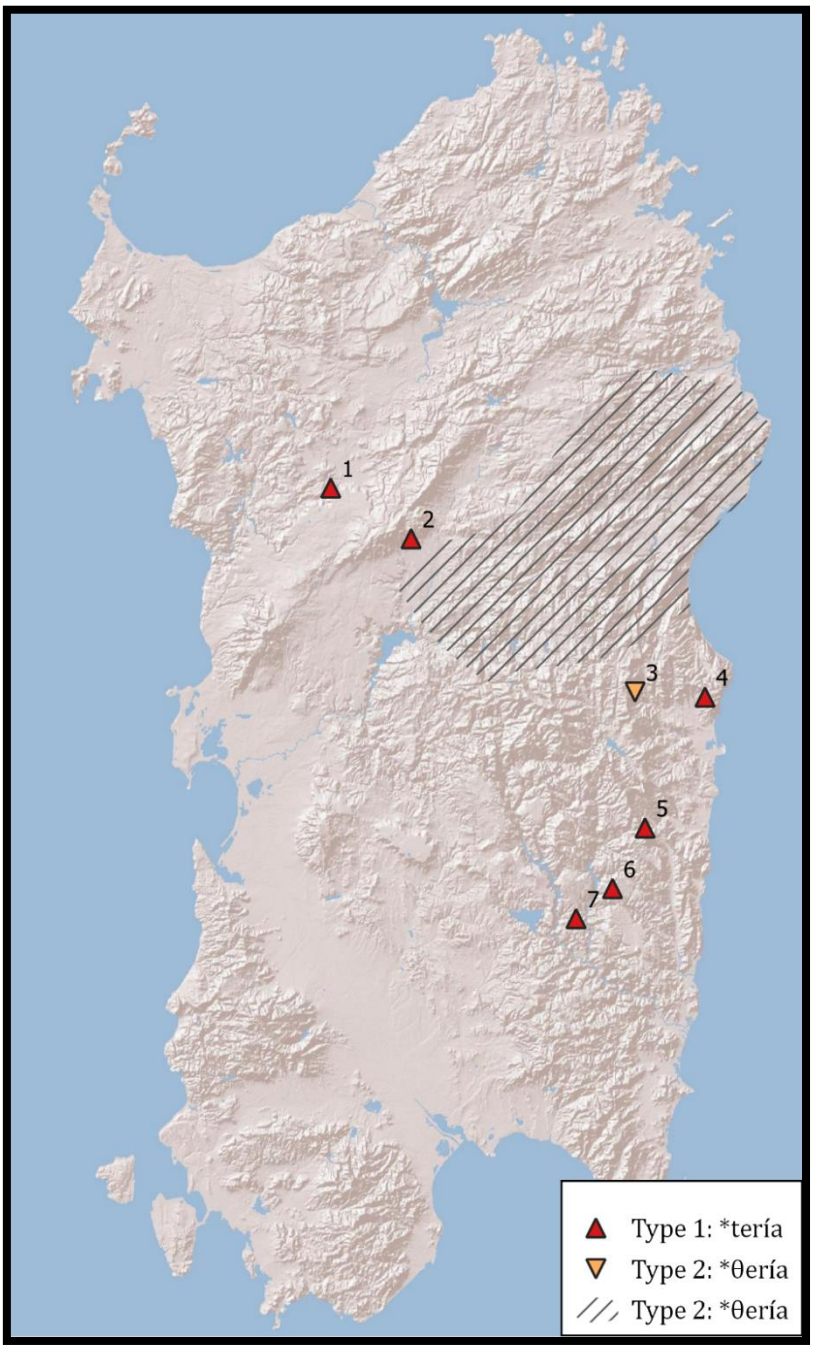

Figure 3.3.2: Spiny broom (*Calicotome sp.*, *Genista sp.*)

The analysis of *ϑ-* (which is common in words of suspected pre-Roman origin) as secondary, removes an important factor that may have led Wagner to designate this word as "probably pre-Roman". Paulis (1992: 299–300) analyzes Srd. **tiría*, **tería* as the outcome of Lat. *aetheria* 'ethereal, heavenly'. This word, in turn borrowed from Gr. αἰθέριος, is attested in Caelius Aurelianus (1757: 332): "the seed of the aethereal/heavenly herb that is called *eryngion*".[92] *Eryngion* refers to species in the plant genus *Eryngium* L. 'eryngo', which are similar to thistles or teasels. According to Paulis (1992: 300) eryngo would have been considered to keep away lightning and hailstorms (whence the name *aetheria* 'heavenly/aethereal'), and its name was later in Sardinian transferred to species of broom because of the superficial similarity between the plants caused by their spiny leaves. There are some factors that cast doubt on this etymology however. For one, the quote from Caelius Aurelianus does in my view not demonstrate that *aetheria herba* was the Latin name for eryngo, which is in fact called *eryngion* in the very same line. Secondly, we would need to assume a semantic shift from 'eryngo' to 'spiny broom' which is possible in theory, but not completely straightforward, since the similarity between both plants is based solely on the prickliness of their leaves.[93] Additionally, the penultimate stress in Srd. *ϑiría*, *tería* etc. does not correspond to the antepenultimate stress that Lat. *aetheria* would have had. Paulis (1992: 300) tries to resolve this by assuming the addition of the Greek suffix *-έα*, which was rendered in Sardinian as *-ía*. If we need to assume that the Sardinian forms go back to an unattested Gr. *αιθερέα, the passage from Caelius Aurelianus containing Lat. *aetheria* cannot be used as conclusive evidence, as the Sardinian word must in that case have been borrowed from Greek independently anyway.

Paulis' etymology is based on a single attestation of a Latin word in the context of a different plant than the Sardinian forms denote. In light of the uncertainties discussed above, I would like to consider another Greek word, already briefly discussed by Pittau (1995: 109), as a possible etymology for Srd. *ϑiría*, *tería* etc. 'spiny broom'. For the passage from Caelius Aurelianus mentioned above, Kühn (1828: 145) proposes a correction **atheria* for attested *aetheria*. Lat. **atheria* would be from Gr. ἀθερής, -ίδος 'prickly' (cf. Liddell et al. 1978: ἀθερής), which would semantically be more straightforward in the context of *eryngion* 'eryngo', which is a particularly spiny plant. Whether or not this proposed connection is

[92] "[A]etheriae herbae semine, quam *eryngion* appellant".

[93] It must be noted that the species *Eryngium maritimum* is also called *sea holly* in English, which may be a parallel for transfer of a phytonym on the basis of the prickliness of the leaves. Still, the shape of the leaves of sea holly is much more similar to actual holly than to the various species of broom.

correct for Caelius Aurelianus' *De Morbis Acutis et Chronicis*, Gr. ἀθερηίς 'prickly' corresponds well with Srd. *tería* 'spiny broom' semantically, since Wagner (DES II: 548), Paulis (1992: 298) and Puddu (2023: s.v. *tería*) all agree that Srd. *ϑiría*, *tería* etc. refer to species of broom that are particularly spiny. Phonologically, ἀθερηίς may explain the penultimate stress of the Sardinian forms. Note that the suffix *-ía* in Sardinian is mainly found on abstract and collective nouns, rendering Sardinian-internal addition of *-ía* to a plant name questionable (Wagner 1952: 14). The question is which variant of ἀθερηίς the Sardinian forms would reflect. For most of the Sardinian dialects, the non-neuter accusative and neuter plural ἀθερηίδα would probably yield the attested forms, with loss of intervocalic **-d-*. In the Nuorese dialects however, **-d-* would have been preserved, so *ϑiría* would have to be analyzed as a later loan from Campidanese or Logudorese. In light of the initial *ϑ-* that is most likely secondary as well (Paulis 1992: 299), this does not pose a real problem.

Whether it is Gr. ἀιθέριος 'heavenly/aethereal' >> *ἀιθερέα (as per Paulis 1992: 299–300) or Gr. ἀθερηίς 'prickly' that is the origin of Srd. *tería* etc., it seems that Greek θ- was borrowed as *t-*. This runs counter to the treatment of Greek θ- observed by Wagner (1941a: 109). Wagner cites Nuor. *ϑíu*, Camp. *tsíu* 'uncle' << Gr. θεῖος and Old Camp. *hanazzema* << Gr. ἀνάθημα, both of which suggest that Gr. -θ- was borrowed as Sardinian **-ϑ-* > Nuor. *-ϑ-*, Camp. *-ts-*, Log. *-t-*. A similar treatment to *tería* is however found in Nuor., Log. *timándza*, Camp. *timònğa* 'incense' << Gr. θυμίαμα 'id.' For this word, REW (8722) and Wagner (1997: 173) assume an intermediate Vulgar Latin form **thymania* (with **th-* realized as /t/). However, contrary to the reflexes of **thymania*, which is attested in multiple Romance languages, *ϑiría, tería* etc. is only attested in Sardinian, rendering the assumption of an unattested Latin form **at(h)ería* rather dubious.

Despite the many uncertainties, it remains attractive to posit a Latin **at(h)ería*, be it from Gr. ἀθερηίς 'prickly' or Gr. ἀιθέριος 'heavenly'. If both scenarios are to be rejected however, a (non-Greek) substrate origin for the name of these plant species would remain the only plausible option. Note however that this would purely be based on negative evidence, i.e. the absence of an alternative etymology, rather than positive evidence pointing to a pre-Roman origin.

3.3.3 Dog rose (*Rosa canina*)

Reconstruction	**Forms**
1: *kulári	DES (II: 195): [***orròsa***] ***ɣullári*** (1: Busachi) 'dog rose (*Rosa canina* L.)'
2: *ro/ulári	DES (II: 195): ***orrolárị̊u*** (2: Desulo, 6: Laconi), ***orrulárị̊u*** (3: Villagrande Strisaili, 7: Perdasdefogu), ***rullári*** (4: Aritzo), ***orrolári*** (5: Meana Sardo) 'id.'

Wagner does not provide an etymology for these words. Paulis (1992: 359) analyzes the forms in type 2 as a conflation of *rú*, *orrú* < Lat. *rubus* 'bramble' and an element **kulári* derived from *kúlu* 'arse' because of this plant's laxative properties (cf. e.g. Fr. *gratte-cul* 'dog rose'). There are some phonological issues with this etymology. One is Paulis' assumption that **kulári* was derived from *kúlu* 'arse' by means of the inherited suffix *-ale* < Lat. *-alis*. While it is true that *-ale* would likely dissimilate to *-are* due to the *-l-* in *kúlu* (Wagner 1952: 40), we expect the final vowel to yield ***-e* rather than attested *-i*. With the exception of Laconi and Perdasdefogu, all localities in which the words listed above are attested, preserve the distinction of word-final *-e* vs. *-i* and *-o* vs. *-u* (Wagner 1907b: Map 1). This, and the fact that we also find the ending *-ị̊u*, renders it doubtful that these words contain the adjectivizing suffix *-ale*, which in turn casts doubt on its derivation from *kúlu* 'arse'. If we analyze the forms of type 2 at face value, they require an original **rolári* or **rulári*. Paulis suggests that the velar of original **kulári* was lost in the same way as in the word *lú*(*o*) 'pergola of grapevines' < Lat. *lucus*, which occurs roughly in the same dialects as *orrolári* etc. (Wagner 1928: 35; DES II:47). Although the development of *lucus* > *lú* is not quite expected (we would rather expect ***lúɣu*)(Wagner 1941a: 69), it provides a possible parallel for hypothetical *orrolári* < **rú kulári*.[94] While there is no problem in identifying the first syllable of these forms as (*or*)*rú* 'bramble' — cf. *rúβu* [*berbeɣínu*] (Orgosolo) and *arrú* [*ɣráβĩu*] (Mogoro) 'dog rose' (DES II: 372) — the derivation of *kulári* < **cularis*, *-em* remains problematic.

[94] There is an ongoing process of further weakening of *-ɣ-* < **-k-* in Campidanese dialects, mainly in the adjectival suffix *-icus* and its derivations (Wagner 1941a: 75). This process has not been completed in all instances of original **-k-* though.

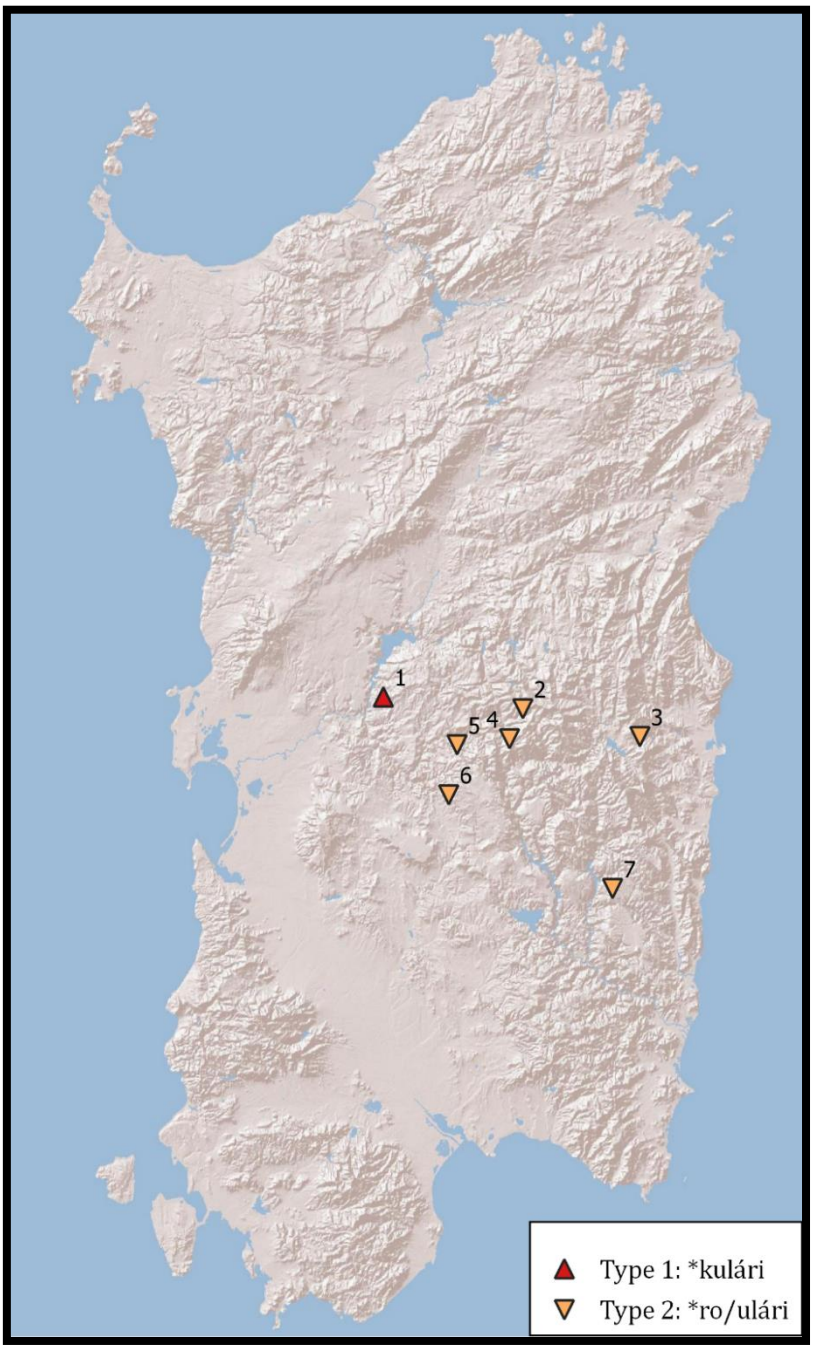

Figure 3.3.3: Dog rose (*Rosa canina*)

We thus face a dilemma between assuming **ro/ulári*, which is etymologically obscure, and assuming **rú kulári*, the second part of which would be morphologically obscure. In light of the fact that the expected outcome of the initial **k-* in **kulári* is found in only one variant (i.e. Busachi [*orròsa*] *ɣullári*), we may consider the possibility that this is in fact a folk-etymological back-formation from original **ro/ulári*, on the basis of *kúlu* 'arse', rather than an archaic formation. Irrespective of whether we reconstruct **ro/ulári* or **kulári*, it seems likely that this word does not belong to the inherited portion of the Sardinian lexicon, as Latin words with nominative in *-ĭs* are expected to be continued in Sardinian by its original accusative in *-em* > Srd. *-e* (cf. § 8.1.5.1). Since it is not obvious that this word was borrowed from another language in a more recent stage either, it is possible that it originally pertained to a pre-Roman language.

3.3.4 Heather (*Erica scoparia*)[1]

Reconstruction	**Forms**
1: *fla/umíu	DES (I: 539): ***framíu*** (Ogl.) 'bundle of dry twigs for the oven', ***frumíu*** (1: Paulilatino) 'small wild olive tree (*Olea europaea* L.)'
2: *flámmi̯u	***frámi̯u*** (Ogl.), ***frámmi̯u*** (2: Tonara) 'heather (*Erica scoparia* L.)'

Wagner (DES I: 539) believes that these words are pre-Roman in origin, with possible influence of *frámma* 'flame' < Lat. *flamma* in the form *framíu* 'bundle of dry twigs for the oven'. He compares the forms referring to 'heather' to Sp. *firma* (Salamanca) 'heather (branch)'. Paulis (1992: 415) does not agree with a pre-Roman origin for these forms, and instead derives all of them from Lat. *flammeus* 'flaming, fiery'. I agree that this is attractive at least for the forms in type 2, which are in fact the regular outcome of Lat. *flammeus*. Semantically this derivation is probable because of heather's qualities as firewood. For the forms in type 1, proposing a direct continuation from *flammeus* is more difficult. Accented *-íu* in *framíu* 'bundle of dry twigs for the oven' and in *frumíu* 'small wild olive tree' cannot go back to Lat. *-́eus*.

As such, the forms in type 1 need a different explanation. If *framíu* and *frumíu* are the same formation, either the *-a-* or the *-u-* is secondary. Whereas pretonic *-a-* > *-u-* occurs occasionally in the presence of a labial consonant (cf. Wagner 1941a: 28–29), *-u-* >> *-a-* by influence of *frámma* 'flame' is in this case also very well imaginable for *framíu* 'bundle of dry twigs for the oven' (as already suggested in DES I: 539), making it difficult to judge which one is original.

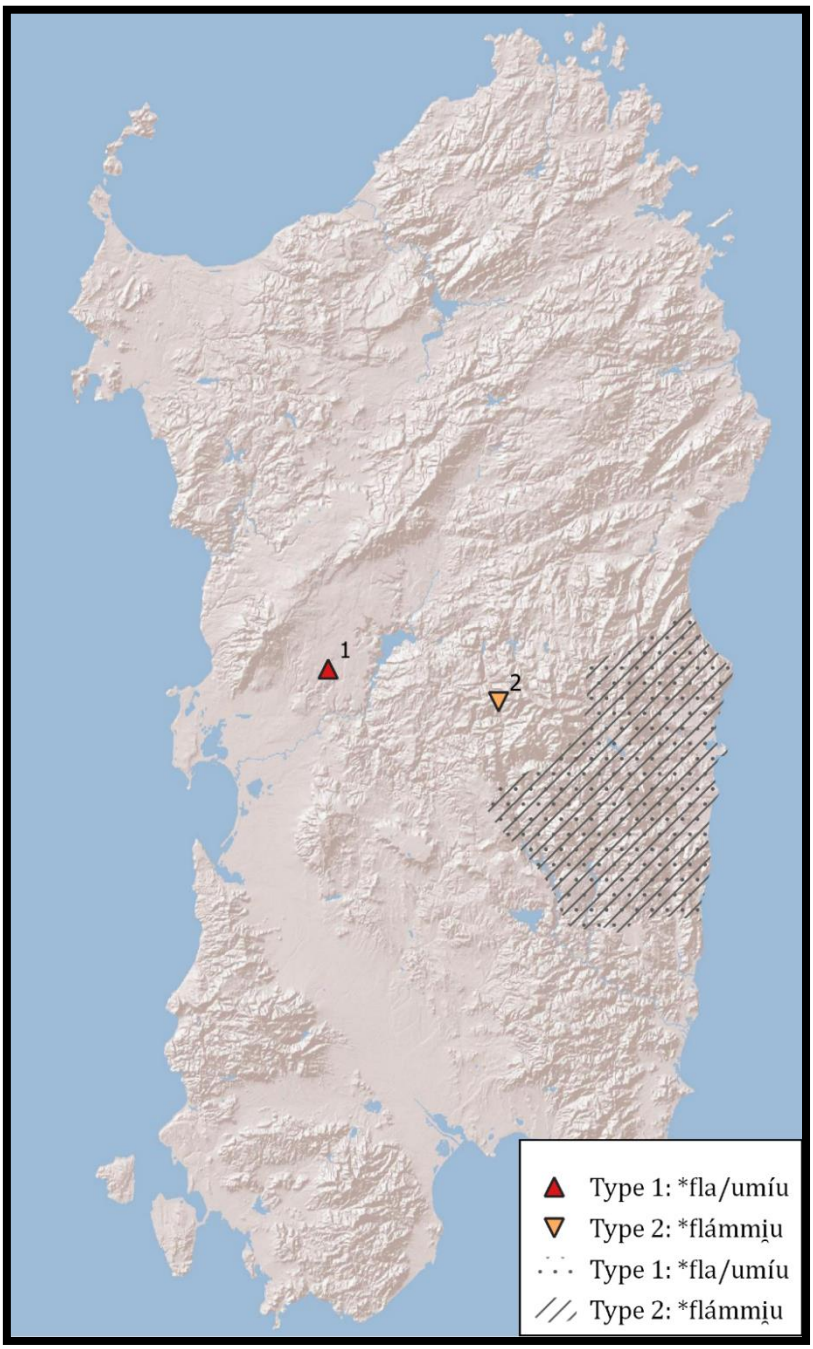

Figure 3.3.4: Heather (*Erica scoparia*)

Whatever the original vowel was, we are left with a form **fra/umíu*,[95] that refers to types of plant growth typical of Mediterranean vegetation, and which cannot be derived from Lat. *flammeus* for formal and semantic reasons. Wagner's (DES I: 539) analysis of this group of words as pre-Roman thus remains a possibility. If we assume original **fra/umíu* 'heather/some maquis shrub', remodeling after *frámma* 'flame' may have led to the creation of *frám*(*m*)*i̯u* 'heather' (type 2), due to heather's inflammable properties. The semantic development to 'small wild olive tree' and to 'bundle of dry twigs for the oven' can both best be understood from 'heather' as the original meaning; the first by association to olives' and heather's shared habitat and the second again because of the common use of heather as firewood. Beside the lack of a convincing etymology there are no real diagnostic features for or against a pre-Roman origin of **fra/umíu*, as the vowel alternation can be explained as secondary. Wagner's (DES I: 539) comparison to Sp. *firma* (Salamanca) 'heather (branch)' could maybe work if we assume the Sardinian forms to go back to an original **fa/urmíu*, with metathesis. Perhaps it is no more than a chance correspondence however.

3.3.5 Heather (*Erica scoparia*)$_2$, tree heather (*Erica arborea*)$_1$

Reconstruction	**Forms**
1: *túfVr-	DES (II: 540): ***túora*** (1: Dorgali), ***túfera*** (2: Urzulei), ***túv(v)era*** (4: Baunei, 5: Desulo, 6: Villagrande Strisaili), ***túvara*** (7: Seulo, 9: Seui, 10: Sadali, 12: Perdasdefogu, 13: Escalaplano) 'heather (*Erica scoparia* L.), tree heather (*Erica aborea* L.)'
	AIS (617): [***sa***] ***ðűvera*** (4: Baunei, 5: Desulo, 8: Laconi), [***sa***] ***túvera*** (11: Mogoro), [***sa***] ***ðűvara*** (10: Perdasdefogu, 11: Escalaplano) 'heather (*Erica scoparia*)'
2: *úfVr-	DES (II: 540): ***úvvara*** (3: Busachi, 14: Guspini, 15: Samassi, 16: Villacidro, 17: Villaputzu, 18: Narcao) 'heather (*Erica scoparia*), tree heather (*Erica aborea*)'

[95] This reconstructed form **fra/umíu* could in turn go back to earlier **fla/umíu*, **fra/umíu*, **fa/ulmíu* or **fa/urmíu*. The distinction between postconsonantal *-l-* and *-r-* is only preserved in Baunei, for which this specific word is unfortunately not attested (Wagner 1941a: 154–155).

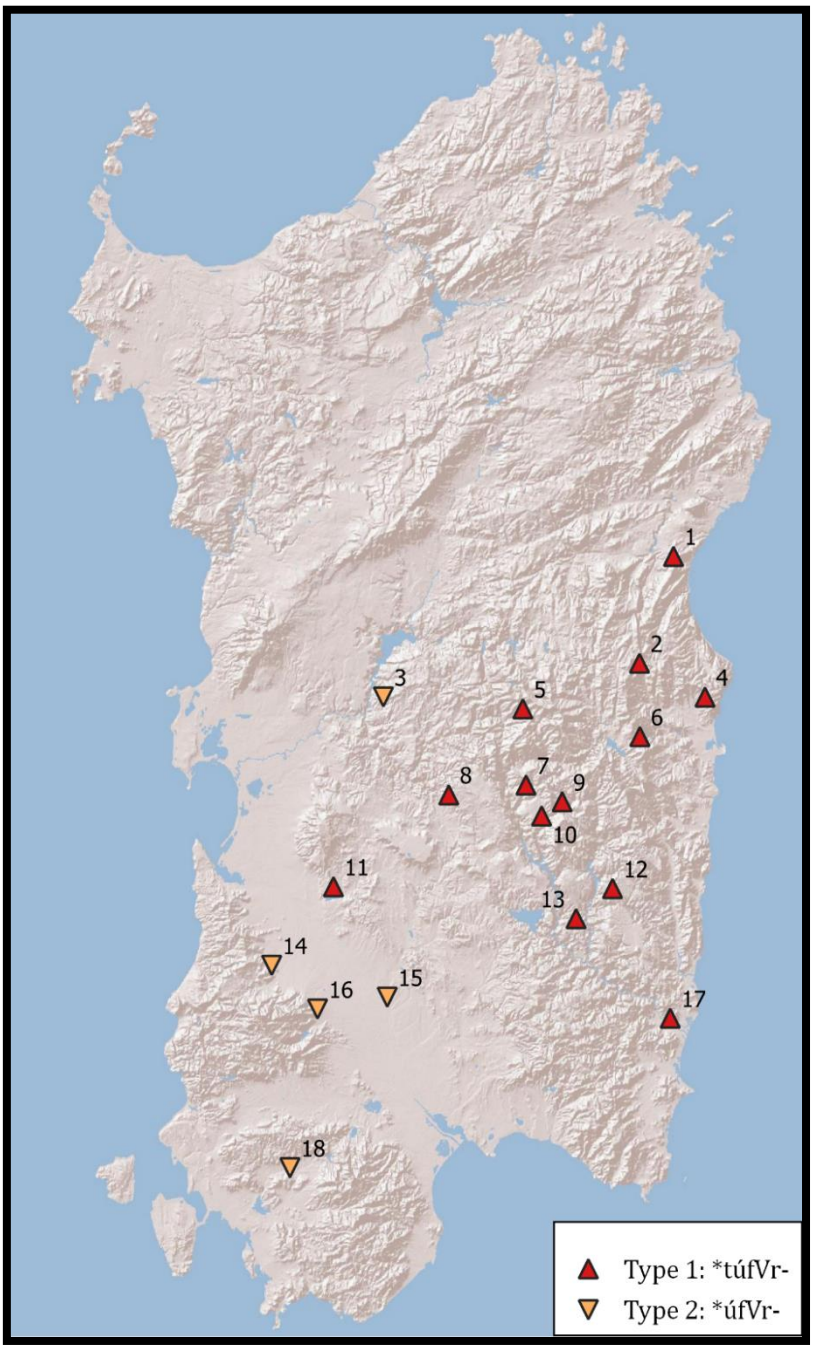

Figure 3.3.5: Heather (*Erica scoparia*), tree heather (*Erica arborea*)

Wagner (DES II: 540) rightly separates these words from Srd. *túvara* 'species of truffle' < Lat. **tūfer/tūber* on semantic grounds. He mentions Bertoldi's (1934: 77 fn. 1) observation that Lat. *tūber* 'truffle' exists next to *tuber* 'species of medlar', but neither Bertoldi nor Wagner explicitly state whether they think that Lat. *tuber* 'medlar' is the origin of Srd. *túvara* etc. 'heather'. I agree in any case with Paulis (1992: 415) that it cannot be, since we would expect the *-b-* of *tuber* to be lost in all Logudorese and Campidanese dialects, and because of the rather great semantic distance between 'medlar' and 'heather'. Paulis instead proposes to connect the Sardinian forms listed above to Log. *túvu* 'thicket', which according to Wagner is from Lat. *tūfa* 'shrub, tuft of fronds'. According to Paulis (1992: 415), the originally pre-Roman suffix **-ár-* (§ 9.2.6.2), with its putative collective meaning, was added within Sardinian to Log. *túvu* 'thicket' << Lat. *tūfa*. A remaining problem is the fact that the forms in type 2 (*úvvara*) lack the initial *t-* found in type 1, which is difficult to explain formally if Lat. *tūfa* is indeed the source of all these forms. I have not been able to provide a plausible solution to this issue.

3.3.6 Holly (*Ilex aquifolium*)$_1$

Reconstruction	**Forms**
1: *alaθúli	DES (I: 67): ***alaθúli*** (1: Urzulei) 'holly (*Ilex aquifolium* L.)/butcher's broom (*Ruscus aculeatus* L.)'
	Blasco Ferrer (2019: 40): ***balatúle*** (1: Urzulei) 'holly (*Ilex aquifolium*)'
	Puddu (2023: s.v. *alatúli*): ***alatúli*** 'laurel (*Laurus sp.* L.)'
2: *laθòre	Blasco Ferrer (2019: 40): ***latθòri*** (2: Villagrande Strisaili) 'holly (*Ilex aquifolium*)'

Wagner (DES I: 67) notes that Pellis (ALIT 4076)[96] has registered this word in Urzulei as 'laurel', which is also the meaning found in Puddu (2023: s.v. *alatúli*). However, Wagner suggests that it might rather refer to 'butcher's broom', which is at times called *láu spinósu* ("spiny laurel") in Sardinian. Paulis (1992: 417) on the other hand thinks that the translation 'laurel' is mistaken for 'holly' instead of 'butcher's broom'. The meaning 'holly' also matches that of the form *latθòri* recorded by Blasco Ferrer (2019: 40) in Urzulei. Wagner, Paulis, and Blasco Ferrer agree that this word is likely of pre-Roman origin, presumably based on its peculiar shape and presence of **-θ-*.

[96] Non vidi.

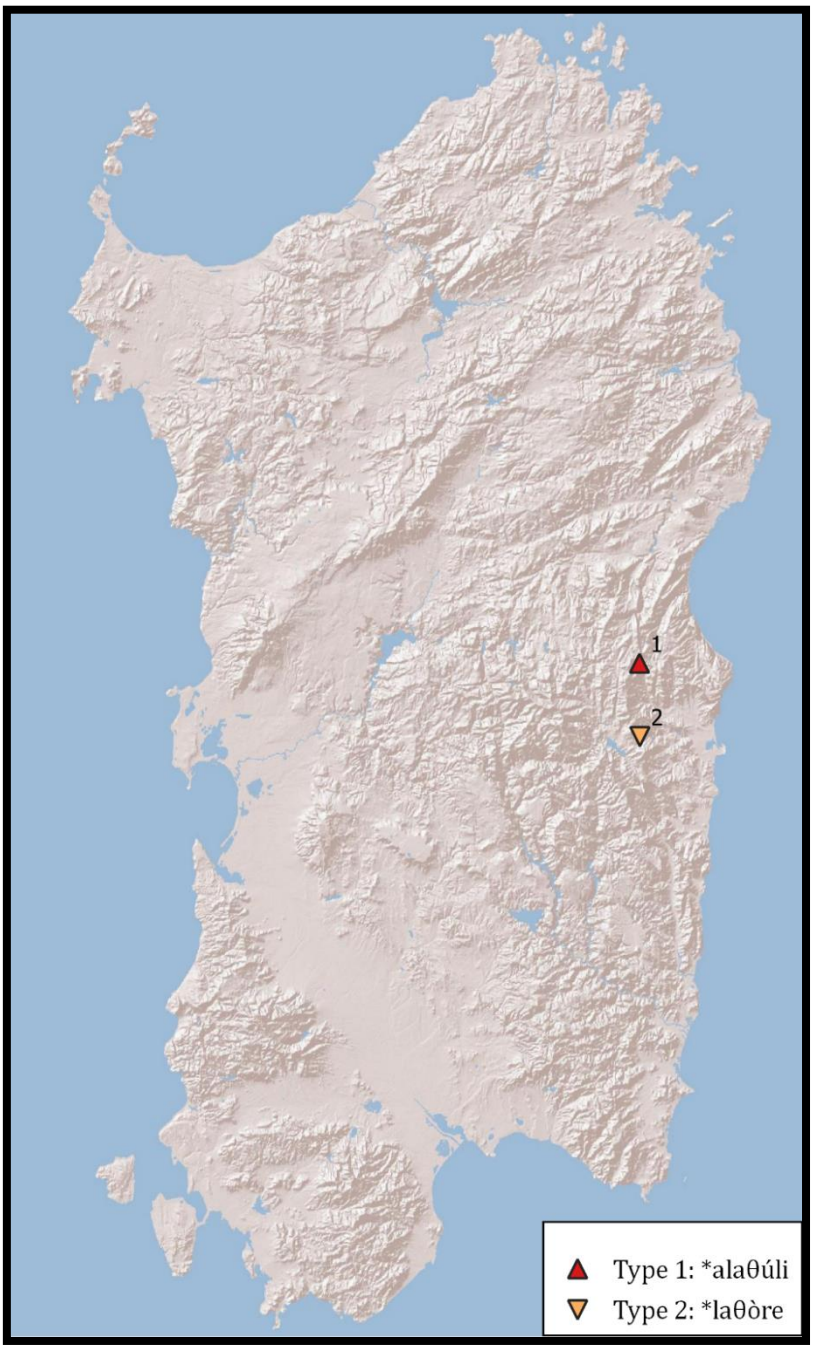

Figure 3.3.6: Holly (*Ilex aquifolium*)

Blasco Ferrer (2019: 40) proposes *alaϑúli* and *latϑòri* to be derived from a base *(*a*)*laϑ*- by means of two different suffixes (**-úle* and **òro/i*). It is possible however that the ending *-òri* is due to influence from the inherited agentive suffix *-òre/i*.[97] Both Wagner (DES I: 67) and Paulis (1992: 417) discuss *alaϑúli* in the context of *aláṡe* etc. 'butcher's broom; holly' (§ 3.2.1), without specifying how these forms would be related. I do not know of any other instance of a word ending in *–(ϑ)úli*, so a connection between *aláṡe* and *alaϑúli* is uncertain.

Similar in shape to this word is Gr. ὀνοθοῦρις 'oleander (*Nerium oleander* L.)' (Dioscorides Anazarbei 1907: 267; Dioscorides 2005: 295), attested as *onotirum* in the Latin translation of *De materia medica* (Stadler 1901: 56). The Latin form cannot have yielded the Sardinian forms regularly, so we would need to assume a Byzantine loan with subsequent deformations. Both holly and oleanders are designated as a kind of laurel in various languages: e.g. Srd. *láu spinósu* 'holly' (Paulis 1992: 417); Fr. *laurier-rose*, Sp. *laurel de flor* 'oleander'. On the basis of only two forms this connection is speculative however, and a pre-Roman origin remains a possibility.

3.3.7 Holly (*Ilex aquifolium*)$_2$

Reconstruction	**Forms**
1: *golóstru	DES (I: 584): ***olóstru*** (1: Oschiri, 2: Mores, 4: Giave, 5: Bonorva, 11: Borore) 'holly (*Ilex aquifolium* L.)'
	Paulis (1992: 362): ***kolóstru*** (10: Orani) 'dog rose (*Rosa canina* L.)'
2: *golósti	DES (I: 584): ***golóstḭu*** (3: Bitti, 7: Orosei), ***olóstḭu*** (6: Bono), ***golóstri*** (8: Nuoro), ***olósti*** (9: Dorgali), ***olóstri*** (15: Fonni, 16: Mogoro) 'holly (*Ilex aquifolium*)'
3: *gósti	DES (I: 584): ***kóstḭu*** (10: Orani) 'holly (*Ilex aquifolium*)'
4: *golóstike	DES (I: 584): ***ʔolóstiʔe*** (12: Orgosolo), ***golóstiʔe*** (13: Olzai), ***olóstriɣe*** (14: Santu Lussurgiu) 'holly (*Ilex aquifolium*)'

[97] Even though Villagrande Strisaili usually preserves word-final *-e* (cf. Wagner 1907b: Map 1), the open *-ò-* shows that *-òri* must go back to **-òre*.

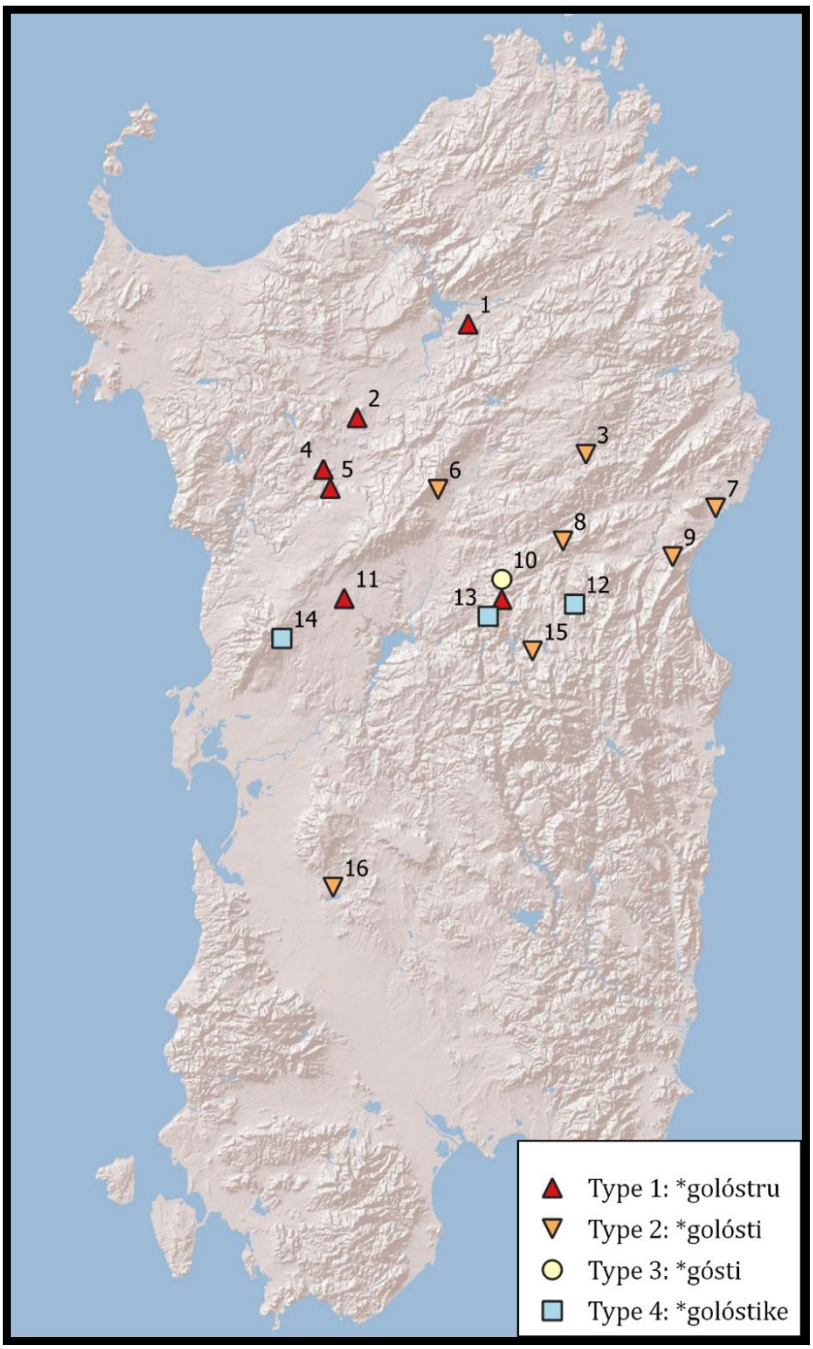

Figure 3.3.7: Holly (*Ilex aquifolium*)

All Sardinian forms of this word can be derived from an original **golósti*. The variations in the onset between *k-*, *g-* and *Ø-* are due to reinterpretations of post-vocalic forms as pausa forms and vice versa (cf. Wagner 1941a: 205–213). Type 3 *kóstị̊u* (10: Orani) shows irregular loss of intervocalic *-l-*; perhaps it was influenced by Nuor. *kòstike*, Log. *kòsti* 'maple' (§ 3.3.9). In all types, we find various adaptations of the final syllable, all of which can be understood from original **-i*, which is rare as an ending of inherited nouns (§ 8.1.5.1). In type 1, **-i* was replaced by the masculine suffix *-u*. In types 2 and 3, it was preserved in some cases and extended with *-u* in others.[98] In type 4 we find an additional suffix **-(i)ke*, which is again reminiscent of *kòstike* 'maple' (§ 3.3.9). All Logudorese forms for 'holly', as well as Nuoro *golóstri* and Orani *kolóstru*, have introduced an *-r-* in the final syllable, which may be due to analogy to the suffix *-ástru* which is present in various 'wild' plants: e.g. *mentástru* 'wild mint' << *mènta* 'mint'; *odzástru* 'wild olive' < Lat. *oleástru*, *pirástru* 'wild pear' << *píra* 'pear' (Wagner 1952: 120; DES II: 200, 274). The ending *-ost(r)-* in this word is perhaps the same as the one found in *giḍḍòstre* etc. 'tree heather' (§ 3.3.20, § 9.2.9).

This word is certainly of pre-Roman origin. Bertoldi (1929: 261 fn. 3) noted its perfect correspondence to Bq. *gorosti* 'holly' < **golosti*,[99] and Hofmann (1949: 142) adduced Gr. κήλαστρος 'id' (cf. Theophrastus 1916: chap. 9:3).[100] Jud (1931: 439) connects Lang. *agaloussé*, *agaloussës* 'holly' < **agalústo-* (cf. FEW XXI: 106 for many additional forms). Hubschmid (1953: 29) further compares Navarr. *colostia*, *golostia*, *corostia*, *gorostia*. as well as several Celtic and Germanic forms — e.g. Bret. *kelenn* 'holly', OIr. *cuilenn* 'id.', and OHG *hulis* 'id.' — which must go back to PC **kolinno-* < **kolis-n-* and PGm. **hulisa-* respectively (Kroonen 2013: 253; van Sluis, Jørgensen & Kroonen 2023: 200).

There is evidence for the existence of a (regional) Latin word **celastru-*, formally compatible to Gr. κήλαστρος, continued in Lucanian *cèlastru*, Calab. *cègghiastru*, *cìstriddazzu*, Bov. *azzìddastro* 'holly' (DEI II: 842; Rohlfs 1977: 104, 110, 156, 694, 808).[101] Lat. **celastru-* is considered to be a loan from Gr. κήλαστρος by DEI (II: 842), but it should be noted that the provenance of the Greek form is not well-

[98] Note that the ending *-i̯u* cannot be original, as **golóstị̊u* would have yielded ***golóθθu* vel sim.

[99] Trask (1997: 307) notes that Bq. *gorosti* could have **goloztri* as a pre-form, in which case the presence of *-r-* in the Sardinian forms might be original too.

[100] Hesychius also has κηλάστραι 'vessels; or trees' (Hesychius Alexandrinus 2020: 596).

[101] Perhaps Bret. *kelastrenn* 'flexible (holly) branch' derives from Lat. **celastrum* too (cf. Deshayes 2003: 380).

understood. Gr. κήλαστρος 'holly' is only attested in Theophrastus (1916: chap. 9:3; Hofmann 1949: 142), and in Hesychius (2020: 596) as κήλαστραι 'trees'. The evidence for Gr. κήλαστρος being a plausible source for Lat. **celastru-* and its descendants is therefore dubious. On the basis of the scarce Greek attestations of this form, it cannot be excluded that Gr. κήλαστρος belonged to the Greek varieties of Magna Graecia, which would explain the southern Italian attestations of this word. At any rate, Lat. **celastru-* cannot account for the vocalism of the Sardinian, Basque, Navarrese and Lengadocian forms. We are thus dealing with multiple independent loans of this pre-Roman word into the languages of the Mediterranean.

The Celtic and Germanic forms both appear to contain an element **kulis-*, which is indeed similar to the sequence **golost-/galost/kēlast-* required for Sardinian, Basque, Occitan and Greek, but it is somewhat more removed phonologically and morphologically. The potentially European-wide distribution of this group of words has interesting implications for our understanding of pre-Indo-European Europe (§ 11.2.1).

3.3.8 Hop-hornbeam (*Ostrya carpinifolia*)

Reconstruction	**Forms**
****a(C)urri*/ ****áruli***	DES (I: 153): ***aúrri*** (Barb.) 'hop-hornbeam (*Ostrya carpinifolia* Scop.)'
	Puddu (2023: s.v. *àrrui*, *àurri*): ***àrrui***, ***aúrri*** 'hop-hornbeam (*Ostrya carpinifolia*)', ***àurri*** 'hornbeam (*Carpinus betulus* L.)'[102]

Between the attestations in DES (I: 153) and Puddu (2023), we find various accent shifts and vowel metatheses. Inherited *-au-* was monophthongized to *-a-* in Sardinian (Wagner 1941a: 13–14). These forms must therefore go back either to **aCurri*, with an elided consonant, or perhaps to **áruli* (cf. Srd. *meúrra* < Lat. *mérula* etc.; Wagner 1941a: 5). Wagner (1932: 231) compares it to Bq. *aurri* 'plant with long roots and wide leaves' (De Azkue 1905: 109) and to Bq. *aurrede* 'osier bed, willow plantation' (De Azkue 1905: 108). However, in EH (75), Bq. *aurri* does not occur in the meaning given by De Azkue, but rather means 'ruins', which would significantly weaken its comparison to Sardinian.

[102] The hornbeam (*Carpinus betulus*), does not occur in Sardinia (Sikkema, Caudullo & de Rigo 2016), so this form probably also refers to the 'hop-hornbeam (*Ostrya carpinifolia*)'.

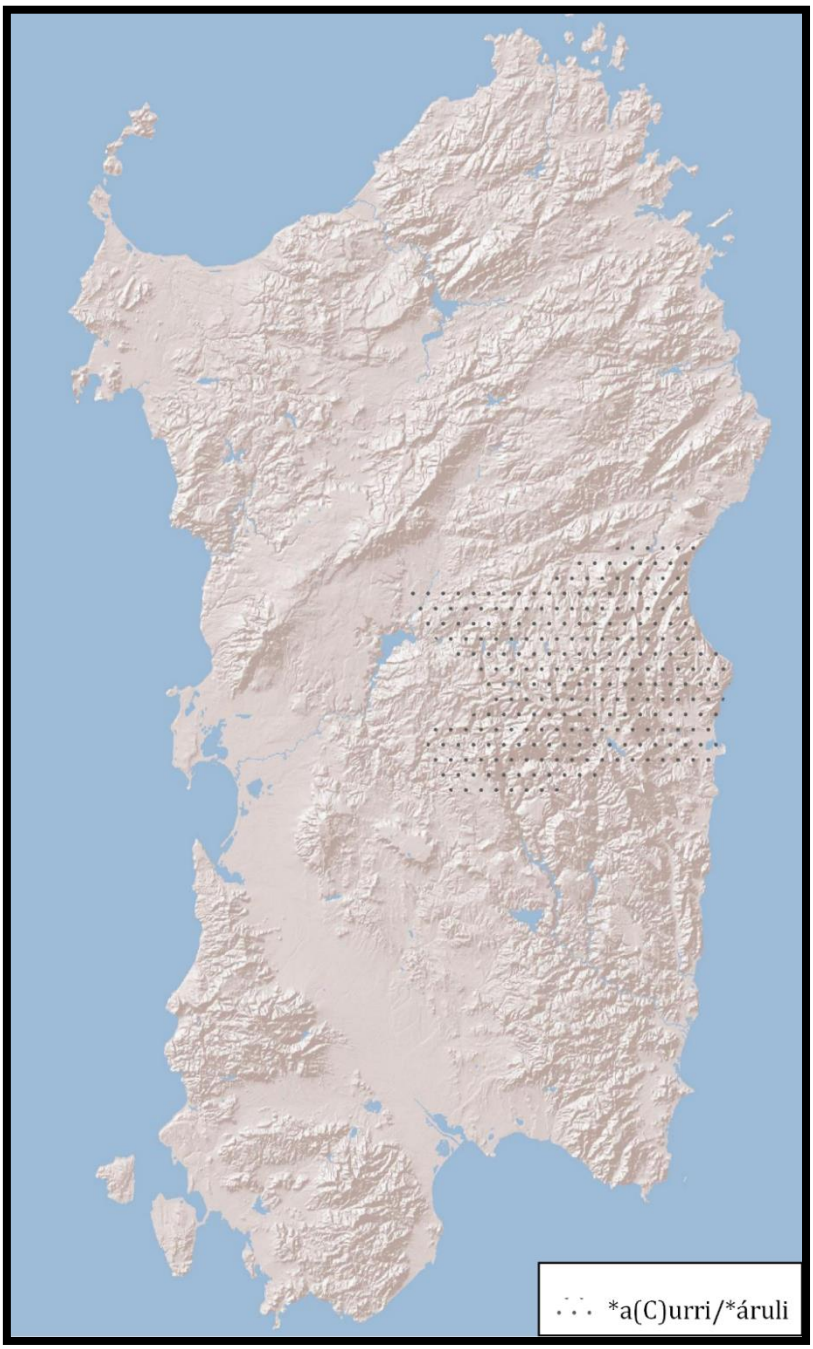

Figure 3.3.8: Hop-hornbeam (*Ostrya carpinifolia*)

Furthermore, Michelena (1958: 405–406) points out that Bq. *aurrede* is likely from Gasc. *aubredé* 'id.' (FEW XXIV: 295), leaving Bq. *aurri* isolated. I agree with Bouda (1949: 334) and Michelena (1958: 406) that the uncertain attestation and vague meaning of the Basque word makes the comparison to Srd. *aúrri* 'hop-hornbeam' rather difficult (cf. Michelena 1959: 524), even more so because **aúrri* cannot be the original shape of the Sardinian forms. Even though a pre-Roman origin for the Sardinian word is plausible on the basis of its peculiar structure, its connection to Bq. *aurri* 'plant with long root and wide leaves' cannot be upheld.

3.3.9 Maple (*Acer monspessulanum*)

Reconstruction	**Forms**
1: *kósti	DES (I: 392–393): **kòsti** (Log.) 'Montpellier maple (*Acer monspessulanum* L.)'
	Spano (*apud* DES I: 393):[103] **kòsti** 'tree that has hard wood and clover-like leaves'
	Hubschmid (1953: 81): **costi** (2: Desulo) 'solitary and colossal box tree on the flanks of the Gennargentu-mountains (*Buxus sempervirens* L.)'
	Blasco Ferrer (2019: 40): **kóstiu** (1: Orani), **kósti** (3: Villagrande Strisaili)
2: *kóstike	DES (I: 392–393): **kòstike** (Nuor.), **kòstiɣe** (Log.) 'Montpellier maple (*Acer monspessulanum*)'
	Blasco Ferrer (2019: 40): **kóstike** (3: Villagrande Strisaili)

According to DES (I: 393), all forms can be explained from an original **kóstike*, with secondary loss of the second *-k-* in *kòsti* (cf. Wagner 1941a: 76). However, a suffix **-(i)ke* must be assumed for *ʔolóstiʔe*, *golóstiɣe* 'holly' (§ 3.3.7), so perhaps *kòstike* may be analyzed as **kòsti-ke*/**kòst-ike* (Blasco Ferrer 2019: 40). In some cases, *kòsti* also refers to the box tree (*Buxus sempervirens*). The open *-ò-* of these words is surprising, as it is expected to be raised to *-ó-* due to the *-i-* in the following syllable. Blasco Ferrer's (2019: 40) attestations do have closed *-ó-*, suggesting that *-ò-*in the DES may be mistaken. This word, which does not have a Latin etymology, has long been considered of pre-Roman origin, and has been compared to various phytonyms in other Mediterranean languages.

[103] In Spano's handwritten notes (*non vidi*) on his own dictionary (i.e. Spano 1851).

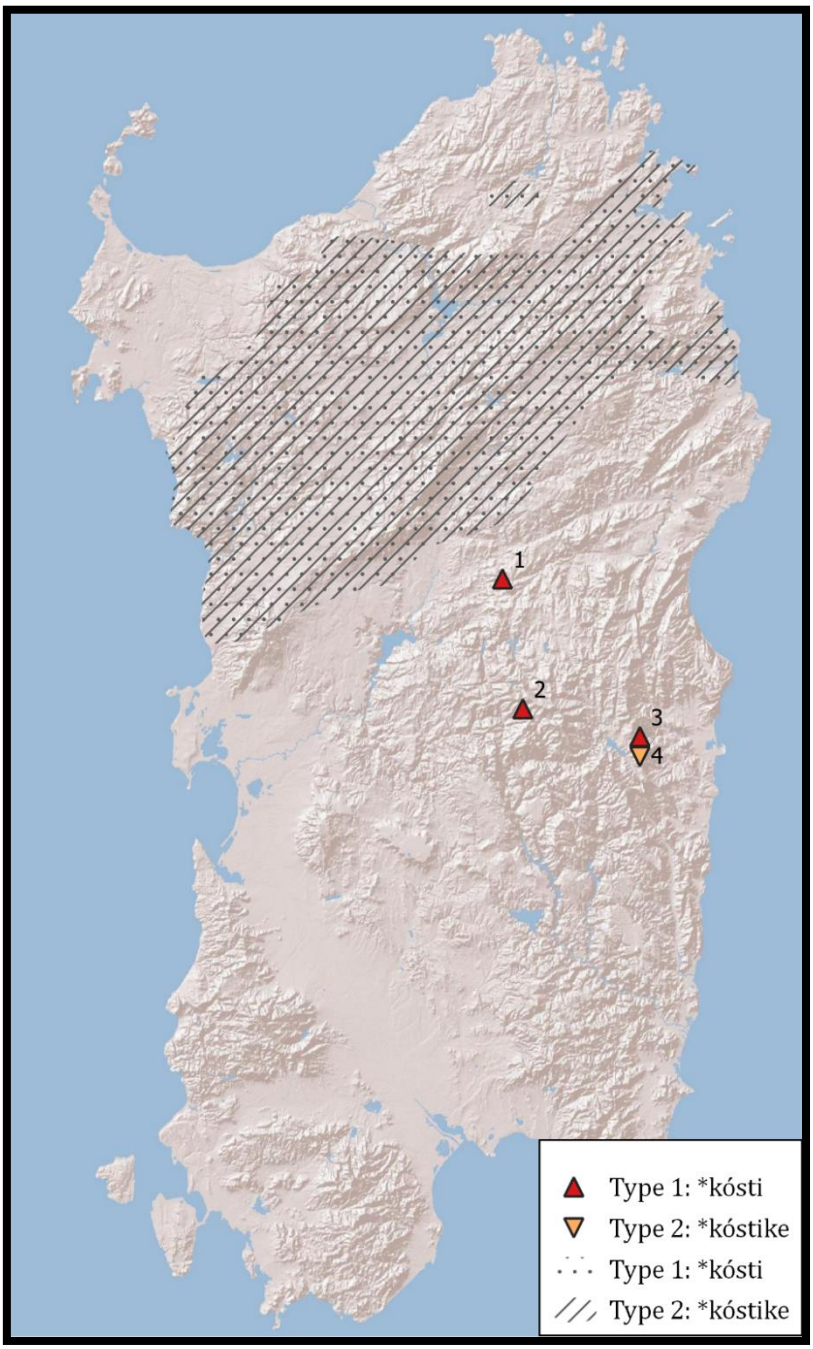

Figure 3.3.9: Maple (*Acer monspessulanum*)

Bertoldi (1931: 131) compares Bq. *gastigar, gaztigar astigar* 'maple'. This word is probably a compound of *(g)azt-* and *ihar* 'maple' (Michelena 1953: 48). The element *(g)azt-* is not attested on its own, and Michelena (1959: 525; 1990: 253 fn. 49) wonders whether its similarity to the Sardinian forms is not due to chance. If the connection between Bq. *(g)azt-* and Srd. *kòstike* 'maple' is correct however, this would be evidence in favor of the ending *-ik-* being a suffix (§ 9.2.2.2) (cf. Hubschmid 1953: 80–81).[104] Alessio (1949: 132), on the other hand, compares Tusc. *còstolo* 'ash tree (*Fraxinus ornus* L.; *Fraxinus excelsior* L.) and Gr. Hes. ἄκαστος 'maple' (Hesychius Alexandrinus 2018: 109). Alessio hypothesizes that this gloss is originally a Massiliote Greek loan from Ligurian. This is possible, though the only evidence in favor of this is the existence of OProv. *agast* 'maple' and comparable forms in other Occitan dialects (FEW XXIV: 276). There is no mention in Hesychius that this was a Massiliote form. Hubschmid (1953: 82) argues that ἄκαστος consists of a prefixed *a-* and a root **kast-*, which he thinks is also present in Greek κάστανον 'chestnut'.

In my view, a connection between Gr. ἄκαστος 'maple' to Srd. *kòstike* etc. 'id.' and possibly Bq. *(g)azt-* is formally more convincing than the often proposed relation between Gr. ἄκαστος and Lat. *acer* 'maple' (cf. Hubschmid 1953: 82 fn. 3; Wigman 2023: 213). Tusc. *còstolo* 'ash tree' is semantically slightly further removed, but in light of the Sardinian forms referring to both maples and box trees, this need not be a problem. If these connections are correct (and even if Greek ἄκαστος is of Massiliote origin), this word shows an interesting distribution across the western Mediterranean that is found in various other words as well (§ 11.3).

3.3.10 Osier (*Salix viminalis*)

Reconstruction	**Forms**
****ϑò(C)a***	DES (II: 551): ***tòa*** (Log.), ***tsòa*** (Camp.), ***attòa*** (1: Monti, 5: Bonorva), ***ϑòa*** (2: Siniscola, 6: Orosei), ***ϑòga*** (4: Bitti, 8: Orani), ***ϑòβa*** (7: Nuoro), ***šòva*** (9: Fonni) 'osier (*Salix viminalis* L.)'
	Jud (1942: 352): ***tòa*** (3: Alghero) 'id.'
	Puddu (2023: s.v. *atòa, sciòva, tzòa*): ***atòa, atzòa, tòa, tzòa, tzòba, tzòga, tzòva*** 'common sallow (*Salix cinerea* L.)'; ***sciòva*** 'osier (*Salix viminalis*)'

[104] But note that, following Michelena's (1953: 48) analysis, this suffix is probably not present in the Basque form, as Hubschmid (1953: 80) proposes.

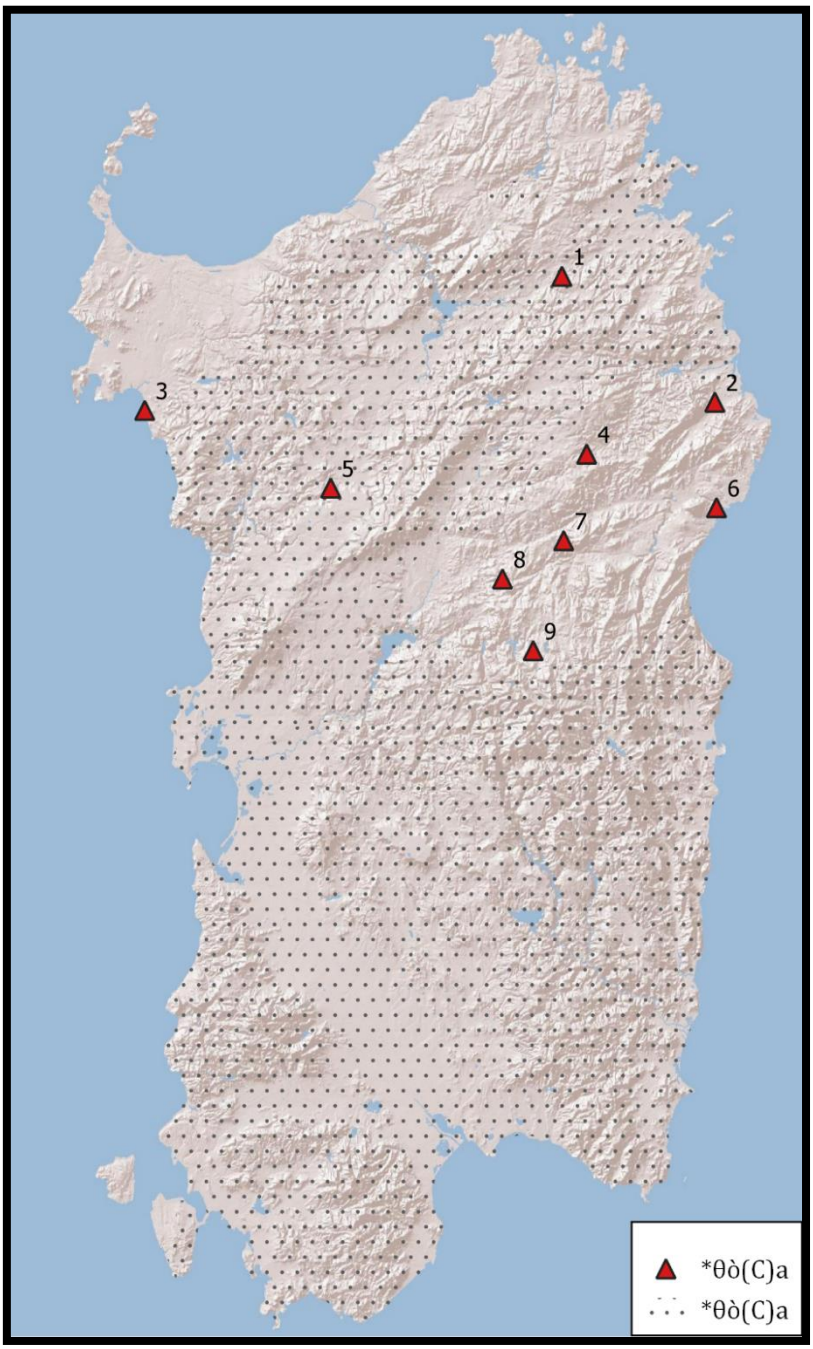

Figure 3.3.10: Osier (*Salix viminalis*)

This word for 'osier' is found all over Sardinia. Jud (1942: 352) compares Bq. *zume* 'willow', Laz. *zułę* (Castro dei Volsci) 'osier', Sp. *zade* (Salamanca) 'species of willow with thin stems that is grown on the banks of rivers and streams'. Wagner (DES II: 551) reasonably doubts these connections, as the comparison is essentially limited to the first consonant.[105] He nevertheless thinks the Sardinian word is of pre-Roman origin. Paulis (1992: 437), however, argues to have found a Greek etymology. He cites the construction "*dia iteon implastrum*" 'a poultice made from willow leaves', the first part of which is a Latinization of Greek διὰ ἰτεῶν 'made of willows'. This syntagm, which was supposedly borrowed into Romance as **di'iteón* (vel sim.) would subsequently have been reanalyzed as **di teón*. This last part would, according to Paulis, have developed regularly as follows: **teón* > **tɨ́ó* > **ϑò* >> *ϑòa* 'osier'.

This etymology is to be rejected on several grounds. Even if the Latinized construction *dia iteon* passed from the purely medical register into the common speech, it would require the unlikely scenario that speakers of Sardinian Romance using this syntagm knew enough Greek to be able to recognize Gr. διὰ as a preposition that could be detached, but did not know enough Greek to realize that ἰτεῶν was a genitive plural. It is unlikely that Sardinian speakers knowingly borrowed an oblique case form. Additionally, loss of final *-n* is irregular. We would rather expect **teón* to develop into ***ϑòno*, with a paragogic vowel as in (Log.) *sámbene* 'blood' < *sanguen*, *nòmene* 'name' < *nōmen* and *báttoro* 'four' < *quatt*(*u*)*or*. In absence of a better etymology, and because of the presence of the phoneme **ϑ-* which is present in many Sardinian words of obscure origin, I still hold Wagner's hypothesis that this is a word of pre-Roman origin as the most attractive.

As Wagner (DES II: 551) notes, it is difficult to say whether the intervocalic consonants of forms like *ϑòɣa* (4: Bitti, 8: Orani), *ϑòβa* (7: Nuoro) and *šòva* (9: Fonni) are old, or merely a device to resolve vowel hiatus (cf. also Wagner 1941a: 41–42). Moreover, if they are old, it is impossible to establish whether the original consonant was labial (as in Nuoro and Fonni) or velar (as in Bitti and Orani). We must therefore content ourselves with reconstructing **ϑò*(*b/g/v*)*a*.

[105] Moreover, Bq. *zume* 'willow' has a good internal Basque etymology: *zur*, *zu-* 'wood' + *mehe* 'slim' (Michelena 1990: 412).

3.3.11 Oleander (*Nerium oleander*)

Reconstruction	**Forms**
1: *neuláke	DES (II: 165): ***neuláke*** (Nuor., 1: Nuoro), ***neuláɣe*** (Log.), ***neuláχe*** (2: Dorgali), ***neuláǧi*** (3: Talana), ***neulaǧǧi*** (4: Triei) 'oleander (*Nerium oleander* L.)'
2: *leonárṛu	DES (II: 165): ***lonáržu*** (5: Oristano), ***solịonážu*** (7: Escalaplano), ***leonáržu*** (9: Sant'Antioco) 'id.'
3: *leonáke	DES (II: 165): ***leonáǧi*** (6: Perdasdefogu), ***launáži*** (8: Guspini, 10: Teulada) 'id.'

Although not explicitly, Wagner (DES II: 165) seems to consider these words for 'oleander' to be pre-Roman. The ending *-áke* would be parallel to words for the same plant found in Latin glosses as *biblace* etc., and in Greek and Romance dialects of southern Calabria: Bov. *plembáći*, *pilimbáći*, *glimbáći*, Calab. *limbáći* (Melito di Porto Salvo) (Rohlfs 1932: 413; 1964: 85). As Rohlfs remarks, only the the endings of these forms can be compared; the bases **bibl-/*plimb-* and **neul-/*leon-* must be kept separate. Alessio's (1937: 313–314) reconstruction of Srd. *neuláke* < **nibláke* << **liblace* << *biblace* does not work phonetically, as **nibláke* would yield Nuor. ***niβráke/**nirβáke* vel sim. Moreover, his assumption that *biblace* 'oleander' would be reformed into **liblace* on the basis of the association between Gr. βίβλος and Lat. *liber* 'book', is rather far-fetched. The morphophonological variation in the Sardinian forms makes it clear that some folk-etymological processes must have been involved. Wagner (DES II: 165) believes that the forms in types 2 and 3 have been influenced by It. *oleandro* 'oleander'.[106] Paulis (1992: 419) suggests that there was influence from Srd. *oliòni* 'strawberry tree' (e.g. in *solịonážu*; 7: Escalaplano; § 3.3.16), as well as from *láu* 'laurel' (in *launáži*; 8: Guspini, 10: Teulada).

Both Wagner (DES II: 165) and Paulis (1992: 419–420) regard the forms of type 1 as the original ones, whence types 2 and 3 arose through folk etymology. However, whereas Wagner sees *neuláke* etc. as originally pre-Roman, Paulis proposes it to be derived from Srd. *néula* 'mist'. The reason for this would be that, from oleander, a type of eye drops was produced that served to dilate the pupils. On the basis of this, and the fact that certain eye conditions are called *sas néulas 'essos òkros* (litt.: 'mists of the eyes') in the Nuorese dialects, Paulis argues for a scenario in which Nuor. *néula* < Lat. *nebula* 'mist' was extended with a pre-Roman suffix *-áke* and applied to 'oleander'.

[106] Salvioni (1909: 220) believes all the Sardinian forms are learned in origin, but this is unlikely.

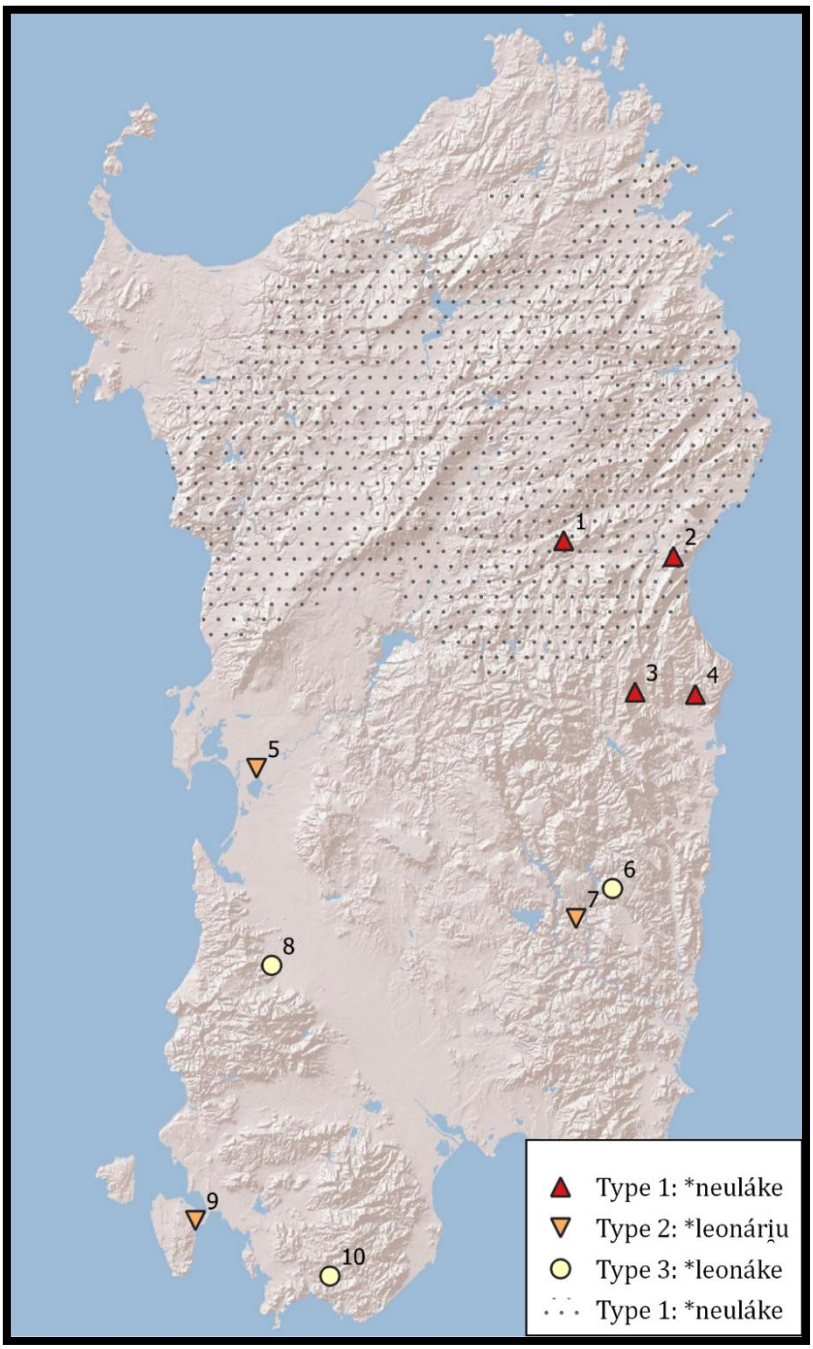

Figure 3.3.11: Oleander (*Nerium oleander*)

The problem is that, if the association between oleander and *néula* 'mist' indeed exists, it is still rather difficult to explain the forms of types 2 and 3. As we have seen, the vocalism of *soḷ̣onážu* (7: Escalaplano) on the one hand, and of *launáži* (8: Guspini, 10: Teulada) can most likely be explained by influence from *oliòni* 'strawberry tree' and *láu* 'laurel' respectively (1992: 419), but neither these words nor It. *oleandro* 'oleander' can satisfactorily explain the vocalism *-o-*, *-eo-* that we find in *lonáržu* (5: Oristano), *leonáği* (6: Perdasdefogu) and *leonáržu* (9: Sant'Antioco). With this in mind, we may consider whether **leonáke* was instead the original form, with the other forms being altered by association with *néula* 'mist', *oliòni* 'strawberry tree', *láu* 'laurel'. This would still likely be of pre-Roman origin, as already suspected by Wagner (DES II: 165) for this word. On the hypothetical suffix **-áke*, see § 9.2.2.1. In the forms of type 2, this ending has been replaced with the reflexes of inherited Lat. *-arius*.

3.3.12 Pine tree (*Pinus sp.*)[107]

Reconstruction	**Forms**
***kampínḭu**	DES (I: 278): ***kompínğu*** (1: Iglesias) 'pine (*Pinus sp.* L.)', ***kampínğu*** (Camp., 2: Cagliari)
	Puddu (2023: s.v. *campíngiu*): *campíngiu*, *compíngiu*, *cumpíngiu*, *cumpríngiu* 'pine'

This cluster of words refers to various species of pine-tree (*Pinus*), along with words of the type *pínu* and *oppínu*, respectively from Lat. *pīnus* and *sappīnus* 'kind of pine/fir'.[108] The element *-pínğu* is the regular outcome of **píneu* < Lat. *pīneus* 'of the pine'. The variation between **píneu* ~ **kampíneu* thus looks like an element *kV-* was added to a Latin root (§ 9.1.1). It is not immediately clear whether the *-m-* would originally have been part of the prefix, or whether it arose through secondary *m*-insertion, like in *sampunáre* (Orgosolo) << *sapunare* 'to wash' (Wagner 1941a: 223).

A completely different interpretation is by Paulis (1992: 443), who prefers to explain *kampínğu*, *kompínğu* as a semantic extension of original **conuspineus* 'pine cone'. If this is correct, this word obviously does not constitute an example of **ka*-prefixation. There are two possible objections to this etymology. One is the fact that Lat. *conus* '(pine) cone' is as such not preserved in Sardinian, although *kampínğu* etc. could of course be the result of a very early compound.

---

[107] These data are discussed in a similar fashion in Swanenvleugel (2024: 230–232).

[108] Srd. *oppínu* is due to reanalysis of the *s-* in Lat. *sappīnus* as the definite article (DES II: 272).

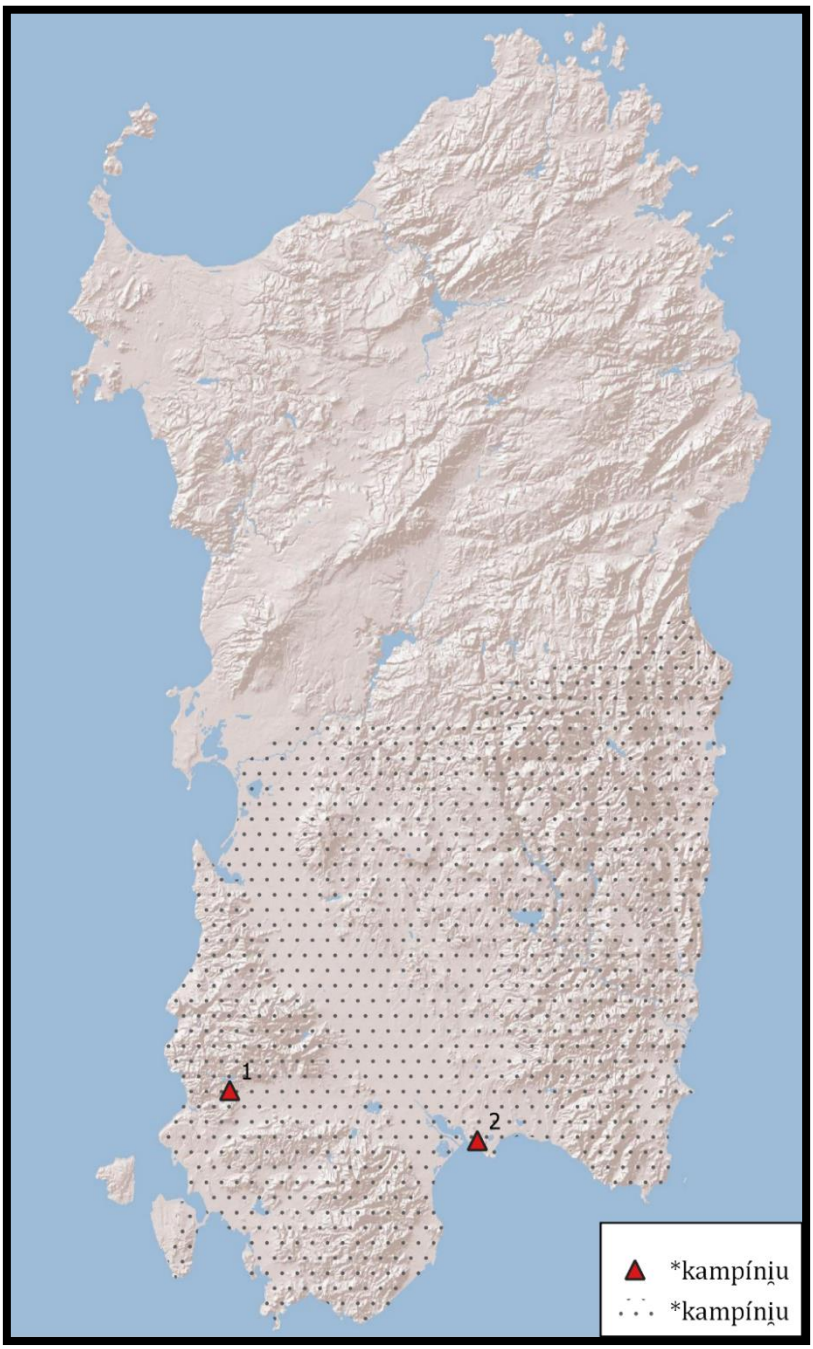

Figure 3.3.12: Pine tree (*Pinus sp.*)

The other issue is the phonetic development. We would have to assume loss of the final vowel of **kónu* (**kónu píneu* > **kónpíneu* > **kompíneu*), which is not quite regular, and subsequently the shift from pretonal *-o-* to *-a-*. This latter development, as well as the opposite development (pretonal *-a-* to *-o-*) is not without parallels (cf. Wagner 1941a: 28, 32), but if Wagner (DES I: 278) is right in attributing the form *kampíŋ̆gu* to most of Campidanese, it would seem that geographically restricted forms of the type *kompíŋ̆gu* (1: Iglesias) (and Puddu's *cumpíngiu* etc.) represent secondary alterations of an original **kampíneu*. Although Paulis' analysis of *kampíŋ̆gu* etc. as regular reflexes of **conus pineus* 'pine cone' is semantically very attractive, I am not convinced that it could work phonologically.

3.3.13 Rockrose (*Cistus sp.*)

Reconstruction	**Forms**
1: *mutéklu	Wagner (1928: Map 25): ***muðéyu*** (1: Luras, 11: Ozieri), ***murðéǧ̆gu*** (5: Osilo), ***murdéɣu*** (7: Ploaghe, 10: Torpè, 34: Norbello, 36: Bonarcado, 40: Milis, 41: Narbolia, 42: Busachi, 44: Sorgono, 45: Tonara, 49: Aritzo, 51: Meana Sardo, 52: Arzana, 54: Gadoni, 55: Lanusei, 56: Seulo, 57: Laconi, 58: Gairo, 59: Seui, 60: Ulassai, 61: Isili, 64: Perdasdefogu, 66: Siurgus, 68: San Nicolò Gerrei, 71: Cagliari, 73: Domus de Maria), ***muðéɣu*** (8: Olmedo, 14: Mores, 47: Atzara), ***muðréɣu*** (9: Posada, 16: Bono, 24: Macomer, 26: Scano di Montiferro, 29: Cuglieri, 33: Santu Lussurgiu, 39: Seneghe, 43: Baunei, 50: Cabras, 62: Terralba, 63: Mogoro, 65: Escalaplano, 69: Samassi), ***muðéǧ̆gu*** (12: Pattada, 18: Padria, 27: Sennariolo), ***mutréku*** (13: Siniscola, 15: Bitti, 17: Orune, 19: Orosei, 20: Lollove, 21: Nuoro), ***muðréχu*** (22: Dorgali), ***muðréku*** (23: Oliena, 30: Olzai, 31: Ollolai, 32: Gavoi, 35: Fonni, 37: Ovodda), ***muðéčču*** (26: Orani), ***murtélʔu*** (28: Orgosolo), ***muðréhu*** (38: Urzulei), ***muðéɣru*** (46: Desulo, 50: Samugheo), ***murdéɣru*** (48: Villagrande Strisaili), ***muddéɣu*** (67: Guspini, 70: Villacidro), ***muréɣu*** (72: Sant'Antioco) 'rockrose (*Cistus sp.* L.)'
	DES (II: 150–151): ***muteclu*** (OLog.), ***mudeglu*** (OCamp.), ***muðréɣu, murdéɣu*** (Log.), ***murdéɣu***, ***muðéɣu*** (Camp.) 'id.'
2: *mút(u)lu	Wagner (1928: Map 25): ***múćću*** (2: Tempio Pausania), ***múǧ̆gu*** (3: Sorso), ***múčču*** (4: Sennori, 6: Sassari)

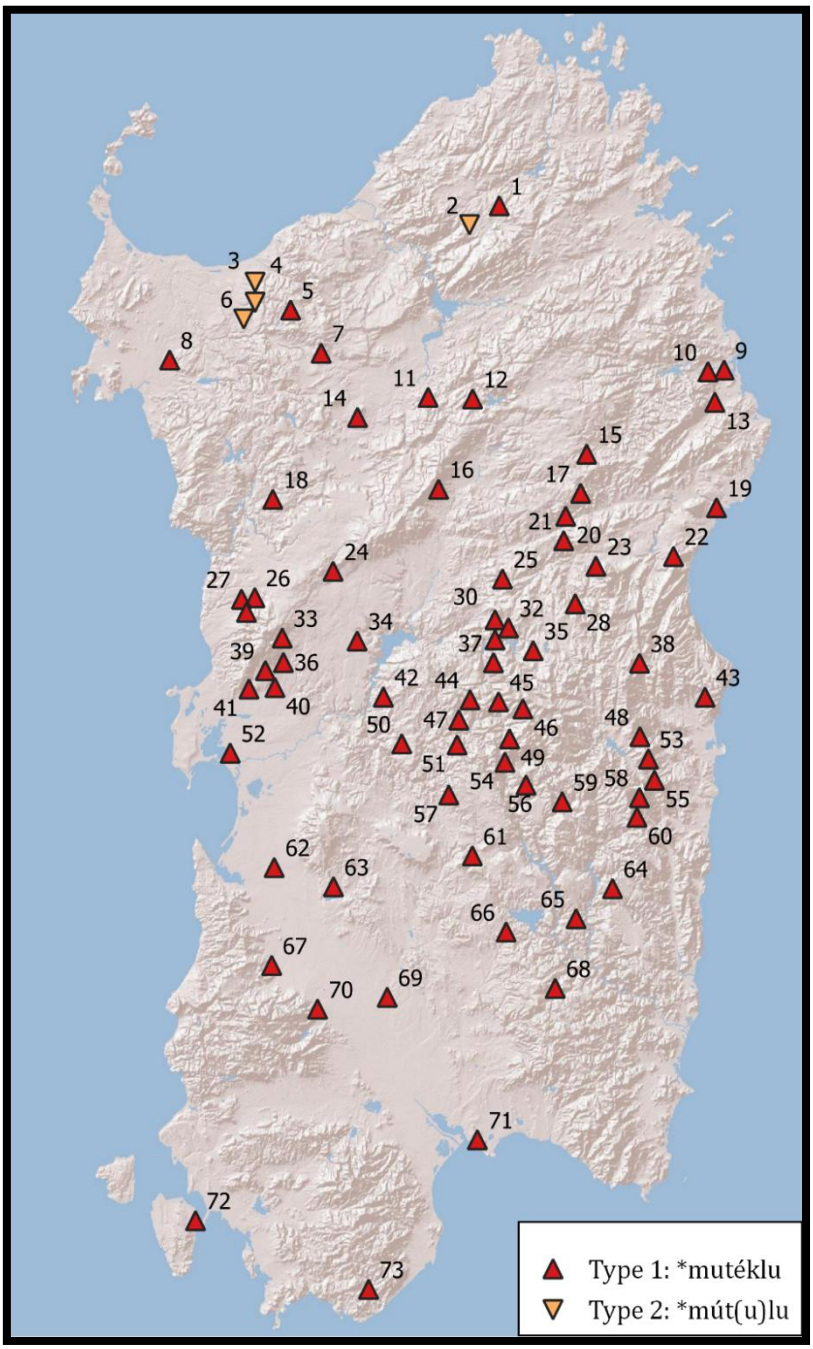

Figure 3.3.13: Rockrose (*Cistus sp.*)

All forms of type 1, which is attested all over Sardinia, can regularly go back to *muteclu*, attested in the Condaghe di San Pietro di Silki (Bonazzi 1900: secs. 207, 418). Type 2 can go back to **mútulu*, and is found in other Romance varieties as well (DES II: 150). This word has been connected by Bertoldi (1936: 312 fn. 1) to Etr. μούτουκα 'thyme, rockrose' (attributed to the "Θοῦσκοι" by Dioscorides Anazarbei 1906: 48; cf. Bonfante & Bonfante 1983: 190).[109] Words for 'rockrose' in other Romance languages that have been compared, include: Calab. *mútaca*, *múndaca*, *amùndaci*, *amùndici* etc. 'narrow-leaved cistus (*Cistus monspeliensis* L.)' (Bertoldi 1936: 312; Alessio 1949: 114; Rohlfs 1977: 439, 447),[110] Lang. *mouche* (dep. Aveyron), *mouge* (dep. Gard) etc., Cat. *módega* 'id.' (Hubschmid 1953: 36; FEW VI/III: 306; DECLC V: 715-716) < **mútVka*; Cat. *motxera* (reg. Roussillon) 'id.' (FEW VI/III: 306) < **mutukária*; Lang. *moudré*, *moudre* (St-Sernin) 'id.' (FEW VI/III: 306) < **mútaro-*; Tusc, It. *mucchio* 'id.' (Bertoldi 1936: 311), Cors. *mucchju*, *múćću* (Bertoldi 1936: 312 fn. 1; DES II:150), Gallur. *múćću*, Sass. *múčču*, Calab. *mucchiu*, Salent. *muk'k'u* (Francavilla Fontana) 'id.' (DES II: 150) < **mútulu*.

The connection between these forms is semantically and formally likely. We appear to be dealing with a root **mut-*, extended with different suffixes. Srd. *muteclu* etc. is especially close to Etr. μούτουκα 'thyme' and the Romance forms going back to **mútVka*.

3.3.14 Rosemary (*Salvia rosmarinus*)

Reconstruction	**Forms**
****ϑíppiri***	DES (II: 593): ***tsíppiri*** (Camp., 4: Desulo), ***ϑíppiri*** (1: Urzulei, 2: Triei, 3: Baunei), ***síppiri*** (5: Gairo, 6: Perdasdefogu) 'rosemary (*Salvia rosmarinus* Spenn.)'

This word for 'rosemary' is present in the southern half of Sardinia, where Wagner (DES II: 593–94). It is first attested in the Condaghe di Santa Maria di Bonarcado as *cipiri* (Virdis 2003: sec. 1c). This word is likely a loan from Pun. *zibbir* 'rosemary', attested in Pseudo-Apuleius (80: 31; Howald & Sigerist 1927: 145; Bertoldi 1947;[111] DES II: 594).

---

[109] Wagner has: "Etr. μούτουκα '*Cytinus hypocistis* (L.) L.'". *Cytinus hypocistis* is a species of parasitic plant that parasitizes rockroses. This meaning for μούτουκα is however not actually attested, but inferred by Bertoldi (1936: 311) on basis of the fact that Dioscorides (1907: 87; 1906: 48) gives Gr. θύρσιον as a synonym for both 'thyme' and '*Cytinus hypocistis*'. Wallace (2008: 9) gives the meaning '*Anagallis*'.

[110] Rohlfs (1934: 73) is less certain about the precise meaning of this word however.

[111] *Non vidi*.

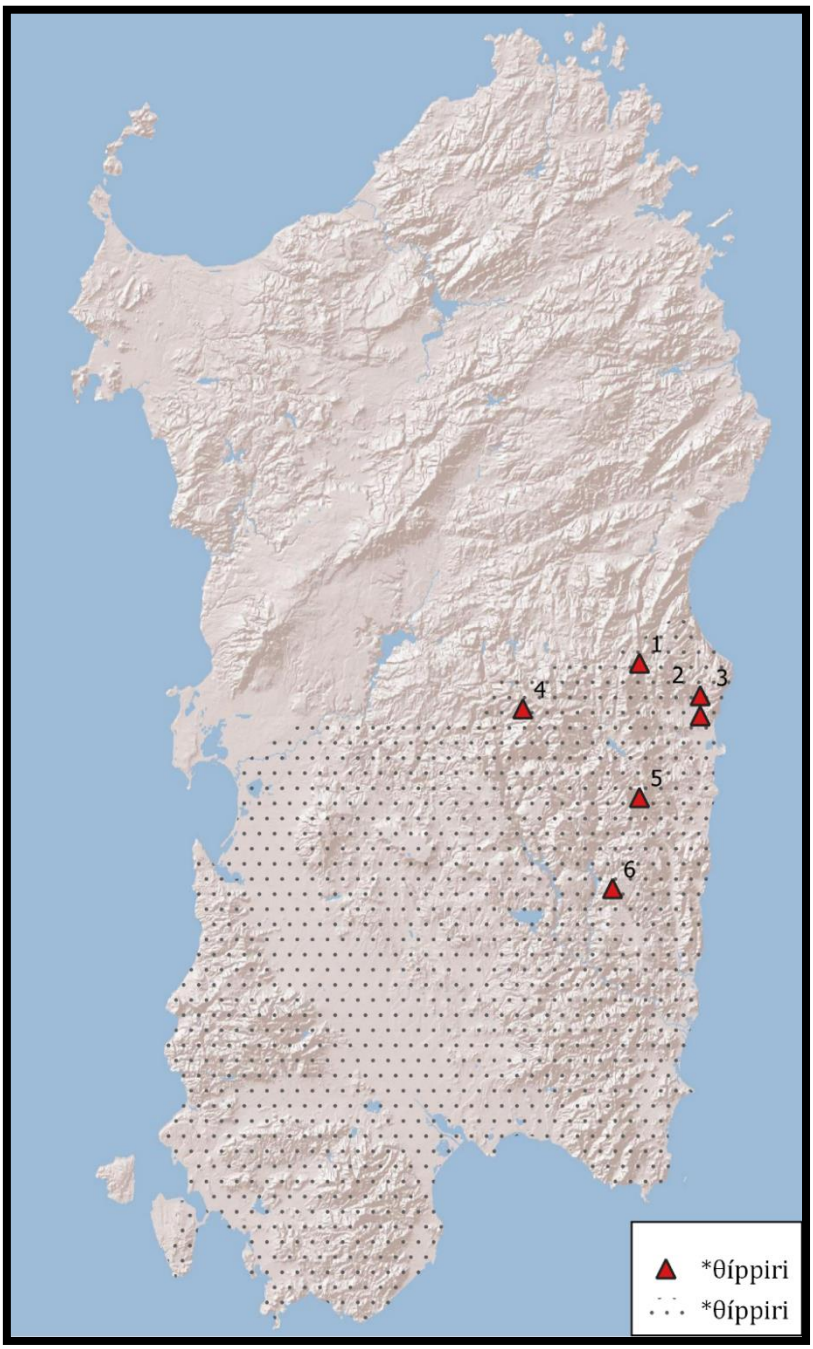

Figure 3.3.14: Rosemary (*Salvia rosmarinus*)

As in other Punic loans, the sibilant *z-* has been borrowed as **ϑ-* in Sardinian (cf. § 8.1.2); devoicing of geminate **-bb-* > *-pp-* is regular (Wagner 1941a: 192–193). Paulis (1990: 602) further notes toponymic attestations of *típpiri*, the expected Logudorese correspondent to Camp. *tsíppiri* etc., in Tresnuraghes, Bosa, Borutta, Cheremule and Torralba. For more detailed discussions of associated toponyms and personal names, cf. Wagner (DES II: 594) and Paulis (1990: 602–603). Múrcia Sànchez (2011: 428–429) has compared Berber forms like <*āzīr*> 'rosemary' (*'Umdat aṭ-ṭabīb*; Tilmatine & Bustamante 2002: 429), <*iazir, aziir*, *ouzbir*> 'rosemary' (Trabut 1935: 223), Rif., Tmz. *azir* 'rosemary, lavender; wormwood', which could go back to **ā-zV̌βīr*.[112] If this connection is correct, we are likely dealing with a word of North-African origin that was introduced in Sardinia during the Punic colonization. Paulis (1992: 133) suggests that this Punic loan may have taken place in the context of trade in aromatic and medicinal herbs.

3.3.15 Smilax (*Smilax aspera*)

Reconstruction	**Forms**
1: *tétti	DES (II: 480): ***tét(t)i*** (1: Villanova Monteleone, 4: Orosei, 7: Dorgali, 10: Jerzu) 'smilax (*Smilax aspera* L.)
	Paulis (1992: 220): ***tèttu*** (Gallur.), ***tét(t)i̯u*** (6: Bosa) 'id.'
2: *tetti̯òne	DES (II: 480): ***titti̯òne*** (2: Bonorva, 3: Padria, 5: Sindia, 9: Laconi), ***teti̯òne*** (8: Santu Lussurgiu), ***tintiòi*** (10: Guasila), ***tintiòni*** (11: Sarroch) 'id.'

Representatives of this word are found all over Sardinia, including in Gallurese. The basis of the various forms appears to be **tétti*, which is directly continued by the forms in type 1. Final *-u* is added secondarily in *tét(t)i̯u* (6: Bosa) and replaces original final *-i* in Gallur. *tèttu*. The forms of type 2 must have received the suffix *-òne* after the common Sardinian development **-ti̯-* > **-ϑ-*. In the Condaghe of San Pietro di Silki (Bonazzi 1900: sec. 10), we also find *tectinosu* 'full of spiny herbs like *tét(t)i*' (DES II: 480), but the spelling *-ct-* may be an unetymological Latinizing spelling (cf. Wagner 1941a: 170).

[112] Note that most forms can, but need not contain **β*. The form <*ouzbir*>, recorded by Trabut (1935: 223) and of unclear provenance, is the only form that contains a labial consonant. Maarten Kossmann informs me, however, that *b* cannot go back to **β*. Kossmann (p.c.) moreover prefers a reconstruction **zβīr* over Múrcia Sànchez' (2011: 428–429) **zV̌βīr*.

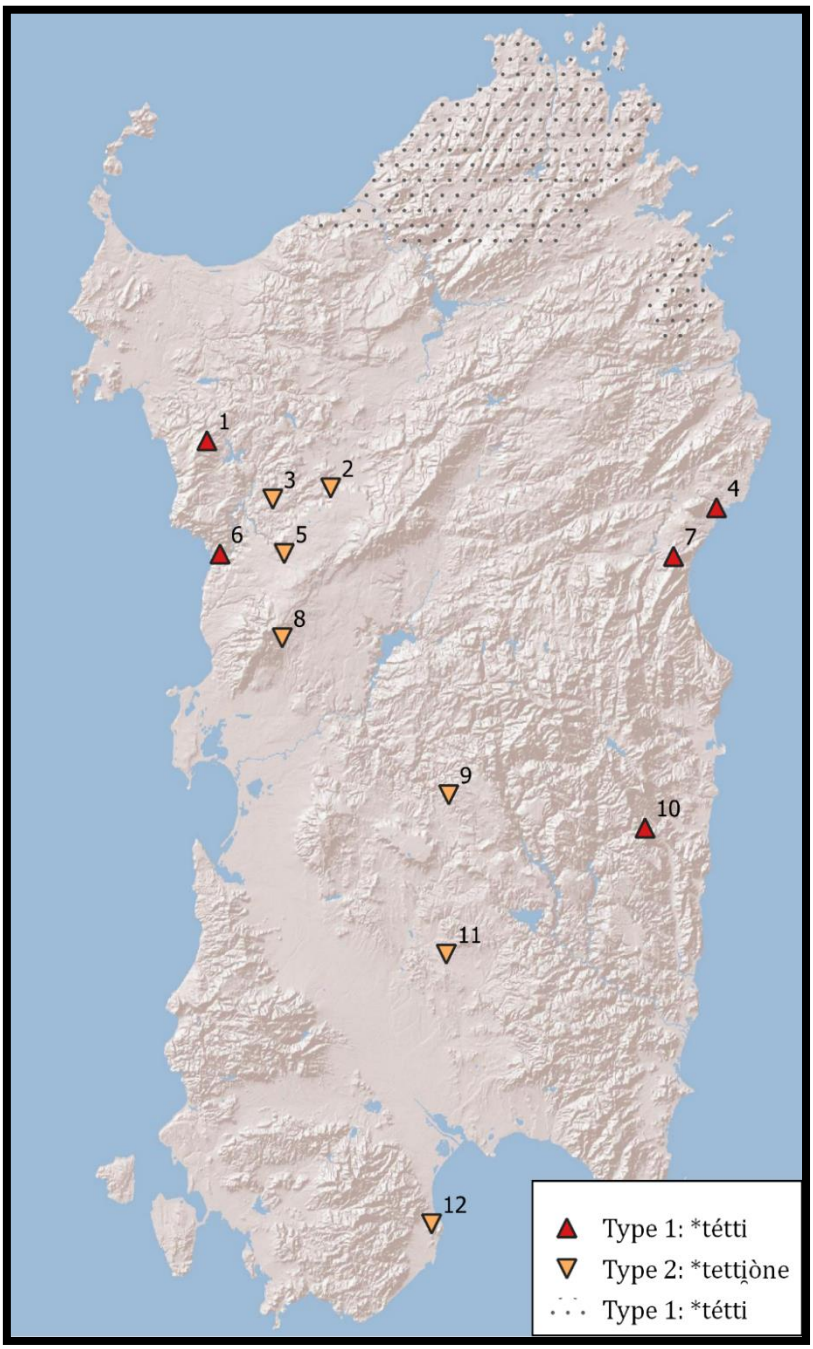

Figure 3.3.15: Smilax (*Smilax aspera*)

DES (II: 480) considers this word "probably pre-Roman" and compares various toponyms that supposedly contain the same element: *Teti* (a *nuraghe* near Ploaghe), *Teti* (a village in the Barbagia di Mandrolisai), *Tettène* (a *domus de gianas* near Perfugas) and *Tettinosa* (a nuraghe near Perfugas). Whether or not these toponyms actually contain **tétti* 'smilax', it is likely that Wagner is right in attributing it to pre-Roman Sardinian, because of its lack of an etymology and the fact that it ends in *-i*, for which there is no good source in Latin nominal morphology (§ 8.1.5.1).

3.3.16 Strawberry tree (*Arbutus unedo*)

Reconstruction	**Forms**
1: *olidòne	DES (II: 480): ***oliòni*** (Camp.), ***liðòne*** (3: Siniscola, 5: Bitti, 6: Nule, 8: Nuoro, 12: Fonni), ***oliðòne*** (13: Urzulei, 16: Baunei), ***oliòne*** (18: Laconi), ***liòni*** (19: Seui, 21: Perdasdefogu), ***oĩ̯ði*** (20: Mogoro), ***oβiòi*** (23: Villacidro) 'strawberry tree (*Arbutus unedo* L.)'
	AIS (1286): [***lu baga***]***liòni*** (1: Tempio Pausania), [***su***] ***liðòne*** (9: Dorgali), [***s***] ***iliòne*** (17: Desulo), [***s***] ***olliòni*** (22: Escalaplano), ***oβiði̯*** (23: Villacidro), [***su***] ***liòni*** (24: Cagliari), [***s***] ***oriòni*** (25: Sant'Antioco) 'strawberry tree (*Arbutus unedo*)'
2: *olitòne	DES (II: 480): ***liðòne*** (7: Bonorva, 10: Macomer) 'strawberry tree (*Arbutus unedo*)'; [***melu***]***liðòne*** (4: Villanova Monteleone)
	AIS (1286): [***s***] ***oliðòne*** (2: Ploaghe), [***su***] ***liðòne*** (11: Santu Lussurgiu, 15: Busachi), [***su***] ***liðði̯*** (14: Milis) 'id.'

This word for 'strawberry tree', an evergreen shrub or small tree that is characteristic of the Mediterranean maquis biome, is attested all over Sardinia, and is usually taken to be from Lat. *ūnedō* 'strawberry tree', attested multiple times in Pliny's *Naturalis Historia*. The forms of type 1 listed above all require a pre-form **olidone*, which in some cases lost its initial vowel (e.g. *liðòne*, *liòni* etc.). The forms in type 2 are synchronically identical to some of those attested in type 1. However, they occur in the Logudorese dialects, where original intervocalic **-d-* should have been lost (Wagner 1941a: 79). The easiest explanation is to assume some sort of restoration of etymological *-d-*, although the circumstances under which this could have happened are not fully clear to me.

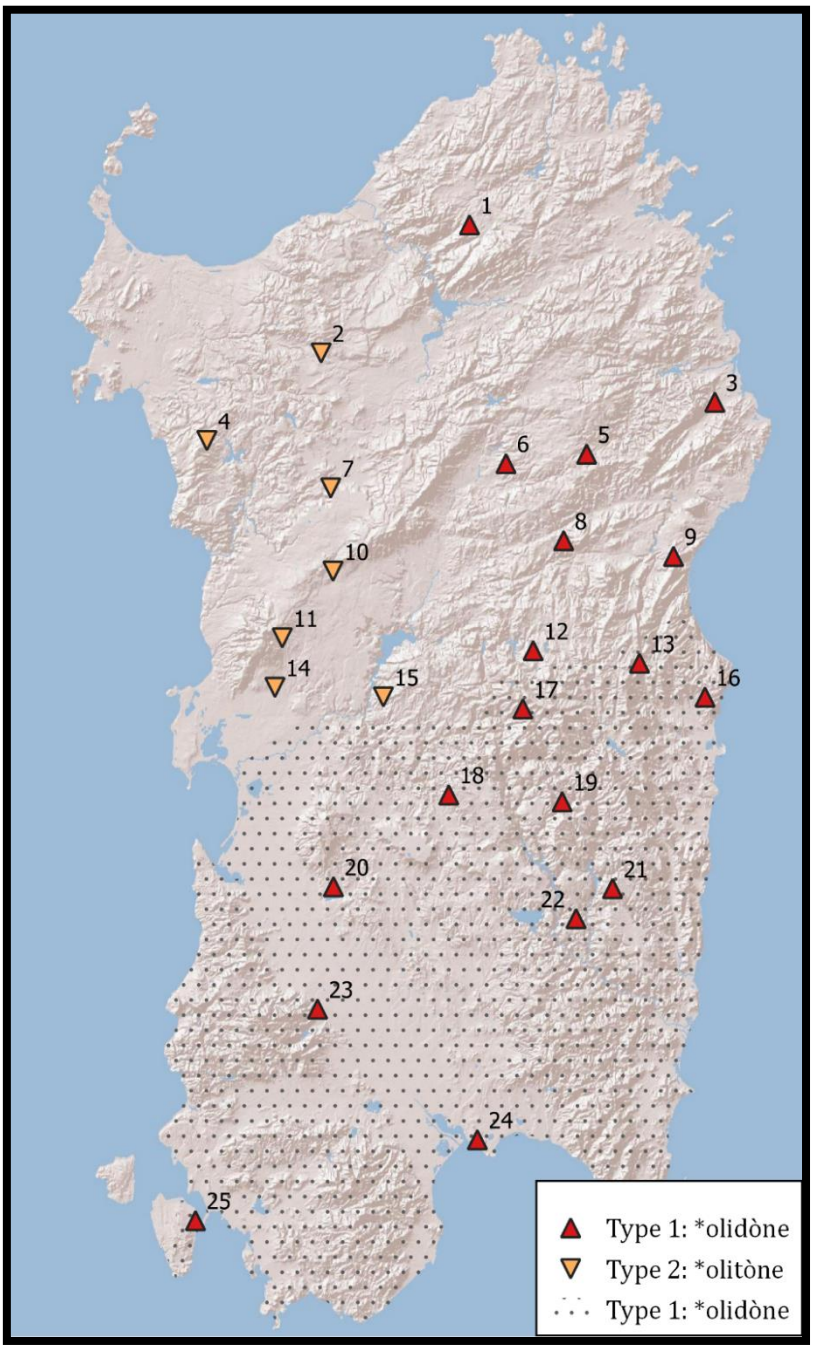

Figure 3.3.16: Strawberry tree (*Arbutus unedo*)

Lat. *ūnedō* is only attested in Pliny's *Naturalis Historia* (and later sources referring to it), where its name is said to be from *ūnum edō* 'I eat one', because of this fruit's poor taste (cf. Plinius Secundus Maior 1857: chap. 15:28). This is an evident folk etymology. Besides Srd. *oliðòne* etc., the only other Romance reflexes of this word are found in Gallo-Italian and Gallo-Romance dialects: Pied. (*l*)*ürion*, Saintongeais *olon*, Gasc. *auledun* (dep. Landes), Lang. *leduno* (reg. Guyenne) 'strawberry tree' (REW 9068; cf. also FEW XIV: 35). All Romance forms can be explained more easily from **olidōne* than from *ūnedō* attested in Pliny.[113] In light of Pliny's folk etymology from "*ūnum edō*", Lat. *ūnedō* 'strawberry tree' may perhaps be from **olidō*. It is also remarkable that this word was preserved on such a limited territory, which raises the question what the relation was between Lat. *ūnedō* and *arbutus*, both of which refer to the strawberry tree. Given the marginal dispersion of the descendants of *ūnedō* (as opposed to those of *arbutus*, which are more widespread) and the fact that *ūnedō* is only attested in one Latin source and is likely influenced by a folk etymology, it seems that *ūnedō* (or indeed **olidō*) was not a common name for this plant. Perhaps **olidō* was a pre-Roman word that entered some varieties of Vulgar Latin, and happened to be recorded by Pliny. Its geographical distribution is similar to that of various other words of suspected pre-Roman origin, some of which are nevertheless attested in Classical sources (§ 11.2.1).

3.3.17 Strawberry tree flower (*Arbutus unedo*)

Reconstruction	**Forms**
****a(C)íϑϑo***	DES (I: 64): ***aíϑϑo*** (1: Urzulei), ***aíϑϑu*** (2: Talana), ***aíssu*** (3: Perdasdefogu) 'flower of the strawberry tree (*Arbutus unedo* L.)'

I agree with Wagner (DES I: 64) that these forms are probably of pre-Roman origin. This is not only likely because of its lack of an etymology, but also because of its peculiar phonological structure. The final *-o* in *aíϑϑo* (1: Urzulei) can hardly be explained from Latin. In *aíϑϑu* (2: Talana) this unusual ending has been replaced with the inherited masculine suffix *-u*.[114]

[113] The observation that the attested Romance forms correspond regularly to Sardinian *oliðòne* etc., was already made by Schuchardt (1904: 194).

[114] In *aíssu* (3: Perdasdefogu) it is impossible to see whether this form contains original *-o* or *-u*, because *-o* > *-u* in unaccented position in this area (cf. Wagner 1907b: Map 3; Virdis 1988: 908).

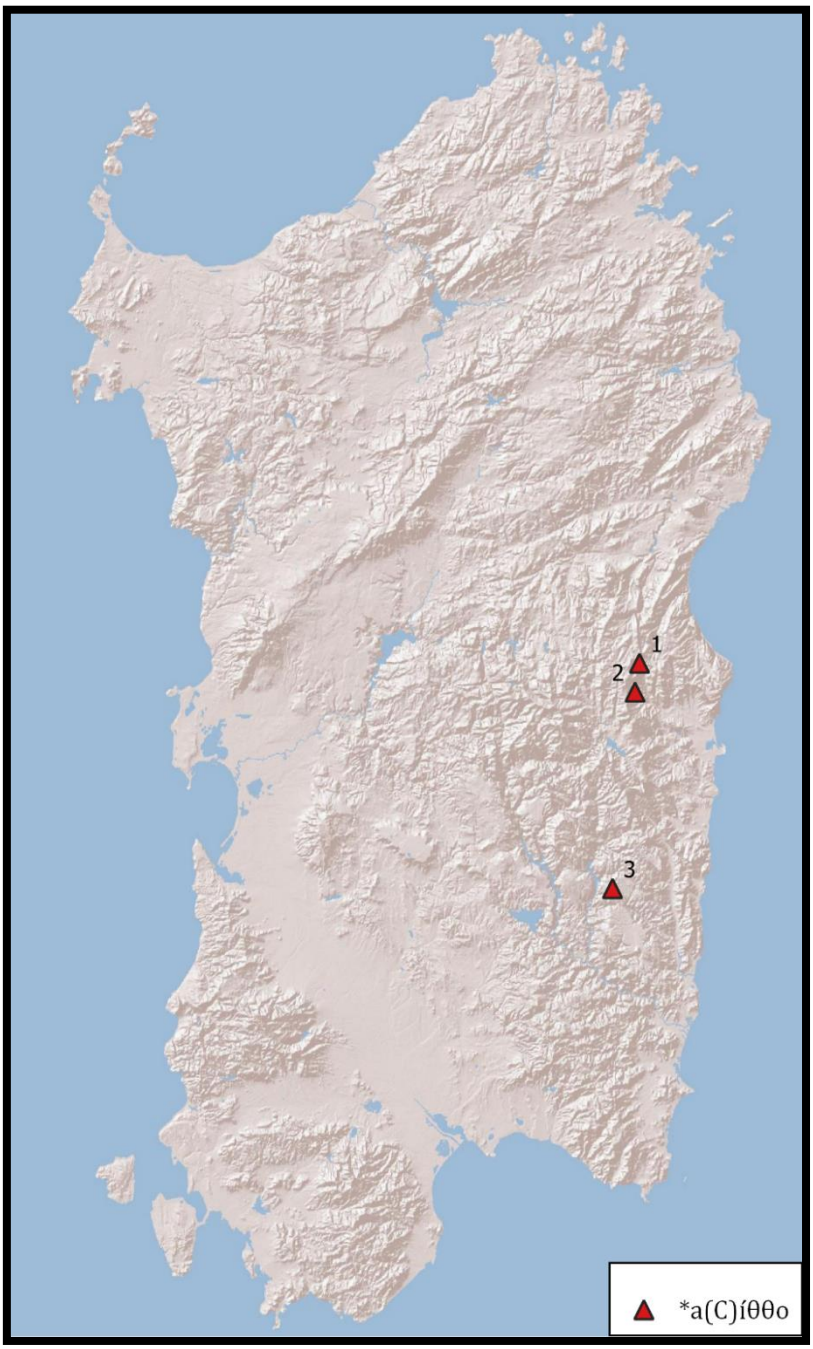

Figure 3.3.17: Strawberry tree flower (*Arbutus unedo*)

Wagner (DES I: 64) speculates that these words would have been the pre-Roman name of this plant or its flower, but given the fact that in *oliðòne* and *liòni* (§ 3.3.16) are attested respectively in Urzulei and Perdasdefogu as the name for the tree (DES II: 185), it is more likely that the modern restriction of *aíϑϑo* etc. to the flower of this plant reflects its original semantic value.

3.3.18 Strawberry tree fruit (*Arbutus unedo*)

Reconstruction	**Forms**
1: *bilisòne	DES (II: 185): ***bilišòne*** (1: Urzulei), ***pilišòne*** (3: Baunei), ***bilišòni*** (6: Gairo), ***gilišòni*** (7: Seui) 'fruit of the strawberry tree (*Arbutus unedo* L.)'
2: *lisiriòne	DES (II: 185): ***liširiòne*** (2: Busachi) 'id.'
3: *kele(C)isòne	DES (II: 185): ***kelleišòne*** (4: Tonara) 'id.'
4: *lisòne	DES (II: 185): ***lišòne*** (5: Villagrande Strisaili) 'id.'

These forms, that specifically refer to the strawberry tree's fruit, bear a superficial similarity to *oliðòne* etc. 'strawberry tree' (§ 3.3.16), but cannot be derived from it. Wagner (DES II: 185) and Paulis (1992: 413) believe these forms are of pre-Roman origin. Because of the variation attested among these forms, it is difficult to establish what the original base was. The forms in type 1 may have been influenced by Srd. *gilindzòne*, *elindzòne*, *bilindzòne*, *pilindzòne* etc. 'barley bran' < **iligione* << **siligione* < Lat. *siligō* 'winter wheat' (cf. DES I: 578), or have at least undergone a similar process of reinterpretation of the initial consonant. Perhaps the absence of *bi-/gi-/pi-* in type 4 *lišòne* (5: Villagrande Strisaili) is an archaism, a suspicion that is reinforced by type 2 *liširiòne* (2: Busachi). The element *-iri-* in *liširiòne* (2: Busachi) is also found in several other words of potential pre-Roman origin (§ 9.2.6.2), suggesting **liš-iri-òne* as an internal analysis. I do not know how to explain type 3 *kelleišòne* (4: Tonara). The sequence *kellei-* too gives the impression of being secondary, but it is unclear what its source could be. Interestingly, the Ogliastran dialects, and specifically that of Urzulei, have three etymologically unrelated words for 'strawberry tree' (*oliðòne*; § 3.3.16), 'flower of the strawberry tree' (*aíϑϑo*; §3.3.17) and 'fruit of the strawberry tree (*bilišòne*), of which at least the last two (and perhaps all three) are likely of pre-Roman origin.

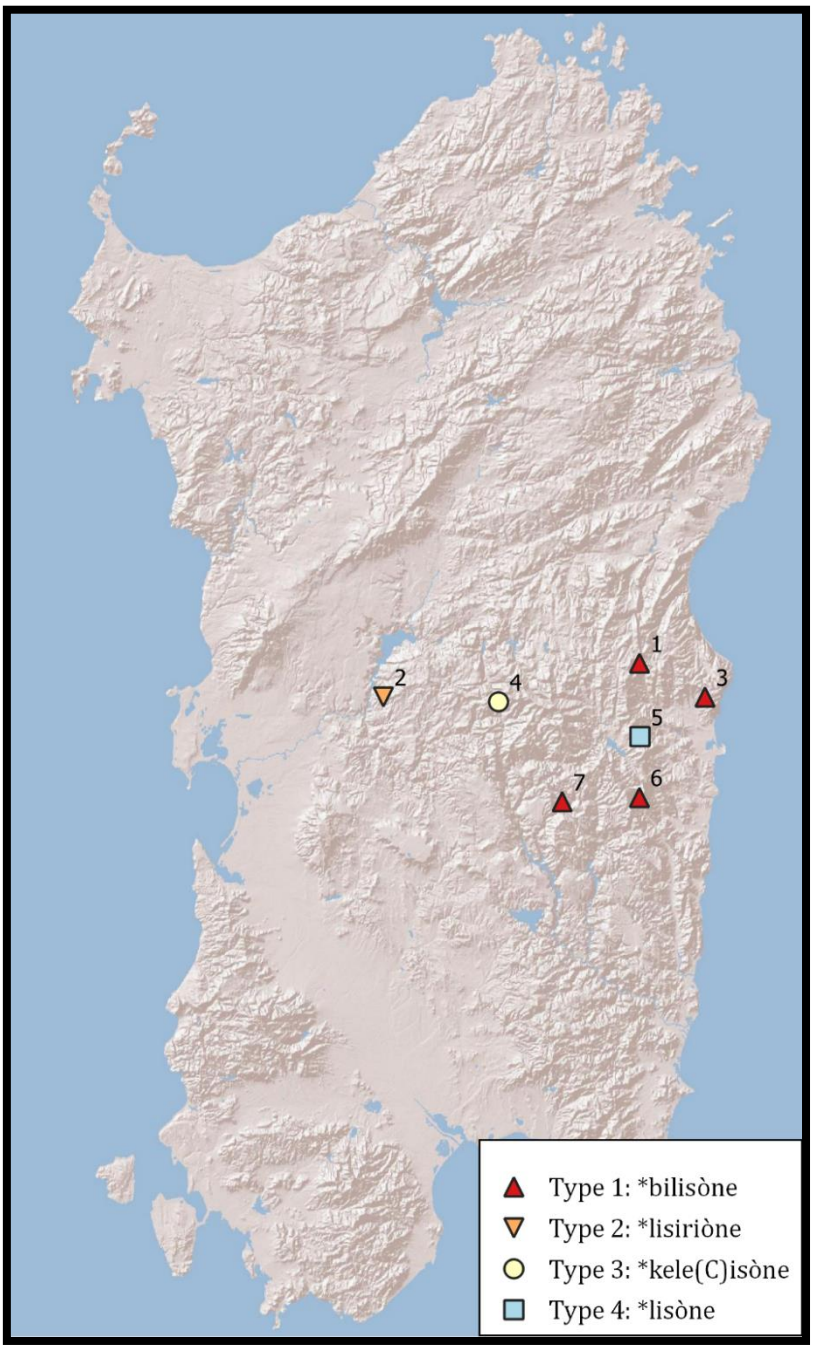

Figure 3.3.18: Strawberry tree fruit (*Arbutus unedo*)

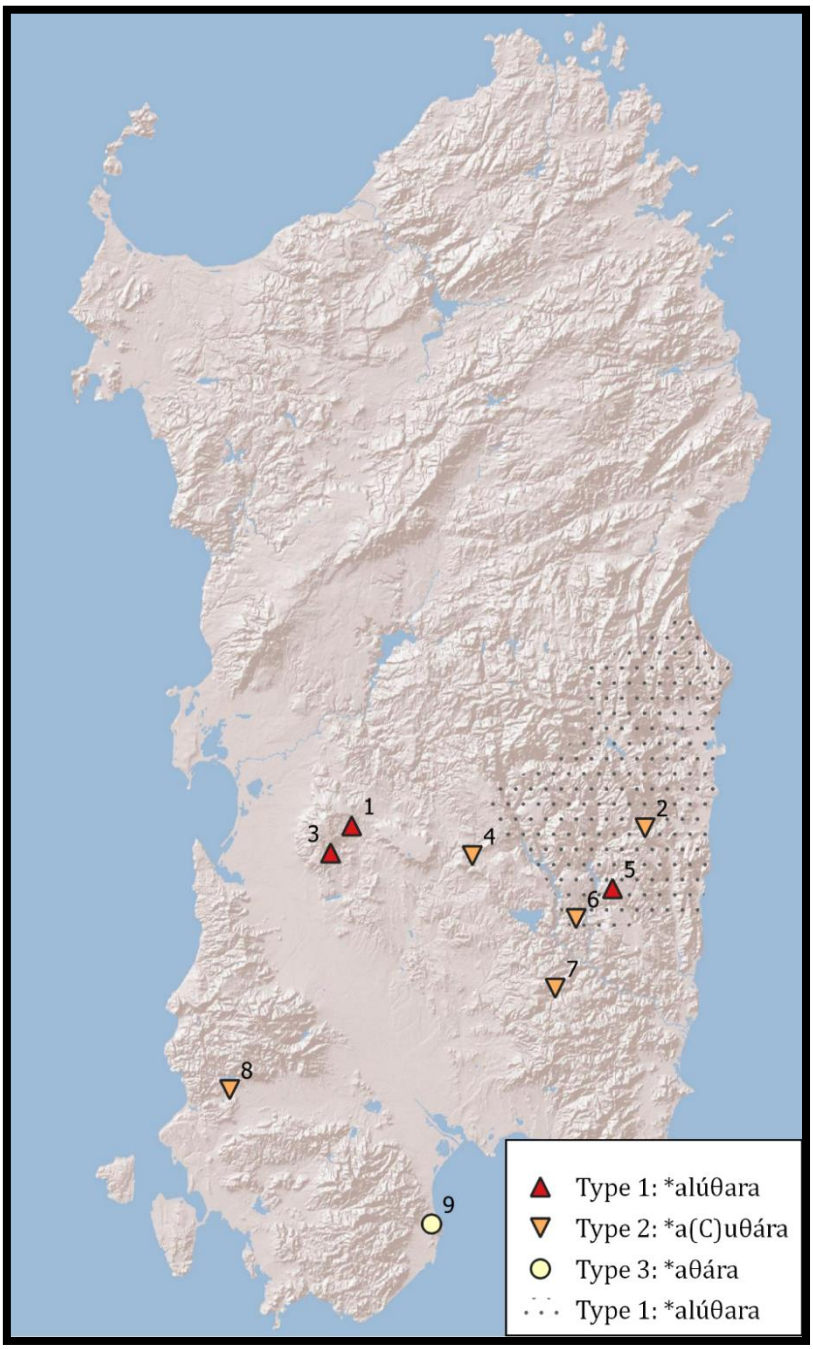

Figure 3.3.19: Traveller's joy (*Clematis vitalba*)

3.3.19 Traveller's joy (*Clematis vitalba*), fragrant virgin's bower (*Clematis flammula*)

Reconstruction	Forms
1: *alúθara	Blasco Ferrer (1989: 18 fn. 6): ***a(l)úsara*** (Ogl.) 'fragrant virgin's bower (*Clematis flammula* L.)'
	Paulis (1992: 182): ***luttsára*** (1: Villaverde, 3: Morgongiori), ***alússara*** (5: Perdasdefogu) 'id.'
2: *a(C)uθára	DES (I: 153): ***aussára*** (2: Jerzu), ***auttsára*** (4: Isili, 6: Escalaplano, 7: San Nicolò Gerrei, 8: Iglesias) 'traveller's joy (*Clematis vitalba* L.)'[115]
3: *aθára	DES (I: 153): ***attsára*** (9: Sarroch) 'id.'

Beside the forms listed above, *tsára* and *auttsáða* have been attested in the same meaning, but no location is known for these forms (cf. DES I: 153). This word has long been considered to be of pre-Roman origin (Wagner 1941c: 321; DES I: 153; Paulis 1992: 182–183). Among the Sardinian forms there are some variations that cannot easily be explained formally, mainly concerning the shape of the first syllable. For type 1, Paulis (1992: 183) suggests influence from Camp. *lúži* 'light', because its wood would be used to start fires.[116] Types 2 and 3 respectively have *au-* and *a-* as their initial vowel. Paulis (1992: 182) holds type 2 in *a-* to be original, because of its supposed relation to the toponym *Atzara*. While we cannot be certain of the idea that *Atzara* was indeed named after this plant, I agree that type 2 in *au-* may be an innovation, since original *-au-* (both stressed and unstressed) should regularly have yielded *-a-* (Wagner 1941a: 13–14, 22–23). Initial *au-* can therefore only be the result of the loss of an intervocalic voiced stop, or due to secondary addition of *a-* to *u-* after the common Sardinian shift from **au-* to *a-* was complete.

One possible solution to unite the three different variants could be to reconstruct **uθθára* as the original form of this word. In type 2, the *a-* may have been reinterpreted from the feminine definite article: **sa uθθára* >> *s'auttsára*. In type 3, pretonic **u-* > *a-* could have taken place due to assimilation to the stressed *-á-* in the following syllable. In type 1, Paulis' proposed analogy to *lúži* 'light' is easier to understand if the two words' vowels already matched. It is problematic that a variant with initial *u-* is not actually attested however. Finally, it is striking that the forms containing *-l-* (e.g. *luttsára*; type 1) are geographically rather dispersed

[115] Wagner (DES I: 153) notes that these words also refer to smilax (*Smilax aspera*), but Paulis (1992: 182) doubts this.

[116] I have not been able to confirm this property elsewhere.

(from Morgongiori in the west to the Ogliastra in the east), which raises the question whether an *-l-* was perhaps present in some pre-Roman forms. Yet, loss of **-l-* cannot easily be explained for the forms of type 2 and 3 either.[117]

3.3.20 Tree heather (*Erica arborea*)$_2$

Reconstruction	Forms
***gillóstr-**	DES (I: 578): ***giḍḍòstre*** (1: Oliena), ***giḍḍóstru*** (2: Orgosolo, 3: Olzai), [***sa***] ***ʔiḍḍòstra*** (4: Fonni), [***su***] ***iḍḍòstro*** (5: Busachi) 'tree heather (*Erica arborea* L.)'
	Wagner (1928: 24): ***giḍḍòstre*** (3: Olzai), [***sa***] ***ʔiḍḍòsta*** (4: Fonni) 'id.'

These words for 'tree heather' are likely of pre-Roman origin (Wagner 1928: 24). The Sardinian forms require an original **gillostr-*, according to Bertoldi (1928: 233) with the same suffix found in *golóstri* etc. 'holly' (cf. § 9.2.9). There is a striking amount of variation in the listed forms' final vowel (*-e*, *-u*, *-a*, *-o*) (cf. § 8.4.1.2). Hubschmid (1953: 29) compares the Sardinian forms to Bq. *gillar*, *gilharre*, *ilharri* etc. 'heather', which would consist of the same root **gill-* with a different suffix *-ar*(*r*)*-*. However, in a later publication, Hubschmid (1958: 214) expresses doubts about the feasibility of his own comparison, as the Basque forms seem to point to an original singleton **-l-* rather than the **-ll-* that is needed for the Sardinian forms (cf. also Michelena 1990: 316–317). Trask (1997: 306) and Santano Moreno (2000: 170) reconstruct Proto-Basque **inar*, which is even further removed from the Sardinian forms. So, although the ending of *giḍḍòstre* etc. would suggest that it belonged to the same pre-Roman layer as *golóstri* etc. 'holly', which has a plausible Basque comparandum in *gorosti* 'id.', a connection between Srd. *giḍḍòstre* etc. and Bq. *gillar* etc. is unlikely.

[117] The forms *urtsúla, úrtsula* 'fragrant virgin's bower (*Clematis flammula*), smilax (*Smilax aspera*)', discussed in DES (II: 564) and by Paulis (1992: 221), have been identified by Blasco Ferrer (2010a: 55) as a ghost word.

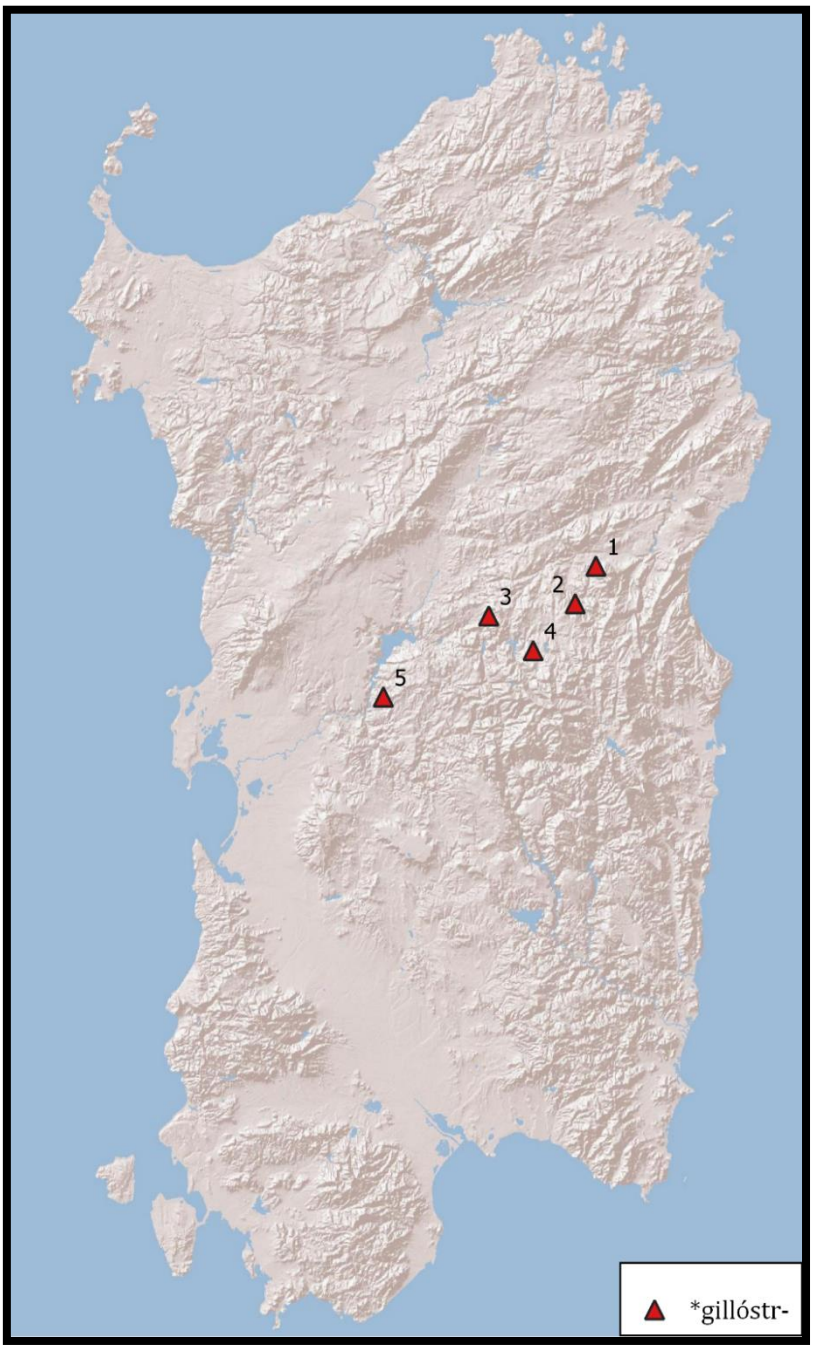

Figure 3.3.20: Tree heather (*Erica scoparia*)

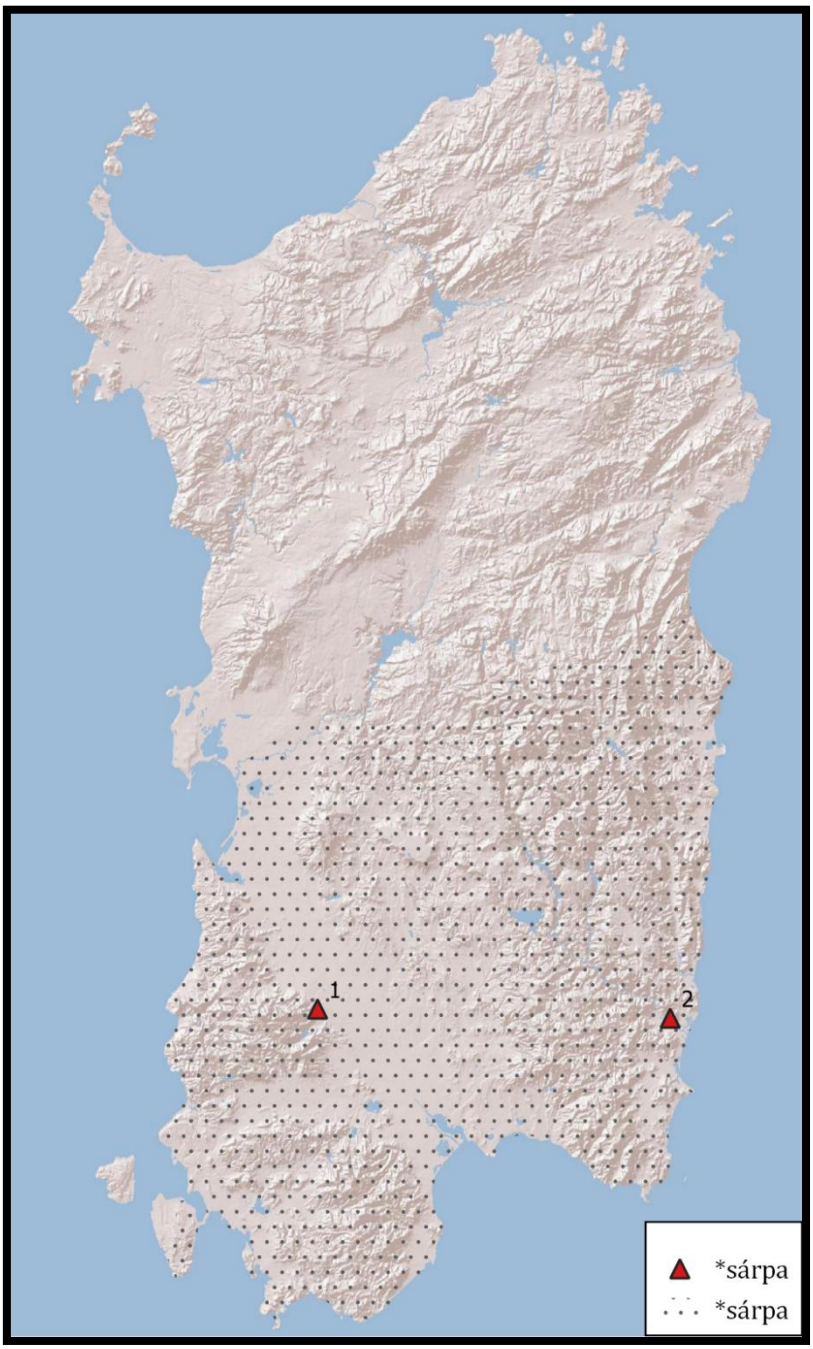

Figure 3.3.21: Willow (*Salix sp.*)

3.3.21 Willow (*Salix sp.*)

Reconstruction	Forms
***sárpa**	DES (II: 384): ***sárpa***, ***ṡráppa*** (Camp.), [***sa***] ***ʂráppa*** (2: Villaputzu) 'various species of willow (*Salix sp.* L.)'
	AIS (600): [***sa***] ***srápa*** (1: Villacidro) 'id.'

Wagner (DES II: 384) is unable to clarify the origin of this word, but he does reject on semantic grounds Jud's (*apud* DES II: 384) proposal that it is somehow related to Sp. *serpa, jerpa* 'infertile grapevine shoot'. Paulis' (1992: 437–438) hypothesis that Camp. *sárpa* etc. is from Cat. *sarga* 'willow' with influence from Sp. *zarpa* 'paw' is not convincing either. Pittau (1995: 227) proposes a connection between Srd. *sárpa* and Lat. *sirpus, scirpus* 'bulrush', both used as weaving materials. A more straightforward solution is to analyze Srd. *sárpa* 'willow' as a deverbal "τομή-type" derivation of Lat. *sarpō* 'to cut off, to prune' (cf. Weiss 2009: 300). The deverbal meaning of "a cutting" could explain both Sardinian. *sarpa* 'willow' and Sp. *serpa* 'infertile grapevine shoot', although the Spanish vocalism remains problematic.[118]

3.3.22 Yew (*Taxus baccata*)

Reconstruction	Forms
***éni**	DES (I: 489): ***éni*** (Ogl., 1: Dorgali, 2: Orgosolo) 'yew (*Taxus baccata* L.)'

This word for 'yew' has been connected by Bertoldi (apud Wagner 1941a: 45) to Bq. (*h*)*agin* 'id.', requiring a development **ágin* > **ágini* > **áini* > *éni* on the Sardinian side. This is theoretically possible, though it must be noted that the contraction *-ai-* > *-e-* is a tendency rather than a pervasive sound law; cf. e.g. *ispaḭnare* (Fonni), Camp. *spaḭnai* 'to spread out' < **expaginare* and Camp. *máiða* 'type of vessel' < *magida*, where it does not happen (DES I: 673; II: 55; cf. Wolf 1998a: 218). Even though it is attractive to provide these Sardinian forms with a Basque cognate, **ágin* is only one of several possible reconstructions for the Sardinian forms, and the correspondence is essentially limited to a single consonant. It is therefore no more than a possible instance of a Basque-Sardinian lexical correspondence (cf. § 10.3). Hubschmid's (1953: 25–26) and Alessio's (1957: 243) connection of these words to Tashelhiyt Berber forms like Ihahan *taga*, Ouarzazate *tiqqi*, Igliwa *tiqi* etc. 'juniper (*Juniperus sp.* L.)' (cf. Laoust 1920: 490), and to Lat. *acinus* 'berry', are rather too far-fetched both semantically and phonetically.

[118] Cf. FEW (XI: 234–236) for formally similar derivations from this verb, with the meaning 'billhook'.

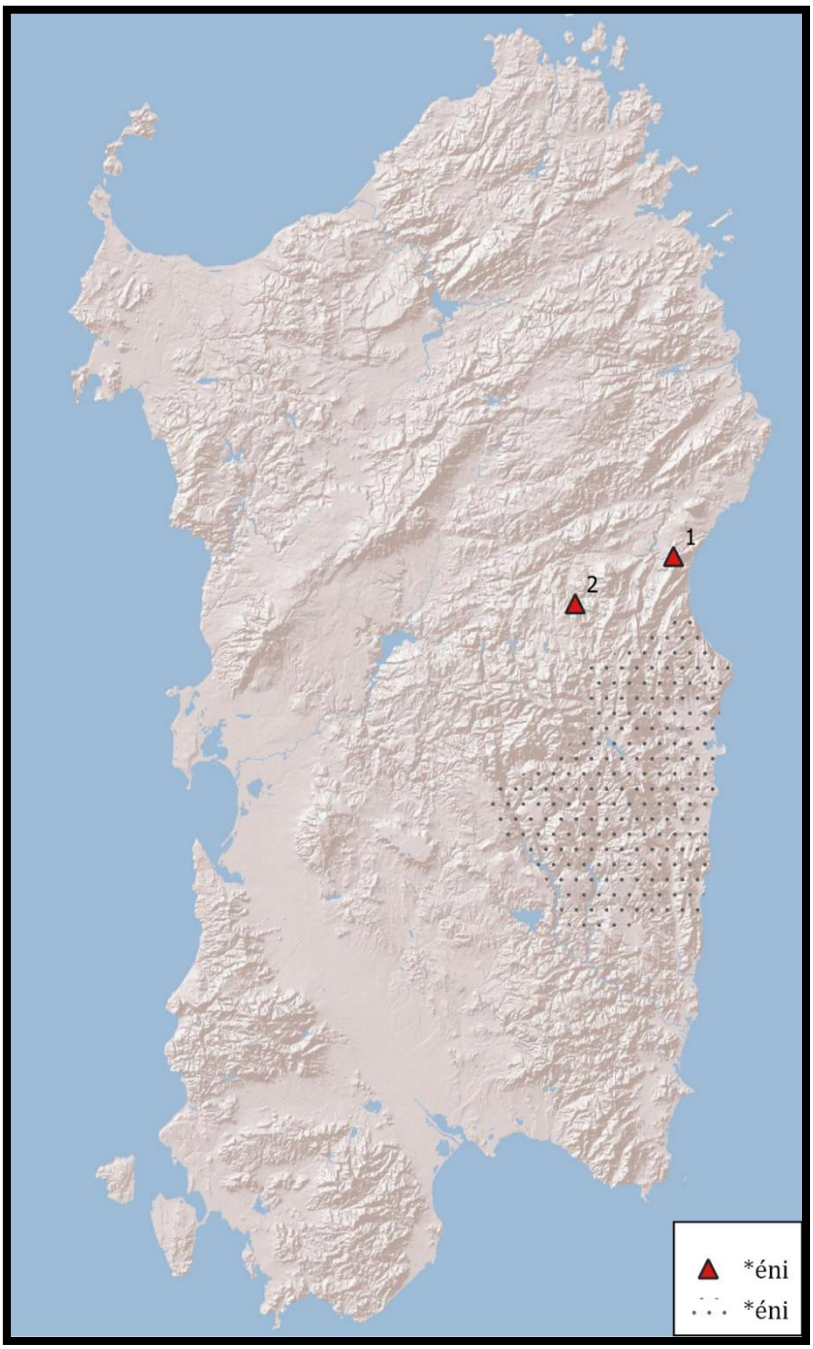

Figure 3.3.22: Yew (*Taxus baccata*)

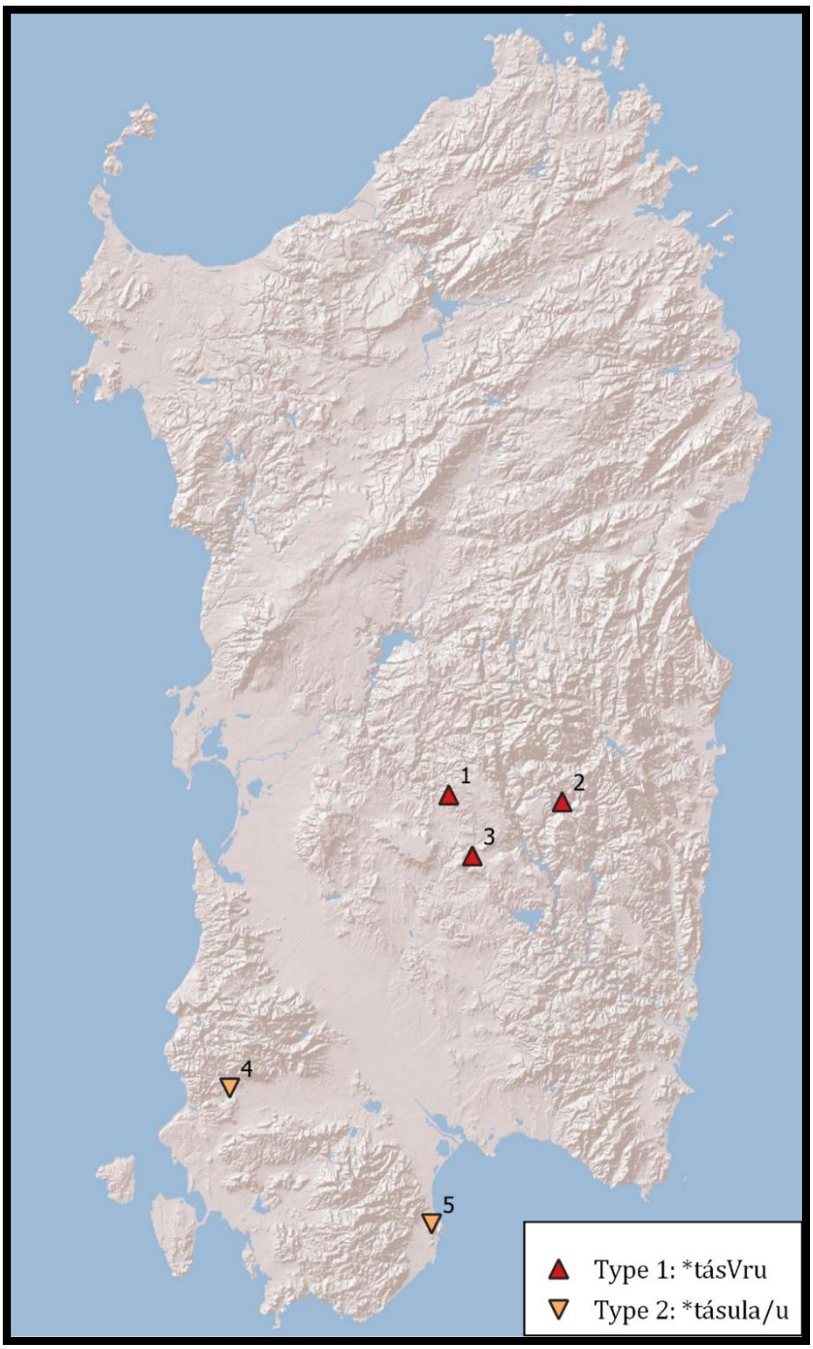

Figure 3.3.23: Yew (*Taxus baccata*); buckthorn (*Rhamnus alaternus*); spindle (*Euonymus europaeus*)

3.3.23 Yew (*Taxus baccata*); buckthorn (*Rhamnus alaternus*); spindle (*Euonymus europaeus*)

Reconstruction	Forms
1: *tásVru	DES (II: 467): ***táširu*** (1: Laconi), ***tášuru*** (2: Seui) 'yew (*Taxus baccata* L.)'; ***tášuru*** (3: Isili) 'buckthorn (*Rhamnus alaternus* L.)'; ***tášaru*** (Trexenta reg.) 'spindle (*Euonymus europaeus* L.)'
2: *tásula/u	DES (II: 467): ***tášua*** (4: Iglesias), ***tášulu*** (5: Sarroch) 'yew (*Taxus baccata*)'

Yew, buckthorn and spindle are all shrubs or small trees, with dark-green glossy leaves (or in the case of yew, needles) and fleshy, round, red-coloured fruits; both yew and spindle are poisonous (cf. Paulis 1992: 407–410). The fact that the meaning 'yew' is attested most widely, suggests that the meanings 'buckthorn' and 'spindle', which occur locally, are secondary from original 'yew'. The forms listed above are also reminiscent of the inherited word for 'yew': Log. *tássu* < Lat. *taxus* 'yew' (DES II: 468). However, as Wagner (DES II: 467) observes, *tášaru* etc. cannot be derived from Lat. *taxus*, because of the voiced *-š-*.

Paulis (1992: 410) proposes a derivation from Lat. *Thasia*. This form is discussed in Pliny's Naturalis Historia as a kind of laurel (viz. "*laurus Thasia*"), but it probably refers to 'butcher's broom (*Ruscus aculeatus*)' (cf. André 1985: 140).[119] Although the semantic shift from 'butcher's broom' to the several plant species denoted by *tášaru* etc. is imaginable (butcher's broom too has dark-green waxy leaves and round red berries), and even though the presence of *-si̯-* could explain the voiced *-š-* in the Sardinian forms (Wagner 1941a: 150), this etymology does raise some questions. The Latin adjective *Thasius* is generally taken to refer to the Greek island of Thasos, in the northern part of the Aegean sea (Lewis & Short 1879: s.v. *Thasus*). It is unclear what exactly would be the connection between a Greek island and a species of butcher's broom. If we accept this meaning at face value though, it still remains puzzling that this specific word survived on Sardinia only, where its meaning shifted to 'yew'. What makes it even more difficult, is the presence of what appears to be a suffix *-́Vr-* (where *-V-* stands for any vowel). This suffix is not present on the proposed Latin comparandum, nor can it be explained in terms of inherited Latin/Romance material. Paulis (1992: 410) explains this ending as an adaptation of the inherited diminutive suffix **-ulu* (as attested in the forms in type 2) to the ending of a supposed pre-Indo-European

[119] André (1985: 149) mentions *Ruscus angustifolia*, which does not exist. However, *Ruscus aculeatus* var. *angustifolius* Boiss. is a synonym of *Ruscus aculeatus* (WFO 2023: s.v. *Ruscus aculeatus*).

word **ebura* 'yew' (as per Alessio 1957: 263). This idea is problematic, as there is no actual evidence for the existence of this word on Sardinia.[120]

As there is no reason to assume the existence of **ebura* 'yew', this can not be used to explain the presence of **-Vr-* on the actually attested forms *tášaru*, *-uru*, *-iru*. Moreover, the presence of diminutive *-ula/u* on *tášulu* and *tášua* is easier to explain as the result of reinterpretation of non-inherited **-Vr-*, than the other way around. If *tášuru* etc. 'yew', which is the *lectio difficilior*, is to be explained from Lat. [*laurus*] *Thasia*, we need to accept the secondary addition of a pre-Roman Sardinian suffix *-Vr-* (§ 9.2.6.2). This word may be thus an instance of an inherited Latin word that received native Sardinian morphology. Finally, Paulis (1992: 407, 410) notes that Srd. *tášuru* etc. cannot be separated from Lat. *taxus* 'yew', despite not being regularly inherited. However, *taxus* does not seem to play a role in any part of his account of the origin of Srd. *tášuru* 'yew', which only involves a semantic shift from 'butcher's broom' to 'yew' and addition of a pre-Roman suffix to inherited *Thasia*. Perhaps he believes the necessary semantic shift may have been aided by the similarity of *tássu* 'yew' < *taxus* to *tášuru*, but this is not strictly necessary.

3.4 Plant parts

3.4.1 Awn of wild oats (*Avena fatua*)

Reconstruction	**Forms**
1: *salaCúspe	DES (II: 378): ***salaúspu*** (1: Osilo, 3: Bonorva), ***salaúspe*** (3: Bonorva) 'awn of wild oats (*Avena fatua* L.) and ryegrass (*Lolium sp.* L.) that attaches itself to shoes and pants'
2: *salaCúspe	***saraúipu*** (2: Sassari) 'butcher's broom (*Ruscus aculeatus* L.)'

[120] Alessio's (1957) idea that prior to the advent of Indo-European languages in Europe, a form **ebura* must have denoted the 'yew' over a vast swath of the European continent, is based on the existence of OIr. *ibhar* 'yew' and a series of assumptions on historically attested toponyms. In light of our general lack of knowledge about the original semantics and *Benennungsmotive* of these presumably pre-Indo-European toponyms, I tend to disagree with much of the methodology employed by Alessio, as well as with his conclusions (cf. § 1.3.2, § 2.1.2.2.1).

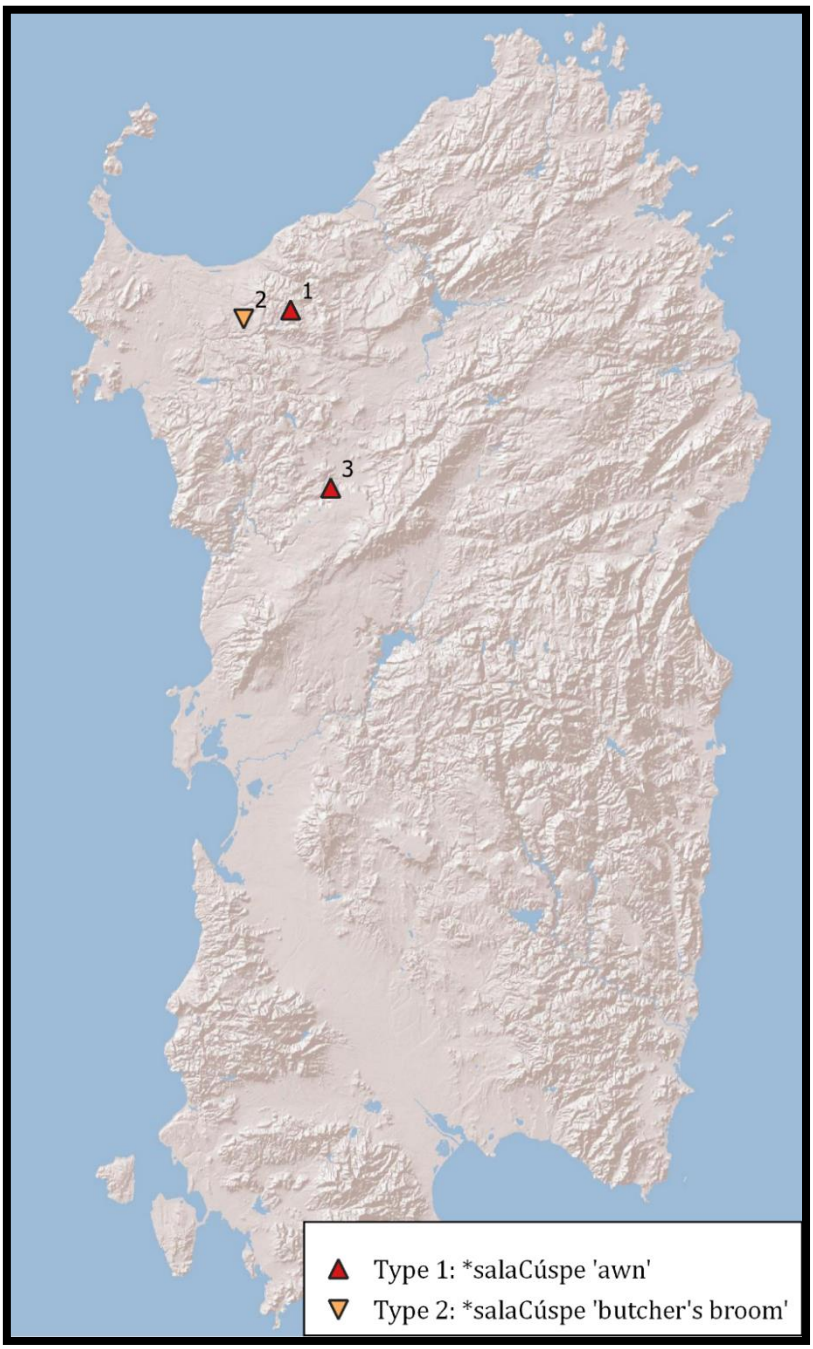

Figure 3.4.1: Awn of wild oats (*Avena fatua*)

The difference in meaning between *salaúspu/-e* 'awn of wild oats' in various Logudorese dialects and *saraúipu* 'butcher's broom' in Sassarese is so difficult to understand, that one is tempted to see it as a mistake of some kind.[121] Wagner (DES II: 378) suggests a pre-Roman origin for this word. Paulis (1992: 264), on the other hand, proposes an etymology in **ala-ulpe*. This compound, to which the article *s-* was added by misinterpretation of the definite article *su*, would consist of *ala* 'wing', due to the fact that this plant's seeds disperse through the air, and of **gulpe* 'fox' < Lat. *vulpēs* to indicate that it is "wild" oats as opposed to domesticated oats. Such a semantic development, i.e. from a compound literally meaning 'fox-wing' to 'awn of wild oats', requires a few steps too many to be convincing. Moreover, we would also need to assume that **ala-ulpe*, which would have been feminine, secondarily acquired the masculine gender that is attested for *salaúspu*. Phonologically, a derivation from *vulpēs* 'fox' is all but obligatory, as the cluster *-sp-* (and *-ip-* in Sassarese) can go back to **-rp-*, **-lp-* or **-sp-* (Wagner 1941a: 190). Although there is not much positive evidence for Wagner's classification of this word as originally pre-Roman, I do not think Paulis' proposed Latin etymology is to be favored. The variant with final *-e* is most likely original, whereas *-u* was likely introduced because of the masculine gender of this word.

3.4.2 Branch

Reconstruction	**Forms**
1: *kárva/ *kárba	DES (I: 311): ***kárva*** (Nuor., 1: Dorgali), [***sa***] ***hárβa*** (2: Urzulei), [***una***] ***hárva*** (3: Baunei) 'tree branch'

This word for 'branch', which in Sardinia is restricted to the Nuorese dialects, has been compared to a range of phonologically and semantically similar words outside Sardinia, especially in northern Spain and southern France. Wagner (1932: 231; DES I: 311) compares Ast. *garbu* 'small firewood', *carba* (Lena) 'scrubland', *garbu*, *gárabu* (Cabranes) 'small branch used as firewood'; Sp. *escarabajas* (Salamanca) 'small firewood used to light a fire'; Bq. *karbasta* 'stick with branches', *karbaza* 'stem, trunk'; Prov. ("Alpine") *garbo* 'hollow tree trunk'.

[121] The two types in this lemma do not reflect different formations, but rather the vast difference in meaning between 'wild oats' on the one hand and 'butcher's broom' on the other. For a similar apparent homonymy, cf. *alášu* 'Bermuda grass (*Cynodon dactylon*)' vs. *aláše* 'butcher's broom (*Ruscus aculeatus*)' (§ 3.2.1), which may however originally be two unrelated words.

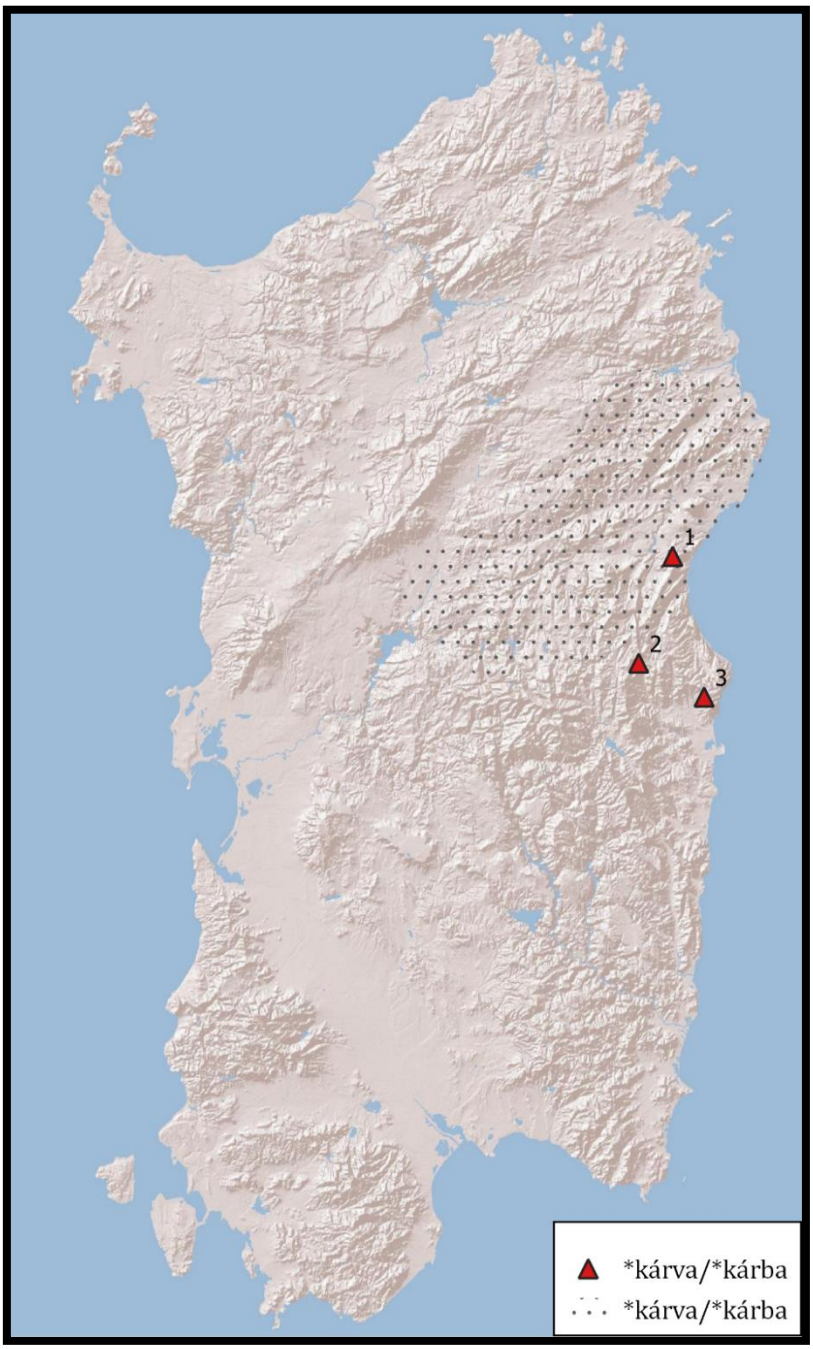

Figure 3.4.2: Branch

DCEC (I: 856) adds Sp. *carba* (Salamanca) 'scrubland full of coarse oak trees', *gáraba* 'gorse (*Ulex sp.* L.) in its most woody part'; Ast. *carbana* 'sorrel (*Rumex sp.* L.)', *garbu, garbucu* (Rato) 'small firewood'; Leon. *carvayo, -ajo* 'oak'; Pg. *carvalho* 'oak', *gravulha* 'dry pine needles', *garvalha, gravalho* 'fallen branches and leaves'; Gal. *carvallo* 'oak', *gar*(*r*)*abullos, carabullos, garabellos, garamallos, garrampallos, garabujos, garabizos* 'dry tips of tree branches used as small firewood'; Gasc. *garbout* (Lavedan) 'stunted plant', *garboucha* 'copse of trees', *garbalhe* (dep. Landes) 'pine needles'; Lang. *garbasso, garbassat* 'young oak'. From northern Italian varieties, Jud (1931: 439) compares *skalvắ* (Pescarolo ed Uniti), *ṣkravắ* (Carpaneto Piacentino) etc.[122]

The meanings of these forms can be divided into two general groups: one the one hand those referring to 'oaks', on the other hand those referring to 'dry branches, leaves etc.' and 'firewood' etc. In the Ibero-Romance languages, where both meanings occur, the forms meaning 'oak' etc. most often have a voiceless *c-*/*k-*, whereas the words referring to dry combustible plant material generally start in voiced *g-*; cf. Pg. *carvalho* 'oak' vs. *gravulha* 'dry pine needles', *garvalha* 'fallen branches and leaves'. This distribution could indicate that we are dealing with two etymologically unrelated sets of words that have phonologically influenced each other. It is attractive both phonologically and semantically to connect the Western Romance forms in the meaning 'dry branches, leaves; firewood' to Srd. *kárva* 'branch'. Sardinian 'branch' is likely the result a semantic broadening of the meaning attested in the Iberian and Occitan dialects.

Phonologically, the Sardinian forms presuppose a base **kárva* or **kárba*. The Romance forms cannot be derived from Latin, and given their Mediterranean distribution and irregular phonetic correspondences a pre-Roman origin is plausible. If the Sardinian forms are indeed related to those found in the Iberian Peninsula and southern France, the wide distribution of **karb-*/**garb-* has interesting implications for prehistoric language relationships in the Mediterranean. Similar in form and meaning is also Lat. *carbō, -ōnis* 'piece of charcoal' (de Vaan 2008: 91–92), which also refers to burnable material and which lacks a good Indo-European etymology. If these comparisons are correct, we are dealing with a family of words for 'fuel for fire/firewood' that is represented in the Iberian Peninsula, southern France, Sardinia, and the Italian Peninsula. The distribution of this word is similar to that of several other pre-Roman words attested in Sardinian and could be indicative of prehistoric linguistic ties in the Mediterranean basin (§ 11.2.1).

[122] Cf. AIS (543) and Hubschmid (1953: 19-20).

Finally, regarding the comparison of Bq. *karbasta* etc. 'stick with branches' to this group of potentially pre-Roman words, it should be noted that there is also another Basque word *abar* that covers more or less the same range of meanings as the forms containing **karb-/*garb-* discussed above. Lakarra et al. (2019: 3–8) prefer to derive Bq. *garbasta*, *karbazta*, *arbazta* etc. from Bq. *arba* 'branch' etc., which would be a metathesized form of *abar*. Since the addition of the initial velar in *karbazta* etc. would be irregular, and *abar* cannot explain the various other Mediterranean forms of the type **karva*/**karba*, perhaps the Basque forms *arba*, *arbazta*, *karbazta* etc. are the result of a conflation of inherited Bq. *abar* and the various forms of the type **karva/*karba* in the surrounding Romance languages.

3.4.3 Juniper berry (*Juniperus sp.*)

Reconstruction	**Forms**
****bollVro***	DES (I: 215): ***bòḍḍoro*** (1: Dorgali), ***bòḍḍero*** (2: Baunei) 'juniper berry' (*Juniperus sp.* L.)

The final vowel *-o* of these forms is difficult to explain from an inherited point of view (§ 8.1.5.2). DES (I: 215) and Pittau (1995: 202) compare it to Sic. *bóḍḍaru* 'berry, grain; bump' and Tusc. *bòllora* 'gall (on plants)'. These comparanda are formally and semantically similar, and all three may go back to a shared pre-Roman origin. However, Paulis (1992: 444) suggests that this and other names for 'juniper berry' in Sardinian may originally have referred to generic 'round things'.

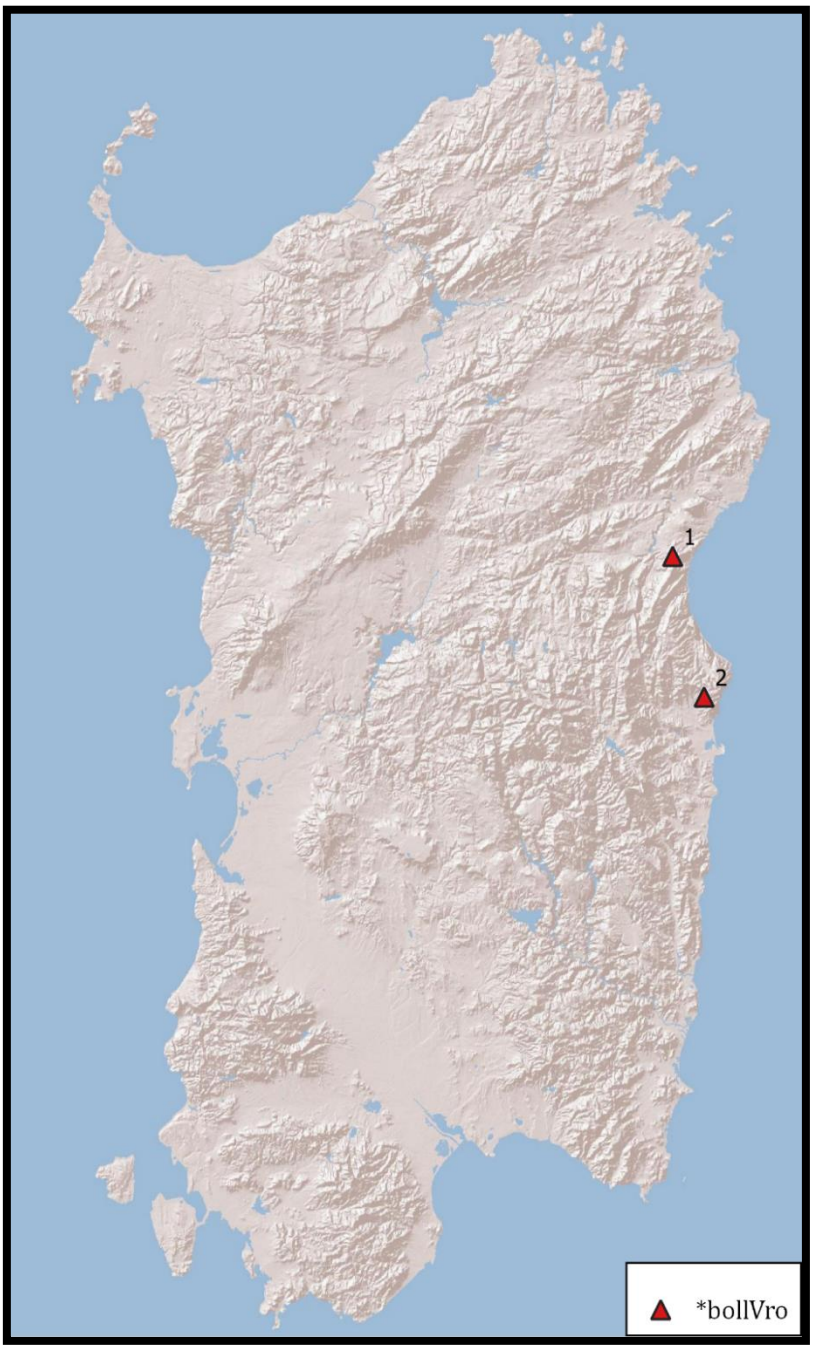

Figure 3.4.3: Juniper berry (*Juniperus sp.*)

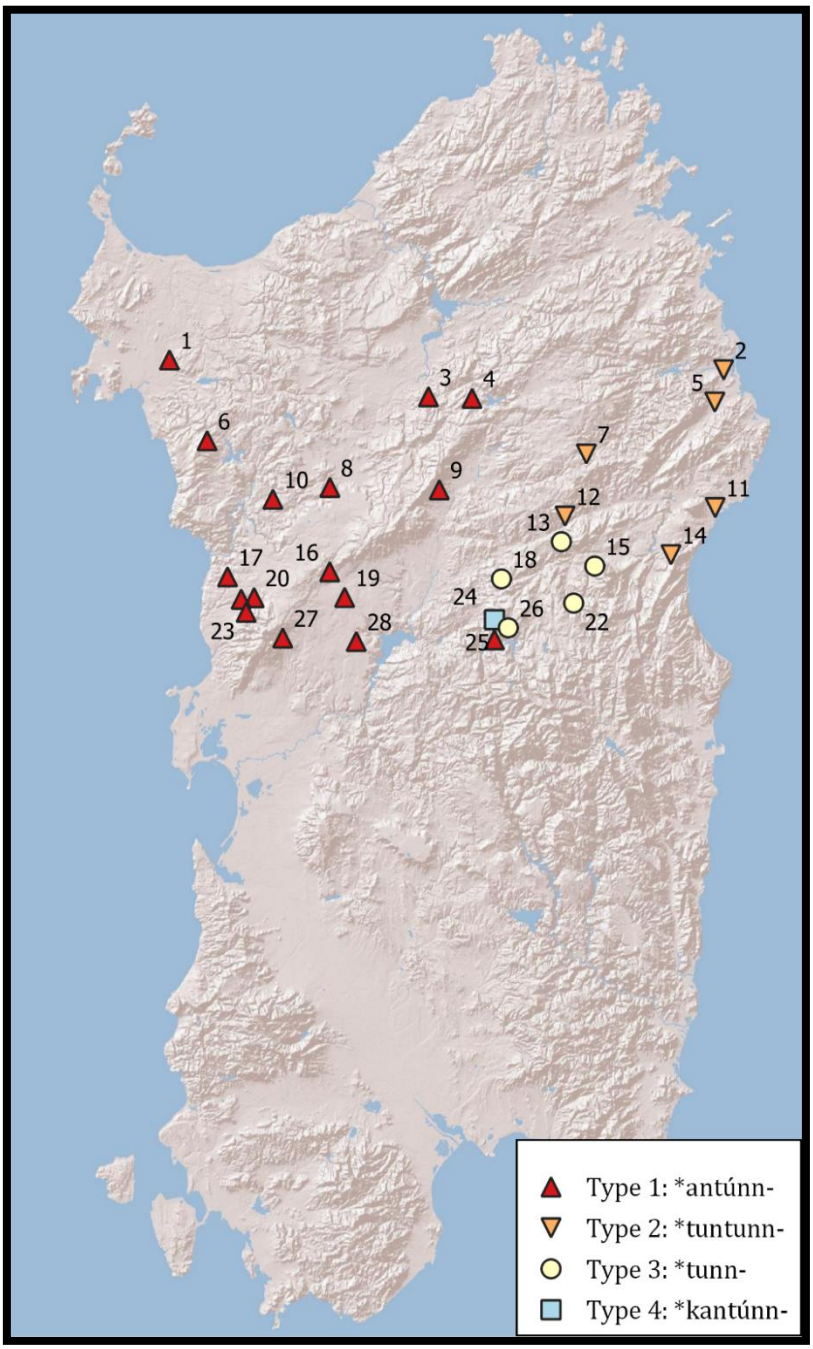

Figure 3.4.4: Mushroom

3.4.4 Mushroom[123]

Reconstruction	Forms
**1: ****antúnn-*	DES (II: 533–534): ***antúnnu*** (1: Olmedo, 3: Ozieri, 4: Pattada, 6: Villanova Monteleone, 10: Padria, 20: Scano di Montiferro, 23: Cuglieri, 21: Sennariolo, 17: Tresnuraghes), ***antúnna*** (8: Bonorva, 9: Bono, 16: Macomer, 19: Borore, 25: Ollolai, 27: Santu Lussurgiu, 28: Norbello) 'mushroom'
	AIS (621): [***s***] ***antúnnu*** (Santu Lussurgiu) 'id.'
**2: ****tuntunn-*	DES (II: 533–534): ***tuntunnúu*** (2: Posada, 14: Dorgali), ***tuntúnnu*** (5: Siniscola, 7: Bitti, 11: Orosei, 12: Lollove) 'id.'
	AIS (621): [***su***] ***tuntúnnu*** (Bitti), [***su***] ***tuntunnúu*** (Dorgali) 'id.'
	Puddu (2023: s.v. *tontonnúu*): ***tontonnúu***, ***tuntúnna***, ***tuntunnína***, ***tuntunnúu***, ***tuntúnnu*** 'id.'
**3: ****tunn-*	DES (II: 533–534): ***túnn̥u*** (13: Nuoro), ***tuññúu*** (15: Oliena), ***túnna*** (18: Orani), ***tunnúu*** (22: Orgosolo, 26: Gavoi) 'id.'
	Puddu (2023: s.v. *túgnu*): ***túgnu***, ***túnna***, ***tunnío***, ***tunnúu***, ***túnniu*** 'id.'
**4: ****kantúnn-*	DES (II: 533–534): ***kantúnna*** (24: Olzai) 'id.'

All of the forms listed above contain the element *-tunn-*. Four types can be distinguished, on the basis of the elements present before *-tunn-*. All Logudorese dialects have type 1: *antunn-*. Type 2 **tuntunn-* appears to be the reduplicated variant of type 3 **tunn-*; both occur in the Nuorese dialects, with type 2 in a more northern distribution than type 3. Finally, type 4 is only found in Olzai, very close to the isogloss running between the types *antúnn-* and *túnn-*. Given the isolated attestation of type 4, it cannot be ruled out that it is a local and secondary variant of *antúnna*, in which case it does not tell us anything about the presence of a pre-Roman prefix **kV-* that has been identified in other words (cf. Paulis 1992: 19). This leaves us with the opposition between type 1 in *antunn-* and type 3 in *tunn-*.

Paulis (1992: 19) derives the words for 'mushroom' discussed here from Lat. *autumnus* 'autumn'. But even if the phonetic variation can be explained away (the expected outcome is Nuor. ***atúnnu*, Log. ***aðúnnu*), the semantic development from 'autumn' to 'mushroom' without any identifiable derivational devices seems too big a leap. Nevertheless, even if these forms do not directly go back to Lat. *autumnus*, influence of this word is imaginable.

[123] These data are discussed in a similar fashion in Swanenvleugel (2024: 241–242).

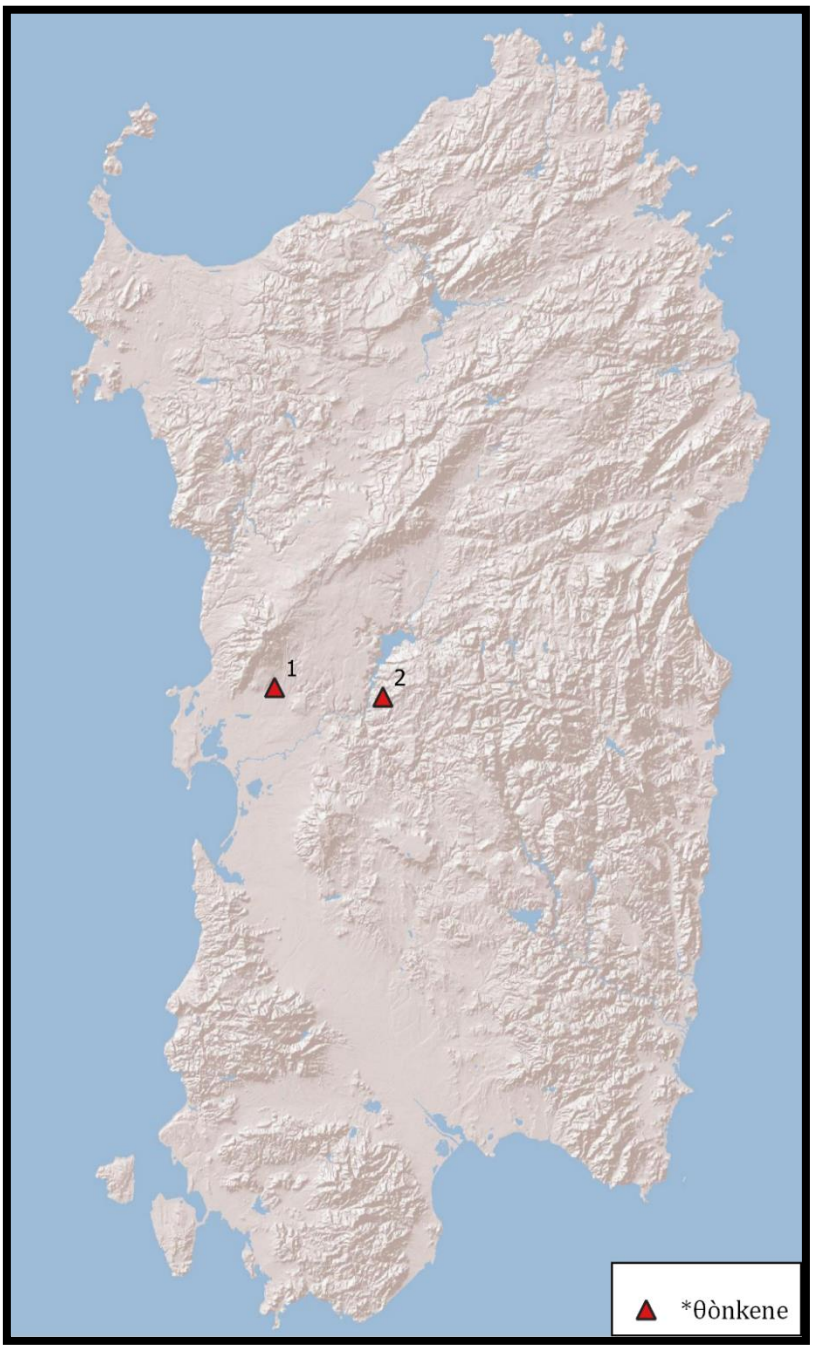

Figure 3.4.5: Resin

Note that the Sardinian words for 'autumn' (Nuor. *atóndzu*, Log. *atúndzu*, Camp. *atúnğu*, *atónğu* etc.) cannot be inherited and were likely remodeled after Sp. *otoño* or It. *autunno* 'autumn' (DES I: 150). Thus, Lat. *autumnus* does not have regular descendants in the Sardinian dialects. The phonological alternations found in the words for 'mushroom', such as prothetic *(k)V-* and variation between final *-u*, *-iu*, *-íu*, are reminiscent of the alternations found in other words suspected of a pre-Roman origin (§ 8.1.5; § 9.1.1). For irregular insertion of *-n-*, cf. Wagner (1941a: 219–223).

3.4.5 Resin

Reconstruction	**Forms**
***ϑònkene**	DES (II: 599): ***tsònki*** (1: Milis), ***tsònkine*** (2: Busachi) 'tree resin'
	AIS (568): ***tsónkine*** (2: Busachi) 'id.'

Both of these forms can go back to **ϑònkine*. Milis *tsònki* < **tsònkîi* < **tsònkini* has regular loss of intervocalic *-n-* and raising of final *-e* to *-i* (cf. e.g. *pắi* < *pane* 'bread'). Paulis (1992: 345) proposes a derivation from *tsònka* 'scops owl (*Otus scops*)' (§ 4.4.11), caused by a complex analogy to the similarity between *kukkumèle* 'resin' and *kukkéḍḍu*, *kukkúi* 'cuckoo (*Cuculus canorus* Linnaeus, 1758)' in various Sardinian dialects. I think this is rather too far-fetched, and it is moreover not clear what the ending *-ine* (< **-ene*) would be. I regard its etymology as not yet clarified.

3.4.6 Sprout, bud, shoot

Reconstruction	**Forms**
***ϑérula**	DES (II: 589): ***tséurra*** (Camp.), ***dzeúrra*** (1: Cagliari), ***ntsaúrru*** (2: Domus de Maria) 'sprout, bud, shoot'

These Sardinian forms go back regularly to original **ϑérula* (DES II: 589). Whereas Wagner (DES II: 589) does not quite know what to derive this from, Paulis (1990: 616) plausibly proposes it to be a loan from Pun. ζερα 'seed',[124] cognate to Hebr. זֶרַע *zéraʿ*, Akk. *zēru*, *zarʾu* etc. 'seed', which was extended in Sardinian Latin with diminutive *-ula*. For the representation of Pun. *z-* by Srd. **ϑ-*, cf. § 8.1.2.

[124] Attested in the compound ζεραφοίστ 'flax seed' (Dioscorides Anazarbei 1907: 177).

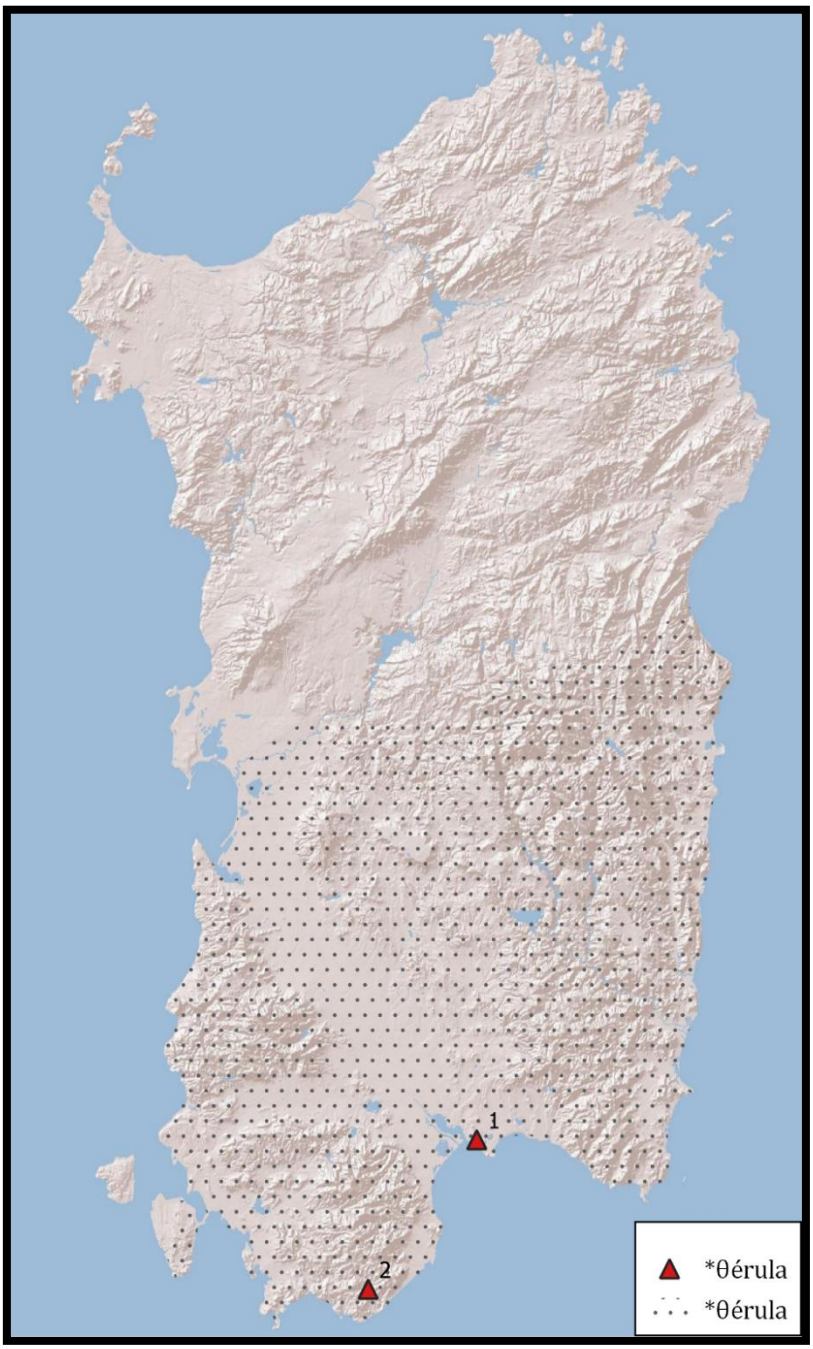

Figure 3.4.6: Sprout, bud, shoot

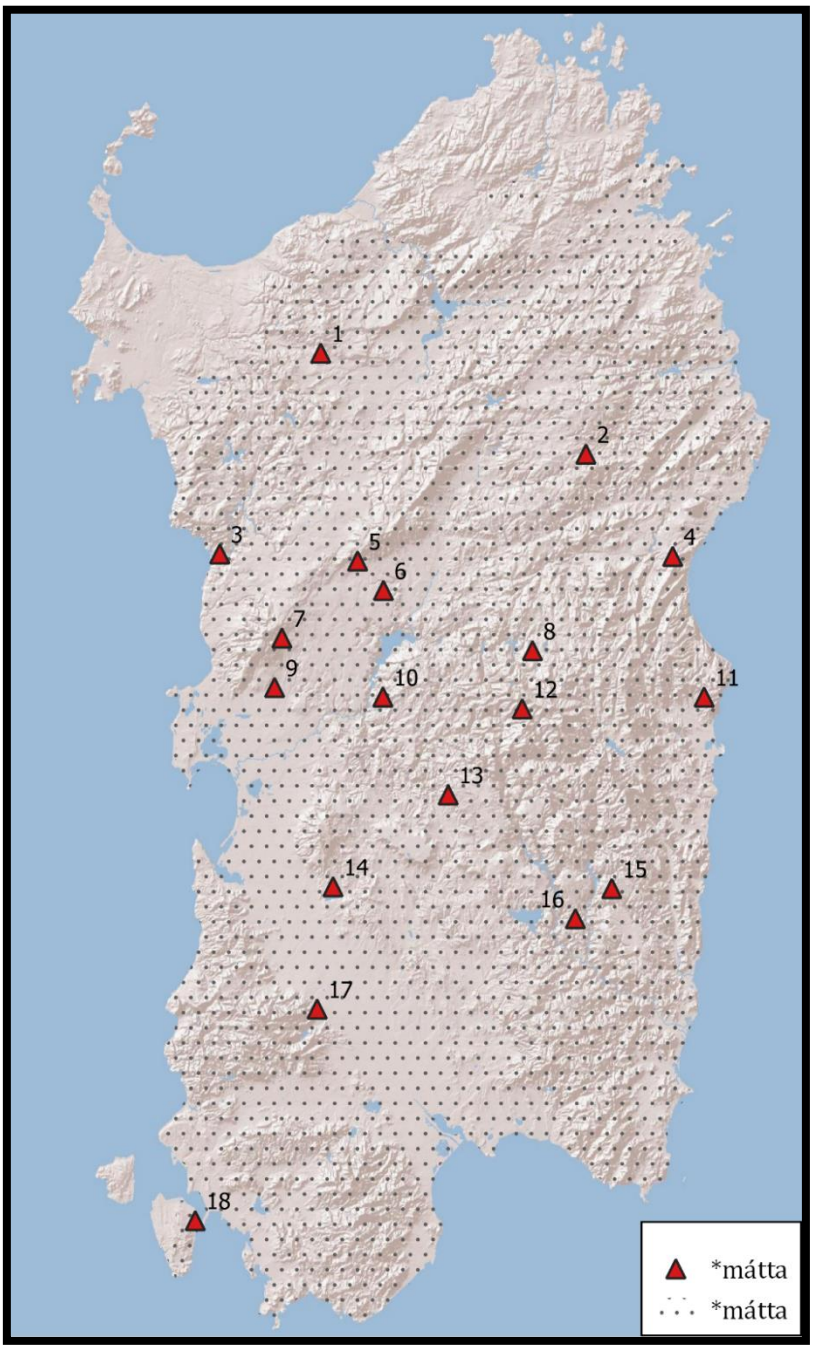

Figure 3.4.7: Tree, shrub

3.4.7 Tree, shrub

Reconstruction	Forms
***mátta**	DES (II: 87): ***mát(t)a*** (Log., Camp.) 'tree, plant, shrub'
	AIS (531, 533): ***mátta(s)*** (4: Dorgali, 7: Santu Lussurgiu, 8: Fonni) 'shrub(s)'; ***mátta*** (9: Milis, 10: Busachi, 11: Baunei, 13: Laconi, 14: Mogoro, 15: Perdasdefogu, 16: Escalaplano, 17: Villacidro, 18: Sant'Antioco) 'tree'; ***mattéðu*** (1: Ploaghe) 'shrub'; ***mattiɣèḍḍas*** (9: Milis, 10: Busachi), ***mattièḍḍas*** (12: Desulo), ***matižèḍḍas*** (13: Laconi, 14: Mogoro, 16: Escalaplano, 17: Villacidro), ***matiǵèḍḍas*** (15: Perdasdefogu) 'shrubs'
	Hubschmid (1953: 33): ***mata*** (2: Bitti, 3: Bosa), ***matta*** (5: Bortigali, 6: Dualchi) 'shrub'

This word has been connected with a large number of forms outside Sardinia, from the Iberian Peninsula, France, Italy and north Africa. The attested meanings range from 'tree, shrub' to 'land covered by shrubs' to 'pile, bundle of wood/plants' to 'herd of animals'. REW (5424) and DCEC (III: 874–875) argue that all these forms go back to (late) Lat. *matta* 'mat', which in turn is perhaps a loan from Punic; cf. Hebr. מִטָּה *mita* 'bed' (DCEC III: 875). In Berber, Hubschmid (1953: 33) compares Snh., Rif. *tamaṭṭa*, 'stack of sheaves that are to be threshed' (cf. Renisio 1932: 385; Serhoual 2002: 329),[125] but this is most likely a loan from Moroccan Ar. *məṭṭa* 'id.' (de Prémare 1999: 222). Regardless of the ultimate origin of Lat. *matta*, it is most likely that the Sardinian forms, as well as all other Romance forms cited by REW and DCEC, simply go back to Latin, with various semantic developments.

3.4.8 Walnut (*Juglans regia*)

Reconstruction	Forms
***kòkkoro**	DES (I: 356): ***kòkkoro*** (1: Sorgono, 2: Tonara, 3: Desulo, 4: Atzara, 5: Belvì, 6: Aritzo, 8: Seulo), ***kòkkolo*** (7: Meana Sardo) 'walnut (*Juglans regia* L.)'

These forms have been compared to Tusc. *cóccoro* 'big, round hazelnut' (Castelnuovo di Garfagnana), Nap. *còccola* 'walnut husk', Laz. *kùkkẹrẹ* 'half of a walnut husk' (Castro dei Volsci), Romagnol *cocla* 'walnut'; Arag. *coca* 'walnut'; Pg. *cócora* 'cooked chestnut' etc. (DES I: 356; Pittau 1995: 207).

[125] Cf. also Argiolas (2020: 49).

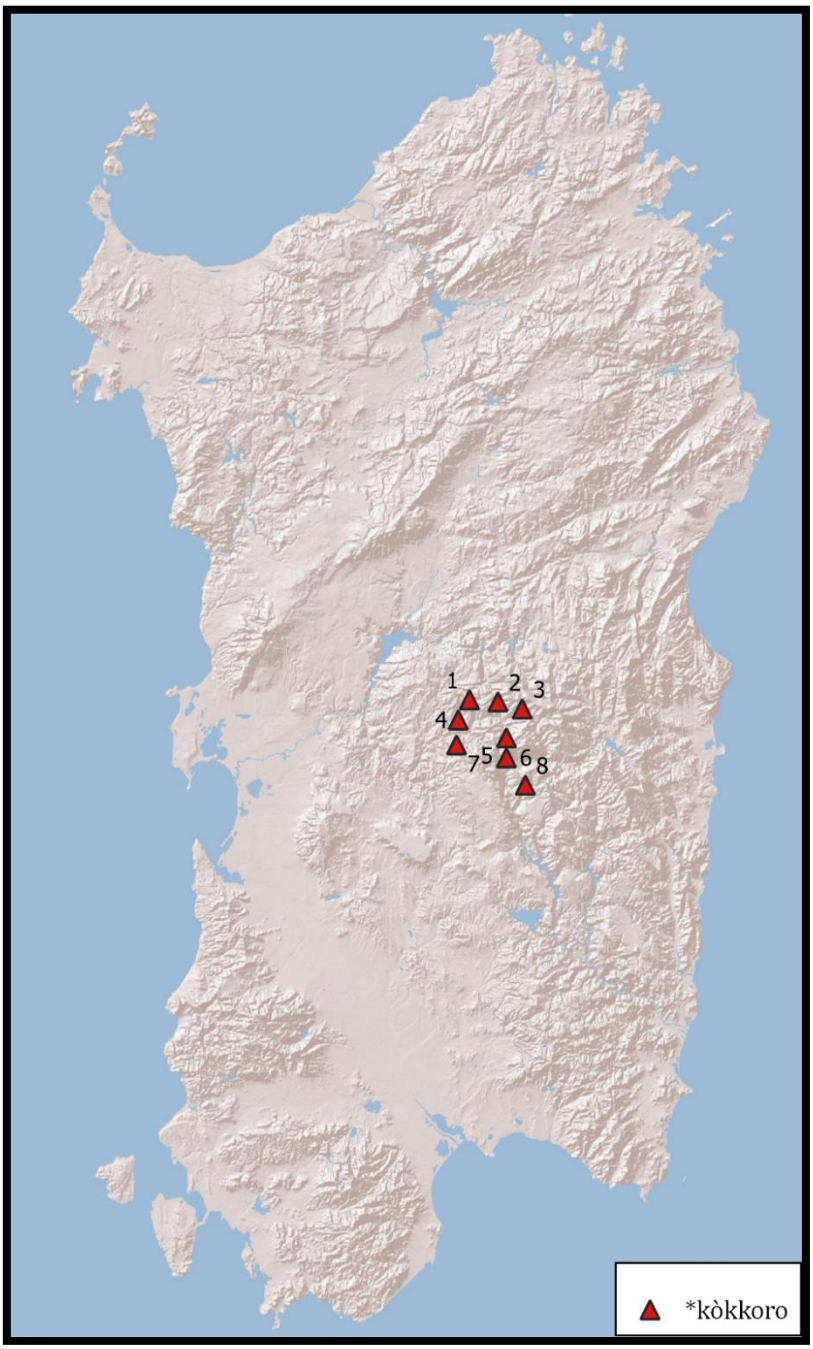

Figure 3.4.8: Walnut (*Juglans regia*)

However, on the basis of similar Sardinian forms that refer to generic round objects (e.g. Ogl., Barb. *kokoròni* 'round object; bump' etc.), Wagner (DES I: 357) proposes the Sardinian forms to ultimately go back to Lat. *coccum* 'scarlet oak gall'. The non-inherited ending *-oro* could be an instance of productivity of an originally pre-Roman suffix (cf. § 9.2.6.2.5). The ending *-olo* in *kòkkolo* (7: Meana Sardo) is probably secondary.

4 Fauna

Like plant names, animal names figure prominently among the Sardinian lexicon of putative pre-Roman origin. The lemmata discussed in this chapter are divided into invertebrates (§ 4.1), fish (§ 4.2), reptiles (§ 4.3), birds (§ 4.4), and mammals (§ 4.5). The scientific names of the animal species discussed have been provided wherever possible.[126]

4.1 Invertebrates

4.1.1 Ant

Reconstruction	Forms
***kaḷ̰iar-**	DES (I: 270): ***kallarèḍḍa*** (1: San Nicolò Gerrei) 'small light-coloured ant ("*piccola formica bionda*")', [***sa vrúmmiɣa***] ***ɣallái*** (2: Sant'Antioco) 'small black ant'
	Puddu (2023: s.v. *callarèdha*): ***callài***, ***callagàlla***, ***callarèdha*** 'type of ant; small ant'

These forms appear to go back to **kaḷ̰iáre/i*, which has been extended with diminutive *-èḍḍa* in San Nicolò Gerrei (1). Wagner (DES I: 270) suspects that this is a pre-Roman word. Apart from the lack of an etymology, there is no positive evidence for this.

[126] All zoological binomial names follow the naming conventions of the International Code of Zoological Nomenclature (ICZN 2024).

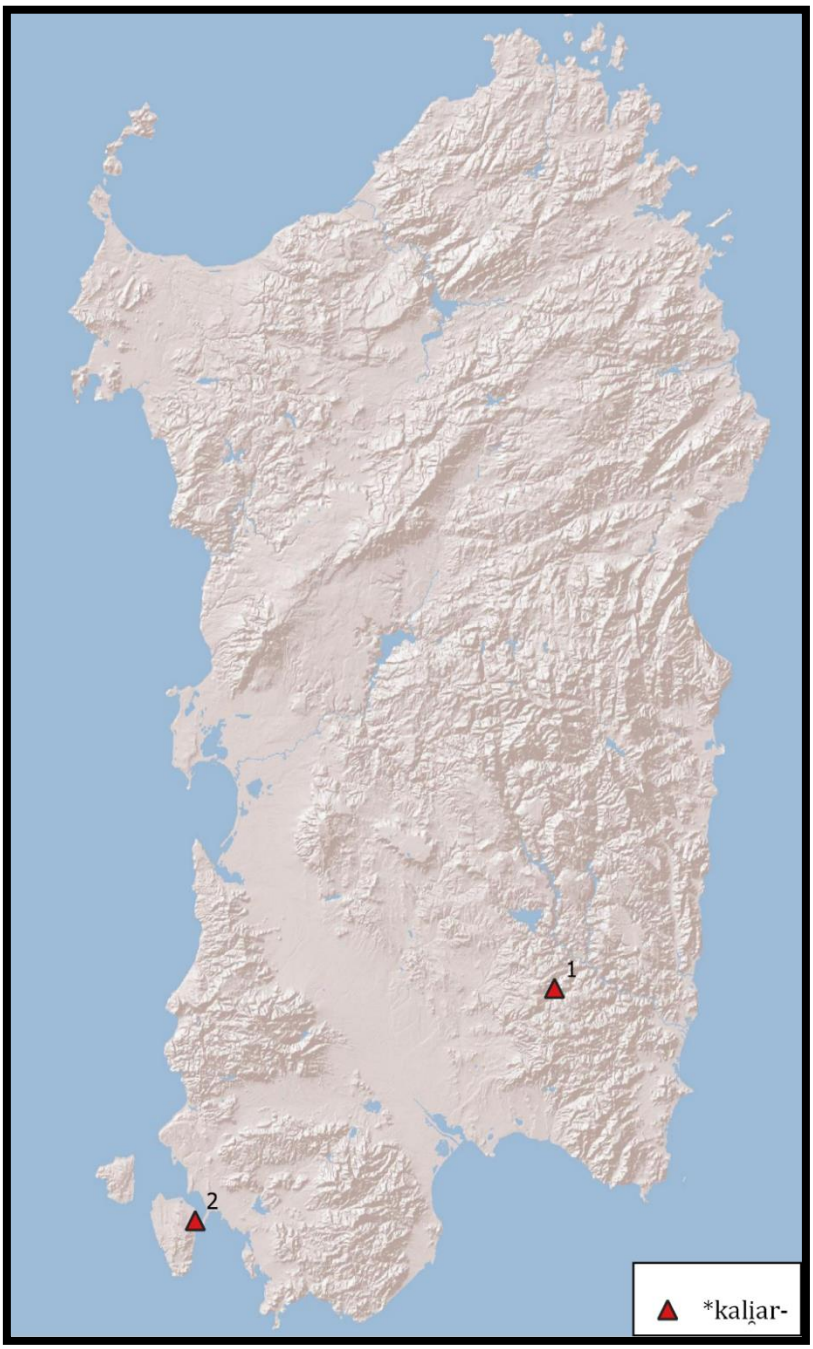

Figure 4.1.1: Ant

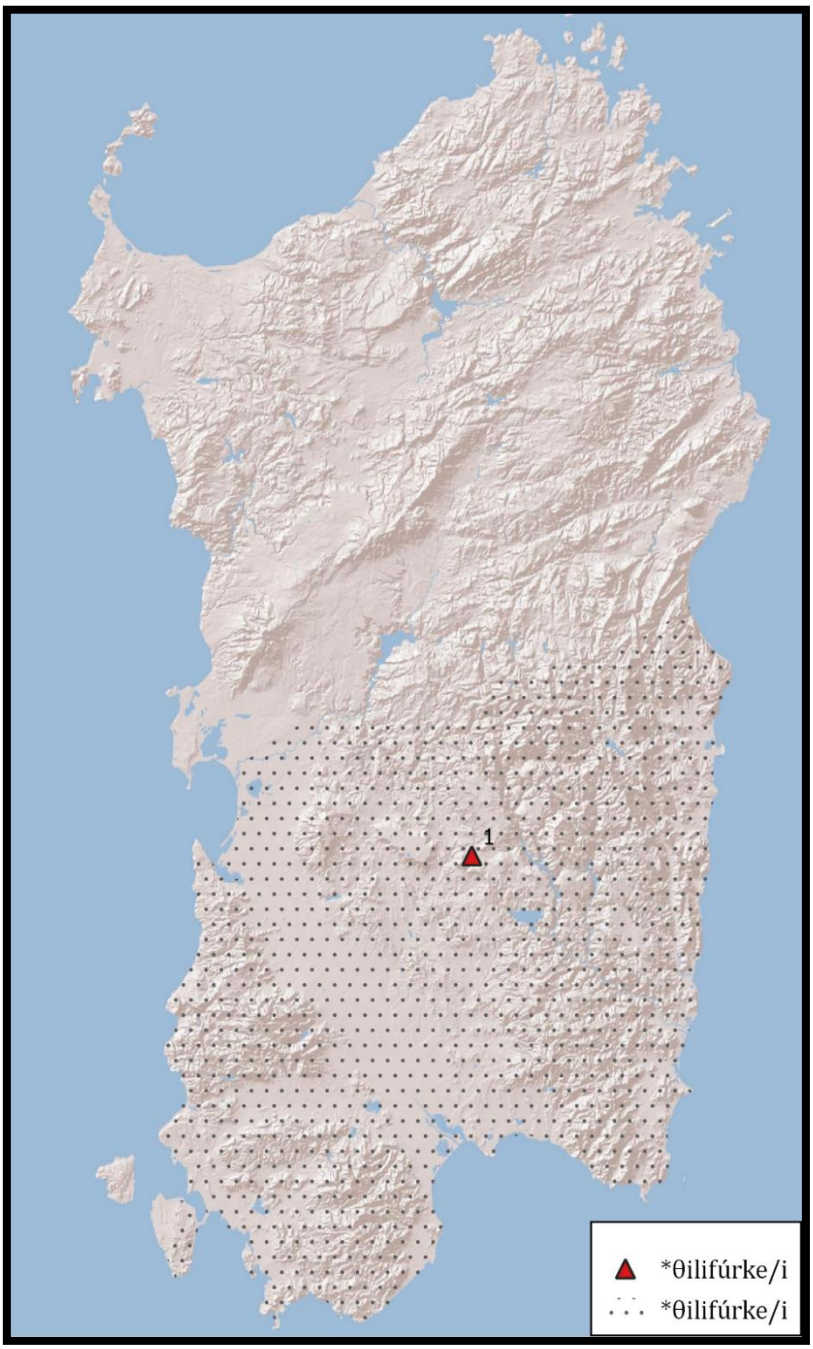

Figure 4.1.2: Backswimmer (*Notonecta glauca*)

4.1.2 Backswimmer (*Notonecta glauca*)

Reconstruction	**Forms**
***ϑilifúrke/i**	DES (II: 595): ***tsirifrúči*** (Camp.) 'backswimmer (*Notonecta glauca* Linnaeus, 1758)'
	Marcialis (2005: 70): ***zirifrúci*** (1: Isili) 'id.'

DES (II: 595) does not attribute this word to a specific location, other than Campidanese in general. Marcialis (2005: 70), however, specifically mentions that it is a word from Isili. Wagner (DES II: 595) suggests that the second part of this word (*-frúči* << **-fúrči*) is derived from Srd. *fúrka* 'pitchfork' (cf. mainly Camp. dim. *furčíḍḍa* 'fork' and *fručiḍḍai* 'to bifurcate'; DES I: 559).[127] Although this is semantically imaginable on the basis of the shape of this animal, it is not clear why the final vowel would become *-i* or *-e* (both of which could be the source of *-i*). The initial part *tsiri-* is reminiscent of various other presumed pre-Roman words starting in *ϑili-*, *tili-*, *tsili-* etc., that often refer to native fauna (§ 9.1.2). In Isili, original intervocalic *-l-* is realized as pharyngeal /ʁ/ (Contini 1987: 559), which may have been transcribed by Marcialis as *-r-*.[128] In light of the typically pre-Roman element *tsiri-* < **ϑili-* and the formal difficulties in regularly deriving *-fruči* from Lat. *furca* 'fork', I suspect that this word is the result of a folk-etymological alteration of a substrate word starting in **ϑili-* and perhaps ending in *-(č)i*, with secondary remodeling after *fúrka*, dim. *furčíḍḍa* etc. 'fork'.

4.1.3 Bed bug (*Cimex lectularius*)[129]

Reconstruction	**Forms**
1: *rústa	DES (I: 428): ***rústa*** (1: Luras, 2: Olbia, 3: Monti, 4: Berchidda, 7: Pattada, 8: Buddusò, 11: Nule, 12: Orune, 14: Nuoro), ***rúłła*** (6: Ozieri) 'bed bug (*Cimex lectularius* Linnaeus, 1758); noxious animal, fox'
2: *kurústa	DES (I: 428): ***kurústa*** (5: Posada, 9: Siniscola, 10: Bitti, 13: Orosei), [***sa***] ***ʔurústa*** (15: Oliena), ***koròsta*** (NLog.) 'bed bug (*Cimex lectularius*)'

This word generally refers to 'bed bug', but in some places to 'pest', and more specifically 'fox'. The forms can be reduced to two bases: **kurústa* and **rústa*, neither of which can be explained in terms of Romance material. The vocalism *-o-* instead of *-u-* in *koròsta* is difficult to explain however. Wagner notes that

[127] Cf. Wagner (1941a: 235) for the common occurrence of *-r-*metathesis in Sardinian.

[128] Cf. Wagner's (1941a: 122–121) transcription of this sound as *-ṛ-*, and his description of it as "a shortly rolled uvular trill (*einen kurzgerollten uvularen Zitterlaut*)".

[129] These data are discussed in a similar fashion in Swanenvleugel (2024: 239–241).

koròsta is attested by Casu in North Logudorese. But Casu (2011: s.v. *coròsta*) does not specify any information of this kind, so it is difficult to establish what the exact provenance within Logudorese of this word could be. Note that in North Logudorese we would also expect *-st-* to yield *-łt-*, which would most likely have been written as ***koròlta*. This form and its localization are thus somewhat uncertain.

Paulis (1990: 614) proposes to derive Srd. *kurústa* and *rústa* 'bed bug' from Lat. *crusta* 'crust', possibly influenced by *kúrma*, *kúruma* 'rue (*Ruta graveolens*)'. Besides the semantic issues discussed in DES (I: 438), there are some formal objections to deriving *kurústa*, *rústa* from Lat. *crusta*. The regular outcome of Lat. *crusta* would be ***krústa* (cf. Srd. *krás* < Lat. *cras* 'tomorrow'). Paulis argues that *rústa* could have behaved like Lat. *crux* 'cross', whose onset consonant was voiced early on to *gruke*, eventually yielding Nuor. *rúke*, Log. *rúɣe*, Camp. *grúži*. However, there is only a handful of words that, like *crux*, show initial voicing early enough to be continued as such in all Sardinian dialects (Wagner 1941a: 206). Moreover, neither *crusta* nor **grusta* can easily explain *kurústa*. Paulis' suggestion that this word would have undergone influence from *kúrma*, *kúruma* 'rue' (because this plant was presumably used to fight bed bugs), does not solve much, because we would rather expect initially stressed ***kúrusta* vel sim. Moreover, Wagner observes that Srd. *kurústa*, *rústa* does not refer to 'bed bugs' everywhere. This makes a derivation from Lat. *crusta* 'crust', as well as influence from *kúruma* 'rue (used to fight bed bugs)' semantically more difficult. As such, Lat. *crusta* can in the end probably not be the source of this word. I rather prefer to follow Wagner's suggestion that this may be a word of pre-Roman origin. Taken at face value, the coexistence of *kurústa* and *rústa* suggests the presence of an element **kV-*, which is also found in other words of suspected pre-Roman origin (§ 09.1.1).

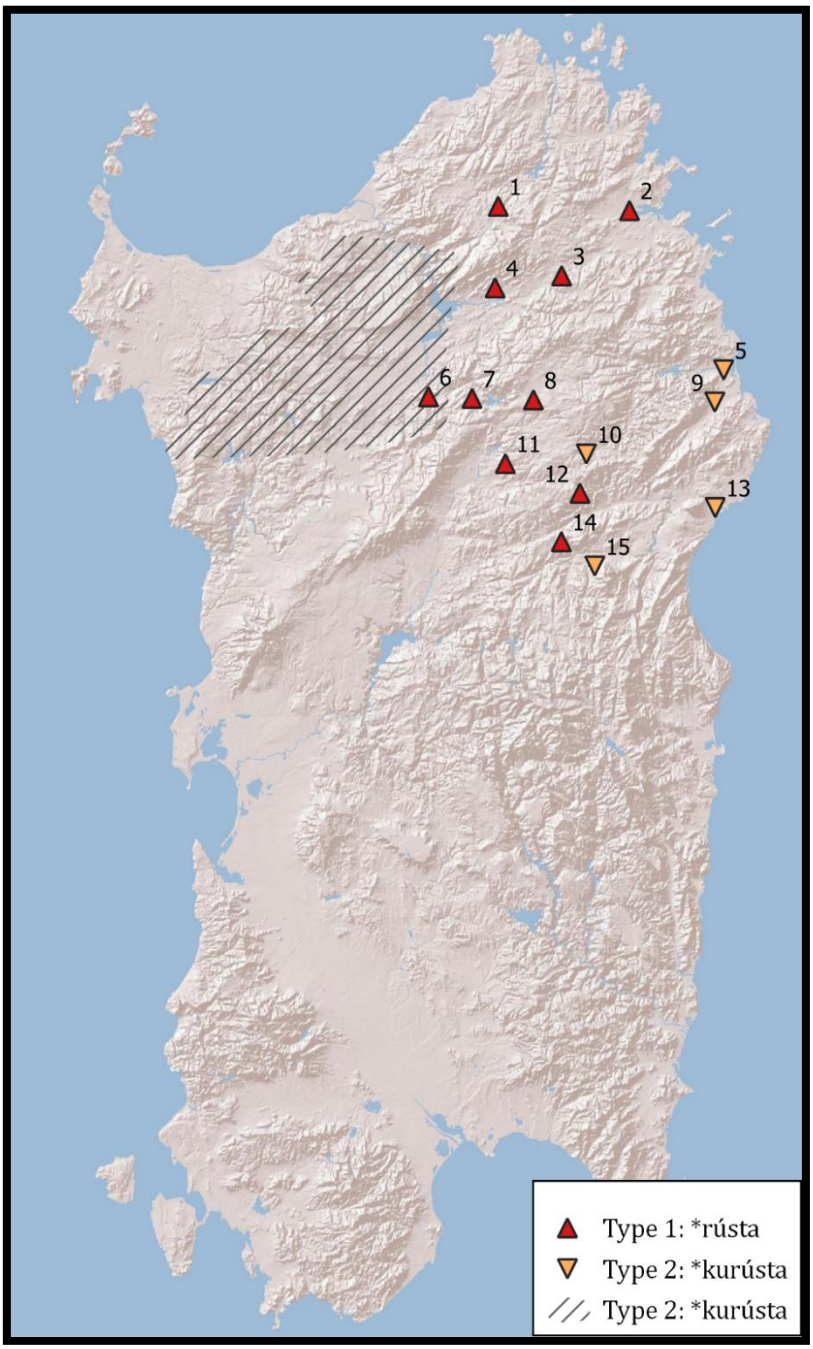

Figure 4.1.3: Bed bug (*Cimex lectularia*)

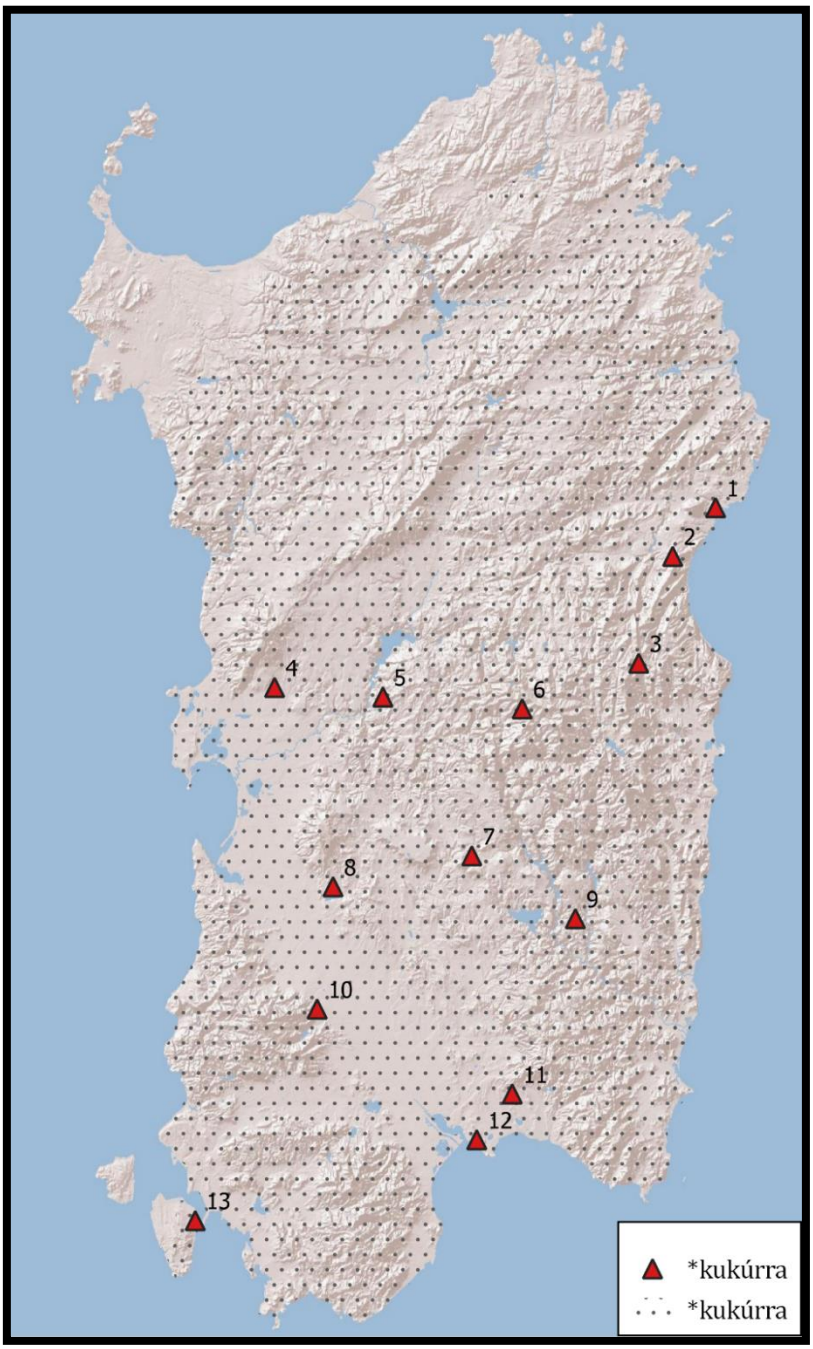

Figure 4.1.4: Caterpillar; earwig (*Forficula auricularia*)

4.1.4 Caterpillar; earwig (*Forficula auricularia*)

Reconstruction	Forms
***kukúrra**	DES (I: 419–420): ***kukúrra*** (Nuor.), ***kuɣúrra*** (Log., Camp.) 'caterpillar, earwig (*Forficula auricularia* Linnaeus, 1758)'; ***kukúrra***[***vóina***] (1: Orosei), ***kuhúrra***[***bóina***] (2: Dorgali), [***sa***] ***huhurra***[***óina***] (3: Urzulei) 'scorpion'; ***kuɣúrra*** (Planargia reg., Camp.) 'security knot on a noose that prevents it from strangling a draft animal; slipknot; tangle'
	AIS (468; 481): [***sa***] ***ɣuɣúrra*** (4: Milis, 5: Busachi, 6: Desulo, 8: Mogoro, 10: Villacidro, 12: Cagliari), [***sa***] ***ɣuɣúrra*** [***vručiḍḍáða***] (9: Escalaplano) 'earwig'; [***sa***] ***ɣuɣúrra*** (8: Mogoro, 10: Villacidro, 13: Sant'Antioco) 'caterpillar'
	Marcialis (2005: 34): ***cugúrra*** 'caterpillar, earwig', ***cugurra*** [***a coa frucidada***] (7: Isili) 'large earwig', ***cugurra*** [***de centu peis***] (11: Sinnai) 'millipede's
	Casu (2011: s.v. *cugùrra*): ***cugùrra*** 'noose, knot that prevents a noose from tightening; earwig'

The basic meaning of this word appears to have been 'caterpillar' (Schuchardt 1910; DES I: 419–420). It was extended to other oblong arthropods such as 'earwigs' (at times specified by *fručiḍḍáða* 'forked'), 'scorpions' (with addition of *bóina* 'bovine') and 'millipedes'. The meaning 'tangle; security knot on a noose' can probably be understood as a further semantic extension. Wagner (DES I: 419–420) rejects Schuchardt's (1910) proposal to derive all of these forms from Lat. *cōchlea* 'snail', since this word is already continued regularly in Sardinian as *kròka*, *kròɣa* 'snail'. Instead he proposes a pre-Roman origin, because of the presence of the supposed suffix *-úrra* and of the fact that *Cugurra* is attested in the Condaghe di Santa Maria di Bonarcado as a personal name (Virdis 2003: secs. 104, 115, 136). These are not the strongest arguments, but the peculiar structure of this word makes a pre-Roman origin plausible nonetheless. Wagner (DES I: 419–420) further compares Prov. *coucouro* 'cockchafer (*Melolontha melolontha*)' and Prov., Sp. *cuco* 'caterpillar'. However, Gallo-Romance forms similar to *coucouro* are derived by FEW (II: 1563) from Lat. *curculiō, gurguliō* 'corn-worm'. For the comparison to Sp. *cuco*, an analysis of Srd. *kukúrra* as **kuk-úrra* is essential, but the evidence for **-úrra* as a suffix is very slim (cf. § 9.2.7). As such, Srd. **kukúrra* remains isolated. Superficially similar in form is Lat. *curculiō*, *gurguliō* 'corn-worm; weevil', which is regularly inherited in Sardinian as Log. *gruɣudzòne*, Camp. *gruɣullòni* etc. 'weevil' (cf. DES I: 594), but this may be due to coincidence.

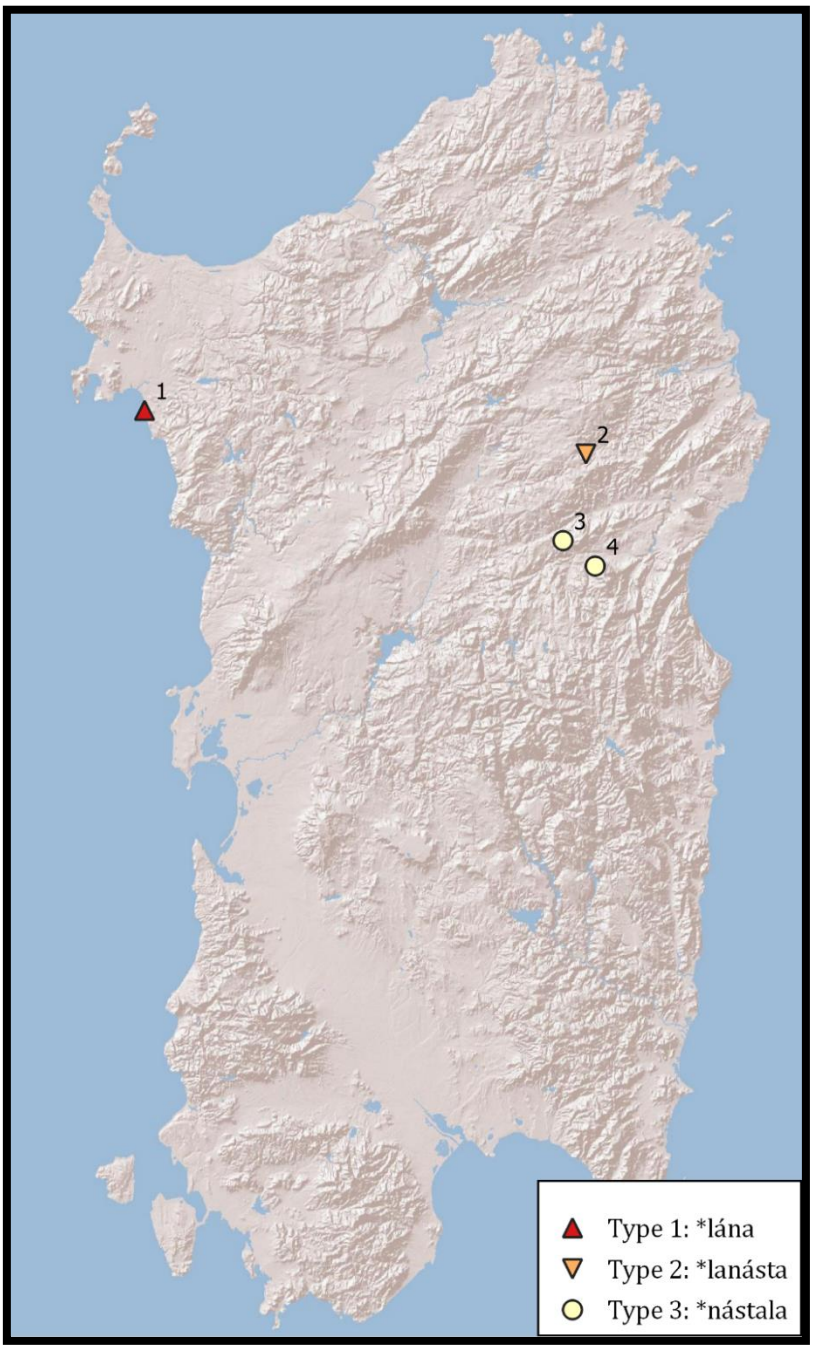

Figure 4.1.5: Clothes moth (*Tineidae sp.*)

4.1.5 Clothes moth (*Tineidae*)

Reconstruction	Forms
1: *lána	DES (II: 9–10): ***lána*** (1: Alghero) 'woodworm'
2: *lanásta	DES (II: 9–10): ***lanásta*** (2: Bitti) 'clothes moth (*Tineidae* Latreille, 1810)'
3: *nástala	DES (II: 9–10: ***nástala*** (3: Nuoro, 4: Oliena), ***nástula*** (3: Nuoro) 'clothes moth (*Tineidae*)'

Hubschmid (1949: 245) analyzes type 2 as a derivation from type 1, with addition of a pre-Roman suffix *-st-* (§ 9.2.9). *Nástala* 'clothes moth' would be the result of a metathesis from *lanásta*. One can but agree with Wagner (DES II: 9–10), however, that it is more attractive to assume that inherited *lána* 'wool' played some role in *lanásta* 'clothes moth', as these animals feed on wool. Since type 2 is somewhat more restricted in its distribution and since the proposed development *lanásta* >> *nástala* is more complex than only a metathesis, it is tempting to see *lanásta* as an innovation from original *nástala*, aided by the association of these animals to *lána* 'wool'. If this is correct, the connection to *lána* 'woodworm' must be abandoned, as well as the identification of a **-st-* suffix. It must be stressed however that the attestations for all three forms are rather restricted, which makes detailed analysis of the possible origins of these words difficult. It is in any case clear that neither *lanásta* nor *nástala* goes back to a Latin source. Nor can *lanásta* have been derived from *lána* wool, since *-st-* is not a productive suffix in Latin or Sardinian (and even the evidence for it being a pre-Roman suffix is slim; § 9.2.9). A pre-Roman origin for this word is possible, but there is little positive evidence for it. Similar in structure are the forms going back to **lakasta* and **ladaska* 'tick' found in Occitan dialects, whose origins are not fully clarified (FEW V: 127-128; Hubschmid 1949: 244–245; Bertoldi 1952). It is formally difficult, however, to explain them and the Sardinian forms from a common source.

4.1.6 Cockroach (*Blatta orientalis*)$_1$; cockchafer (*Melolontha melolontha*); darkling beetle (*Blaps sp.*)

Reconstruction	Forms
***meláke**	DES (II: 101): ***meláye*** (2: Oschiri) 'cockroach (*Blatta orientalis* Linnaeus, 1758)'
	Casu (2011: s.v. *melàghe*): ***melàghe*** 'cockchafer (*Melolontha melolontha* Linnaeus, 1758), beetle'
	Marcialis (2005: 43, 44): ***malangoni*** (1: Tempio Pausania), ***melogi*** (3: Ozieri) 'darkling beetle (*Blaps sp.* Fabricius 1775)'

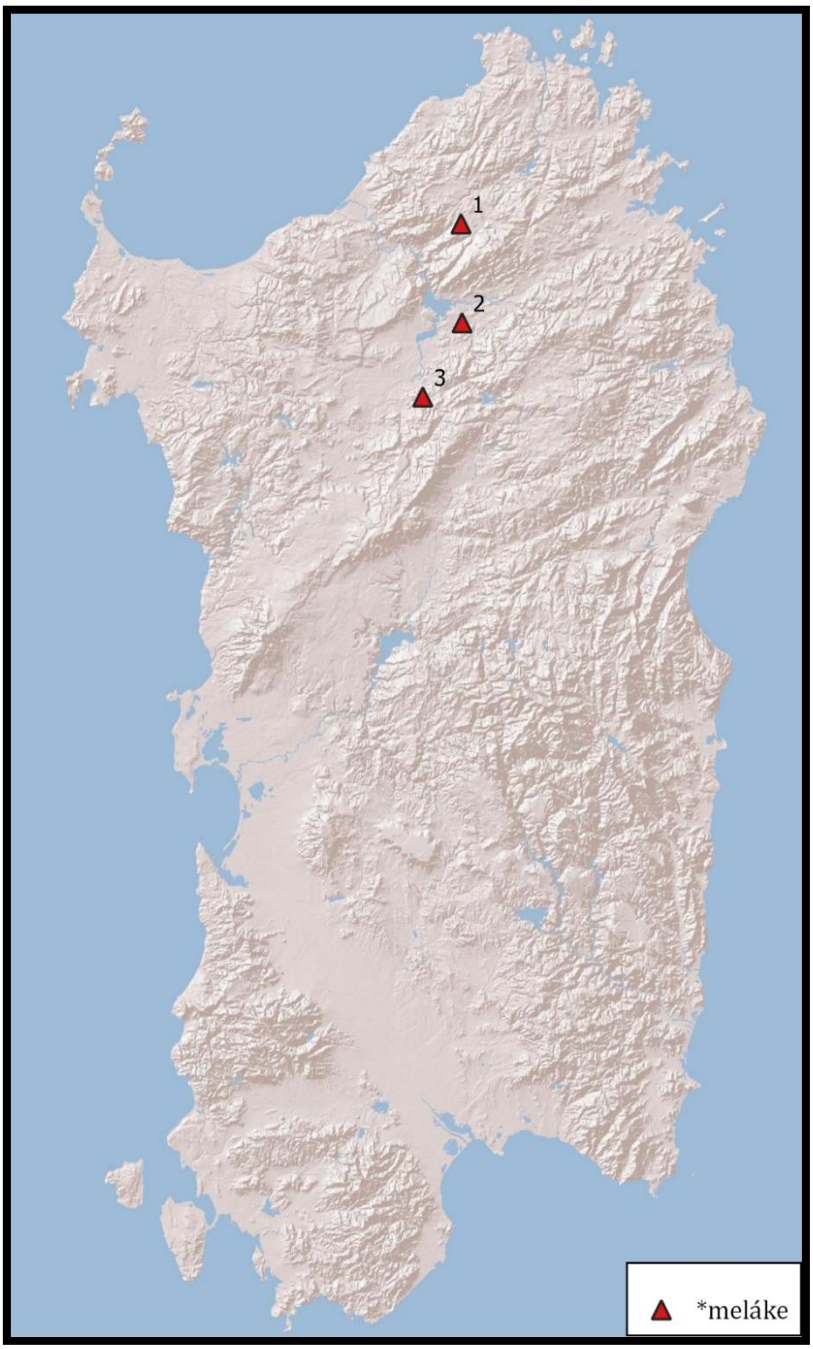

Figure 4.1.6: Cockroach (*Blatta orientalis*); cockchafer (*Melolontha melolontha*); darkling beetle (*Blaps sp.*)

Wagner (DES II: 101) proposes the words above to be derived from Gr. μέλας by means of a pre-Roman suffix *-ake*, supposedly from the ancient Greek that was spoken in the colony of Olbia, in the island's northeast corner. The element *-áke* appears in several other words of potential pre-Roman origin as well (§ 9.2.2.1).[130] However, there are several problems with this proposal. The stem of Gr. μέλας is not μέλα- or μέλ-, but μέλαν- (cf. gen.sg. μέλανος), so we would rather expect ***melanáke vel sim.*, if the derivation was made within Greek. Moreover, Greek μέλλαξ, μέλακες is in fact attested, but it means 'young man' (Chantraine 1933: 379–380), which is difficult to connect with 'beetle/cockroach'.[131] Paulis (1983: 52), on the other hand, suggests a Byzantine Greek origin, noting that similar forms refer to the same animal species in southern Italy: Salento *melòta*, *malòke* etc. 'cockroach' (cf. also AIS 472). This scenario is probably more attractive than a loan from ancient Greek, as Wagner (DES II: 101) proposes.

4.1.7 Cockroach (*Blatta orientalis*)$_2$; giant darkling beetle (*Blaps gigas*)

Reconstruction	**Forms**
1: *sisái̯a	DES (II: 421): ***siṡáya*** (1: Olbia, 3: Monti, 5: Siniscola, 8: Bitti, 14: Nuoro, 16: Dorgali, 17: Oliena), ***seṡáya*** (4: Ploaghe, 10: Bonorva), ***saṡáya*** (6: Mores, 12: Orosei), ***siṡáḍḍza*** (9: Nule, 13: Bolotana), ***siṡáğa*** (11: Bono), ***saṡáğa*** (15: Bosa, 18: Macomer) 'cockroach (*Blatta orientalis*); giant darkling beetle (*Blaps gigas* Linnaeus, 1767)'
	Marcialis (2005: 59, 60): ***saṡáya*** ("northern"), ***sasağa*** (15: Bosa), ***sisaia*** (17: Oliena) 'darkling beetle (*Blaps sp.*)'
2: *sei̯sèi̯	DES (II: 402, 412): ***sèṡè*** (18: Macomer) 'butterfly', ***seiṡèi*** (19: Seui) 'longhorn beetle (*Cerambycidae sp.* Latreille, 1802)'
	Marcialis (2005: 60): ***sei sèi*** (Ogl.) 'longhorn beetle (*Cerambycidae sp.*)'
3: *babbasái̯u	DES (II: 421): ***babbaṡáyu*** (2: Tempio Pausania), ***babbaṡáu*** (7: Thiesi) 'cockroach (*Blatta orientalis*); giant darkling beetle (*Blaps gigas*)'

The base form of these words meaning 'cockroach, darkling beetle' is likely continued by type 1, which is most widely attested.

[130] Greek -αϰ-, -ᾱϰ- too is considered to be pre-Greek by Chantraine (1933: 376) and Beekes (2014: 31).

[131] For the homonymy between cockroaches and beetles, going back to classical times, see Beavis (1988: 80–81).

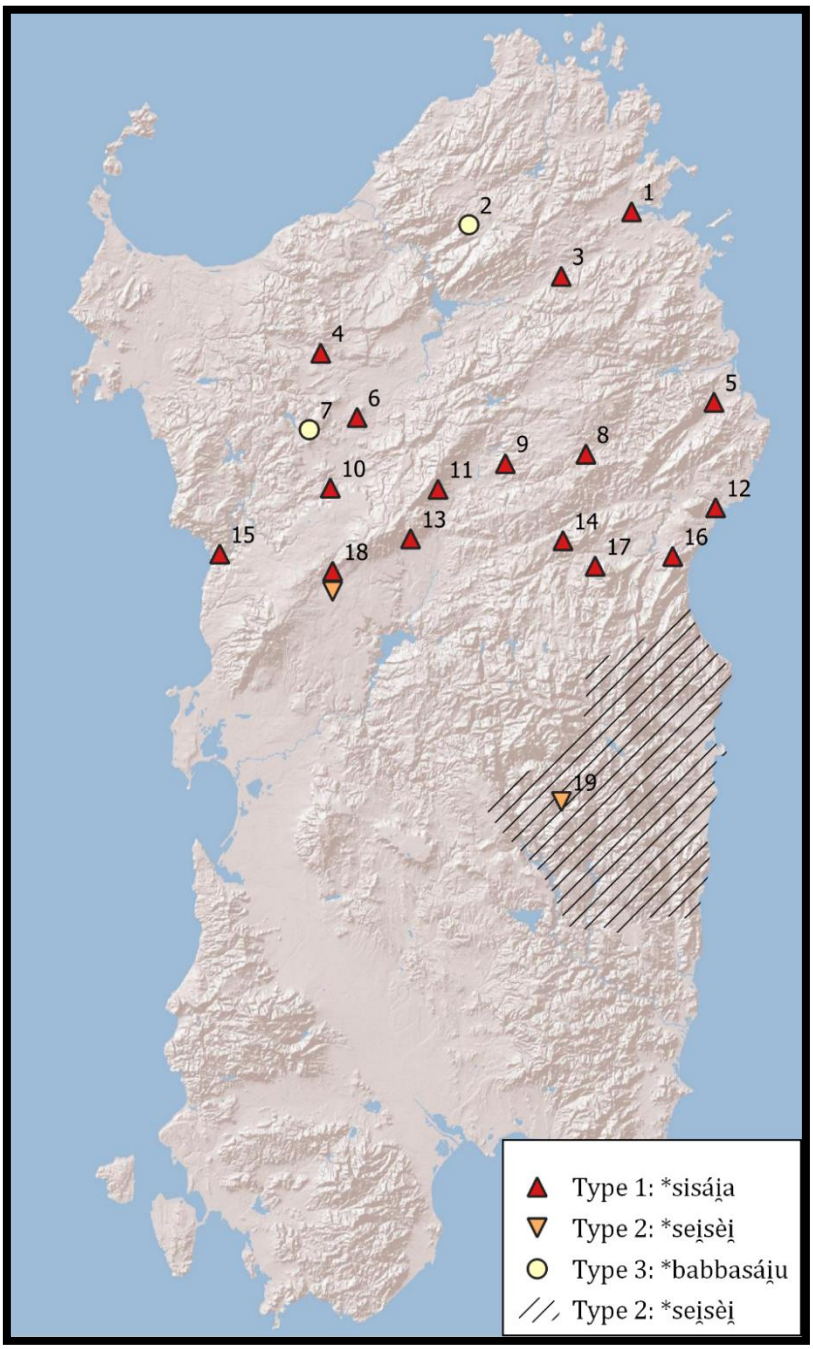

Figure 4.1.7: Cockroach (*Blatta orientalis*); giant darkling beetle (*Blaps gigas*)

The pretonic vowel alternations found in *seśáya*, *saśáya* etc. can be explained as the result of assimilation of the *-i-* to the accented syllable. *Babbaśáyu* and *babbaśáu* contain the element *babba-* that is often found in words for insects (cf. DES I: 161–62). Wagner (DES II: 402) considers this word pre-Roman, with which I tend to agree, on the basis of its peculiar phonological shape. Rather similar in shape is type 2. Wagner (DES II: 402) explains it as an "infantile formation". However, in light of the similarities of this word to *siśáya* etc. and their close semantic proximity, it is unattractive to separate these two words. The variant *seiśèi* may indicate that this word originally ended in a diphthong *-éi*, which is also found in numerous pre-Roman Sardinian toponyms (§ 9.2.1) (cf. Wolf 1998a: 53–55). Finally, there is also *sèśè* 'butterfly' in Macomer (18), which is similar to Ogl. *seiśèi* 'longhorn beetle'. However, the meaning 'butterfly' is further removed, and in Macomer we also find *saśáǧa* 'cockroach'. We must probably depart from an original **sisèi̯*/**sisái̯* 'cockroach, beetle' to explain the forms discussed above. Secondary remodelings of "infantile" nature (cf. DES I: 402) cannot be excluded. Wagner (1997: 256–257) compares Bq. *sits, sats* '(clothes) moth' and *sats* 'trash'.[132] Mayer Modena (1989: 11–12) further compares Akk. *sāsu*, Hebr. *sās* 'moth' < PSem. **sā/ūs-* (cf. Kogan 2012a: 213).[133] Semantically, the Sardinian forms are farther removed, referring to cockroaches and beetles rather than to moths, although associations between cockroaches and moths are historically known (Beavis 1988: 80–84). Formally, the corresponding sequence is too short (viz. **sVs-*) to confidently rule out chance correspondence between the forms involved.

[132] For the meaning of these words, cf. EH (566b, 574b).

[133] Gr. σής 'moth' may be a Semitic loan (Lewy 1895: 16–17).

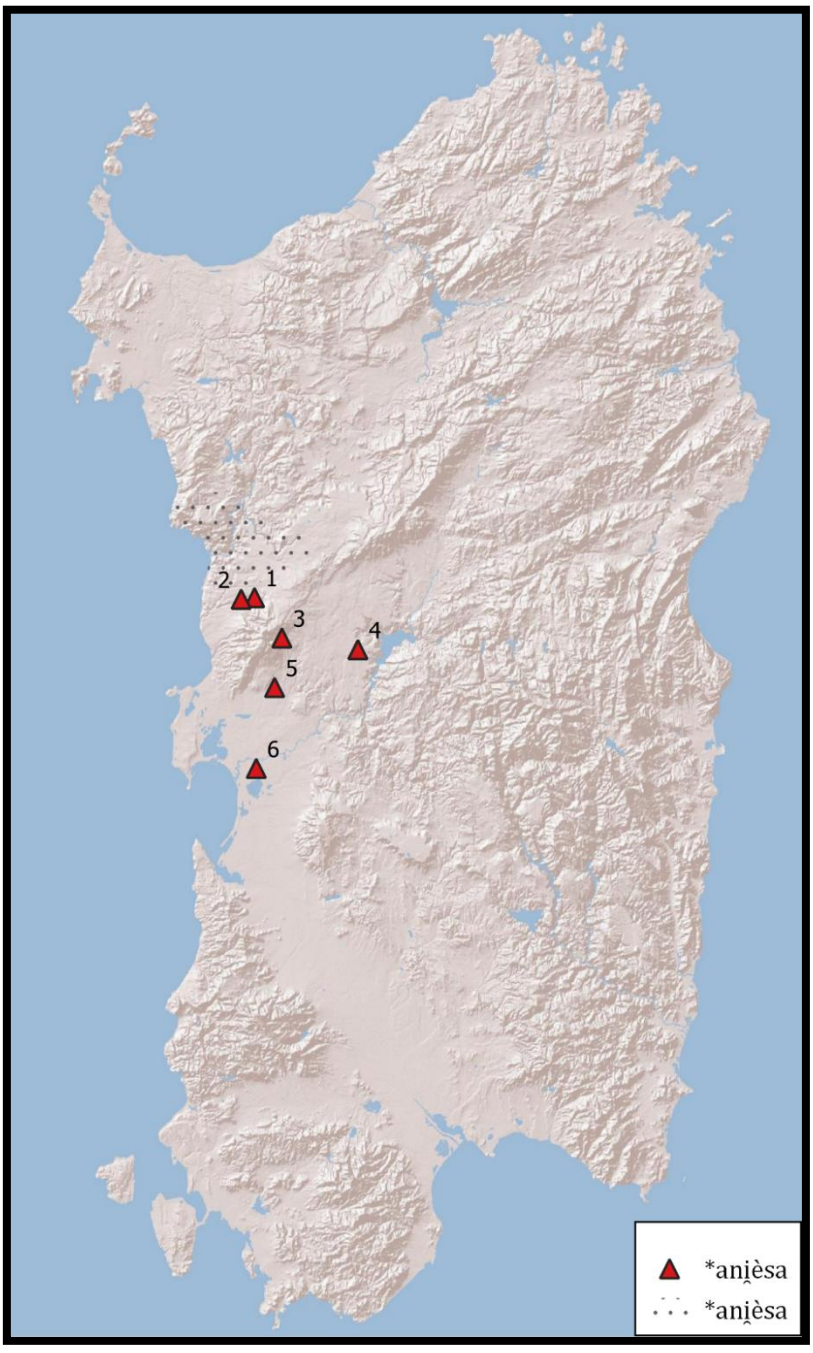

Figure 4.1.8: Cricket

4.1.8 Cricket

Reconstruction	**Forms**
***anḭ̀esa**	DES (I: 97): ***andzèṡa*** (Planargia reg., 1: Scano di Montiferro, 2: Sennariolo, 6: Ghilarza, 7: Milis) 'great green bushcricket (*Tettigonia viridissima* Linnaeus, 1758)', ***andzèṡa*** (3: Santu Lussurgiu) 'type of cricket (*Gryllus sp.* Linnaeus, 1758)', ***anǧèṡa*** (6: Oristano) 'cicada (*Cicada sp.* Linnaeus, 1758)'
	Marcialis (2005: 14): ***angesa*** (6: Oristano) 'oil beetle (*Meloe proscarabaeus* Linnaeus, 1758)'

All forms listed above are from the central portion of the Sardinian west coast, and they all go back to **anḭ̀esa*, without any phonological variation. Wagner (DES I: 97) considers this word "probably pre-Roman", probably because of the lack of an etymology. The homogeneity of the Sardinian forms makes any further analysis impossible.

4.1.9 Dragonfly; grasshopper

Reconstruction	**Forms**
1: *pilíke	DES (II: 265): [***káḍḍ' e***] ***pilíkke*** (1: Siniscola) 'dragonfly'
2: *ϑimpilíke	DES (II: 547): ***tsimpilíɣe*** (2: Ovodda) 'grasshopper'

The name for the dragonfly in Siniscola (2) literally means "horse of *pilíkke*". This formation corresponds to other insect names such as *káḍḍu 'e déus* (Désulo) 'praying mantis' (litt. "horse of God"), *káḍḍu 'e déu* (Luras), *káḍḍu 'e demóniu* (Posada) 'dragonfly' (respectively "horse of god" and "demon horse"), and many others (DES I: 323–324). Outside Sardinia, Wagner (DES I: 323–324) cites Ar. فرس الشيطان (*faras aš-šayṭān*) 'praying mantis' (litt. "devil's horse"), and various similar Romance formations can also be found (FEW II: 12); cf. further also It. *cavalletta* 'grasshopper' (litt. "little horse").[134] Whatever the origin of this *Benennungsmotiv*, it raises the suspicion that the element *pilíkke* (< **pilíke*) in *káḍḍ' e pilíkke* did originally not necessarily mean 'dragonfly', but that it could refer to some spiritual entity, be it a god or a demon. On the other hand, we may compare *tsimpilíɣe* (2: Ovodda), which means 'grasshopper' and which looks like **pilíke* preceded by a **ϑi*-prefix (§ 9.1.2).[135]

[134] In light of these formations, an interpretation of *čerβúsku* (Desulo) 'grasshopper' as a derivation from *čérβu* 'deer' is not at all "bold", as Wagner (DES I: 446) calls his own proposal. A pre-Roman origin fort his word is not necessary.

[135] For nasal insertions, cf. Wagner (1941a: 223).

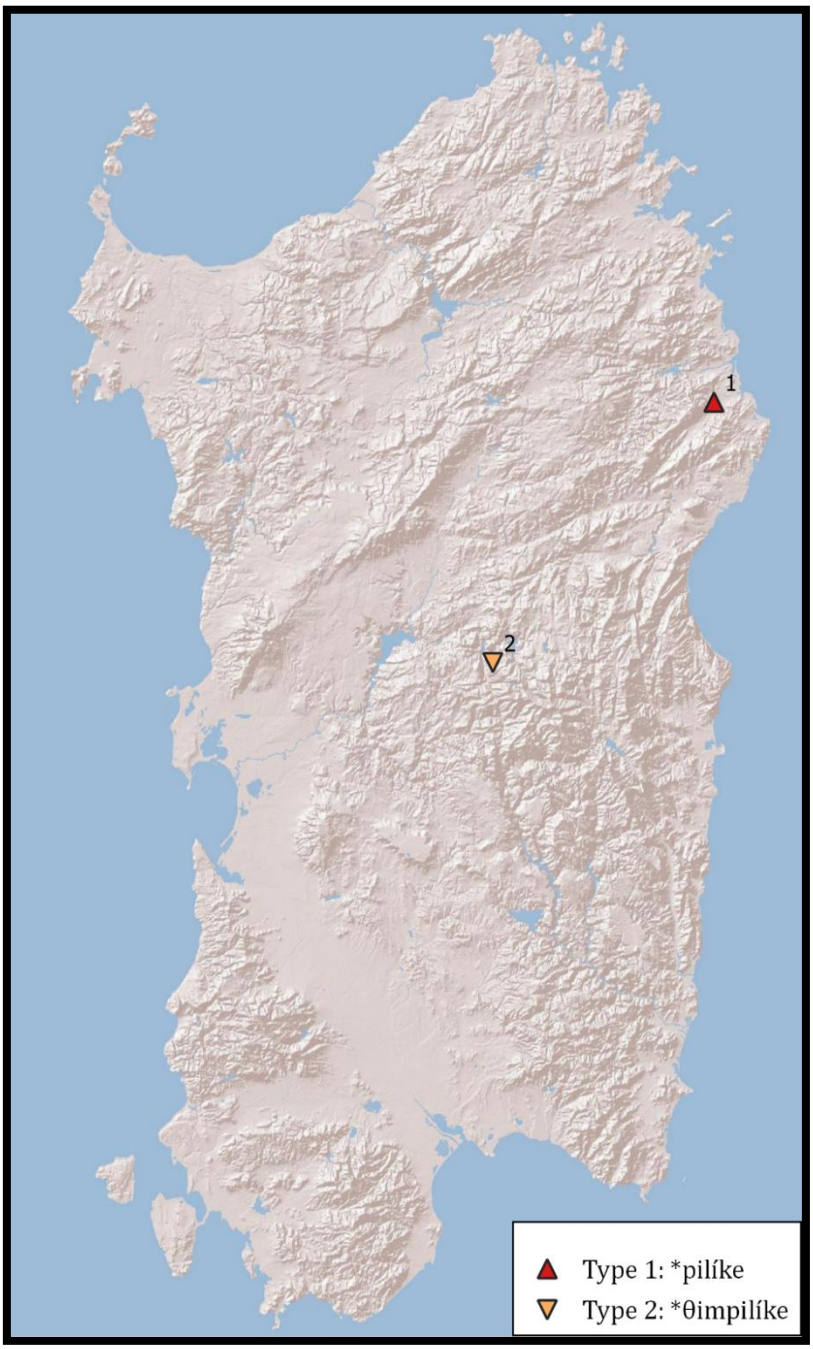

Figure 4.1.9: Dragonfly; grasshopper

If this form is related, **pilíke* could be referring to an insect rather than to a spiritual entity after all. Besides *tsimpilíɣe*, there are many forms of the similar type **ϑilipírke* referring 'grasshopper' in the Sardinian dialects (§ 4.1.12). The fact that *káḍḍ' e déu* etc. refers both to dragonflies and praying mantises, makes an etymological connection between [*káḍḍ' e*] *pilíkke* 'dragonfly' and *tsimpilíɣe* 'grasshopper' plausible as well. If they are indeed related, the presence of **ϑi-* in the second form is strong evidence that these words entered Sardinian from a pre-Roman language spoken on the island.

4.1.10 Earthworm (*Lumbricus terrestris*)[136]

Reconstruction	**Forms**
1: *ϑiling(ul)òne	DES (II: 555–556): ***tsirinòni*** (1: Tempio Pausania), ***attaltingòne***[137] (2: Sennori), ***tiringòni*** (3: Sassari), ***tilinǧòne*** (5: Ploaghe, 11: Bonorva), ***attilinǧòne*** (8: Thiesi), ***attolinǧòne*** (9: Villanove Monteleone), ***atterringòne*** (13: Padria), ***ϑilingròne*** (14: Lollove, 15: Nuoro), ***tilingròne*** (18: Macomer), ***tsilingòne*** (19: Santu Lussurgiu, 23: Busachi), ***tsiaʁingồi*** (22: Milis), ***tsiringòne*** (24: Laconi), ***tsirringồi*** (25: Mogoro), ***tsiringồi*** (26: San Gavino Monreale), ***sittsiringòni*** (27: Guspini), ***tsiringồi*** (28: Villacidro), ***tsirringòni*** (30: Sarroch), ***tserringòni*** (31: Domus de Maria) 'earthworm (*Lumbricus terrestris* Linnaeus, 1758)'
	AIS (457): [***lu***] ***tsiriñòni*** (1: Tempio Pausania) 'earthworm (*Lumbricus terrestris*)'
	Marcialis (2005: 65, 114): ***tilingòne*** (16: Bosa) 'earthworm (*Lumbricus terrestris*)', ***zirringòni*** (29: Cagliari) 'various kinds of marine worms; fish bait'
2: *ϑilíng(ul)a	DES (II: 555–556): ***tilínǧa*** (4: Oschiri, 7: Mores), ***attulínǧa*** (6: Ozieri), ***tsilíngu*** (20: Norbello) 'earthworm (*Lumbricus terrestris*)'
3: *ϑulung(ul)òne	DES (II: 555–556): ***ϑurunkròne*** (10: Bitti), ***ϑulunkòne*** (12: Orune), ***ϑulungròne*** (17: Dorgali, 21: Fonni) 'earthworm (*Lumbricus terrestris*)'
	Marcialis (2005: 64): ***sullungòne*** 'earthworm (*Lumbricus terrestris*)'

[136] These data are discussed in a similar fashion in Swanenvleugel (2024: 251–252).

[137] Like this in DES. I suspect that the *-lt-* in *attaltingòne* is a typographical error for *-l-*.

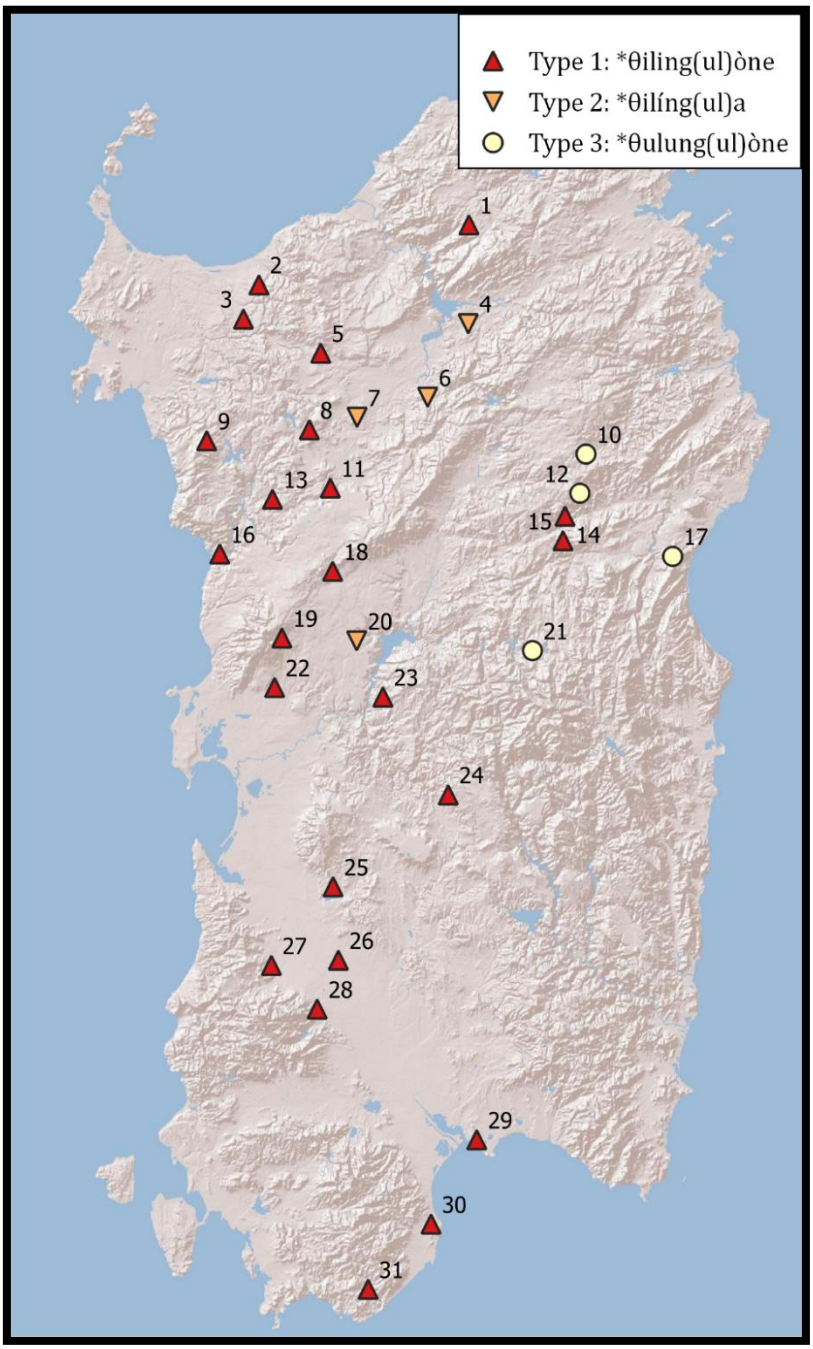

Figure 4.1.10: Earthworm (*Lumbricus terrestris*)

This word for 'earthworm' is widely distributed on Sardinia. All forms contain **ϑiling-* (type 1, 2) or **ϑulung-* (type 3). In all types, most forms have been suffixed with **-(u)l-*; e.g. *tilinǧòne* < **ϑiling(u)lòne* and *ϑulungròne* < **ϑulung(u)lòne*. It is possible that the presence of **-(u)l-* is original, and that it has been lost in the few forms that lack it. In the Campidanese dialects, loss of *-r-* is common (Wagner 1941a: 161). Most forms contain the suffix *-òne*, with the exception of the forms in type 2. As for the vocalism, it is attractive to explain the difference between *-i-i-* (type 1) and *-u-u-* (type 3) as the result of some assimilation or dissimilation of pretonal vowels. The original vocalism is likely that of *tilínǧa* etc. (type 2), where *-í-* is stressed. The *-u-*vocalism of type 3 may be due to assimilation to the rounded vowel of stressed *-òne*. We may thus reconstruct this word as **ϑiling(u)l-(òne)*. There is little doubt that this word is of pre-Roman origin. Although there is no direct evidence for the identification of **ϑi-* as a genuine prefix in this word (§ 9.1.2), it patterns well with the other discussed instances of **ϑi-* both morphologically and semantically. Guarnerio (1904a: 61) compares Lomb. *lingöri* (Verbania), Lad. *lingiola* (Val Gardena) 'lizard'. Although these words are formally similar, their diverging semantics cast doubt on this comparison.

4.1.11 Grasshopper$_1$

Reconstruction	**Forms**
1: *pipiϑíri	DES (II: 259): ***piβintsíri*** (1: Ghilarza, 2: Bonarcado), ***piβirittsíri*** (3: Seneghe), ***piβittséri*** (4: Milis), ***piβittsíri*** (Camp., 5: Busachi, 8: Samugheo, 10: Gadoni) 'grasshopper'
	Marcialis (2005: 51): ***pibizziri*** (20: Cagliari) 'id.'
	AIS (466): [***su***] ***βiβitssíri*** (14: Mogoro), [***su***] ***βippitsíri*** (18: Villacidro), [***su***] ***βiβitsíri*** (20: Cagliari) 'id.'
2: *pipiϑía/u	DES (II: 259): ***piβiϑϑía*** (6: Baunei), ***piβiϑϑíu*** (7: Villagrande Strisaili, 9: Arzana, 19: Villaputzu), ***piβittsíu*** (11: Laconi, 16: Escalaplano, 17: San Nicolò Gerrei), ***piβissíu*** (12: Gairo, 13: Bari Sardo, 15: Perdasdefogu), ***piβisíu*** (21: Sant'Antioco) 'id.'

Both types of this word contain an element **pipiϑϑ(í)-*. Those of type 1 end in *-íri*, while those in type 2 end in *-íu* or *-ía*. The sequence **-ϑi-* that is present in both types, cannot go back to any sequence inherited from Latin. Wagner (DES II: 259) proposes a sound-symbolic origin, after the sound these insects make. Although this cannot fully be excluded, this word exhibits suspiciously little formal variation for a sound-symbolic formation. Moreover, the presence of **-ϑ-* (§ 8.1.1), the ending *-iri* (§ 9.2.6.2), and the peculiar overall structure of this word are reminiscent of other forms for which a pre-Roman origin has been proposed.

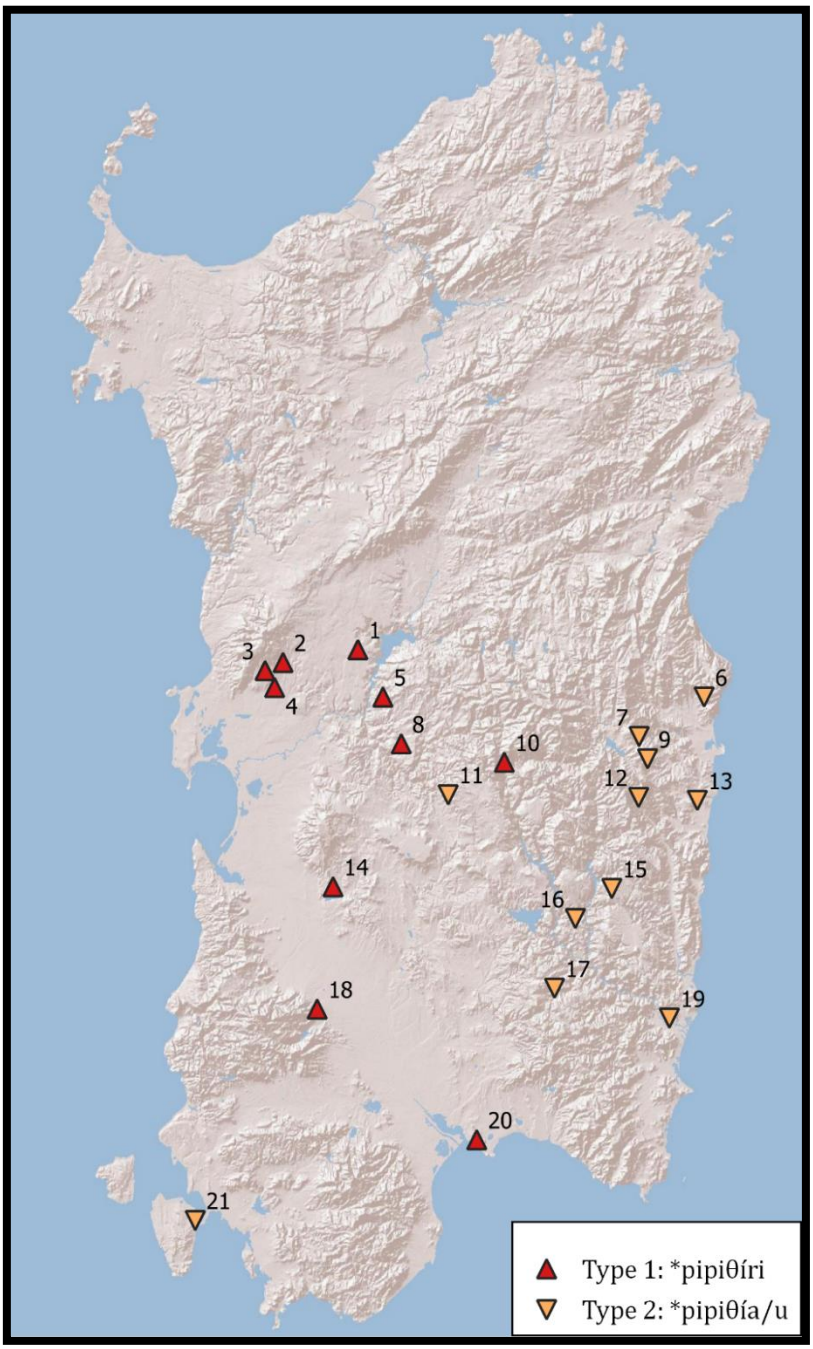

Figure 4.1.11: Grasshopper

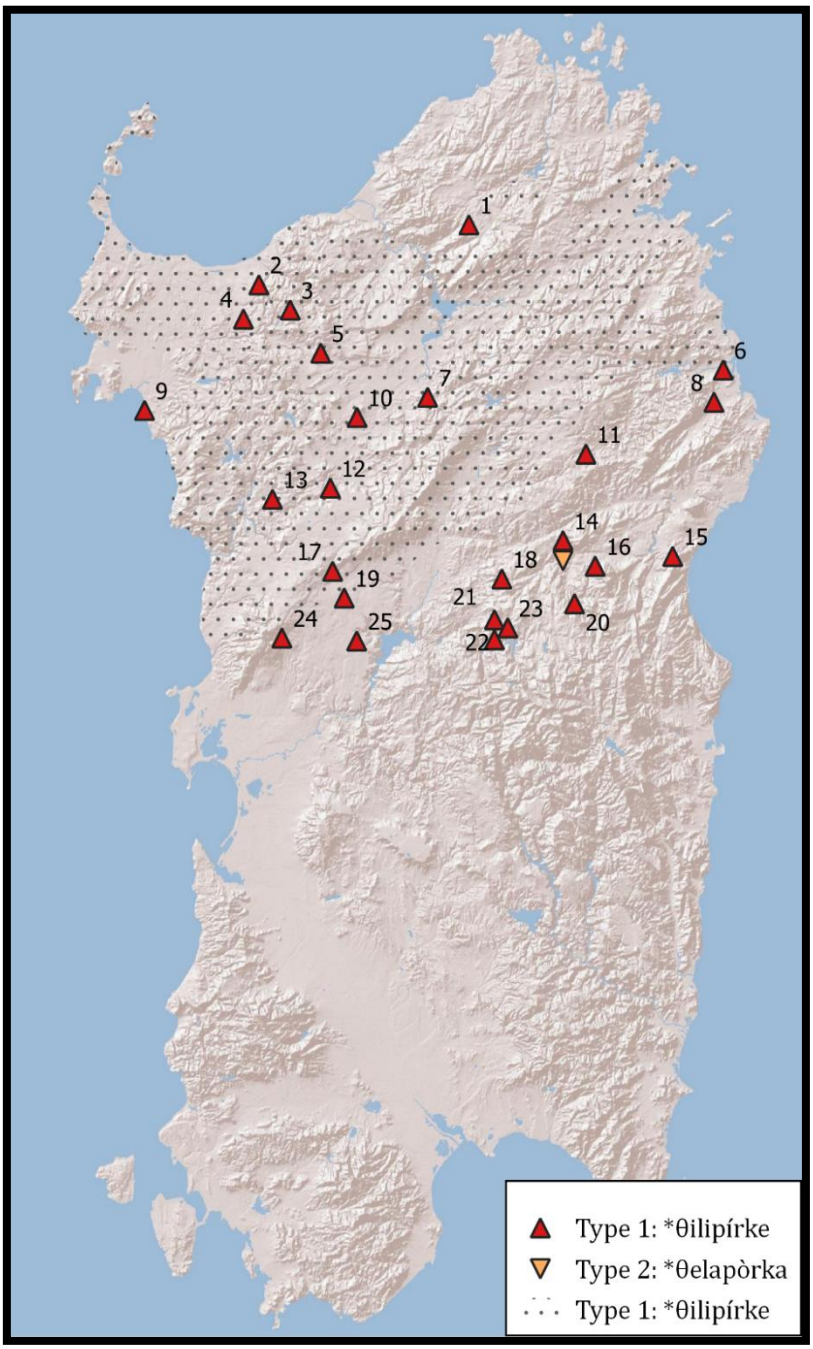

Figure 4.1.12: Grasshopper

4.1.12 Grasshopper$_2$[138]

Reconstruction	Forms
1: **ϑilipírke*	DES (II: 546–547): [*s'*] ***attaliβíχχe*** (2: Sennori), ***attilibíske*** (3: Osilo), ***tiliβíχχe*** (5: Ploaghe, 10: Mores), ***ϑilipírke*** (8: Siniscola, 11: Bitti, 14: Nuoro, 18: Orani), ***tilipírke*** (7: Posada, 15: Dorgali), ***ϑilibríʔe*** (16: Oliena), ***tsilibrítke*** (19: Borore), ***ϑilipríʔe*** (20: Orgosolo, 22: Ollolai, 23: Gavoi), ***tsilipríʔe*** (21: Olzai), ***tsilibrínkidi***, ***zilibrínkiḍḍi*** (25: Norbello), ***tilibítke*** (12: Bonorva), ***attiliβílke*** (17: Macomer), ***atteriβílke*** (13: Padria) 'grasshopper'
	Marcialis (2005: 65, 69): ***zilibricu*** (1: Tempio Pausania), ***tilipische*** (7: Ozieri), ***ziliblich*** (9: Alghero) 'id.'
	Guarnerio (1904a: 61): ***tilibrikku*** (Sass., Log.), ***tilibriku, tilipriku, tilibriki, tilibrirke*** (Log.) 'id.'
	AIS (466): [***lu***] ***tiribíkku*** (4: Sassari), [***su***] ***ðiliβírke*** (17: Macomer), [***ṣu***] ***tsiliβríkke*** (24: Santu Lussurgiu) 'id.'
2: **ϑelapòrka*	DES (II: 546–547): ***ϑelapòrka*** (14: Nuoro) 'large grasshopper'

All forms of type 1 can be reconstructed to **ϑilipírke* or **ϑilipílke*. This word is not attested for the dialects of Baunei and Fonni, which preserve the distinction between **-rk-* and **-lk-* (cf. Wagner 1941a: 176–177). *Ziliblich* (Alghero), already recorded by Schuchardt (1907: 17) is difficult to interpret. According to Blasco Ferrer's historical phonology (1984b: 67–68) **-bl-* should have yielded ***-br-* (e.g. Alghero *dóbra* vs. Cat. *doble*). Nor do syllables of the type **-CVlC-* metathesize, as opposed to syllables of the type **-CVrC-* (Blasco Ferrer 1984b: 93–94). In short, there is no clear regular source for the variant *ziliblich*. A few lines after mentioning *ziliblich*, Schuchardt (1907: 17) mentions *silibrich volador* 'dragonfly', also from Alghero. This form is probably identical to *ziliblich*, and is the expected outcome of **ϑilipírk-*. This is potential evidence that **ϑilipírk-* (rather than **ϑilipílk-*) lies at the base of the Sardinian forms as well. The variant *tsilibrínkidi* has been influenced by the verb *brinkare* 'to hop' (DES II: 547), and *tsimpilíɣe* must be due to metathesis. In Nuoro we find *ϑilipírke* 'grasshopper' next to *ϑelapòrka* 'large grasshopper'.[139] The vocalism in *ϑelapòrka* is similar to that of e.g. *ϑelakúkku* 'skink' (§ 4.3.4); influence by *pòrka* 'sow, female pig' (referring to this grasshopper's larger size) is not excluded.

[138] These data are discussed in a similar fashion in Swanenvleugel (2024: 248–249).

[139] The notion 'large grasshopper' perhaps refers to the great green bush-cricket (*Tettigonia viridissima*) as opposed to smaller kinds of grasshoppers like locusts (*Acrididae* MacLeay, 1819) or pygmy grasshoppers (*Tetrigidae* Rambur, 1838).

Guarnerio (1904a: 61) proposes a derivation from Lat. *lumbricus* 'earthworm' by means of a prefix **ϑi-* (§ 9.1.2), which is unlikely both semantically and formally. Alessio's (1939: 14–15) connection to Gr. ἀττέλαβος 'edible locust', which has been suggested to be a Semitic or Egyptian loan in Greek (cf. Beekes 2010: 166), is rightly rejected in DES (546–547). While there may be a superficial phonetic similarity between the Sardinian and the Greek word, there is no explanation for the Sardinian ending **-írke/-ílke*. Moreover, if the **ϑi-* in Sardinian is a prefix like in *ϑilikèrta* etc., the comparison is reduced to Srd. **-lipírk-* vs. Gr. **attelab-* in which any similarity is hard to find (cf. Paulis 1984a: cvii). Wagner (DES II: 546–547) prefers to follow Schuchardt (1907: 17) instead, who sees Srd. *ϑilipírke* as an internally Sardinian onomatopoeic formation. This is not a satisfying solution either, as the sequence **ϑi*(*li-*) also recurs in words for animals that do not sound like grasshoppers at all (e.g. *ϑilikèrta* 'lizard', *ϑilikúku* 'skink, slug'). In my opinion, it is more reasonable to assume that this word was borrowed from pre-Roman Sardinian. The identification of the initial sequence **ϑi-* as some sort of prefix was already done by Guarnerio (1904a: 61). No alternation of **ϑi-* is attested for this word, however (cf. § 9.1.2). Similar in form and meaning are *tsimpilíɣe* (Ovodda) 'grasshopper', and its un-prefixed counterpart [*káḍḍ' e*] *pilíkke* (Siniscola)(§ 4.1.9). Requiring a reconstruction **pilíke*, however, they must be kept separate from **ϑilipírke* 'grasshopper'.

4.1.13 Rose chafer (*Cetonia aurata*)

Reconstruction	**Forms**
****senèkra***	DES (II: 405): ***senèɣra*** (1: Monti) 'rose chafer (*Cetonia aurata* Linnaeus, 1758)'

This word is restricted to the Logudorese dialect of Monti. Although it would be attractive from a morphological point of view to posit an original **senèkula* (as a derivation from Lat. *senex*), with the common Sardinian suffix *-k*(*u*)*lu/a*, this would be expected to yield ***senèğa* in this variety of Logudorese (Virdis 1988: 907). We thus need to reconstruct **senèkra*, which is morphologically completely opaque. The isolated attestation of this word makes it impossible to draw conclusions about its origin.

Figure 4.1.13: Rose chafer (*Cetonia aurata*)

4.1.14 Spider

Reconstruction	**Forms**
1: *solo(C)ika	DES (II: 424–425): ***suíɣa*** (Camp.), ***solóiɣa*** (1: Bonorva, 2: Macomer), ***soloíɣa*** (4: Norbello, 6: Desulo) 'black widow (*Latrodectus tredecimguttatus* Rossi, 1790), tarantula wolf spider (*Lycosa tarantula* Linnaeus, 1758)'; ***solóiɣa*** (3: Borore), ***suíɣa*** (7: Sardara) 'spider'; ***suíɣa*** (9: Iglesias) 'big ant'
	AIS (485): [***sa***] ***suíɣa*** (8: Villacidro) 'spider'
2: *solórika	DES (II: 424–425): ***solóriɣa*** (5: Ghilarza and surroundings) 'black widow, tarantula wolf spider'
3: *aṇialór(ik)a	DES (II: 424–425): ***andzalóriɣa*** (5: Ghilarza), ***andzalòra*** (5: Ghilarza and surroundings) 'black widow, tarantula wolf spider'

Wagner (DES II: 424–425) argues that these words for a type of venomous spider must be related somehow to Lat. *solifūga*, *solipūga, sālpūga.* These words are attested in Latin texts and refer to an undefined venomous animal apparently uncommon in Italy (cf. Pliny's *Naturalis Historia* 29: 92). Pliny mentions that Cicero calls this animal *solipūga*, but that *sālpūga* is used in Hispania Baetica, while it is described as a Sardinian animal by Isidore and Solinus (cf. Walde & Hofmann 1954: 470; DES II: 424-425). While a connection between the Sardinian and Latin forms is very likely, the forms attested in Latin cannot be the regular source of the modern Sardinian ones. In Sardinian, the most archaic forms appear to be those of type 1, going back to **solo(C)ika*, in which **-C-* can be any voiced stop (or *-v-*). Serra (1954b: 219) explains *andzalóriɣa* (5: Ghilarza) as a compound of *andzáre* 'to give birth (of animals)' and *lóriɣa* 'bell, (ear)ring', because this kind of spider supposedly leaves behind a little drop of water. In light of the existence of the forms of type 1, as well as Lat. *solifūga, solipūga, sālpūga*, it is unlikely that the formation of type 3 is wholly independent. Nevertheless, the semantic associations described by Serra (1954b: 219) may have caused folk-etymological deformations leading to the forms of both type 2 and type 3. The irregular variation between Lat. *solifūga*, *solipūga* and *sālpūga*, referring to a species that was not found in Italy, strongly suggests that these Latin forms were representations of some word from a foreign language. The existence of modern Srd. *solóiɣa* etc. suggests that a pre-Roman Sardinian language was the source, from which Lat. *solipūga* etc. and Srd. *solóiɣa* were borrowed on separate occasions. If Pliny's attribution of *sālpūga* to Baetica is correct, this is possible evidence for pre-Roman linguistic ties between Sardinia and the Iberian Peninsula (§ 11.3.1).

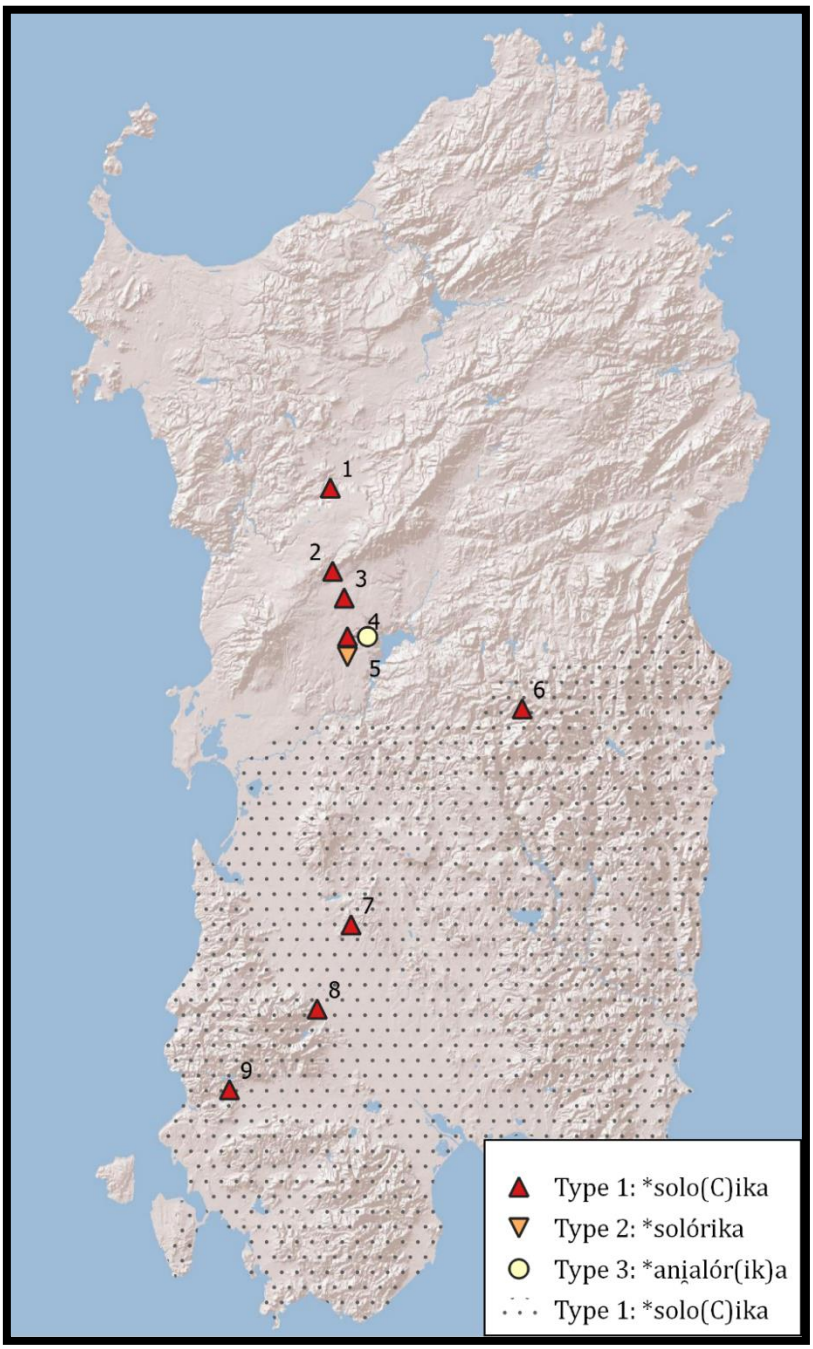

Figure 4.1.14: Spider

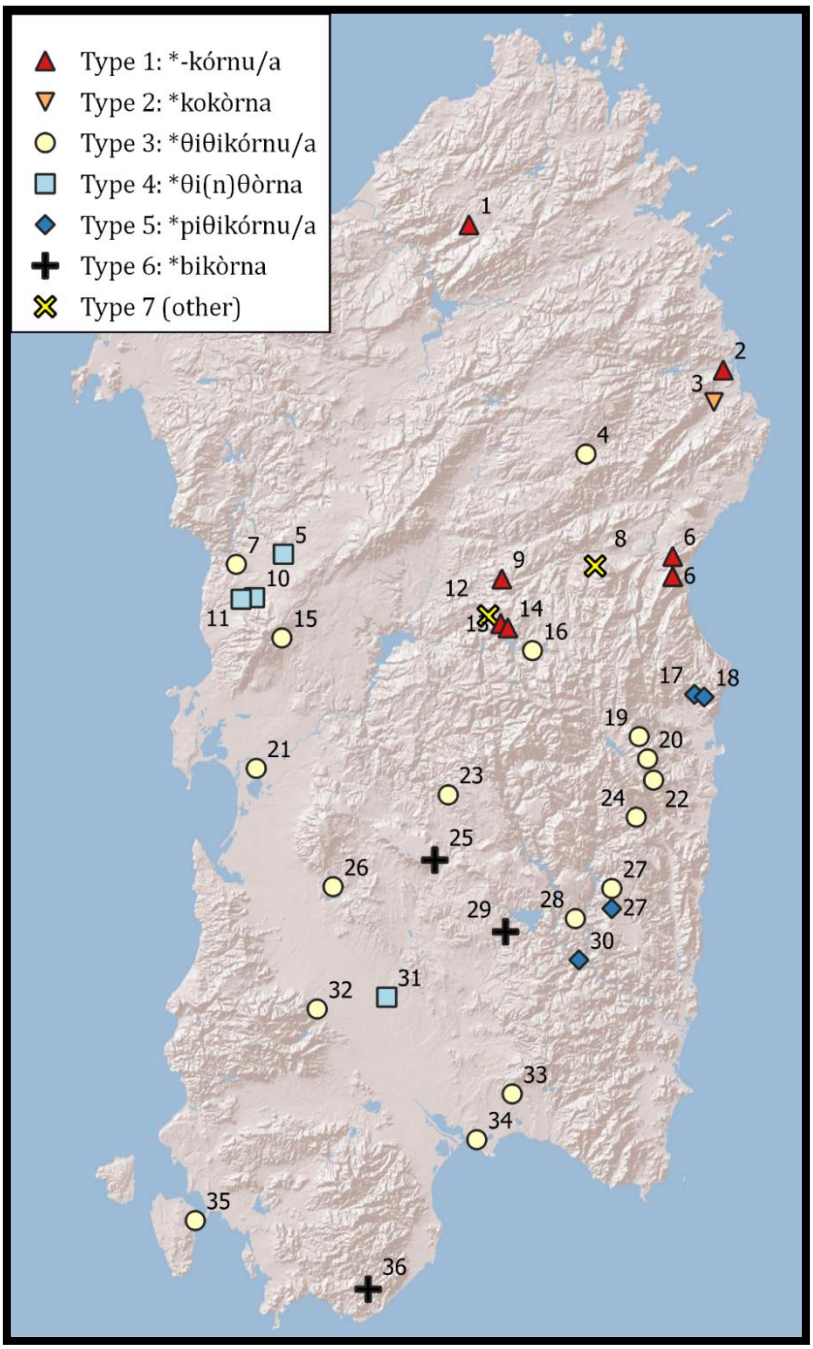

Figure 4.1.15: Slug; snail; cicada

4.1.15 Slug, snail; cicada

Reconstruction	Forms
1: ***-kórnu/a**	DES (II: 422): ***barrakórru*** (1: Tempio Pausania), ***bukkaɣòrra*** (2: Posada), ***tabakòrra, bappakòrra*** (6: Dorgali), ***babbakòrra*** (9: Orani, 13: Ollolai, 14: Gavoi) 'slug, snail'
2: ***kokòrna***	DES (II: 422): ***kokòrra*** (3: Siniscola) 'slug, snail'
3: ***ʔiʔikórnu/a**	DES (II: 422): ***sittsiɣórru*** (Camp.) ***tsintsiɣórru*** (15: Santu Lussurgiu), ***činčiʔórru*** (16: Fonni), ***tiʔiɣórru*** (19: Villagrande Strisaili, 20: Árzana), ***sintsiɣórru*** (21: Oristano), ***tsittsiɣórru*** (23: Laconi) 'slug, snail'; ***sissiriɣòša*** (22: Lanusei), ***sittsiɣárru*** (24: Ulassai), ***sittsiɣòlla*** (33: Sinnai) 'cicada (*Cicada sp.*)'; ***tsontsikòrra*** (4: Bitti) 'earwig (*Forficula auricularia*)'; ***tsintsiɣórru*** (7: Planargia) 'longhorn beetle (*Cerambyx sp.* Linnaeus, 1758)'
	AIS (459, 468): [***su***] ***sittsiɣórru*** (26: Mogoro, 28: Escalaplano, 32: Villacidro), [***su***] ***żissiɣórru*** (27: Perdasdefogu), [***su***] ***sitseɣórru*** (34: Cagliari), [***ṣu***] ***siziɣórru*** (35: Sant'Antioco) 'slug, snail'; [***sa***] ***tsintsikòrra*** (4: Bitti) 'earwig (*Forficula Auricularia*)'
4: ***ʔi(n)ʔòrna***	DES (II: 422): ***tsintsòrra*** (5: Sindía), ***suntsòrra*** (10: Scano di Montiferro), ***tsontsòrra*** (11: Sennariolo), 'slug, snail'; ***tsitárra*** (31: Samassi) 'cicada (*Cicada sp.*)'
5: ***piʔʔikórnu/a**	DES (II: 422): ***piʔʔiɣórru*** (17: Triei, 18: Baunei) 'slug, snail'; ***piśiɣò*** (27: Perdasdefogu) 'longhorn beetle (*Cerambyx sp.*); ***pittsiɣòrra*** (30: Ballao) 'cicada (*Cicada sp.*)'
6: ***bikòrna***	DES (II: 422): ***biɣòrra*** (25: Gesturi), ***bibbiɣòrra*** (29: Siurgus, 36: Domus de Maria) 'cicada (*Cicada sp.*)'
7	DES (II: 422): ***korroβákka*** (8: Oliena), ***pottaminkòrras*** (11: Olzai) 'slug, snail'

The diversity of the forms presented above is to an extent the result of numerous folk-etymologies, but most forms appear to contain *kórru*, *kòrra* 'horn' < Lat. *cornū*, pl. *cornua* 'horn', presumably because of snails and slugs' notable eye stalks.[140] We can distinguish no less than seven different formation types. The forms of type 1 synchronically look like compounds of *kòrra* 'horn' as their second part; among the first parts of these compounds we can identify at least *bukka-* 'mouth' (2: *bukkaɣòrra*) and *babbóị, babba-* which is a generic term for insects and other invertebrate animals (DES I: 161).

[140] The notion 'horn' does probably not refer to snails' shells, as these forms also refer to shell-less slugs and to various species of insect.

The origin of *barra-* in *barrakórru* (1) is less certain, but it could be something like *barrákka* 'shack, hut'. Of type 2, we find only *kokòrra* in Siniscola (3). While Wagner treats it together with the words for 'slug, snail' listed above, the similarity of this word to *kukúrra* 'caterpillar, earwig' (§ 4.1.4) is evident and it is probably best explained as a blend of *kukúrra* with *-kòrra* found in words for 'slug, snail'.

In type 3, we find *-ɣórru* (< *kórru*) preceded by *tsintsi-*, *tiϑi-*, *sintsi-*, *sittsi-* etc. These forms refer to 'slug, snail' in some places, but to 'cicadas', 'longhorn beetles' and 'earwigs' in other places. Wagner (DES II: 422) explains *sittsiɣórru* from *sèttsiri* 'to sit, to put' + *kórru* 'horn', i.e. "put the horns (vel sim." In order for this to be true, we must assume a secondary raising of **-e-* > *-i-* (in all forms) and *-n-*insertion (in many of the forms), which are both common developments. However, more difficult to explain from *sèttsiri* are Ogl. *tiϑiɣórru* 'slug, snail' (19: Villagrande Strisaili, 20: Árzana), which appear to go back to **ϑiϑikórnu*, and Bittese (4) *tsontsikòrra* 'earwig', with an *-o-* of unclear origin. This form from Bitti is rather reminiscent of those belonging to type 4, where we find *tsintsòrra*, *suntsòrra*, *tsontsòrra* 'slug, snail' and *tsitárra* 'cicada'. While *tsitárra* (31: Samassi) was probably influenced by Sp. *cigarra* (DES II: 422; cf. *tabakòrra* [6: Dorgali] with influence from *tabákku* 'tobacco'), this does not work for *tsintsòrra*, *suntsòrra* and *tsontsòrra*. The *-ts-* of these forms can hardly go back to **-ϑ-*, as this would regularly yield ***-t-* in Logudorese (Wagner 1941a: 295). The fact that there is no clear regular source of *-ts-* in Logudorese (other than loanwords), makes it attractive to somehow derive these forms from those of type 3 which are mainly found in Campidanese, where the *-ts-* is an inherited phoneme. Incidentally, Logudorese also has *sèttsere* 'to sit, to put' with *-ts-*, which is probably a loan from Italian dialects (DES II: 399). The development *tsintsòrra* << **tsintsiɣòrra* is not quite regular, however.[141]

In type 5, we find *piϑϑiɣórru* 'slug, snail' and *pittsiɣòrra* 'cicada', which are influenced by *píϑϑu*, *píttsu* 'point, tip'. There is also Perdasdefogu (27) *pisiɣò* 'longhorn beetle' (which can go back to **piϑikò*). It is unclear whether it is identical to *piϑϑiɣórru* with an apocopated final syllable, or whether it is an archaic manifestation that has not been influenced by *kórru*, *kòrra*. Type 6 *biɣòrra* 'cicada' may simply be from Lat. *bicornis* 'with two horns', with reduplication in *bibbiɣòrra*. The two forms in type 7 are unrelated, isolated formations. *Korroβákka* (8: Oli-

[141] The *-u-* and *-o-* in *suntsòrra* and *tsontsòrra* could be the result of pretonal vowel assimilation (Wagner 1941a: 21–22), but this still leaves us with Bittese (4) *tsontsikòrra* 'earwig', for which assimilation is not a satisfactory explanation.

ena) 'slug, snail' is *kórru* 'horn' + *vákka* 'cow', and Wagner (DES II: 422) analyses *pottaminkòrras* (11: Olzai) 'slug, snail' as transparently meaning 'carry me on the horns'.

The question that remains is whether this diversity of forms is the result of several independent formations involving *kórru*, *kòrra* 'horn', or if there is more going on. Some of the forms discussed above are formally rather difficult to explain as a simple compound, such as *tsontsikòrra* (4: Bitti) 'earwig' and *piśiɣò* (27: Perdasdefogu) 'longhorn beetle'. Moreover, while the idea of 'horns' is perhaps understandable in the case of slugs, snails, earwigs and longhorn beetles, cicadas have very small antennae but are nevertheless frequently the referent of the same word. It is possible that the introduction of the word for 'horn' was caused only in part by the semantic association of slugs and snails' eyestalks and the antennae of certain insects to horns, and in part by phonological assonance to the same word for 'horn'. In light of the abundance of forms and the obvious role that folk etymology must have played, it is possible that these forms have one or more pre-Roman words at their base, which were then subject to various secondary remodelings. However, the diversity of forms makes it impossible to establish what the original shape of this word of these words might have been.

4.1.16 Tick (swollen with blood)

Reconstruction	**Forms**
1: *katená(n)kla	DES (I: 319): ***cadinancia*** (Sass.), ***kaðenánča*** (2: Ploaghe, 18: Lanusei), ***kaðranánka*** (Goceano reg., 3: Posada), ***katranákka*** (4: Siniscola, 5: Bitti), ***graðanánka*** (6: Nuoro), ***kaðrenánka*** (Planargia reg., 8: Macomer), ***kaðenánka*** (9: Santu Lussurgiu, 10: Norbello), ***kraðanákka*** (11: Fonni) 'swollen tick (*Ixodida sp.* Leach, 1815)'
	AIS (475): 1: [***lu***] ***ggaddenánču*** (1: Sassari) 'id.'
2: *kánkala	DES (I: 319): ***kánkala*** (7: Dorgali, 14: Baunei, 15: Desulo, 16: Tortolì, 19: Gairo, 21: Perdasdefogu, 22: Escalaplano) 'id.'
3: *katánka	DES (I: 319): ***kaðánka*** (12: Milis, 13: Busachi, 17: Oristano, 20: Mogoro, 23: Villacidro), ***kardánka*** (24: Cagliari), ***garánka*** (25: Sant'Antioco) 'id.'

For the word for 'tick swollen with blood', which is differentiated from Nuor. *erikína*, Log. *eriɣína*, Camp. *arrežíni* etc. 'a tick that is not swollen with blood yet' (DES I: 491–92), we find three different types. The forms of type 1 go back to something like **katená*(*n*)*kla*, which is why Rolla (1895: 22) and REW (1764) derive it from Srd. *katèna*, *kaðèna* etc. 'chain' (with derivations meaning 'necklace'

among other things) < Lat. *catena* 'chain'. Wagner (DES I: 319), however, does not believe that a derivation *catena* can explain all the variation we find among the forms, even though it might have influenced some of them. Type 2 *kánkala* is largely restricted to the Ogliastra (with outliers in 7: Dorgali and 15: Desulo). It is compared by Wagner (DES I: 319) to Sp. *cáncano* 'louse', whose origin is not fully clarified either, but which may be from Lat. *cancer* 'crab; cancer' (DCEC I: 797). Wagner (DES I: 319) thinks this type should be separated from the forms for 'tick' in types 1 and 3, with which I agree. Finally, type 3 is attested in the Campidanese dialects and requires a reconstruction **katánka*.

The question is how much of the similarity between the three different formation types for 'swollen tick' is due to etymological relationships, and how much to folk-etymological influences. It seems in any case sensible to connect type 1 **katená(n)kla* and type 3 **katánka*. It is not clear, however, whether both go back to Lat. *catena* 'chain', in which case type 3 must somehow have been altered, or whether type 3 is in fact more original and type 1 **katená(n)kla* should be explained as the result of folk etymology. The link to *catena* 'chain' may perhaps be explained by an association of swollen ticks to beads on a necklace. For the use of a word for 'necklace' to refer to bodily growths, cf. also Camp. *kannákka* 'necklace', which is also used to refer to the wattles hanging from goats' necks (DES I: 284).

Type 2 **kánkala*, despite its resemblance to Sp. *cáncano* 'louse', cannot easily be explained as a Spanish loan. Nor is a direct inheritance from Lat. *cancer* 'crab, cancer' plausible, for the same reasons mentioned by Wagner (DES I: 281) for *kánkaru* 'cramp', which is from It. *canchero* 'id.'.[142] Wagner (DES I: 281) notes that there is a great deal of confusion between Srd. *kánkaru* 'cramp' and Srd. *kánkaru* 'hinge', which is the reflex of late Gr. κάγχαλος 'hinge' (cf. REW 1575). Incidentally, the meaning 'hinge' is also attested for some Sardinian derivatives from *catena* 'chain': NLog. *kaðenále* 'hinge, metal strip that is part of a hinge'[143]. It is unclear to me whether the formal similarity of the various Sardinian forms for 'tick' to words for 'hinge', 'chain' and 'cramp' is purely due to coincidence or due to a semantic association.

[142] Cf. Craddock (1969: 87–123) for the specificities of the Romance reflexes of Lat. *cancer* 'crab, cancer'.

[143] Cf. AIS (881-883) for detailed explanations and drawings of different types of hinges and their various parts. Unfortunately Sardinia is not represented in the maps in question.

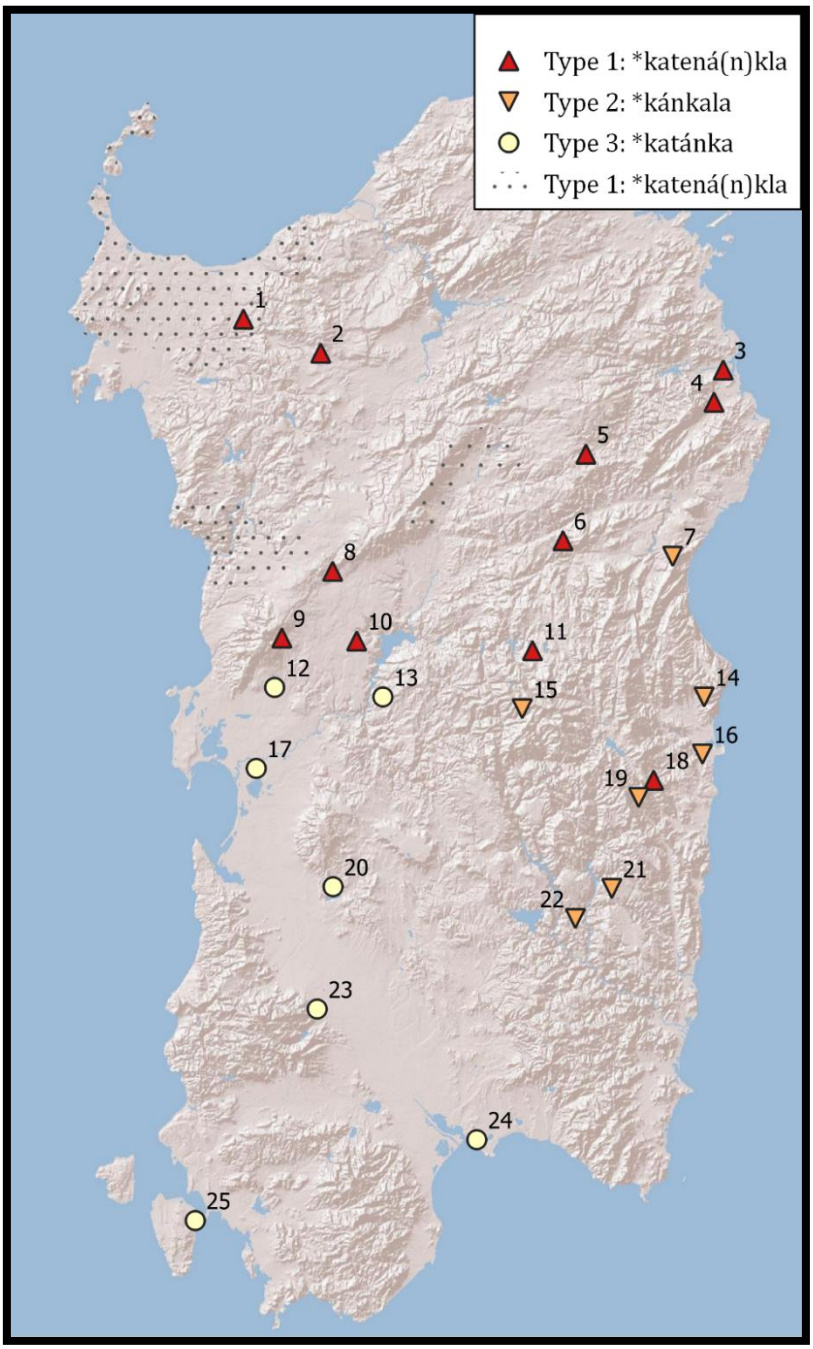

Figure 4.1.16: Tick

I do in any case believe that is not necessary to ascribe the attested diversity in these words for 'tick' to a pre-Roman origin, as Wagner (DES I: 319) tentatively does. Types 1 and 2 probably have a relation to Lat. *catena* 'chain' and *cancer* 'crab, cancer' respectively. Type 3 is more difficult to explain, and might have served as the base form from which type 1 was derived by folk etymology. Besides the absence of a clear etymology, there is no positive evidence for a pre-Roman origin however; none of the forms of type 3 show any irregular correspondences to each other that cannot be explained internally.

4.1.17 Worm

Reconstruction	**Forms**
***ϑVrro(C)íku/a**	DES (II: 552): ***tserríɣu/a, tsirríɣu, tsorroíɣu*** (Camp.), ***tserríɣa*** (1: Norbello), ***ϑorroíɣu*** (2: Baunei), ***tsarríɣa*** (3: Oristano) 'worm living inside horse intestines; worm found in rotting meat'
	Marcialis (2005: 70): ***zorroiga*** (Sulcis reg.) 'earthworm (*Lumbricus terrestris*)'
	AIS (457): [***su***] ***tsorroíɣu*** (4: Escalaplano) 'earthworm (*Lumbricus terrestris*)'

The vowel sequence *-oí-* in the forms *ϑorroíɣu* and *tsorroíɣu* seems original, and was contracted in the other dialects (cf. Wagner 1941a: 46–48). The original vocalism of the pretonic syllable is more difficult to establish, as both assimilations and dissimilations to the stressed syllable are common (cf. Wagner 1941a: 24–33). The vowel hiatus *-oí-* could be the result of the loss of an original *-b-*, *-v-*, *-d-* or *-g-*, but this is not certain.[144] Thus reconstructible as **ϑVrro*(*C*)*íku*, I agree with Wagner (DES I: 552) that this word is probably of pre-Roman origin. Reasons for this are its peculiar syllable structure, as well as the presence of **ϑ-* (§ 8.1.1). However, due to the relative uniformity in the attested forms of this word, it is impossible to reliably discern the presence of the **ϑi-*prefix (§ 9.1.2) or any other morphological elements.

[144] Note the phonetic similarity to forms like *soloíɣa* 'black widow' (§ 4.1.14), where we do have evidence for the presence of an original consonant in this position in classically attested *solipūga*, *solifūga* (which can however not be the direct source of the modern Sard

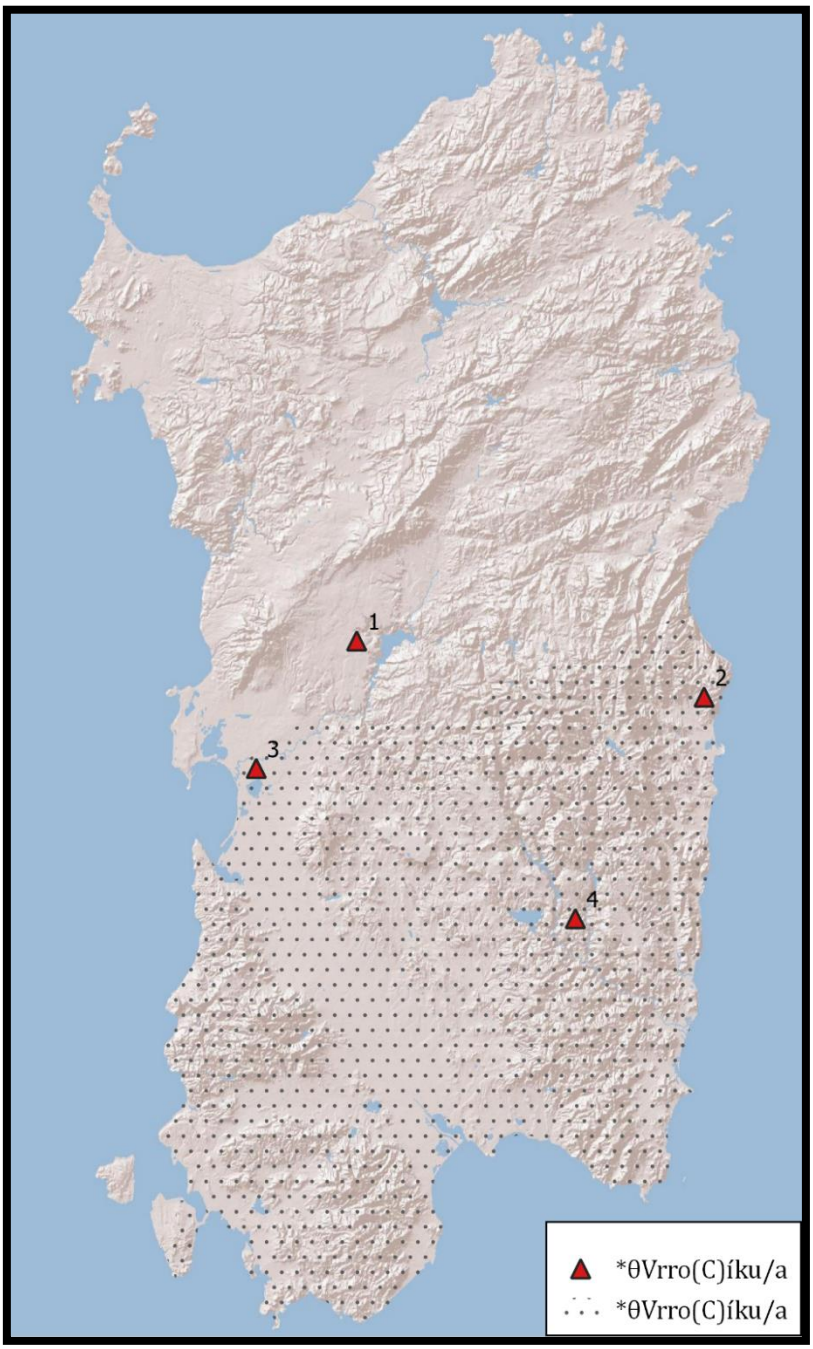

Figure 4.1.17: Worm

4.2 Fish

4.2.1 European eel (*Anguilla anguilla*)

Reconstruction	Forms
***ʔingòrra**	DES (II: 592): ***tsingòrra*** (Camp.), ***dzingòrra*** (1: Cagliari) 'young eel (*Anguilla anguilla* Linnaeus, 1758)'
	Marcialis (2005: 70, 84, 114, 121, 150, 184): ***zingòrra*** (1: Cagliari) 'eel (*Anguilla anguilla*)'[145]
	Puddu (2023: s.v. *tzingòrra*): ***tzingòrra, zangaúrra, zingòrra*** 'young eel (*Anguilla anguilla*)'

Wagner (1997: 160) compares this word to Akk. *singurru* 'a kind of fish' (also *zingurru*, *ziggurru*; cf. CAD XV: 284–285), which Zimmern (1917: 52) connects to Mandaic *singūrā*. Formally these form a near-perfect match, though the meanings of the Semitic forms are only vaguely defined.[146] This leads Wagner (1997: 160; DES II: 592) to conclude that the Sardinian word may be a Punic loan (cf. § 10.1), even though the Punic word for 'eel' is not actually attested. Paulis (1990: 621–623) rejects Wagner's Punic etymology. He objects that the Akkadian word has an unclear meaning and that it could be a Sumerian loan. Moreover, he argues that it is unlikely that a pre-Roman name for a fish would have been preserved for so long, while most of the modern Sardinian maritime vocabulary was borrowed from Italian, Catalan and Spanish. He instead proposes *tsingòrra* to be a loan from It. *ceca* 'young eel' >> Srd. **tsèɣa* >> **tsènga*, which was subsequently suffixed with *-òrra*, present in pre-Roman toponyms and Sardinian surnames. Finally, the pretonic *-e-* of this **tsengòrra* would have been fronted, yielding *tsingòrra*.

The number of assumptions needed to derive the Sardinian word from Italian makes Paulis' etymology rather doubtful. It should moreover be noted that there are no clear cases in which an original voiceless stop is first voiced intervocalically and then acquires a preceding nasal, as would have to be assumed to explain Srd. *tsing-* < **tseng-* from It. *ceca*. Wagner's (1941a: 219–223) section on the topic only contains examples of *-n-*insertion where the original quality of the stop is preserved (e.g. *munkaðòre* vs. *mukkaðòre* << Cat. *mocador* 'handkerchief'), and sometimes even examples where the inserted nasal seems to have blocked the expected intervocalic sonorization (e.g. *kalankáṡu* vs. *kalaɣáṡu* 'butterfly').

[145] Marcialis gives the definition "*Anguilla acutirostro*, Ceriuola. Pesce." *Anguilla acutirostris* Risso, 1827 is a synonym of *Anguilla anguilla*, so this probably refers to the European eel (*Anguilla anguilla*) (cf. Anguilla acutirostris Risso, 1827, in GBIF Secretariat 2023).

[146] I have not been able to verify the Mandaic form, let alone its meaning.

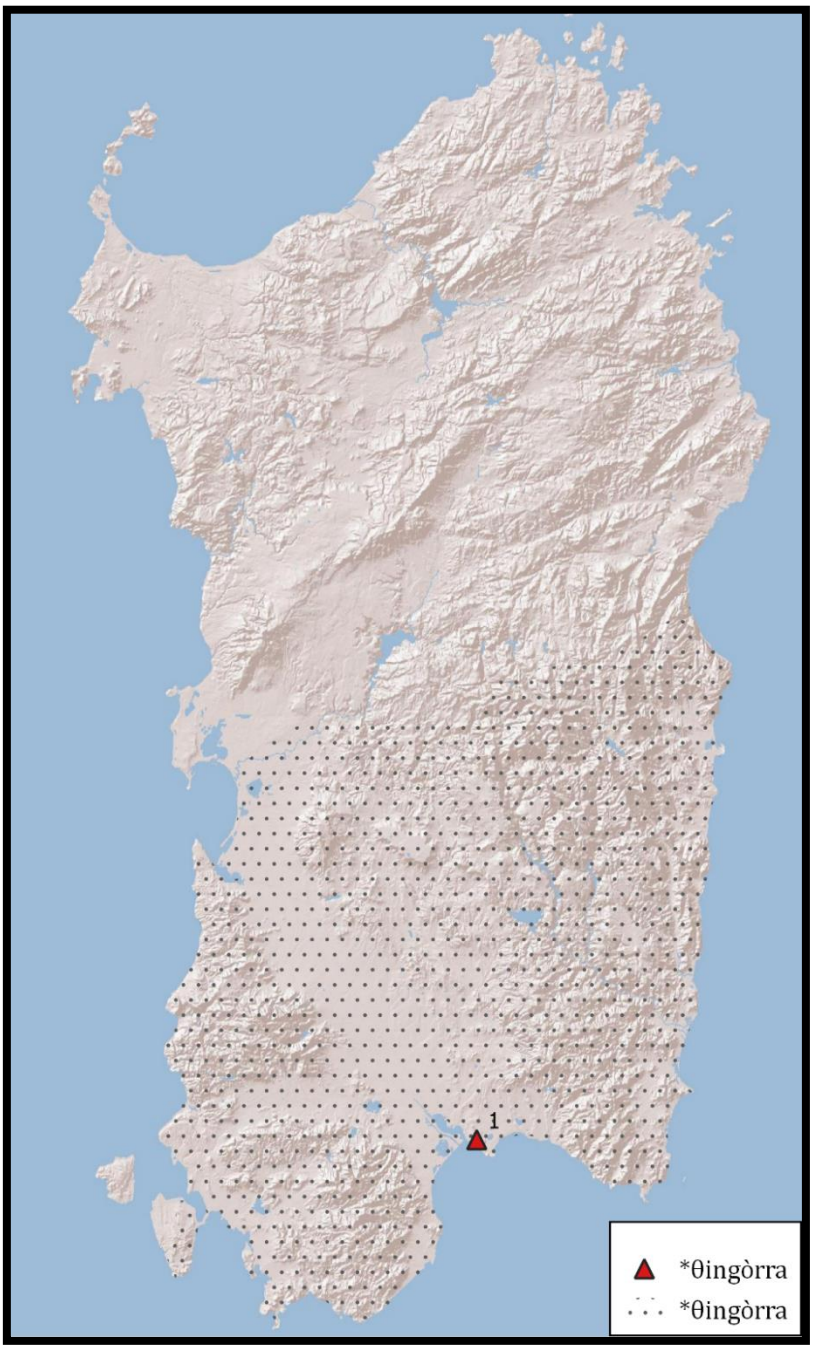

Figure 4.2.1: European eel (*Anguilla anguilla*)

I further think that the analysis of *-òrra* in *tsingòrra* as a "pejorative" suffix is all but evident, especially since the evidence for this proposed suffix is limited to names of places and persons, rather than names of animals. As for the likelihood that a fish name would have survived from (pre-)Roman times up to the modern day, it is important to realize that eels are not primarily maritime fish, but spend the biggest part of their life in fresh water. The irrelevance of the often-cited Sardinian aversion towards the sea (Wagner 1997: 213; Paulis 1990: 621) for the discussion of the designations of eels is demonstrated by the fact that the common Sardinian word for eel, *ambíḍḍa*, *angwíḍḍa*, goes back directly to Lat. *anguilla* 'eel'. If a Punic word for 'eel' entered early Sardinian Latin, it was no less likely to survive than Latin *anguilla*.

I agree with Paulis that the Punic etymology is rather uncertain, as the only pieces of evidence are an Akkadian and a Mandaic word referring to an undefined species of fish. Nevertheless, I do not think a derivation from It. *ceca* 'young eel' is more convincing. Until the Punic word for 'eel' is actually attested, this word must be regarded as no more than a possible example of a Punic loan into Sardinian Romance.

4.2.2 Trout fry (*Salmo cettii*)$_1$

Reconstruction	**Forms**
***grórі**	DES (I: 594): ***gròri, gròli*** (1: Bitti) 'trout fry (*Salmo cettii* Raffinesque 1810)'
	AIS (528): [***su***] ***grórі*** (1: Bitti) 'id.'
	Puddu (2023: gróli): gróli, grórі 'id.'

The forms recorded by DES have open *-ò-*. However, in AIS (528), the Sardinian entries of which were also compiled by Wagner, we find closed *-ó-*. This is also confirmed by Puddu (2023: *gróli*), and is in fact expected, since final *-i* conditions closing of a preceding *-e-* or *-o-* to *-é-*, *-ó-* (cf. Wagner 1941a: 11). The variant *gróli* could be a dissimilation from *grórі*. Wagner (DES I: 594) denotes this word as "probably pre-Roman", without giving any arguments. Besides the lack of an etymology, one argument could be this word's ending in *-i*, which is rare in the inherited lexicon (§ 8.1.5.1). While this could also be the result of a loan, there is no clear source. On the basis of this word's attestation in a single village only, it is hard to say anything definitive about this word. The specific fish species is probably the Mediterranean trout (*Salmo cettii*), a freshwater fish.

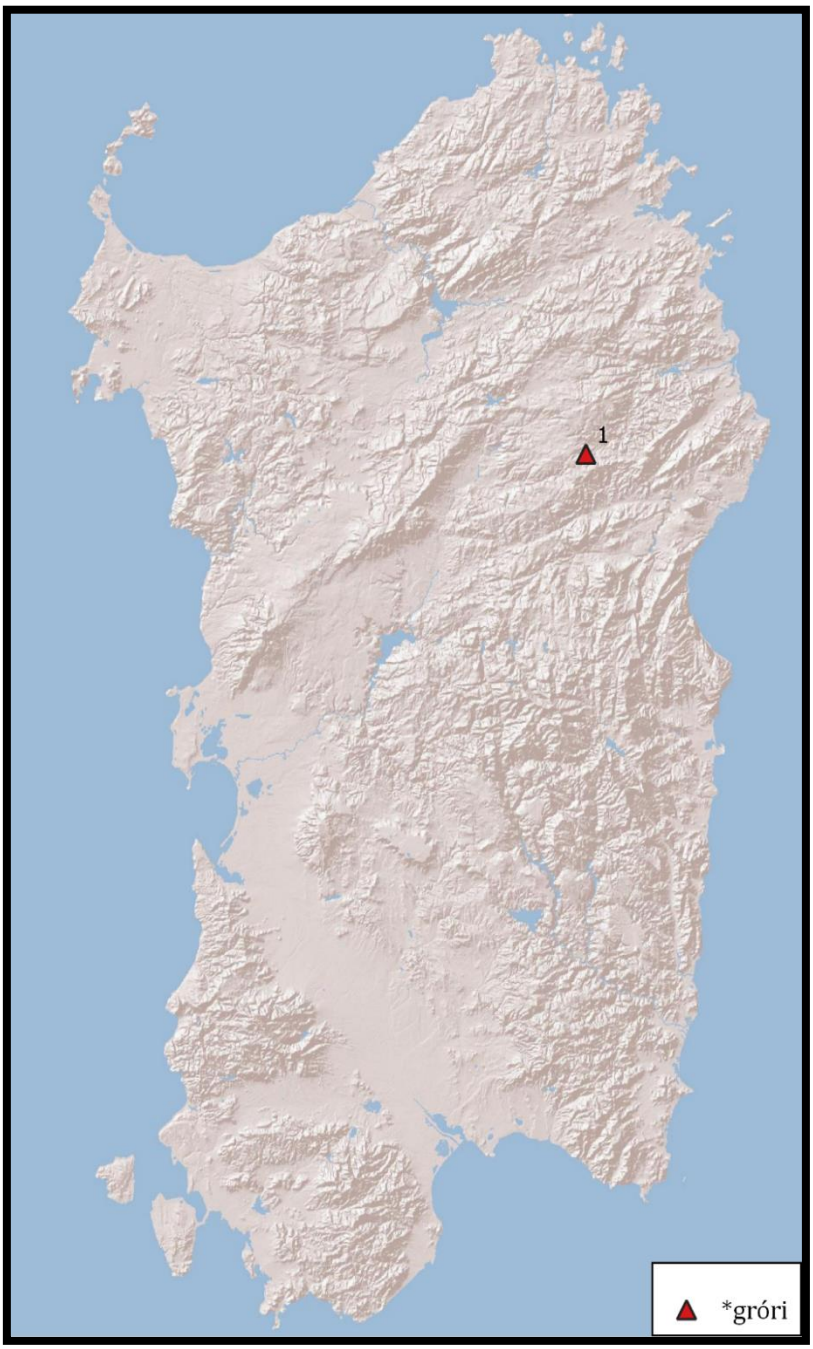

Figure 4.2.2: Trout fry (*Salmo cettii*)

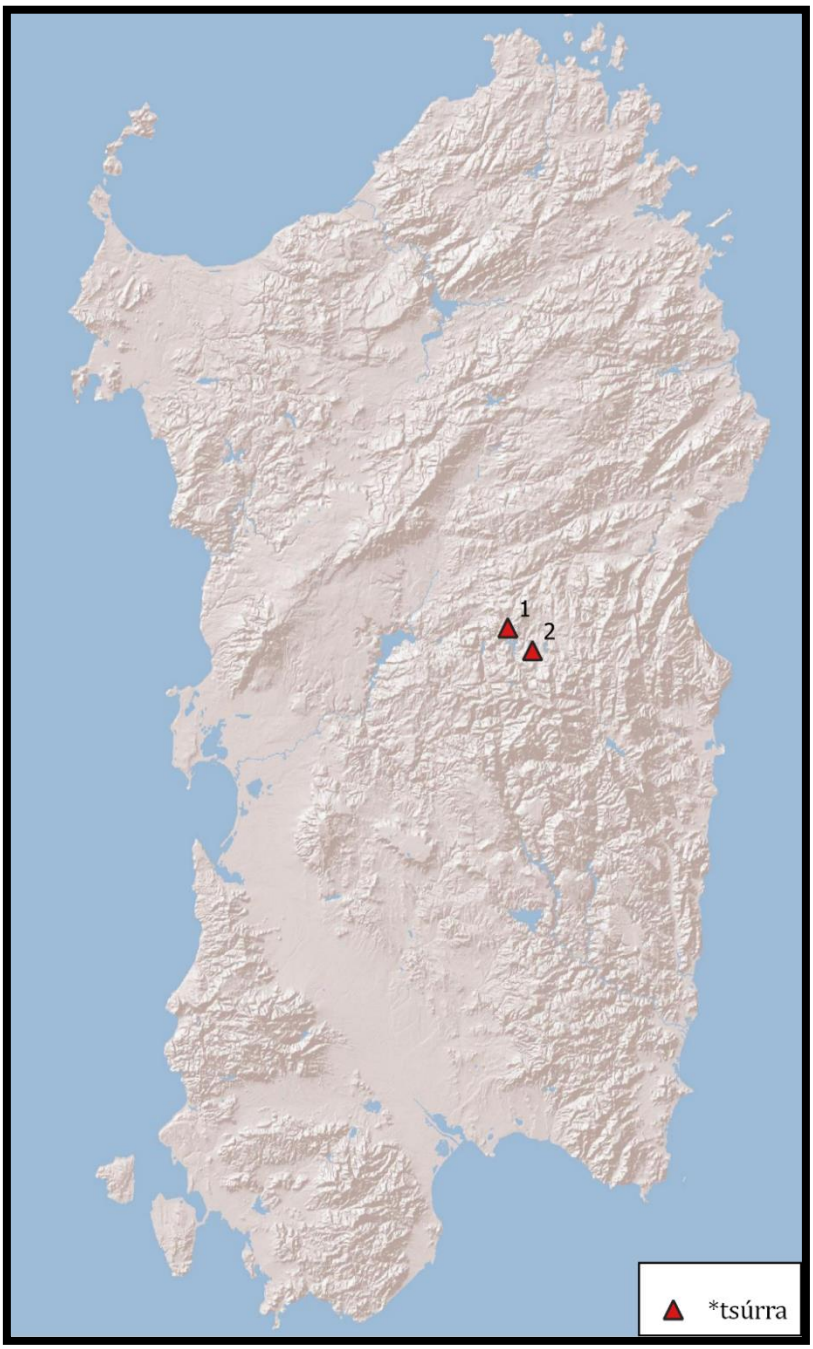

Figure 4.2.3: Trout fry (*Salmo cettii*)

4.2.3 Trout fry (*Salmo cettii*)$_2$

Reconstruction	**Forms**
***tsúrra**	DES (II: 602): ***tsúrra*** (2: Fonni) 'small trout'
	Marcialis (2005: 71): ***tsurríttu*** (1: Gavoi) 'trout fry'

As in the case of Bittese *gróri* 'trout fry' (§ 4.2.2), these words likely refer to *Salmo cettii*, the Mediterranean trout. And like for *gróri*, Wagner (DES II: 602) considers *tsúrra* 'small trout' to be "probably pre-Roman", without clarifying his reason for this. Initial *ts-* does not have an inherited source in these specific dialects (§ 8.1.4), but there is no other positive evidence for a substrate origin of this word. Its limited attestation hinders any further analysis.

4.3 Reptiles

4.3.1 Gecko (*Tarantola mauritanica*)

Reconstruction	**Forms**
***ϑilíkwu**	DES (II: 483): ***tilíbbu*** (1: Dorgali) 'lizard/spider'

This word probably refers to the Moorish gecko (*Tarentola mauritanica* (Linnaeus, 1758)) (cf. Deplano 2020: 811; Puddu 2023: s.v. *tilíbbu*). Its occurrence in AIS (485) as the name of the tarantula spider may have originated in the fact that It. *tarantola* can refer either to various spider species (including *Lycosa tarantula*, the tarantula), or to the Moorish gecko. All other sources agree that it refers to some kind of reptile. The sequence *-bb-* could go back to original *-*kw-*, leading to a reconstruction **ϑilíkwu*.[147] Initial *ti-* could be an instance of the element **ϑi-* that is found in many Sardinian words of pre-Roman origin (§ 9.1.2). Since it is an isolated form, restricted to the dialect of Dorgali, it is of limited value for any morphological or distributional analysis of the Sardinian substrate.

[147] Original *-*bb-* would likely have yielded **-*pp-* (Wagner 1941a: 192–193).

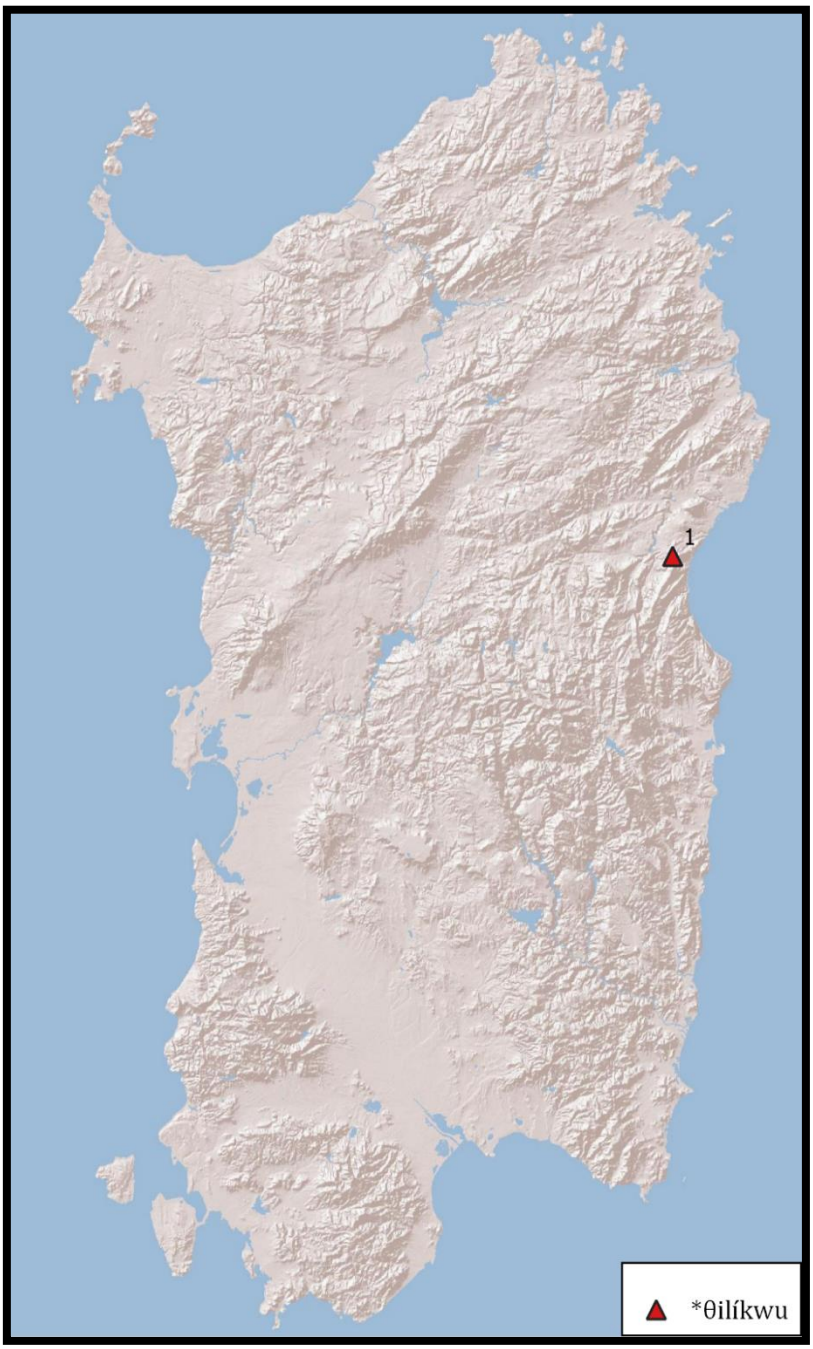

Figure 4.3.1: Gecko (*Tarantola mauritanica*)

4.3.2 Lizard (*Podarcis tiliguerta*)[148]

Reconstruction	Forms
1: *ʤilikèrta	DES (II: 545): ***tiriɣètta*** (Sass.), ***tsirikèlta*** (1: Tempio Pausania), ***tiliɣèɬta*** (2: Ploaghe), ***tiliɣèrta*** (3: Posada, 4: Torpè, 8: Bono, 16: Sennariolo, 17: Cuglieri, 20: Santu Lussurgiu) ***ʤilikèrta*** (5: Siniscola, 6: Bitti, 9: Orosei, 10: Nuoro), ***tiliɣèlta*** (7: Bonorva), ***ʤilikètta***, ***tilikètta*** (11: Dorgali), ***ʤiliʔèrta*** (12: Oliena, 18: Ollolai), ***attiliyèrta*** (13: Macomer), ***ateliyèlta***, ***attaliyèlte*** (15: Scano di Montiferro), ***tsiliyèrta*** (14: Borore, 21: Norbello), ***tsiliyètta*** (25: Busachi, 31: Samugheo) 'Tyrrhenian wall lizard (*Podarcis tiliguerta* (Gmelin, 1789))'
	AIS (449): [***sa***] ***tiliɣèrta*** (13: Macomer), ***tsiliɣèrta*** (20: Santu Lussurgiu), ***tsiʁinyèrta*** (23: Milis) 'id.'
2: *ʤalakèrta	DES (II: 545): ***ʤalaʔèrta*** (19: Gavoi) 'id.'
	AIS (449): ***ʤalaèrta*** (22: Fonni) 'id.'
3: *(*a*)*li/ukèstra*	DES (II: 545): ***luǧèstra*** (24: Triei, 26: Baunei), ***ǧiléstru*** (29: Villagrande Strisaili), ***liǧèstra*** (33: Tortolì), ***arǧiléstru*** (34: Arzana, 36: Lanusei), ***luǧǧèsti*** (38: Gairo), ***čilèstra*** (39: Bari Sardo), ***aliǧèstra*** (43: Perdasdefogu) 'id.'
	AIS (449): ***lučèstra*** (26: Baunei) 'id.'
4: *(*a*)*li/ukèrta*	DES (II: 545): [***s'***] ***aliɣèrta*** (27: Tonara, 28: Desulo), ***lulǧètta*** (30: Aritzo), ***anǧulètta*** (32: Meana Sardo), ***aliɣèlta*** (35: Gadoni), ***aližèrta*** (37: Laconi) 'id.'
5: *kalVkèrta	DES (II: 545): ***kalužèrtula*** (Camp.), ***karǧilètta*** (40: Seui), ***kalužèrta*** (41: Sadali), ***kalažèḍḍa*** (44: Escalaplano), ***kaʁišèḍḍa*** (45: San Nicolò Gerrei), ***kaʔižèḍḍa*** (47: Muravera), ***kožuètta*** (49: Sant'Antioco) 'id.'
	AIS (449): [***sa***] ***ɣawižètta*** (42: Mogoro, 46: Villacidro), [***sa***] ***ɣalužèrtula*** (48: Cagliari), [***sa***] ***ɣalažètta*** (44: Escalaplano) 'id.'

All forms listed above are descendants of Lat. *lacerta* 'lizard'. The reconstruction **ʤilikèrta*, needed for the forms of type 1, is evidence that *ʤi-* may be some kind of prefix (§ 9.1.2). This was first suggested by Guarnerio (1904a: 57–58 fn. 2). Despite the inherited origin of this word, it is hard to explain *ʤi-* in terms of inherited Romance material. However, it should be noted that *-la-* in this word is not preserved in any other Sardinian dialect.

[148] These data are discussed in a similar fashion in Swanenvleugel (2024: 224–226).

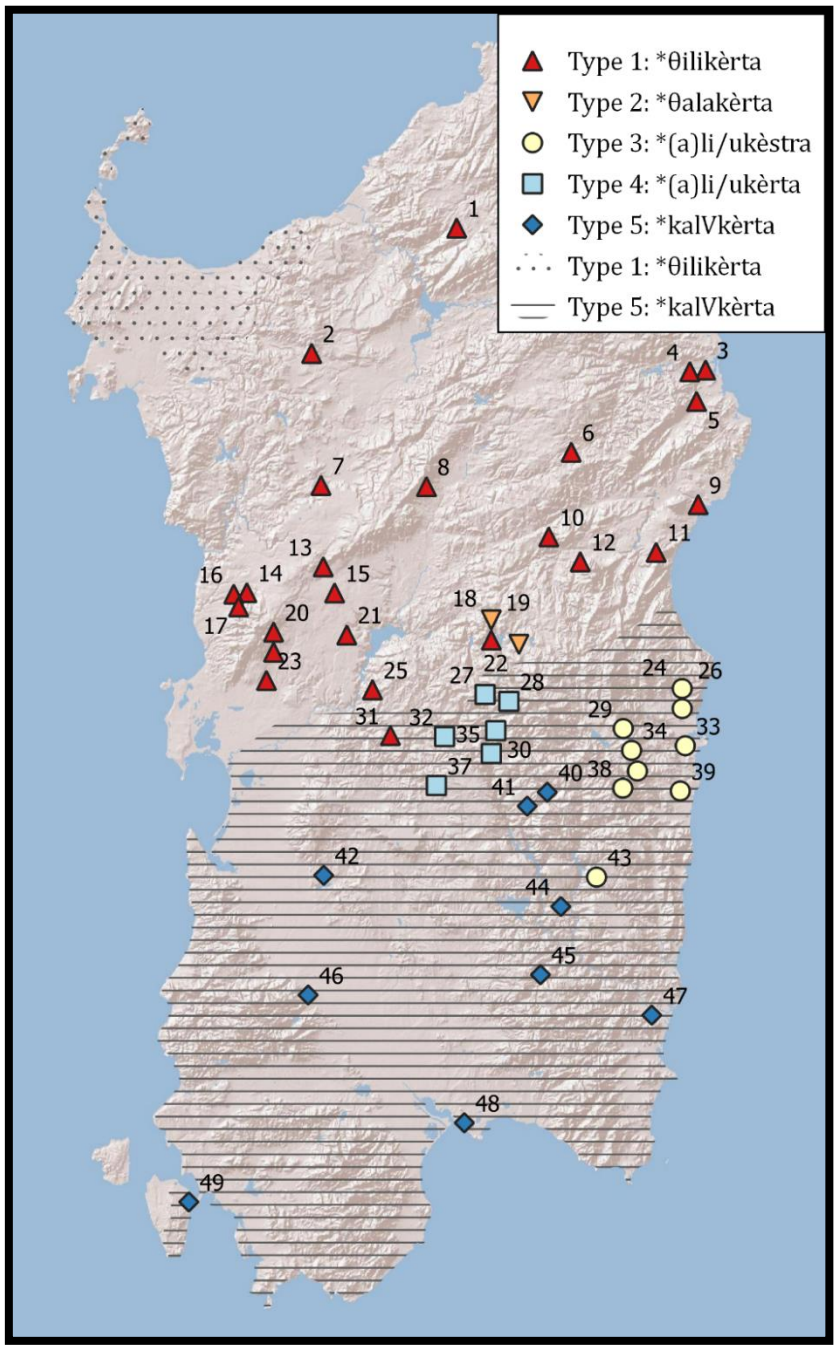

Figure 4.3.2: Lizard (*Podarcis tiliguerta*)

Even the un-prefixed forms have *-li-* or *-lu-*. These aberrant vocalisms are not restricted to Sardinian and are well-attested in French and Italian dialects, where we find similar evidence for **licerta* and **lucerta* (REW 4821; FEW V: 117, 118 fn. 6; Battisti & Alessio 1952: 2276). For Sardinian **ʒilikèrta* too, the sequence *-li-* may thus be of Vulgar Latin origin. Influence from other Sardinian words starting in **ʒili-* is imaginable, cf. *ʒilipírke* (§ 4.1.12), *ʒilikúku* (§ 4.3.4) etc. The forms containing *ka-* (type 5) are most easily explained from an original **kalVkèrta*, again with vowel alternations in the pretonic syllable (cf. § 2.1.1). These forms are found in the Campidanese dialects in the south of Sardinia. Many of the manifestations show various kinds of assimilations and metatheses. It is difficult to assess to what extent the coexistence of forms containing **-la-* (e.g. *kalažèḍḍa*, 44: Escalaplano), **-li-* (e.g. *kaʔižèḍḍa*, 47: Muravera) and **-lu-* (e.g. *kalužèrta*, 41: Sadali) are to be attributed to Vulgar Latin variation between *lacerta*, **licerta* and **lucerta* (REW 4821; FEW V: 117, 118 fn. 6), as pretonic assimilations and dissimilations are very common across Sardinian (Wagner 1941a: 21–33). The remaining forms, that have neither **ʒi-* nor **ka-* (types 3 and 4), can go back to **lukèrta* or **likèrta*. The initial *a-* is due to reanalysis of the vowel of the feminine definite article. In the Ogliastran dialects in the east of Sardinia, *-èrta* has been replaced with *-èstra, -éstru, -ésti* (type 3). This is reminiscent of the sequence *-st(r)-* found in other Sardinian words of uncertain origin, such as *golóstri, golóstịu* 'holly' and *giḍḍóstru/a* 'tree-heath'. This *-st(r)-* has been proposed to be a suffix originating in a substrate language (Bertoldi 1930; van Sluis, Jørgensen & Kroonen 2023: 200–201), but cf. § 9.2.9. Proposals to explain the addition of **ka-* in this word are problematic. Bertoni (1913: 167 fn. 4) suggests that the reflexes of *lacerta* were contaminated with those of *colūbra* 'snake, serpent'. This hypothesis is semantically attractive, but the vocalism does not quite match in all locations (e.g. Escalaplano (44) *kalažèḍḍa, kalažètta* 'lizard' next to *kolóru* 'snake'; AIS 452). Böhne on the other hand (1950: 109) proposes a contamination with *kalái* 'to descend'. This idea is far-fetched, especially on the semantic side, and cannot be accepted.

In conclusion, we need to assume at least three different reconstructions to account for the attested forms: **ʒi-la/ikerta* (types 1, 2), **li/ukèrta* (types 3, 4), and **ka-la/i/ukerta* (type 5). The variants of this word clearly show the secondary addition of **ʒi-* and **ka-*, neither of which are easily explained in Romance terms, to inherited *lacerta* 'lizard'. This word therefore provides plausible evidence in favor of the identification of **kV-* and **ʒi-* as pre-Roman morphemes (§ 9.1.1, § 9.1.2), as well as for the productivity of pre-Roman morphology on inherited words.

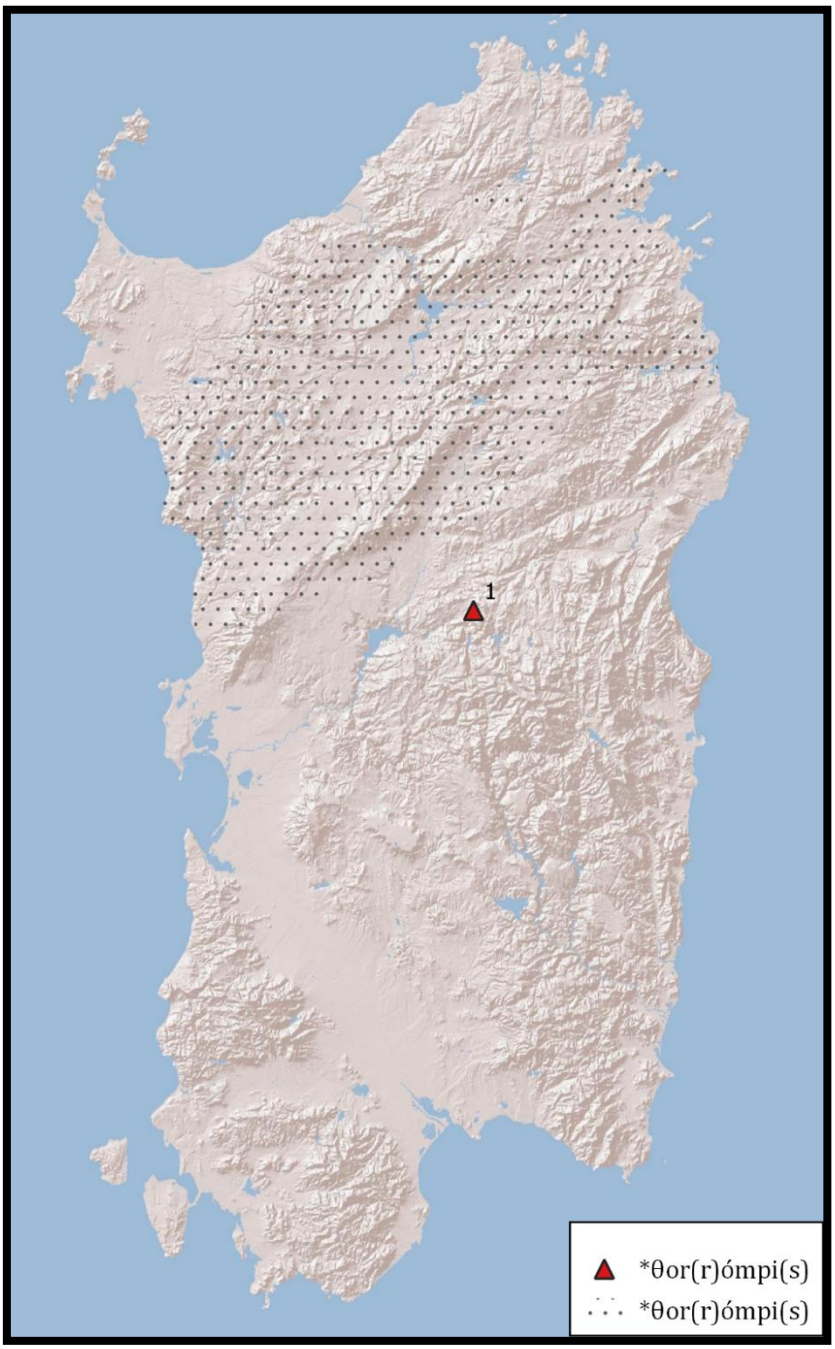

Figure 4.3.3: Lizard; snake

4.3.3 Lizard; snake

Reconstruction	Forms
***ϑor(r)ómpi(s)**	DES (II: 600): ***tsorómpis*** (1: Olzai) 'lizard'
	Marcialis (2005: 70): ***zorrómpi*** (Log.) 'snake'

Unfortunately, Marcialis (2005: 70) does not provide precise geographical information for Log. *zorrómpi* 'snake'. In *tsorómpis* (1: Olzai), initial *ts-* can reflect **ϑ-* (Contini 1987: 548). The ending *-is* in *tsorómpis* (1: Olzai) is also found in some other suspected pre-Roman words (§ 9.2.8). The lack of an etymology and the peculiar phonetic structure of this word point to a pre-Roman origin. I agree with Wagner (DES II: 600) that Jud's (1942: 353) comparison of this word to Arag. *sargantana* and Bq. *suranguilla* 'lizard' is not convincing.

4.3.4 Skink (*Chalcides ocellatus tiligugu*); slug[149]

Reconstruction	Forms
1: *ϑilikúku	DES (II: 546): ***attiliɣúɣu*** (2: Osilo, 5: Ozieri), ***tirigúru*** (3: Sassari), ***tiliɣúɣu*** (4: Ploaghe, 11: Macomer, 8: Bosa), ***ϑilikúkku*** (6: Bitti, 12: Orani), ***attaliɣúɣe*** (Sennori) 'skink (*Chalcides ocellatus tiligugu*)'; ***tsirikúkku***[150] (1: Tempio Pausania), ***tiliɣúɣu*** (11: Macomer, 14: Santu Lussurgiu), ***tsiliɣúɣu*** (13: Borore), ***ϑilikúkku*** (Orani) 'slug'
	AIS (449): [***su***] ***tsiʁiɣúɣu*** (16: Milis) 'skink (*Chalcides ocellatus tiligugu* (Gmelin, 1789))'
2: *ϑalakúku	DES (II: 546): ***ϑalakúkku*** (7: Nuoro), ***ϑalakúkka*** (9: Dorgali), ***ϑelakúkka*** (10: Oliena), ***ϑalakúʔa*** (15: Fonni), ***tsaβaɣúɣu*** (21: Oristano) 'skink (*Chalcides ocellatus tiligugu*)'
	AIS (449): [***sa***] ***ϑaϑϑalakúkka*** (9: Dorgali) 'skink (*Chalcides ocellatus tiligugu*)'
3: *ϑiϑϑilúka	DES (II: 546): ***sintsilúɣa*** (19: Aritzo) 'skink (*Chalcides ocellatus tiligugu*)'

[149] These data are discussed in a similar fashion in Swanenvleugel (2024: 246–249).

[150] Wagner (DES II: 546) gives *tsilikúkku* for Tempio Pausania and cites AIS 461. However, AIS has *tsirikúkku*.

4: *ʤaʤʤalúka	DES (II: 546): ***tsantsallúɣa*** (20: Meana Sardo), ***tsattsalúɣa*** (22: Laconi), ***satsaúɣa*** (27: Villacidro), ***sattsalúɣa*** (Camp., 29: Cagliari), ***sossòɣa*** (Sulcis reg.) 'skink (*Chalcides ocellatus tiligugu*)'; ***sattsalúɣa*** (18: Busachi), ***tsattsalúɣa*** (22: Laconi), ***sassalúɣa*** (23: Gairo), ***sattsaɣúʁa*** (28: Dolianova) 'slug'
	AIS (449, 461): [***sa***] ***ʤaʤʤalúkka*** (17: Baunei), [***sa***] ***sattsaúɣa*** (24: Mogoro), [***sa***] ***ṡassalúɣa*** (25: Perdasdefogu), [***sa***] ***sattsalúɣa*** (26: Escalaplano), [***sa***] ***soṡòɣa*** (30: San Antioco) 'skink (*Chalcides ocellatus tiligugu*)'
5: *babbalúka	AIS (461): [***sa***] ***babbalúɣa*** (26: Escalaplano) 'slug'
6: *alikúkku	Puddu (2023: s.v. *alicúcu* 1): ***alicúcu*** 'slug'

Besides the forms listed, Marcialis (2005: 17, 65, 69) also has *attilingia* (Pattada), *azzacauga* (without location), *tiligulu* (Bosa) and *zaraclua* (Olzai). Although all have a similar structure, they are probably not directly cognate with **ʤilikúku*. *Attilingia* from Pattada probably belongs together with words for 'earthworm', cf. *tilíŋğa* (Oschiri) 'earthworm' (cf. § 4.1.10). *Azzacauga* is difficult to judge, especially because its location is unknown. We might speculate that it goes back to **ʤakalúka*, metathesized from **ʤalakúka* etc., but the loss of the *-l-* (which also occurs in Oristanese *tsaβaɣúɣu*) is difficult to reconcile with the preservation of the voiceless stop *-k-* (spelled *-c-*). *Tiligulu* is cited by Guarnerio (1904a: 66) as "*tilígulu*" (with preparoxytone accent), which makes a connection to **ʤilikúku* difficult, unless some analogy to the diminutive suffix *-ulu* is involved. Guarnerio's (1904a: 61) proposal of *zaraclua* (Olzai) as a metathesis from **zaraluge* is neither convincing, nor would it suffice to fully explain this form.

Of the words that are listed above, the majority can be separated into two main groups. To the first group belong all forms of type 1 **ʤilikúku* and type 2 **ʤalakúka*. These bear an obvious similarity to *ʤilikèrta* and *ʤalaʔèrta* 'lizard' (§ 4.3.2), and are suspect of containing the same pre-Roman *ʤi-*prefix (§ 9.1.2). The second group includes the forms of type 3 **ʤiʤʤilúka* and type 4 **ʤaʤʤalúka*, which are found in the Campidanese dialects. Despite the phonetic similarities between **ʤili-/ʤalakúku* and **ʤaʤʤalúka*, and despite their shared polysemy 'skink; slug', it is very difficult to reconcile the two groups. For this reason, I prefer to treat them as two unrelated words. AIS (461) further presents *babbalúɣa* (26), whose first part is related to *babbói̯*, *bobbói̯* 'generic term for invertebrates' (DES I: 161). Puddu (2023: s.v. *alicúcu*) also lists *alicúcu* 'kind of long snail, but without a shell', which is however absent from Wagner and the AIS.

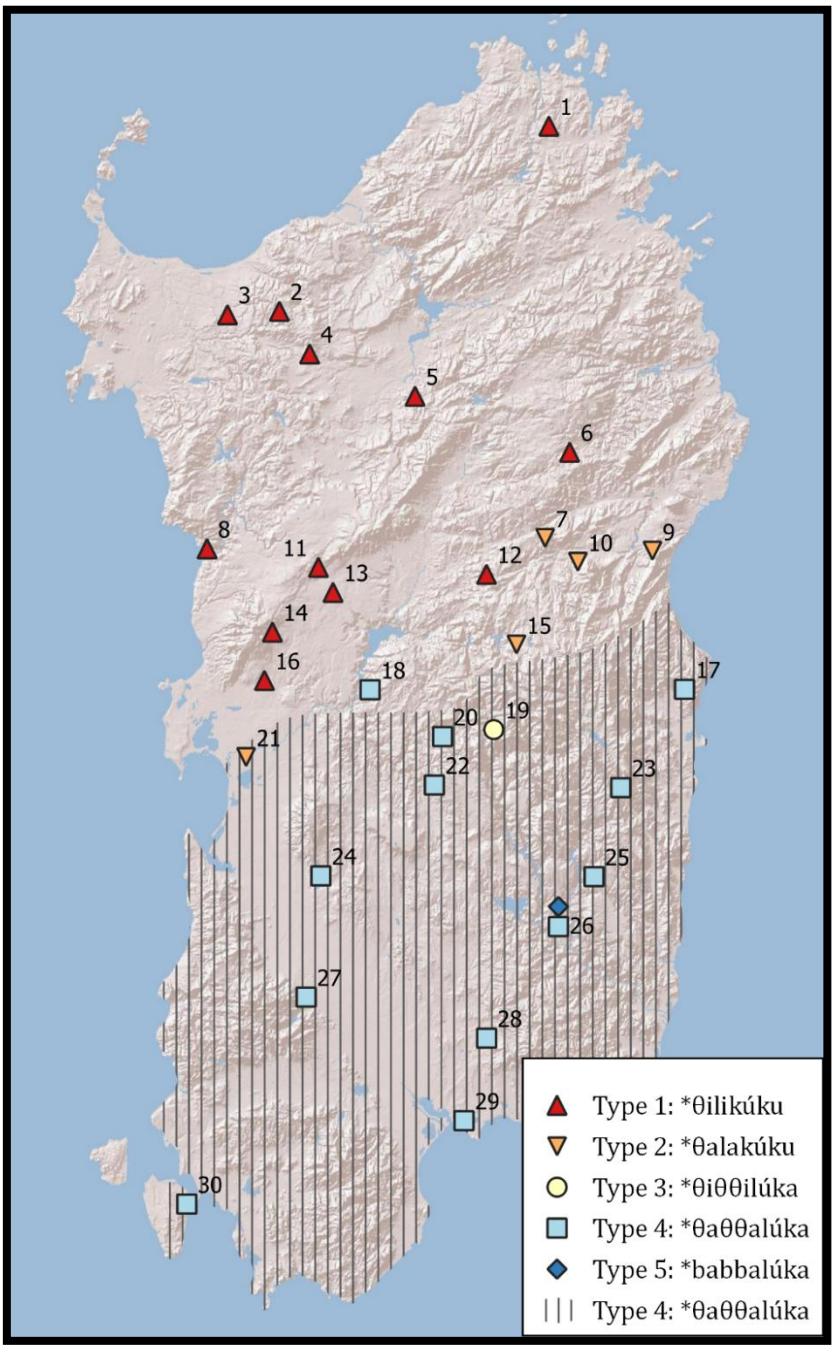

Figure 4.3.4: Skink (*Chalcides ocellatus tiligugu*); slug

Type 2 **ϑalakúka* and type 3 **ϑiϑϑilúka* occur in a strip between the more common types **ϑilikúku* and **ϑaϑϑalúka*. It is therefore attractive to explain **ϑalakúka* (type 2) as originally being identical to **ϑilikúku*, having undergone influence from forms like **ϑaϑϑalúka*. Vice versa, **ϑiϑϑilúka* (type 3) may originally have been **ϑaϑϑalúka* (type 4), influenced by **ϑilikúku* (type 1). This scenario is supported by the form *ϑaϑϑalakúkka* (9: Dorgali), which seems to be a complete blend of the two types. If this analysis is correct, we are left with two etymologically distinct words, **ϑilikúku* and **ϑaϑϑalúka*, both meaning 'skink' as well as 'slug'. This perhaps surprising homonymy is understandable in light of the skink's small, almost vestigial, legs, which give it the looks of a short, fat snake, or indeed a slug. In the meaning 'slug'.

The apparent coexistence of **alikúkku* and **ϑilikúku* in the meaning 'slug' is potential evidence for the identification of **ϑi-* in this word as a separable morpheme, like in *ϑilikèrta*. However, the lack of information on the provenance of the listed forms makes it difficult to use *alicúcu* for the analysis of the distribution of this feature. The part *-kúku* is reminiscent of several formations with the meaning 'slug' containing the element *kokk-*: *kokkóị*, *kokkóile*, *kokkoróị*, *kokkolóḍḍe* etc. (DES I: 356), but it is uncertain whether they are to be connected. There are several other phonetically similar words for slugs and snails, such as *kòrra* 'snail', *sittsi-ɣórru* 'slug, snail' << Lat. *cornus* (DES I: 387; II: 422; cf. § 4.1.15) and *kròka*, *krokkèḍḍa* '(various species of) snail' (DES I: 407). As a result, it is very much possible that the various words for 'snail' and 'slug' influenced each other.

4.3.5 Viperine snake (*Natrix maura*)$_1$

Reconstruction	**Forms**
1: **lírčis*	DES (II: 31): ***lírčis*** (1: Nuoro, 3: Gavoi) 'viperine snake (*Natrix maura* (Linnaeus, 1758))'
2: **lúrtsi*/lúrϑi***	DES (II: 31): ***lúrtsi*** (Nuor.) 'id.'
	Marcialis (2005: 64): ***sulurzi*** (2: Olzai) 'id.'

This word for snake, restricted to the Nuorese dialects, has no clear Romance etymology. We find phonetic alternations in the stressed vowel as well as in the realization of the affricate present in this word. Neither of these is easily explained. The affricates *-č-* and *-ts-* do not have a regular source in these dialects (except for Olzai <-*z*-> /ts/ < *-*ϑ*-), nor do they usually correspond to each other (cf. § 8.1.4). Type 1 has word-final *-s*, which is also found in some other animal names (§ 9.2.8). This might point to a pre-Roman origin of this word. *Su-* in *sulurtsi* (2: Olzai) is due to reinterpretation of the masculine article *su* as belonging to the lexical stem.

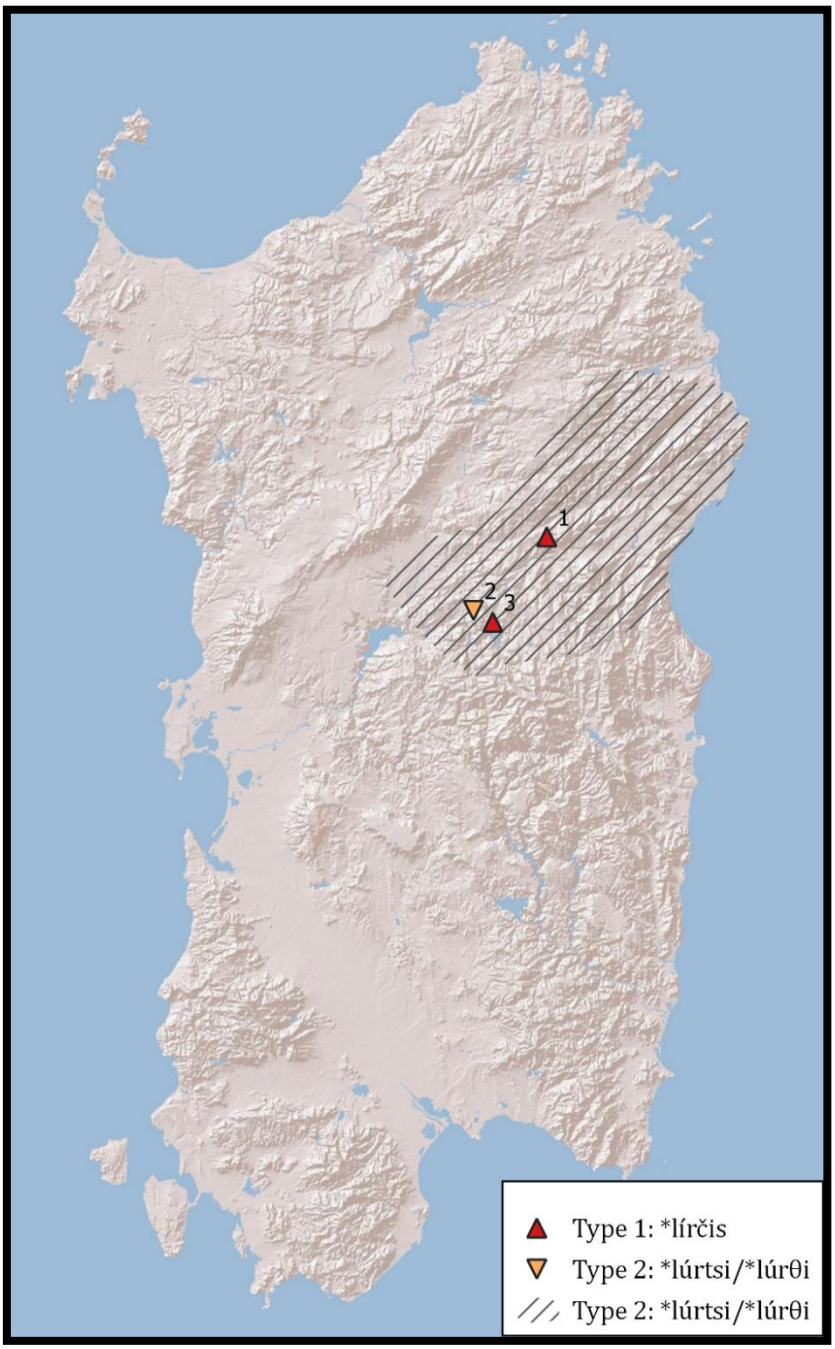

Figure 4.3.5: Viperine snake (*Natrix maura*)

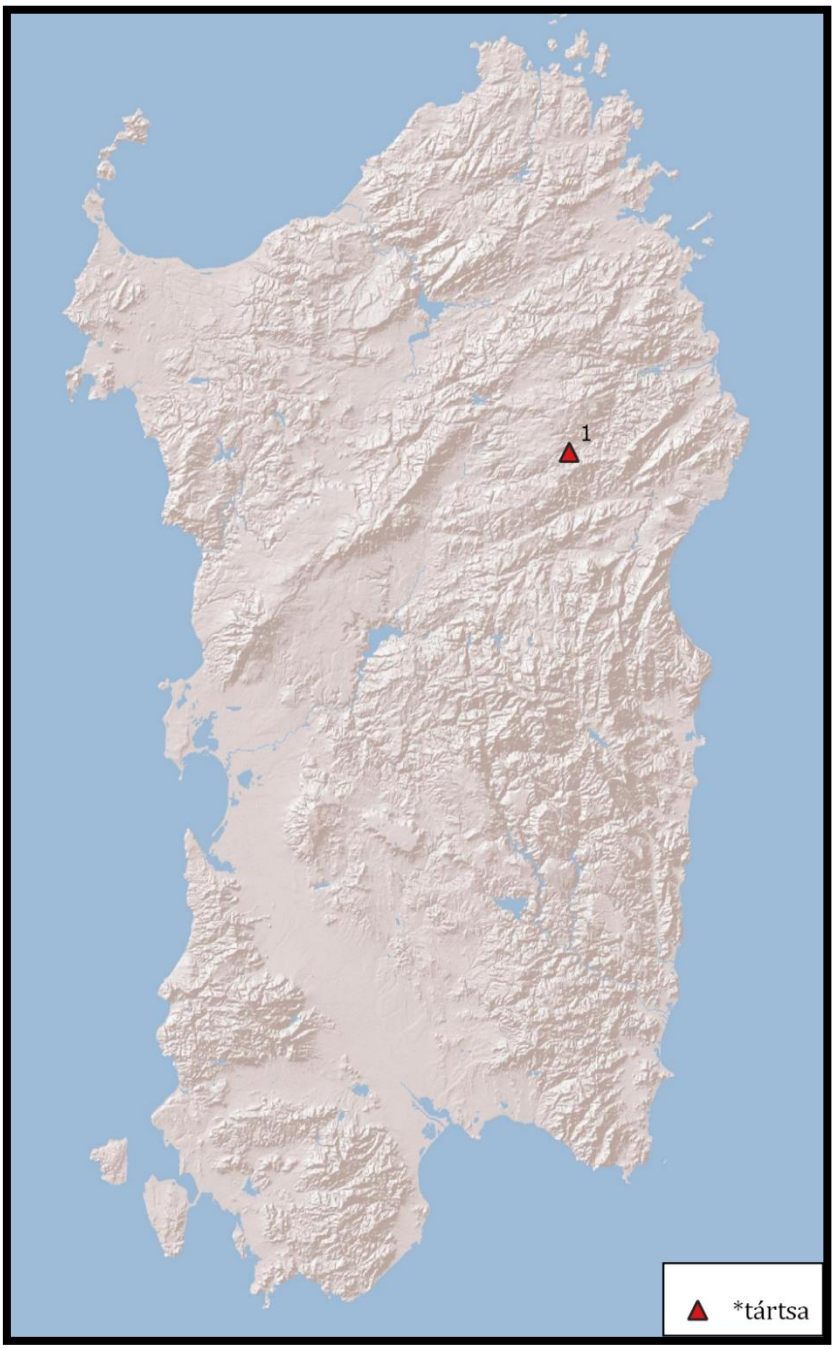

Figure 4.3.6: Viperine snake (*Natrix maura*)

4.3.6 Viperine snake (*Natrix maura*)$_2$

Reconstruction	Forms
1: *tártsa	DES (II: 466): ***tártsa*** (1: Bitti) 'viperine snake (*Natrix maura*)'

This word is only attested in Bitti, which makes its analysis difficult. The presence of *-ts-* is peculiar, as there is no regular phonetic source for it in Bittese (§ 8.1.4). In the same region we find *lírčis* and *lúrtsi* for the same animal and with the same odd cluster *-rts-* (§ 4.3.5). It is impossible to connect *tártsa* to *lúrtsi* formally though.

4.4 Birds

4.4.1 Bee-eater (*Merops apiaster*)

Reconstruction	Forms
1: *marrakátu	DES (II: 71): ***mar(r)ayáu*** (Camp.), ***marayáu*** (1: Triei, 2: Busachi), ***marrangáu*** (3: Laconi) 'bee-eater (*Merops apiaster* Linnaeus, 1758)
2: *marráku	DES (II: 71): ***marráɣu*** (4: Escalaplano, 5: Guasila, 6: Sant'Antioco, 7: Sarroch) 'id.'

The words listed above are found across the Campidanese dialects and refer to the 'bee-eater (*Merops apiaster*)'. The forms come in two types, those derived with the past participle suffix *-áu* < Lat. *-ātus* and those without *-áu*. The un-suffixed forms show us that we must reconstruct a form **marráku* 'bee-eater'. Besides *-n-*epenthesis, which is common, in the variant from Laconi (3), there is not much variation going on in this word. Although it lacks an etymology, there is no positive evidence in favor of a pre-Roman origin.

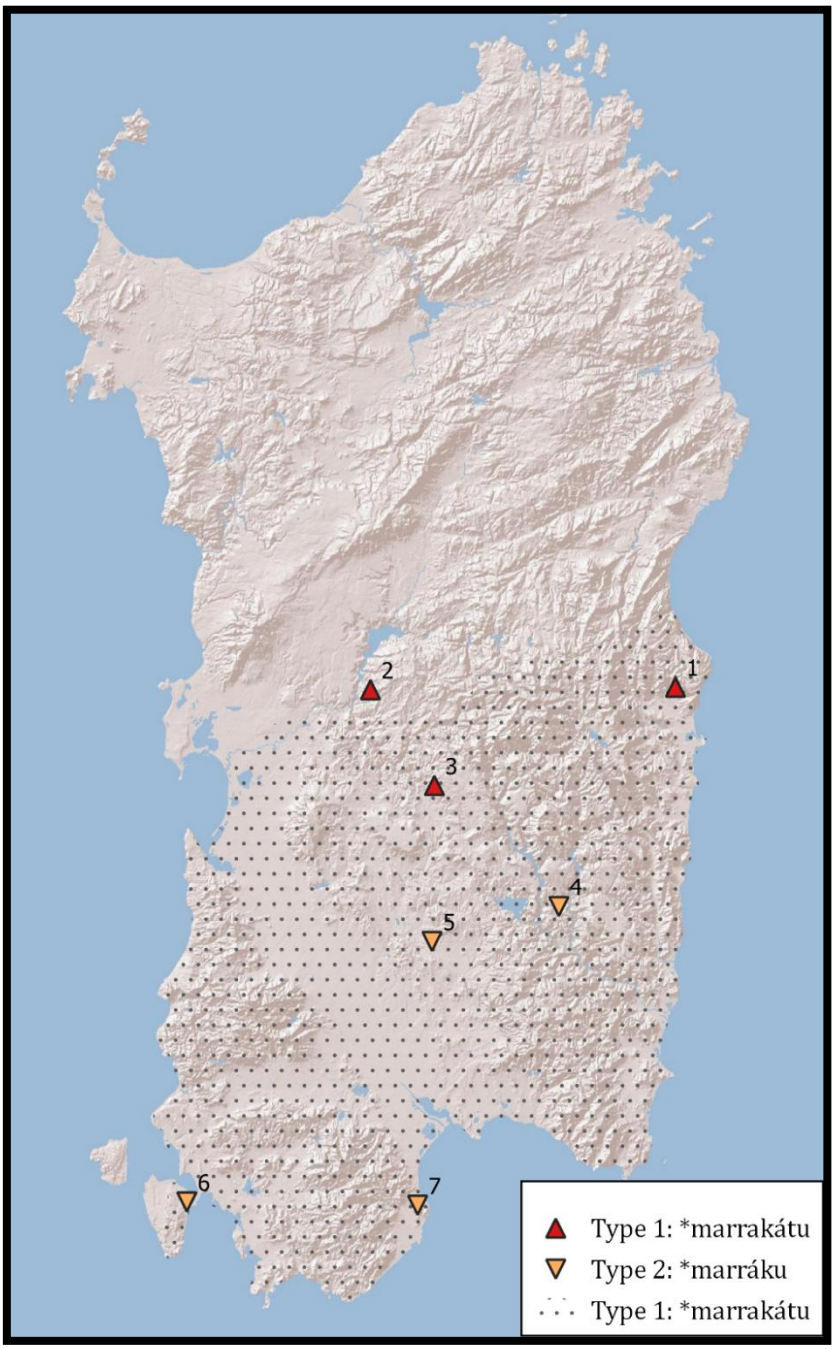

Figure 4.4.1: Bee-eater (*Merops apiaster*)

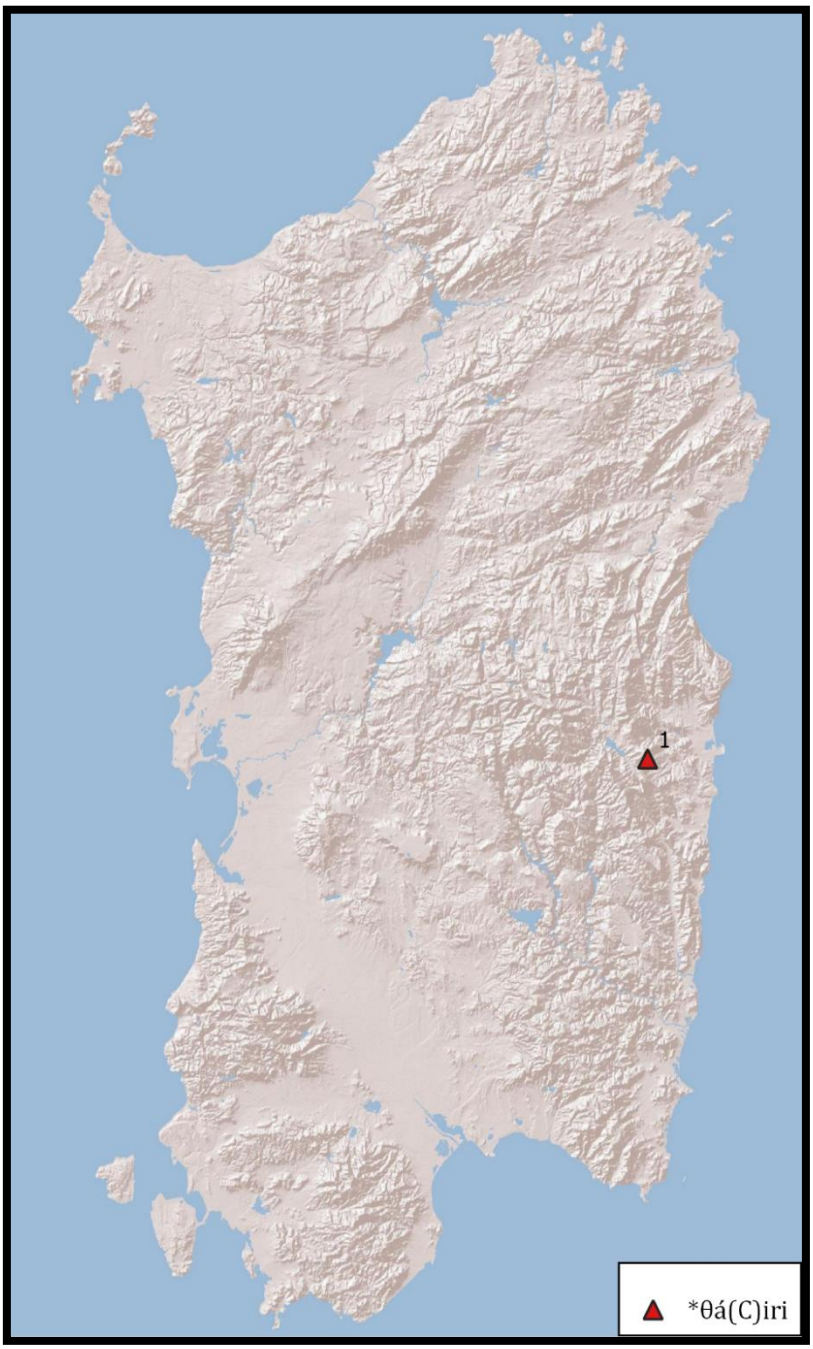

Figure 4.4.2: Bird of prey

4.4.2 Bird of prey

Reconstruction	Forms
***ϑá(C)iri**	Paulis (2008: 24): ***ϑáiri*** (1: Arzana) 'bird(s) of prey; harmful bird(s)'

Paulis (2008: 24) records *ϑáiri* from Arzana, used to "to generally indicate birds of prey, or birds that are harmful and dedicated thieves" (cf. Putzu 2019: 4). This word is apparently not recorded in any lexicon, but may be present in a number of micro-toponyms in Arzana and Gonnosfanadiga (Paulis 2008: 24). Paulis (2008: 24) connects *ϑáiri* to a Basque form *sari* 'vulture'. However, Bq. *sari* 'vulture' is only mentioned by Löpelmann (1968: 1138) as a variant of common Bq. *sai* 'id.' (cf. EH: 558). Trask (2008: 389) reconstructs this word as **sai*. The comparison between Bq. **sai* 'vulture' and Srd. **ϑá(C)iri* 'bird of prey' is formally unconvincing. This word cannot reliably be used in favor of a connection between Basque and the pre-Roman Sardinian substrate. Equally unconvincing is the comparison between Srd. **ϑá(C)iri* 'bird of prey' and Chenoua Bb. *isɣi* 'bird of prey' (Paulis 2008: 24; cf. Laoust 1912: 136).[151]

4.4.3 Bird species (various)

Reconstruction	Forms
***korrokkátu**	DES (I: 387): ***korrokkáu*** (1: Terralba) 'bee-eater (*Merops apiaster*)'; ***kurrukkáu*** (2: Mogoro) 'indeterminate bird'; ***korrokkáu*** (3: Villacidro) 'robin (*Erithacus rubecula* (Linnaeus, 1758))'
	Puddu (2023: s.v. *càu, corrocàu*): ***corrocàu, currucàu*** 'sea gull (*Larus sp.* Linnaeus, 1758)'

The exact meaning of this word is difficult to establish. According to Wagner it means 'bee-eater' in Terralba and 'robin' in Villacidro, whereas in Mogoro the species cannot be determined. The bee-eater and the robin are both small, colorful birds, making any confusion understandable. Puddu, however, defines *corrocàu* as 'sea gull'. Unless one of the two sources is mistaken, I do not know how to explain this semantic variation. The ending *-áu* could be the past participle suffix (< Lat. *-ātus*). The semantic uncertainties and limited attestation of this word make any further analysis difficult. Other names for the robin and the bee-eater are also of uncertain origin (cf. § 4.4.1, § 4.4.9, § 4.4.10, § 4.4.12).

[151] This comparison is based on Löpelmann's (1968: 1138) comparison between Bq. *sari* 'vulture' and Bb. *isɣi* 'bird of prey', which is formally no closer match.

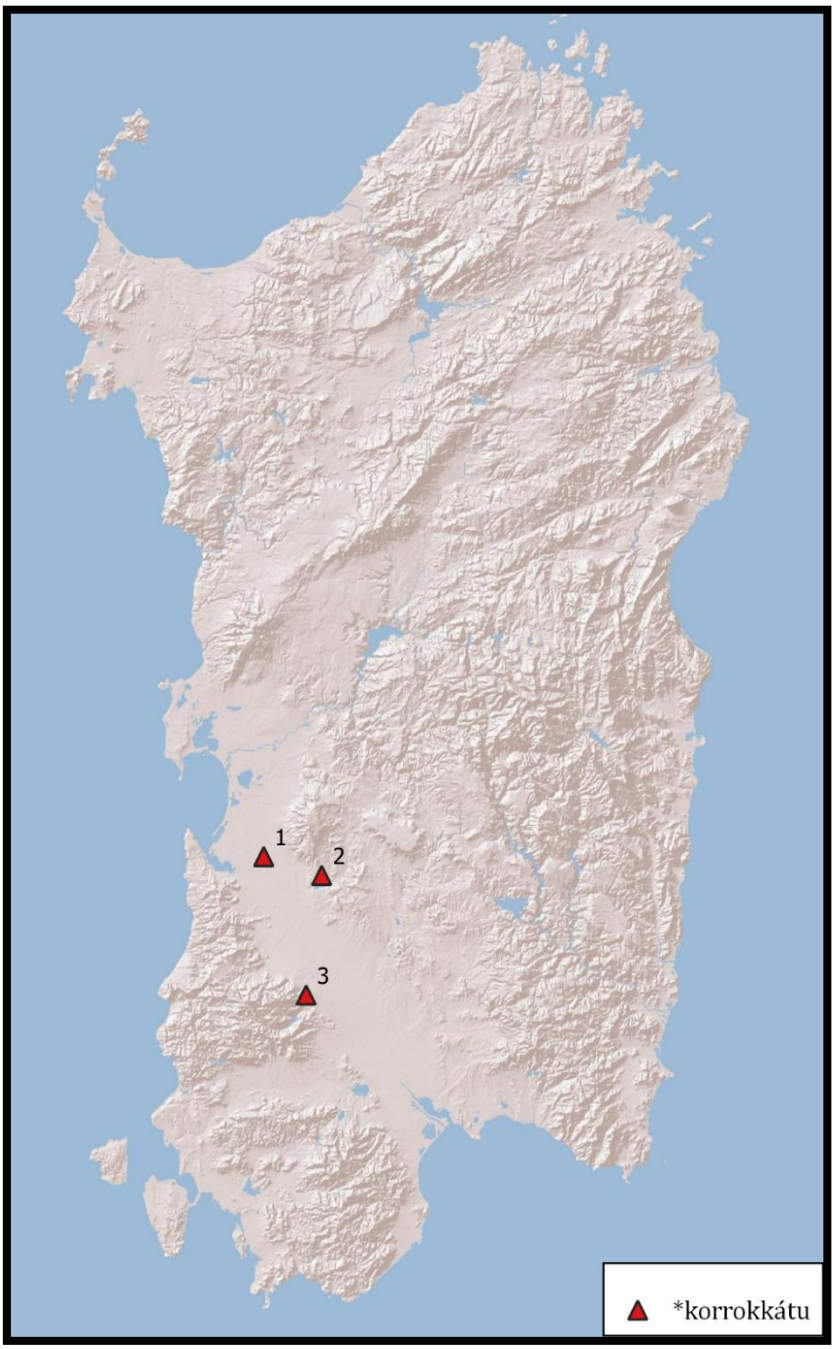

Figure 4.4.3: Bird species (various)

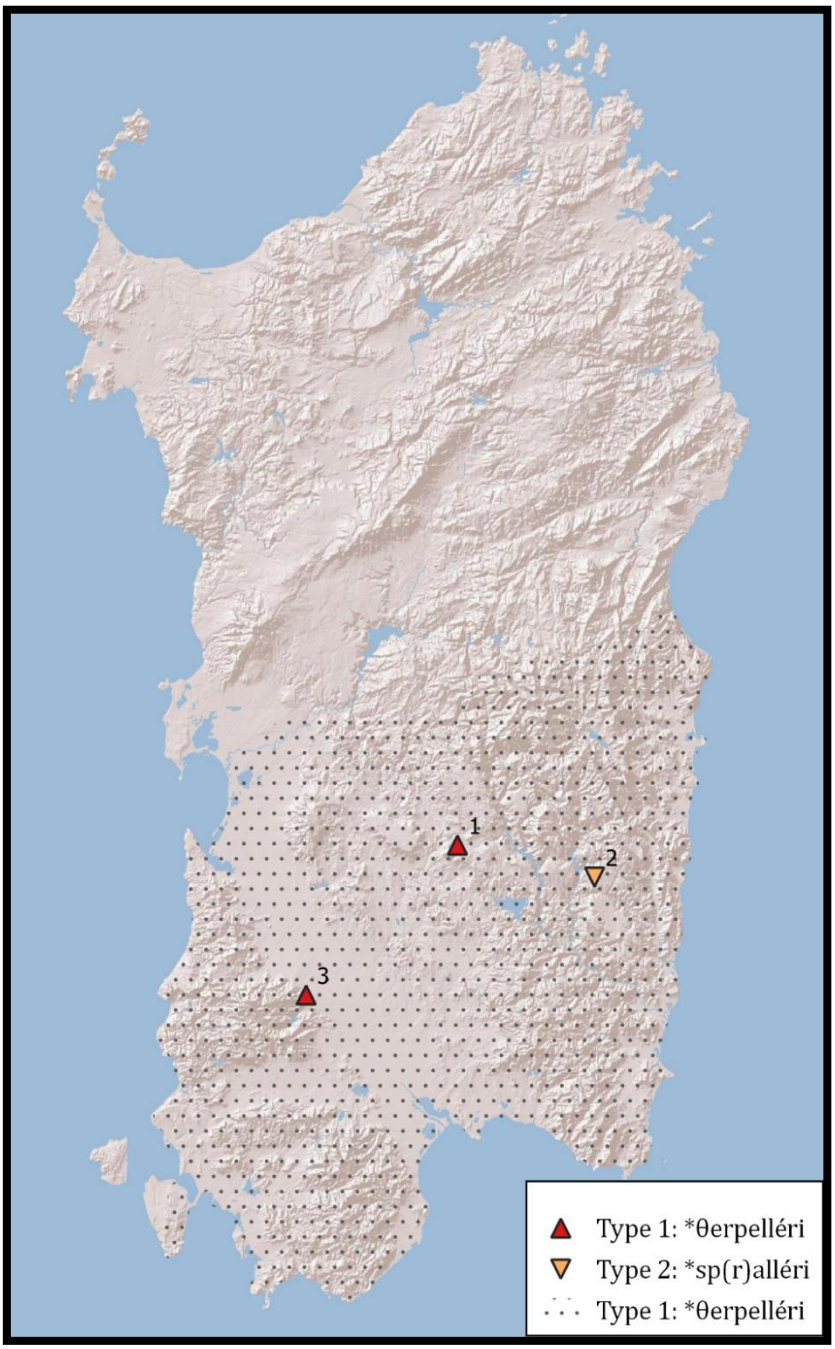

Figure 4.4.4: Kestrel (*Falco tinnunculus*)

4.4.4 Kestrel (*Falco tinnunculus*)$_1$

Reconstruction	Forms
1: *ϑerpelléri	DES (II: 588–589): ***tserpeḍḍéri, tsreppeḍḍéri*** (Camp.), ***tsrappeḍḍéri*** (1: Isili), ***tsarpeḍḍéri*** (3: Villacidro) 'kestrel (*Falco tinnunculus* Linnaeus, 1758)'
2: *sp(r)alléri	DES (II: 588–589): ***spraḍḍéri*** (2: Perdasdefogu) 'id.'
	Cetti (1776: 54): ***spadderi*** 'id.'

Type 2 of this word has clearly been influenced by Camp. *spraβéri* 'sparrowhawk (*Accipiter nisus* (Linnaeus, 1758)); kite (*Milvus milvus* (Linnaeus, 1758))' << Cat. *esparver* 'sparrowhawk' (DES II: 433). The forms of type 1 seem to go back to **ϑerpell-éri*. The suffix *-éri* is usually found in words of Italian or Catalan origin (Wagner 1952: 75–78) and may indicate that the forms of this type also underwent some influence from *spraβéri* 'sparrowhawk, kite'. Wagner (DES II: 589) suggests influence from Camp. *tsèrpi* 'snake, noxious animal' for the forms of type 1, but stresses that this is not enough to explain the whole word. If the word for 'snake' influenced these words for 'kestrel', which seems unlikely to me on semantic grounds, the original shape of this word is impossible to retrieve, since *spraḍḍéri* (type 2) cannot yield any reliable information either. Without any further evidence however, the connection to *tsèrpi* 'snake' is doubtful. This word's etymology remains unclarified.

4.4.5 Kestrel (*Falco tinnunculus*)$_2$[152]

Reconstruction	Forms
1: *ϑiliprí(g)u	DES (II: 483): ***čiliβríu*** (NLog.), ***attiliβríu*** (1: Osilo), ***tilibríu*** (3: Ploaghe, 4: Posada) 'kestrel (*Falco tinnunculus*)'
2: *ϑilipríkku	DES (II: 483): ***tilibríkku*** (Log.), ***tiribíkku*** (2: Sassari) 'id.'
	Marcialis (2005: 65): ***tilipricu*** 'id.'

The forms for 'kestrel' listed above all occur in the northern part of Sardinia. They go back to **ϑiliprí(g)u* (type 1) and **ϑilipríkku* (type 2). The alternation between *-(*g*)- and **-kk-* is unexpected. Like several other animal names, this word starts in **ϑi-*, although we do not find an alternation. Wagner (DES II: 483) cites Sass. *tiribíkku* 'kestrel', which is identical to the Sassarese attestation in AIS for 'grasshopper' (§ 4.1.12). Since neither of the sources have this form in the other meaning, it is possible that they were confused either in DES or in AIS. It is probable that this word originates from the same pre-Roman language as other animal names containing **ϑi-* (cf. § 9.1.2).

[152] These data are discussed in a similar fashion in Swanenvleugel (2024: 250–251).

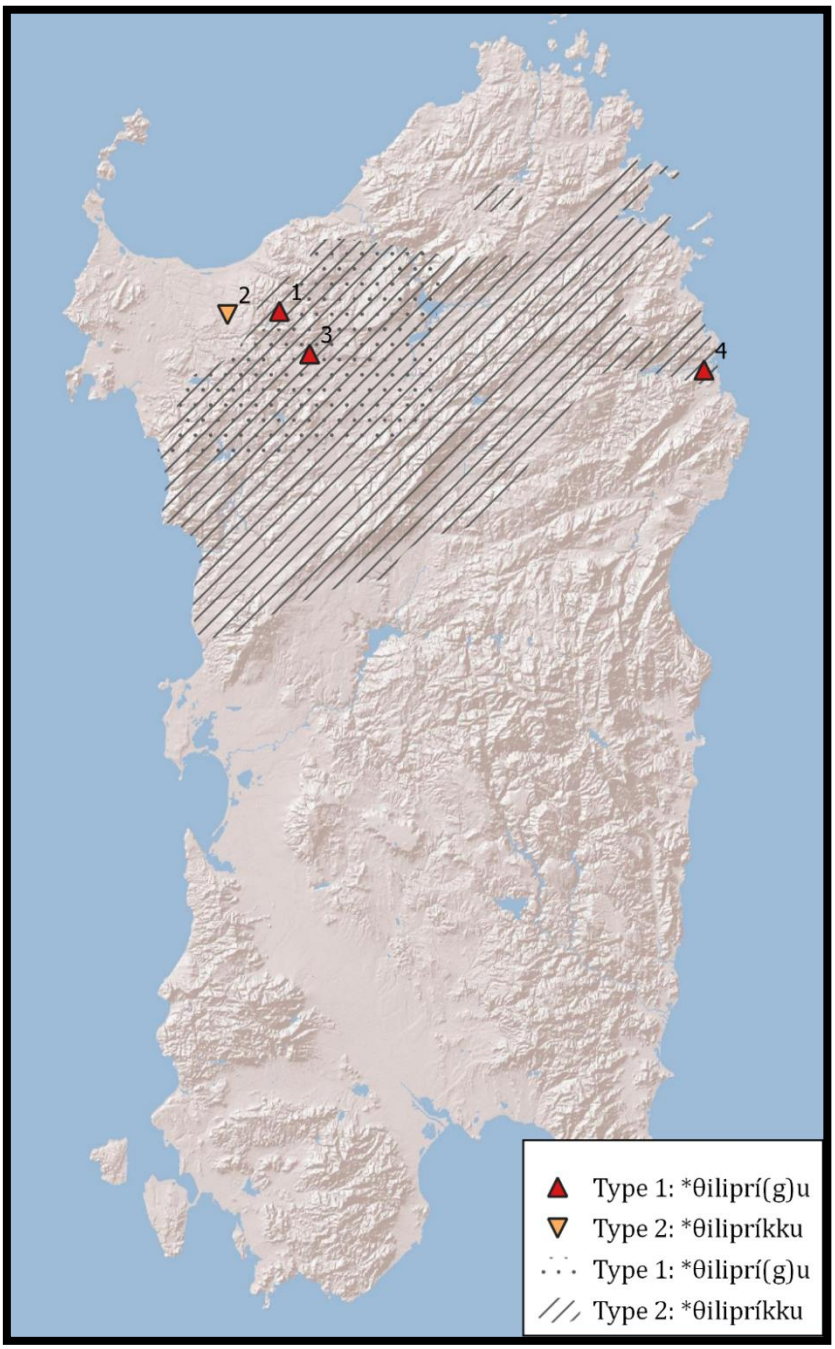

Figure 4.4.5: Kestrel (*Falco tinnunculus*)

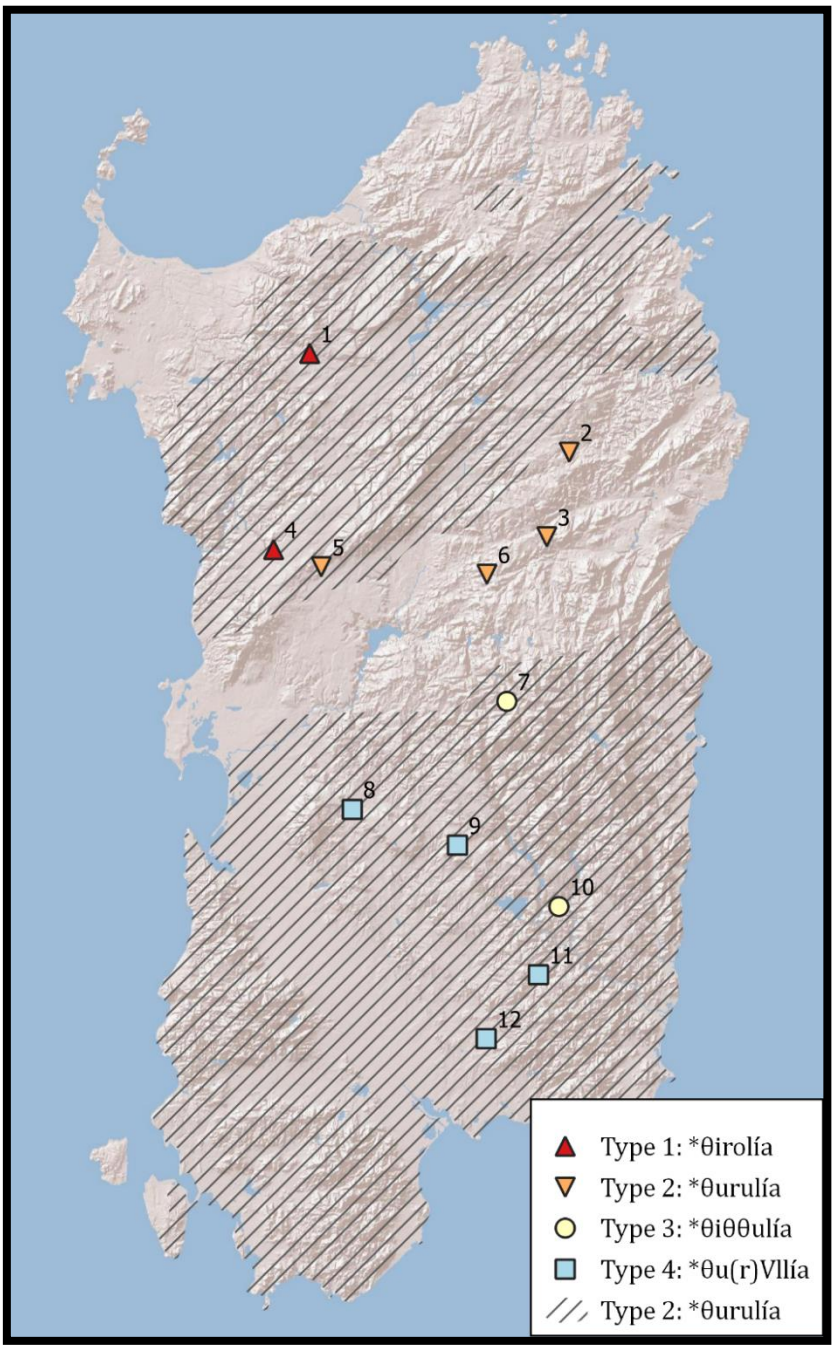

Figure 4.4.6: Kite (*Milvus milvus*)

4.4.6 Kite (*Milvus milvus*)

ReconstructionForms	
1: *ϑirolía	DES (II: 555): ***attiloría, attularia*** (NLog.) ***čirolía*** (1: Ploaghe), ***tirolía*** (4: Sindía) 'red kite (*Milvus milvus* (Linnaeus, 1758))'
2: *ϑurulía	DES (II: 555): ***turulía, asturólia*** (Log.), ***tsurulía*** (Camp.), ***ϑurulía*** (2: Bitti, 3: Nuoro, 6: Orani), ***atturulía*** (5: Macomer) 'id.'
3: *ϑiϑϑulía	DES (II: 555): ***čičču̯lía*** (7: Desulo), ***sittsulía*** (10: Escalaplano) 'id.'
4: *ϑu(r)Vllía	DES (II: 555): ***tsioḍḍía*** (8: Usellus) ***tsuiḍḍía*** (9: Isili), ***tsueḍḍía*** (11: San Nicolò Gerrei), ***tsuaḍḍía*** (12: Dolianova) 'id.'

According to DES (II: 555), these forms all refer to the 'red kite' primarily, but are also often used to refer to the 'common kestrel (*Falco tinnunculus*)'. Of the types that can be distinguished, type 2 **ϑurulía* is most widespread. Type 1, found in Logudorese, goes back to **ϑirolía*, with metathesis of *-l-* and *-r-* in the case of NLog. *attiloría*, *attularia*. Log. *asturólia* has been influenced by *astòre* 'peregrine falcon (*Falco peregrinus* Tunstall, 1771)' << Cat. *astor* 'id.' (DES I: 141; II: 555). Type 3 requires original **ϑiϑϑulía*. Type 4 can go back to **ϑurVllía*, with loss of intervocalic *-r-* that is common in Campidanese dialects (Wagner 1941a: 129). The varying vocalism (*-io-*, *-ui-*, *-ue-*, *-ua-*) must be the result of the dissimilation of vowels in hiatus, but it is difficult to discern what the original vocalism might have been.

Types 1, 2, and 4 mainly differ in their vocalism, and the fact that type 4 contains geminate **-ll-*. We can tentatively reconcile these three types in **ϑurol(l)ía*, since all forms contain either *-u-* followed by another vowel or *-o-* preceded by another vowel. This hypothesized original vocalism is found in Log. *asturólia*, whose vocalism cannot be explained from *astòre*. Still unexplained are geminate **-ll-* in type 4, and the aberrant consonantism in type 3, where the **-r-* found in the other forms appears to have been replaced with **-ϑ-*. It is difficult to explain these phenomena by means of commonly found phonological processes. Wagner (DES II: 555) may be right in attributing the variation in these forms to onomatopoeic deformations, but this is hardly satisfactory. For type 3, it is noteworthy that there are various other Campidanese animal names starting in *sitts-* etc. < **ϑiϑϑ-*, which may have acted as a source of influence (§ 9.1.2.1). The apparent replacement of singleton **-l-* with geminate **-ll-* in type 4 must have pre-dated common Srd. **-ll-* > *-ḍḍ-*. Perhaps this alternation is to be attributed to dialectal variation in the pre-Roman source language of this word. If it is, we would expect to find the same variation in other words of suspected pre-Roman origin (§ 8.2.1.2.2).

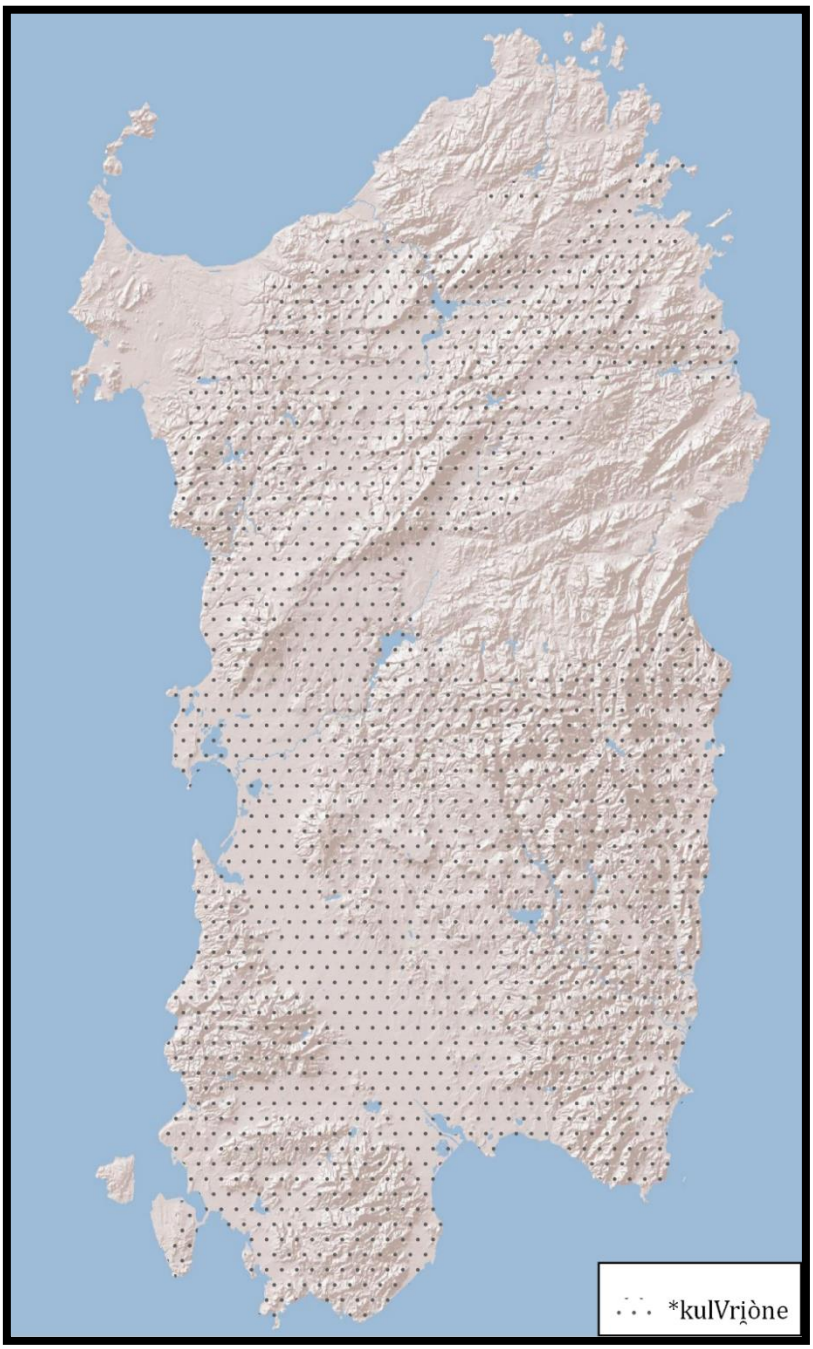

Figure 4.4.7: Plover (*Pluvialis apricaria*)

4.4.7 Plover (*Pluvialis apricaria*)

Reconstruction	**Forms**
1: *kulVr̥ìone	DES (I: 424–25): ***kulurdzòne*** (Log.), ***kulirğòni, kulinğòni*** (Camp.), ***kruβižôi*** [***e ðèrra***] (1: San Gavino Monreale) 'plover (*Pluvialis apricaria* (Linnaeus, 1758))'

All of these forms can go back to **kulur̥ìone* or **kulir̥ìone*.[153] I agree with Wagner (DES I: 424–25) that Subak's (1909: 480) etymology from Lat. *chlōriō, -ōnem* can not be upheld. The Latin form, attested in Pliny, is the same as Gr. χλωρίων 'golden oriole (*Oriolus oriolus* (Linnaeus, 1758))', which is a rather different kind of bird. There are moreover serious formal problems in explaining Srd. **kulVr̥ìone* from Lat. *chlōriō, -onem*. Also strikingly similar is Gr. κολλυρίων 'thrush (*Turdus sp.* Linneaus, 1758)' (Liddell et al. 1978: s.v. κολλυρίων). Thrushes are however hardly more similar to plovers than orioles, so this could be due to a coincidence. The origin of this word remains obscure. In the absence of an etymology and in light of its peculiar structure, the Sardinian word for 'plover' could perhaps go back to a pre-Roman source. Within Sardinian, these forms are largely homophonous with *culurgionis*, a type of ravioli filled with potatoes and sheep cheese, folded in the shape of an ear of grain (DES II: 424). These words have been derived from Gr. κολλύρον, -α, κολλύριον 'type of round bread' (Ettmayer 1906: 656; Rohlfs 1964: 252–253). The semantic relation between the different kinds of birds on the one hand, and ravioli on the other is not clear to me.

4.4.8 Raven (*Corvus corax*)

Reconstruction	**Forms**
****tifáne/i***	DES (II: 492): ***tiváni*** (1: Baunei, 3: Villagrande Strisaili, 4: Arzana, 5: Gairo) 'raven (*Corvus corax* Linnaeus, 1758)'
	Blasco Ferrer (1989: 15 fn. 2): ***tifane*** (1: Urzulei, 2: Baunei) 'id.'

These forms denoting ravens go back to **tifáne/i*, which is attested as a personal name (cf. DES II: 492). Wagner (1928: 21; DES II: 492) is hesitant to decide whether the animal's name is based on the personal name or vice versa. He does however compare Rif. *asiwan, tasiwant* 'bird of prey' (cf. Serhoual 2002: 550).[154] This is unattractive semantically and formally. Blasco Ferrer (1989: 15 fn. 2) considers the personal name to be original and says it is Etruscan, but without specifying any arguments for this statement. Paulis (2017: 111) plausibly proposes these forms to go back to the Greek personal name Θεοφάνης. This renders a pre-Roman explanation unnecessary.

[153] *Kruβižôi* (1: San Gavino Monreale) is due to a metathesis from **kuβiržôi* < **kulir̥ìone*.

[154] I was not able to trace the source of *tsiouânt* 'raven', cited by Wagner (DES II: 492).

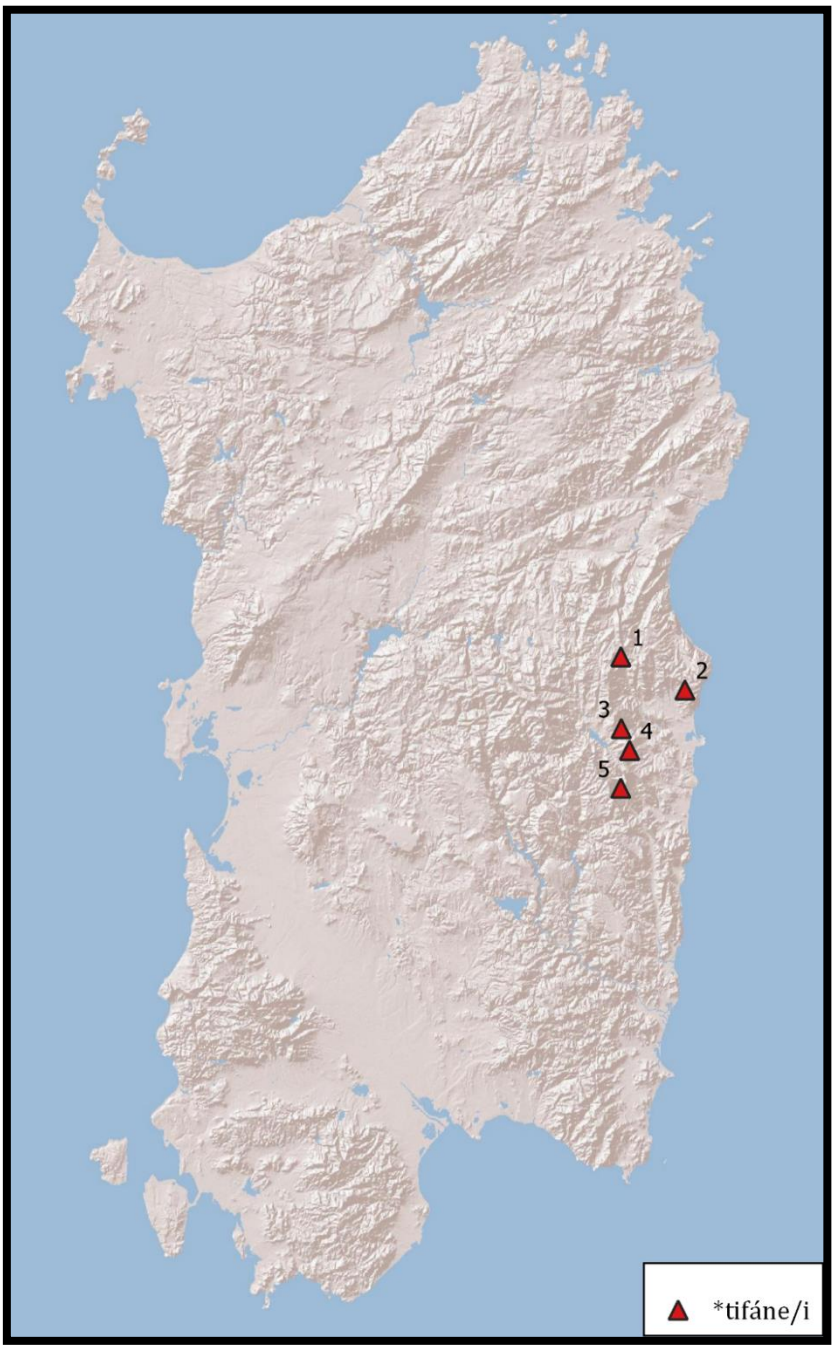

Figure 4.4.8: Raven (*Corvus corax*)

4.4.9 Robin (*Erithacus rubecula*),; yellow wagtail (*Motacilla flava*); hare (*Lepus sp.*)

Reconstruction	Forms
1: *brínčiri	DES (I: 226–227): **príńčiri** (1: Oschiri) 'yellow wagtail (*Motacilla flava* Linnaeus, 1758)'
2: *brinčis	DES (I: 226–227): **brinzis [cannèḍḍa]** (Log.) 'species of worm'; **brínčis** (2: Sindia) 'hare (*Lepus sp.* Linnaeus, 1758)'; **brínčiši** (3: Orani) 'robin' (*Erithacus rubecula*)'
3: *brínči/*brísi	DES (I: 226–227): **bríši** (2: Sindia) 'hare (*Lepus sp.*)'; **brínči** (6: Busachi) 'robin (*Erithacus rubecula* Linnaeus, 1758)'
4: *brinčóttu	DES (I: 226–227): **prinčóttu** (4: Cuglieri) 'robin (*Erithacus rubecula*)'
5: *brinkílli	DES (I: 226–227): **brinkíḍḍi** (5: Norbello) 'robin (*Erithacus rubecula*)'

All of the forms listed above seem to contain an element **brinči/*brinki*, to which different suffixes have been added. Most of the forms refer to the robin or to the yellow wagtail, which are both small colorful birds. The meanings 'hare' and 'worm' are difficult to explain, and one wonders whether these forms should be kept separate, even if they are formally almost identical to those meaning 'robin'.[155] Wagner (DES I: 226–27) wants to explain these forms either as pre-Roman or as onomatopoeic formation. In favor of the latter option could be their reference to wildly different species of animal, as well as the presence of the phoneme *-č-*, which has no clear regular source in these dialects.

On the morphological side, we find the ending *-íri* in type 1 *príńčiri* (cf. § 9.2.6.2). In the forms of type 2, we find the ending **-is*, which also recurs in some other words of potential pre-Roman origin (§ 9.2.8). Type 3 appears lacks any suffixes, and type 4 *-óttu* is of Catalan or Castilian origin (Wagner 1952: 110). The variant *brinkíḍḍi* (5: Norbello) is reminiscent of Lat. *fringilla* 'robin, chaffinch', but cannot be regularly derived from it. Note that, whereas *-illus*, *-a* was inherited in Sardinian as *-iḍḍu/a*, Srd. *-iḍḍi* does not have a good Latin source. Perhaps it represents a contamination of the base **brínči* and Lat. *fringilla*. All forms are reminiscent of *brinkare* 'to hop' (cf. also § 4.1.12). However, the presence of multiple unexplained endings (i.e. *-ís, -íri, -íḍḍi*) speaks against an original derivation from this verb. It is in any case likely that sound symbolism and folk etymology have played a role in the phonetic development of this word.

[155] But cf. § 4.5.8 for a colexification of 'hare' and 'butterfly'. Possibly the association of hares with small flying animals is based on their "hopping" motion.

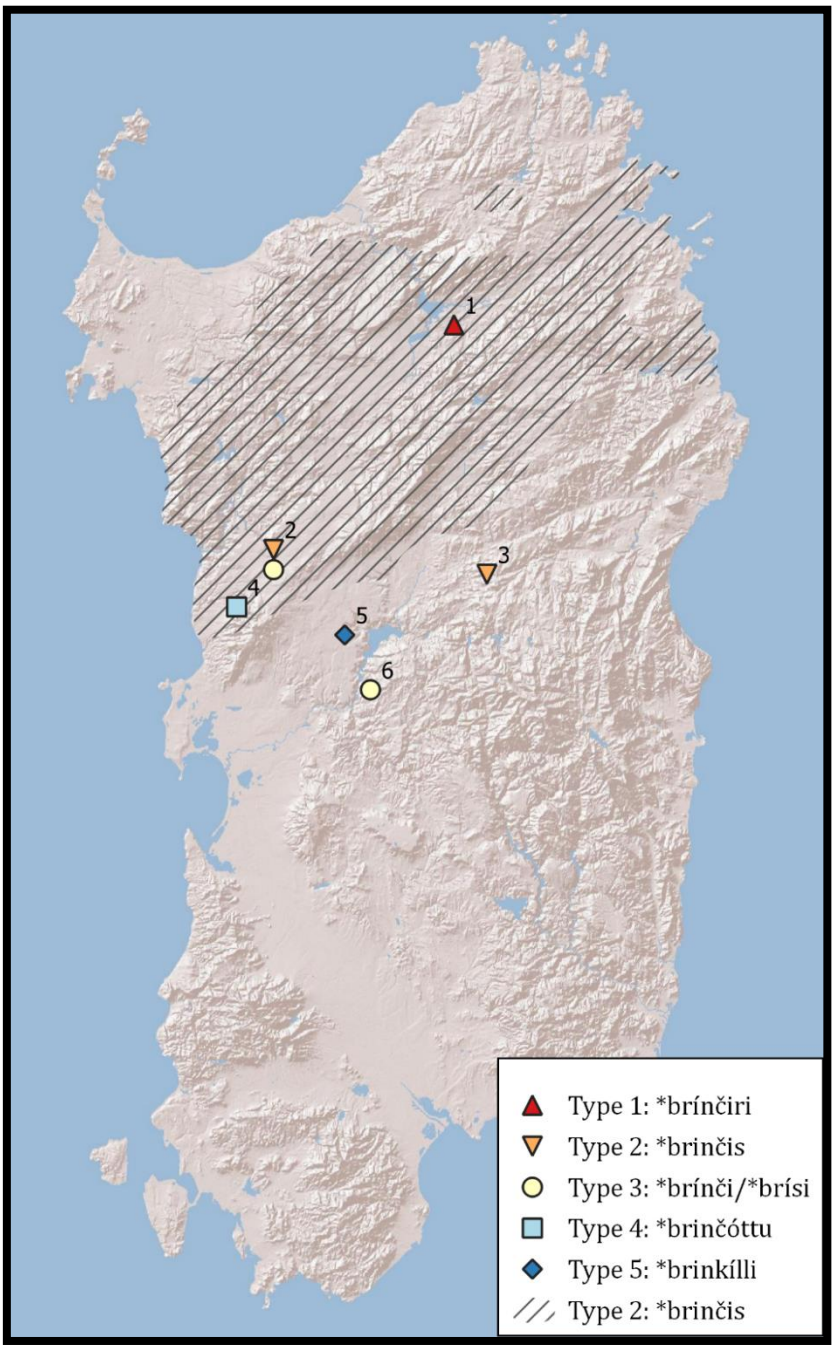

Figure 4.4.9: Robin (*Erithacus rubecula*); yellow wagtail (*Motacilla flava*); hare (*Lepus sp.*)

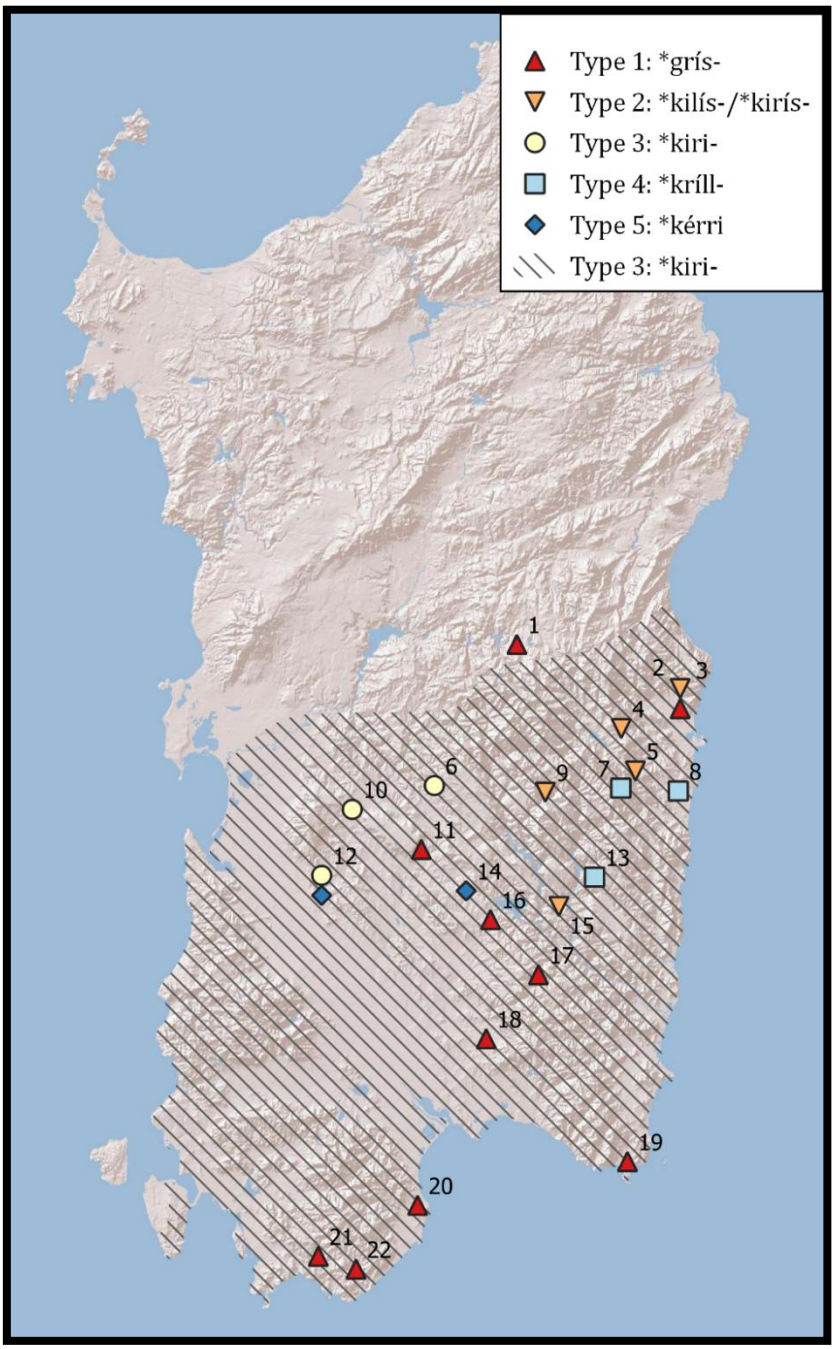

Figure 4.4.10: Robin (*Erithacus rubecula*), wren (*Troglodytes troglodytes*)

4.4.10 Robin (*Erithacus rubecula*)$_2$, wren (*Troglodytes troglodytes*)$_1$

Reconstruction	Forms
1: *grís-	DES (I: 343): ***grísu*** (1: Fonni) 'wren (*Troglodytes troglodytes* (Linnaeus, 1758))'; ***grísi*** (3: Baunei), ***grisòi*** (11: Gesturi), ***grísu*** (16: Siurgus, 17: San Nicolò Gerrei, 18: Dolianova, 19: Villasimius, 21: Teulada, 22: Domus de Maria), ***grisótti*** (20: Sarroch) 'robin (*Erithacus rubecula*)'
	Marcialis (2005: 40): ***grisu*** 'robin'
2: *kilís-/*kirís-	DES (I: 343): ***kiríši*** (2: Triei), ***kilíšu*** (9: Seui), ***kilíši*** (15: Escalaplano) 'robin (*Erithacus rubecula*)'; ***kiríši*** [***e βèrka***] (4: Villagrande Strisaili) 'wren (*Troglodytes troglodytes*)'; ***mameɣeríši*** (5: Lanusei) 'sparrow (*Passer domesticus* (Linnaeus, 1758))'
	Marcialis (2005: 28): ***chirisi*** (Log.) 'robin (*Erithacus rubecula*)'[156]
3: *kiri-	DES (I: 343): ***kiríu*** (6: Laconi), [***su***] ***ggíri*** (10: Usellus), [***su***] ***ɣíri*** (12: Mogoro) 'robin (*Erithacus rubecula*)'; ***gíru*** (Camp.) 'wren (*Troglodytes troglodytes*)'
	Marcialis (2005: 38): ***ghiru*** 'robin (*Erithacus rubecula*)'
4: *kríll-	DES (I: 343): ***kríḍḍi*** (7: Gairo, 8: Bari Sardo), ***kriḍḍólu*** (13: Perdasdefogu) 'robin (*Erithacus rubecula*)'
5: *kérri	DES (I: 343): ***kérri*** (12: Mogoro, 14: Mandas) 'wren (*Troglodytes troglodytes*)'

These words for 'robin' and 'wren' can be divided into five types. The forms of type 1 all contain **gris-* and are mainly found in the southern Campidanese dialects, with more northerly outliers in Fonni (1) and Baunei (3).[157] Type 2 is found in the Ogliastra and is formed from a base **kirís-* or **kilís-*.[158] Type 3 seems to go back to **kir-*. Type 4, containing **krill-* is again restricted to the Ogliastra. Finally, type 5 is only found in two places and goes back to **kerri*. All five types appear to contain **kr-* or **kir-*. Wagner (DES I: 343) assumes that all of these formations are

[156] Marcialis' (2005: 28) attribution to Logudorese of this form is dubious, as it is not mentioned anywhere else, and all forms of type 2 for which we have specific geographical information belong to the Ogliastran dialects. Could this be an error in Marcialis?

[157] The suffixes *-òi* < **-one* and *-ótti* are of recent origin.

[158] The form *mameɣeríši* (5: Lanusei) is analyzable as **mama 'e kiríši* 'mother of the robin/wren'. Similar *Benennungsmotiven* are common in Sardinian: cf. *mamma de sizziɣórru* 'slug (litt.: 'mother of the snail'), *mámma de ambíḍḍa* 'mother of the eel' etc. (cf. DES II: 60).

onomatopoeic, which is imaginable for songbirds. However, I do not think this can fully explain the similarity between these formations as opposed to other words for 'robin' and 'wren' that have been proposed to be of sound-symbolic origin, such as *brínči* 'robin' (§ 4.4.9) and *ḍḍóḍḍi* 'robin, wren' (§ 4.4.12).

Although sound symbolism may certainly have played a role in the formal diversity found in this word, it is possible that the various forms share an original base, which may be of pre-Roman origin. This is supported by the fact that all five types show evidence of originally ending in **-i*, for which Latin nominal inflection does not provide a good source (§ 8.1.5.1). In some cases this final **-i* was replaced with *-u* (e.g. *gríši* vs. *gríšu*), or extended with *-u* (e.g. *su ɣíri* vs. *kiríu*). Although it is difficult to establish how exactly the different types relate to each other, we can make some observations. One is that type 2 **kiríši* is similar to type 3 **kíri*, with an element *-(i)s* that is also found in other words of presumable pre-Roman origin (§ 9.2.8). The sequence *-ḍḍi* found in type 4 also recurs in other words for 'robin': *brinkíḍḍi* (§ 4.4.9), *tsiḍḍí*, *ḍḍóḍḍi* (§ 4.4.12).

4.4.11 Scops owl (*Otus scops*)

Reconstruction	**Forms**
1: *ϑònka	DES (II: 551): ***ϑònka*** (Nuor.), ***tònka*** (Log.), ***tsònka*** (Camp.) 'scops owl (*Otus scops* (Linnaeus, 1758))'
	AIS (508): ***čònka*** (1: Tempio Pausania, 12: Desulo), ***tsònka*** (2: Sassari, 7: Santu Lussurgiu, 9: Milis, 10: Busachi, 13: Laconi, 15: Escalaplano), ***tònka*** (3: Ploaghe, 5: Nuoro), ***ϑònka*** (4: Bitti, 5: Nuoro, 6: Dorgali, 8: Fonni) 'id.'
	Böhne (1950: 115): ***ϑònka*** (17: Muravera) 'id.'
2: *ϑòl̯e	DES (II: 551): ***ϑòǧǧe*** (11: Baunei) 'id.'
3: *anϑònka	DES (II: 551): [***s'***] ***antsònka*** (14: Mogoro, 18: Sarroch, 19: Domus de Maria) 'id.'
	AIS (508): [***s'***] ***antsònka*** (16: Villacidro) 'id.'
4: *ϑòr̯ia	Böhne (1950: 115): ***ϑòrža*** (17: Muravera) 'id.'

Wagner (DES II: 551) thinks that this word may have an onomatopoeic origin. He nevertheless believes that it "belongs to the oldest layer of the Sardinian lexicon", but it is not quite clear if this implies a pre-Roman origin. Although the sound of the scops owl is very characteristic indeed, it is difficult to judge whether **ϑònka* (type 1) is in fact an onomatopoeic rendering of this sound.

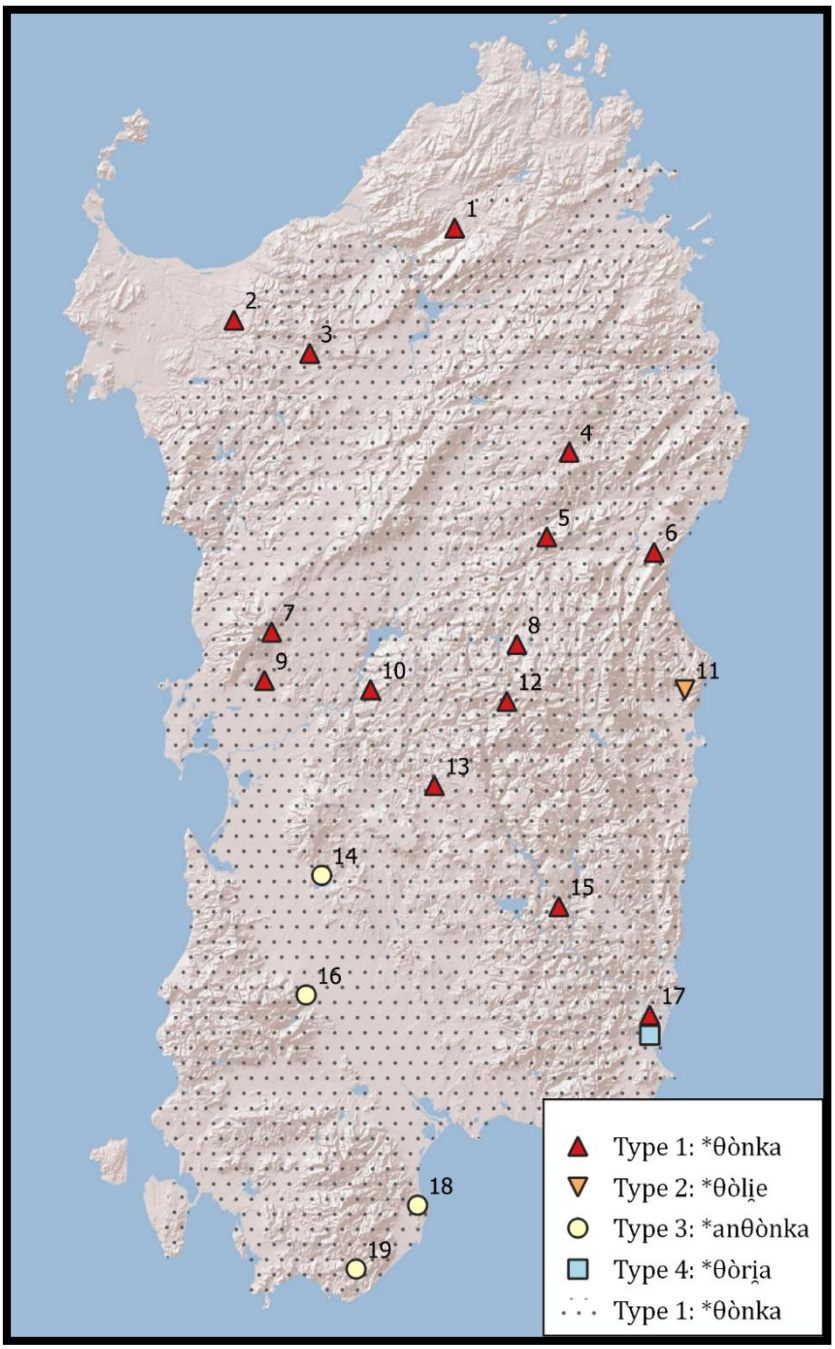

Figure 4.4.11: Robin (*Erithacus rubecula*), wren (*Troglodytes troglodytes*)

Type 3 *antsònka* is found in southwestern Campidanese dialects, and is probably a relatively recent deformation of *tsònka* (type 1). Type 2 *ϑòǧǧe* instead must go back to **ϑòl̥e*, which cannot be reconciled with **ϑònka*, but which is structurally similar to type 4 *ϑòrža* < **ϑòr̥a*. Whether *ϑòǧǧe* and *ϑòrža* are independent onomatopoeic formations or etymologically related to each other and to **ϑònka* is hard to say, especially without direct information about the names for this animal in other Ogliastran dialects. Judging from these words' phonotactics (including the phoneme **-ϑ-*), irregular variation, and semantics relating to the island's native fauna, they may well be of pre-Roman origin.

4.4.12 Wren (*Troglodytes troglodytes*)$_2$

Reconstruction	**Forms**
1: *llólli	DES (I: 593): [***su***] ***ḍḍóḍḍi*** (1: Bitti, 3: Dorgali) 'wren (*Troglodytes troglodytes*); robin (*Erithacus rubecula*)'; [***su***] ***dóḍḍi*** (4: Orgosolo) 'wren (*Troglodytes troglodytes*)'
2: *gròllo	DES (I: 593): ***gròḍḍo*** (2: Nuoro) 'wren (*Troglodytes troglodytes*)'

As Wagner (DES I: 593) notes, these words are very similar to those for 'fox' in the same region: *gròḍḍe*, *gròḍḍo*, *lòḍḍe* etc.; § 4.5.6), but they are not quite identical. In Nuoro for instance we find *gròḍḍo* 'wren' next to *gròḍḍe* 'fox'. The forms of type 1 go back regularly to **llólli*; the initial *ḍḍ-* < **ll-* is likely expressive.[159] It is probably easiest to explain the variation between initial *gr-* and *ḍḍ-* in these words as going back to original **l-*, with expressive gemination in type 1 and with introduction of initial *g-* in type 2. For a discussion of this latter development, cf. § 4.5.6. Perhaps these words too are somehow to be connected with the attested Sardinian personal name *Lollo*, *Lollu*, which Wagner (1933: 503–508; DES I: 593) suggests for *gròḍḍe*, *gròḍḍo*, *lòḍḍe* etc. 'fox'.

[159] We find *-ḍḍ-* in various other words for 'robin' and other songbirds as well (cf. § 4.4.9, § 4.4.10).

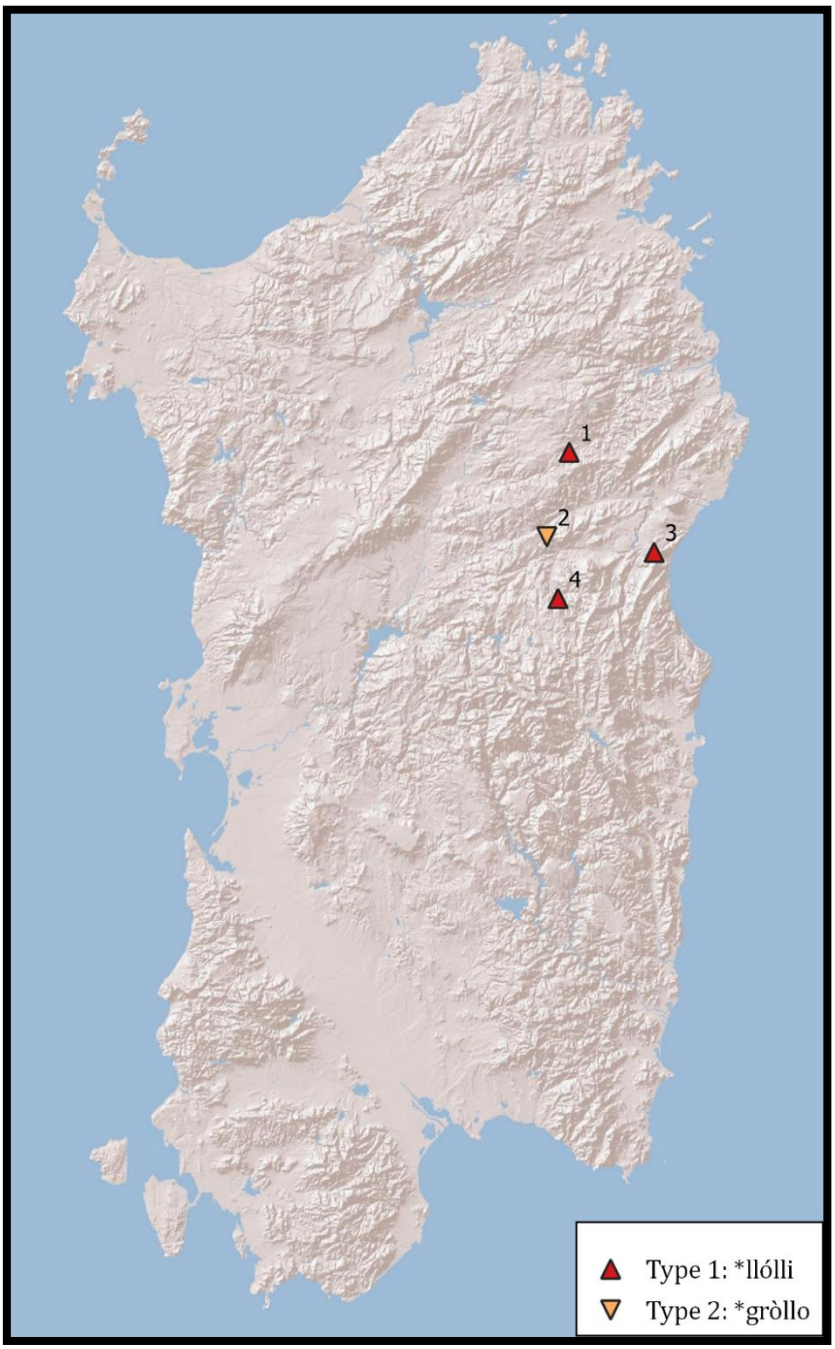

Figure 4.4.12: Wren (*Troglodytes troglodytes*)

4.5 Mammals

4.5.1 Bat$_1$

Reconstruction	Forms
***kikin-**	DES (II: 556): ***činčiréḍḍu*** (1: Urzulei), ***činčiðéḍḍu*** (2: Triei), ***ǧiǧǧinéḍḍu*** (4: Tortolì), ***čížini*** (5: Arzana), ***čižinéḍḍu*** (6: Gairo) 'bat'
	AIS (448): [***ṣu***] ***ǧinǧiréḍḍu*** (3: Baunei) 'id.'
	Terracini (1959: Map 1): ***činčiðéḍḍu*** (1: Urzulei, 2: Triei, 3: Baunei), ***ǧiǧǧinéḍḍu*** (4: Tortolì, 7: Bari Sardo), ***čiǧǧinéḍḍu*** (6: Gairo) 'id.'

The sources cited differ somewhat as to the exact shape of the words of this type. For instance, for Urzulei (1), DES has *činčiréḍḍu*, while Terracini has *činčiðéḍḍu*. The main point of disagreement between the authors and between the forms themselves is the final consonant of the stem, where we find *-n-*, *-ð-*, *-r-*. The only un-suffixed form, *čížini* (5: Arzana), likely preserves the original consonantism in *-n-*. As a result, we can posit an original **kikin-*; the final vowel *-i* can go back to either **-e* or **-i*. In some cases an epenthetic *-n-* appears in the first syllable (e.g. *činčinéḍḍu* (2: Triei); cf. for this process Wagner 1941a: 219–220), which may have been the cause of the dissimilation of final *-n-* to *-ð-* and *-r-*. Wagner (DES II: 556) treats these forms under the same entry as those of the type *ϑuϑuréḍḍu* etc. (cf. § 4.5.2) but calls them a "special type peculiar to the Ogliastra region". With **kikin-* as the most plausible reconstruction for the Ogliastran forms discussed here, it is impossible to connect them etymologically to forms like *ϑuϑuréḍḍu* etc., which probably go back to *(*ϑi-*)*ϑurr-*. Because of the semantic domain, the absence of an etymology, and the formal variation, it is likely that this word is of pre-Roman origin.

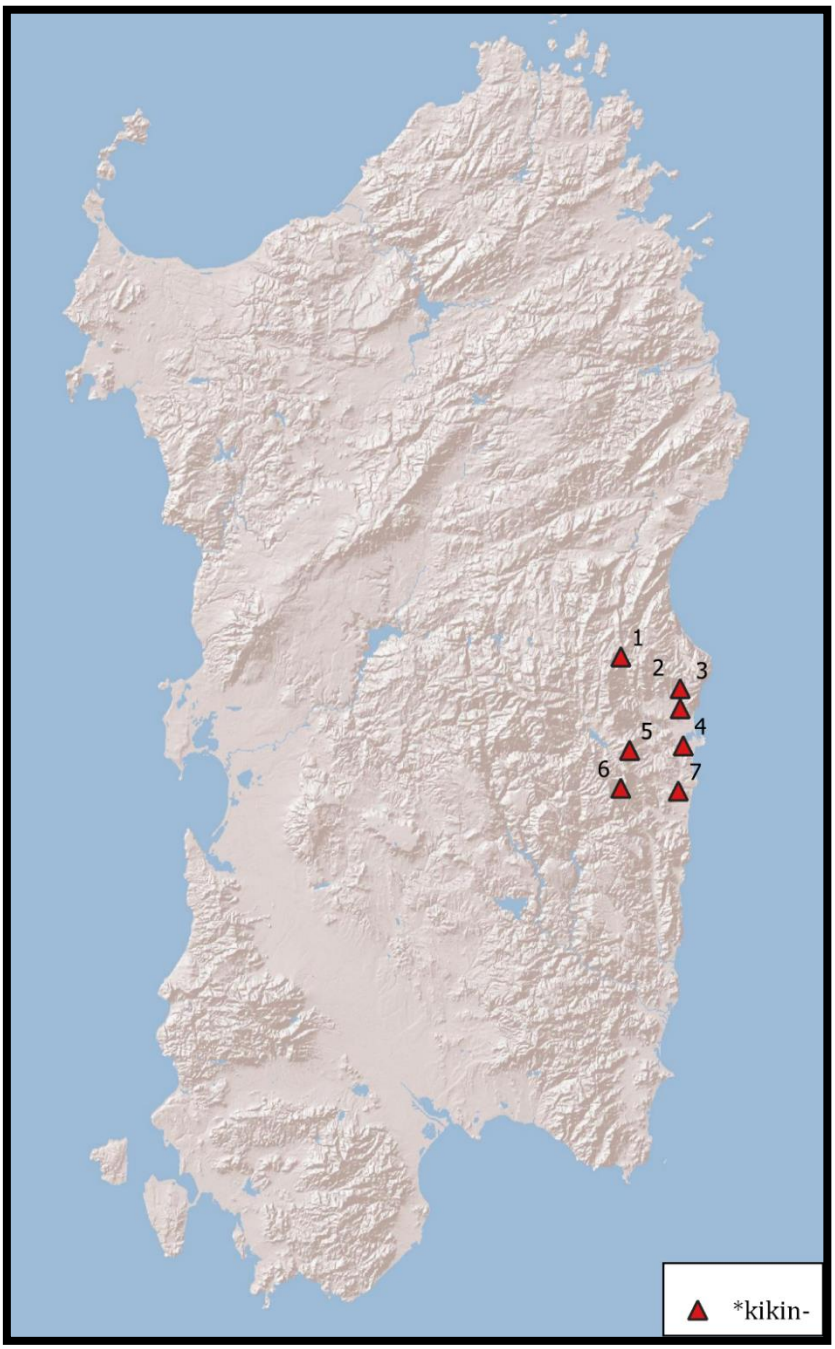

Figure 4.5.1: Bat

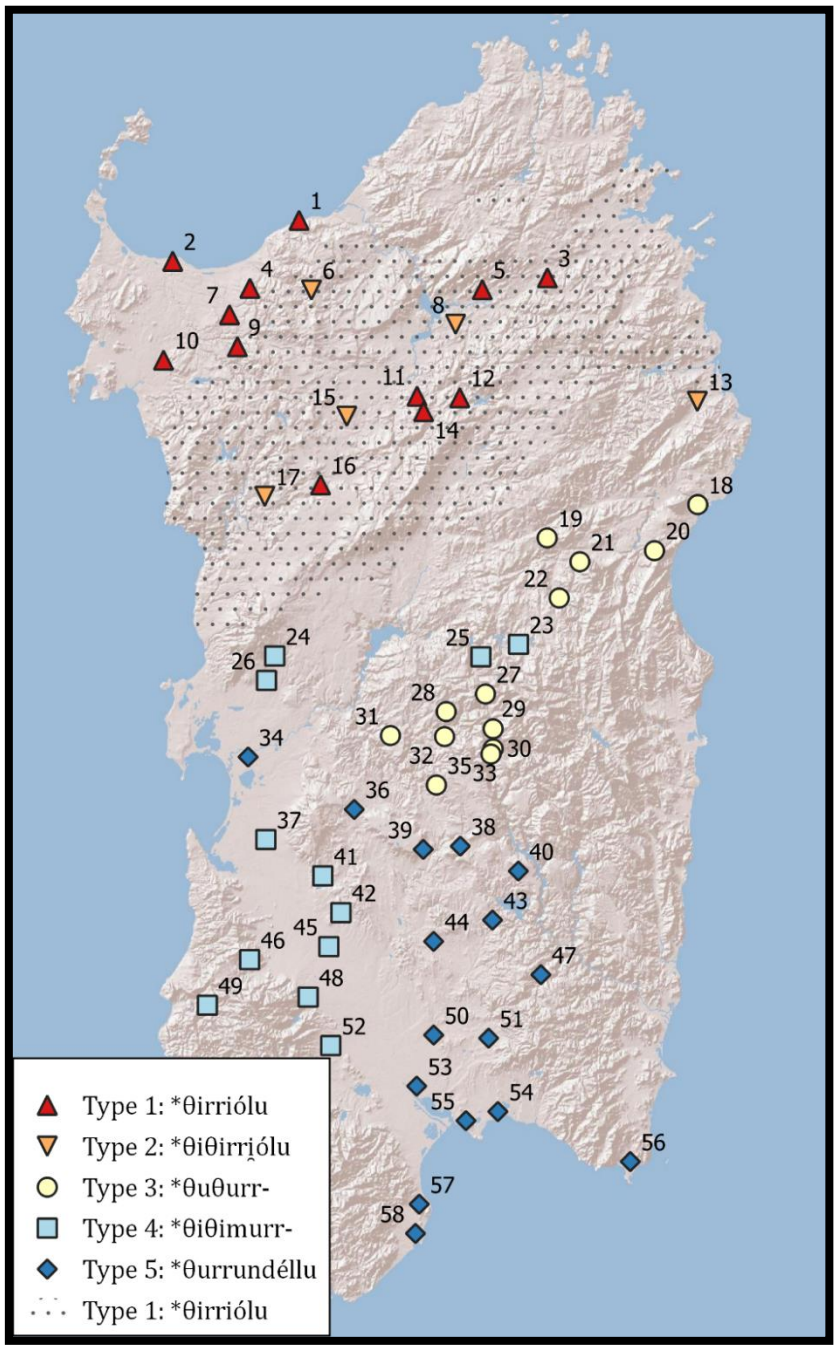

Figure 4.5.2: Bat

4.5.2 Bat$_2$[160]

Reconstruction	Forms
1: **ϑirriólu*	DES (I: 556): ***tirr̰iólu*** (Log.), ***tsirr̰iólu*** (4: Sennori), ***tsirr̰iolu*** [***βèḍḍe***] (10: Olmedo), ***tirr̰iólu*** [***malaβèḍḍe***] (11: Ozieri), ***tirr̰iólu*** [***'e βèḍḍe***] (12: Pattada) 'bat'
	AIS (448): [***la***] ***tsirr̰iòra*** (7: Sassari) 'id.'
	Terracini (1959: Map 1): ***tsirr̰iólu*** (1: Castel Sardo), ***tsirr̰ióra*** (2: Porto Torres, 7: Sassari), [***su***] ***tirr̰iólu*** (3: Monti, 9: Tissi), ***tirr̰iólu***[***βèḍḍe***] (14: Nughedu San Nicolò), ***tsirr̰iólu*** [***βèḍḍe***] (16: Bonorva) 'id.'
2: **ϑiϑirr̰iólu*	DES (I: 556): ***ϑiϑirr̰iólu*** (13: Siniscola) 'id.'
	Terracini (1959: Map 1): ***titirr̰iól'*** [***impeḍḍáðu***] (5: Berchidda), ***tsitsirr̰iólu*** (6: Nulvi), ***tintirr̰iólu*** [***βèḍḍe***] (8: Oschiri), ***činčirr̰iólu*** (15: Mores), ***tintirr̰iólu*** (17: Padria), ***ϑiϑirr̰iólu*** (13: Siniscola) 'id.'
3: **ϑuϑurr-*	DES (I: 556): ***ϑuϑuréḍḍu*** (19: Nuoro), ***ϑuϑuréri*** (20: Dorgali, 21: Oliena, 22: Orgosolo), ***tsuntsuréḍḍu*** (27: Tonara, 31: Samugheo, 32: Meana Sardo), ***tsuntsuríttu*** (28: Atzara, 30: Aritzo), ***tsintsuréḍḍu*** (33: Gadoni) 'id.'
	AIS (448): [***ṣu***] ***tsutsurríttu*** (35: Laconi) 'id.'
	Terracini (1959: Map 1): ***ϑuϑurréri*** (18: Orosei, Orgosolo), [***su***] ***tϑutϑuréḍḍu*** (19: Nuoro), ***tuturréri*** (20: Dorgali), [***su***] ***tϑutϑurréri*** (21: Oliena), ***tsuntsurréḍḍu*** (27: Tonara, 32: Meana Sardo), ***tsuntsurríttu*** (29: Belvì, 35: Laconi) 'id.'
	Marcialis (2005: 67): ***tutturreri*** (21: Oliena) 'id.'
4: **ϑiϑimurr-*	DES (I: 556): ***tsintsimúrru*** (23: Fonni, 25: Ovodda), ***tsintsimurréḍḍu*** (Western Camp.), ***sattsamurréḍḍu*** (48: Villacidro) 'id.'
	AIS (448): [***su***] ***sintsimurréḍḍu*** (26: Milis) 'id.'
	Terracini (1959: Map 1): ***sintsimurréḍḍu*** (24: Bonarcado, 26: Milis), ***sittsimurréḍḍu*** (37: Marrubiu, 41: Mogoro, 42: Sardara, 45: San Gavino, 46: Arbus, 48: Villacidro, 49: Fluminimaggiore,), [***su***] ***sattsemurréḍḍu*** (52: Vallermosa) 'id.'

[160] These data are discussed in a similar fashion in Swanenvleugel (2024: 253–255).

5: **ϑurrundéllu*	DES (I: 556): ***tsur(r)undéḍḍu*** (Western Camp.) 'id.'
	Terracini (1959: Map 1): ***tsurrundéḍḍu*** (34: Oristano, 36: Usellus, 38: Isili, 39: Gesturi, 40: Orroli, 43: Siurgus, 44: Guasila, 47: San Nicolò Gerrei, 50: Monastir, 51: Dolianova, 53: Assemini, 54: Quartu Sant'Elena, 55: Cagliari, 56: Villasimius, 57: Sarroch, 58: Pula,) 'id.'

The words for 'bat' on Sardinia are diverse. Additional etymologically unrelated formations can be found in Terracini (1959). Some of the considerable variation may be explained as the result of folk etymologies or sound symbolism. All forms contain a base **ϑirr-* or **ϑurr-*. Types 1 and 2 may have been influenced by *ϑirrịare* 'to screech' (DES II: 556), suggesting that the forms with *ϑurr-* preserve the original vocalism. Type 4 was likely contaminated with *múrru*, meaning both 'grey' and 'muzzle'. Type 5 was contaminated with *arrúndili* 'swallow' (Guarnerio 1904a: 68; 1904b: 259; Terracini 1959: 15). Thus, underlying the various folk etymologies are **ϑurr-* (types 1 and 5) and **ϑiϑurr-* (types 2, 3, and 4). This is evidence for the presence of the pre-Roman morpheme **ϑi-* (§ 9.1.2).

4.5.3 Dog

Reconstruction	**Forms**
ị́ákaru	DES (I: 602): ***gjagar, jagaru*** (OSrd.), ***ğáɣaru*** (Log.) 'hunting dog'; deriv.: ***ğaɣarare*** (Log.), ***ağğaɣarai*** (Camp.), ***aužaɣarai*** (Ogl.) 'to disperse, to scatter, to assault etc.'

Even though *ğáɣaru* etc. 'hunting dog' is only attested in Logudorese (already as *gjagar, jagaru* in the Carta de Logu de Arborea; cf. Lupinu 2020, secs. 30, 83, 151), the existence of derived Camp. *ağğaɣarai*, Ogl. *aužaɣarai* 'to disperse, to scatter, to assault' suggests that the noun that must have been at the base of these verbal derivations was once more widespread in Sardinia. Wagner (1932: 232; DES I: 602) notes that this word is also widely present in Corsica, as the general word for 'dog': *ghjácaru*, *ghiàgaru*, *ị́ákaru*, *ị́ágaru*.[161] Further comparisons include Bq. *zakur* '(big) dog' and its more common diminutive *txakur* 'dog' (DES I: 602; Trask 2008: 45),[162] Byzantine Gr. ζαγάρι(ον), ζαγάρος 'hound, hunting dog' and Georg. ძაღლი *dzaghli* 'dog' < Proto-Kartvelian **ʒlaɣl-*. Schuchardt (1905: 227) has analyzed the Sardinian forms as a loan from Byzantine Greek, and posits a form **ζάγαρος to account for the Sardinian stress pattern. Byzantine loans in Sardinian are well-known (Paulis 1983; 2017).

[161] Cf. BDLC (s.v. dog).

[162] Michelena (1990: 191) rather sees *zakur* 'big dog' as a back-formation from original *txakur* 'dog'.

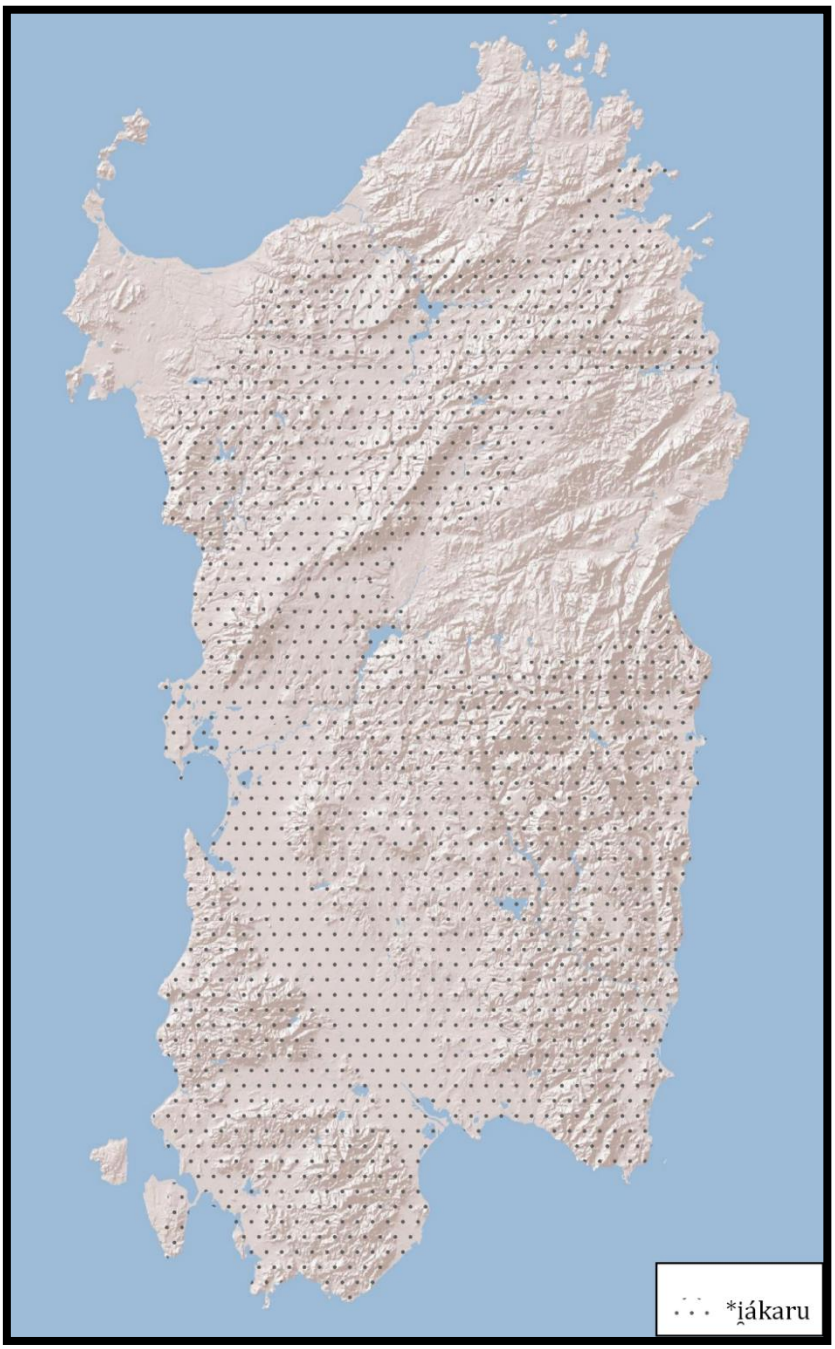

Figure 4.5.3: Dog

Regardless whether the Basque and Kartvelian forms are related, Schuchardt's (1905: 227) proposal is most attractive. Wagner's (1932: 232) identification of this word as pre-Roman is probably not necessary.

4.5.4 Doggie, puppy

Reconstruction	Forms
1: *ka𐞥𐞥éllu	AIS (1097): [***ṣu***] ***χattséḍḍu,*** [***ṣa***] ***χattsèḍḍa*** (1: Dorgali), [***ṣu***] ***ɣattséḍḍu,*** [***ṣa***] ***ɣattsèḍḍa*** (2: Busachi) 'doggie, puppy'
	Böhne (1950: 131): [***su***] ***ɣattséḍḍu*** (10: San Vito) 'id.'
2: *kal(i̯)éllu	DES (I: 270): ***kalléḍḍu*** (3: Baunei, 6: Perdasdefogu, 8: Gúspini), ***kallelléḍḍu*** (4: Oristano) 'id.'
	Böhne (1950: 131): ***kalleléḍḍu*** (10: San Vito) 'id.'
3: *kalili̯onéllu	DES (I: 270): ***kalliğğonéḍḍu*** (3: Baunei) 'id.'
4: *kalú𐞥𐞥u	DES (I: 270): [***su***] ***ɣalluččéḍḍu*** (5: Gairo), ***kallúttsu*** (6: Perdasdefogu), ***kallúčču, kalluččéḍḍu*** (7: Ballao, 9: Armungia) 'id.'

The use of this word as an endearment term explains its variety of formations, in which various diminutive suffixes (such as *-éḍḍu* < Lat. *-ellus* and *-úttsu* < Lat. *-uceus*/It. *-uccio*[163]) have replaced each other or combined with each other. The main problem presented by these words is the origin of *-ll-* in types 2–4. While *-ll-* could go back to *-li̯-* in most southern dialects (viz. 4: Oristano, 7: Ballao, 8: Gúspini, 9: Armungia, 10: Sarrabus), **-li̯-* is expected to yield *-l'-* or *-lğ-* in the Ogliastran dialects (viz. 3: Baunei, 5: Gairo, 6: Perdasdefogu) (Wagner 1941a: 147–148). On the other hand, simple **-l-* would have yielded *-β-* in Guspini (8) and *-ʔ-* in San Vito (10), and inherited **-ll-* regularly yields *-ḍḍ-* in all of Sardinian. We are thus dealing with various instances of *-ll-* that cannot regularly go back to the same phoneme. The most likely solution is to assume that the *-ll-* pronunciation spread from the region of Cagliari, where it is the regular outcome of both **-l-* and **-li̯-*. Among these, original single **-l-* is probably most likely, as its Cagliaritan reflex seems to spread to neighboring dialects more than *-ll-* < **-li̯-* (Wagner 1941a: 125).

[163] Cf. Wagner (1952: 27–28). Camp. *-učču* (in the forms of type 2) is probably rather from Italian *-uccio* than from inherited *-uceus*, which should have yielded *-úttsu*. However, its meaning is closer to the diminutive meaning of inherited *-úttsu* < *-uceus* than to the pejorative meaning of It. *-uccio*. The confusion between these two cognate suffixes, that are phonetically and semantically very similar, is not surprising.

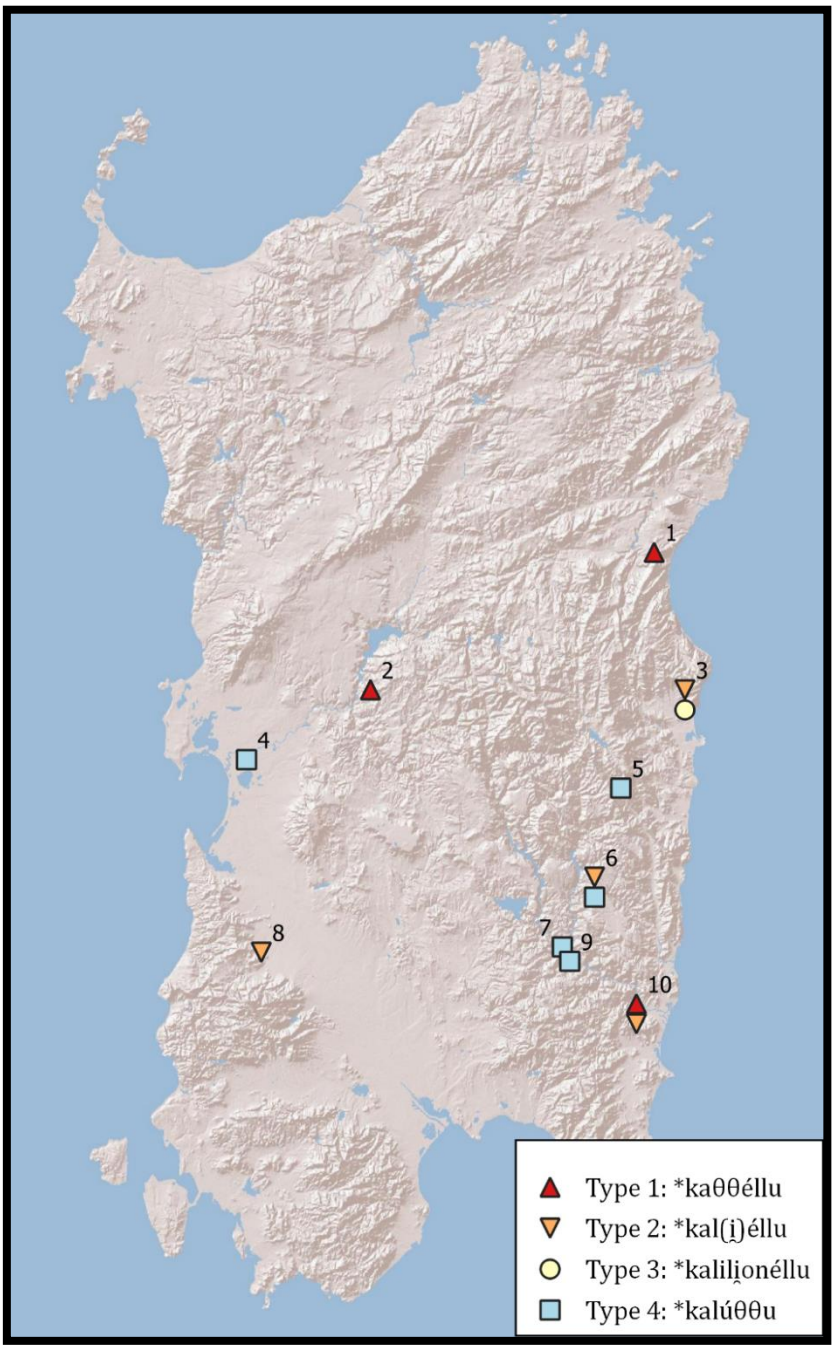

Figure 4.5.4: Doggie, puppy

The forms of type 1 contain *-tts-* instead of *-ll-*. Whereas *-tts-* could perhaps go back to original *-l̨i̯-* in Dorgali, which would correspond to Camp. *-ll-*, this does not work for Busachi (Wagner 1941a: 146–147). Although the superficial resemblance to *kallédḍu* etc. is striking, these words must instead go back to **kaϑϑéllu* and belong together with Log. *kattéḍḍu*, Nuor. *kaϑϑéḍḍu* etc. 'puppy' << Lat. *catellus* (DES I: 320).[164] Wagner (DES I: 320) compares Srd. *kallédḍu* (type 2) to Alb. *këlýsh* 'young animal, especially puppies and kittens' (cf. Demiraj 1997: 218–219). This is in turn compared by Meyer (1891: 186) to Lith. *kalẽ* 'bitch', Welsh *colwyn* and Bret. *kolen* 'puppy',[165] and by Demiraj (1997: 218–219) to Gr. σκύλαξ and the Hesychian gloss (from Elis) κύλλας 'puppy' (Hesychius Alexandrinus 2020: 689). The connection of Lith. *kalẽ* to the Greek forms is rejected by Beekes (2010: 1361).[166] Even if the Albanian and Greek forms are to be connected, a link between them and the Celtic and Sardinian forms is only a possibility at best. It should also be noted that all comparanda show different vocalisms, with **-u-* in Greek and Albanian, **-o-* in Celtic and **-a-* in Sardinian.

4.5.5 Fawn (*Cervus elaphus corsicanus*), lamb

Reconstruction	**Forms**
1: **bètt-/*bítt-*	DES (I: 211–212): ***bíta, bèta*** (Log., Camp.), ***bítti*** (1: Bitti) 'fawn'; ***bètte*** (2: Goceano reg.) 'lamb, small animal'
	Casu (2011: s.v. *bètta, bètte, bìtta, bìtti*): ***bètta*** 'fawn', ***bètte*** 'lamb, small gracious animal', ***bìtta*** 'deer',[167] ***bìtti*** 'deer, fawn'
	Puddu (2023: s.v. *bèta, bétu, bítalu*): ***bèta, bétu, bíta, bíti*** 'fawn'

[164] The development of Lat. *-t-* > Srd. **-ϑϑ-* is irregular however.

[165] Matasović (2008: 213) reconstructs these forms as PC **koligno-*.

[166] For alternative etymologies for Lith. *kalẽ*, cf. Smoczyński (2018: 474).

[167] Casu gives the definition *capriolo* 'roe deer'. However, the roe deer (*Capreolus capreolus* (Linnaeus, 1758)) does not occur on Sardinia. It is thus more likely that this word refers to the Sardinian red deer (*Cervus elaphus corsicanus* Erxleben, 1777). This subspecies of red deer is smaller than the continental red deer, which may be the reason that Casu refers to it as 'roe deer'.

2: *bíttara	DES (I: 211–212): ***bítaru, bitaréḍḍu*** (3: Dorgali) 'suckling calf', [***sa***] ***bíttara*** (4: Milis) 'female mouflon (*Ovis aries musimon* (Pallas, 1811)' 'lamb, small animal'
	Casu (2011: s.v. *bètta, bètte, bìtta, bìtti*): ***bìttera*** 'deer'
	Puddu (2023: s.v. *bèta, bétu, bítalu*): ***bítalu***, ***bíteru***, ***bítulu***, ***bíturu*** 'fawn', ***bítula*** 'doe'

This word is attested in old Sardinian texts: as [*capra de*] *bita*, [*capras de*] *vita* in the Condaghe di San Nicola di Trullas (Merci 1992: secs. 57, 137),[168] and as *bituru* in the Carte Volgari (Solmi 1905: sec. 11: 4). Although DES gives only a few specific locations for the attested forms — Casu and Puddu do not — it is clear that representatives of this word occur all over Sardinia, with 'fawn' as their most common meaning. The endings *-́aru*, *-́ara*, *-́era*, *-́alu*, *-́uru* are likely to be interpreted as representatives of the same non-inherited suffix *-́Vr*, which occurs in other words of obscure origin as well. The suffixes *-́ulu*, *-́ula* could be the reflexes of Lat. *-ul-*. In the un-suffixed forms we find an alternation between final *-a*, *-u*, *-e* and *-i*. The first two of these can easily be explained as secondary in order to express the gender of the animal in question. The correspondence between *-e* and *-i* is irregular however, and is difficult to explain morphologically. The fact that *-e* only occurs in *bètte* and *-i* only in *bítti* (rather than ***bétti* or ***bítte*), makes it look like the paragogic vowel that regularly develops in Sardinian after word-final consonants (Wagner 1941a: 57–62). However, whereas inherited final *-t* behaves like a singleton *-t-* (e.g. Camp. *kántaða* < Lat. *cantat* '(s)he sings'), *bítti/bètte* clearly has a geminate *-tt-*. It is attractive to see the variation in the final vowel of this word as the result of the adaptation of a non-inherited word to the Sardinian phonology. While languages with geminate stops in the syllable coda exist, it remains to be seen whether a **bitt* or **bett* as a pre-Roman source can really be posited for this word (§ 8.4.1). The variation of the tonal vowel *-è-* vs. *-í-* is also irregular (§ 8.3). Wagner (1932: 233) and Hubschmid (1953: 70; 1955: 189–190) compare various extra-Sardinian forms, such as Bq. *pitika* 'kid (young goat)', Gasc. *pito*, Cat. *bita*, *bitè*, For. *bette*, *bitte* 'goat' etc. (cf. also FEW XXII/I: 288). Bouda (1949: 334), Michelena (1959: 525) and DES (I: 212) prefer to interpret these and the Sardinian forms as parallel formations based on the sound a sheep or goat makes. While this may well have been the ultimate origin of the word, the vowel alternations as well as the non-inherited suffix *-́Vr* in Sardinian indicate that this word originated in a pre-Roman language.

[168] These instances of *bita, vita*, recorded in DES (I: 211), have been analyzed by Paulis (1997: 164) as the reflex of Lat. *vita* 'life', in which case they are not relevant to the current discussion.

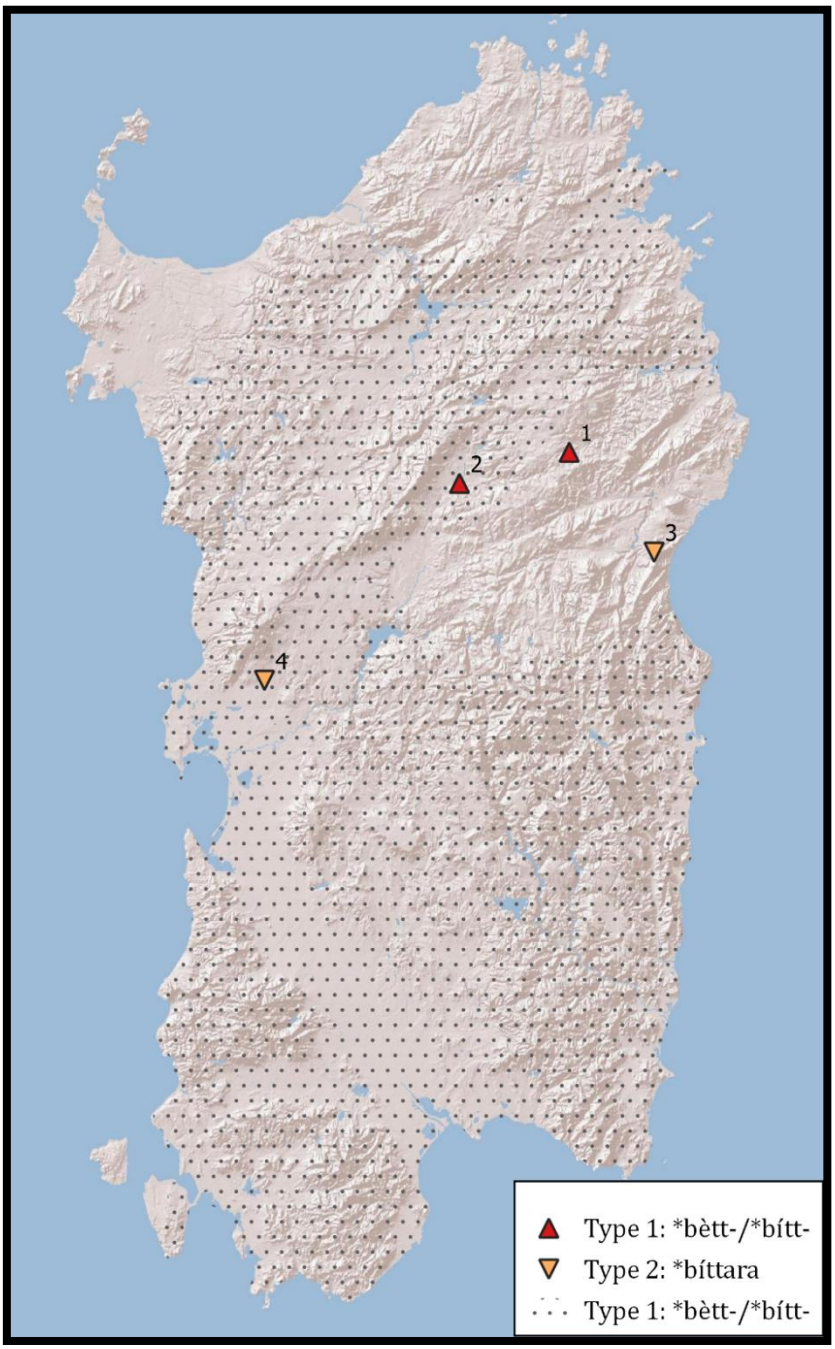

Figure 4.5.5: Fawn (*Cervus elaphus corsicanus*), lamb

Furthermore, the perfect formal correspondence between the Sardinian forms and those found in Catalan and Gallo-Romance is perhaps more than we would expect for independent onomatopoeic formations. Note further that For. *bette*, *bitte* even has the same stressed vowel alternation as Srd. *bètte*, *bítti*. Both could perhaps go back to **bētt-* vs. **bītt-*, which looks distinctly non-Romance.[169]

4.5.6 Fox (*Vulpes vulpes*)$_1$

Reconstruction	**Forms**
1: *glóll-	DES (I: 593): ***gròḍḍe*** (Nuor., Log.), ***gróḍḍu*** (8: Olmedo), ***gròḍḍo*** (18: Dorgali) 'fox (*Vulpes vulpes* (Linnaeus 1758))'
	AIS (435): [***su***] ***gròḍḍe*** (7: Ploaghe), [***su***] ***ɣròḍḍe*** (19: Macomer) 'id.'
	Wagner (1933: 503): ***gróḍḍẹ*** (1: Castelsardo), ***gróḍḍe*** (2: Sedini, 3: Bulzi, 4: Laerru, 5: Martis, 9: Ozieri, 10: Pattada, 11: Siligo, 12: Mores, 13: Ittireddu, 14: Giave, 15: Anela, 16: Bono, 20: Fonni) 'id.'
	Casu (2011: s.v. *gròddhe*): ***gròddhe*** 'id.'
2: *lóll-	DES (I: 593): ***lòḍḍe*** (6: Osilo), ***loḍḍósu*** (17: Orune) 'id.'
	Wagner (1933: 503): ***lóḍḍe*** (1: Castelsardo) 'id.'
	Casu (2011: s.v. *lòddhe*): ***lòddhe*** 'id.'

In this word, we find *gròḍḍe* (type 1) next to *lòḍḍe* (type 2).[170] Wagner (1922: 251; 1933: 503–508; DES I: 593) suggests a relation to the personal name *Lòḍḍo* (attested in older sources as *Lollo*, *Lollu*). The use of personal names to refer to animals, and especially foxes, is rather common in Sardinian and elsewhere in Romance (e.g. Srd. *marğáni* 'fox' < Lat. *Marianus*, Fr. *renard* 'id.' << *Reinhard*). The advantages of treating this word as an original personal name are that we may explain the variation between final *-o* and *-e* as the regular outcomes of the original nominative and vocative case in *-o/u* and *-e* respectively, and the same name could lie at the base of (Nuoro) *gròḍḍo* 'wren' (§ 4.4.12) (Wagner 1933: 503).

[169] The length of the vowels is not needed for Sardinian. In most other Romance languages however, *-ī-* is the reflex of Lat. *-ī-*. If this word entered the languages in question after the Romance vowel developments, we could of course simply posit **bett-* vs. **bitt-*.

[170] Wagner's (1933: 503) *gróḍḍe* and *lóḍḍe* must probably be interpreted as having an open *ò*. There is no source for a closed *ó* preceding an *-e-*, and in AIS (435) and DES (I: 593) he also writes open *ò*.

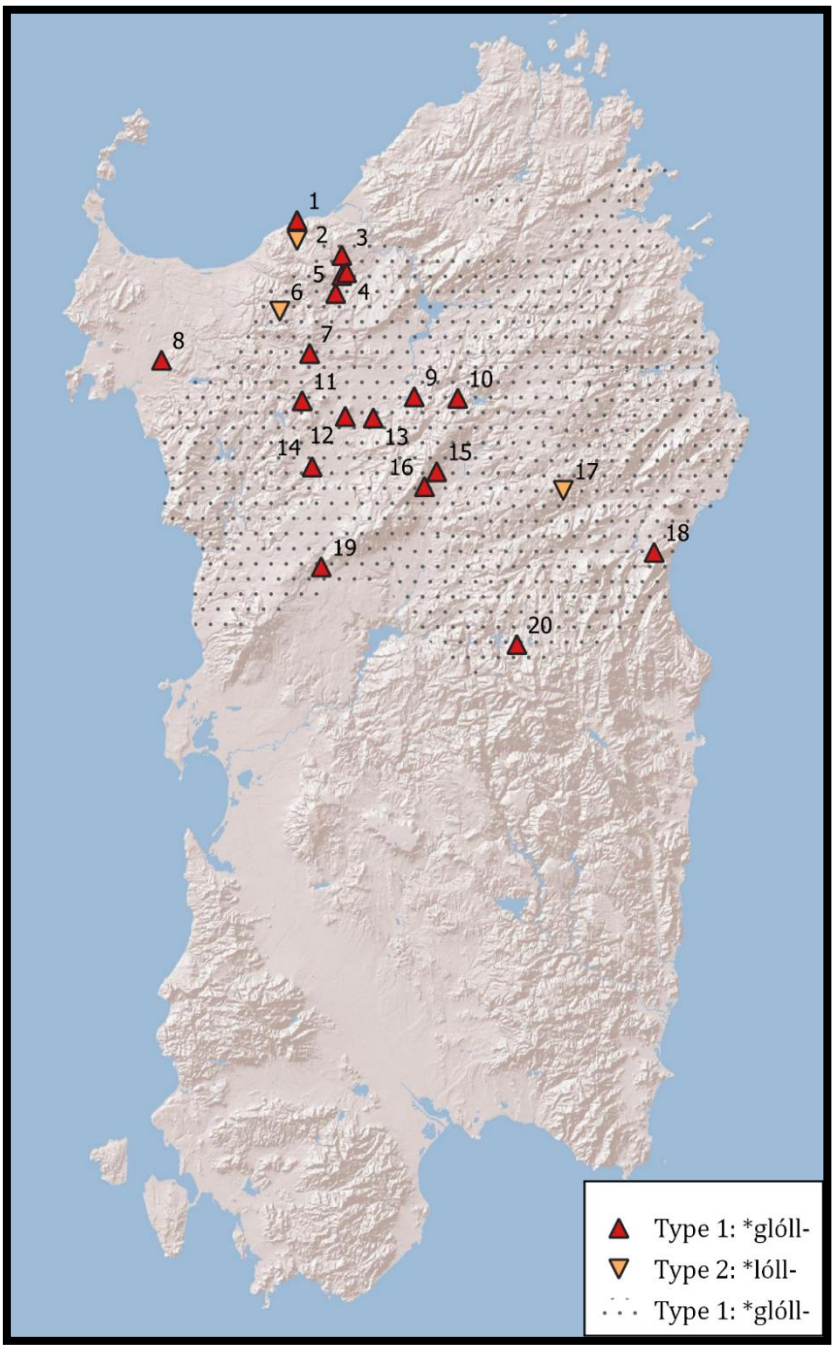

Figure 4.5.6: Fox (*Vulpes vulpes*)

A difficulty is the variation between initial *gr-* and *l-*. This is reminiscent of the outcomes of inherited initial *gl-*; e.g. Srd. *gránde* vs. *lánde* < Lat. *glans*, *-ndis* 'acorn' (Wagner 1941a: 162). But, although the evidence for this development is limited to only a handful of words, the reflex *l-* appears to be much more widespread than *gr-*, and is especially common in Logudorese (Wagner 1941a: 162). For the word for 'fox' however, which is limited to the Logudorese and Nuorese dialects, *gr-* is much better represented than *l-*, which is only found in three geographically discontiguous places (1: Castelsardo, 6: Osilo, 17: Orune). It is in any case difficult to explain *gr-* regularly from **l-*. In order to keep the personal name *Lollo* as an etymology for this word we need to assume the secondary introduction of *gr-*, probably caused by association to another word starting in *gr-*. The exact source is unclear however. The name *Lollo* itself is not from Latin and is perhaps of pre-Roman origin.

4.5.7 Fox, vixen (*Vulpes vulpes*)$_2$

Reconstruction	**Forms**
****kalíʒʒ-***	DES (I: 270): ***gallíssi*** (1: Gairo), ***kal(l)íttsa*** (2: Isili, 4: San Nicolò Gerrei, 5: Villasalto, 6: San Vito, 7: Villaputzu, 11: Villasimius) 'vixen', ***gallítsu*** (9: Dolianova) 'tabooistic name of the fox (*Vulpes vulpes*)'
	Wagner (1933: 508–509): ***kal(l)íttsa*** (3: Armungia, 8: Muravera, 10: Serdiana) 'vixen'
	Puddu (2023: s.v. *calítza*, *gallícciu*): ***calítza, callítza, galítza*** 'vixen', ***gallíciu, gallíssi, gallítzu*** 'fox'

These forms probably go back to a base *kal-, with the inherited suffix **-iʒʒu/a* < Lat. *-icius* (cf. Wagner 1952: 26–27). Wagner (DES I: 270) connects these forms to *kallédḍu* etc. 'dog' (§ 4.5.4), which is plausible. The words for 'fox/vixen' have a more restricted distribution than those for 'puppy' etc.: the former seem to be limited to the southeast corner of the island. Puddu's form *gallíciu* can only regularly be from Desulo (in the Barbagia) or from the Sulcis region (in the farthest southwest of the island), since these are the only dialects that have *-č-* as the regular outcome of **-ʒ-* (Wagner 1941a: 295). If this form is taken seriously it implies a slightly more northern (if from Desulo) or western (if from Sulcis) extension for this word. The variant *gallíssi* (1: Gairo) has an irregular final *-i*, which cannot be from *-icius*, unless it is from an original vocative (cf. § 4.5.6). The initial *g-* in several of the forms is due to reanalysis from syntactic constructions in which the initial *k-* preceded by a vowel regularly lenites. For the phonetic details of *-l(l)-* and external comparanda of this word, cf. § 4.5.4.

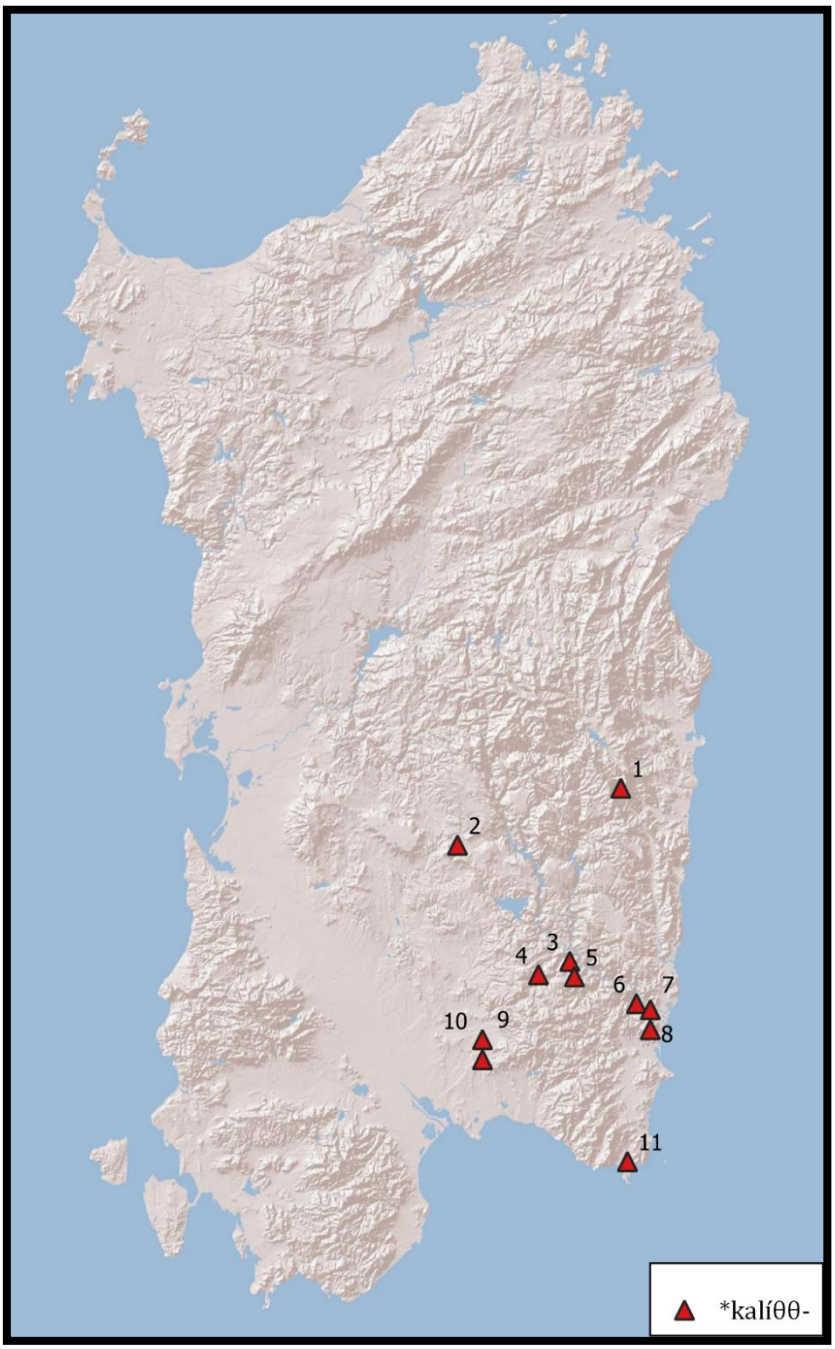

Figure 4.5.7: Fox, vixen (*Vulpes vulpes*)

4.5.8 Hare (*Lepus sp.*)

Reconstruction	Forms
1: *lèppar-	AIS (521): ***lèpparu*** (1: Tempio), ***lèpari*** (2: Sassari) 'hare'
2: *lèpper-	DES (II: 22–23): ***lèppere*** (Log.), ***lèppiri*** (Camp.) 'hare', ***lepperéḍḍu*** (7: Macomer, 8: Orani, 9: Orgosolo, 11: Ollolai, 19: Tonara) 'butterfly'
	AIS (521): ***lèpere*** (3: Ploaghe, 7: Macomer), ***lèppere*** (13: Santu Lussurgiu), ***lèppiri*** (16: Milis, 26: Mogoro) 'hare'
	Böhne (1950: 131–132): [***su***] ***ʔèppiri*** (29: Muravera) 'hare'
3: *lèppor-	DES (II: 22–23): ***lèppore*** (Nuor.), ***lèppuri*** (Camp.) 'hare', ***lepporèḍḍa*** (6: Dorgali), ***lèporo, lepporéḍḍu*** (12: Gavoi, 14: Fonni, 15: Urzulei, 17: Triei, 18: Baunei) 'butterfly'
	AIS (480, 521): ***lèppore*** (4: Bitti, 6: Dorgali, 18: Baunei), ***lèpore*** (5: Nuoro, 14: Fonni) 'hare'; ***lèppore*** (20: Desulo) 'butterfly'
4: *léppuri	DES (II: 22–23): ***léppuri*** (Camp.) 'hare'; ***leppuréḍḍu*** (10: Olzai, 22: Meana Sardo, 23: Gadoni), ***léppuri*** (21: Aritzo) 'butterfly'
	AIS (521): ***léppuri*** (18: Escalaplano, 24: Perdasdefogu), [***su***] ***ʁépuị*** (30: Sant'Antioco) 'hare'
5: *léppi	DES (II: 22–23): ***léppi léppi*** (24: Seui, 25: Sadali) 'butterfly'

In various Sardinian dialects, the word for 'hare' is etymologically identical to the word for 'butterfly', the only distinction being that the latter often has the diminutive suffix *-éḍḍu/-èḍḍa*. Wagner (DES II: 23) considers two scenarios for this peculiar *Benennungsmotiv*. One is the idea that the fluttering of butterflies was perceived as similar to the jumping of hares. The other is that it is sound-symbolic. Since all listed words for 'butterfly' can synchronically be perfectly understood as simple derivation from the word for 'hare', I prefer not to assume an independent origin of these words, as suggested by Wagner's second scenario. The fact that the 'lapwing' too is in some places called *lèppere de ardzòla* (Wagner 1928: 30), demonstrates that the word for 'hare' could semantically be extended to certain flying animals (cf. § 4.4.9).

As for the forms themselves, there is obviously a connection to Lat. *lepus*, *-oris* 'hare'. However, this would have yielded Log. ***lèβore*, Camp. ***lèβuri*, neither of which is attested. The voiceless *-p(p)-* that is present in the Sardinian forms for 'hare' can only be from original geminate **-pp-* or **-bb-*. Moreover, the different formation types found across Sardinia present varying vocalism in the second syllable.

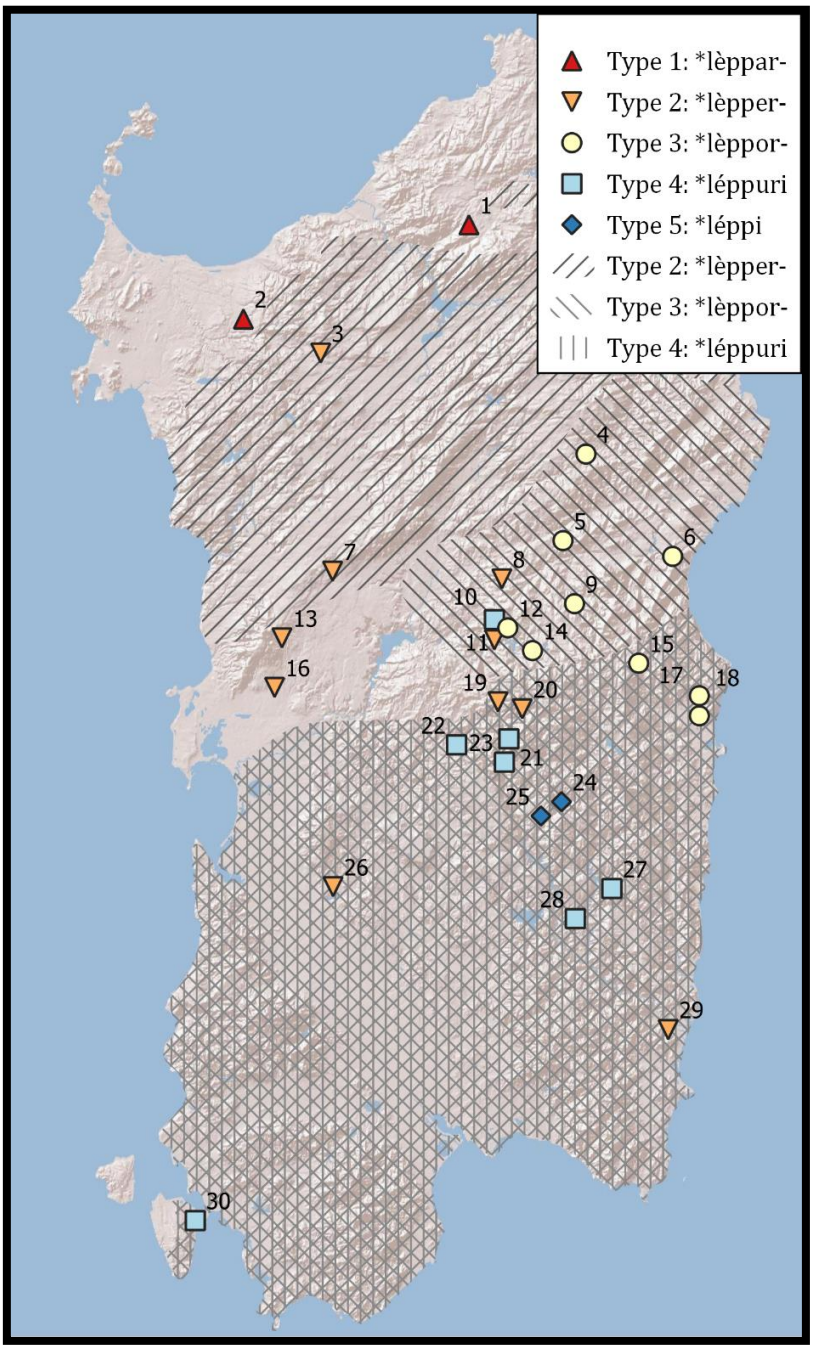

Figure 4.5.8: Hare (*Lepus sp.*)

Wagner (DES II: 22) rejects various proposed folk etymologies that could explain the unexpected *-pp-* and instead suggests influence from a pre-Roman **lappar-*. This latter form is posited by Bertoldi (1937a: 146) as the source of Lat. *lepus, -oris* 'hare', of Siculian λέποριν 'hare' (Varro 2015: 108), and of λεβηρίς, denoting the rabbits in Turdetania in Strabo (1877: chap. 3.2.6) and apparently also in use in Massilia (Liddell et al. 1978: s.v. λεβηρίς). It is not clear to me why the aforementioned forms need **lappar-*, as none of them contains an *-a-*, but this reconstruction is in any case needed for Pg. *láparo*, OProv. *laperel* 'young rabbit (*Oryctolagus cuniculus* (Linnaeus, 1758))', OFr. *lapriel* 'rabbit'. These and various other forms for rabbit (e.g. Pg. *lapim,* Arag. *llapí* 'young rabbit', Fr. *lapin* 'rabbit') are compared by Hubschmid (1943: 246–250). The lack of a clear etymology for Lat. *lepus* 'hare' (cf. de Vaan 2008: 335), combined with the irregular correspondence between it and the various Romance forms denoting rabbits and hares, suggests that this word entered Latin and Romance independently on multiple occasions. The vocalism of the Sardinian forms is clearly influenced by Latin *lepus, -oris,* but the presence of geminate *-pp-* brings it closer to e.g. Pg. *láparo* 'young rabbit'. Hubschmid's (1943) derivation of the various words for 'rabbit' etc. from pre-Roman **lappa* vel sim. 'cave; stone slab' is perhaps possible, but difficult to verify. His proposal that **lappar-* spread from Iberia with the domestication of rabbits, which are native to the Iberian Peninsula, does not immediately follow from the attestation of similar forms in the meaning 'hare' in Latin, Sardinian and colonial Greek.

4.5.9 Marten (*Martes martes*)[171]

ReconstructionForms	
1: *bassíle	Wagner (1934: 483–489; DES I:139): ***bassíli*** (1: Santa Teresa Gallura, 6: Calangianus, 7: Nuchis, 9: Tempio Pausania), ***bassíle*** (2: Luogosanto, 10: Bortigiadas, 12: Bulzi, 15: Laerru), ***basíli*** (8: Castelsardo, 13: Porto Torres, 17: Sorso, 25: Sassari), ***vassíli*** (11: Sedini) 'marten (*Martes martes* (Linnaeus, 1758))

[171] These data are discussed in a similar fashion in Swanenvleugel (2024: 232–236).

2: *_assíbile_	Wagner (1934: 483–489; DES I:139): ***assíle*** (3: Luras, 5: Olbia, 14: Perfugas, 16: Monti, 18: Sennori, 19: Berchidda, 20: Nulvi, 21: Martis, 22: Chiaramonti, 23: Osilo, 24: Tula, 26: Oschiri, 29: Cargeghe, 30: Ploaghe, 33: Olmedo, 34: Alà dei Sardi, 35: Florinas, 37: Torpè, 38: Ardara, 41: Ozieri, 42: Pattada, 43: Buddusò, 44: Siligo, 46: Banari, 47: Putifigari, 48: Bessude, 50: Ittireddu, 51: Bonnanaro, 55: Cheremule, 61: Monteleone, 63: Nule, 64: Bultei, 65: Benetutti, 66: Giave, 67: Anela, 68: Cossoine, 69: Bonorva, 70: Bono, 75: Pozzomaggiore, 76: Semestene, 78: Bottida, 79: Burgos, 80: Esporlatu, 83: Montresta, 85: Illorai, 86: Bolotana, 88: Lei, 90: Sindia, 91: Bosa, 94: Bortigali, 95: Suni, 98: Flussio, 99: Magomadas, 100: Macomer, 102: Tresnuraghes, 105: Dualchi, 107: Noragugume, 108: Scano di Montiferro, 109: Borore, 111: Sennariolo, 129: Bonarcado, 130: Ovodda), ***asíli*** (4: Aggius), ***ansíle*** (27: Muros, 28: Ossi, 31: Usini, 32: Codrongianus, 39: Ittiri, 49: Mores, 52: Thiesi, 53: Borutta, 54: Torralba, 58: Romana, 77: Padria, 113: Cuglieri), ***essíle*** (56: Villanova Monteleone, 93: Silanus, 101: Birori), ***asseíle*** (89: Orotelli), ***assaíle*** (120: Santu Lussurgiu), [***sa***] ***síβile*** (126: Fonni) 'id.'
3: *_grassíbile_	Wagner (1934: 483–489; DES I:139): ***grassíle*** (36: Posada, 40: Lodè, 45: Siniscola, 57: Onanì, 59: Garofai, 60: Bitti, 62: Lula, 71: Irgoli, 72: Orune, 73: Onifai, 74: Loculi, 81: Galtellì, 82: Orosei, 92: Dorgali, 110: Mamoiada), ***grassíβile*** (87: Nuoro, 96: Oliena, 97: Oniferi, 103: Orani, 106: Sarule) 'id.'
4: *_kassíbile_	Wagner (1934: 483–489; DES I:139): ***kassíβile*** (104: Ottana), ***ʔassíβile*** (112: Orgosolo, 114: Olzai, 117: Ollolai, 118: Gavoi), ***kassíle*** (115: Sedilo, 116: Aidomaggiore, 119: Domusnovas Canales, 121: Norbello, 122: Soddì, 123: Abbasanta, 124: Ghilarza, 125: Zuri, 127: Bidonì, 128: Boroneddu, 131: Paulilatino, 132: Seneghe, 133: Milis, 134: Narbolia), ***kassíβi*** (135: Bauladu, 137: Tramatza, 138: Villanova Truschedu, 140: Siamanna, 141: Oristano), ***kassí*** (136: San Vero Milis), ***kassíli*** (139: Cabras) 'id.'

All of these forms can be derived from a base **assíbile*, which is not of Latin origin. The forms of type 1 are restricted to Gallurese and Sassarese, and were likely influenced by the personal name *Basilio* (DES I: 139). Types 3 and 4 have added initial **gr-* and **k-* respectively. Type 3 is probably a variant of type 4, possibly remodeled after *grássu* 'greasy' (DES I: 139). The opposition between

**assíbile* (type 2) and **kassíbile* (type 4, and probably type 3) is difficult to explain away within Sardinian. The addition of an initial sequence **kV-* occurs in several other Sardinian words of possible pre-Roman origin too, and may originally have been some kind of morpheme (§ 9.1.1).

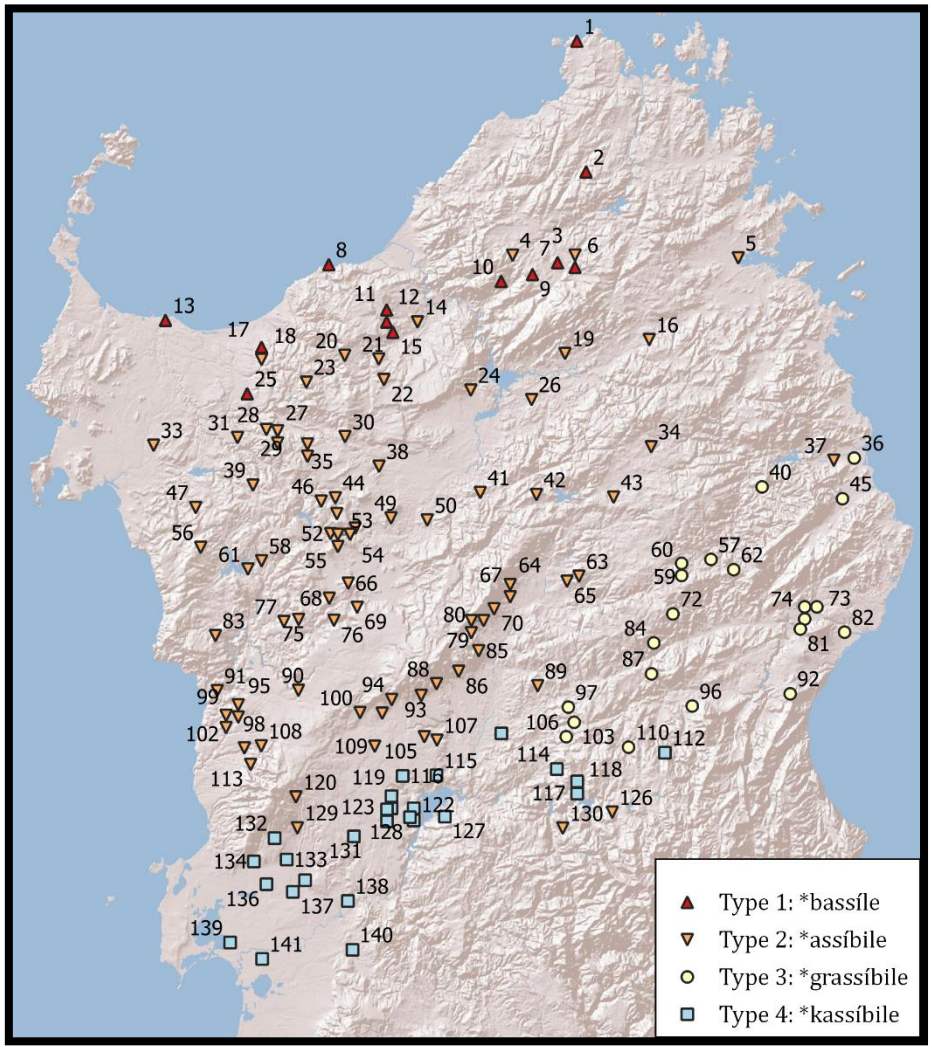

Figure 4.5.9: Marten (*Martes martes*)

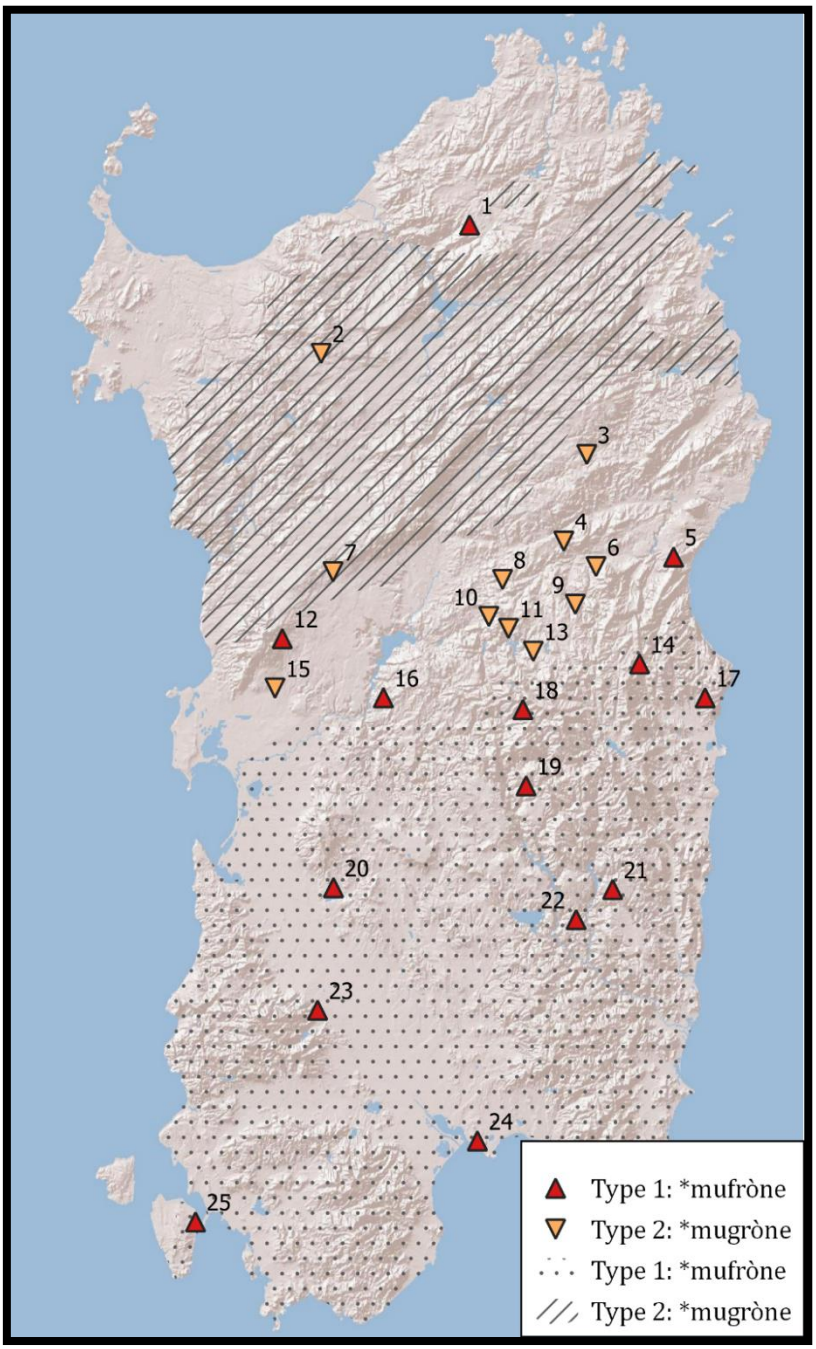

Figure 4.5.10: Mouflon (*Ovis aries musimon*)

4.5.10 Mouflon (*Ovis aries musimon*)

Reconstruction	**Forms**
1: *_mufròne_	DES (II: 134–135): ***murvòni***, ***murvồi***, ***mruvồi*** (Camp.), ***muvròne*** (5: Dorgali, 21: Perdasdefogu), ***murvòne*** (12: Santu Lussurgiu), ***mufròne*** (14: Urzulei, 18: Desulo) 'mouflon (*Ovis aries musimon* (Pallas, 1811))'
	AIS (519): ***muvròni*** (1: Tempio, 19: Seulo), ***mruvòne*** (16: Busachi), ***muvròne*** (17: Baunei), ***murvồi*** (20: Mogoro, 23: Villacidro), ***mruvòni*** (22: Escalaplano), ***murvòni*** (24: Cagliari, 25: Sant'Antioco) 'mouflon'; ***múvra*** (5: Dorgali), ***múrva*** (12: Santu Lussurgiu; 20: Mogoro, 23: Villacidro), ***mrúva*** (16: Busachi, 22: Escalaplano), ***múvra*** (17: Baunei, 21: Perdasdefogu), ***múfra*** (18: Desulo) 'mouflon (female)'
2: *_mugròne_	DES (II: 134–135): ***muròne*** (Log.), ***muɣròne*** (3: Bitti, 4: Nuoro, 6: Oliena, 8: Orani, 9: Orgosolo, 10: Olzai, 11: Gavoi), ***murgòne*** (13: Fonni) 'mouflon'
	(AIS 519): ***muròne*** (2: Ploaghe, 7: Macomer), ***murồi*** (15: Milis) 'mouflon'; ***múɣra*** (3: Bitti, 4: Nuoro), ***múrga*** (13: Fonni) 'mouflon (female)'

The mouflon is a type of wild sheep native to Sardinia and Corsica. The Sardinian name for this animal is mirrored by Cors. *muvrone*, *-i*, feminine *muvra*, *mufra* (DES II: 134; BDLC: s.v. mouflon), whence it entered French as *mouflon* and was subsequently adopted by other European languages (FEW VI/III: 186). In Latin sources we find *mufron* (Polemius Silvius, 5th c. CE), referring to a species of wild sheep in Sardinia in Corsica (TLL VIII: 1557; Ernout & Meillet 2001: 417).

As Wagner (DES II: 134) notes, the Sardinian forms require two pre-forms: **mufròne* for type 1 and **mugròne* for type 2.[172] In light of the wider distribution of type 1 **mufròne* in the Sardinian dialects, its presence in Corsican and its early attestation in 5th century Latin, it is attractive to assume that **mufron-* is original. Type 2 **mugron-* is restricted to Logudorese and Nuorese and may have been an early innovation. This word has no Latin etymology and, especially since it specifically refers to animals endemic to the Tyrrhenian islands, it is suspect of being of pre-Roman origin. Besides Lat. *mufron*, there is also Lat. *musmō* 'wild sheep with skin like a goat, especially common in Corsica', and Pliny mentions

[172] Wagner (DES II: 134) states that Log. *muròne* can go back both to **mugròne* and to **mufròne*, but post-vocalic *-fr-* would likely have yielded ***-vr-* (Wagner 1941a: 90), indicating that the Logudorese forms go back to **mugròne*.

that "the ancients called the offspring of the *musmō* and of sheep *imbrōs*" (Walde & Hofmann 1954: 118). It is not clear what the relation between Lat. *musmō*, *mufron* and *imbro-* is. Whereas *musmō* and *mufron* share their syllable structure and vocalism, the consonant clusters *-sm-* and *-fr-* are hard to reconcile. *Imbro-* on the other hand is closer with its sequence *-br-*, but its vocalism differs from that of *mufron*. The late attestation of *mufron* indicates that it may have been borrowed directly from Sardinian or Corsican Romance. Since Lat. *musmō* and *imbro-* appear to refer to the same animal (or its offspring), it is attractive to assume that these forms too entered Latin from a language spoken in one of these islands, but the modern Corsican and Sardinian dialects show no evidence for the existence of similar forms.[173]

Next to Srd. **mufròne*, **mugròne*, we find **múfra*, **múgra* everywhere, referring specifically to a female mouflon. This shows that *-on-* is a suffix, but it is unclear whether it was added by Latin or if it was a feature of the original source language (§ 9.2.4.2). Finally, Wagner (DES I: 134) mentions Georg. *maklo* 'deer', without making his ideas on its relation to the Sardinian words explicit. Besides the obvious semantic discrepancy between 'mouflon' and 'deer', the comparison runs into phonetic trouble too if we depart from **mufr-* as the most likely base form of the Sardinian and Corsican words.

[173] Pittau (1984: 199–200), however, argues that Lat. *musmō* and Lat. *mufron* are unrelated.

5 Geography

This chapter contains words referring to geographical items. Many of the meanings of the forms discussed are highly specific, and at times particular to the variety in which they are attested. Included are features of the natural landscape (§ 5.1), as well as landmarks related to human activity (§ 5.2).

5.1 Natural features

5.1.1 Chasm; pile

Reconstruction	Forms
1: *_núrra_	DES (II: 176–179): ***núrra*** (1: Posada) 'pile of wood or rocks', (3: Bitti) 'heap (of fruit etc.) sticking out above the rim of a basket'
2: *_núrra_	DES (II: 176–179): ***núrra*** (2: Siniscola, 5: Nuoro, 7: Orgosolo) 'ravine in the shape of a pit',[174] (4: Lula, 6: Oliena) 'chasm, crack'

Although all of these forms are formally identical, they have been subdivided into two types on the basis of their diverging meanings (i.e. 'pile' next to 'ravine, chasm'), which do however co-occur in the same geographical area. This odd fact was proposed by Hubschmid (1953: 45) to be the result of the original meaning being something like 'cup'. This would subsequently have come to denote depressions in the terrain on the one hand, and elevations or heaps (in the shape of an upside-down cup) on the other. He provides several parallels from other languages. In a later article, Hubschmid (1958: 220) abandons this idea and opts for 'elevation' as the original meaning, stating that "the word would simply have denoted elevations most of the time, and more rarely depressions". This is not in line with the Sardinian evidence however, where *núrra* refers to some kind of depression more often than to and elevation. Hubschmid's original explanation seems more plausible.

[174] "[B]urrone a forma di pozzo" (DES II: 177).

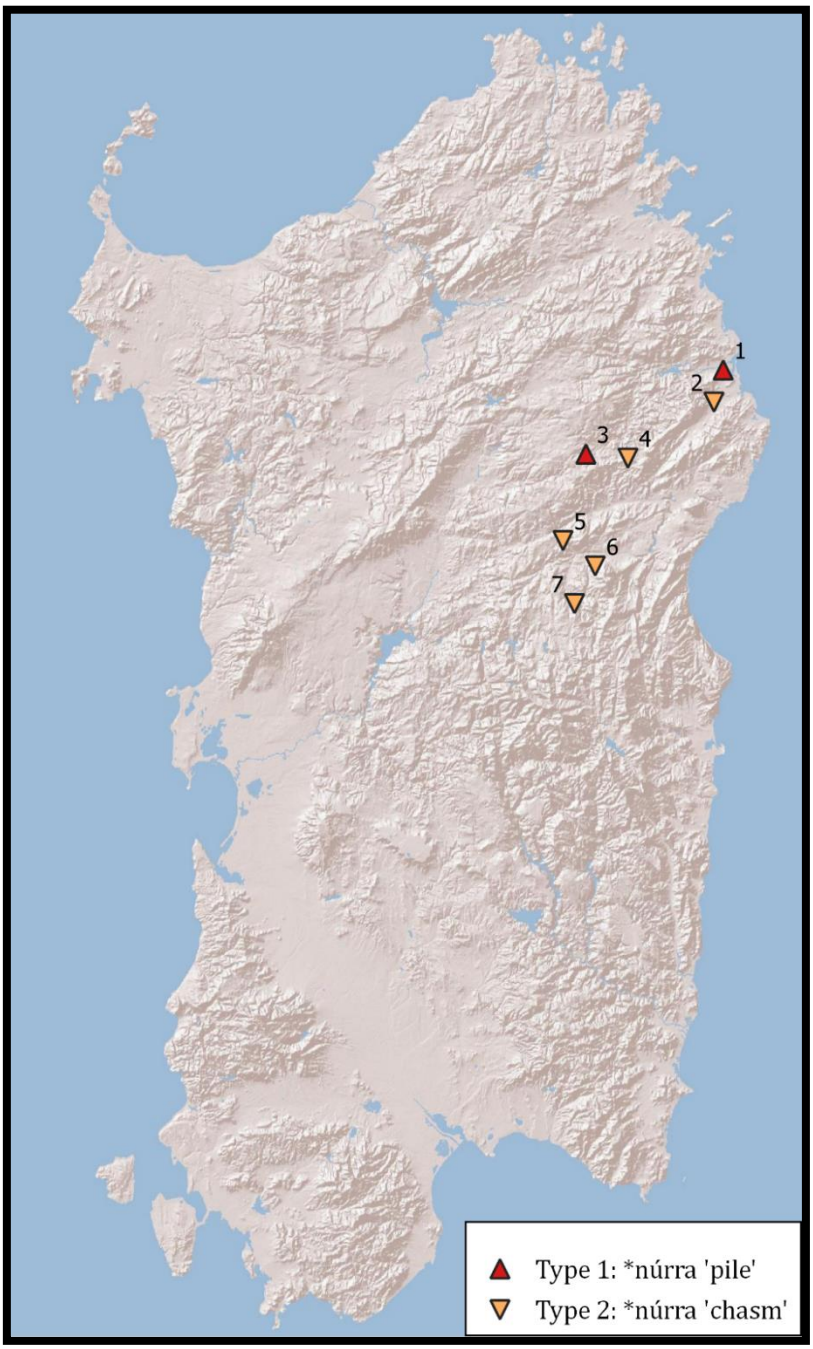

Figure 5.1.1: Chasm; pile

In DES (II: 176-79), these forms are discussed under the lemma *nuráke* 'nuraghe' (§ 5.2.2). The idea that *nuráke* 'nuraghe' would be a pre-Roman derivation of *núrra* was first proposed by Pais (1910: 190), who further compares the toponyms *Nurra* (a historical region), *Nurri* (a town), and others. Wagner (1932: 230–231) has some justified doubts as to the semantic and formal feasibility of connecting *núrra* 'pile, heap; ravine, chasm' with *nuráke* 'nuraghe'. Semantically, it seems by no means self-evident that the builders of these monumental structures would have referred to these feats of early architecture as a 'heap' or 'elevation' — or as comparable to a 'cave', as Terracini (1936: 55–56) suggests. Formally, *nuráke* invariably has a singleton *-r-*, while *núrra* always has geminate *-rr-*. Hubschmid (1953: 47; 1958: 219) attributes this to degemination rules related to stress placement in the pre-Roman substrate language and provides parallels from the Iberian Peninsula. He does not provide evidence for a similar process in Sardinian however. While it is possible that Srd. *nuráke* and *núrra* originally belonged to the same pre-Roman language, there is in my view not enough evidence to identify the former as a derivation from the latter.

This reduces the comparanda of Srd. *núrra* to various toponyms within and outside Sardinia (Pais 1910: 189–190; Wagner 1932: 229; Hubschmid 1953: 45–47; 1958: 219–220), and to various forms in Ibero-Romance varieties mentioned by Hubschmid (1953: 47; 1958: 219): e.g. Ast. *noriu* 'pile of rocks made in an estate', *noiro* (Boal) 'bump of earth with undergrowth', Galic. *nórrago* 'big knot in the root of a tree' etc. While the meaning of the first of these forms is close to that attested in Posada (1), the large semantic distance to the other forms (e.g. from 'chasm' in Sardinia to 'knot in the root of a tree') make an etymological connection doubtful.

5.1.2 Clay

Reconstruction	**Forms**
***loḷ̭iána**/ ***luḷ̭iána**	DES (II: 47): ***ludzána*** (Nuor., Log., Sass., 4: Scano di Montiferro), ***lodzána*** (Campidano di Milis reg.), [***terra***] ***luzana*** (3: Nuoro), ***lurdzána*** (5: Norbello), ***ložána*** (8: Desulo) 'clay'; ***ludzána*** (1: Posada, 2: Siniscola) 'layer of black sediment that is deposited by a river'
	AIS (416): ***lužă̂a*** (6: Milis), ***luǧána*** (7: Busachi) 'clay'

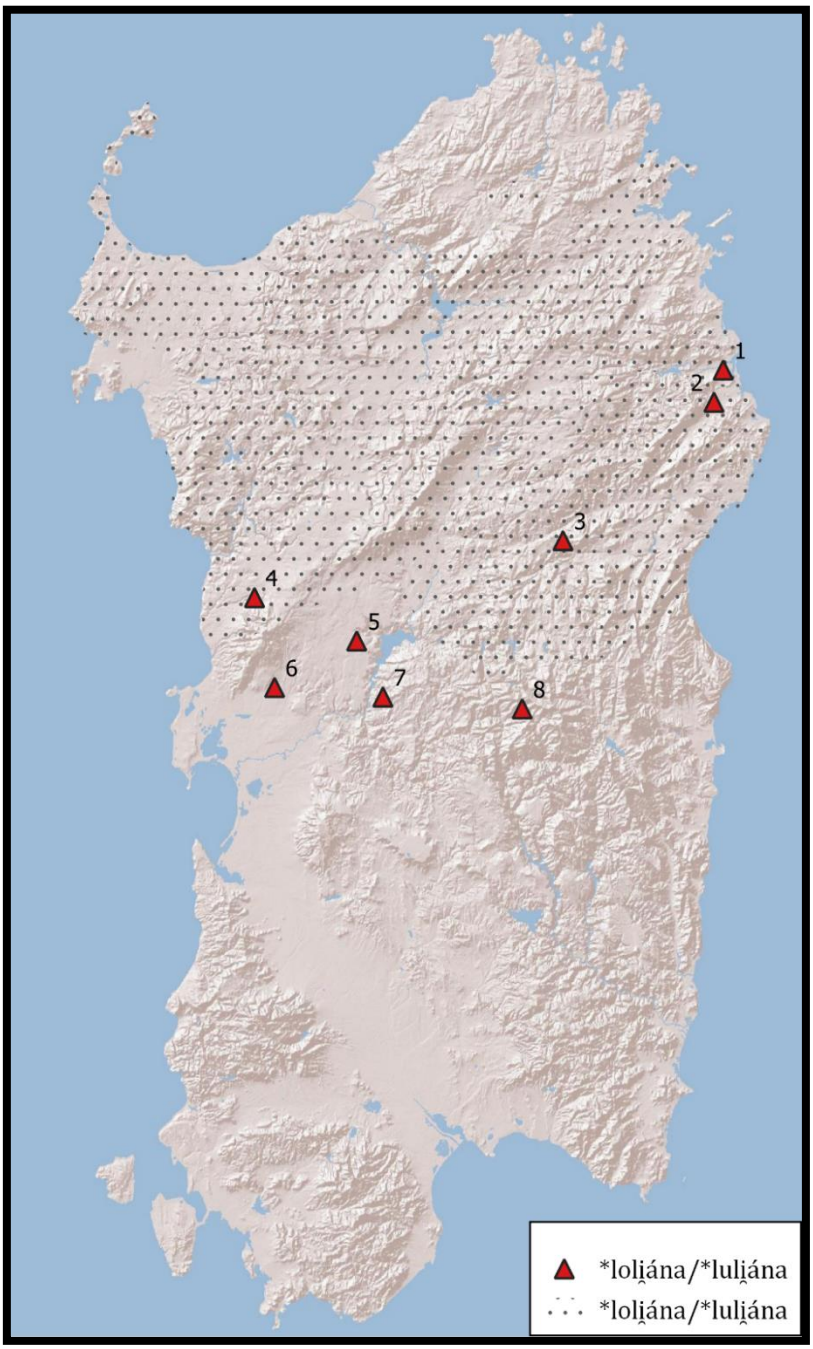

Figure 5.1.2: Clay

These words for 'clay', found in the northern half of Sardinia and attested as *Luianathos* in the Condaghe of San Pietro di Silki (Bonazzi 1900: sec. 202) and as *luiana* in the Condaghe of San Nicola di Trullas (Merci 1992: sec. 125), cannot go back to any Latin source. Wagner (DES I: 47) reconstructs **luḷiána* and considers a pre-Roman origin. The alternation between pretonic *-u-* (*ludzána* etc.) and *-o-* (*lodzána* etc.) could be of recent date (cf. Wagner 1941a: 28–29). Paulis (1992: 261) proposes a derivation from Lat. *lolium* 'darnel', because the type of clay in question would be conducive to the growth of this plant. Formally, a derivation **loḷiana* << *lolium* is attractive, but I am not convinced by the various semantic steps required for Paulis' etymology. The etymology of this word remains uncertain.[175]

5.1.3 Den (for animals)

Reconstruction	**Forms**
1: *kála	DES (I: 264–265): ***kála*** (Log., Camp.) 'shelter where river eels take refuge'
	Wagner (1932: 242): ***kála*** (Log., Camp.) 'cove, inlet'
2: *kála	DES (I: 264–265): ***kála*** (Nuor.), [***sa***] ***kála*** [***e maṛiáne***] (2: Dorgali), [***sa***] ***ʔála*** [***essu marğáne***] (4: Fonni) 'fox den'; ***kála*** [***e vormíkas***] (1: Nuoro), [***sa***] ***kála*** [***e vrummíka***] (2: Dorgali), [***sa***] ***ʔála*** [***e sa omílʔa***] (3: Orgosolo), [***sa***] ***ʔála*** [***e bobborissína***] (4: Fonni), ***kála*** [***e fromíha***] (5: Urzulei) 'anthill'

Although all abovementioned forms are formally identical, they have been divided into two types according to their rather divergent meanings. Those in type 1, attested in Logudorese and Campidanese, denote some kind of inlet or cove in the shore along a body of water. Those in type 2, on the other hand, are restricted to the Nuorese dialects and refer to animal dens — specifically fox dens and anthills according to DES (I: 264–265). I agree with Wagner (1932: 242) that *kála* in the meaning 'cove, inlet' (type 1), must be a loan from either Sp., Cat. or It. *cala* 'cove, inlet, creek', but that such a recent loan scenario is semantically difficult for the Sardinian forms referring to 'fox dens' and 'anthills'.

[175] Equally unsatisfactory is the idea that homonymous *ludzána* 'type of red ant' would be derived from Lat. *lolium* because the redness of the ants was supposedly associated to the redness of the soil in which darnel grows (Paulis 1992: 261).

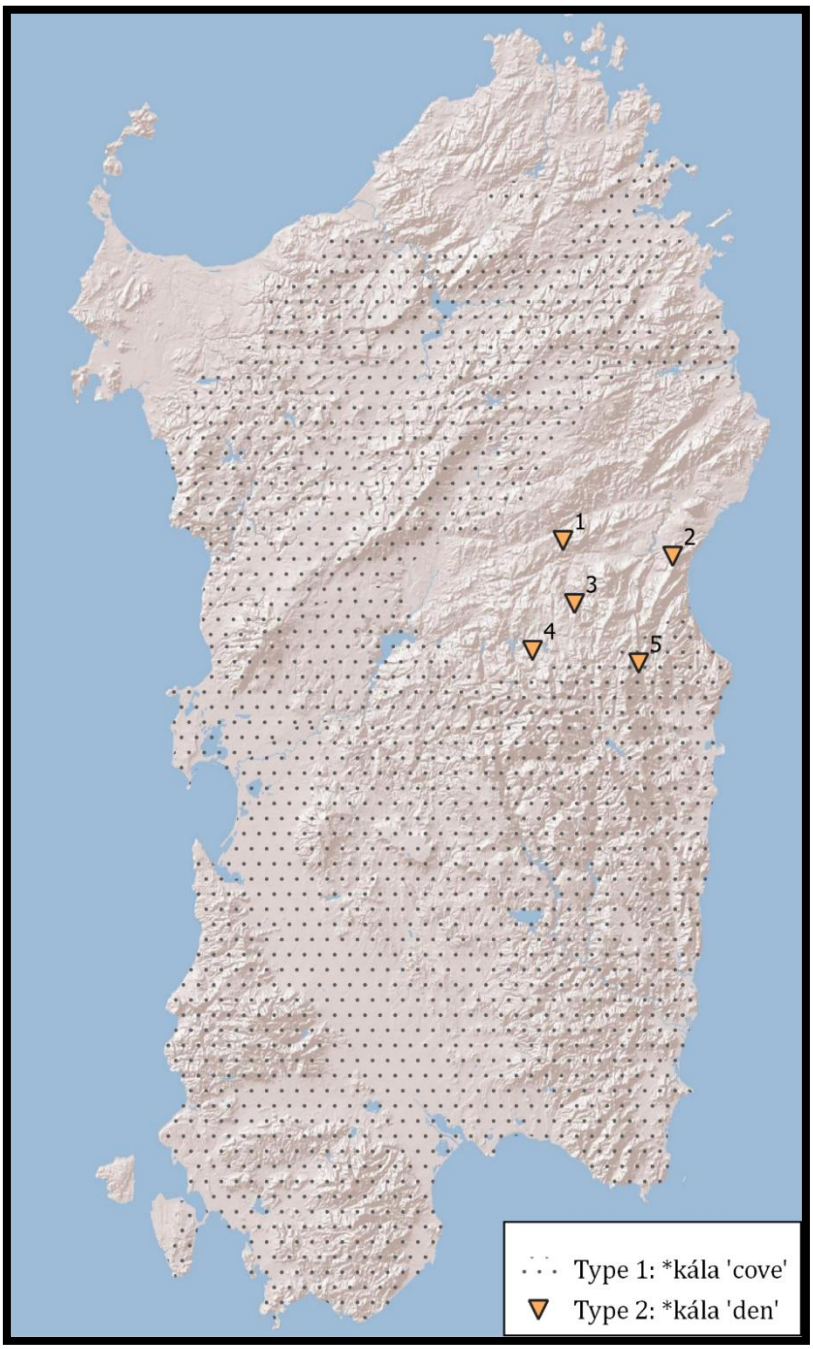

Figure 5.1.3: Animal den

Wagner (1932: 242) believes that Srd. *kála* 'den, anthill' (type 2) is of pre-Roman origin and that it is ultimately related to similar forms elsewhere in the western Mediterranean region, such as Lang. *câlo* 'shelter (from the wind)', Sav. *ϑalę* 'chalet (Alpine shepherd's hut)', Sp., Cat., OProv., Cors., It., Sic. *cala* 'cove, inlet, creek' (and many other forms; cf. FEW II: 50–51; BDLC: s.v. Cors. *cala*), which was eventually introduced in Sardinia a second time in the maritime meaning of type 1 (also Paulis 1987: 476). Whether Srd. *kála* (and *cala* in other languages) can be shown to be in any way related to Srd. *kalánka* 'rock crevice' (§ 5.1.18), as Wagner (1932: 242) argues, depends on the evidence of **-anka* as a suffix, of which there is little (cf. § 5.1.18). For widespread Mediterranean *cala* in the sense 'cove, inlet, creek', a derivation from Lat. *calāre* 'to lower, to let down' << Gr. χαλάω 'id.' cannot be excluded (cf. REW 1487; FEW II: 58–61). For the other meanings, this is more difficult. Note also that Gr. χαλάω was originally borrowed into Sardinian as *faláre* 'to descend' (FEW II: 60; DES I: 449), precluding a derivation of the Sardinian forms from this verb.

5.1.4 Forest

Reconstruction	**Forms**
1: *líttu	DES (II: 33): **líttu** (Nuor., 1: Bitti, 2: Orune, 3: Nuoro, 4: Orani, 6: Fonni) 'large expanse of forested terrain; dense forest'
	Paulis (1987: 509): **líttu** (Log.) 'id.'
2: *líϑϑu	DES (II: 33): **lítϑϑu** (5: Orgosolo) 'id.'

Wagner calls this word 'probably pre-Roman', likely because of the absence of a clear etymology. This word, which is also attested in toponyms, is derived by Pittau (2013: s.v. *littigheddu*) from Lat. *īliceus* 'of the holm-oak, oaken'. This is formally impossible, since Lat. *-ceu* is expected to yield > **-cu̯* > ***-ϑϑu* in the Nuorese dialects. This is attested in Orgosolo (5) *lítϑϑu*, but all other forms show *-tt-*. Moreover, this word is attested as *littu* in the Condaghe di San Pietro di Silki (Bonazzi 1900: sec. 96), which usually represents the reflexes of **-cu̯-* as *-th-* (Maulu 2016: 543–544). This means that *líttu* must be taken as the base form, and that Orgosolo (5) *lítϑϑu* is due to an irregular, secondary shift from *-tt-* to *-tϑϑ-*, which is not uncommon (Wagner 1941a: 119). Wagner (1932: 242) tentatively compares Bq. *lito* 'deep place', but does not mention it anymore in the DES (II: 33). Besides the semantic discrepancy between' Srd. 'forest' and Bq. 'deep place', Bq. *lito* 'deep place' is a variant of *lütho* and may be internally derived from Bq. *lu*(*r*) 'earth' + *it*(*h*)*o* 'drowned, choked' (Bouda 1949: 335; Michelena 1990: 81 fn. 17).

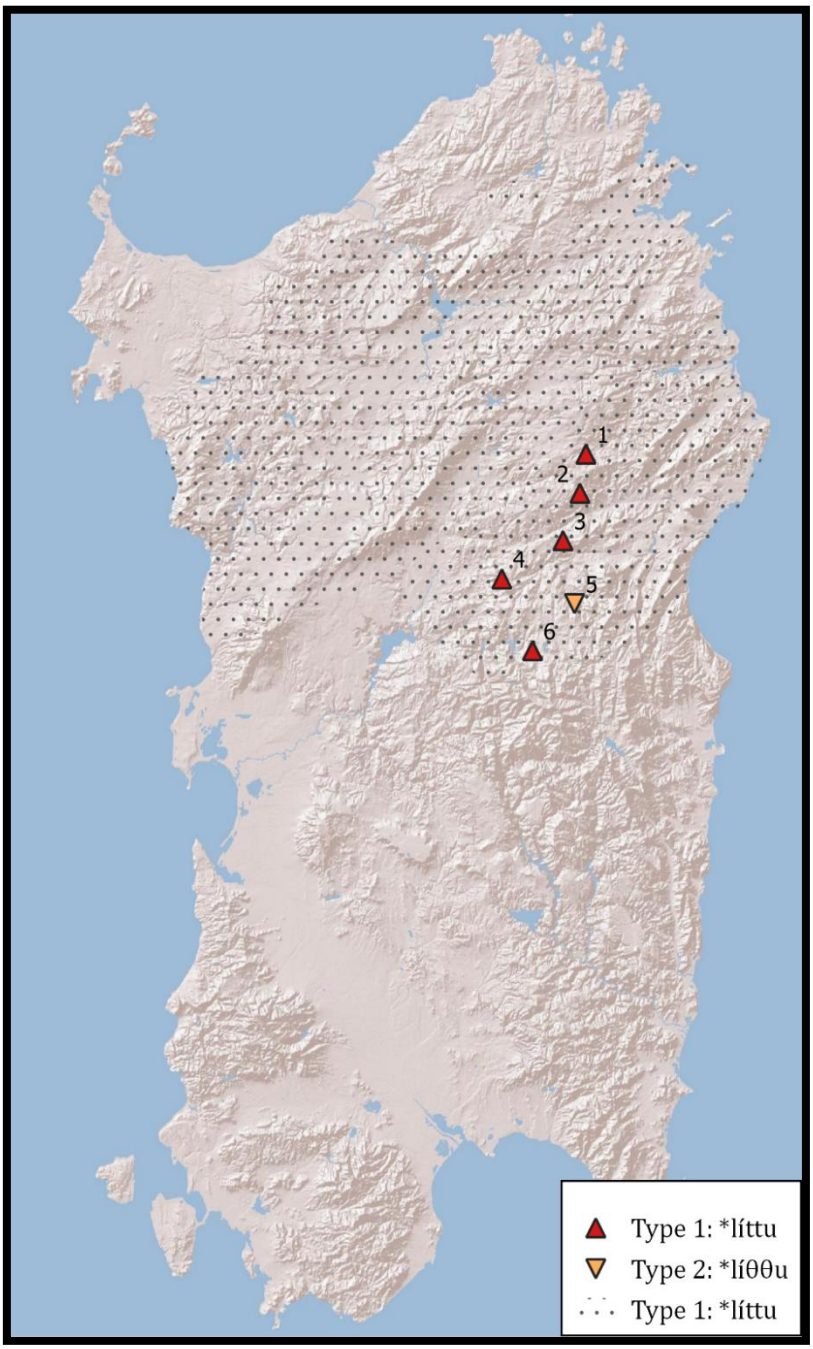

Figure 5.1.4: Forest

Other than the lack of an etymology, there is no real positive evidence for Wagner's attribution of this word to a pre-Roman language. Mastrelli (1970) posits a derivation **ēlictum* 'oak grove' << Lat. *ēlex*/*īlex* 'holm oak',[176] but this use of *-tum* instead of usual *-ētum* to form collectives from plant names is unparalleled in both Sardinian and Latin (Wagner 1952: 88–89; Weiss 2009: 293–294).[177] One wonders whether *líttu* could be the result of the same development we find in Log., Nuor. *fríttu* 'cold' < **frigdu* < Lat. *frīgidus* 'id.' In that case *líttu* might be the regular reflex of **libdu* < Lat. *līvidus* 'bluish, dark, malicious'; cf. the use "*lividissima* […] *vorago*" 'the darkest depth' in Catullus 17. Although this Lat. *līvidus* is already continued with a different phonetic development in Nuorese as *líðu* 'livid, bruised' (< **lividu* < Lat. *lividus*), it should be noted that we also find Nuor. *fríɣiða* 'the sound hot iron makes when it is put in water to cool down' next to *fríttu* 'cold' (DES I: 546). The question is whether such a double phonetic development could also have happened for *līvidus* and whether the semantic development of 'bluish, dark' > 'dense forest' is acceptable.

5.1.5 Highland

Reconstruction	**Forms**
***ĵára**	DES (I: 603): ***ǧára*** (Camp.) 'basaltic and granitic highlands'

This word, which mainly occurs in toponyms referring to rocky plateaus, is proposed in DES (I: 603) to be of pre-Roman origin, but no further are provided. Pittau's (2013: s.v. *Giorrè*) suggestion that it has anything to do with Lat. *glārea* 'gravel' and descendent Tusc. *ghiara*, *iara* (standard It. *ghiaia*) 'id.' must be rejected. The Tuscan form has been borrowed into Sardinian as Camp. *ǧár*(*r*)*a* 'gravel' (DES I: 708), but according to the AIS (417), *ǧárra* with geminate *-rr-* is found in most of Campidanese, Cagliari being the exception with *ǧára*. Pittau suggests that Camp. *ǧára* 'rocky highland' goes back to an Etruscan form that would also have been the origin of Lat. *glārea* 'gravel', but this is formally impossible, as a cluster **gl-* would have yielded ***gr-* in Campidanese. Moreover, the semantic distance between 'rocky highland' and 'gravel' is too great for any plausible connection. Paulis (1994; cf. Putzu 2019: 6) proposes Camp. *ǧára* to be a loan from a Punic cognate of Hebr. יער *yáʿar* 'wood, bush'.[178] Although the Punic form is not attested, this comparison is formally and semantically possible.

[176] *Non vidi*. Cf. Blasco Ferrer (1989: 20).

[177] The toponym *s'elíttu* from Bottidda, cited as evidence for **ēlictum*, could have received its initial *e-* by analogy to Log. *eliɣéðu* 'oak grove' < Lat. *ēlicētum* (cf. Wagner 1952: 88).

[178] I have unfortunately not been able to consult Paulis (1994) myself.

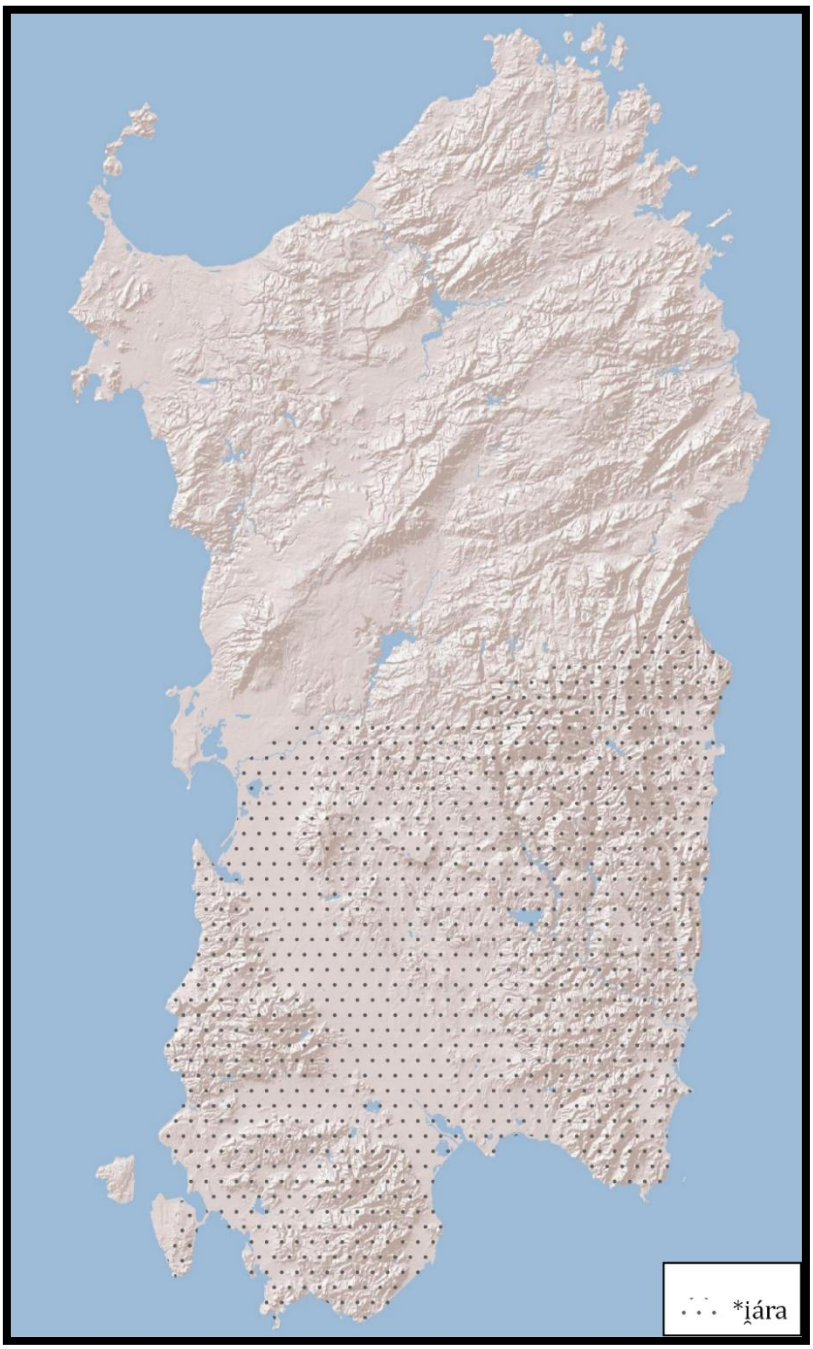

Figure 5.1.5: Highland

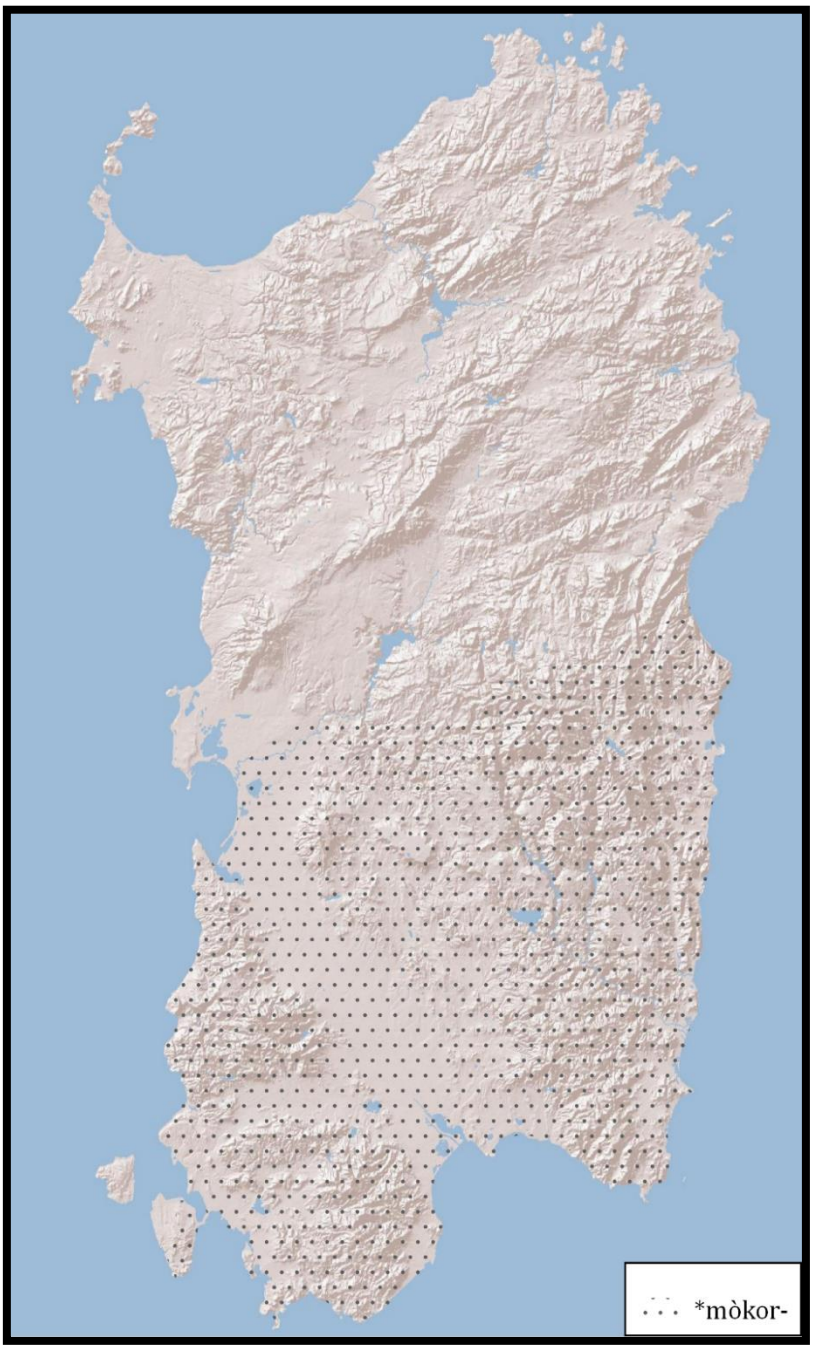

Figure 5.1.6: Hillock

5.1.6 Hillock

Reconstruction	Forms
mòkor-	DES (II: 122): ***mòɣoro*** (Camp.) 'hillock, low hill'
	Paulis (1987: 514): ***mòɣoro, mòɣuru, mòɣora*** (Camp.) 'id.'

Wagner (1932: 227; DES II: 122) provides for this word the rather unspecific distribution: "rural Campidanese, especially in the northern part". It is in any case clear that it is the same as the name of the towns of *Mòɣuru* (It. Mogoro) and *Moɣorèḍḍa* (It. Mogorella). Paulis (1987: 514) adduces the Ogliastran toponyms *mògola*, *mógula*. Wagner (1932: 227–228; DES II: 122) thinks this word is of pre-Roman origin, going back to original **mókor* and remarks that the preservation of word final *-o* in Campidanese points to a pre-Roman origin, because inherited **-o* would have yielded ***-u* (Wagner 1932: 227 fn. 3). This reasoning is flawed, as we would in fact expect an old final *-o* (of whatever origin) to participate in this more recent sound law. Nevertheless, forms with expected *-u* are attested (Puddu 2023: *mògora*, *mògoro*) and word-final *-o* (which is clearly original in this word) can indeed point to a pre-Roman origin, because of its infrequency in the Sardinian nominal system (§ 8.1.5.2). In this case, final *-o* may be of paragogical origin, as Wagner (1932: 227 fn. 3) notes (cf. § 8.4.1). This word does not have a good Latin etymology. Santano Moreno's (2000: 158–160) derivation from Lat. *mucrō* 'sharp point' does not work formally.

This word has been compared to a large number of words of dubious origin in other languages, spanning the width of the Mediterranean. Roughly from west to east, these comparisons include:

- Bq. *mokor* 'tree trunk, crust of bread, lump' (Wagner 1932: 227), *mukurru* 'climax, pinnacle (in the abstract sense)' (Hubschmid 1953: 49). However, Bq. *mokor* 'tree trunk, crust of bread, lump' is analyzed by Michelena (1990: 272) as a formation containing an *m*-prefix and perhaps a variant of *sokor* 'lump', and/or *okor* 'slice' (Michelena 1990: 273). Bq. *mukurru*, *mukuru* 'climax, pinnacle' on the other hand seems to be metathesized from Lat. *cumulus* 'heap, pile', next to Bq. *gonburu*, *bonburu* 'id.' (Michelena 1990: 331). If Michelena's analyses are correct, Bq. *mukur(r)u* needs to be separated from Srd. *mòɣoro* etc., which cannot go back to Lat. *cumulus*. Trask (2008: 276) mentions Gipuzkoan Bq. *mokoru* 'height, summit'. If this is not the same as Bq. *mukurru*, Wagner's (1932: 227) comparison may still stand. Hubschmid (1953: 50) also compares the Sardinian forms to Bq. *moko* 'beak, tip, peak'.

- Old Cat. *mocorró*, *moguró*, *muguró*, Cat. *mogró*, *mugró* 'nipple', *mugaró* (Barcelona) 'pile of sheaves made on the land'; Sp. *mugrón* 'layering (i.e. a branch of a plant taking root after touching the soil') (Hubschmid 1953: 49–50). These comparisons are all rejected by Corominas (DCEC IV: 184), who derives the Catalan words for 'nipple' from Lat. *mucrō*, *-ōnis* 'sharp point' and Sp. *mugrón* 'layering' from Lat. *mergere* 'to plunge, immerge'. However, Corominas is more favorable towards the connection of Sp. *mogote* 'isolated cone-shaped hillock', which he speculates could be from an unattested pre-Roman word **mokoti*, which looks like a productive derivation from Bq. *moko* 'beak, tip, peak'.
- Prov. *maguelo*, Lang. *maghêlo*, *maghiêlo* 'steep hill, steep place', Gasc. (Barèges) *maguèro* 'stony hedge,[179] bush in the middle of cultivated land', (Gèdres) *maguèro* 'little wall of dry stones' (FEW XXI: 15b). According to the FEW, all these forms presuppose a pre-Roman form **maguella*.
- Alb. *magulë* 'small pile, heap', Rom. *măgură* 'mountain, hill', MGr. μαγούλα, PSl. **mogyla* 'burial mound' (Wagner 1932: 228; Ciorănescu 1958: s.v. *măgură*). According to Orel (1998: 109), Alb. *magulë* 'heap' is the result of a metathesis from *gamule*, *gamulë* 'id.', which was borrowed from PSl. **gomola* 'lump'. This would subsequently have been borrowed into Romanian as *măgură* 'hill'. Ciorănescu (1958: s.v. *măgură*) notes however that the change **-l-* > *-r-* implies that this word cannot be a recent loan in Romanian, and hypothesizes that at least Rom. *măgură* and Gr. μάγουλα must be independent loans from an earlier source, but does not draw any conclusions about the nature of their relation to Alb. *magulë* or PSl. **mogyla* 'burial mound'. PSl. **mogyla* is of unknown origin (Vasmer 1955: 143–144).

In light of the fact that several of the proposed connections can be explained internally in the respective languages, it is wise to include in the comparison only those forms that are close to Srd. *mòɣoro* 'hillock' in form and meaning, without having to resort to hypothetical derivational mechanisms in unattested languages. This excludes Hubschmid's (1953: 50) connection to Bq. *moko* 'beak, tip, peak', for which we would need to assume suffixation of some *-r-*element in all other forms. Even after the rejection of many forms on the Iberian Peninsula by Michelena (1990) and the DCEC however, we are left with a large series of words of similar form and meaning to Srd. *mòɣoro* etc. < **mòkor*, in southern

[179] "[*H*]*aie caillouteux*" (FEW XXI: 15b).

France and the Balkan peninsula, and maybe in Bq. *mokoru* if it is not from Lat. *cumulus*.

5.1.7 Humid place

Reconstruction	Forms
1: *_orgòs-_	DES (II: 192): ***orgòsa*** (1: Orgosolo) 'humid, marshy place',[180] ***orgòše*** (2: Fonni) "not well-defined (small forest?)"
	Wagner (1997: 266): ***orgòsa*** (1: Orgosolo) 'humid terrain'
	Paulis (1987: 520): ***orgòsa*** (Barb.) 'humid, marshy place'
	Blasco Ferrer (1988: 175): ***orɣòsa, uryúsa*** (2: Fonni) 'meander', (3: Urzulei, 4: Triei, 5: Baunei) 'open-air source, water pool, place where water flows and where herds are taken to drink'
2: *_org-_	Blasco Ferrer (1993: 180; 2011: 79): ***òrga, òrge*** (Alta Ogliastra reg., Baronie reg.) 'open-air source, water pool, place where water flows and where herds are taken to drink'

This word is formally close to the toponym *Orgòsolo* (It. Orgosolo), a town in which this word is actually attested. Wagner (1997: 267) and Blasco Ferrer (1988: 175; 1993: 180) point out several other Sardinian toponyms containing the element *orgos-* or *org-*, many of which refer to similar watery places. This word's lack of a Latin etymology, as well as its phonotactic structure and its apparent occurrence in the island's toponymy point toward a pre-Roman origin, as proposed by Wagner (1997). Wagner's (1997: 267–268) tentative connection with Gr. ὀργάς, -άδος 'humid, fertile (of earth, land)' must be rejected, as this eventually goes back to PIE **uerǵ-* 'swell of juice, strength, anger' (Beekes 2010: 1097). On the basis of the forms of type 2, and of toponyms like *untána òrɣa* (Dorgali) and *báðu oryé* (Urzulei) etc. (Blasco Ferrer 1988: 175), the ending *-òsa* can be interpreted as a suffix (Wagner 1997: 266). Although it is homophonous with the reflex of Lat. *-ōsus, -a* (Wagner 1952: 83–84), the non-inherited ending *-olo* in *Orgòsolo* speaks against an a Latin origin for *-òsa* in the nouns of type 1. Blasco Ferrer's (1993: 181) proposal to interpret *-òsa* as the reflex of a pre-Indo-European root **ausa* is not convincing.[181]

[180] Actually as *orgòśa* in DES, but in light of its rendering as *orgòsa* in Wagner (1997: 266), Paulis (1987: 520) and Puddu (2023: s.v. *orgòsa*), this is probably a typographical error.
[181] Cf. § 3.1.6.

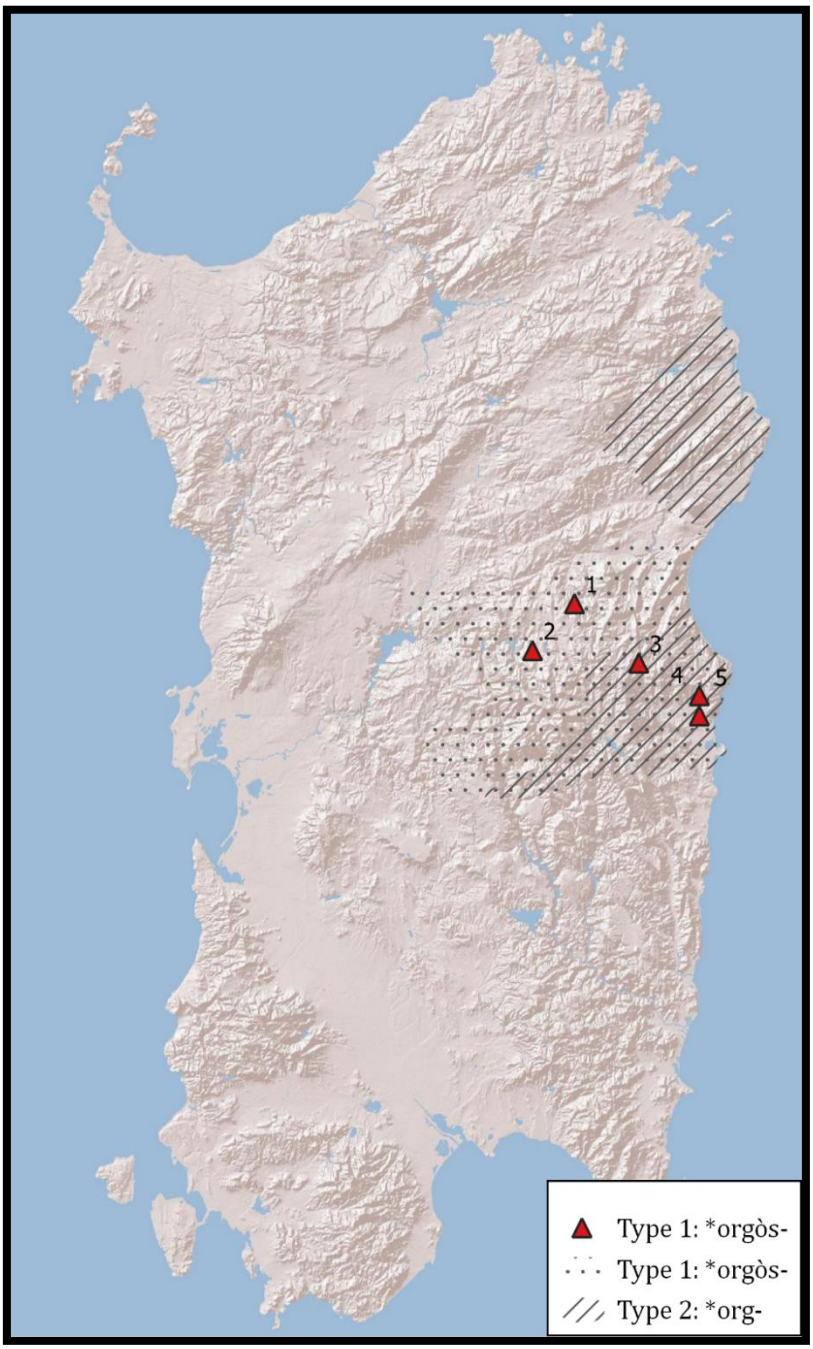

Figure 5.1.7: Humid place

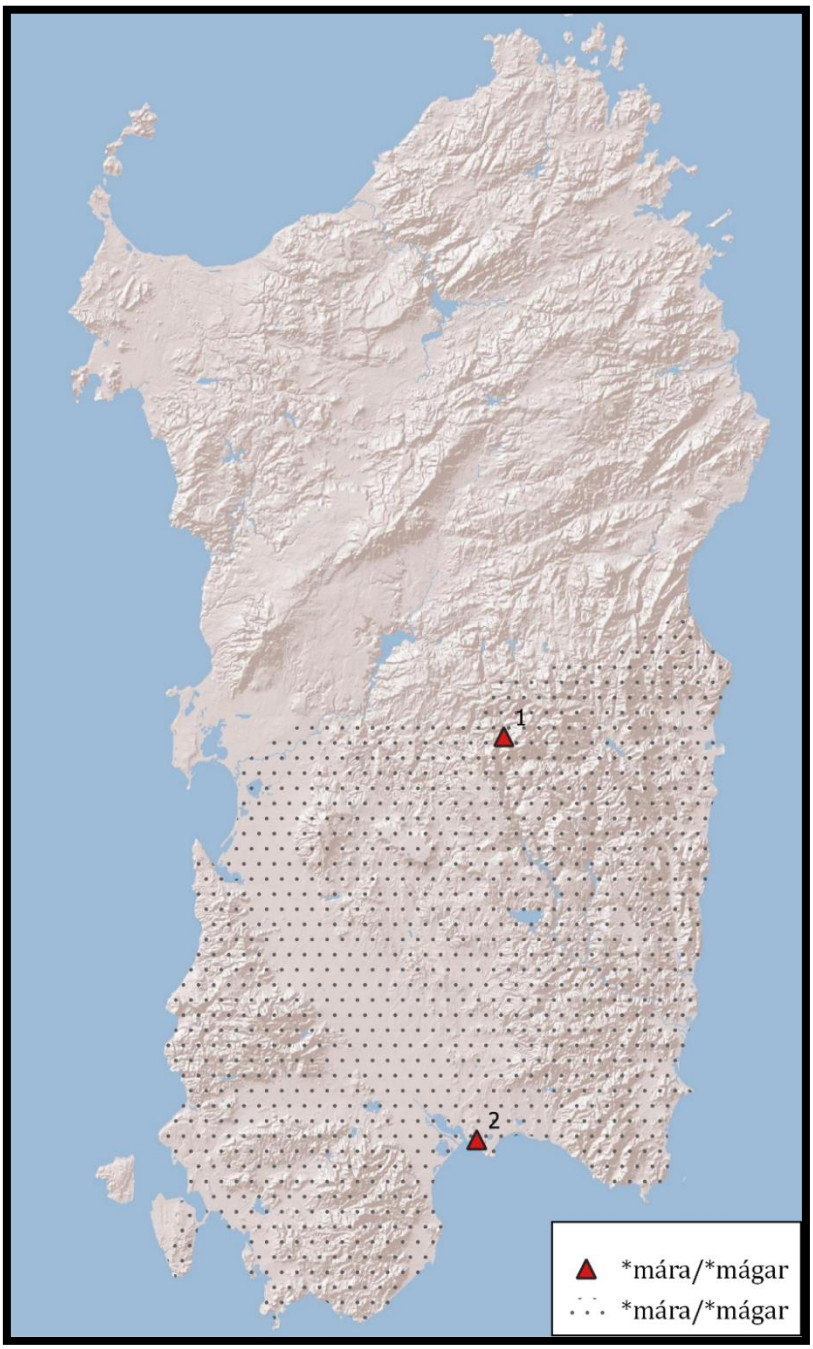

Figure 5.1.8: Marsh

5.1.8 Marsh

Reconstruction	**Forms**
****mára*/**mágar***	DES (II: 71): ***mára*** (Camp.) 'marsh, swamp, bog', (2: Cagliari) 'sewer, drain'; deriv.: ***marètta*** (1: Belvì) 'permanent puddle of water in a cave'

This word, which refers to various kinds of watery places in the Campidanese dialects, is attested in toponyms, such as *sa mara de Urzulei*, *sa mara 'e Usíni*, *sa mara 'e Padria*, as well as the town names *Mara Arborei* (It. Villamar) and *Mara* (It. Maracalagonis)(DES II: 71). All of these fall within the Campidanese dialect area, with the exception of the toponym near Padria, which is significantly farther to the north. The name of the town *Mara* is attested as *Magar* in documents from the 12th century (Wolf 2011: 612). It is formally possible that lexical *mára* 'marsh, swamp, bog' goes back to **magar* too, but it is not certain that the name *Mara* contains the same lexeme.

This word has had a long history of discussion. Terracini (1927: 144) first noted the similarity between the Sardinian and various North Italian forms, especially regarding their denotation of swamp-like places etc. He, and Wagner (1932: 241) and Hubschmid (1953: 67) with him, convincingly argue that the Sardinian forms cannot be from Latin, but must rather be from a pre-Roman source.[182] The North Italian forms adduced by Hubschmid (1953: 63–65) include Lad. *mara* 'place subjected to landslides', *lamár*, [*la*] *mar* 'swampy place; landslide caused by a spring of water',[183] Lad. *marena* 'landslide caused by a spring of water', Ven. *marecia* 'marshy terrain', It. *marazzo* 'puddle, marsh'. Wagner (1932: 241) further compares Laz. *marana* (Velletri) 'artificial waterway' and Cors. *mara* 'deviation of water to irrigate gardens', which resemble the meaning 'sewer, drain' found in Cagliari (2). Also similar in form and meaning is Gr. ἀμάρα 'trench, channel' (Alessio 1944: 117–118; Hubschmid 1958: 223), which is of uncertain origin and may or may not be related to Alb. *amë* 'river bed'.

Because of the obvious formal and semantic similarity between Srd. *mára* 'marsh, swamp, bog' and Lat. *mare* 'sea', Hubschmid (1953: 67) proposes that the Sardinian word is of pre-Roman origin, but originally somehow related to the Latin word and its Indo-European cognates (cf. de Vaan 2008: 365). This would

[182] As opposed to REW (5349), where many of these forms are subsumed under Lat. *mare* 'sea'.

[183] Hubschmid (1953: 64) notes that forms like Lad. *la mara* could in theory have been secondarily formed from *la mar* (which is also attested) < Lat. *mare* 'sea', but not in Srd. and Cors., where *máre* 'sea' is preserved as such.

suggest the presence of an unattested pre-Roman Indo-European language in Sardinia, as Lat. *mare* 'sea' and its cognates are generally taken to be of Indo-European origin (de Vaan 2008: 365). Hubschmid (1953: 120; 1958: 223) however proposes a rather intricate scenario by which Indo-European elements entered some non-Indo-European languages in the Mediterranean region, whose speakers then brought them to Sardinia during their hypothetical migrations. This is impossible to verify on the basis of the available data.

The large amount of similar forms in various languages confuses the picture of the origin of this word. It has on the one hand been compared to Indo-European words for 'sea, lake', on the other hand to words for 'landslide' and 'marsh' in northern Italy, and to words for some kind of 'artificial waterway' in Corsica, central Italy and Greece. It is difficult to uphold an Indo-European origin of Srd. *mára*, as proposed by Hubschmid (1953; 1958); chance similarity with forms like Lat. *mare* 'sea' < PIE **mori-* is more likely. The comparisons with Gr. ἀμάρα 'trench, channel', Cors. *mara* 'deviation of water to irrigate gardens', as well as the various northern Italian words referring to marshes etc. are formally and semantically more attractive. From the Sardinian evidence, it is clear that semantic shifts between 'marsh, swamp' on the one hand and 'sewer, drain' (or other kinds of artificial waterways) are possible, but if Cors. *mara*, Gr. ἀμάρα and other forms are indeed somehow related to the Sardinian forms, it is difficult to establish whether the word originally referred to swampy areas or artificial waterways, or whether the polysemy is old.

Crucially, all proposed connections only work if the Sardinian lexical forms go back to **mára*. If, on the other hand, the lexical forms are cognate to the place name *Mara* (It. Maracalagonis), attested in the 12th century as *Magar* (Wolf 2011: 612), none of the putative cognates can be upheld.

5.1.9 Marshy valley

Reconstruction	**Forms**
****bèka***	DES (I: 191-192): ***bèɣa*** (Camp.) 'marshy valley'
	AIS (428): **[*un*] *èɣa*** (1: Sant'Antioco) 'id.'

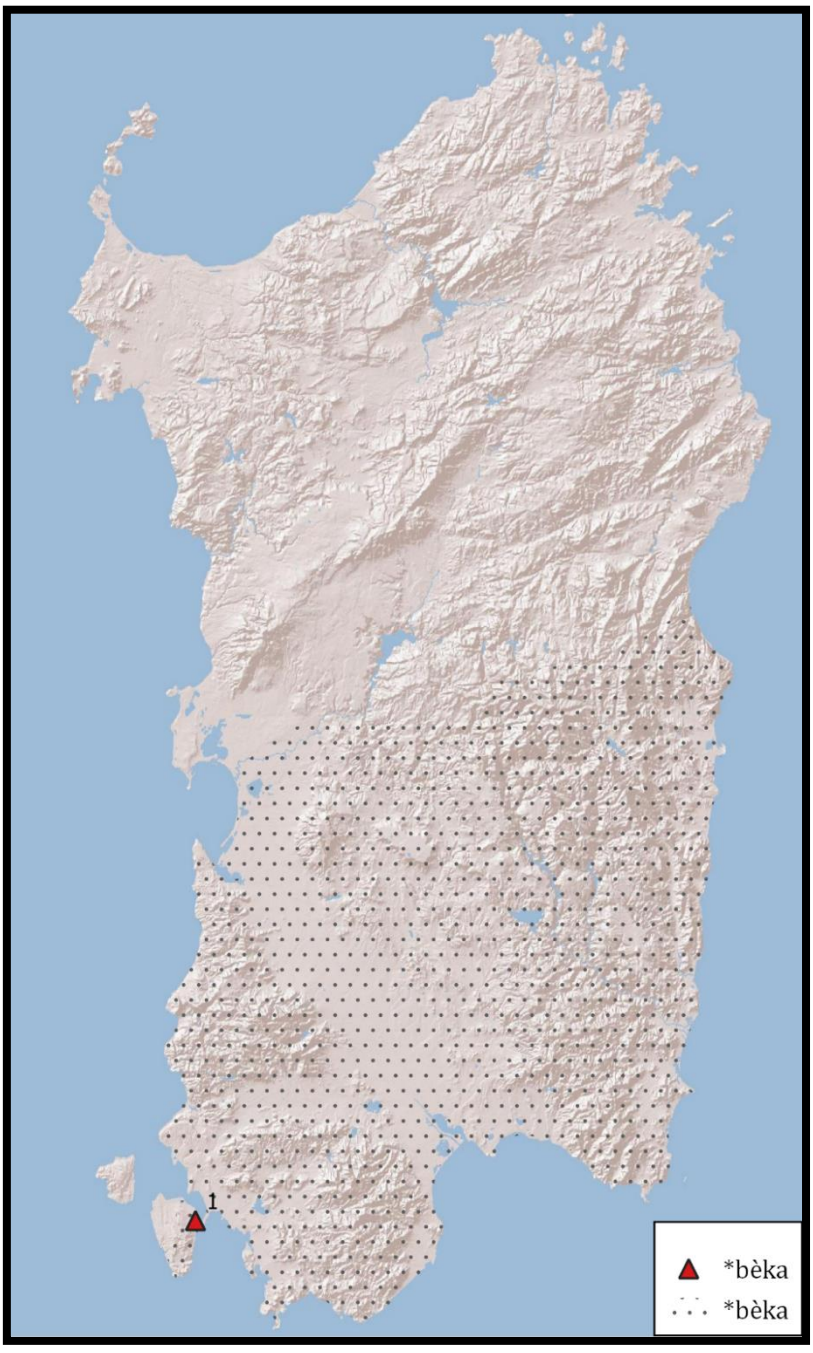

Figure 5.1.9: Marshy valley

Wagner (1997: 254; DES I: 192) recognizes the similarity between this Sardinian word on the one hand, and Sp. *vega* 'meadow, fertile plain' on the other.[184] Yet, he argues that the Sardinian word cannot be the result of a Spanish loan, as *bega* is attested in the *Carte Volgari* (Solmi 1905: sec. 2: 2). This work has been identified as an authentic document from the 12th century (Solmi 1905: 273, 283), which would be several centuries too early for linguistic influence from Castilian on Sardinian (Wagner 1907a: 408). This leads Wagner (1997: 254) to propose a pre-Roman origin for Camp. *bèɣa*, shared with Sp. *vega*, from hypothetical **baika*, presumably related to Bq. *ibai* 'river'. This view is shared by Hubschmid (1953: 38), Corominas (DCEC V: 753-54), and Paulis (2008: 37), but Trask (1997: 420–421) argues that a putative pre-Roman form **baika* is not explainable from a Basque perspective.

It is important to note, however, that Paulis (1997: 133–139) has argued against an early date for the attested version of the text in which Srd. *bega* is first attested. In fact, Paulis (1997: 135–137) demonstrates the presence of two likely Catalanisms: *implassandullu* << Cat. *embraçar* 'to embrace' and *cumenzat*, *cumenzzat*, *cumençat* etc. << Cat. *comensar* 'to begin', as well as the Catalan spelling of [k] as *-qu-* in *pasquiri*. We would not expect to see this kind of kind loans before the presence of the Aragonese on Sardinia, some centuries later. For scroll XI therefore, containing the form *implassandullu*, Paulis (1997: 135) proposes an origin somewhere in the 15th century at the earliest, based on the hypercorrection from **imprassandullu*. The rhotacism of post-consonantal *-l-* that is the cause of this hypercorrection starts showing up sporadically from the 15th century onward (Paulis 1997: 135). The word currently under discussion, *bega*, is attested in scroll II, which contains the Catalanism *cumenzzat* (Paulis 1997: 137a). While an exact dating of this piece is absent, it unmistakably shows that certain Ibero-Romance elements had already penetrated Campidanese Sardinian. This raises the question whether Camp. *bèɣa* is in fact not a loan too (cf. Blasco Ferrer 2010a: 55–56). Corominas (DCEC V: 753-54) notes that *vega* is rare in Aragonese and absent from modern Catalan. As there is no evidence in the *Carte Volgari* for Castilian Spanish linguistic influence on Sardinian, this would speak against a recently borrowed origin of Camp. *bèɣa*. Nevertheless, DCEC (V: 755 fn. 3) does mention various sporadic attestations of *vega* in Catalan and Valencian linguistic contexts, including on the Balearic islands, which are explained as loans from Castilian.

[184] For more Ibero-Romance cognates of Sp. *vega*, cf. DCEC (V: 753-56).

The pre-Roman origin of Camp. *bega*, *bèya* has been based on its attestation in the *Carte Volgari*. While an Ibero-Romance loan is still problematic because of this word's apparent non-existence in Catalan, and the absence of evidence for Castilian elements in the *Carte Volgari*, Wagner's argument has become significantly weaker. Whether or not this word may be considered to be of shared pre-Roman origin with Sp. *vega*, depends on the origin date of scroll II of the *Carte Volgari* and of the original existence of *vega* in Catalan. If it truly is of pre-Roman origin, Corominas' (DCEC V: 755b fn. 4) objection that the **-k-* in hypothetical pre-Rom. **baika* should have been preserved is unjustified. In the *Carte Volgari*, there is sufficient evidence showing that the sonorization of post-vocalic voiceless stops had already taken place, e.g. *fegit* < Lat. *fecit*, *iugi/juigi* < Lat. *iudicem*, *logu* < Lat. *locus* etc. The development of **-ai-* (as in **baika*) in Campidanese as *-e-* is attested too; e.g. Camp. *méstu* 'master' < Old Camp. *maistru* < Lat. *magistrum*; Camp. *arrèga* 'radish' < **radica* (Wagner 1941a: 44). However, these forms often coexist with variants in which the hiatus has been preserved, e.g. Camp. *maístu*, *arraíɣa* (ibid.). It is implausible that the contraction of *-ai-* > *-e-* would have been completed in this specific word already in Old Campidanese, as it is not even completed in the modern dialects. Corominas (DCEC V: 755 fn. 4) suggests that pre-Roman **baika* could have entered Sardinian Latin as **baeca*, regularly yielding Srd. *bèɣa*, which would solve this issue. In summary, a shared origin of Srd. *bèɣa* and Sp. *vega* in a pre-Roman form **baika* is formally possible. However, an early Ibero-Romance loan into Sardinian is not out of the question.

5.1.10 Pit

Reconstruction	**Forms**
****kèi̯a***	DES (I: 326): ***kèya, kèa*** (Log., Nuor.), ***čèa*** (Camp.) 'valley or plain in between mountains; ditch, pit (specifically the big pit in which charcoal burners stack wood in order to carbonize it)'; [***sa***] ***ɣèa*** [***dessu attíle***] (1: Osilo), [***sa***] ***χèa*** [***essu aθíle***] (3: Dorgali) 'the hollow of the neck"[185]
	Böhne (1950: 77): [***sa***] ***žèa*** 'plain'
	Hubschmid (1953: 38): ***chea*** (2: Bosa), ***cheia*** (4: Desulo) 'pit, grave'

[185] All these forms are also listed in Wagner (1930: 70–71; 1997: 269–270).

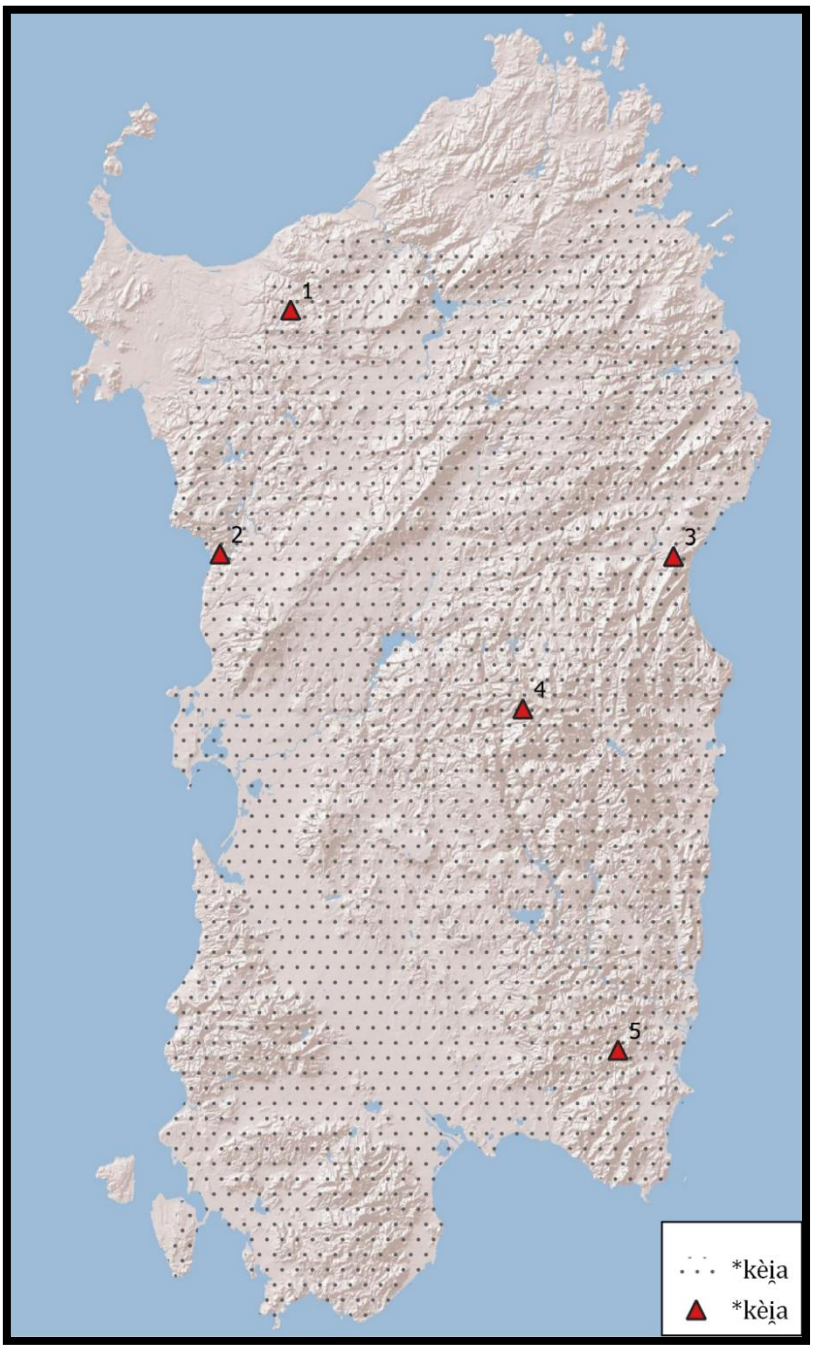

Figure 5.1.10: Pit

Besides the forms listed above, DES (I: 326) also notes the presence of this word in Old Sardinian texts as part of toponyms, and once as a noun in the syntagm "*sa keia dessu lauore de Gosantine Locco*" 'the grain *keia* of Gosantine Locco', found in the Condaghe di San Pietro di Silki (Bonazzi 1900: sec. 241), where it appears to refer to some kind of pit to store grain (DES I: 326). For this word, which appears to go back to **kèi̯a*, the notion of 'pit, hollow' may explain the various other attested meanings.[186] Wagner (1930: 71 fn. 2; 1997: 270; DES I: 326–27) rejects several previously proposed etymologies and argues for a pre-Roman origin.

Wagner's position is in agreement with the views of Hubschmid (1953: 38–40) and Corominas (DCEC II: 58–59), who compare a range of Aragonese and Catalan forms, including Arag. *cea*, *cija* 'silo; underground cavity used for grain storage and to imprison slaves'; Cat. *cía* (western Ribagorza) 'chasm', *ϑía*, *síya*, *sía* (prov. Huesca) 'grain silo' (DCEC II: 58-59); Cat. *sitja* 'pit covered with earth for storing grain', *sitges* (Mallorca) 'silos, hewn out in the rock inside houses', OProv. *setgia* (Narbonne), Lang. *siejo* (Montréal-de-l'Aude) 'grain silo' (Hubschmid 1953: 38–39). All of these forms provide perfect formal and semantic matches with the Sardinian ones, allowing us to reconstruct pre-Roman **kĕi̯a*,[187] which appears to have referred to some kind of pit used for storage purposes. DECLC (VII: 959) further compares various suffixed forms: Elban *čiálę* 'shelter under a protruding rock', Dauphinois *sialet* 'chasm; land collapse in high plateaus' (cf. FEW XXI: 18), and Friul. *céole, cévole* 'chasm'.

DCEC (II: 59) and DECLC (VII: 959) see a link with Gr. Hes. καίατα 'pits (split by earthquakes)', Gr. καιάδᾱς 'pit or cavern in Sparta, into which convicted persons were thrown' etc. (cf. Beekes 2010: 615; Hesychius Alexandrinus 2020: 499). These Greek attestations are semantically close to the Romance forms, but there are some formal problems. One is the fact that Greek intervocalic -ι- cannot go back to **-i̯-*, which would have been lost. It could probably go back to **-i̯u̯-* or **-u̯i̯-* (cf. Beekes 2010: 615), but this is incompatible with the Romance forms. Another is the presence of -ατ- in the Greek form, which would have to be some kind of suffix. Beekes (2010: 615) posits a pre-Greek origin for καίατα, καιάδᾱς. For the wide geographical distribution of this potential pre-Roman word, see § 11.2.1.

[186] Cf. Puddu (2023: s.v. *cèa*) for more examples and uses of this word, including the meaning 'grave'.

[187] On the length of the vowel, see DECLC (VII: 957).

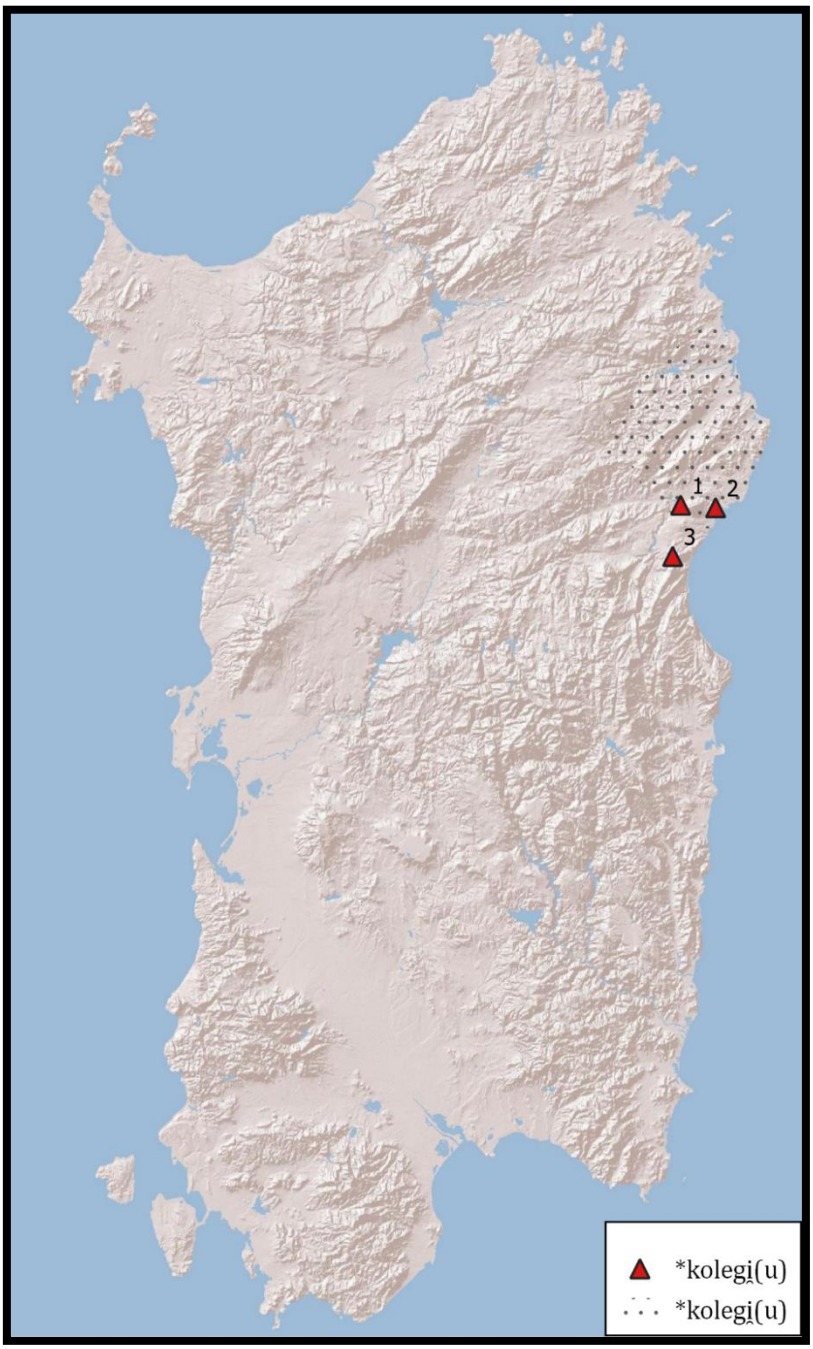

Figure 5.1.11: Plateau

5.1.11 Plateau

Reconstruction	Forms
kolegi̯(u)	DES (II: 608): ***golléi̯, gulléi̯*** (1: Galtellì) 'plateau'
	Spano (1872: 58): ***Guillèi***, ***Gullèi***, ***Gulai***, ***Gullai*** (2: Orosei) 'name of a mountain near Orosei'
	Blasco Ferrer (1989: 61–62): ***golléi*** (Baronie reg.)

Wagner (DES II: 608–609) regards this word as "certainly pre-Roman", probably on the basis of its irregular phonological alternations and its occurrence in local toponymy. The ending *-éi̯* (and *-ai̯*) is frequently found in toponyms, especially in the eastern part of Sardinia; e.g. Orosei, Urzulei, Baunei, Lanusei, Ulassai etc. (cf. Terracini 1927: 139; Wagner 1997: 265). The geminate *-ll-* in *golléi̯*, *gulléi̯* etc. cannot be original, as this would have yielded **-ḍḍ- (Wagner 1941a: 195). However, Blasco Ferrer (1989: 61–62) has convincingly shown that this noun can be from **colēgium* << Lat. *collēgium*, with the same irregular degemination of *-ll-* as in the verb *kollíre* 'to collect' < Lat. *colligere*.

5.1.12 Plateau, characterized by jagged rock outcroppings

Reconstruction	Forms
tákk-	DES (II: 458): ***tákku*** (northern Camp.,[188] Barb.) 'name for small plateaus consisting of jagged dolomitic formations'
	Blasco Ferrer (1988: 176): ***tákka*** (1: Baunei), ***tákka, tákku*** (Ogl., Barb. di Seulo) 'piece of sandy or calcareous terrain; square-shaped rock; plateau with jagged rock formations'

This word refers to a type of landscape typical for certain parts of Sardinia, and is as such on the border between being a proper noun and being a toponym.[189] DES (II: 458) considers it to be of pre-Roman origin. This is possible, but there is no evidence for it besides its lack of an etymology.

[188] Wagner's (DES II: 458) attribution of this word to "northern Campidanese" is rather non-specific and is probably meant to denote those regions where this type of landscape occurs. In Figure 5.1.12, all of Campidanese is indicated.

[189] See Spano (1872: 112) for instances of toponyms containing *tákku*.

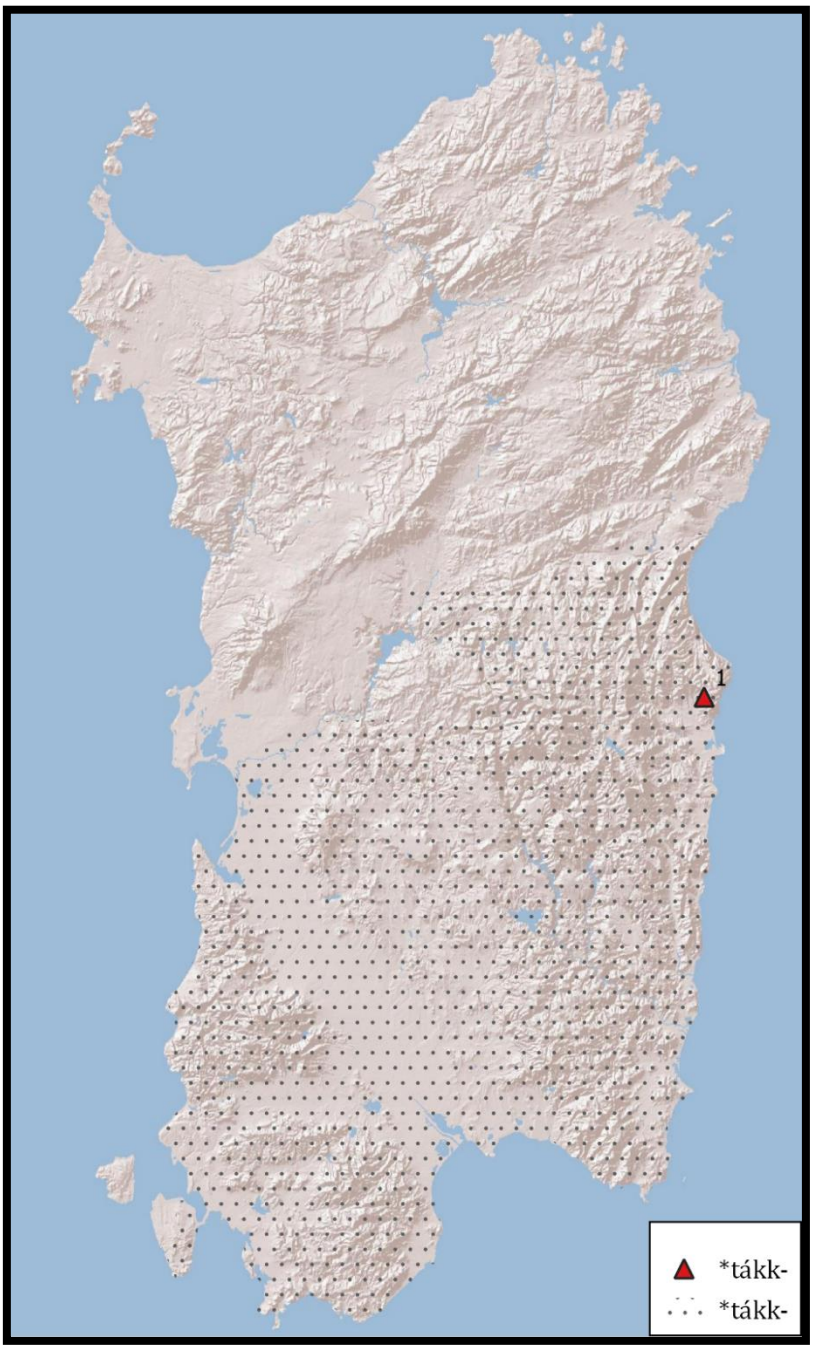

Figure 5.1.12: Plateau, characterized by jagged rock outcroppings

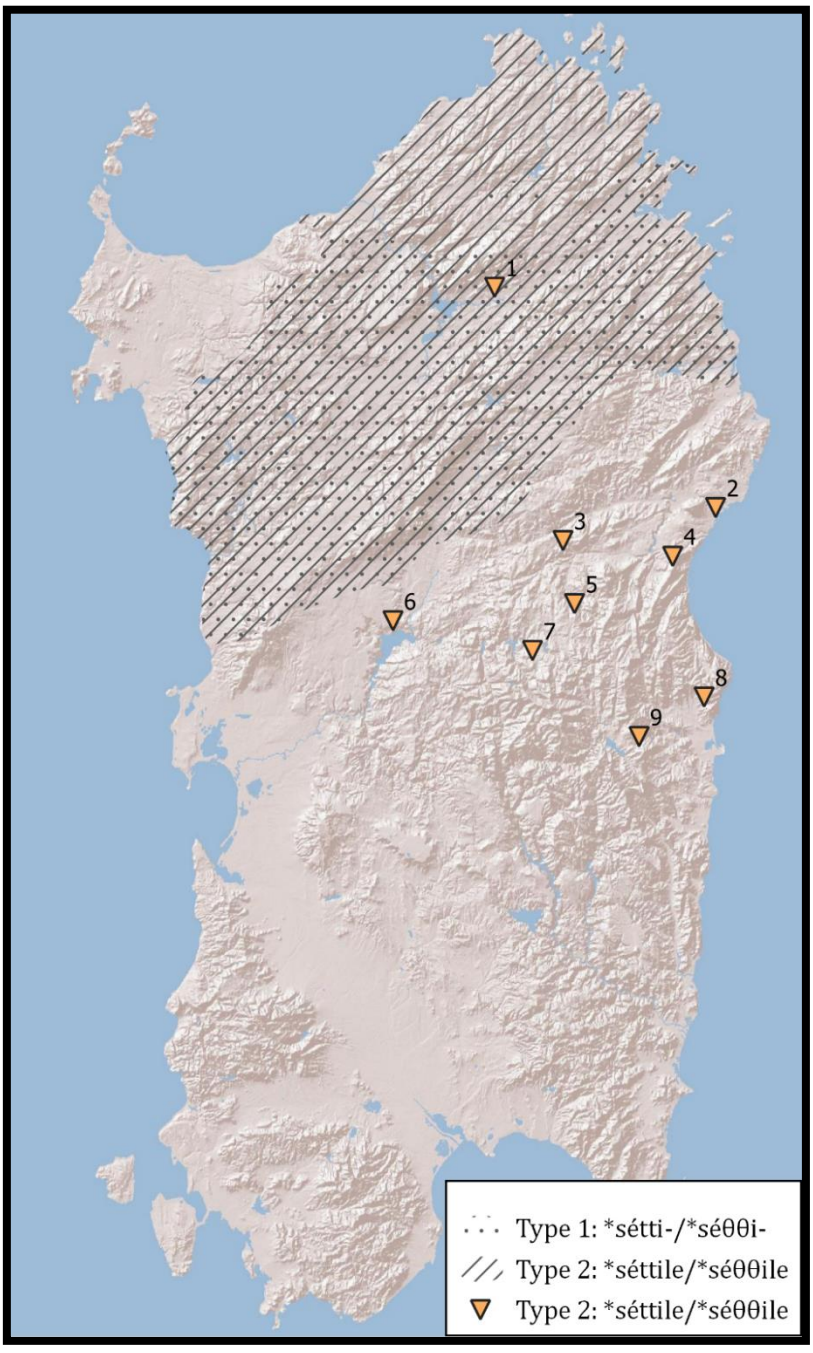

Figure 5.1.13: Plateau, flat terrain

5.1.13 Plateau, flat terrain

ReconstructionForms	
1: *sétti-/*séϑϑi-	DES (II: 413): 1: ***séttị̊a, séttị̊u*** (Log.) 'small elevation, small protrusion'; deriv.: ***settị̊áre*** (Log.) 'to watch, observe from above'
2: *séttile/ *séϑϑile	DES (II: 413): ***séttile*** (Log.) 'small depression in the terrain on a plateau', (7: Fonni) 'hillock', (8: Baunei) 'valley', (9: Villagrande Strisaili) 'plains'; deriv.: ***assettiliɣáðu*** (Log.) 'small, flattened depression in the terrain on a plateau'
	Pittau (2013: s.v. *séttiles*): ***séttile, séthile*** (Log.), (Gallur.) 'plateau on top of a hill, (1: Berchidda) 'small depression on a plateau', (2: Orosei) 'rolling plateau, wide hillock', (3: Nuoro, 9: Villagrane Strisaili) 'plains', (4: Dorgali) 'small plateau in a steep area', (5: Orgosolo) 'flat and rolling terrain in mountainous areas', (7: Fonni) 'hillock', (8: Baunei) 'valley'; ***séttile*** [***'e linna***] (6: Sedilo) 'expanse of wood'

All of these forms refer in one way or another to a place with relatively flat terrain in a mountainous area. Wagner (DES II: 413) considers a derivation from *séttiðu*, *séttị̊u*, the past participle of *sèðere* 'to sit', but believes this to be morphologically and semantically difficult. Because of the lack of an etymology he keeps open the possibility that these forms could be of pre-Roman origin. Pittau (2013: s.v. *séttiles*) cites Nuor. *séthile* next to *séttile*, but the existence of the former is confirmed neither by DES (II: 413), nor by Puddu (2023: s.v. *sétile*, *sétili*). Pittau derives this word from Lat. *sectilis* 'cut, cleft, divided', through the idea of a 'section of terrain'. While this is formally possible for *séttile*, its meanings seem too specific to go back to 'section of terrain'. Moreover, Lat. *sectilis* can formally explain neither Log. *séttị̊a*, *séttị̊u*, nor Nuor. *séthile*. Perhaps Wagner is right in suggesting a pre-Roman origin, but there is no positive evidence.

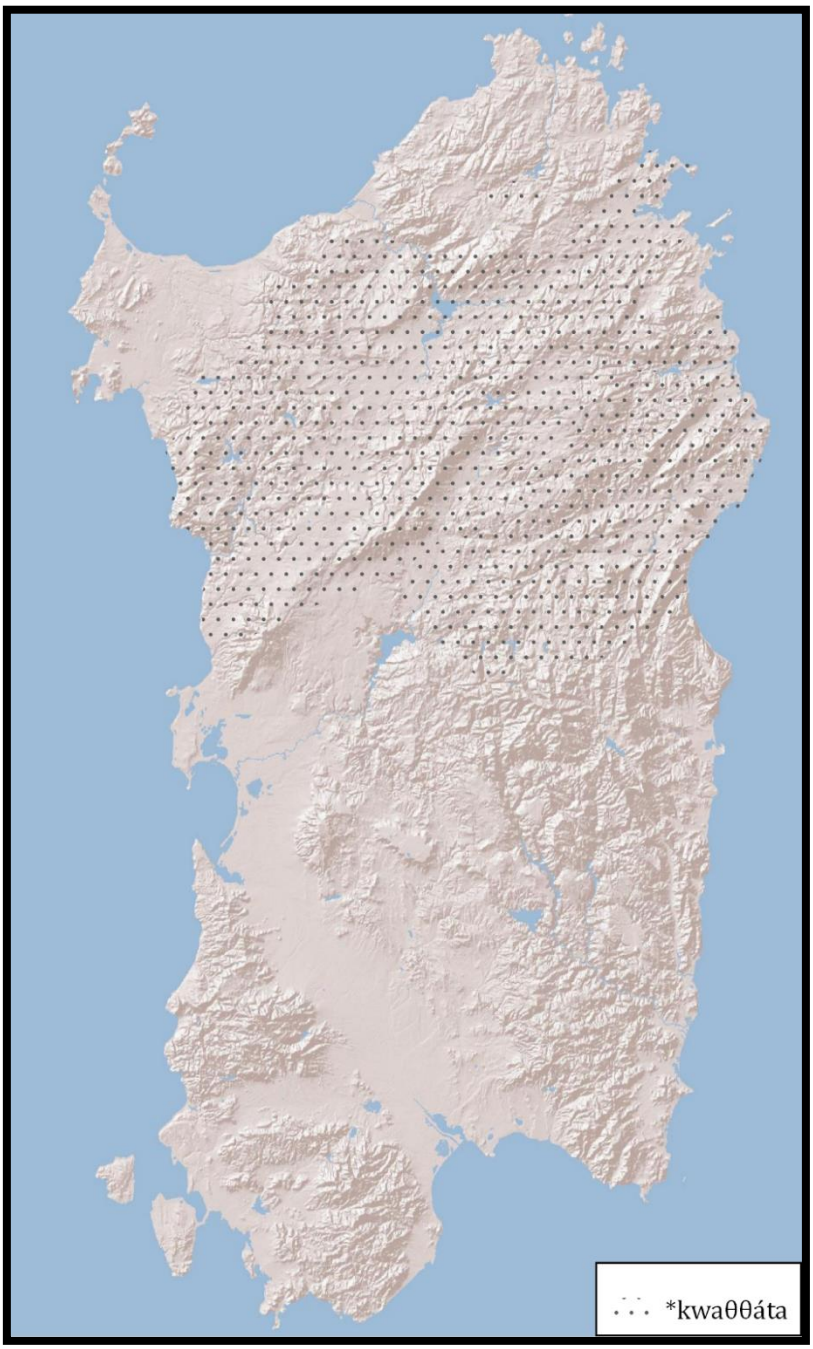

Figure 5.1.14: Small plateau

5.1.14 Plateau, small

Reconstruction	Forms
***kwaʢʢáta**	DES (II: 236): ***pattáða*** (NLog.) 'small plateau; elevated place from where one has a wide view'[190]
	Paulis (1987: 522): ***pattáda*** (Log.), ***pattháta*** (Nuor.) 'id.'

Wagner does not provide an etymology besides noting: "perhaps pre-Roman?". He does mention various toponyms of the same shape, found in the same area. Paulis (1987: 522) proposes it to go back to **coactiata*, derived from **coactiare* 'to press together'. Although this presupposes a relatively abstract shift in meaning, parallels can be found in Gallo-Romance, e.g. dial. Fr. *casse* 'clod of earth' << **coactiare*, Aostan *katę* 'id.' << **coactitare* (FEW II: 806a, 813b). Pittau (2013: s.v. *pattada*) prefers to derive it from **pactiare* 'to make a pact'. Semantically, Paulis' etymology is more likely. For the development of Lat. **coa-* > Srd. **pa-*, cf. Srd. *báttile* 'cloth put between a horse's or donkey's back and the saddle' < Lat. *coactilis* 'felt' (DES I: 187–88), Log. *padzáre* 'add rennet to milk' < Lat. *coagulāre* 'to cause a fluid to curdle' (DES I: 237).[191] Wagner's suggestion of a pre-Roman origin for this word is in any case not necessary.

5.1.15 Precipice, cliff, ravine

Reconstruction	Forms
***péntum-**	DES (II: 244): ***péntuma*** (Log.) 'chasm', (1: Bonorva) 'cavern of which the bottom is not visible', ***ispéntumu*** (Log., Camp.) 'precipice, cliff', ***ispéntuma*** (2: Orgosolo) 'ravine'

Forms of similar shape and meaning have been documented for much of southern Italy: cf. Laz. *pèndima* (Nemi) 'sloping rock face', Campan. *péndəm^{a}* (Formicola), *péngima* (Ausonia) 'steep rock', Salent. *péntima* (Manduria, Taranto), *pèntuma* (Calimera) 'isolated boulder; cliff' etc. (Parodi 1907: 772; Rohlfs 1927a: 163 fn. 1; Ribezzo 1933; Aebischer 1934).[192] Parodi, Rohlfs and Aebischer regard this as an originally Latin word. According to Parodi (1907: 772) it could be a derivation from **penditus* 'hanging', with the Latin suffix *-ĭmus*, *-ŭmus*. Rohlfs (1927a: 1963 fn. 1) considers it a specifically southern Italian "provincial term" in Latin, without offering a more specific etymology.

[190] Freely translated from: "luogo eminente dal quale lo sguardo spazi in un ampio orizzonte".

[191] The development of **-g(ŭ)l-* > *-dz-* in Srd. *padzáre* < Lat. *coagulare* is not regular, but may be influenced by Log. *kadzáre* 'to curdle' << It. *cagliare* 'id.' (Wagner 1941a: 163; DES I:396–97).

[192] Cf. these sources also for toponymic comparisons.

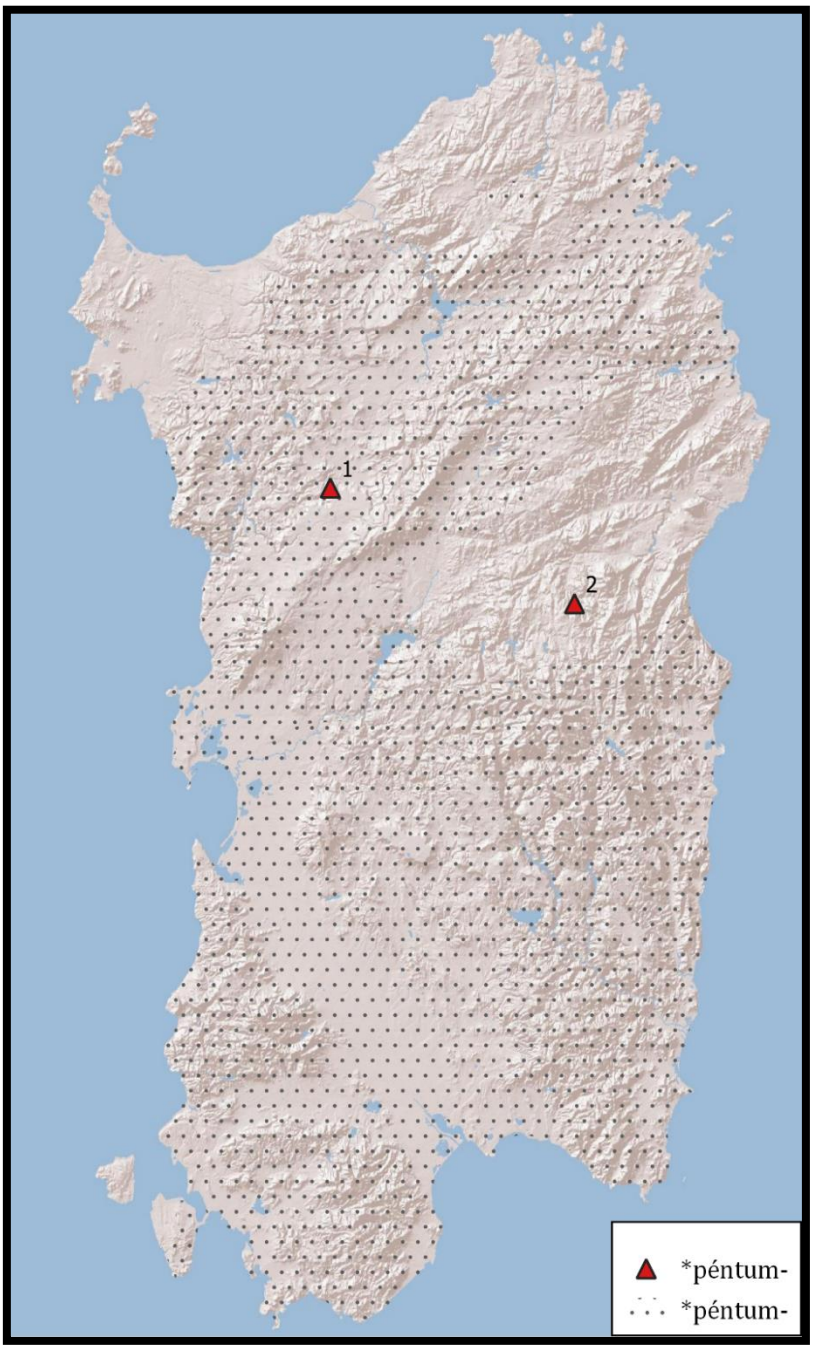

Figure 5.1.15: Precipice, cliff, ravine

Ribezzo (1933; 1953: 119 fn. 9) and Hubschmid (1953: 59) instead prefer a pre-Roman origin, potentially related to Etr. *penθuna*, *penθna* 'stele, cippus' (Ribezzo 1953: 119 fn. 9).[193] I am more convinced by the analysis of this word as an originally Latin formation particular to southern Italy and Sardinia, even if its formation within Latin is not fully understood (cf. Wagner 1932: 243). Perhaps the various forms listed by Hubschmid (1953: 59–60) — e.g. Campan. *pentúoškə* (Benevento) 'clod', Cors. *péntica* 'steep slope' — are derivations of the same Latin root with different suffixes.[194] If this analysis is correct, Srd. *péntuma* etc. does not provide us with any information of the pre-Roman languages of Sardinia. The comparison of these words meaning 'slope; cliff' etc. to Etr. *penθuna*, *penθna* 'stele, cippus' is, in any case, a rather large semantic leap.

5.1.16 Puddle

Reconstruction	**Forms**
***bibitíle**	DES (I: 204): ***biðíle*** (Barb.), ***iðíle*** (Log.), ***biðíli, iðíli*** (Camp.), [***unu***] ***iðíle*** (1: Orgosolo, 2: Norbello, 3: Urzulei, 4: Baunei, 5: Desulo), ***biðíri*** (6: Isili), ***biðílli*** (7: Escalaplano), ***irí*** (9: Narcao), ***ǧiríli*** (10: Sant'Antioco), ***biðíli*** (11: Teulada) 'puddle of water in the mountains; marshy place in a field'[195]
	Böhne (1950: 77): [***su***] ***iðíʔi*** (8: San Vito) 'fish pond'

Wagner (1932: 227) compared this Sardinian word to Bq. *itil* 'puddle', but in DES (I: 204; II: 607) he notes that a formation of the type **bibit-íle*, derived by means of inherited morphemes, is phonetically also possible. Furthermore, Michelena (1959: 525–526) doubts the reliability of the Basque word, and identifies it as a variant of Bq. *istil*, *tistil*, *txistil* 'puddle, gutter, mire etc.', which does not appear to be old in Basque. Wagner (DES I: 204), followed by Paulis (1987: 472), partly bases his preference for a pre-Roman origin of Srd. *biðíle* etc. on the attestation of a toponym *Gitil*, attested in various old Sardinian sources, whose form cannot go back to **bibitíle*. However, as Wagner himself notes (DES I: 204), the meaning of the toponym *Gitil* is not known. This means that its connection with *biðíle* etc. cannot be established with confidence, and *Gitil* cannot be used as an argument for a pre-Roman origin of Srd. *biðíle* etc.

[193] For the meaning of the Etruscan form, cf. Breyer (1993: 347 fn. 515).

[194] Forms like Fr. *pente*, Prov. *pendo* are from **pendita* << Lat. *pendere* 'to hang' (FEW VIII: 181, 183).

[195] DES (I: 204) mentions some further specificities of this word's meaning in its various places of attestation. It sometimes seems to refer also to 'marsh, bog', while in some other places (2: Norbello, 6: Isili, 8: San Vito) it specifically denotes a permanent and walled pool of water.

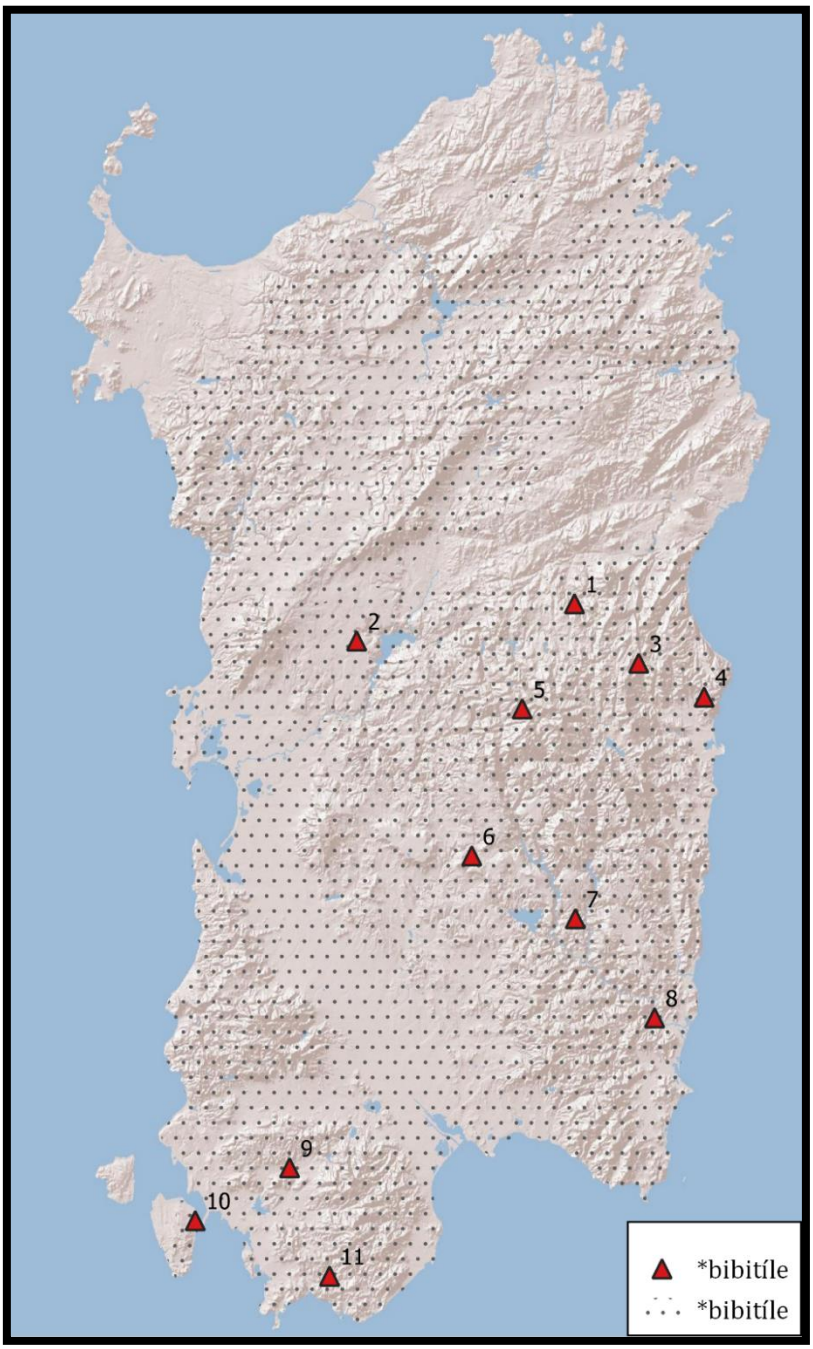

Figure 5.1.16: Puddle

Since the Basque connection is to be rejected too, nothing stands in the way of simply accepting **bibit-íle* as an etymology for this word, which is what Wagner (DES II: 607-8) settles on as well.

5.1.17 Ravine, gorge

Reconstruction	Forms
tórku	DES (II: 522): ***trókku*** (1: Fonni) 'ravine, gorge'
	Paulis (1987: 543): ***trókku*** (Barb.) 'id.'

Besides the lexical form from Fonni, DES (II: 552) lists several toponyms containing this etymon, near the towns of Talana and Tonara. Wagner (1997: 269) considers it pre-Roman, and suggests a shared origin with *tròkko*, *tòrko* 'clay used to make acorn bread' (§ 6.2.2), found in the Ogliastra (DES II: 522), which is semantically quite a stretch. Hubschmid (1953: 42) compares some formally similar and semantically closer forms from the Iberian Peninsula, such as Bq. *troka* 'ravine', dialectal Sp. *torco* (Navarre) 'chasm caused by water erosion; unstable terrain, on which it is dangerous to put one's feet', (Álava) 'pothole', (Bilbao) 'pit for carbonizing wood' etc. (cf. also DCEC V: 551b-52a for more examples, including feminine *torca* in the same meanings). While Hubschmid (1953) is convinced of a pre-Roman origin for both the Sardinian and Iberian forms, Corominas (DCEC V: 521b-522) hesitates between a pre-Roman origin and an etymology in Lat. *torques* 'necklace, collar', which would be semantically motivated by the Iberian words' reference to depressions with circular edges (cf. also Santano Moreno 2000: 166–167).

While the formal difficulties highlighted in DCEC (V: 552) — notably the lack of diphthongization of Lat. *-ŏ-* in the Spanish varieties — might favor a pre-Roman origin, we do in fact find a good semantic parallel for the polysemy between 'gorge' and 'necklace, collar' in Srd. *gútturu* 'throat, mountain gorge' and derived *gutturáta* 'horse collar' (DES I: 599).[196] Hubschmid (1953: 43) and Corominas (DCEC V: 553 fn. 2) agree that Bq. *troka* likely goes back to **torka*, and Corominas (DCEC V: 553 fn. 2) suspects that the word is not originally Basque, due to the low frequency of initial *t-* in the Basque inherited lexicon.

[196] This is an example of the extension from 'throat' to 'gorge' on the one hand, and to 'collar' on the other. The question is whether the opposite semantic extension, from 'collar' to 'gorge' (perhaps with 'throat' as an intermediate step') is also possible.

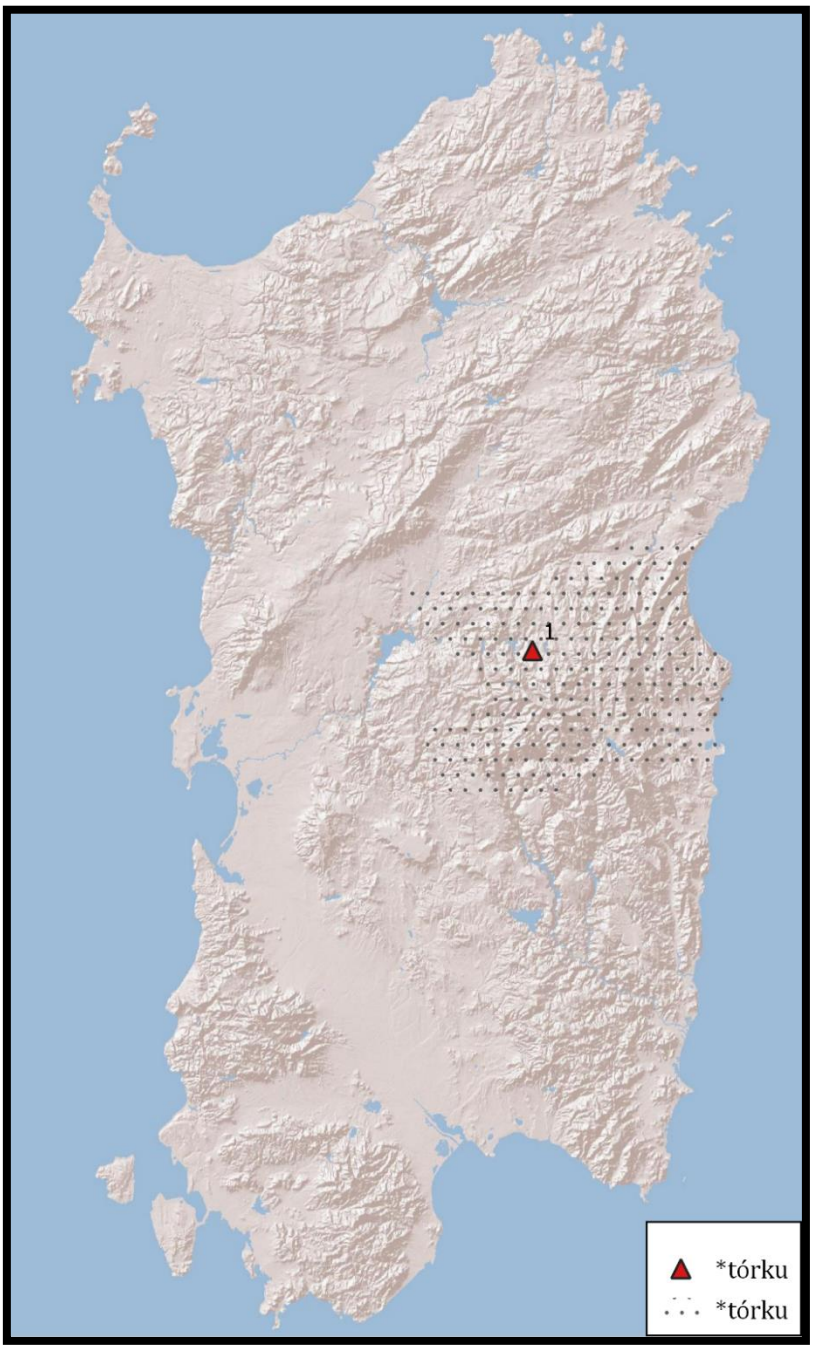

Figure 5.1.17: Ravine, gorge

Sardinian *trókku* too can very well go back to **tórku*, a metathesis that is in fact attested for near-homonymous *tròkko*, *tòrko* 'clay used to make acorn bread' (§ 6.2.2), raising the question whether Srd. *trókku* 'ravine, gorge' too could be from Lat. *torques*, for which only a replacement of final **-e* with *-u* needs to be assumed.[197] Although an inherited origin of Srd. *trókku* 'ravine, gorge' and dial. Sp. *torco*, *torca* 'ravine, pit, pothole etc.' is far from certain, we cannot reliably use Srd. *trókku* to say anything about the linguistic substrate in Sardinia.

5.1.18 Rock, boulder

Reconstruction	**Forms**
****marrárḭu***	DES (II: 78): ***marrárdzu*** (Log.), ***marráğğu*** (NLog.) 'rocky place', ***marráğğu*** (Sass.) 'big rock', ***marráldzu*** (1: Sennori) 'boulder'

These forms, going back to **marrárḭu*, are derived by means of inherited **-arḭu* from a base **marr-*, which Hubschmid (1953: 52) identifies as a pre-Roman word for 'rock'. He lists a large amount of examples of words containing **marr-* in Romance varieties ranging from the Pyrenees (e.g. Gasc. *marràlh* (Lourdes) 'big stone') to the eastern Alps (e.g. Lomb. *maròɡn* (Bergamo) 'boulder') and southern Italy, interspersed with similar formations referring to various kinds of piles/stacks (of wood, hay etc.) (Hubschmid 1953: 52–57; cf. also FEW VI/I:368-373).[198] If these connections are correct, this provides interesting evidence for pre-Roman linguistic connections in the western Mediterranean (§ 11).

[197] For the development of **-qu-* > *-k-* needed for this form, cf. Srd. *tòrkere* < Lat. *torquere* (DES II: 496).

[198] I am less convinced by the forms Hubschmid (1953: 55) adduces for Ibero-Romance. They are limited to a toponym and the name of a game.

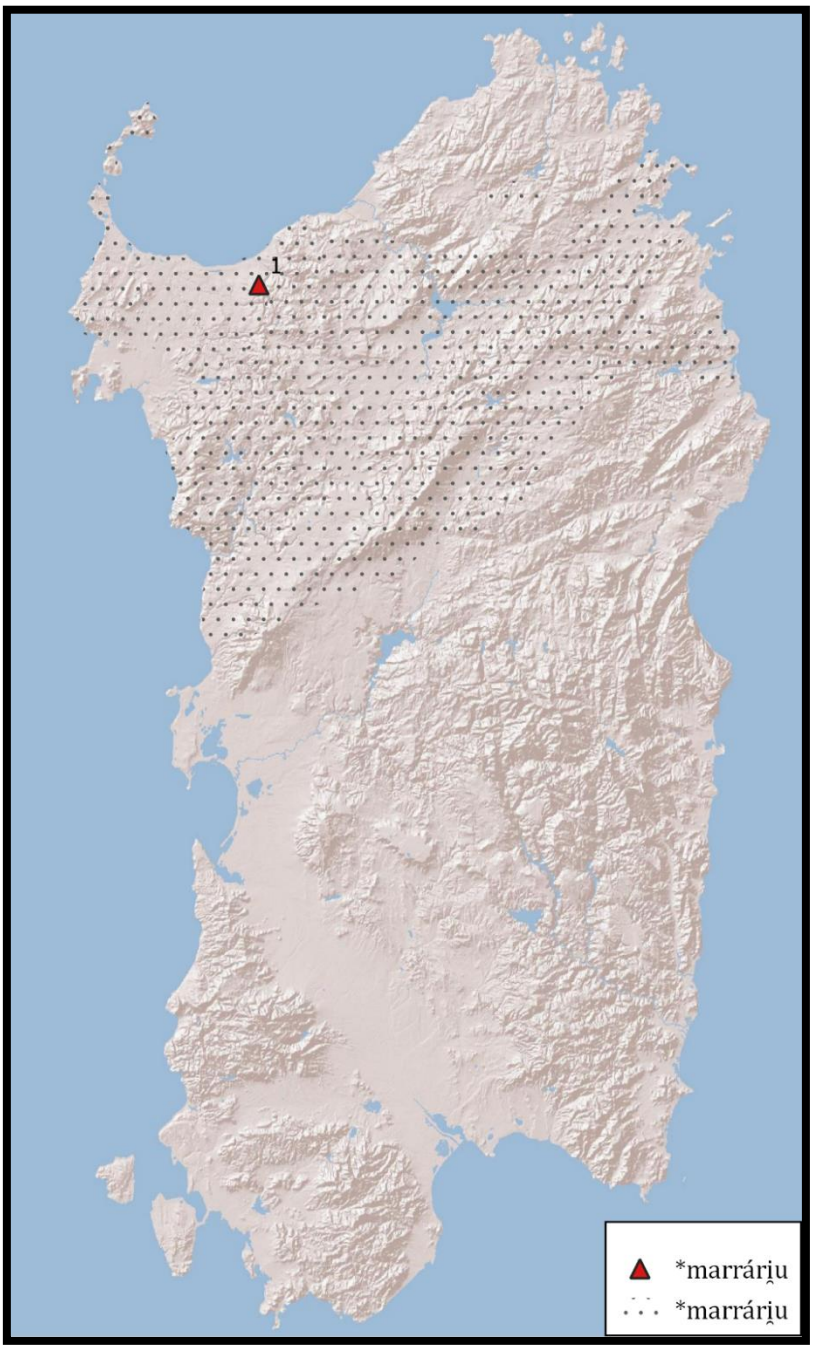

Figure 5.1.18: rock, boulder

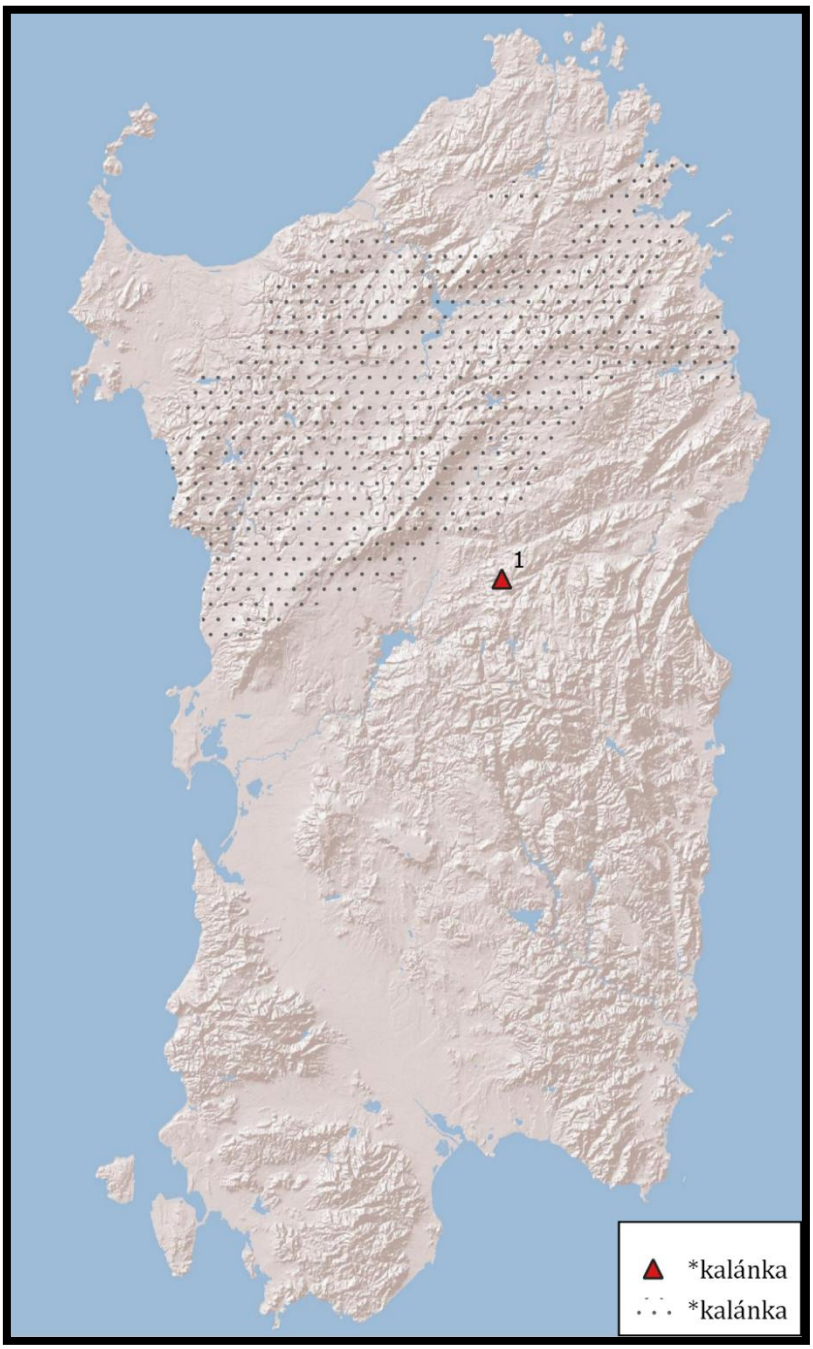

Figure 5.1.19: Rock crevice

5.1.19 Rock crevice$_1$

Reconstruction	Forms
kalánka	DES (I: 266): ***kalánka*** (Log.) 'rock crevice; precipice; cave'
	Blasco Ferrer (1988: 174): ***kolonkòne*** (1: Orani), ***kalánka, kalankòne*** 'large inlet; rock crevice; precipice; eye socket'

This word has good comparanda in Cors. *calanca*, *calônca* 'sheltered place between mountains or hills; rocky inlet', OProv. *calanca* 'narrow alley', Prov. *caranco* 'narrow bay where ships take shelter; deep spot along the sea or a river, where one can find fish', *kalãnko* 'rock crevice' (dep. Var), *kanãnko* 'caverne' (dep. Alpes-Maritimes), Tusc. *calanco* 'rockslide', Calab. *kalankə* 'landslide' etc., and is also attested in multiple toponyms in these regions (cf. Rohlfs 1932: 139; FEW II:56; DES I:266; Ambrosini 1983: 12; BDLC: s.v. Cors. *calanca*). The various different meanings attested united by their common reference to a narrow and often steep place in rocky terrain. Wagner (1932: 242; DES I:266) and FEW (II: 56) believe that Srd. *kalánka* and its apparent cognates are of pre-Roman origin, and that they are derivations from **kála* 'secluded, sheltered place', whence Srd. *kála* 'animal den; cove, inlet' (§ 5.1.3). Since *-(*a*)*nka* is neither a productive nor an inherited suffix in any of the languages in which these words occur, this derivation would have to have taken place in its hypothetical pre-Roman source language. While there may be evidence for such a suffix in some of the other languages mentioned (FEW II: 56), we do not have this evidence for Sardinian and I am not aware of other proposed pre-Roman Sardinian words ending in –(*a*)*nka/u* besides *kalánka* 'rock crevice' and *katánka* etc. 'tick' (§ 4.1.16). Wagner (1932: 242) and FEW (II: 56) are probably correct in positing a pre-Roman origin for this word. However, the restriction of this word to Logudorese, in the north of Sardinia, raises the question whether its Sardinian representative could somehow be a more recent acquisition from Cors. *calanca* 'sheltered place between mountains or hills', whose meaning is very close to the Sardinian one. Whether **kalánka* 'narrow place between rock (vel sim.)', found in France and Italy, is indeed eventually a derivation from **kála* 'sheltered place (vel sim.)', found in more or less the same regions (§ 5.1.3), depends on the evidence for *-(*a*)*nka* as a suffix in the originally pre-Roman lexicon of these other languages, but there does not seem to be good evidence that *-(*a*)*nka* was ever a productive suffix in any language spoken on Sardinia (§ 9.2.5).

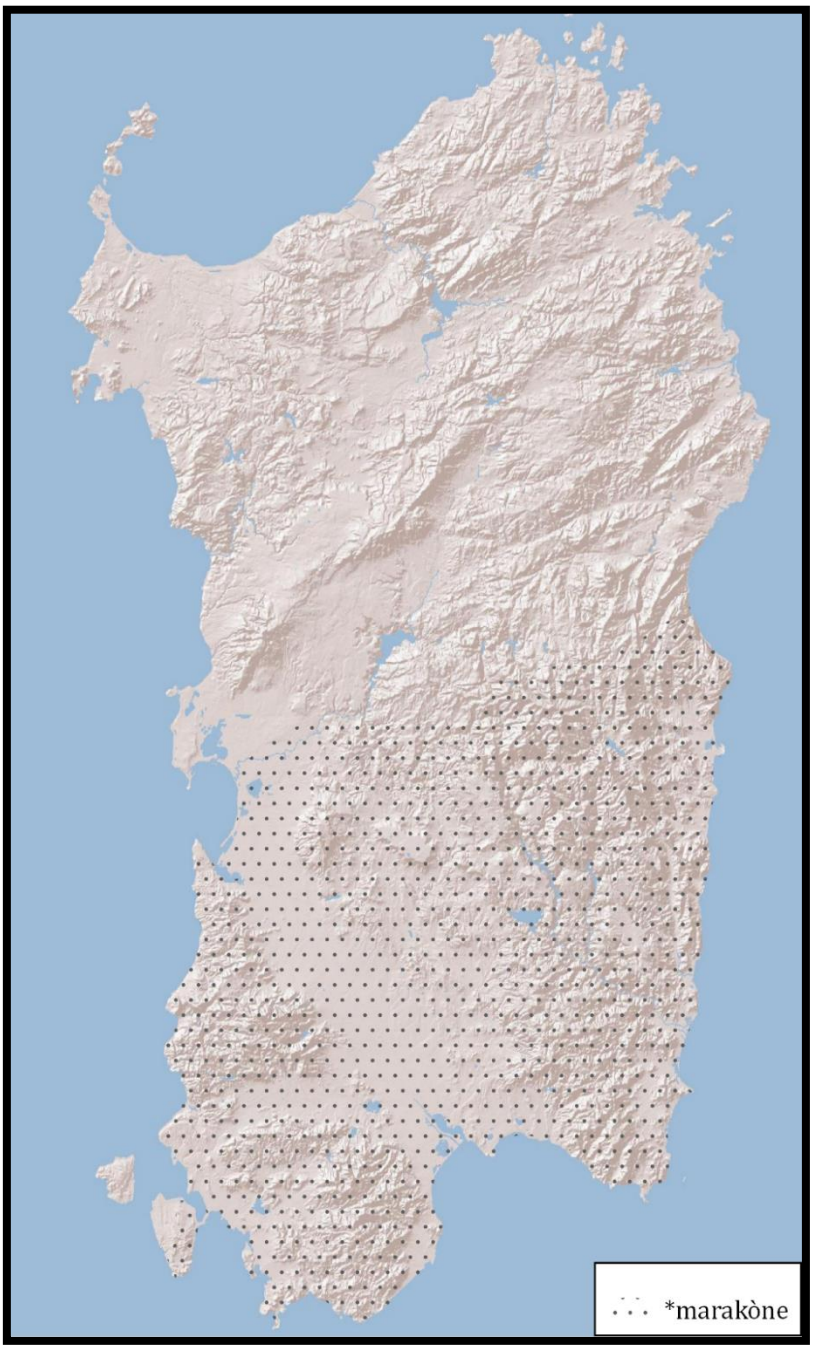

Figure 5.1.20: Rock crevice

5.1.20 Rock crevice$_2$

Reconstruction	**Forms**
marakòne	DES (II: 72): ***marayòni*** (Camp.) 'rock crevice'

This word has been much discussed in the literature. Terracini (1927: 144–145) derives the mountain name *Monte Maragonis* from Srd. *mára* 'marsh' (§ 5.1.8), about which Wagner (1932: 241 fn. 1) expresses his doubts on semantic grounds. Hubschmid (1953: 65–66) argues that semantic shifts between 'marsh' and 'rock crevice' or 'precipice' are not uncommon, and provides some Gallo-Romance parallels. He connects Srd. *marayòni* 'rock fissure' to Abruzz. *maragone* which in some parts means 'large quagmire, formed by flooding' and in others 'precipice'.[199] Later however, Hubschmid (1958: 223) retracts his idea of the cognacy between Srd. *marayòni* 'rock fissure' and Abruzz. *maragone* 'large quagmire', as the latter was explained by Alessio (apud Hubschmid 1958: 223) as a reflex of **margone* << Lat. *marga* 'marl'. It is not clear whether Hubschmid also rejects the connection between Srd. *marayòni* and the South Italian forms in the meaning 'precipice, cliff; ravine'. This latter connection is in any case still continued in DES (II: 72) and in DCEC (III: 782), who further adds Sp. *malecón* 'embankment, dike made to prevent water damage' and proposes all forms to eventually go back to **muriconem* << Lat. *mūrex, -icis* 'pointed rock, caltrop'. This semantic development may work for the Spanish word, and with some imagination perhaps even for the South Italian forms in the meaning 'precipice'. However, a development from 'pointed rock, caltrop' to 'rock crevice' seems rather implausible, not to mention the irregular vowel developments that are needed to derive Srd. *marayòni* from **muriconem*. Wagner (DES II: 72) finally also mentions Cors. *varangonu*, *vangaronu* 'ravine; landslide', whose comparison to Srd. *marayòni* is not evident either.

In summary, the only connection that withstands scrutiny is the one to Abruzz. *maragone* (Gallo, prov. Caserta) in the meaning 'precipice' and perhaps to Lucanian *ɣarammȍnu* (San Chirico Raparo, prov. Potenza) 'precipice, cliff' and *ɣarammȍnə* (Oriolo, prov. Cosenza) 'ravine' further south, with apparent metathesis (AIS 428). That is, only if these forms are not in fact identical to Abruzz. *maragone* 'large quagmire', in which case they too go back to Lat. *marga*, leaving Srd. *marayòni* isolated. This seems semantically less likely, though.

[199] Cf. AIS (428), where in the meaning 'precipice, cliff' we find *marayōne* (Gallo, prov. Caserta).

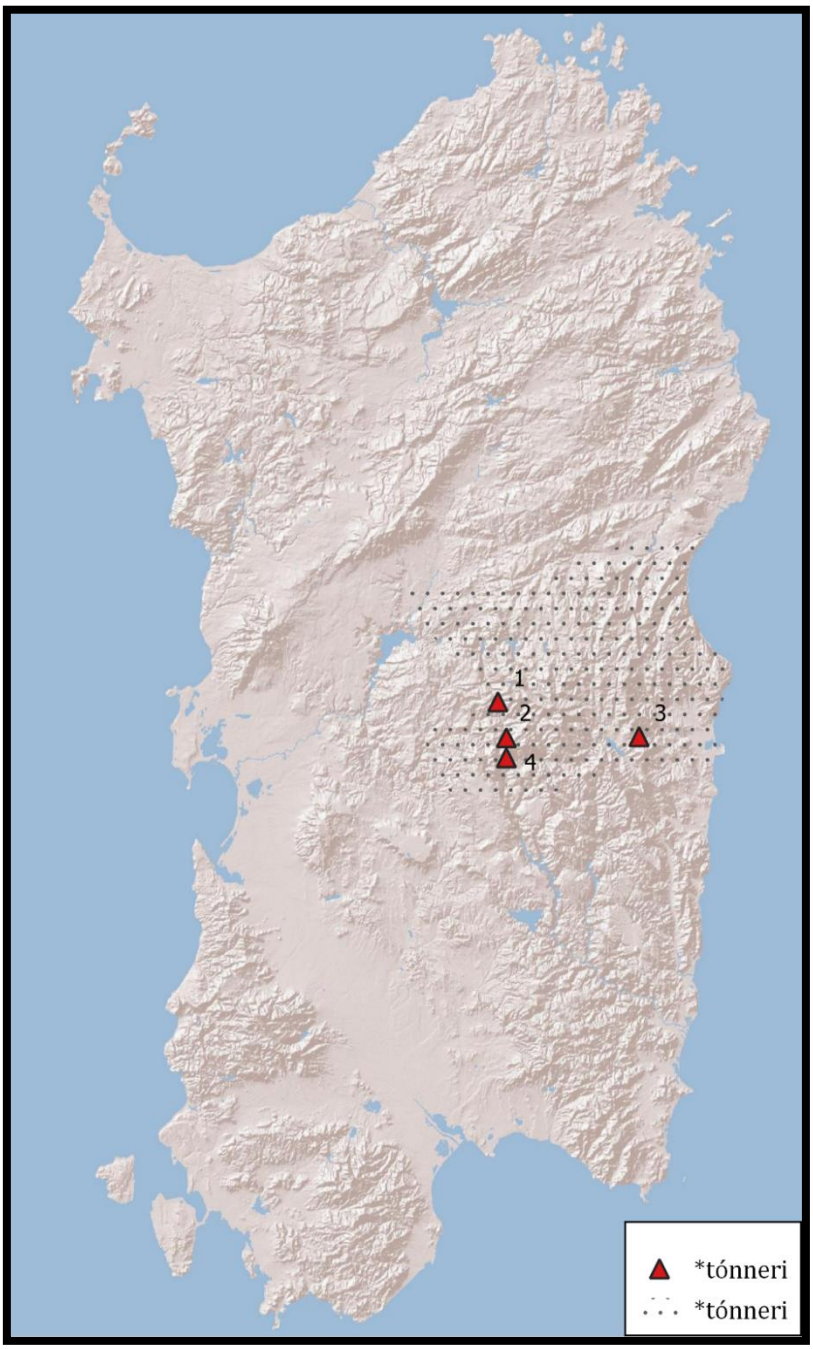

Figure 5.1.21: Rock outcropping

5.1.21 Rock outcropping$_1$

Reconstruction	**Forms**
tónneri	DES (II: 495): ***tón(n)eri*** (Barb., 1: Tonara, 2: Belvì, 4: Aritzo) 'type of conical rock outcroppings'; (3: Villagrande Strisaili) 'vertical rock wall'

According to Wagner (DES II: 495), this word is of pre-Roman origin. This is plausible due to its phonological shape (cf. § 8.1.5.1), and its reference to regionally specific geological features. Wagner (DES II: 495) further notes that an element **ton(n)-* is found in several toponyms in the region where the outcroppings referred to as *tónneri* are found (e.g. Tonara). While this is certainly suggestive, the semantic opacity of these toponyms does not allow us to draw any conclusions about this potential link. Finally, Wagner (DES II: 495) mentions a Corsican hydronym *Tónnari* collected by Bottiglioni (1929: 82 fn. 2), who derives it from Etruscan. In this case, too, the opaque nature of the semantics of toponyms prevents any firm conclusions.

5.1.22 Rock outcropping$_2$

Reconstruction	**Forms**
karaCile	Puddu (2023: s.v. *caraíle*, *caràili*, *caraíli*): ***caraíle, caràili, caraíli*** 'rock outcropping that is difficult to traverse etc.'
	Paulis (2008: 23): ***caràili*** (1: Ilbono) 'big rock'

The vowel hiatus in *caraíle*, *caràili* suggests the original presence of a consonant (possibly **-b-*, **-v-*, **-d-*, **-g-*) between the antepenultimate and the penultimate syllable. Paulis (2008: 22, 23) believes this word to be of pre-Roman origin, and hypothesizes that it might be the origin of *Caralis*, the Latin name of modern day Cagliari. Moreover, *caràili* etc. would be proof of the existence of the pre-Roman root **kar(r)-* 'rock' in Sardinia (cf. Hubschmid 1953: 108–109). However, the origin of the ending **-aCile* is not clear. I am not convinced by Paulis' (2008: 22) comparisons of this word to Cat. *quer* 'rock' etc.

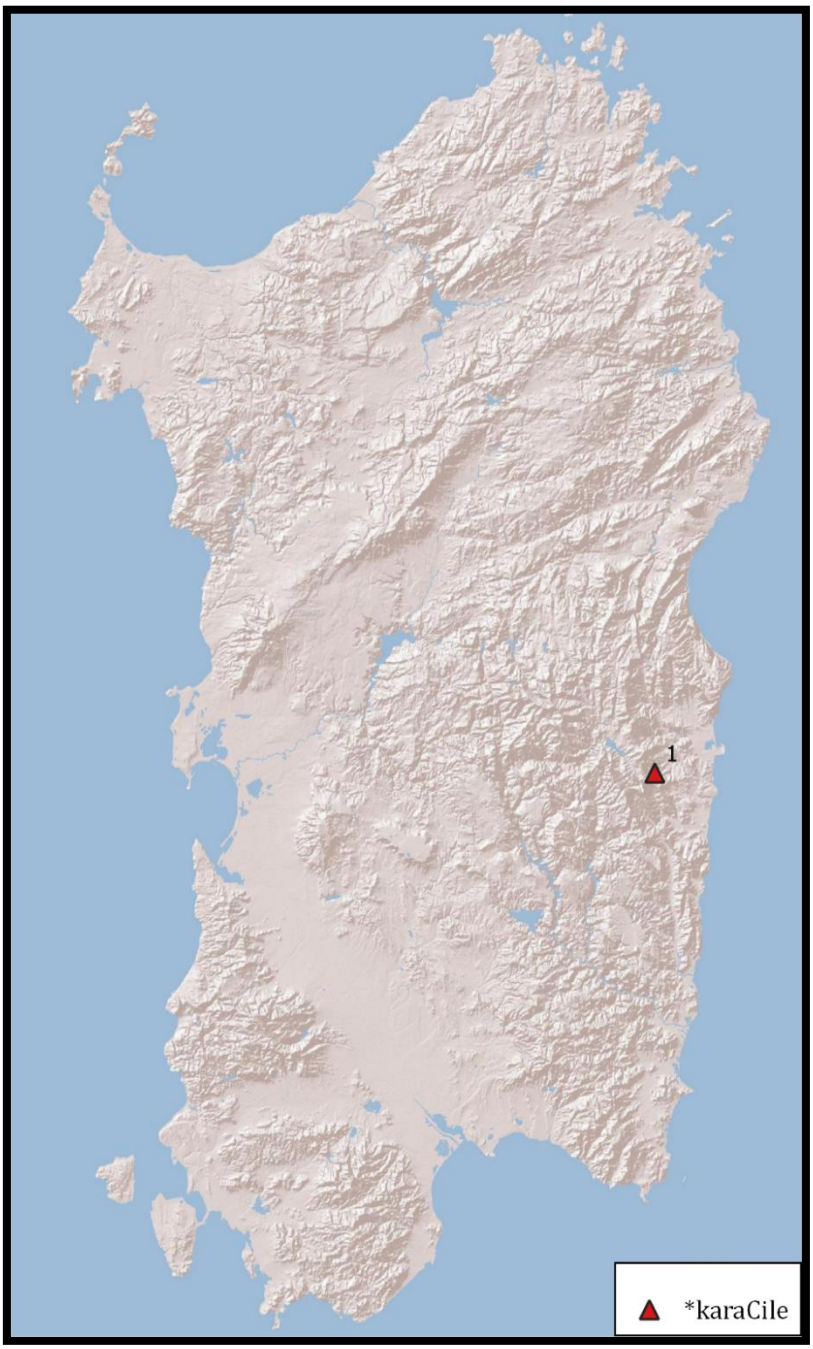

Figure 5.1.22: Rock outcropping

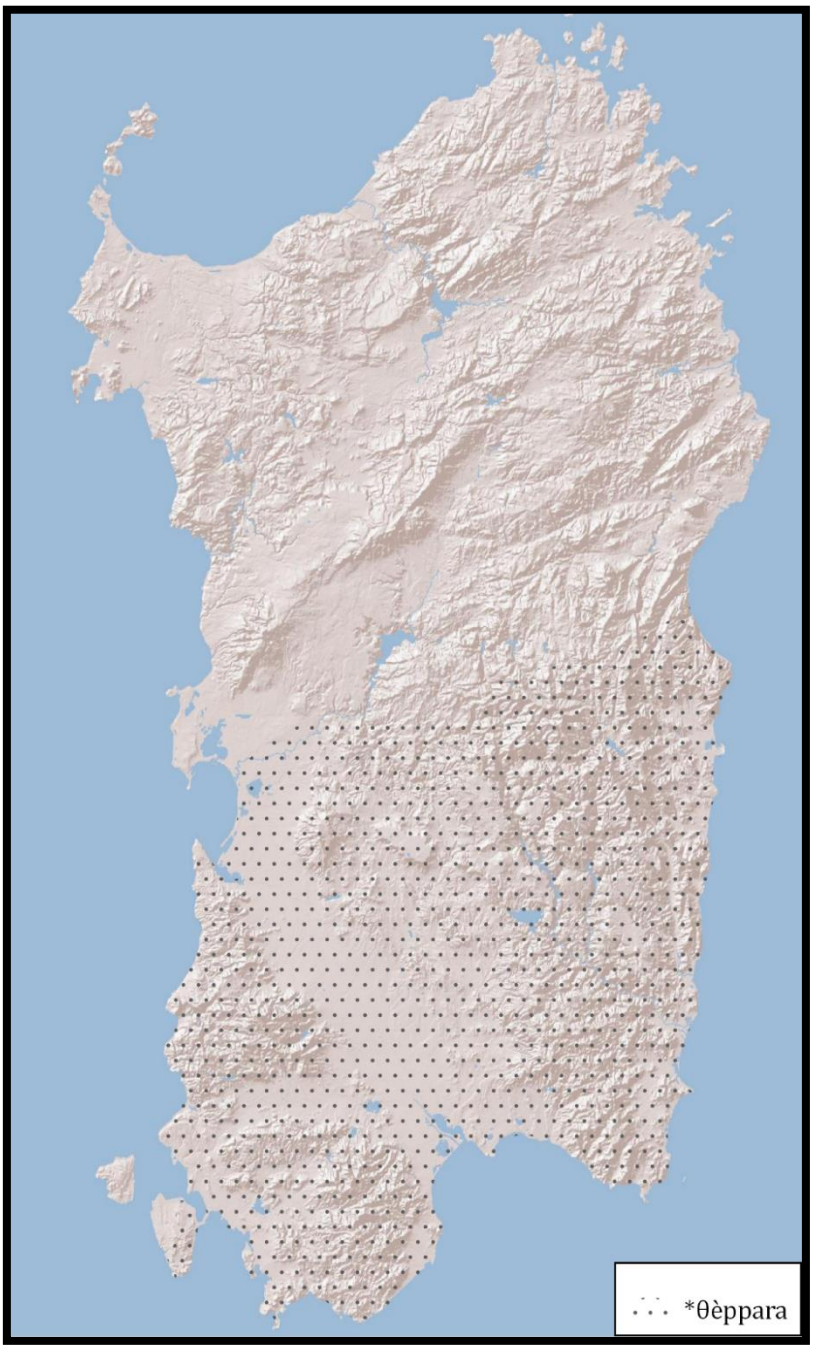

Figure 5.1.23: Rocky plain

5.1.23 Rocky plain

Reconstruction	**Forms**
***ϑèppara**	DES (II: 588): ***tsèppara*** (Camp.) 'rocky plain; hill, summit'

Hubschmid (1953: 47–48) and Wagner (DES II: 72) consider this word to be of pre-Roman origin. Paulis (1985: 232a) agrees and suggests that the form *Theppar* attested in the Condaghe di Santa Maria di Bonarcado (Virdis 2003: sec. 173) may belong here too, as well as the toponym *Theppage* attested in the Condaghe di San Pietro di Silki (Bonazzi 1900: sec. 295), be it with a different suffix. Among Hubschmid's (1953: 47–48) comparisons are Cors. *sápara* 'cave' and Bq. *zapalda* 'natural shelter for lifestock', but he later retracts the Basque comparison (Hubschmid 1958: 220). I agree with Battisti (1955: 477) that the connection of the Sardinian word to Cors. *sápara* 'cave' is difficult as well, from a semantic point of view. Hubschmid (1953: 48) and Wagner (DES II: 72) think that this word contains the pre-Roman suffix *-ára* that must be posited on the basis of other words (§ 9.2.6.2), and which may have had a plural or collective function (Terracini 1927: 139). For *tsèppara*, the toponym *Theppage* could be evidence for a root **ϑepp-* with a suffix *-ára*, although it is not certain that these forms are cognates.

5.1.24 Secluded place

Reconstruction	**Forms**
***óspile**	DES (II: 198): ***óspile*** (1: Lollove, 2: Nuoro) 'enclosure in an estate where calves are kept', (3: Orgosolo) 'small, undeep gorge', (4: Villagrande Strisaile) 'gorge, natural cavern hidden by branches'

The exact meanings of this word are rather intricate and specific, and expanded upon by Wagner (1932: 226). The common semantic feature of the forms listed is that of some secluded place. Wagner (1932: 227; DES II: 198) compares these words to Bq. *ospel* 'shady place' and (*h*)*ozpil* 'cold, chilly'. These two Basque words are probably not etymologically related to each other however. According to Trask (2008: 320), Bq. *ospel* 'shadowy place' may contain **bel* 'dark', while the first part is obscure. Bq. (*h*)*ozpil* 'cold', on the other hand, likely contains Bq. *hotz* 'cold' (DES II: 198). As a consequence, the comparison is restricted to Srd. *óspile* and Bq. *ospel* 'shady place' only. Wagner may be right in considering this word a pre-Roman remnant, although there is no positive evidence that *-ile* is a suffix in this word, as he suggests (cf. § 9.2.3.3).

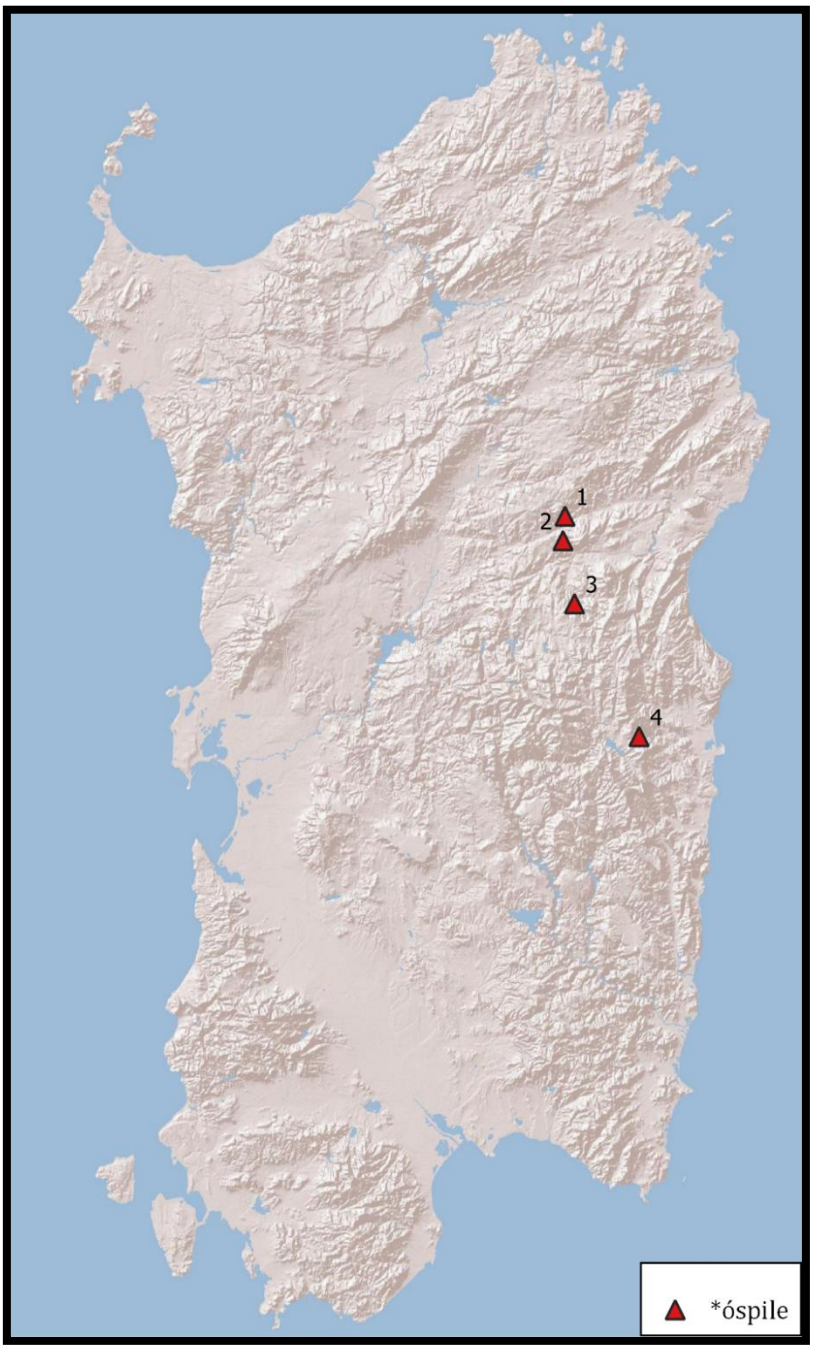

Figure 5.1.24: Secluded place

The formal and semantic proximity between Srd. *óspile* 'secluded place etc.' and Bq. *ospel* 'shady place' is close enough to tentatively posit a shared origin for these words. It remains to be seen on the basis of more evidence what this potential Basque-Sardinian connection means for the prehistoric linguistic situation of Sardinia (§ 10.3, § 11).

5.1.25 Stream, torrent

Reconstruction	**Forms**
***rò̯ia**	DES (I: 127): ***ròya*** (Camp.) 'downfall of water', ***arróyu, arròya*** (1: Bitti) 'low and watery place', ***orròa*** (2: Tonara) 'stream, torrent', (4: Belvì) 'valley', (5: Laconi) 'marshy valley', ***arròa*** (3: Desulo) 'watery ravine, gorge'
	Porru (1866: 1085): ***ròja*** 'downfall of water'

All Sardinian forms can go back to **rò̯ia*, with regular vowel prothesis before initial *r-* (cf. DES I: 127).[200] This word has been much discussed in the literature. Wagner (1940: 152) compares the Sardinian forms to Sp. *arroyo*, Pg. *arroio* 'stream, brook, creek', which in turn have been derived from Lat. *arrugia* 'tunnel' (REW 678) — cf. also Cat., Val. *roll* 'stream of water, irrigation ditch', Gasc. *arrouio*, *arrouil* 'gutter, rivulet', Bearn. *arroulho* 'irrigation ditch' etc. < **arrugula* (DCEC I: 359; FEW XXV: 343-44), Bq. *arroil*, *arrolla* 'gutter, ditch, trough' (DCEC I: 359);[201] Piem. *ruğa* (Valsesia), Lomb. *róğa* (Como), Ven. *roza*, Lad. *roia* (Val Gardena), *roa* (Val di Fassa), Friul. *roye* 'millrace' < **rugi̯a* (REW 678; Bertoldi 1932: 400; Hubschmid 1950c: 34–35). However, Lat. *arrugia* is only attested in Pliny (*Nat*. 33.25), referring specifically to shafts in gold mines in Hispania, indicating that this word entered Latin (and Romance) from a local Hispanic language (cf. Hubschmid 1950c: 35; 1953: 67). A direct derivation from Lat. *rūga*, as proposed by Walde and Hofmann (1938: 69, 849), is not possible, as the majority of Romance forms point to *arrugia* with a short **-ŭ-* (Bertoldi 1932: 401).

[200] Camp. *ròja*, cited by Porru (1866: 1085) (and as *ròya* in DES I: 127), is probably to be interpreted as a graphic rendering of /arròya/. In Campidanese, vowel prothesis before initial *r-* is productive (Wagner 1941a: 53–54).

[201] Bq. *arroil* is likely a loan from Gascon (FEW XXV: 343b).

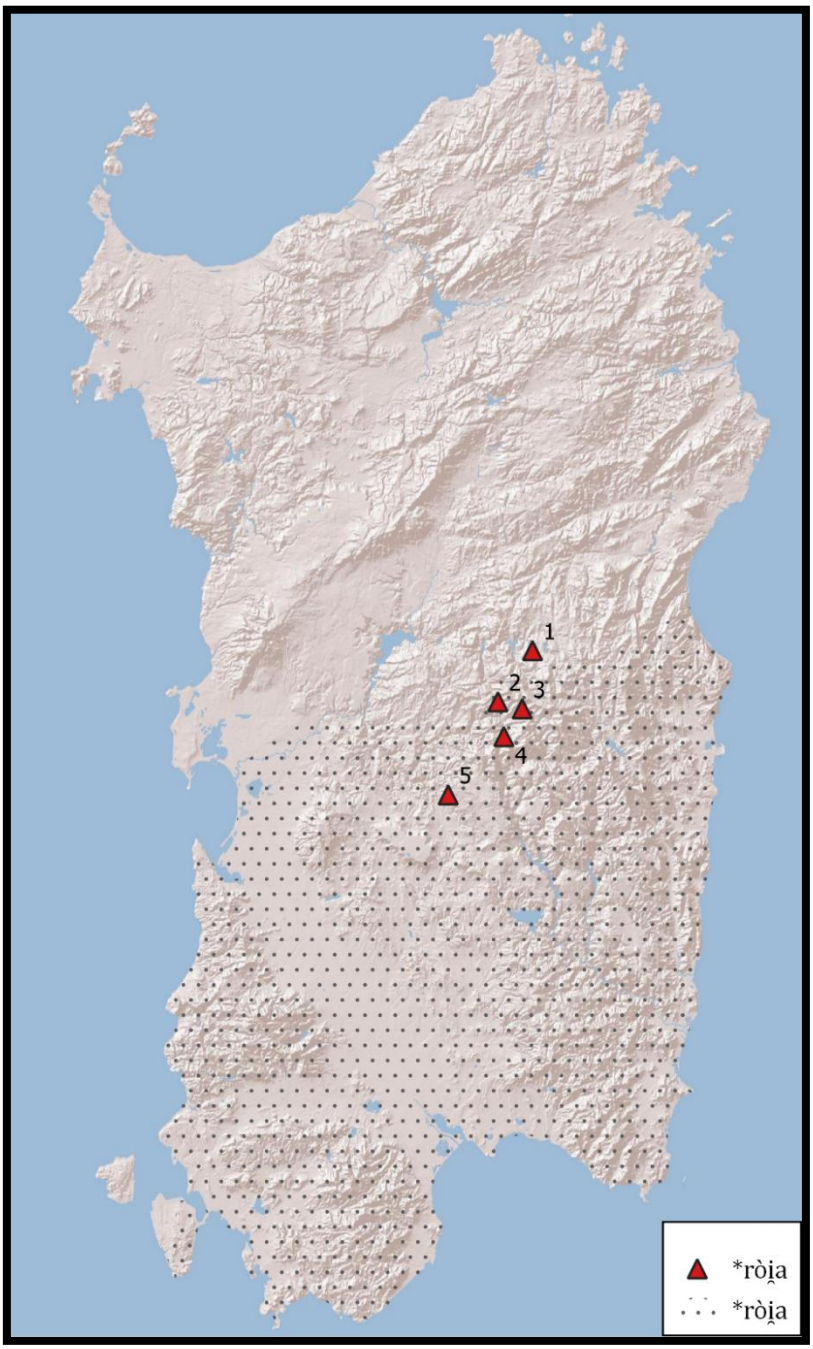

Figure 5.1.25: Stream, torrent

Nevertheless, Hubschmid (1953: 68–69) is favorable to an Indo-European, albeit not a Latin, origin of *arrugia*, on the basis of the attestation of *corrugos* (acc. pl.) 'pipelines to wash fallen material' (in the same context, Plin. Nat. 33.25; cf. Adams 2003: 452),[202] and the comparison to Gr. ὀρύσσω 'to dig', ὀρυχή, ὀρυγή 'digging' < PIE *$h_3$reu-k- (cf. Beekes 2010: 1113).[203] The forms *arrugia* and *corrugus* have been seen as cognates, due to their shared Hispanic context and similar meanings (e.g. DCEC I: 359b; Adams 2003: 452 fn. 135). If this is correct, it is attractive to analyze *corrugus* as **kon-rug-*, suggesting an Indo-European origin. As Corominas (DCEC II: 277) notes however, there are no clear Italic or Celtic formations that could account for both *arrugia* and *corrugōs*, and it is indeed uncertain that these forms are in fact related (DCEC II: 277b).

Returning to the Sardinian forms, it is clear that they cannot go back directly to *arrugia*, as *-u-* would have been maintained in Sardinian (Wagner 1940: 152); the same issue applies to Gr. ὀρυγή 'digging' (cf. Hubschmid 1953: 69). Wagner (1940: 152; DES I:127) also rules out a loan from Sp. *arroyo*, as Srd. *arroia* is attested already in the Condaghe di Santa Maria de Bonarcado (cf. Virdis 2003: sec. 119). This work was composed in the 12th and 13th centuries (Virdis 2003: 11–12), and as such predates Spanish political as well as linguistic influence in Sardinia.[204] Similarly, although Cat. *rolh* 'stream of water; irrigation ditch' would formally work as a source of borrowing for the Sardinian forms, with the same development of Cat. *-lh-* as in Srd. *sakkáyu* 'one-year old lamb, kid' << Cat. *segalh* 'id.' (DES II: 374; DCEC V: 189), the first attestation of Srd. o*rroia* predates the Aragonese presence in Sardinia. This leads Hubschmid (1953: 69) and Wagner (DES I: 127) to propose a pre-Roman origin for the Sardinian forms as well, with an original variant **arrŏgia* with a short **-ŏ-* that is corroborated by Western Leon. *arruoyo* (DCEC I: 359). As of yet, this appears to be the most plausible scenario.

The *orroia* mentioned in the Condaghe di Santa Maria di Bonarcado (Virdis 2003: sec. 119) is described as being in the vicinity of *Sancte Eru* 'San Vero Milis', which is in the Campidano di Milis in the west of the island. This shows that the word in question must at one point have had a wider distribution than its modern representatives suggest.

---

[202] Probably from **cŏrrŭgus*, judging by its Ibero-Romance descendants like Sp. *cuérrago*, Pg. *córrego*, Cat. *còrrec* (DCEC II: 276-77).

[203] In his earlier article however, Hubschmid (1950c: 35) notes that a Celtic origin is unlikely based on the geographic distribution of these words.

[204] The form *orroia* is attested on page 48v, which is dated to 1171-1184 (Virdis 2003: 11).

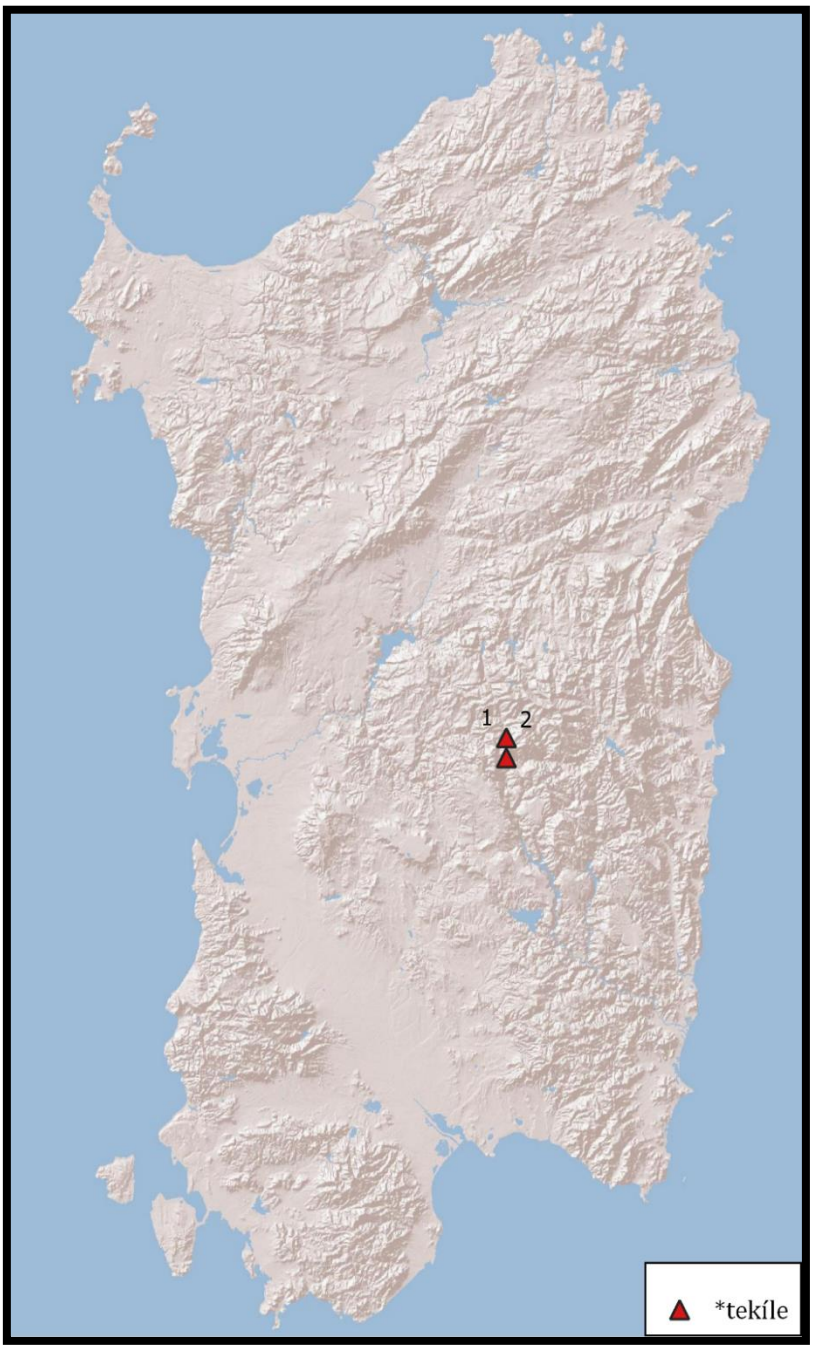

Figure 5.1.26: Isolated summit

5.1.26 Summit, in isolation

Reconstruction	Forms
tekíle	DES (II: 481); Spano (1872: 114): ***težíle*** (1: Belvì, 2: Aritzo) 'isolated summit'

This word, which goes back to **tekíle*, is restricted to the region around Belvì. Wagner (DES II: 481) calls it "certainly pre-Roman", without further argumentation. The ending *-íle* is found in other Sardinian words of uncertain origin as well (§ 9.2.3.3), but this word does not show evidence of it being a suffix.

5.1.27 Water basin; source; stream

Reconstruction	Forms
ϑúrgalu	DES (II: 555): ***ϑúrgalu*** (Nuor.) 'downpour of water', ***túrgalu*** (Log.) 'a groove cut out in the mountain by rainwater'; deriv.: ***inturgaláre, intulgaláre*** (Log.) 'to channel water into gutters', ***istrugaláða, istrúgulu*** (NLog.) 'shower of rain'
	AIS (852, 854): ***ϑúrgalu*** (1: Baunei) '(basin of a) water source'

Wagner considers these forms to be of pre-Roman origin, on the basis of the presence of **ϑ-*, and rightly notes that a connection to It. *trogolo* 'trough' is phonetically impossible. Pittau (2013: s.v. *Tulgaru*) also connects forms like *trúbulu* (Oliena) 'temporary stream of water', *troccale* (Mamoiada) 'little stream of water' and Log. *trógliu* 'trough, basin', as well as the toponym *Durgále* (It. Dorgali). None of these forms contain *ϑ-* however, which was the reason to posit a pre-Roman origin in the first place, and which would be expected at least in Oliena and Mamoiada.[205] This and other phonetic objections make it more sensible to keep the forms listed by Pittau apart from **ϑúrgalu* for now. Log. *tról'u*, *dról'u* (written as *trógliu* in Pittau) is explained by Wagner as a loan from Cat. *toll, doll* 'deep trough in a river bed, small pool of water', influenced by It. *trogolo*.

[205] The form *troccale* (Mamoiada) 'little stream of water', more likely belongs with Barb. *trókku* 'ravine, gorge' (§ 5.1.17).

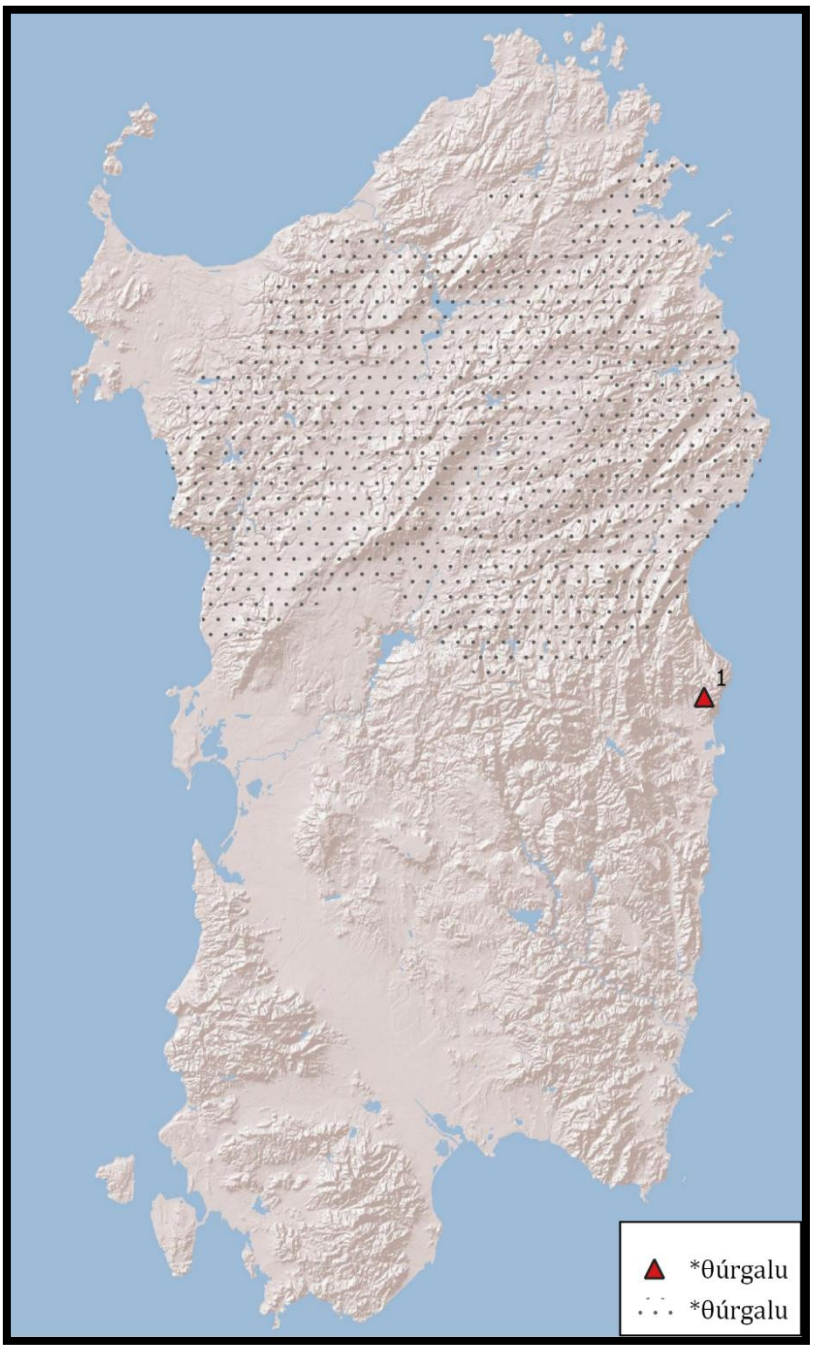

Figure 5.1.27: Water basin; source; stream

5.1.28 Whirlpool; crevice

Reconstruction	Forms
***garrópp-**	DES (I: 308): ***korróppu, gorróppu*** (Barb.), ***karróppu, garróppu*** (Camp.) 'whirlpool', ***garóppu*** (1: Oliena) 'crevice', [***unu***] ***ɣarróppu*** (5: Mogoro) 'whirlpool, whirlwind'; deriv.: ***ṡgarroppaiṡì*** 'to fall down, to plunge'
	AIS (429): [***ũu***] ***garóppu*** [***di ákwa***] (6: Villacidro) 'whirlpool'
	Wagner (1932: 243): [***su***] ***garróppu*** (3: Villacidro) 'water well in the countryside'
	Blasco Ferrer (1988: 174): ***korróppu, gorróppu*** (2: Urzulei, 3: Triei, 4: Baunei)

This word is also represented in toponymy, including in the name *Gola di Gorropu*, one of Europe's largest canyons, between Orgosolo and Urzulei. Within Sardinian we find some phonetic alternations, such as pretonic *-o-* next to *-a-*, and initial *k-* next to *g-*. Both of these phenomena can be explained as relatively recent developments however (cf. Wagner 1941a: 24–27, 206). Santano Moreno's (2000: 163) derivation from Lat. *rūpes* 'cliff' is semantically attractive, but formally impossible. Wagner holds this word to be of pre-Roman origin and compares it to similar forms outside Sardinia: Irpino *sgarrupá* , Abruzz. *scarrupà* 'to fall down, to plunge', Laz. *carapone* (Castelmadama) 'whirlpool, swamp', Nap. *scarrupo*, *sgarrupo* 'cliff, precipice', Calab. *garroppa* (Petronà) 'trough' etc. (Rohlfs 1932: 336; Wagner 1932: 243; DES I:308). Hubschmid (1953: 58) connects even more forms, predominantly from Italo-Romance. Several of them are semantically and/or formally too far removed to be useful, but potentially relevant are South Tusc. *karravone*, Cors. *karravóni* 'ravine',[206] Calab. *karafúnu* 'puddle'.[207] From Gallo-Romance we may add Lang. *caraven* 'chasm, precipice' and Bearn. *garimbáut* 'precipice, ravine, dangerous pass' (FEW XXI: 20).

Hubschmid (1953: 58) also mentions dialectal Greek κάραβος 'gorge, cliff, trench'. Although close in form and meaning, it is not clear to me what the provenance of this form is and whether it is ancient or modern Greek. Wagner (1932: 244) and Hubschmid (1953: 58) believe that this Sardinian word and its cognates contain some kind of "*p*-suffix", but there is no clear evidence that would suggest that these forms are indeed built up of more than one morpheme.

[206] BDLC (s.v. Cors. *caravone*) has Cors. *caravone* 'precipice' (Quenza).

[207] Sic. *garbera* 'ravine' (listed by Hubschmid 1953: 58) has a similar meaning and contains similar consonants, but its syllable structure is rather different than that of the other forms mentioned above.

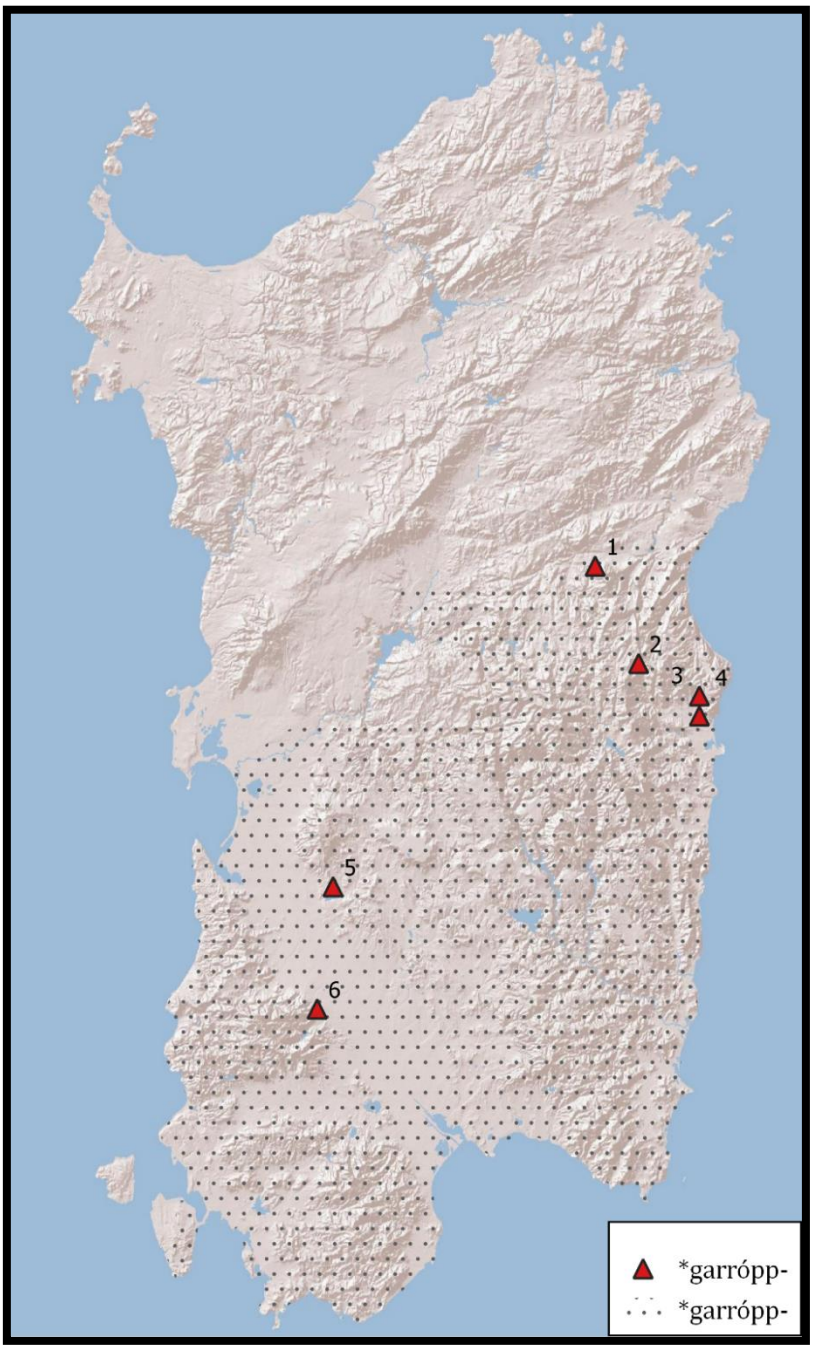

Figure 5.1.28: Whirlpool; crevice

The meanings of the various adduced comparanda ranges from 'puddle' to 'whirlpool' to 'crevice, ravine', all of which are also attested for the Sardinian forms, and it is difficult to establish which might have been the original, or if the polysemy is old. I agree with Wagner (1932: 244) that it is plausible that Sardinian *gorróppu* etc. 'whirlpool, puddle, crevice' shares a pre-Roman origin with the Italo-Romance forms mentioned above. Yet, Wagner (1932: 244) is right that there is some evidence for contamination from other word families, such as It. *groppo di vento* 'whirlwind', *dirupo* 'cliff, precipice' etc. One may also consider Sp., Pg. *barranco* 'trench, ravine, gully' and Gr. φάραγξ 'id.'.[208] These forms cannot be demonstrated to be related either to each other, or to the Sardinian and Italo-Romance forms discussed here (DCEC I: 522-23). Yet, they are strikingly similar in structure, the only substantial difference being the order of the consonants.[209] This raises the question to what extent these formations are somehow subject to some cross-linguistic tendencies of sound-symbolism regarding 'ravines' etc. This problem is outside the scope of this study however.

5.2 Man-made structures

5.2.1 Border, limit

Reconstruction	**Forms**
1: *lákkana	DES (II: 2–3) ***lák(k)ana*** (Log., Camp.) 'border, limit (of a territory)', [***r'a***] ***r'ákana*** (Sass.) 'boundary', [***prèða e***] ***láʔna*** 'border stone' (8: Fonni), ***allákkana*** [***de íḍḍa***] 'boundary marker between villages' (14: Crucuris)'
	AIS (1421): [***sa***] ***lákkana*** (1: Ploaghe, 4: Nuoro, 5: Dorgali, 6: Macomer, 7: Santu Lussurgiu, 9: Milis, 11: Baunei, 12: Desulo, 13: Laconi, 15: Mogoro, 16: Perdasdefogu, 17: Escalaplano), [***sa***] ***láʔana*** (8: Fonni), [***sa***] ***lákkara*** (18: Villacidro) 'border, limit (of a territory)'; [***sa***] ***lákkana*** (2: Bitti) 'border between a community meadow and private property'; [***sa***] ***lákkana*** (10: Busachi) 'border between villages'
2: *lákuna	DES (II: 2–3): ***lákuna*** (3: Orosei) 'trough'

[208] Similar in form is Srd. *barrókku, borrókku* 'cliff, ravine, gully, gorge', cited by Paulis (1987: 470) and Pittau (1995: 202), but absent from DES. Its relation to Srd. *gorróppu* etc. 'whirlpool, crevice' and to Sp., Pg. *barranco* 'trench, ravine, gully' is unclear.

[209] I am explicitly not suggesting a common origin for the Sardinian and Italo-Romance forms on the one hand, and the Greek and Ibero-Romance forms on the other.

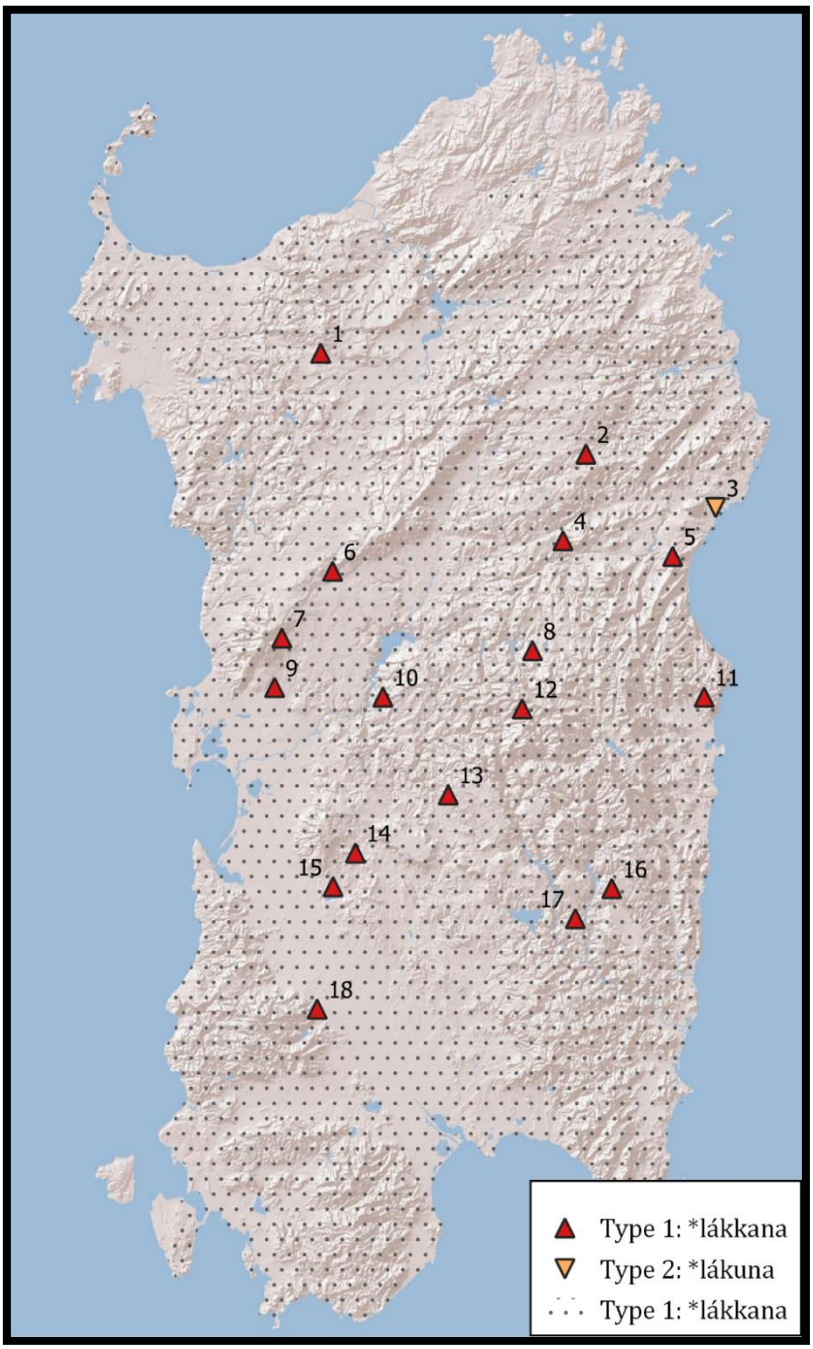

Figure 5.2.1: Border, limit

Wagner (DES II: 2) considers this word, which is attested all over Sardinia, to be of pre-Roman origin. As Wagner notes, a derivation from Srd. *lákku* 'basin, trough, puddle' is difficult, because *-ána* and *-úna* are not productive suffixes in Sardinian. I do think that type 2 may rather be from Lat. *lacūna* 'ditch, pool, pond', with an accent shift to the first syllable (cf. Wagner 1941a: 1–5). However, it is not possible to derive the forms of type 1 from Lat. *lacūna* too, since inherited singleton *-c-* would have sonorized to **-ɣ- in most dialects. Geminate *-kk-* in Srd. *lákku* is attested already in late Lat. *laccus* (DES II: 4), but there is no evidence for a geminate stop in *lacūna*. Semantically the meaning of Lat. *lacūna* 'ditch, pool, pond' is also closer to type 2 'trough' than to type 1 'border, boundary'. The connection by DEI (III: 2197) to Tusc. *legone* 'big rock' and Bq. *legarra*, *lakarra* 'gravel, pebble' does not convince either (cf. DES II: 3).

5.2.2 Nuraghe

Reconstruction	**Forms**
**1: ****nuráke*	DES (II: 176–179): ***nuráke*** (Nuor.), ***nuráɣe*** (Log.), ***nuráži*** 'nuraghe' (Camp.)
**2: ****runák(k)e*	DES (II: 176–179): ***runáke*** (1: Ardara, 4: Siligo, 5: Ittireddu), ***runáɣe*** (2: Ittiri, 3: Ozieri, 6: Thiesi) 'id.'
**3: ****murák(k)e*	***muráke*** (7: Silanus, 8: Macomer, 9: Abbasanta, 10: Bonarcado, 11: Paulilatino) 'id.'

Nuraghes are by a wide margin the most visible and recognizable human-made structures on Sardinia, and have been so roughly for the past 3500 years. Thousands of these tower-like buildings still exist in Sardinia, and they range in size and complexity from single towers to enormous complexes with up to 21 towers (Nuraghe Arrubiu, Orroli). Built mainly between 1700 and 1200 BCE, their use rapidly declined in the start of the 1st millennium BCE, although a number of them were reused for agricultural practices, as sanctuaries, or as tombs in Punic and Roman times (cf. Lilliu 1963). No doubt in part due to their monumentality and their importance for the Sardinian historical identity, many proposals concerning the etymology of the word *nuraghe* have seen the light. It is clear that the forms of type 1 are original, while those of type 2 represent a metathesis and those of type 3 a contamination with *múru* 'wall' (cf. DES II: 176-179).

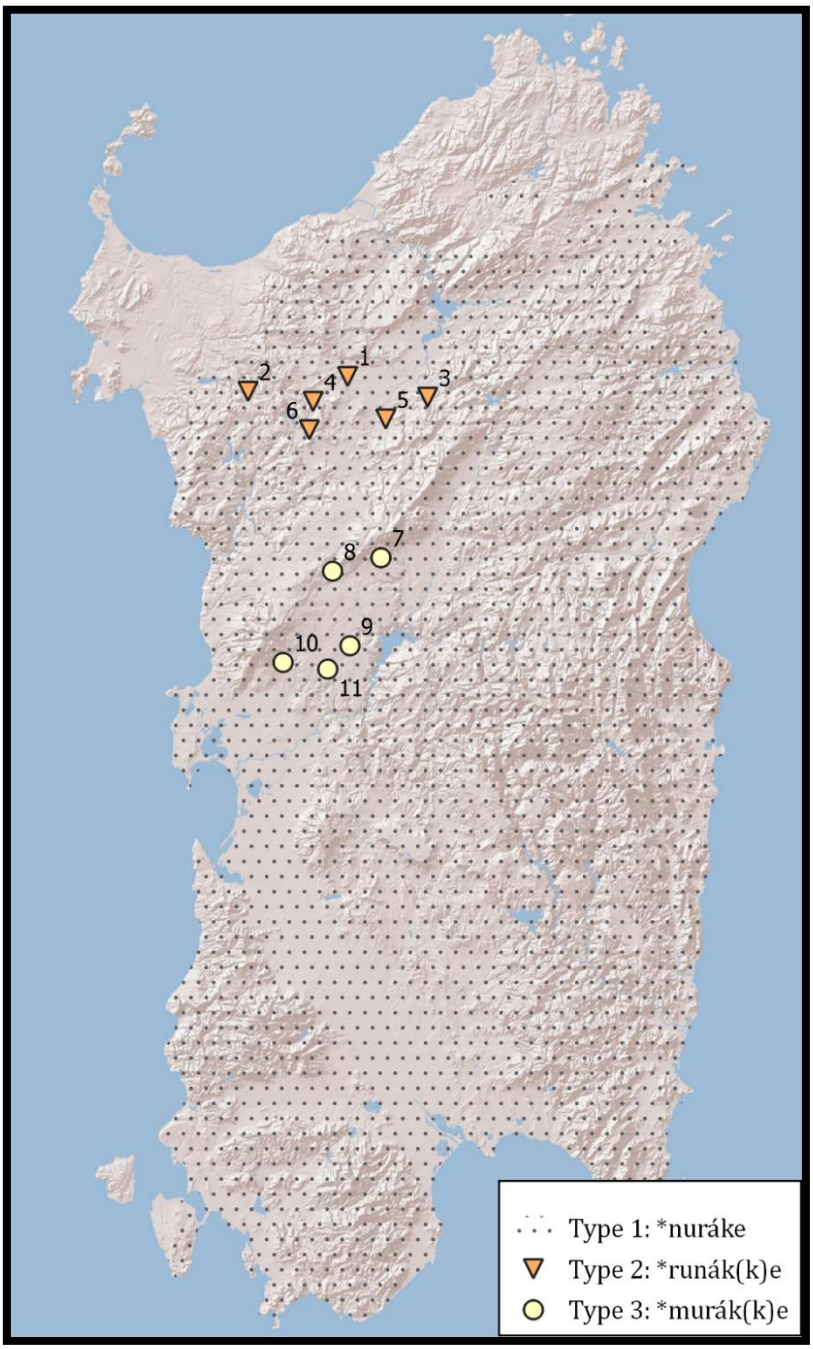

Figure 5.2.2: Nuraghe

That **nurák(e)* must have been the original form of this word, is demonstrated by the word *NURAC* found in a Roman inscription on the nuraghe *Aidu Entos* in Bortigali (Gasperini 1992: 303–306). This inscription also makes **nuráke* one of the very few Sardinian words of proposed pre-Roman origin of which we can be certain that they were actually in use in a form and meaning similar to that attested in modern Sardinian. Wagner (DES II: 176–179) discusses and rejects a range of proposed etymologies that need not all be repeated here. At any rate, etymologies deriving **nurák(e)* from a Latin formation, such as from *mūrus* 'wall' (cf. Flechia 1872: 9), or from *honor* 'honor' (cf. REW 4171), can be put aside.

The pre-Roman etymologies are not without problems either. Wagner's (DES II: 176) analysis of *nuráke* containing a pre-Roman suffix *-áke* is exclusively based on the presumed presence of this suffix in other words (§ 9.2.2.1). While it is not impossible that *nuráke* originally consists of more than a single morpheme, there is no direct evidence for this — unless one accepts Terracini's (1936: 55) and Hubschmid's (1953: 45–47) connection of *nuráke* to Srd. *núrra* 'chasm, shaft; pile' (§ 5.1.1), which is not particularly convincing semantically. Blasco Ferrer's (2010a: 108–109, 115) analysis of **nuráke* as consisting of **nur* '(pile of) rocks' and the Basque plural suffix *-ak* is to be rejected in any case, because nominal plurality in Basque is originally exclusively marked on definite determiners and not on nouns (Trask 1997: 90, 198–200).

For methodological reasons outlined in § 2.1.2.1, I do not think the various toponyms starting in *Nor-*, *Nur-* etc. that have been compared to *nuráke*, are indicative of anything (cf. DES II: 176–178 and refs.). In conclusion, Battisti's (1955: 477) and Wagner's (DES II: 178-179) observation that the etymology of *nuráke* has not yet been satisfactorily clarified, still applies.

5.2.3 Pile of rocks

Reconstruction	**Forms**
****marganị̀one***	DES (II: 74): ***marganǧòni, mraganǧǫ̀i*** (Camp.) 'pile of rocks'

This word is formally identical with Camp. *marganǧòni* 'large eel' << Sic. *maraghiuni* 'male moray eel' (DES II: 74). This must be coincidental however. Wagner (DES II: 74) considers this word "probably pre-Roman" and compares several similar Sardinian toponyms, such as *Morgongiori* and *Margangioni*. I am not convinced by Wagner's (DES II: 72) suggestion that it is related to Camp. *marayòni* 'rock crevice' (§ 5.1.20), as there are significant semantic and formal discrepancies. On the basis of the current data, not much more can be said about this word.

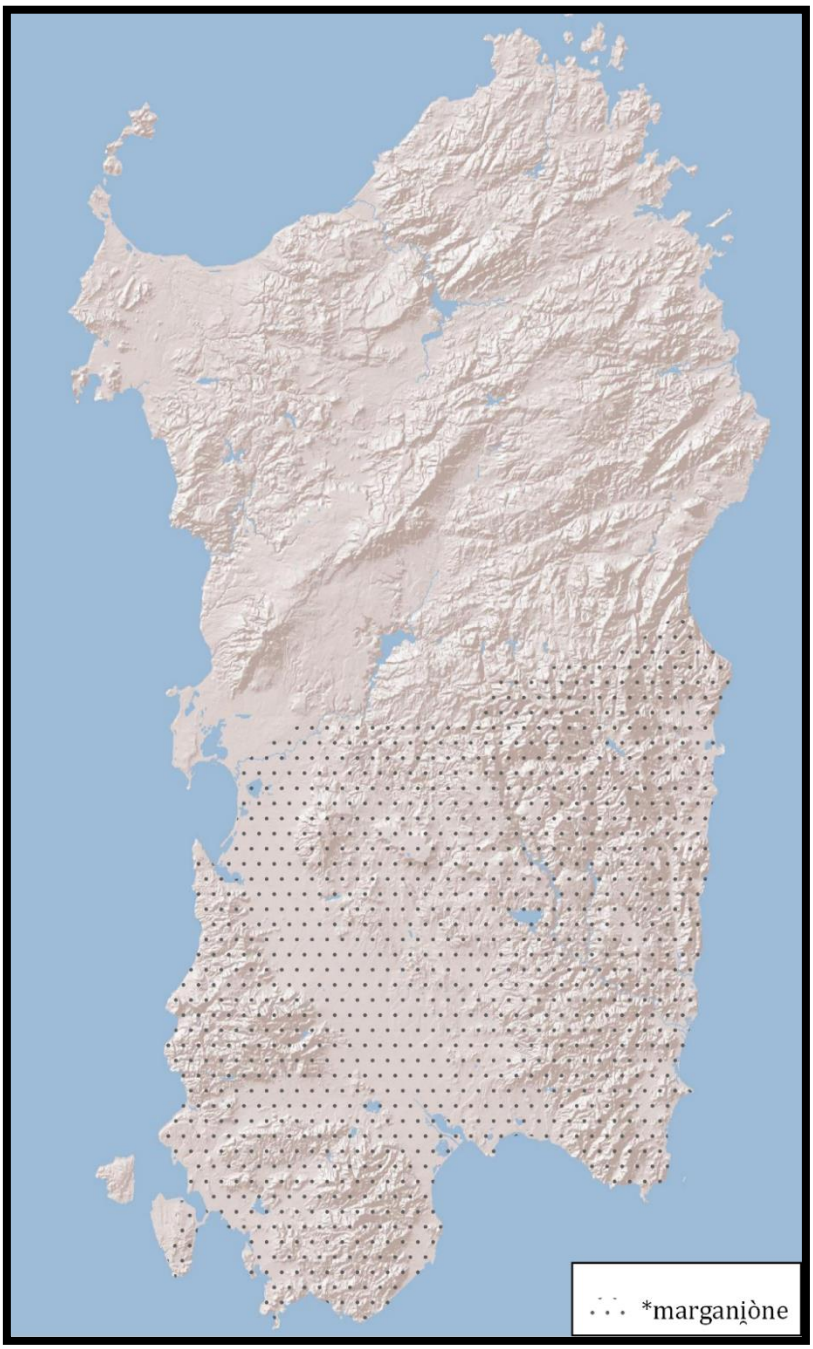

Figure 5.2.3: Pile of rocks

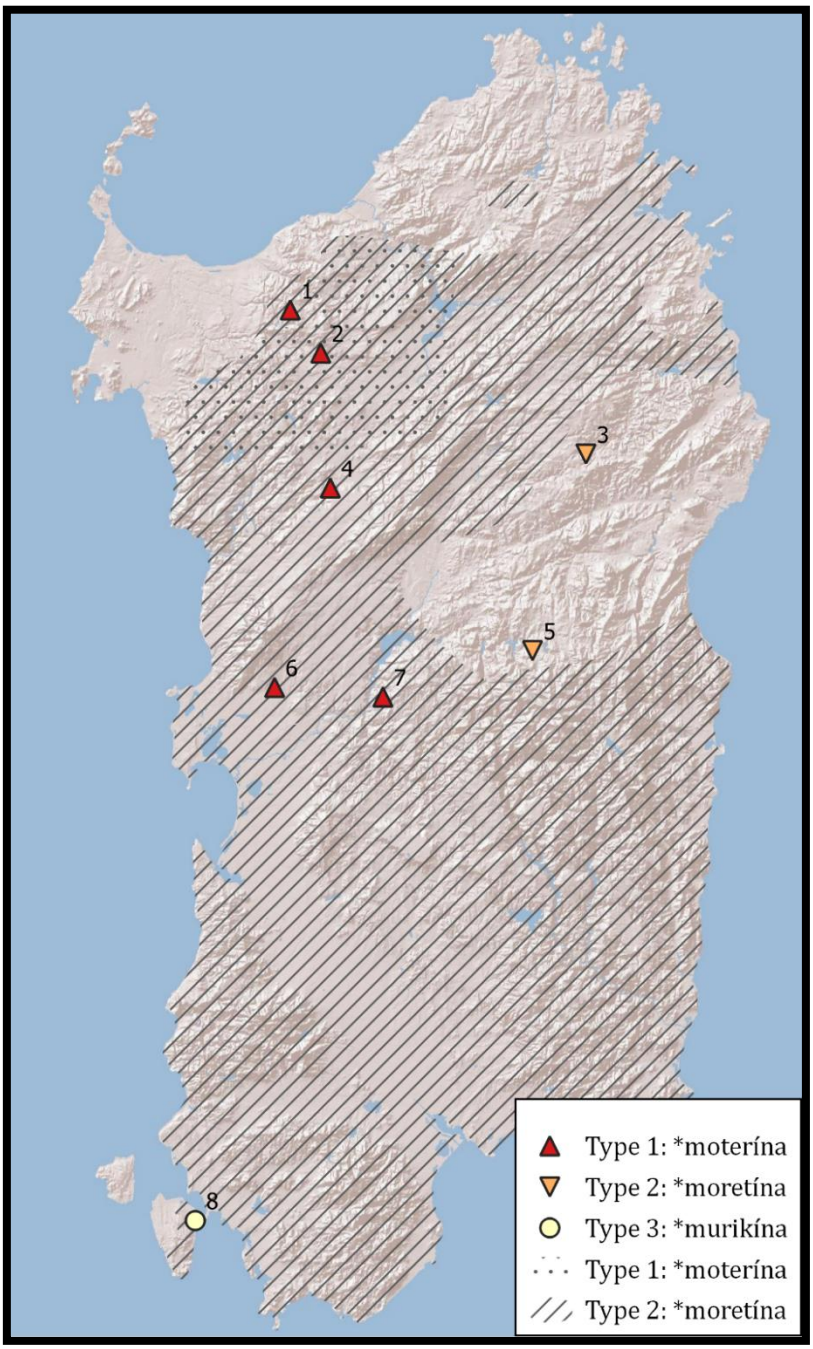

Figure 5.2.4: Pile of rocks; dry stone wall

5.2.4 Pile of rocks; dry stone wall

Reconstruction	Forms
1: *moterína	DES (II: 127–128): ***moðerína*** (NLog., 1: Osilo, 2: Ploaghe, 4: Bonorva), ***muðerĺ́a*** (6: Milis), ***muðerína*** (7: Busachi) 'pile of rocks, dry stone wall'
2: *moretína	DES (II: 127–128): ***moreðína, mureðína*** (Log., Camp.), ***moretína*** (3: Bitti), ***muretína*** (5: Fonni) 'id.'
	Paulis (1987: 515): ***moridína*** (Log., Camp.), ***muritína*** (Nuor.) 'id.'
3: *murikína	DES (II: 127–128): ***muriɣína*** [***de βèrða***] (8: Sant'Antioco) 'id.'

These words can probably all be derived from original **moretína* (type 2), as it is still found in Bitti (3). The shift of pretonic *-o-* > *-u-* could be due to influence from *múru* 'wall' (DES II: 128). The forms in type 1 show a metathesis, which is common in Sardinian (Wagner 1941a: 242–244). Wagner (DES II: 128) derives these forms from a hypothetical pre-Roman base **mora*, for which he adduces as evidence Cors. *mora*, *mola* 'pile of rocks or wood',[210] Tusc. *mora* 'id.', Laz. *mora* (Castelmadama) 'boulder, very big rock', Sp. *morón* 'small hillock', Pg. *moroiço*, *moiroiço* 'pile, mound' etc. (cf. Jud 1911: 11 fn. 2). The ending **-etína* found in the Sardinian forms could have been extended from Lat. *parietina* 'of walls; old fall-en-down wall', as Corominas (DCEC IV: 147 fn. 8) proposes. DCEC (IV: 146) also compares Sp. *morena*, Pg. *moreia* 'pile of plant material', Gal. *moreas* 'the little piles of rocks that are made on plots of land when they are being worked',[211] Frpr. *morêna* (Savoy), *moŕāina* (Vaux) 'bump which forms at the lower edge of a sloping field; glacier moraine', and Bq. *more*(*a*) (found in toponyms), *mora* 'cairn, pile of rocks; dolmen' (cf. Michelena 1969: 23–25). Hubschmid (1943: 150–151) and FEW (VI/III: 237b) want to derive all Romance forms of the type **mor-ena* from a base **mŭrr-* 'snout, muzzle', but the single *-r-* in many of the forms denoting 'pile' etc. as well as the *-o-* vocalism in Basque and in Sardinian make this rather problematic. I agree with Corominas' (DCEC IV: 146a) suggestion that the two word families may partly have been conflated by the rather common semantic development of 'nose, muzzle' to 'muzzle-shaped rock, rounded hill', as found for instance in Prov. *mourre* (DCEC IV: 146a).

Wagner's (DES II: 130) link to the Corsican and Tuscan forms is formally and semantically convincing, and we may add Bq. *more*(*a*), *mora* 'pile of rocks', as

[210] Also 'haystack', 'heap of wheat after threshing' and 'mound of earth' (BDLC: s.v. Cors. *mora*).

[211] "[L]os montoncitos de piedras que se hacen en las heredades cuando las trabajan."

well as the various Ibero-Romance and Gallo-Romance formations of the type **mor-ena* denoting artificial piles and/or natural piles like moraines. The most widely dispersed meaning is that of an artificial stacking of rocks. Thus, while a derivation from Lat. *mūrus* would seem attractive for the Sardinian forms like Log., Camp. *mureðína*, the difficulty of explaining *-o-* in the other forms, as well as the existence of semantically and formally very similar formations requiring **mora* in other Romance languages, makes a shared pre-Roman origin of these forms more probable. It is nevertheless important to be aware of chance similarities in such a phonologically simple base as this one, the risk of which is clearly pointed out by Corominas (DCEC IV: 146) for Sp. *morón* 'hillock',[212] and of which unrelated Lat. *mūrus* 'wall' < PIE **(H)moi-ro-* (de Vaan 2008: 396) is an example as well.

5.2.5 Source

Reconstruction	**Forms**
1: *míϑϑa	DES (II: 120): ***míttsa*** (Camp.), ***míntsa*** (3: Gesturi, 5: Escalaplano, 6: Siurgus, 7: Suelli), ***míϑϑa***[213] (9: Villaputzu) 'source, spring'
	AIS (852): ***míttsa*** (1: Milis, 2: Busachi, 4: Mogoro, 8: Villacidro), ***mítsa*** (10: Cagliari), ***míća*** [***i akwa***] 'id.'

This word was originally proposed by Guarnerio (1906: 245) to continue **mītia* 'wilted' << Lat. *mītis* 'ripe, mellow etc.', which is widespread in Romance (cf. REW 5614). Wagner (1997: 159; 1957: 105–106; DES II: 121) expresses his doubts about this etymology on the basis of the semantic distance between 'wilted' and 'water source', and instead proposed a derivation from Pun. **miṣṣā*, which would be a deverbal noun from the root **yṣ'* 'to exit'. This etymology was discussed from a Semitist point of view by Friedrich (1957: 223), who argued that a more likely Punic source form would be the cognate to Hebr. מוֹצָא *mōṣā* 'exit (of water), from the same root **yṣ'* 'to exit'. According to Friedrich (1957: 223), the *-i-* vocalism of Srd. *míttsa* could reflect raising and fronting of original **mōṣā* to **mūṣā* > **mṻṣō*, supposedly attested as *mysehi* in Plautus' *Poenulus* (931) (cf. Argiolas 2010: 27). This specific form however is analyzed by Krahmalkov (2000: 303; 2001: 159) as *mysethi* 'I have come'.[214]

[212] DCEC (IV: 146a) explains this as a back-formation from *desmoronar* 'to crumble'.

[213] The dialect of Villaputzu, in the Sarrabus, does not have *-ϑ-* as a phoneme (Contini 1987: 560). This form should probably be *míttsa* too.

[214] For yet other interpretations, see Mensching (1997: 295 fn. 10).

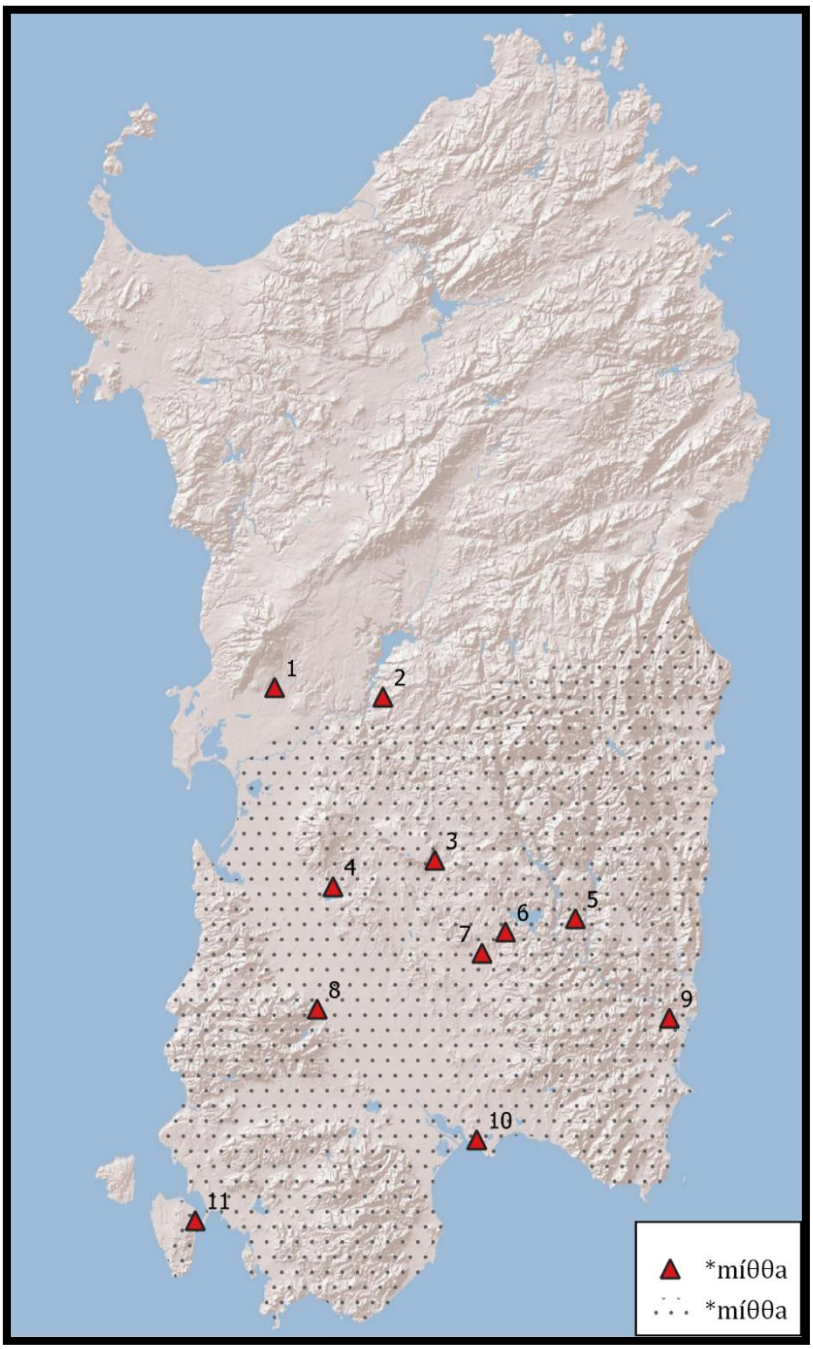

Figure 5.2.5: Source

On the basis of the uncertainties regarding the reading of the form *mysehi* and phonology of potential Punic source word, Mensching (1997) returns to Guarnerio's original etymology, and shows parallels within Romance for the use of this very etymon in the meaning 'wet, swampy area' or even 'pond' (Mensching 1997: 298–301). It thus appears that a Punic origin is not necessary after all.

6 Culture and economy

This chapter contains words referring to human activity. This includes agricultural terms (§ 6.1), and more general material and cultural concepts (§ 6.2).

6.1 Agriculture

6.1.1 Bran

Reconstruction	Forms
1: *ϑelá(*C*)*u*	DES (II: 541): ***teláu*** (1: Luras, 2: Monti, 3: Berchidda, 4: Oschiri, 6: Pattada), ***taláu*** (7: Buddusò) 'bran'
2: *ϑála(*C*)*u*	DES (II: 541): ***ϑálau*** (5: Posada, 8: Siniscola, 9: Orosei, 10: Dorgali) 'id.'

In this word, we find two different stress patterns. One has the stress on the penultimate syllable (type 1), the other on the antepenultimate (type 2). Wagner (DES II: 541) proposes a pre-Roman origin for this word, which is plausible in light of initial **ϑ-* and the impossibility to derive this word from a Latin source. For the ending *-au*, we need to assume the loss of an original intervocalic consonant, as Lat. *-au-* regularly yields *-a-* in Sardinian (Wagner 1941a: 14). Type 2 most likely represents the original stress pattern, as there are several examples of final *-au* secondarily receiving the stress by analogy to other words ending in stressed *-áu*; e.g. Log. *kannáu* < *kánnau* (both are attested) < Lat. *cannabu* < *cannabis* (Wagner 1941a: 2–3). Wagner (1941a: 115; DES II:541) tentatively considers a link between these Sardinian forms, and Bq. *zalauts* 'tanbark', *za(h)i* 'bran' and *zalki* 'dandruff'.[215] These forms are derived from a root **zal-* (Santano Moreno 2000: 168), or **azal-* (Lakarra et al. 2019: 57–59) 'skin, bark, surface'. The only Basque word directly comparable to the Sardinian forms is *zalauts* 'tanbark', which is probably a compound of *azal* 'skin' and *hauts* 'dust' (Lakarra et al. 2019: 58). In my view, the formal and semantic discrepancy between the Sardinian and the Basque forms is too great to confidently posit a shared origin.

[215] I have not been able to verify Bq. *zalauts*. In EH (638) 'dandruff' is given as *zalgi*, rather than *zalki*. This does not influence the argument however.

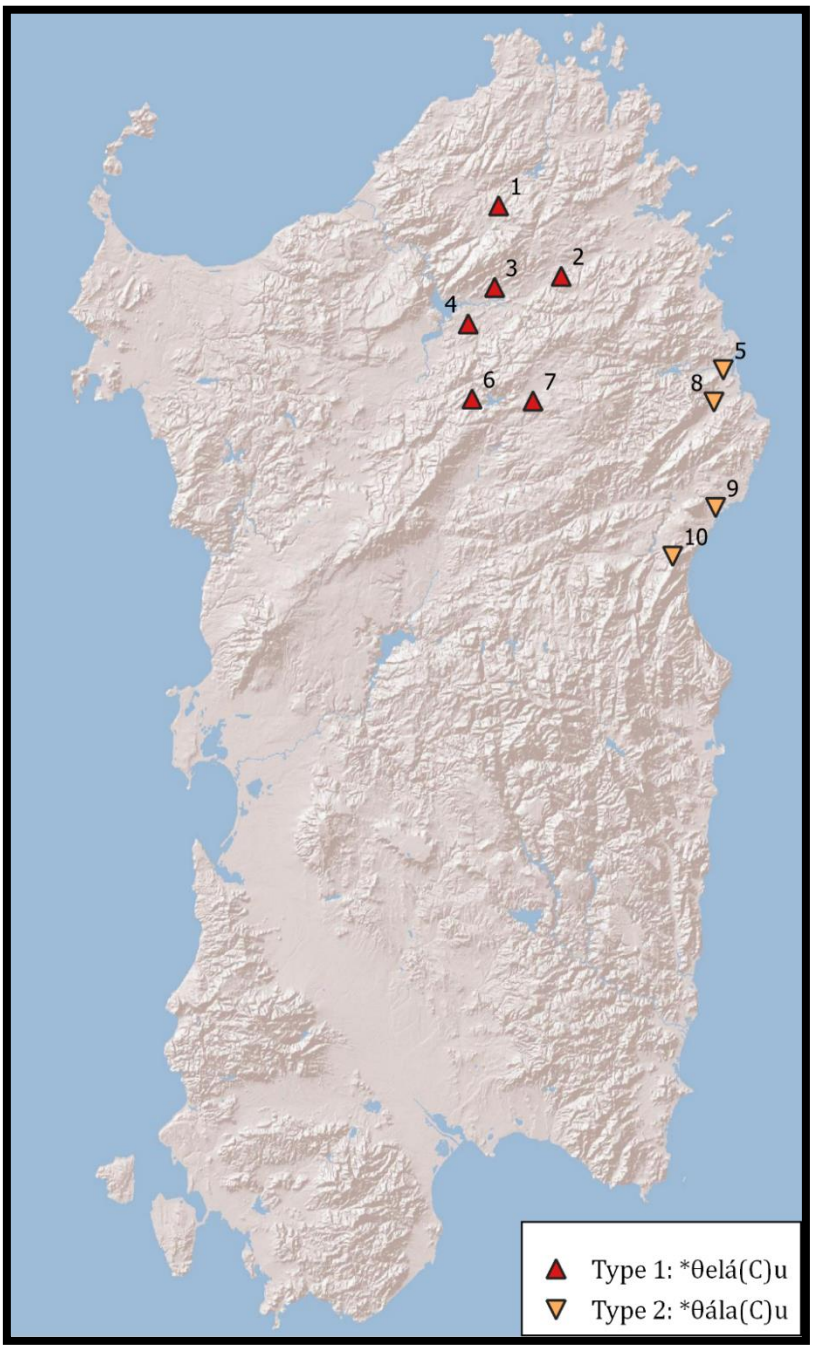

Figure 6.1.1: Bran

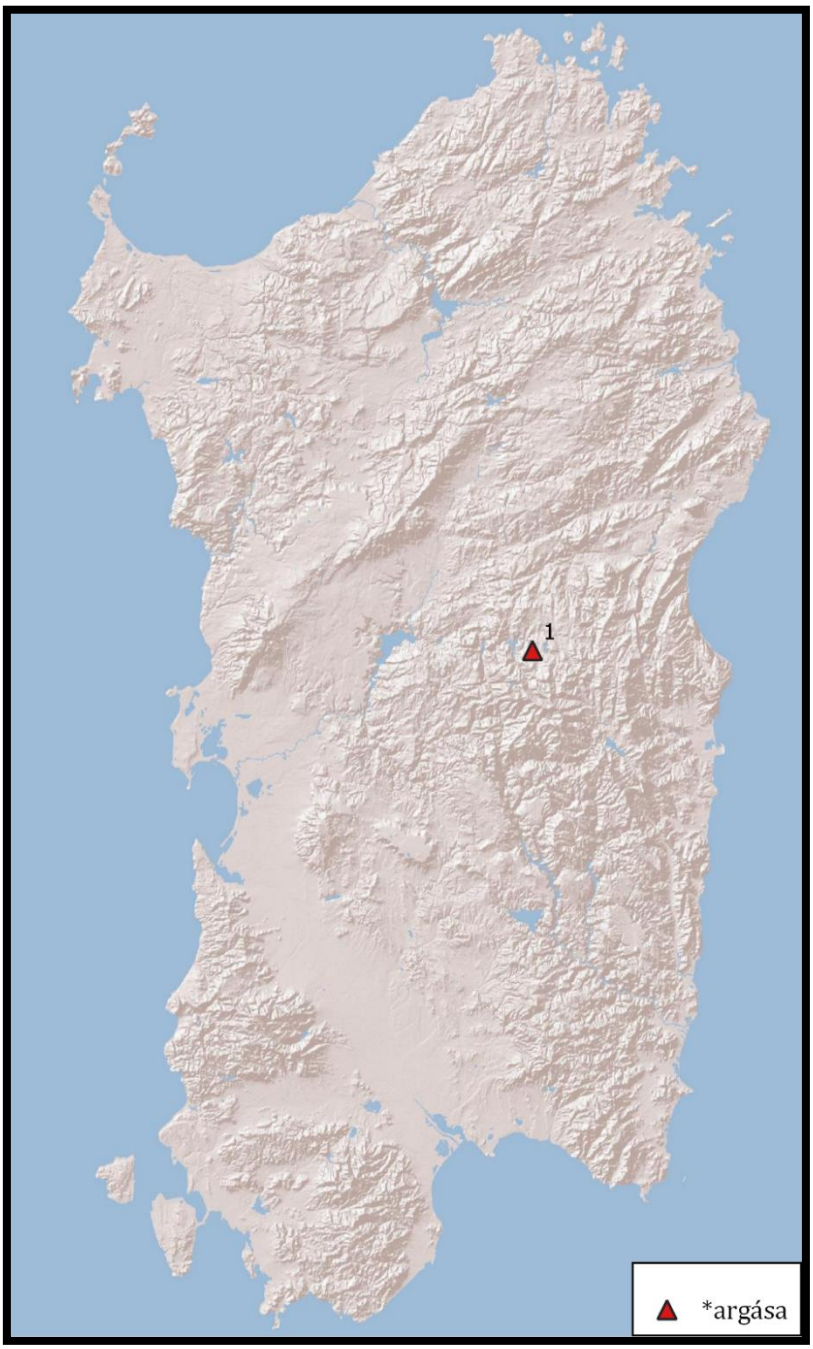

Figure 6.1.2: Goat or sheep that lost its young

6.1.2 Goat or sheep that lost its young

Reconstruction	**Forms**
***argása**	DES (I: 110): ***argáša*** (1: Fonni)[216] 'goat or sheep that has lost its young'; deriv.: ***argašare*** (1: Fonni) 'remove the young from goats or sheep'

This word is only attested in Fonni. Wagner (DES I: 110) proposes a pre-Roman origin. On the basis of a single occurrence this is hard to substantiate however. Similar in form and meaning are forms like *argáğğu* 'one-year old goat that has not had young yet' (Urzulei), *argárğu* 'id.' (Desulo), but these go back to Lat. *gregārius* 'of the flock' (DES I: 589). This is impossible for *argáša* though, as Lat. *gregārius* is reflected in Fonni as *grayárğu*, *-a* 'one year old goat' (DES I: 589).

6.1.3 Herd

Reconstruction	**Forms**
1: *grústu	DES (I: 595): ***grústu*** (Nuor., 4: Fonni, 5: Desulo), ***rustólu*** (1: Bitti), ***vrústu*** (3: Orani) 'herd (of mouflons, goats, sheep, pigs); group (of people)'
2: *grústi(C)u	DES (I: 595): ***grústi̯u*** (2: Nuoro) 'id.'
3: *rustía	Casu (2011: s.v. *rustìa*): ***rustìa*** (Log.) 'group of children, especially poor ones'

The forms in type 1 can all be derived from *grústu*, with a suffix *-ólu* in Bitti (1), and with replacement of initial *g-* with *v-* in Orani (3) due to reanalysis of the onset in lenited contexts (cf. Wagner 1941a: 207). The ending *-i̯u* in type 2 is the past participle suffix. Wagner (DES I: 595) considers these forms to be of "probable pre-Roman" origin. The evidence for this is mainly the lack of an etymology. Because of the semantic distance, I am not convinced by his connection of this word to *kurústa*, *rústa* 'bed bug, vermin, fox etc.' (§ 4.1.3). The word *rustìa* 'group of children, especially poor ones', mentioned by Casu (2011) may perhaps rather be a derivation of *rústa* in the sense 'vermin' (§ 4.1.3).

[216] This form is actually cited in DES as *argáśa*, in which *-ś-* would be an alveolar sibilant (DES I: 33). However, with the exception of this specific lemma, Wagner consistently uses *-š-* instead of *-ś-* to represent the alveolar realization of sibilants in Fonni.

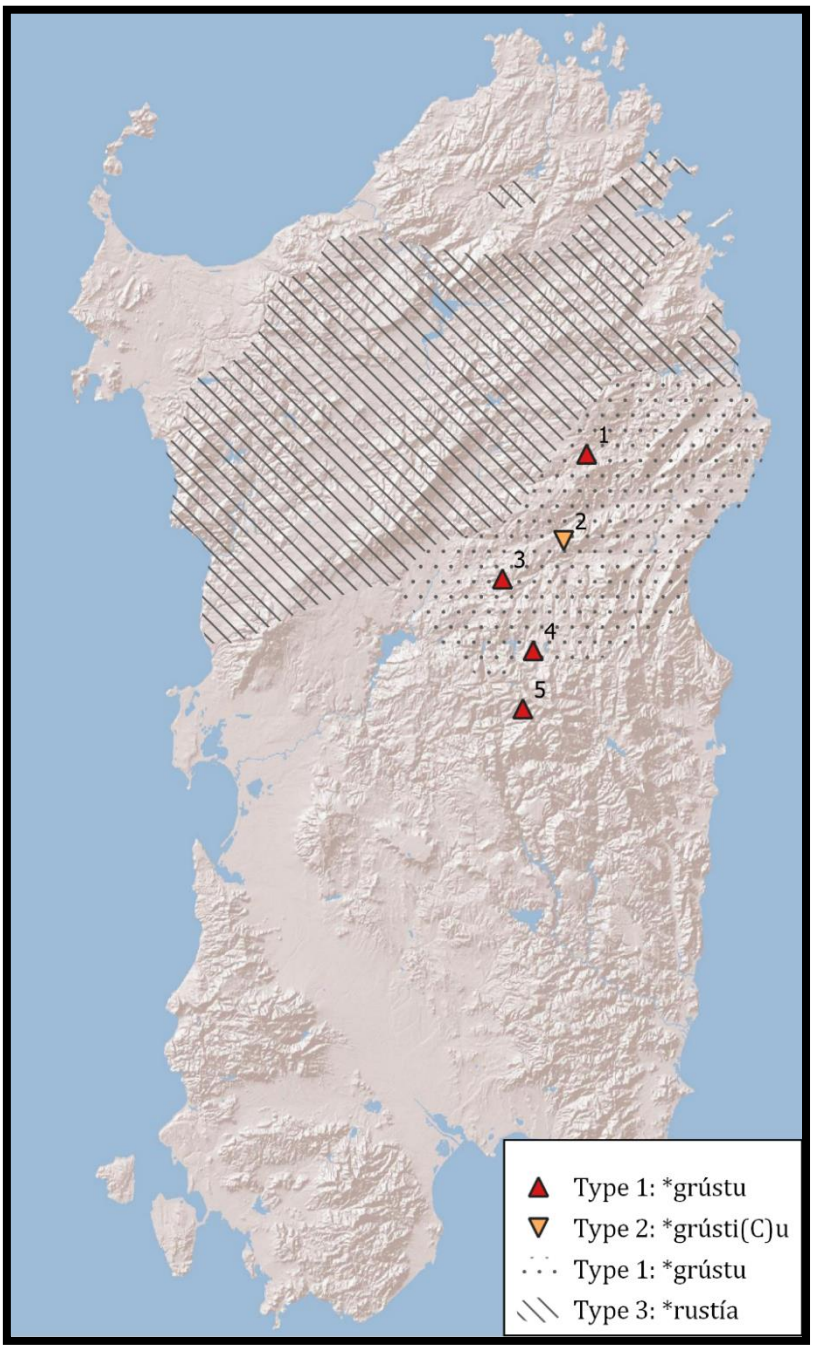

Figure 6.1.3: Herd

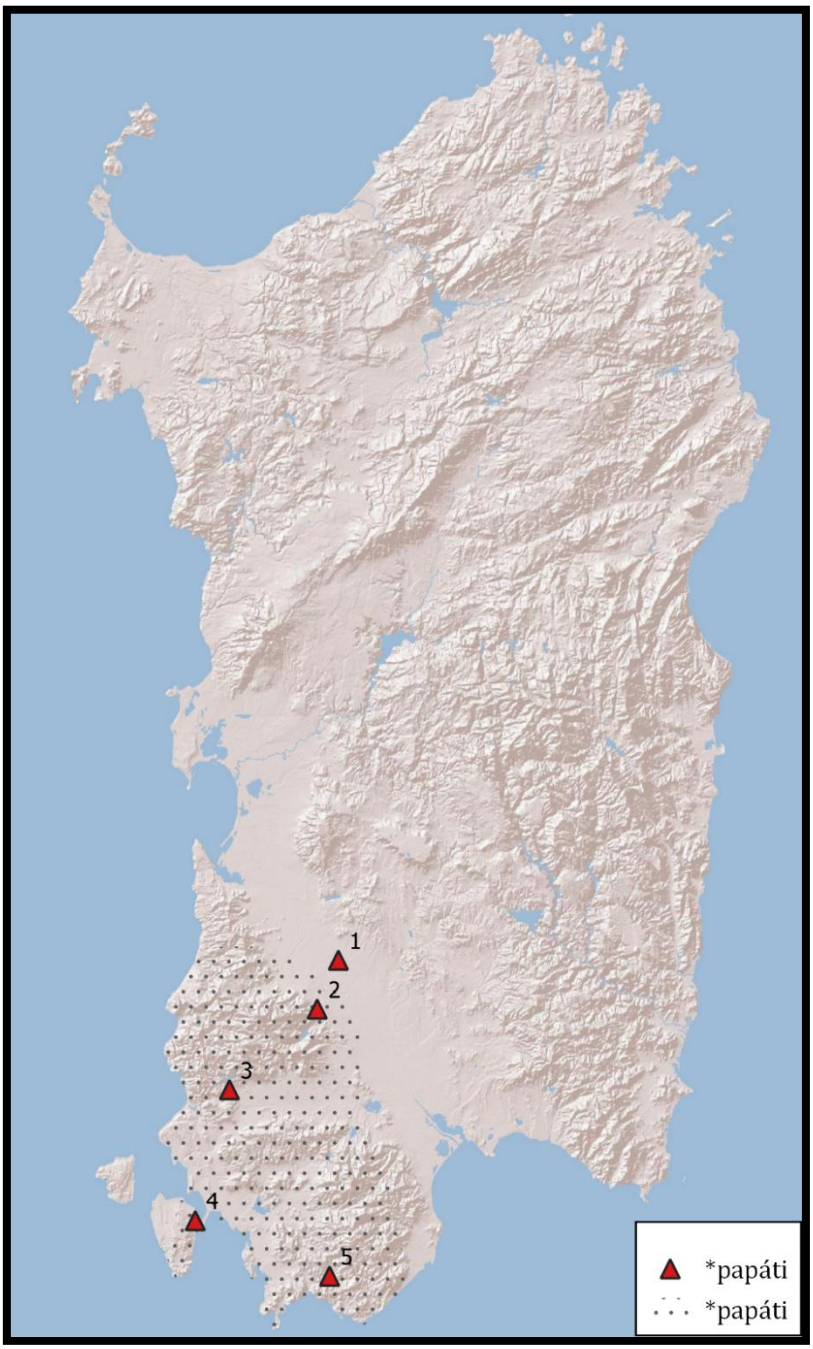

Figure 6.1.4: Hornless (of goats)

6.1.4 Hornless (of goats)

Reconstruction	**Forms**
***papáti**	DES (II: 200): ***paβáða*** (1: San Gavino Monreale), ***paβáði*** (Sulcis reg., Iglesiente reg., 2: Villacidro, 4: Teulada), ***paβári*** (3: Iglesias, 4: Sant'Antioco) 'hornless (of goats)'

This adjective, that is used for hornless goats, is found in the Sulcis and Iglesiente regions in the southwest of Sardinia. The ending *-i* in *paβáði* etc. could go back either to original *-*i* or *-*e*. It is uncertain whether the final *-a* in *paβáða* (1: San Gavino Monreale) is secondary due to its reinterpretation as a feminine participle, or whether the ending *-i* was added secondarily (cf. Wagner 1941a: 218–219).

Wagner (DES II: 200) rightly rejects Serra's (1954b: 219) proposed derivation from **vacativus* on formal grounds. Moreover, the actual reflex of **vacativus* is attested as *baɣaðíu* 'bachelor(ette), unmarried'. Serra (1954b: 219) cites the forms *baβáðia*, *baβaðía*, but Wagner (DES II: 200) questions their reality, and they are not included in Puddu (2023) either.

As such, the forms discussed are of unknown origin, and seem to go back to **papate/i*. The matter is complicated by the fact that similar formations are found in southernmost Calabria, e.g. Calab. [*crapa*] *papáttsa* (Cataforio, Melito di Porto Salvo, San Pantaleone), Greco-Calab. *popáttso* (Ghorio di Roghudi), Abruzz. [*la krápa*] *kəkǫ́ćća*, *pəkǫ́ttsa* (Scanno) etc. 'hornless' (AIS 1082; Rohlfs 1932: 120). It is difficult to establish the relationship between the Sardinian and the southern Italian forms, as the latter appear to lack any cognates elsewhere. If the Sardinian and south Italian forms are to be connected, they perhaps share a pre-Roman origin. Alternatively, one wonder whether we could be dealing with independent expressive formations.

6.1.5 Sheep (old, skinny)

Reconstruction	**Forms**
1: *čúrra	DES (II: 602): ***tsúrra*** (Marghine reg., Planargia reg., 1: Padria), ***čúrra*** (2: Fonni), ***giurrítta***[217] (Log.) 'old, skinny sheep'

[217] In Spano's handwritten notes (*non vidi*) on his own dictionary (i.e. Spano 1851).

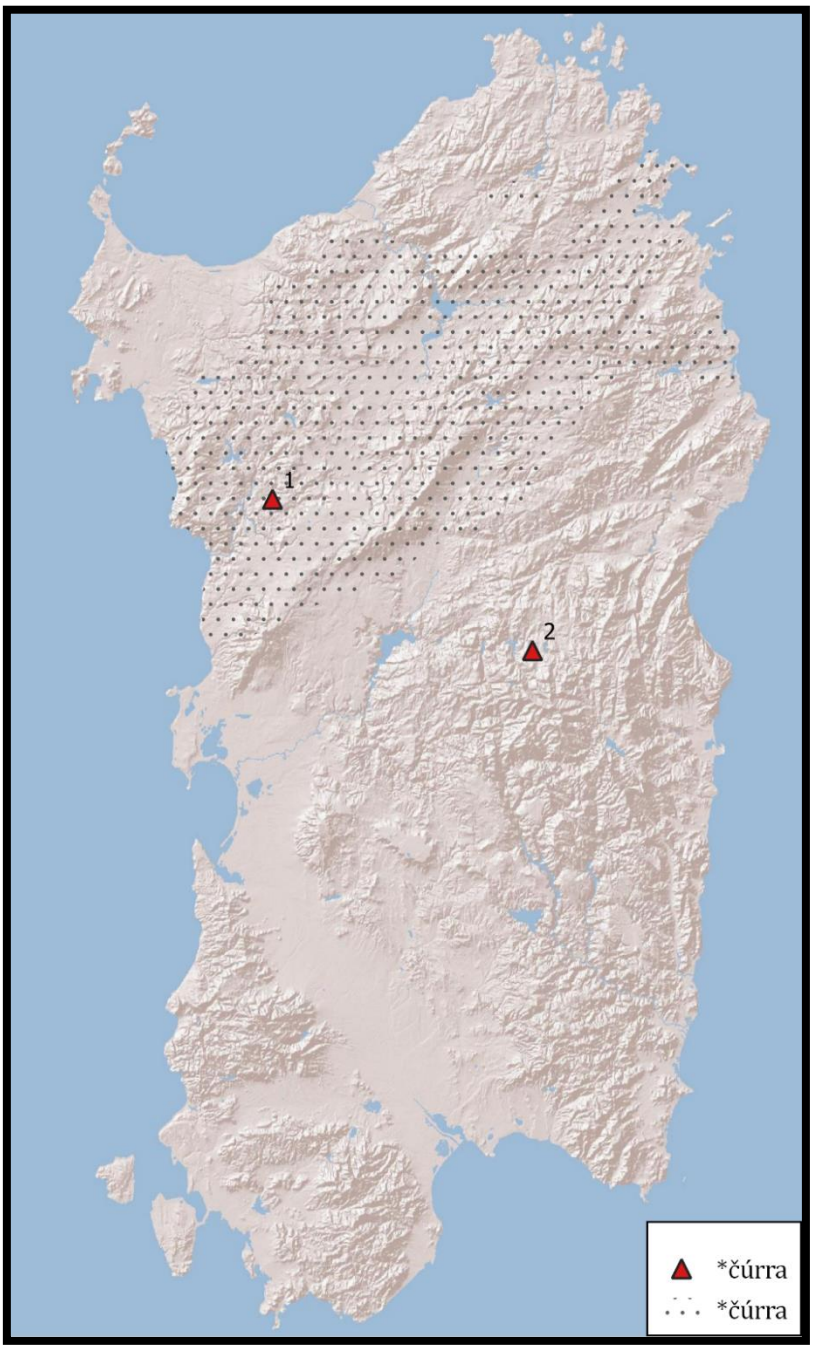

Figure 6.1.5: Old, skinny sheep

These Sardinian forms contain initial *ts-* and *č-*, which do not have a clear inherited source in the Logudorese dialects in which this word is found (§ 8.1.4). Jud (1937: 312) compares Trent. *ciorla* 'skinny cow', Lomb. *ciorla* (Como) 'cow of little value', Calab. *zurra* 'hermaphrodite goat'. Hubschmid (1950c: 14–15) adds Lad. *čorla* 'skinny cow' (Val di Fassa), Abruz., Nap., North Pugl. *tsurrə* 'he-goat' (cf. AIS 1080), (modern) Umbrian *ciorro* (Amelia) 'ram' (cf. AIS 1069), and MGr. τζοῦρος (Cyprus) 'he-goat'.[218] All of these words refer to livestock animals that are somehow of reduced value, be it by being of poor health or by being infertile. In DES (II: 602), these forms are compared to Sp. *churra* 'sheep with *churra* wool'. This adjective belongs to a much larger group of Ibero-Romance forms referring to such negative notions as 'filth, dirty, greedy etc.' (DCEC II: 409-12).

Whereas the connection between Srd. *tsúrra*, *čúrra* and Sp. *churra* is attributed to a shared pre-Roman origin in DES (II: 602), I see no reason why the Sardinian forms could not simply represent loans from Spanish (cf. Virdis 2017). This could also account for the irregular correspondence of initial *ts-* and *č-* in the Sardinian forms.

6.1.6 Sheep (old)

Reconstruction	**Forms**
***bòlla**	DES (I: 215): ***bòḍḍa*** (Barb., 1: Orani) 'old sheep'

This word can go back either to **bolla* or to **volla*, neither of which are explained in terms of inherited material. Wagner does not propose an etymology for this word, but does compare Calab. [*crapa*] *uoḍḍa* (Pentedattilo) 'goat whose horns curve behind its ears' (Rohlfs 1932: 351). There are two Berber etyma that are comparable in shape and meaning. One is Bb. *ullī*, *wəllī* 'small livestock (i.e. sheep and goats)' < **wV̆llī*/**wV̆llē* (Haddadou 2007: 85; cf. Kossmann 2020a: 28). The other is Ghadames *taβali*, Mali Tuar. *tehăle*, Nef. *tili* etc. 'sheep' < **tē-ḫălē* (Kossmann 1999: 97).[219] It is difficult to decide which Berber etymon offers the better comparison. It is moreover unlikely that the Sardinian word is related to both the Calabrese and the Berber forms.

[218] For additional forms, see Rohlfs (1925: 673). Rohlfs' explanation involves the lexicalisation of a call for goats of the shape *zurr-*. However, it is hard to imagine that animal calls referring to infertile/worthless lifestock arose parallelly in multiple areas and with such formal similarity.

[219] The reconstructions of both Berber etyma are by Maarten Kossmann, p.c.

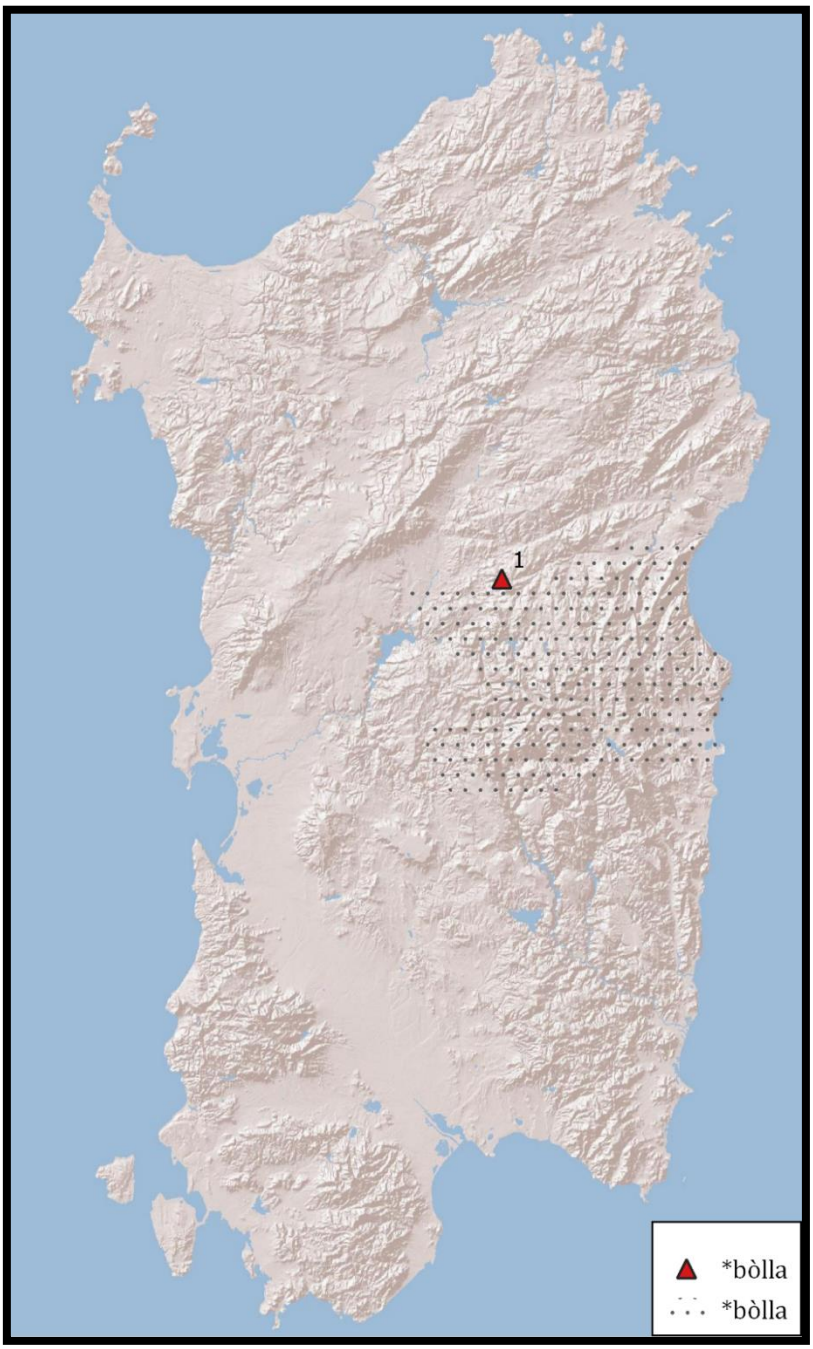

Figure 6.1.6: Old sheep

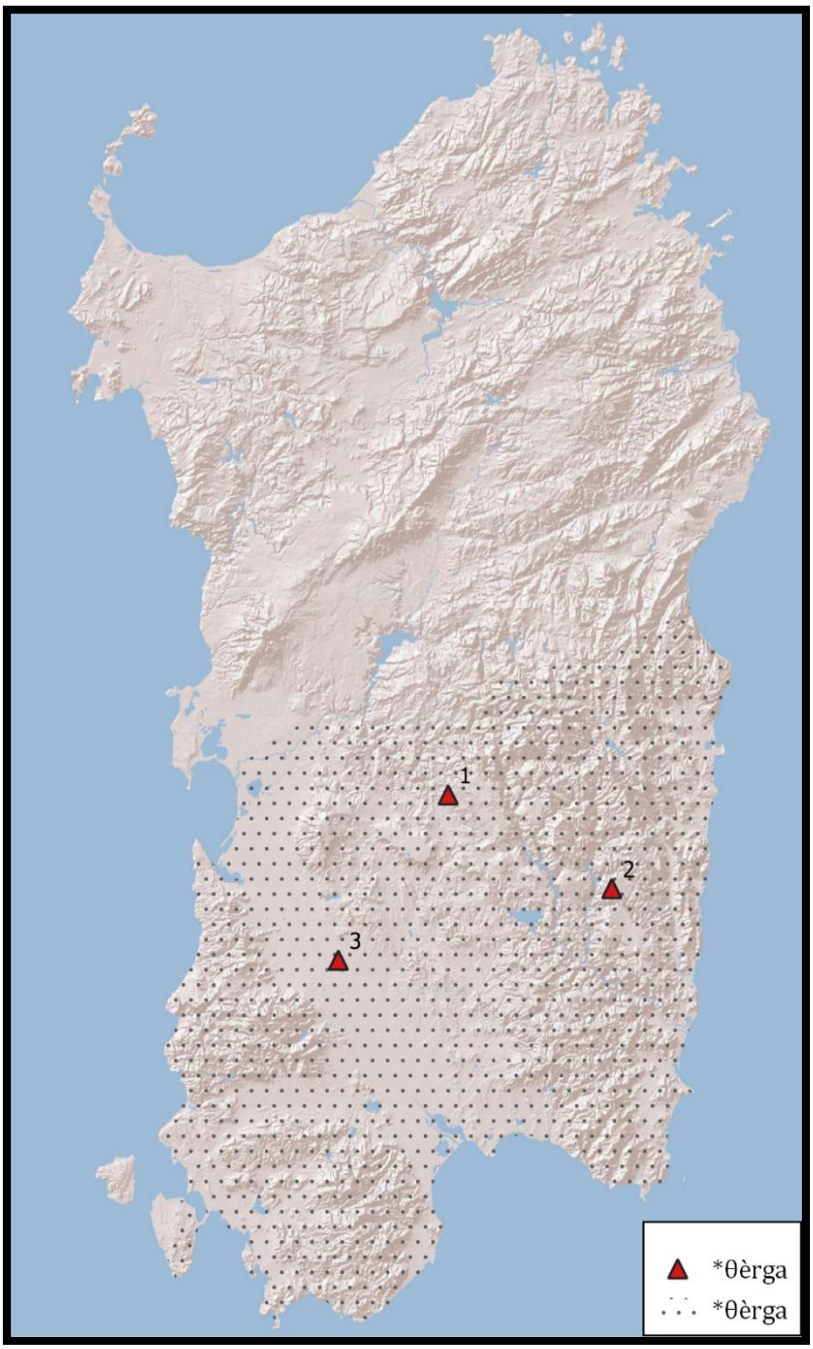

Figure 6.1.7: Stack of wheat

6.1.7 Stack of wheat

Reconstruction	Forms
ϑèrga	DES (II: 588): ***tsèrga*** (Camp.), ***tsrèɣa*** (1: Laconi, 3: San Gavino Monreale), ***sèrga*** (2: Perdasdefogu) 'stack of wheat'

Wagner (DES II: 588) does not offer an etymology for these forms, which go back regularly to **ϑèrga*. He does mention Srd. *sèrra* 'saw', in the construction *sèrra de ðríɣu* 'pile of grain' (cf. DES II: 408-9), but a link between *tsèrga* and *sèrra* is phonologically impossible. In my view, initial *tse-* < **ϑe-* makes an inherited origin unlikely. As a consequence, a pre-Roman origin is a possibility, but besides the onset there is not much direct evidence. Perhaps it is the same word as *ϑèrga, tsèrga* 'cloths that a master gives to his servants on holidays or at the end of the year' (§ 6.2.3).

6.1.8 Tilled land

Reconstruction	Forms
1: **tèle*	DES (II: 481): ***tèle*** (1: Lodè) 'scrub'; (2: Bitti, 3: Lollove, 4: Nuoro, 5: Dorgali) 'steep and wooded terrain, that has recently been tilled for the cultivation of grain'; (5: Dorgali) 'field sown with oats, beans or flax'.[220]
2: **tevèle*	DES (II: 481): ***tevèle*** (6: Oliena, 7: Orgosolo) 'steep and wooded terrain, that has recently been tilled for the cultivation of grain'

Despite some differences in meaning, this word refers mostly to a patch of land that has recently been prepared (i.e. either tilled or sown) for the cultivation of crops. Spano (*apud* DES II: 481) mentions *tuvèle* 'threshing floor, sown field', but without specifying its place of attestation. Wagner (DES II: 481) thinks a pre-Roman origin is likely on the basis of the ending *-èle*, which he identifies as a suffix that also occurs in toponyms like *Sarunele* (a nuraghe), *Gabutele* (a mountain), *Neoneli* (a town)(cf. Wolf 1998a: 64). However, there is no direct evidence that *-èle* is a suffix in *tèle*, *tevèle* 'tilled land', nor do we have any clues as to what the function of this suffix would be. Serra (1956: 203) suggests that *tevèle* could be cognate to Log., Camp. *túla, tulèḍḍa* 'furrow' < Lat. *tubula* (REW 8968), with secondary introduction of *-èle* by analogy to pre-Roman toponyms. This is rather far-fetched, and does not explain the variant *tèle*, as Lat. *tubula* is expected to yield ***túβula* in the dialects where type 1 *tèle* is found.

[220] "Campo seminato a biada, o fave o lino."

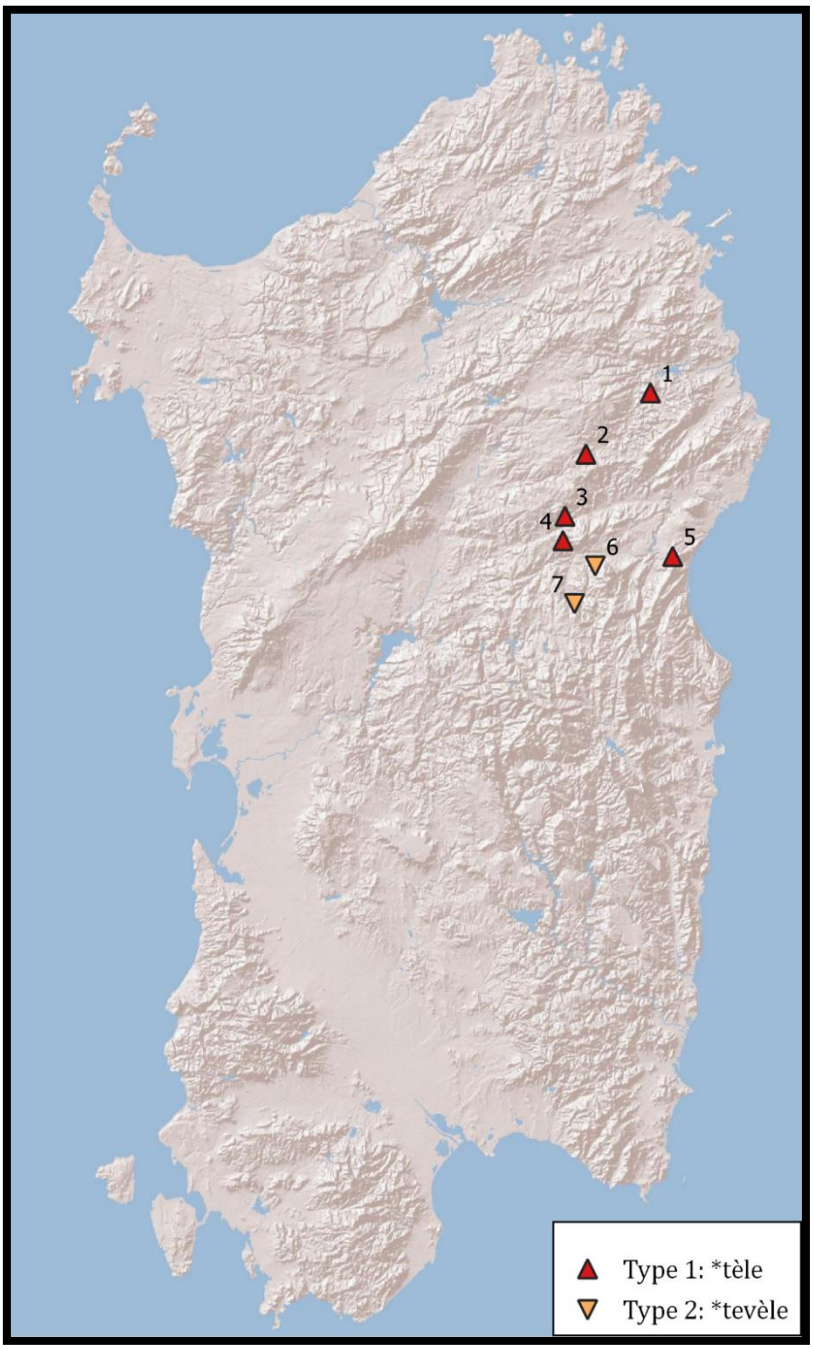

Figure 6.1.8: Tilled land

Another indication of pre-Roman origin is the apparent irregular correspondence between the attested forms. For type 2 we find *tevèle*, with an intervocalic *-v-*, which is absent in the forms of type 1. In DES (II: 481) this is explained as an instance of irregular and hypercorrect *-v-*deletion in words where the Campidanese dialects do have an intervocalic *-v-* (cf. Wagner 1941a: 100–101). However, it is not clear from Wagner's description of this phenomenon that the geographical extent of this phenomenon extends to the dialects of Bitti (2) and Nuoro (4). Moreover, Paulis (1984b: 533) points out that this word is not attested in Campidanese. Nor can the loss of *-v-* in type 1 be attributed to influence from Logudorese forms, as it is not attested there either.

Paulis (1992: 305) compares the Sardinian forms to It. *debbio* 'slash and burn', which would be from ancient Ligurian **debelo-* < PIE **d*^{h}*eg*^{wh}- 'to burn'. While the semantic and formal similarity between the Sardinian and Italian forms is undeniable, a hypothetical form in an unattested language of uncertain affiliation does not constitute an acceptable etymology for the Italian or the Sardinian forms (cf. § 1.3.2.1).

6.2 Material culture and food

6.2.1 Bread made of acorns and clay

Reconstruction	**Forms**
****ispéli***	Wagner v: [***pan'***] ***ispéli*** (1: Urzulei, 2: Talana, 3: Triei, 4: Baunei, 5: Villagrande Strisaili, 6: Gairo) 'bread made of acorn flour and clay'

This word, restricted to the northern part of the Ogliastra, refers to a highly specific and peculiar type of bread that is made of acorn flour paste mixed with clay. Its consumption has declined to the point that it is now restricted to special festivities, and its production process is only remembered by a few elderly women (Pinna 2013: S72).[221] Whereas it may have originated as a "poor man's" substitute of cereal flour, as Wagner (1921: 60) suggests, Pinna (2013: S73) notes that its production has come to be associated with certain ritual and religious practices.

[221] According to Pinna (2013: S72), the existence of this pastry may once have been more widespread than the Ogliastran villages cited by Wagner (1921: 60) where it is still remembered. However, even if *pan' ispéli* was once known in towns like Orani and Sarule in the Barbagia, the local terminology associated with this bread has been lost to us.

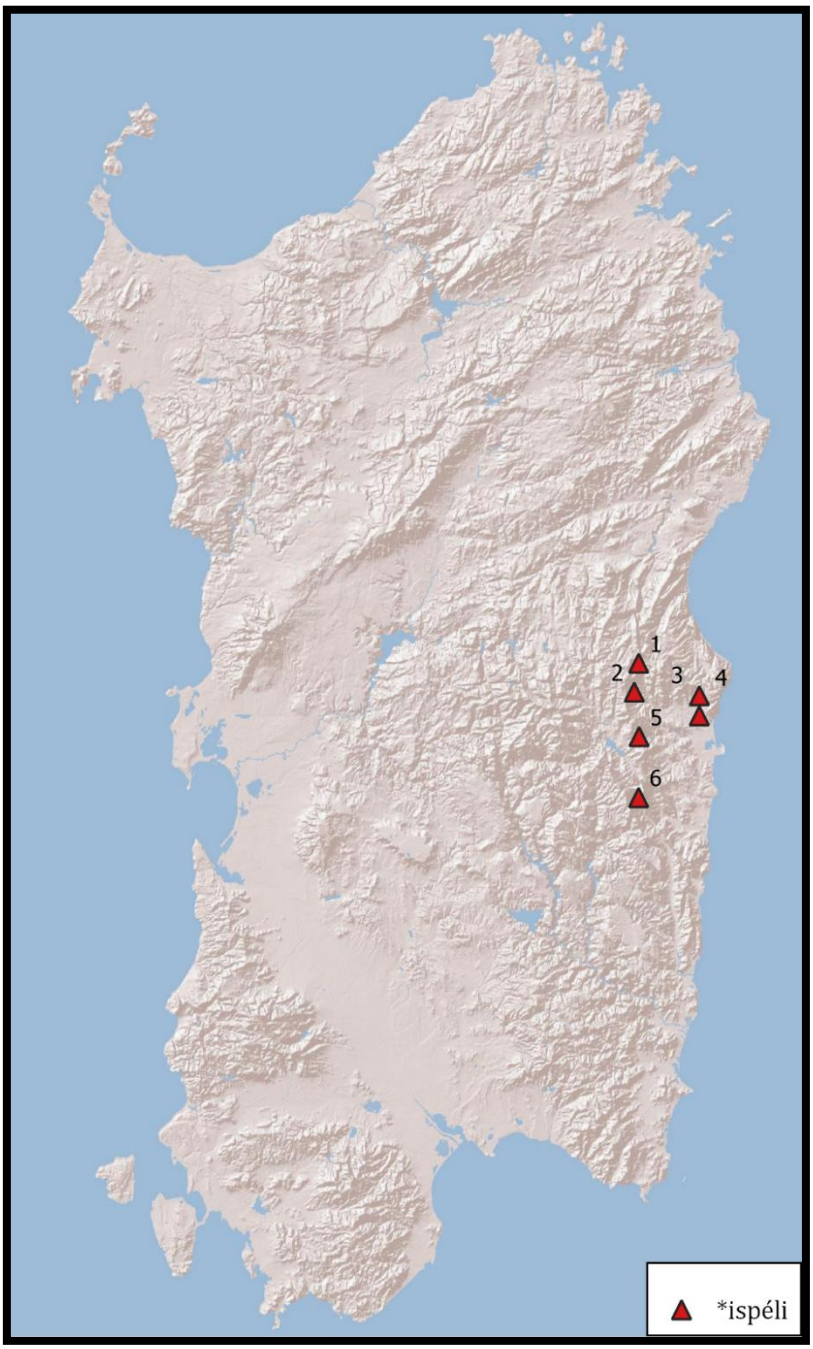

Figure 6.2.1: Bread made of acorns and clay

Wagner (DES I: 676) considers the word *ispéli* to be of "probable pre-Roman" origin, in part because of its semantics, being highly specific to local culture, and because the consumption of acorns is considered an "ancient" and "primitive" practice (cf. Wagner 1921: 60). Although objective notions such as "primitivity" do not substantiate a pre-Roman origin, the cultural specificity of this word and its relation to certain rituals may indeed be indicators of an older local tradition. A formal feature indicating a pre-Roman origin is word-final *-i*, which is rare in the inherited nominal and adjectival lexicon (§ 8.1.5.1). On the basis of this we may conclude that *ispéli* is likely the original shape of this word. Although some of the locations where *ispéli* is attested exhibit regular (or at least frequent) *i*-prothesis in word-initial clusters of the type *sC-* (Baunei, Triei, Urzulei, Talana), this process is exceptional in Villagrande Strisaili and absent in Gairo, indicating that initial *i-* is original (cf. Wagner 1984: map 2).

It is noteworthy that the bread is referred to as *pan' ispéli*, i.e. '*ispéli* bread'. The form *ispéli* therefore appears to denote a certain quality of the bread, but it is impossible to establish whether it refers to acorns, to clay or to some other feature of the bread altogether. Paulis (2011; cf. Putzu 2019: 4) notes that other types of pastry are called *pistíḍḍu* < Lat. *pistillum* 'pestle'.[222] Paulis hypothesizes that pre-Roman *ispéli* referred to a 'pestle' as well. From there, he compares *ispéli* to Bq. *ezpel* 'box tree', based on the idea that pestles are frequently made of box wood. This comparison is problematic. It requires the double assumption that both Srd. *ispéli* and Bq. *ezpel* originally meant 'pestle', rather than 'clay bread' and 'box tree' respectively, which significantly weakens the comparison. Moreover, Bq. *ezpel* 'box tree' probably contains the element *ez-*, which occurs as the first part of various tree names, and **-bel* 'dark' (Trask 1997: 256; 2008: 188). The etymology of Bq. *ezpel* as originally meaning 'dark tree' renders any connection to pestles, and therefore to Srd. *ispéli*, unlikely.

The specific type of clay that is used to make *pan' ispéli* is called *trókko* or *tòrko* (DES II: 522), which may also be of pre-Roman origin (§ 6.2.2). The other terms related to the production of *pan' ispéli* cited by Pinna (2013) are inherited.

[222] I have unfortunately not been able to consult Paulis (2011) myself.

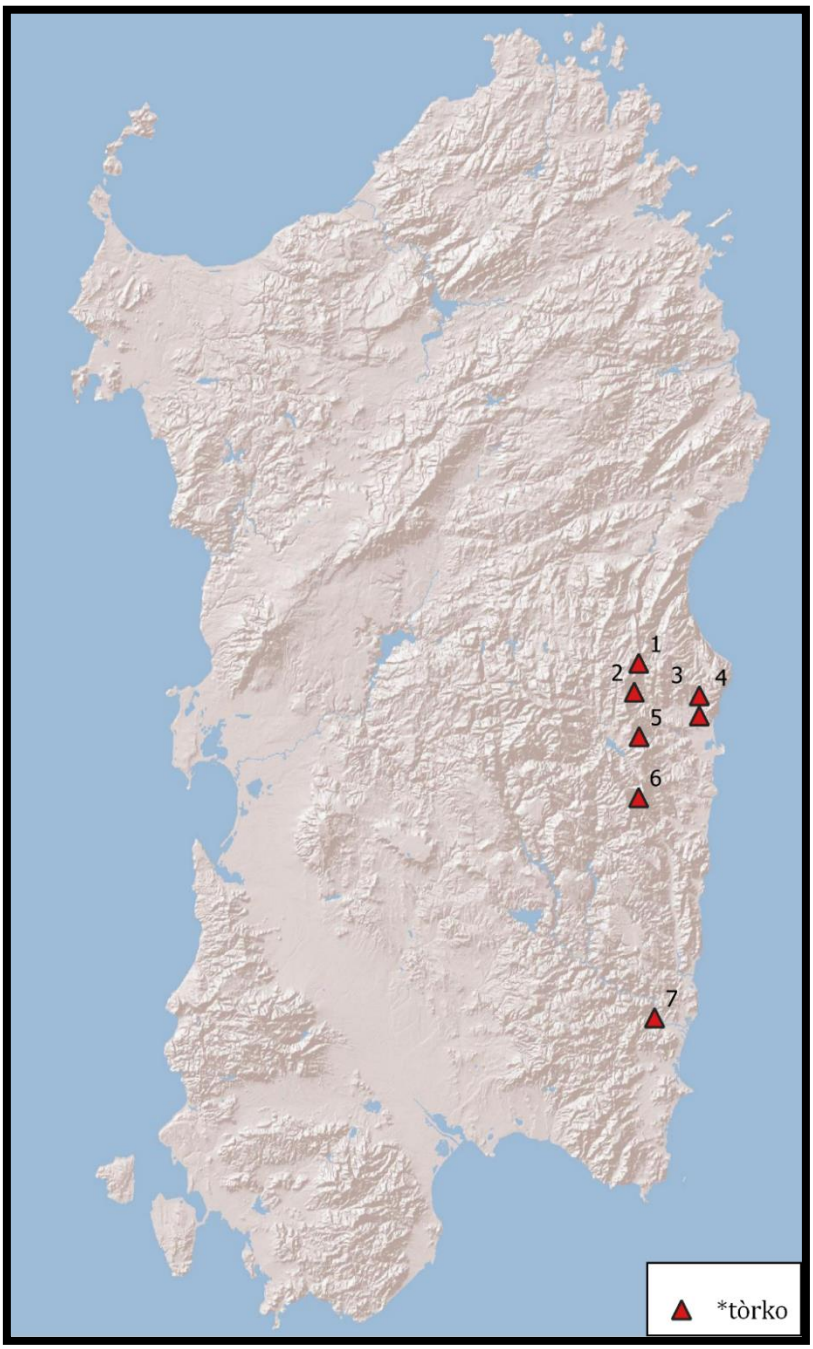

Figure 6.2.2: Clay used to make acorn bread

6.2.2 Clay used to make acorn bread

Reconstruction	Forms
tòrko	DES (II: 522): ***tròkko, tòrko*** (1: Urzulei, 2: Talana, 3: Triei, 4: Baunei, 5: Villagrande Strisaile, 6: Gairo) 'type of red clay used to prepare acorn bread'
	Böhne (1950: 69): ***trókku*** (7: San Vito) 'clay'
	AIS (416): ***tòrko*** (4: Baunei) 'clay'
	Blasco Ferrer (1988: 176): ***tròχo*** (1: Urzulei)

The form *tòrko* is likely original, with *tròkko* and *trókku* as the result of a metathesis. In the Ogliastran dialects, this word refers specifically to a type of clay used in the production of *pan' ispéli*, a kind of bread made from acorns (§ 6.2.1). In the Sarrabus (represented by San Vito), it has a more general meaning 'clay' (Böhne 1950: 69), which is also recorded by AIS (416) for Baunei (4). The ending *-o*, as well as the reference to a specific and likely ancient cultural practice, make it plausible that this word goes back to some pre-Roman source. Wagner (DES II: 522) considers a relation to *trókku* 'ravine' found in Fonni (§ 5.1.17), but this is semantically unattractive.

6.2.3 Clothes given to servants as a gift on certain holidays

Reconstruction	Forms
ϑèrga	Wagner (1921: 158; DES II: 544): ***ϑèrga*** (Nuor.), ***tèrga, attèrga*** (Log.), ***tsèrga, tsrèɣa*** (Camp.) 'cloths that a master gives to his servants on holidays or at the end of the year'

Wagner (1921: 158 fn. 2; DES II: 544) compares this word, which is found all over Sardinia, to Alb. *çergë, cergë* 'blanket of coarse wool', which is cognate to Turkish *çerge* 'type of tent', Rom. *cergă* 'blanket; tent canvas', MGr. τσέργα etc. 'id.' (cf. Ciorănescu 1958: s.v. *cergă*). These words are probably all loans from Turkish *çerge* 'type of tent' (Meyer 1891: 440; Ciorănescu 1958: s.v. *cergă*).[223]

[223] Contrary to the claim in DES (II: 544) that Albanian would be the donor to the other Balkanic languages.

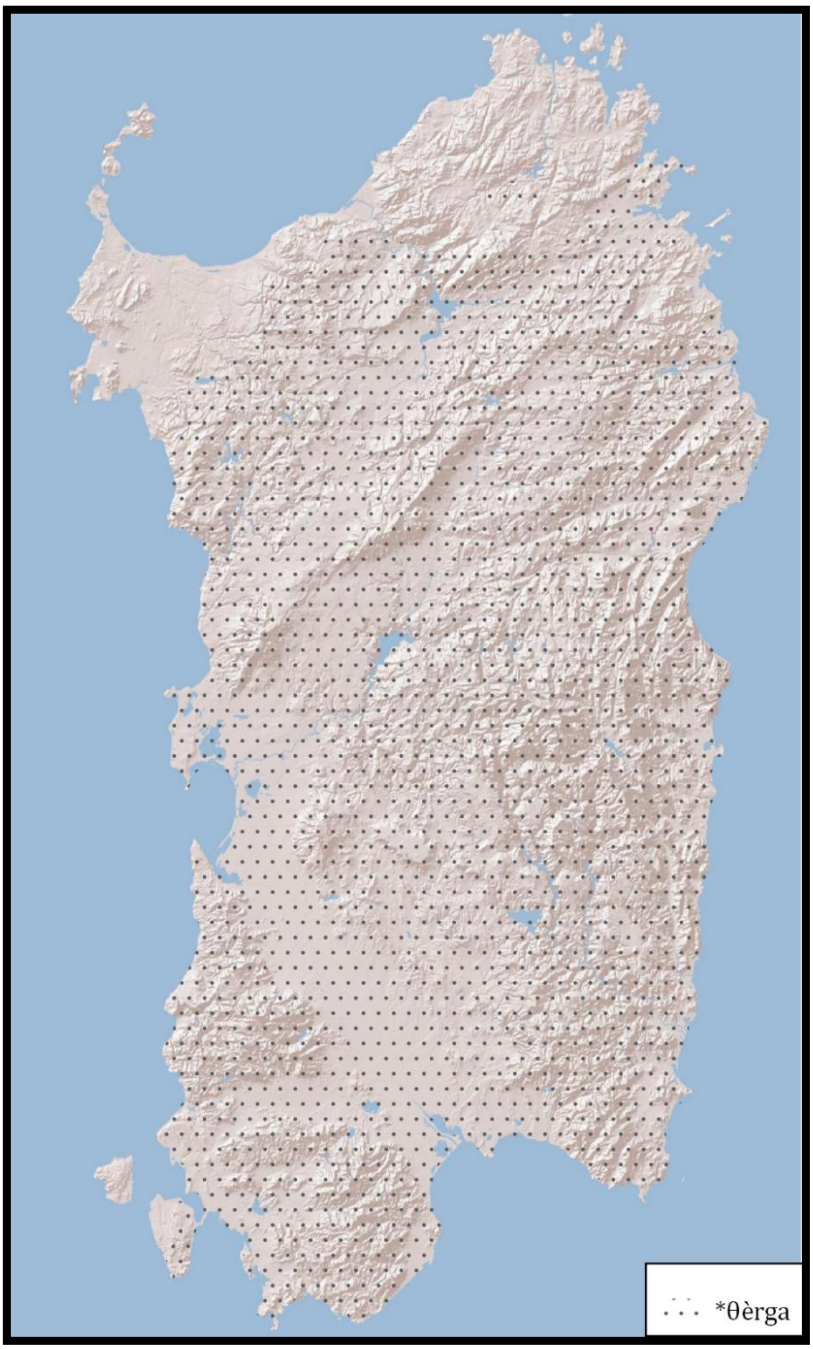

Figure 6.2.3: Clothes given to servants as a gift on certain holidays

Since for Srd. **ϑèrga*, both a derivation from Lat. *sērica* 'silk' (cf. REW 7848) and a loan from its apparent continuant Sp. *jerga* (cf. DCEC III: 507) are formally difficult,[224] one wonders whether **ϑèrga* could be a loan from MGr. τσέργα. This would have to be of Byzantine date. It is in any case clear that these Sardinian forms are not of pre-Roman origin.

6.2.4 Mat$_1$

Reconstruction	Forms
lóssi̯a	DES (II: 39): ***lúša*** (Log., Gallur.), ***rúša*** (1: Sassari), ***lússa*** (3: Siniscola), ***lússa*** (4: Bitti, 5: Bonorva), ***lússi̯a*** (6: Nuoro), ***lóssi̯a*** (Campidano di Milis, 8: Milis) 'mat made of reed or rushes; cylinders made of these mats used to store grain'
	AIS (1488): ***rúša*** (1: Sassari), ***lúša*** (2: Ploaghe), ***lússias*** (pl.) (7: Macomer) 'cylinders made of reed mats used to store grain'
	Böhne (1950: 98–99): ***lóssi̯a*** (10: San Vito) 'big reed mat used in carts; the plant used to make these mats'

Wagner (1921: 36 fn. 3) argues that most of the Sardinian forms presented above are interdialectal loans from *lóssi̯a* (8: Milis). Vendors from Milis and its surroundings, where reeds and rushes are common, would have travelled the island to sell their woven mats, thus dispersing their word for it. Wagner (1921: 36 fn. 3) initially suggested an origin in Lat. *lōdīx* 'cover, blanket, counterpane', by a derivation **lōdīcia*. This etymology has been taken over by FEW (V: 397) to account for OProv. *lodissa* 'blanket'. In DES (II: 39) however, Wagner questions his own etymology because of the fact that the same form *lóssi̯a* in San Vito does not only refer to mats, but also to the plants from which they are made (Böhne 1950: 98). This would make a derivation from Lat. *lōdīx* 'cover, blanket, counterpane' semantically more difficult. However, whereas a semantic shift from 'mat made of rushes' to 'rush' is perhaps imaginable, it is rather the formal development that makes a derivation from *lōdīx* problematic.

[224] The large amount of Romance forms that are similar to each other and to Lat. *sērica* 'silk', but that do not correspond regularly amongst themselves or to the Latin form (e.g. REW 7848; DCEC III: 507-08; FEW XI: 511-513) could suggest that we are dealing with some kind of medieval *Wanderwort*. This process is clearly continued by OFr. *sarge*, Fr. *serge*, whatever its exact origin, which was taken over in multiple languages: It. *sargia* 'light fabric', Flemish *sargie*, Westfalian *Schassen* 'blanket', Bret. *charj*, *serch* 'twill' etc. (FEW XI: 513a).

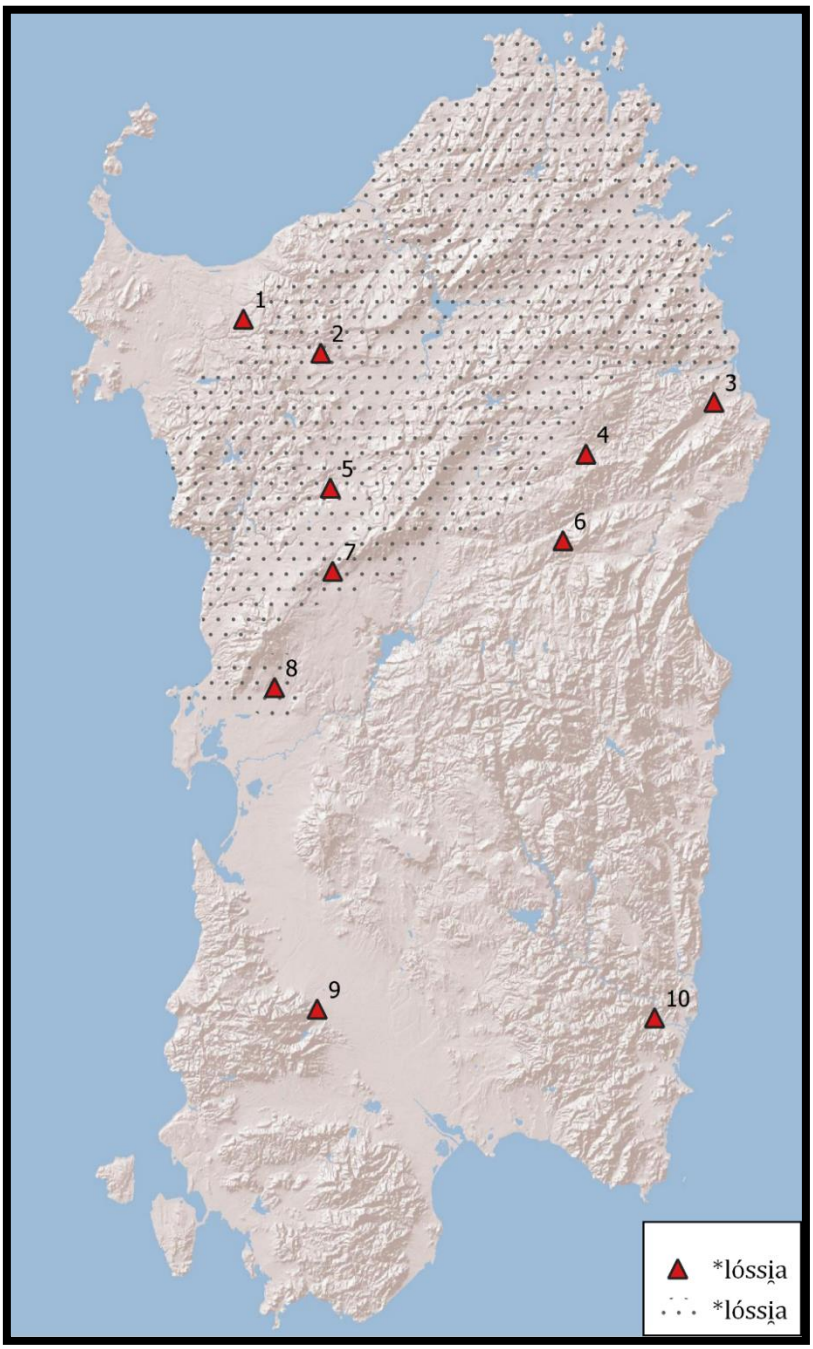

Figure 6.2.4: Mat

A form of the shape **lōdīcia*, as proposed by Wagner (1921: 36 fn. 3), would rather be expected to yield ***loíttsa*, as he himself also notes (Wagner 1921: 36 fn. 3; cf. also Wagner 1941a: 109–110). As a parallel development, he compares Srd. *ğustítsi̯a*, *ğustíssi̯a* 'justice', but this is clearly a learned borrowing from It. *giustizia* 'id.', rather than a form inherited from Lat. *iustitia*. It can therefore not be used as an argument concerning of the development **lōdīcia* > *lóssi̯a*. In order to derive *lóssi̯a* from *lōdīcia*, the retraction of the accent and the irregular development of *-ci̯-* remain unaccounted for.[225] At face value, *lóssi̯a* (8: Milis, 10: San Vito) seems to reflect **lóCsia/*lóCsea* (in which *-Cs-* may also reflect **-ss-*; cf. Wagner 1941a: 152–153, 170–172). This does not bring us any closer to a convincing etymology however. Perhaps Wagner (DES II: 39) is right in positing a pre-Roman origin for this word, in which case the forms from Milis (8) and San Vito (10) could be independent continuations, while the other forms are still probable loans from the form from Milis. The original meaning would have been 'rush', which was extended to 'reed/rush mat' in the area of Milis, and from there was introduced to the rest of the island. However, as Milis *lóssi̯a* 'mat made of reed or rushes' and San Vito *lóssi̯a* 'mat made of reed or rushes; rush (*Juncus sp.*)' are formally identical, there is no evidence for a pre-Roman origin other than the absence of a convincing etymology. Because of the formal issues discussed here, OProv. *lodissa* 'blanket' mentioned by FEW (V: 397), whether it is from **lōdīcia* or not, can probably not be upheld as a cognate to the Sardinian forms.[226]

6.2.5 Mat$_2$

Reconstruction	**Forms**
1: *kètra	DES (I: 446–447): ***čèrda*** (Camp.), [***sa***] ***žèðra*** (1: Mogoro) 'mat woven from straw or rushes that serves to spread out fruits on and that is also put in carts, so that the transported straw or fruits cannot fall out'[227]
2: *čèrda	DES (I: 446–447): ***ğèrda, ğèlda*** (Log.), ***dzèrda*** (Nuor.) 'id.'

[225] Loss of intervocalic **-d-* (or of *-b-*, *-g-* or *-v-* for that matter) would be regular (Wagner 1941a: 79). However, if **lōdīcia* is given up as an etymology for Srd. *lóssi̯a*, there is no reason to assume the contraction of two syllables on the basis of the attested evidence.

[226] Cf. however Van Sluis (2024: 100), who cites similar Gallo-Romance and Germanic forms for sedge: Clairvaux *lôche*, Dauphinois *lātša* Middle Low German *lusch* 'sedge'. If and how these are related to the Sardinian forms remains to be investigated.

[227] See Wagner (1921: 71) and Argiolas (2012: 2737) for images of carts with *čèrdas*.

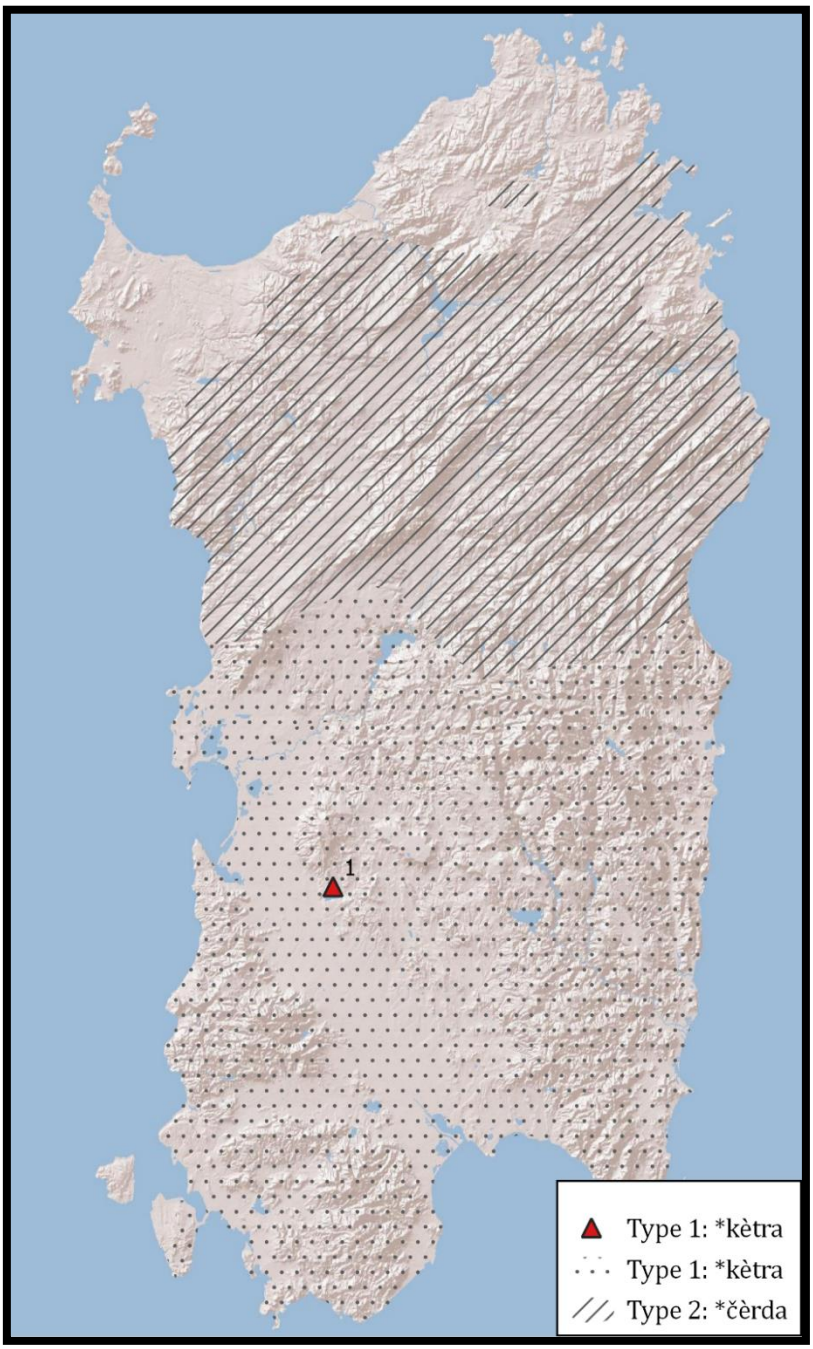

Figure 6.2.5: Mat

Like in the case of *lóssi̯a* (§ 6.2.4), this word was borrowed from Campidanese (type 1) into Logudorese and Nuorese (type 2). The explanation for this is that the mats in question were traded predominantly by merchants from Milis, in the Campidano (Wagner 1921: 71). The variant *čèrda* (and *čèðra*, 1: Mogoro) can go back either to **ketra* or **kerda*, and has been connected to Lat. *caetra*, *cētra* 'type of Spanish and/or African shield' (cf. Walde & Hofmann 1938: 135). Wagner's (1943: 196) connection of the Sardinian word to Ast. *sardu*, *sarda*, *zarda* 'wicker-work placed over the kitchen stove'[228] is questioned by Hubschmid (1953: 22–23) and DCEC (VI: 102a) on justified formal grounds. Hubschmid (1953: 22–23) instead regards both the Sardinian and the Asturian forms as independent pre-Roman loans from an "Afro-Hispanic" language. In this, he accepts Bertoldi's (1947b: 201) comparison to Tmz., Tsh., Rif. *ažərtil* 'big mat made of wool and esparto grass' (cf. Haddadou 2007: 81),[229] which can be reconstructed as **gʸV̌rtīl*.[230] A connection between the Sardinian and the Berber forms suffers from some formal objections however. These include the lack of *-l-* in the Latin and Romance forms and the irregular voicing of **-rt-* that would have to be assumed to arrive at Srd. *čèrda* (as well as Ast. *sardu*, *sarda*, *zarda*, which Bertoldi (1947b: 201) still connects).[231]

Argiolas (2012: 2737–2741) alternatively proposes a derivation from the Berber root **ɣrd*, found in Tmz. *ɣərd* (Ayt Izdeg) 'to lie down on the floor' (cf. Taïfi 1991: 202), Kb. *əɣṛuṛəd* 'to sleep without being able to wake up' etc. On the basis of these attestations, she proposes a base meaning 'mat' for this root. Its supposed Libyco-Berber predecessor **ɣerd* would have entered African Latin first as *cherda* 'some plant' (probably eryngo, cf. André 1985: 62) and from there as *caetra*, *cētra* 'woven shield'. There are some semantic issues with this account, however. First, the motivation for a semantic shift from 'mat' to 'eryngo' to 'woven shield' is not evident. Moreover, assuming an original meaning 'mat' for **ɣrd* on the basis of verbs meaning 'to lie down, etc.' is not evident either. A semantic shift like 'to lie down' > 'mat' > 'woven shield' would perhaps be more plausible, but this still does not resolve the formal issue that the metathesis and devoicing of *-rd-* >> *-tr-*, required to explain Lat. *caetra* from either Bb. **ɣrd* or from African Lat. *cherda*, are irregular and unexpected.

---

[228] "Tejido de mimbres que se coloca sobre el llar, para abrigar la cocina y donde se ponen las avellanas para turrar" (Wagner 1943: 196).

[229] Cited by Bertoldi (1947b: 201) as *ažerϑil*.

[230] Maarten Kossmann, p.c.

[231] Cf. also Argiolas (2012: 2736), who rather sees a link between the Berber forms and Gr. κάρταλλος, Lat. *cartallus*, which are of unknown origin (cf. Beekes 2010: 650).

For Camp. *čèrda*, I agree with Wagner (DES I: 447) that it is most economical to simply assume inheritance from Lat. *caetra* 'type of shield', whatever its origin.[232] The phonological development is as expected and for the semantic shift from 'shield' to 'type of mat', we may compare Gr. γέρρον 'oblong shield; wicker body of a cart; wattled screen'. Although Latin *caetra* is evidently of foreign origin (Walde & Hofmann 1938: 135),[233] it is inherited in Sardinian, and does not offer any information about the pre-Roman language(s) of Sardinia.

[232] This Latin word potentially belongs together with some other words referring to weaving or to woven products without a clear etymology. Cf. Lat. *cartallus* 'basket', *gerdius* 'weaver', which are identical to Gr. κάρταλλος 'basket with a pointed bottom' and γέρδιος 'weaver', but which lack good etymologies (Beekes 2010: 268, 650). Because of the obvious phonological similarity, Gr. γέρδιος has been connected by Argiolas (2012: 2738) to the Berber root *yrd* 'to spread out' as well.

[233] Cf. Múrcia Sànchez (2011: 423).

7 Miscellaneous

This chapter contains all words that are not easily categorized into one of the earlier chapters. Those in § 7.1 refer to body parts or bodily features in the broadest sense. The rest is subsumed under § 7.2.

7.1 Physical

7.1.1 Blind

Reconstruction	Forms
ϑúrpu	DES (II: 555): ***ϑúrpu*** (Nuor.), ***túrpu*** (Log.), ***tsrúppu*** (Camp.) 'blind'
	AIS (188): ***ϑúrpu*** (1: Bitti, 2: Nuoro, 3: Dorgali, 5: Fonni, 8: Baunei), ***túrpu*** (2: Nuoro), ***tsúrpu*** (4: Santu Lussurgiu, 13: Escalaplano, 15: Cagliari, 16: Sant'Antioco), ***tsrúppu*** (6: Milis, 7: Busachi, 10: Laconi, 11: Mogoro, 14: Villacidro), ***čúrpu*** (9: Desulo), ***issúrpu*** (12: Perdasdefogu) 'id.'

This word for 'blind' is attested already in the Condaghe di San Pietro di Silki, in the personal name *Petru Natale su thurpu* (Bonazzi 1900: sec. 433). In the northern part of the Logudorese-speaking area, we find forms of the type *tséɣu* 'blind' (AIS: 188; DES: 555). These must be loans from It. *cieco*, which suggests that cognates of *ϑúrpu* were once present in the north of the island too. REW (9006) considers Lat. *turpis* 'ugly, repulsive etc.' as a possible origin, but notes that it can only work if the initial *t-* of **turpu* was assimilated to the *ϑ-* of *ϑóppu* 'lame' in an original syntagm like **túrpu e ϑóppu* 'blind and lame'. This is formally and semantically unconvincing, and Wagner (1930: 137) therefore rejects the etymology. On formal grounds, Wagner (1930: 137) also rejects a derivation from Lat. *orbus* 'bereft, devoid', which is the source for other words meaning 'blind' in various Romance languages (REW 6086).

On the basis of the onset in **ϑ-*, Wagner (1930: 137; DES II: 555) takes this word to be of pre-Roman origin. Blasco Ferrer (1999: 68) instead proposes it to be the reflex of **supp(u)rus* << Lat. *suppūrāre* 'to form pus, to fester'. This requires a syncope to **supprus* and a subsequent metathesis to **surpus*. Additionally, a development of **s-* to **ϑ-* is needed to explain its outcomes *ϑ-*, *ts-*, *t-*.

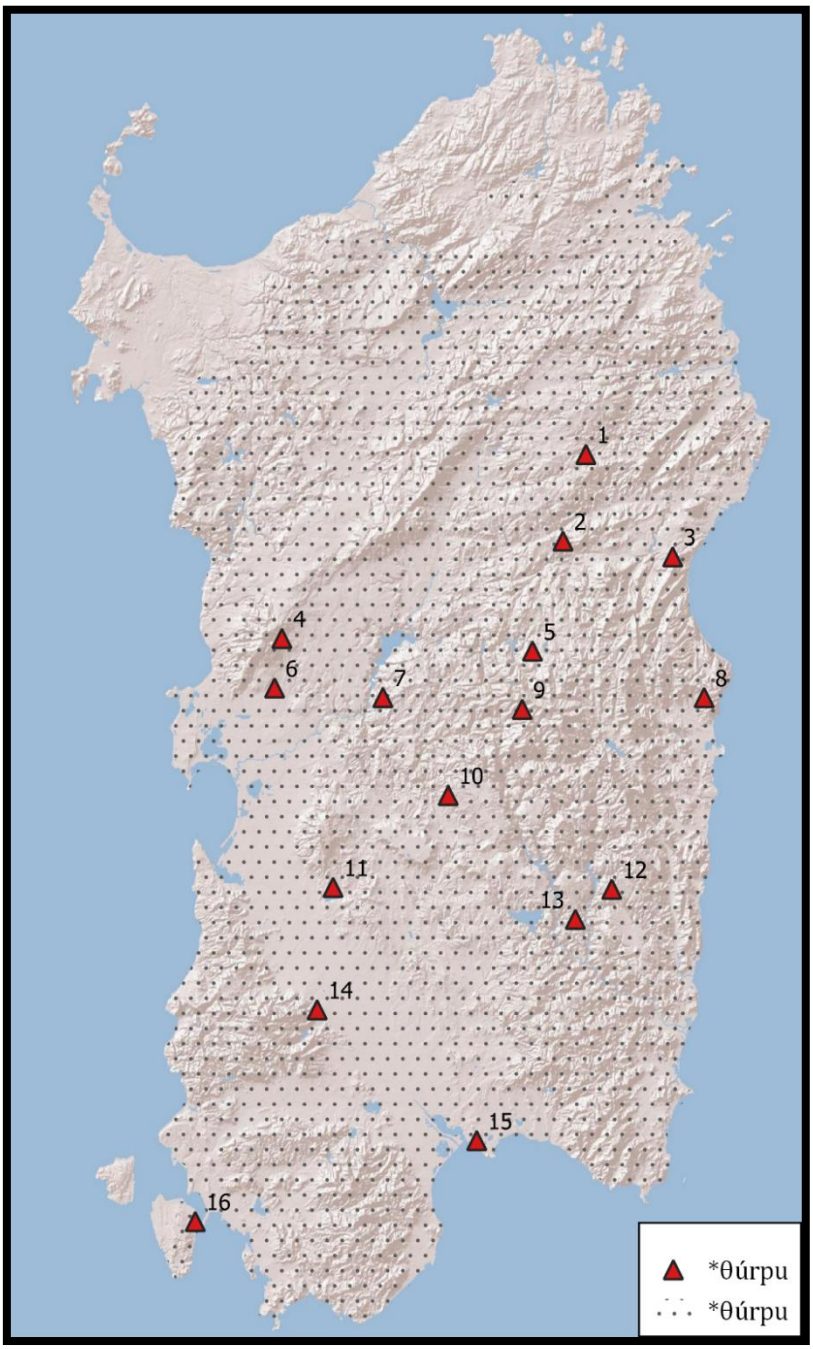

Figure 7.1.1: Blind

Whereas *ϑ- is admittedly attested as an irregular reflex of Latin *s-* in some words (cf. Paulis 1984b: 536–537), I agree with Areddu (2022: 263) that the hypothesized syncope would be unexpected in this position and that the semantic development from 'build-up of pus, festering' to 'blind' is not straightforward. Combined with the fact that the required Latin form **suppūrus* is not actually attested, Blasco Ferrer's (1999: 68) etymology fails to convince.

Areddu (2022: 267, 277) returns to a substrate origin, as proposed by Wagner (DES II: 555), assuming it is the outcome of a pre-Roman introduction of Gr. τυφλός 'blind' in Sardinia.[234] In order to make this work formally, he argues that there must have been a by-form *θυπλός, with a different dissimilation of the aspiration in original *θυφλός. While the similarity between Srd. *ϑúrpu* etc. and Gr. τυφλός is indeed striking, there are some issues with this account as well. One is the development of *-*l-* > *-r-* in the hypothetical form **ϑulpu*, which purportedly arose after the metathesis of *-*l-* of Gr. *θυπλος.[235] Whereas *-r-* is the regular outcome of *-*l-* before a consonant in almost all dialects of Sardinian, *-*l-* is usually preserved in all contexts in Baunei; e.g. *linnálβu* 'white poplar' (< Lat. *lignum album*), *splíyu* 'mirror' (< Lat. **spiculum*), *dulče* 'sweet' (< Lat. *dulcis*) (AIS: 585, 675, 1266; DES I: 106, 483, 682). Areddu (2022: 269) adduces some cases in which there is an alternation between *-r-* and *-l-* in non-inherited words shared between Latin and Greek, but Wigman (2023: 346–349) argues that these are likely due to the intervention of a third language rather than direct Greek loans in Latin.

An additional problem is the rendering of hypothetical Gr. θ- as Srd. *ϑ-. There is indeed good evidence for this in words of Byzantine origin (Wagner 1941a: 109), but by this time Greek -θ- would have been fricative [θ], whereas a much earlier pre-Roman Classical Greek loan would still have had the aspirated plosive realization [t^h]. In light of the fact that in many early Greek loans in Latin original Gr. -θ- is realized as Lat. *-t-* (Wagner 1941a: 109), it is not at all self-evident that a Classical Greek loan containing [t^h] would appear with *ϑ in Sardinian. Moreover, the evidence for a substantial linguistic presence of Greek before the Byzantine

[234] Gr. τυφλός 'blind' < *d^hub^h-*lo-* itself has a well-established IE etymology. Its root *d^heub^h- is further attested in OIr. *dub* 'black', OHG *toub* 'deaf' etc. (Beekes 2010: 1521).

[235] Note that in Areddu's (2022) scenario this metathesis must have been of pre-Roman date. If it had taken place within Sardinian, where metathesis is very common, we would expect the *-*p-* to have lenited to **-*β-* in its original post-vocalic position in Campidanese and Logudorese; cf. Lat. *pētra* 'rock' > Log. *pèðra*, *prèða*, Camp. *pèrda* 'id.' (DES II: 254).

period is scarce. In conclusion, while a comparison between Srd. *ϑúrpu* 'blind' and Gr. τυφλός 'id.' might seem attractive at first sight, it is fraught with problems. Quite possibly, we are rather dealing with a chance similarity. So far therefore, Wagner's (DES II: 555) proposal that Srd. **ϑúrpu* originated in a pre-Roman language remains the best option.

7.1.2 Bristle (of pigs)

Reconstruction	**Forms**
***ϑúlla**	DES (II: 553): ***tsúḍḍa*** (Gallur., Camp.), ***tsúḍḍu*** (Sass.), ***túḍḍa*** (Log.), ***ϑúḍḍa*** (Nuor.) 'bristle hair (of pigs and wild boars)'; [***sar***] ***ϑúḍḍar*** [***dessu póḍḍike***] (pl.) (6: Nuoro) 'hangnail', ***túḍḍa*** (Log., 5: Nule) 'but, shoot, sprout (of plants)'[236]
	AIS (1093): ***tsúḍḍi*** (pl.) (1: Tempio, 2: Sassari), ***túḍḍa*** (3: Ploaghe, 5: Nuoro, 8: Macomer), ***ϑúḍḍa*** (4: Bitti, 10: Fonni, 13: Baunei), ***tsúḍḍa*** (7: Dorgali, 9: Santu Lussurgiu, 11: Milis, 12: Busachi, 15: Mogoro, 17: Escalaplano, 19: Villacidro, 21: Cagliari), ***čúḍḍa*** (14: Desulo), ***súḍḍa*** (16: Perdasdefogu), ***zúḍḍa*** (22: Sant'Antioco) 'bristle hair (of pigs and wild boars)'
	Böhne (1950: 115): [***iz***] ***itsúḍḍas*** (20: San Vito) 'bristle hairs (of pigs and wild boars); inflamed eyelash'

The forms presented above correspond regularly to each other and go back to **ϑúlla*.[237] Wagner (1941a: 115 fn. 1; DES II: 554) does not accept Salvioni's (1914: 585) derivation from **seϑϑúḍḍa* < **saeteulla* << **saeteus* << Lat. *saeta* 'bristle'. He also opposes a derivation from *saetula* (Jud *apud* DES II: 554), or from the Latin plant name *sulla* (Jud *apud* DES II: 554).[238] He likewise rejects a connection with Gr.Hes. θύλλος, θύλλα 'shoots', since the meaning 'bud, shoot, sprout', found in the Sardinian forms, is an extension of the original meaning 'bristle' and is restricted to Logudorese. According to Wagner (DES II: 554), this word has "every appearance" of being of pre-Roman origin. This opinion was presumably inspired by the initial **ϑ-* and the lack of a convincing etymology.

[236] With many derivations; see DES (II: 553).

[237] One exception is *tsúḍḍa* in Dorgali (7). If from **ϑúlla*, we would rather expect ***túḍḍa* or ***ϑúḍḍa*. I have no explanation for this.

[238] I have not been able to trace Wagner's (DES II: 554) reference to a publication by Jud in *Romània* (43: 194).

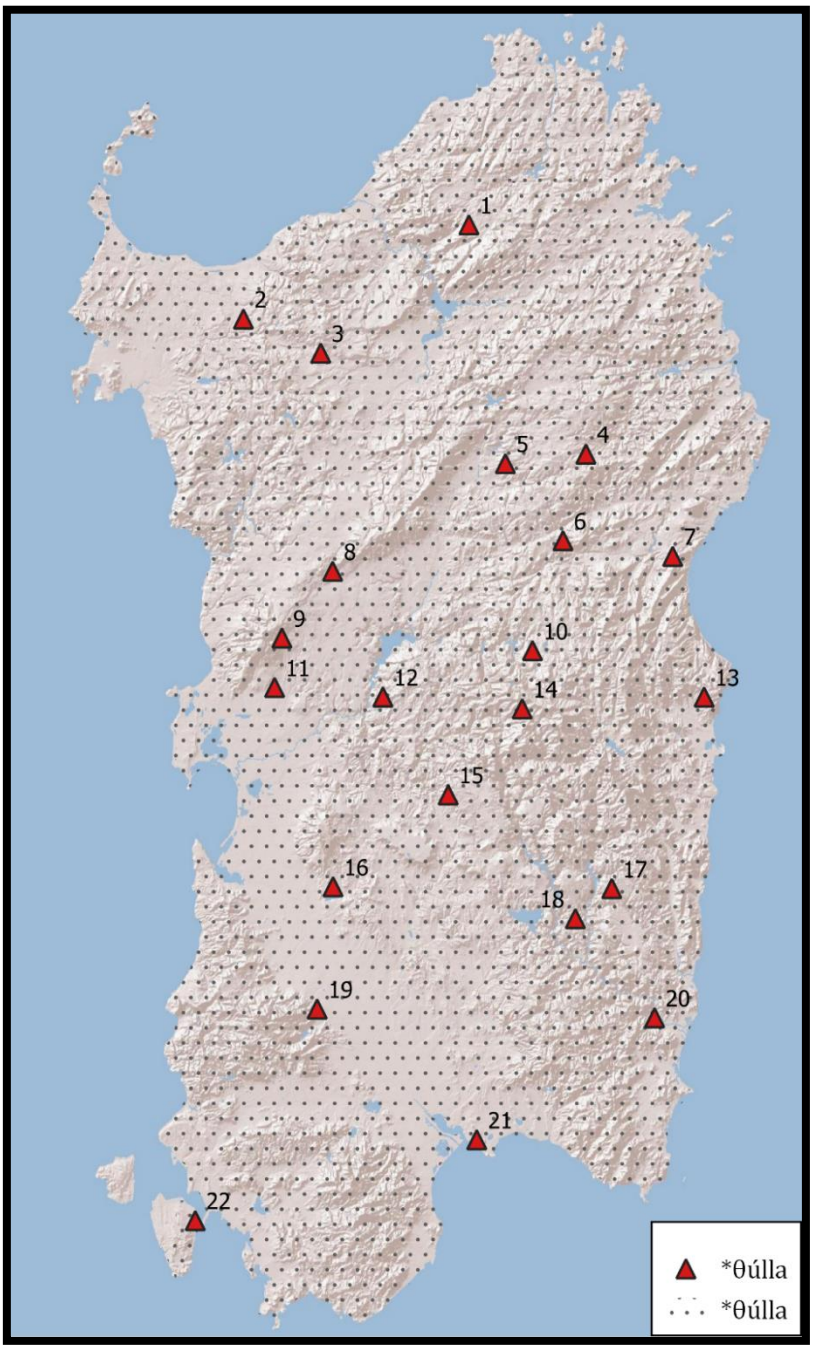

Figure 7.1.2: Bristle

Wagner (DES II: 554) instead proposes a connection to Snh. *əzzuṭ* 'pubic and armpit hairs' (cf. Renisio 1932: 320). Formally, however, the comparison between this word and Srd. *ϑúḍḍa* etc. < **ϑúlla* is by no means more attractive than the etymologies proposed by Salvioni and Jud. The etymology of this Sardinian word remains obscure.

7.1.3 Cheek

Reconstruction	**Forms**
***tuttúrru**	DES (II: 537): ***tuttúrru*** (Log., 2: Orosei) 'cheek', (3: Olzai) 'pig missing one ear'; ***tutturròne*** (1: Siniscola) 'mumps'

The original meaning of this word appears to be 'cheek'.[239] Wagner (DES II: 537) follows Hubschmid (1953: 75–76) in connecting it to Bq. *tutur* (Biscay) 'crest', *tuturru* (Navarre) 'point, summit' etc. According to Trask (2008: 344), however, these Basque forms belong to a larger group of expressive formations with reduplication of *-t-* or *-k-*. Thus, next to *tutur* 'crest' we find *kukur* (Biscay, Navarre) 'id.', and next to *tuturru* 'point, summit' we find *kukutz* (Biscay), *kokor* (Labourd) 'id.' (Trask 2008: 344). This indicates that the Basque formations adduced by Hubschmid are rather recent. Moreover, the comparison between Srd. 'cheek' and Bq. 'crest; summit' is not at all straightforward semantically. In my view, a chance similarity between the Basque and Sardinian forms is more likely. This does not exclude a pre-Roman origin for Srd. *tuttúrru*, but the evidence is restricted to the absence of an etymology and perhaps its somewhat peculiar phonological structure. At any rate, this word must be kept separate from Srd. *tútturu* 'round object; rolling pin etc.' (DES II: 538).

[239] Cf. also the various verbal derivations listed in DES (II: 537), meaning 'to slap on the cheek'.

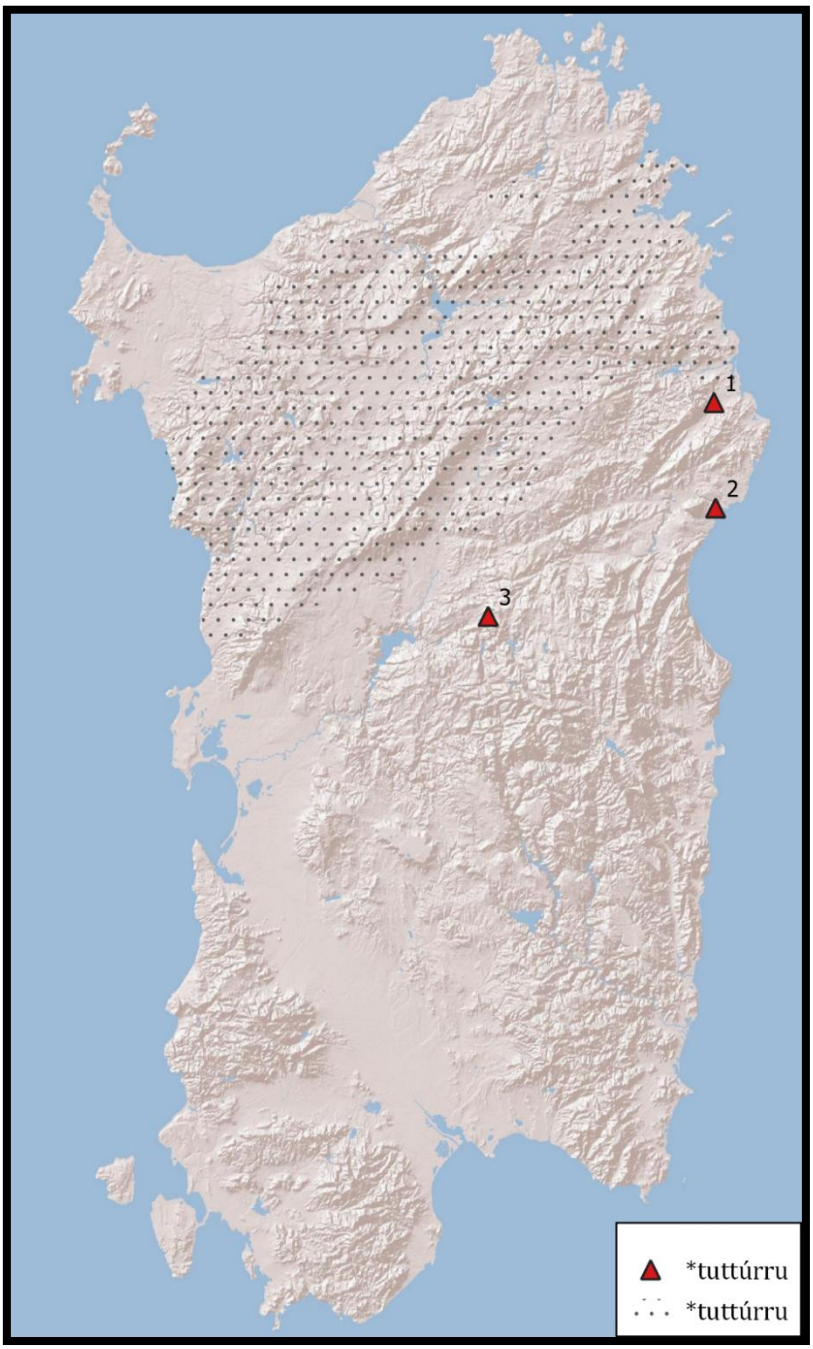

Figure 7.1.3: Cheek

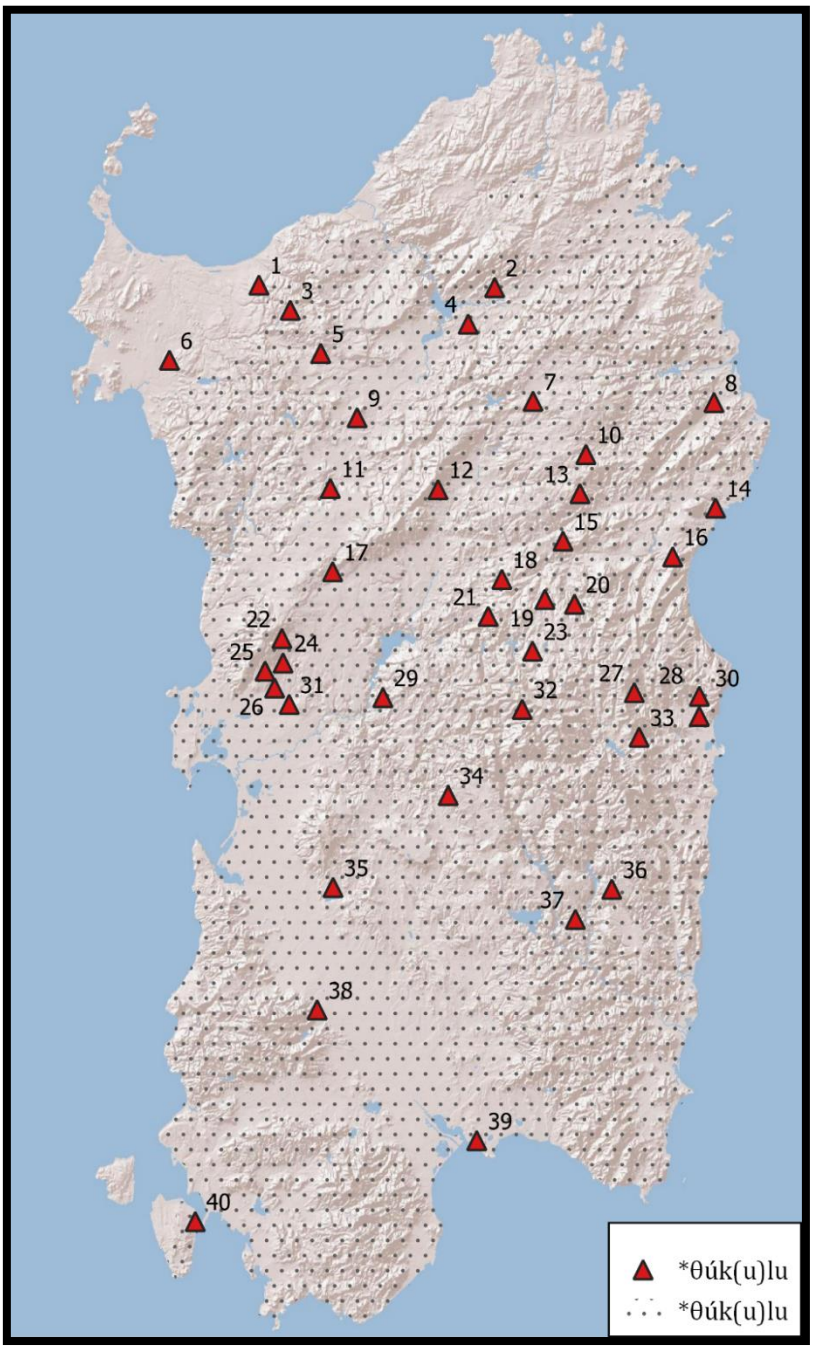

Figure 7.1.4: Neck

7.1.4 Neck[240]

Reconstruction	Forms
ϑúk(u)lu	DES (II: 553): ***túǧu*** (NLog., 1: Sennori, 4: Oschiri, 7: Buddusò, 9: Mores, 11: Bonorva), ***túǵu*** (2: Berchidda), ***túyu*** (3: Osilo, 5: Ploaghe), ***(b)utúyu*** (6: Olmedo), ***ϑrúk(k)u*** (8: Siniscola, 10: Bitti, 13: Orune, 14: Orosei, 15: Nuoro), ***trúɣu*** (Log., 12: Bono, 17: Macomer, 25: Seneghe), ***ϑúčču*** (18: Orani) ***ϑúkru*** (19: Mamoiada, 23: Fonni), ***ϑúlʔu*** (20: Orgosolo), ***tsúkkru*** (22: Olzai), ***tsúɣu*** (Camp.), ***tsrúɣu*** (24: Bonarcado, 31: Bauladu) ***ϑúɣru*** (27: Talana, 28: Triei), ***ϑúɣulu*** (30: Baunei), ***ϑrúɣu*** (33: Villagrande Strisaili) 'neck'
	AIS (118): ***ϑrúχu*** (16: Dorgali), ***tsúɣu*** (22: Santu Lussurgiu, 29: Busachi, 35: Mogoro, 37: Escalaplano, 38: Villacidro, 39: Cagliari), ***tsrúɣu*** (26: Milis, 34: Laconi), ***čúɣu*** (32: Desulo), ***súɣu*** (36: Perdasdefogu), [***su***] ***zúɣu*** (40: Sant'Antioco) 'id.'

This is the common Sardinian word for neck. It is present in all Sardinian dialects, but not in the Sardo-Corsican dialects spoken on the island (cf. Sass. *lu ggóḍḍu*, Gallur. *lu kóḍḍu*, AIS 118). All forms go back to earlier **ϑúk(u)lu*. Despite the perfect phonological regularity between the forms and their omnipresence in Sardinian, both of which might be indications for an inherited origin, **ϑúk(u)lu* cannot easily be shown to derive from any Latin formation. However, there are various etymologies that try to do exactly that.

Wagner's (DES II: 553) derivation from Lat. *iugulum* 'collarbone', with a prefixed **ϑ-*, requires a number of assumptions. First of all, we would have to assume the prefixation of an inherited word with a non-inherited and probably pre-Roman element **ϑ(i)-* (cf. § 9.1.2). Second, it requires the assumption of a semantic development from 'collarbone' to 'neck', which is possible. Finally, it is necessary to assume the devoicing of *-g-*: *jugulum* >> *juculum* >> **ϑ-juculum* > **ϑúkulu*. This too is rather trivial in Sardinian, where many words get the ending **-k(u)lu* secondarily on the basis of the frequent inherited suffix *-culus* (cf. Wagner 1952: 29–30).

Blasco Ferrer (1999: 67) instead believes Srd. **ϑúk(u)lu* 'neck' to regularly go back to Lat. *sucula* 'screw/mechanism for lifting a wine/oil press' (cf. Lewis & Short 1879: s.v. *sucula*). Not only is this semantically quite a leap, his argument that *ϑ-* would be the regular outcome of **s-* is exclusively based on *suglu* he adduces

[240] These data are discussed in a similar fashion in Swanenvleugel (2024: 255–256).

from Baunei. However, not only is the form *ϑúyulu* with *ϑ-* documented for Baunei in AIS (118) and DES (II: 553), the *s-* in *suglu* could be from the neighbouring Ogliastran dialects, where it is in fact the regular outcome of **-ϑ-* (Wagner 1941a: 107). Blasco Ferrer's (1999) etymology does not resolve all problems surrounding this word.

If Wagner's etymology is correct, *ϑúkru* etc. 'neck' is another possible candidate for the secondary *ϑ(i)*-prefixation of a word inherited from Latin. However, Sardinian *-ϑ-* is the outcome of *-tɟ-* or *-kɟ-* (Latin <ci>) before a vowel, so in the case of *jugulum*, an initial **t-* or **k-* would work equally well. Additionally, we have no information about what the regular outcome of **ϑju-* or **ϑiju-* would be, since a sequence **tɟiu-/*kɟiu-* did not occur in Latin. We are, as a result, forced to choose between **tɟúkulu/*kɟúkulu* that is phonetically warranted but morphologically opaque, and **ϑɟiúkulu* which is phonetically uncertain, but at least has some parallels when it comes to its derivation. For now it seems best to regard this word as no more than a possible example of secondary **ϑi-*addition, but at least a partial pre-Roman origin is plausible.

7.1.5 Rash, eczema

Reconstruction	**Forms**
***ϑèrna**	DES (II: 544–545): ***attèrra*** (Log.), ***tsèrra*** (Camp.), ***ϑèrra*** (Nuor., 1: Bitti, 2: Dorgali, 4: Fonni, 7: Baunei) 'rash, eczema'
	AIS (683): ***tsèrra*** (3: Santu Lussurgiu, 5: Milis, 6: Busachi, 9: Laconi, 10: Mogoro, 13: Villacidro), ***čèrra*** (8: Desulo), ***sèrra*** (11: Perdasdefogu), ***itsèrras*** (pl., 12: Escalaplano), [***sa***] ***zèrra*** (14: Sant'Antioco) 'id.'

Lat. *zerna* 'rash, eczema etc.' is attested in various late Latin sources (cf. DES II: 544–545) and can be the regular source of the Sardinian forms. Wagner (DES II: 545) posits a pre-Indo-European origin for this Latin word. This is based on its initial *z-*, and on its irregular correspondence to *sarna* found in Isidore of Seville and continued by Sp. *sarna* (cf. DCEC V: 163–166), and by Bq. *sarra* 'rust, slag'. However this may be, it is clear that similar forms were widespread in the vulgar Latin of Roman Africa, Italy and the Western Mediterranean (DCEC V: 164). I see no reason to assume that Lat. *zerna* would have belonged to a pre-Roman Sardinian language. Srd. **ϑèrna* is therefore most likely inherited from late Latin *zerna*.

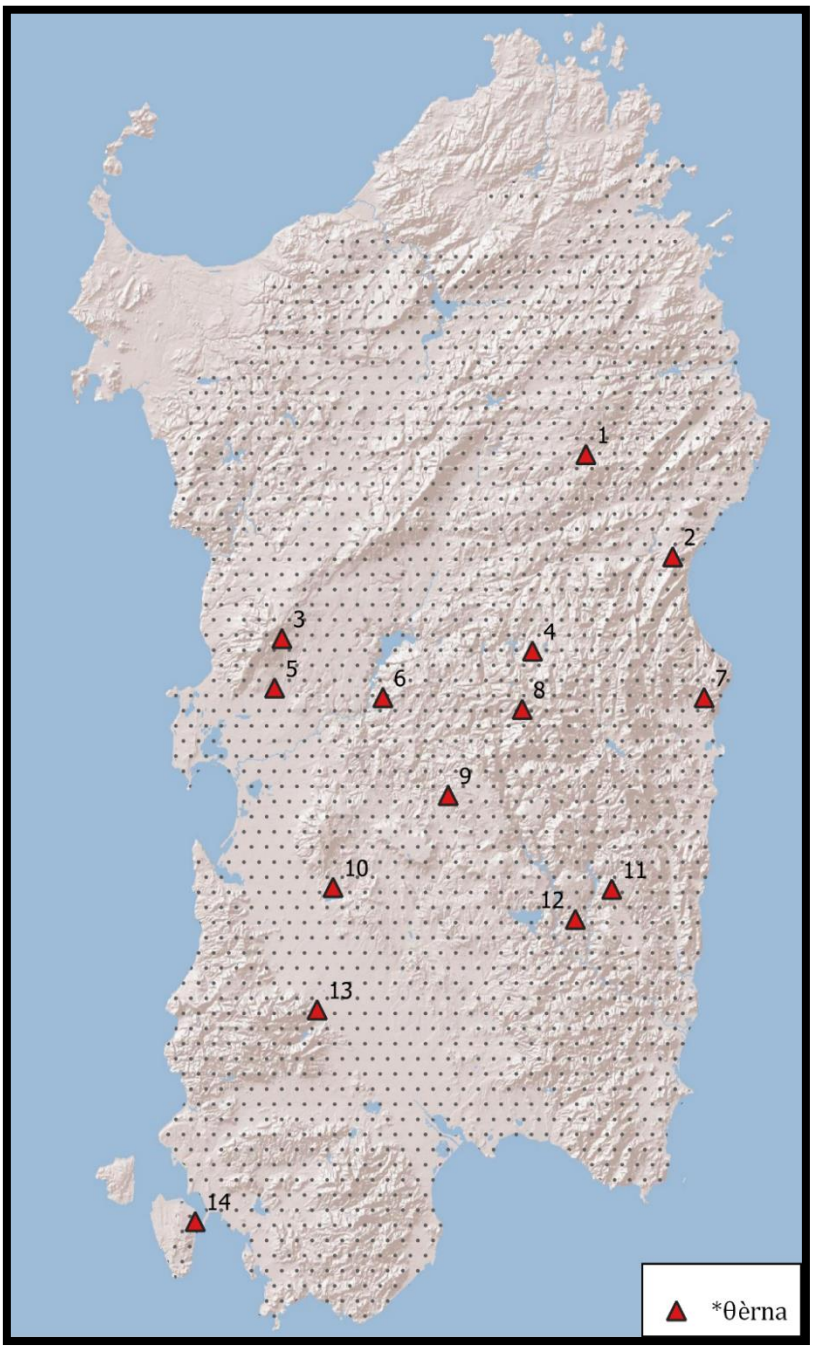

Figure 7.1.5: Rash, eczema

7.2 Other

7.2.1 Evil spirit

Reconstruction	**Forms**
***turlío**	DES (II: 526): ***trullío*** (1: Orgosolo), ***turrío*** (2: Fonni) 'evil spirit'

Wagner (DES II: 526) proposes a pre-Roman origin for this word. He does so for two reasons. One is its reference to a presumably pre-Christian superstition. The other is its final *-o*, which is rare in Sardinian nouns inherited from Latin (§ 8.1.5.2). These considerations are valid, and Wagner (DES II: 526) is likely right. The phonological discrepancy between the two attestations is reminiscent of Sardinian doublets like Log. *trúḍḍa*, Camp. *túrra* < **trúlla*, **túrla* < Lat. *trulla* 'ladle' (Wagner 1941a: 173; DES II: 526). This suggests that the variation between *trullío* and *turrío* too is the result of a metathesis. It is not easy to establish the base form, but the fact that we find geminate *-ll-* in Orgosolo rather than ***-ḍḍ-* < **-ll-* suggests that *trullío* is the result of a metathesis from **turlío* after the change of original **-ll-* > *-ḍḍ-*.[241] The regularity of the development **-rl-* > *-rr-*, needed to explain *turrío* (2: Fonni) < **turlío*, is difficult to prove, since *-rl-* is a consonant cluster unknown to Latin. Still, *-rr-* is an outcome of **-rl-* in Campidanese forms like *méurra* 'blackbird' < **méurla* < *mérula*, after metathesis. It appears that a reconstruction **turlío* most easily explains all attested forms. Whether this in turn goes back to **turílo* vel sim. (like **méurla* < *mérula*) is impossible to say.

[241] It is possible that influence from forms like *trullíu* 'rotten (of eggs)' played a role in the hypothetical metathesis from **turlío* to *trullío* (as suggested by DES II: 536).

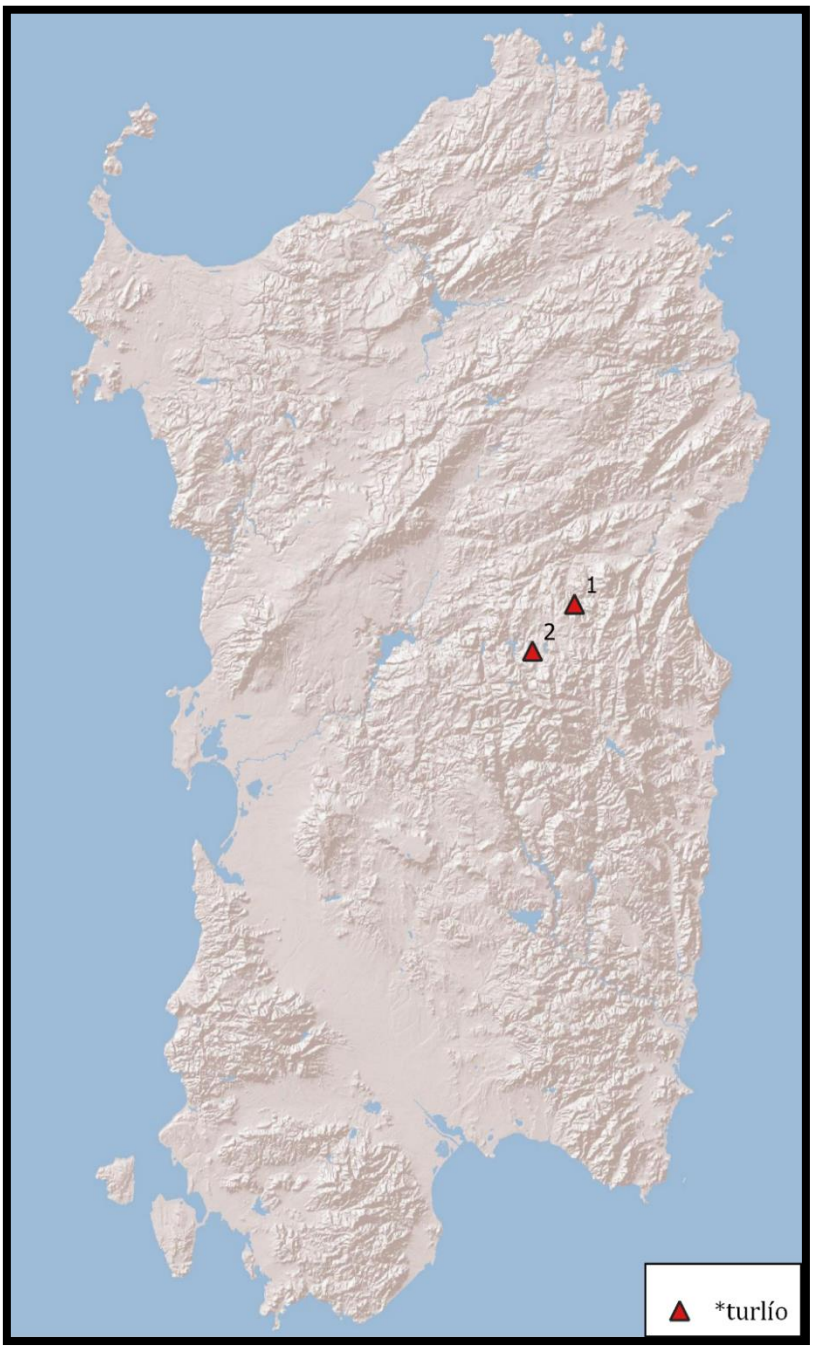

Figure 7.2.1: Evil spirit

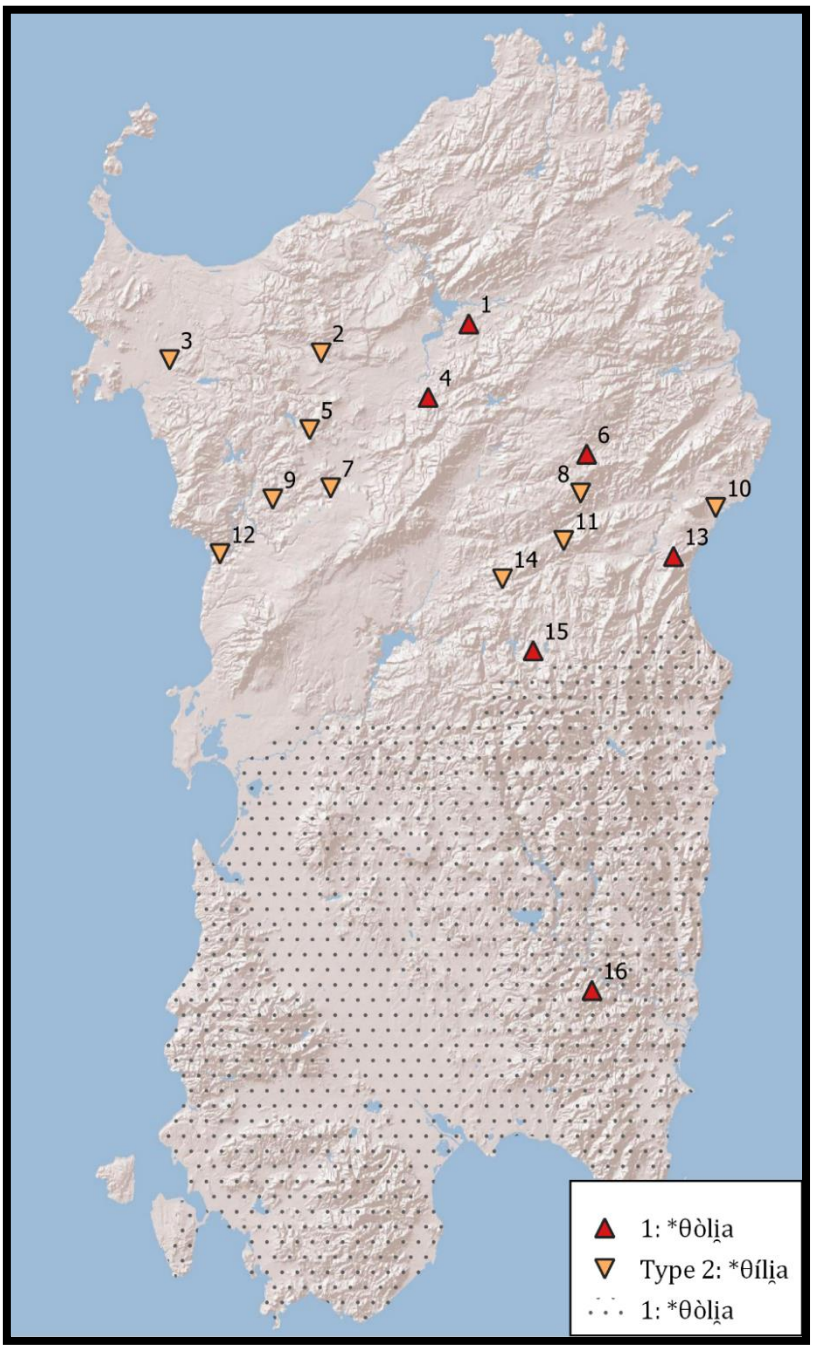

Figure 7.2.2: Film, layer

7.2.2 Film, layer

Reconstruction	Forms
1: *ϑòl̥ia	DES (II: 549-50): ***tòdza*** (1: Oschiri, 4: Ozieri), ***ϑòdza*** (6: Bitti, 13: Dorgali), ***ϑòǧǧa*** (15: Fonni), ***sòlla*** (Camp., 16: Villasalto) 'filth on one's head, dandruff; any type of thin layer (e.g. of bread)'; in Dorgali also: 'the woody part of hemp'; Villasalto: 'snow flake'
2: *ϑíl̥ia	***tídza*** (2: Ploaghe, 3: Olmedo, 5: Thiesi, 7: Bonorva, 9: Padria), ***ϑídza*** (8: Orune, 10: Orosei, 11: Nuoro, 14: Orani), ***attídza*** (12: Bosa) 'filth on one's head, dandruff; any type of thin layer (e.g. of bread)'

These forms for 'dandruff; thin layer etc.' are found across Sardinia. They occur in two types. Type 1 suggests a reconstruction **ϑòl̥ia*, type 2 instead **ϑíl̥ia* (cf. DES II: 550). The apparent vocalic alternation is hard to account for by any scenario involving a single etymology. Especially puzzling is the fact that, in certain regions, both types are found in neighboring towns (e.g. Bitti *ϑòdza* vs. Orune *ϑídza*). Jud (1914: 453 fn. 1) proposes a connection with Abruz. *zélle* 'ringworm', Nap. *zella* 'disease of the skin on the head', which are of unknown origin. It is possible that these forms are somehow related to Srd. **ϑíl̥ia*, but a relation to **ϑòl̥ia* is formally difficult. All of this raises the suspicion that we might be dealing with the partial conflation of two etymologically originally distinct words, but their ultimate origin remains obscure.

7.2.3 Forked juniper trunk

Reconstruction	Forms
****dúri***	DES (I: 483): ***dúri*** (1: Nuoro, 2: Orgosolo) 'forked trunk of a juniper tree with many side branches used to hang meat from'
	Puddu (2023: s.v. *dúle*): ***dúle*** 'dried or forked juniper tree used to hang tools from in a stable'

Wagner (1921: 153–154) describes various types of wooden frames found mainly in *pinnetas* (i.e. shepherd's huts) with horizontal bars to hang meat or other things from. Of these, the *dúri* is a specific type that is usually made of a juniper stem. In absence of an etymology, it is considered to be "possibly pre-Roman" by DES (I: 483). An indication for this might be its ending in *-i*, which is not frequent in the inherited lexicon (§ 8.1.5.1). The alternation between *dúri* and *dúle* is unexpected. Puddu (2023: s.v. *dúle*) does not provide a location for the form *dúle*. For lack of sufficient evidence, no more can be said about this word.

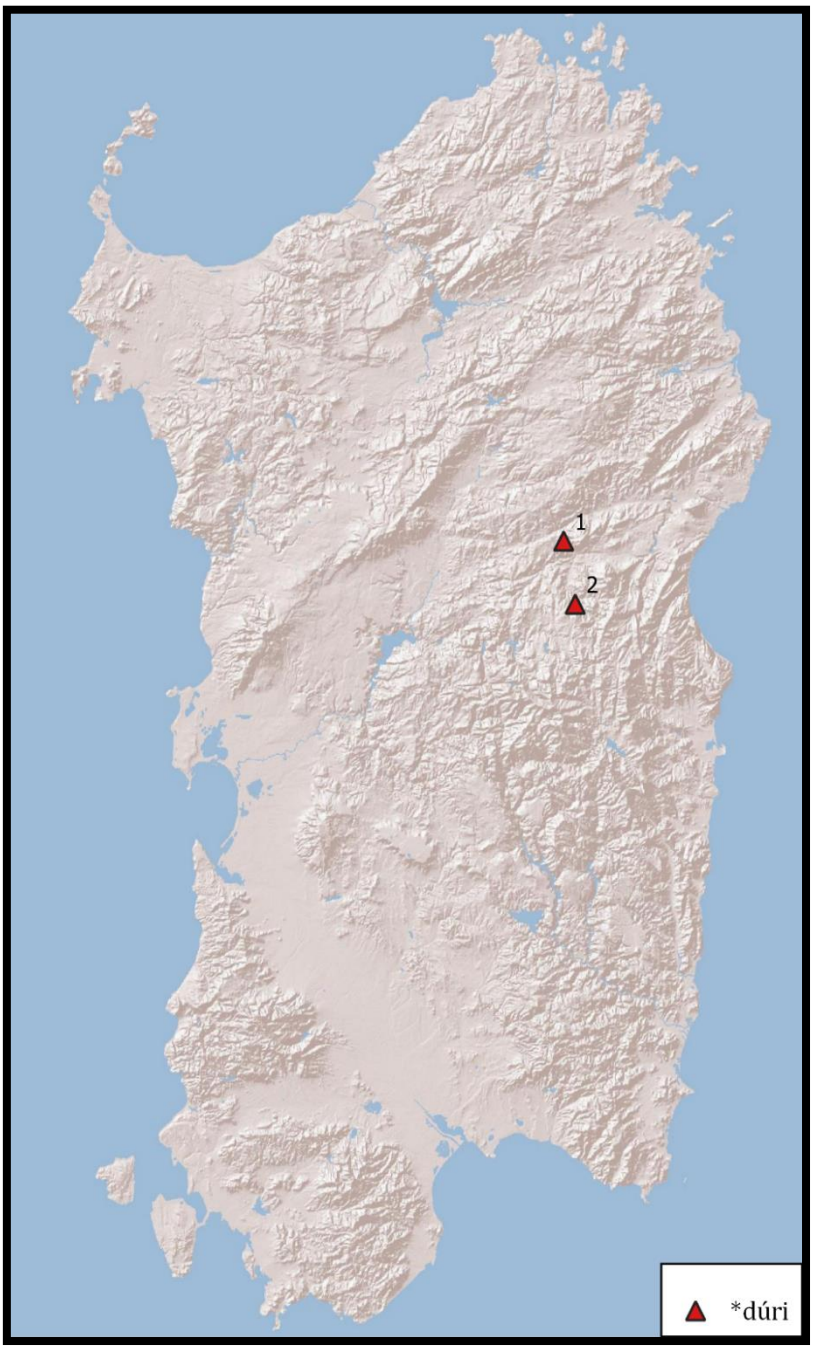

Figure 7.2.3: Forked juniper trunk

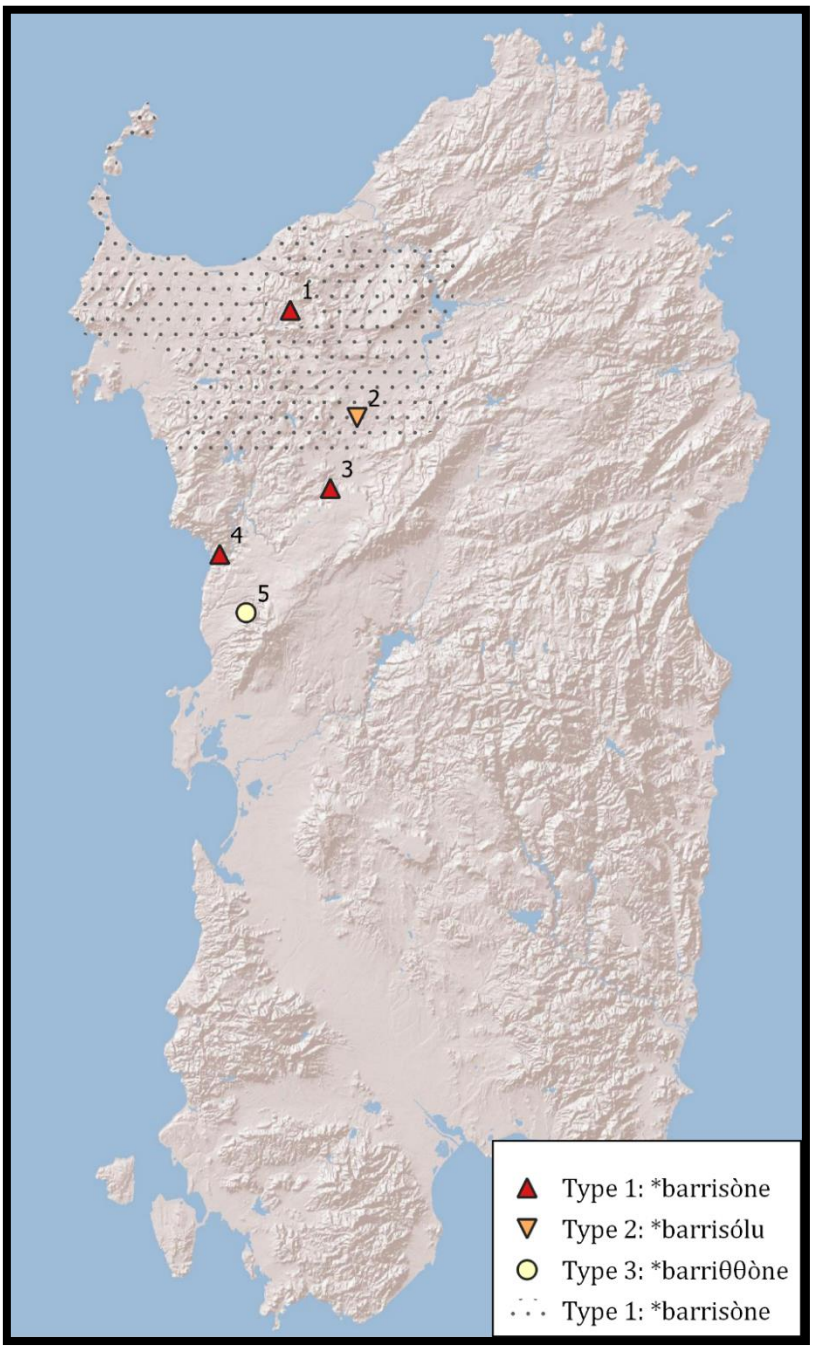

Figure 7.2.4: Tangle, bundle of wood

7.2.4 Tangle, bundle of wood

Reconstruction	Forms
1: *barrisòne	DES (I: 181–182): ***barraṡòni*** (Sass.) 'tangle of brushwood or shrubs used to seal off a gate', ***barraṡòne, bariṡòne*** (NLog.), ***berriṡòne*** (1: Osilo, 3: Bonorva) 'bundle of thorns; quantity of brushwood used to seal off breaches or entrances in the walls of a homestead'
	Spano (1851: 203): ***erisòne*** (4: Bosa) 'bundle of tangled wood used to seal off a passage'
2: *barrisólu	DES (181–182): ***barraṡólu*** (2: Mores) 'id.'
	Casu (2011: s.v. *arresólu*, *barrasòne*): ***arresólu*** 'bundle of thorns'; ***barrasólu*** 'hawthorn (*Crataegus monogyna*)'
3: *barriʒʒòne	DES (181–182): ***barrittsòne*** (5: Cuglieri) 'small load of wood that is transported by donkey'

Most forms listed above refer to tangles or bundles of spiny or impenetrable brushwood that are used to block some passage.[242] The forms in type 1 can most easily be derived from **bar*(*r*)*isòne*, with various vowel assimilations (e.g. in *barraṡòne* and *berriṡòne*) and loss of the initial *b-* due to generalization of post-vocalic forms (e.g. *erisòne*). Type 2 contains the same element **barris-*, but suffixed with *-ólu* < Lat. *-eolus* instead of the *-òne* found in type 1. Type 3 *barrittsòne* (5: Cuglieri) is likely the same as type 1, but with replacement of *-iṡ-* by diminutive *-itts-* (cf. Wagner 1952: 28).

These words, which Wagner (DES I: 182) considers to be "certainly pre-Roman", have previously been compared to Ibero-Romance forms of similar form and meaning. Notable comparanda are Pg. *barda* 'fence made of intertwined branches and brambles', Sp. *barda* 'cover put over the fences of corrals', Cat. *barda* 'portable barrier to lock up livestock' etc., Cat. *bardissa* 'thicket; hedge', among various others (DCEC I: 510; DECLC I: 644 45; Paulis 1992: 22). After Rohlfs (1927b: 402), Corominas (DCEC I: 510b) additionally compares Bq. *barra* 'bar', hypothesizing its original meaning to have been 'barrier' and noting that the alternation *-rr-* ~ *-rd-* is common in Iberian words of pre-Roman origin. However, Michelena (1990: 269) regards *barra* as an originally Romance word. At any rate, further probable comparanda are Occitan forms like OProv. *barta* 'bush, brushwood', Lang. *barto* (Lot-et-Garonne dep.) 'hedge', *bartas* (Aude dep., Hérault dep.) 'hawthorn' etc., which FEW (I: 262) reconstructs as **barrăt-*/**barrit-*. The proposal by

[242] Paulis (2008: 18–19) argues that the medieval Sardinian personal name *Barisone* is to be identified with this word.

FEW (I: 262) that this word is from Gaulish **barros*, and cognate to MW, OIr. *barr* 'point, top, crest' is semantically far-fetched. More likely is Rohlfs' (1927b: 402; cf. also DCEC I: 510b, DECLC I: 645b) suggestion that it is of pre-Indo-European origin.[243]

The semantic correspondence between the Sardinian and the Ibero-Romance and Occitan forms is striking. For one, they all refer to fences or barriers made of tangled or intertwined plant material. If we assume a primary base **barr-/*bard-*, it must have been extended with an element **-is-* in Sardinia. This apparent suffix is attested in several other words of pre-Roman origin (§ 9.2.8).[244] A Catalan loan origin for Srd. *barrišòne* etc. can at any rate be excluded on formal grounds, not least in view of the Algherese manifestation *baldíssa* (Alcover & Moll 2002: s.v. *bardissa*). As a consequence, the idea of a shared pre-Roman origin of the various forms discussed is plausible. We may conclude that this word forms part of a set of lexical correspondences connecting Sardinia, the Iberian Peninsula and southern France (§ 11.2.1).

[243] Specifically of "Iberian" origin. Rohlfs (1927b) does not make a distinction between Basque and the Iberian language. Cf. § 1.3.2.2. on the use of the term "Iberian".

[244] The suffix *-issa* of Cat. *bardissa* is of Romance origin (< Lat. **-icia*) and is thus not the same as Srd. **-is-*.

III
Structural features

8 Phonology

8.1 Non-inherited phonemes

One indication that a word is non-inherited, is when it contains a phoneme that cannot regularly continue a sequence present in the ancestor language (cf. § 2.1.1). This includes phonemes that occur exclusively in loanwords (such as /g/ and /θ/ in English loans in Dutch), and also native phonemes in contexts that usually preclude them, e.g. non-rhotacized *-s-* in Lat. *rosa* 'rose'. The Sardinian evidence for non-inherited phonemes consists mainly of the second category. Forms containing phonological irregularities are unlikely to be inherited from Latin. If these forms cannot be shown to have entered Sardinian from one of its historically known contact languages (Greek, Catalan, Spanish, Italian etc.), they are suspect of being pre-Roman in origin.

8.1.1 Non-inherited *ϑ

One of the phonemes that often occur in words of presumed pre-Roman origin, and which is at times used by DES as an indicator of "pre-Roman-ness", is Proto-Sardinian *ϑ. Its main reflexes across the Sardinian dialects are ϑ(ϑ) in the Nuorese dialects and the northern part of the Ogliastra, *t*(*t*) in Logudorese, and *t*(*t*)*s* in Campidanese. Geographically more restricted reflexes are *č* in Desulo and the Sulcis region, and *s*(*s*) in the southern part of the Ogliastra (Wagner 1941a: 107–109, 295). It is attested in Old Sardinian sources mainly as <th> or <ç>, and as <τζ> in the records using the Greek alphabet (Wagner 1941a: 106–107). In the body of inherited lexical items, *ϑ is the regular outcome of the (vulgar) Latin sequences *$t\underset{\cdot}{\imath}$ and *$c\underset{\cdot}{\imath}$, in which $\underset{\cdot}{\imath}$ can be either from prevocalic *ĕ* or *ĭ*; cf. Srd. *fáϑϑo*, *fátto*, *fáttsu* 'I make' < Lat. *faciō*, Srd. *púϑϑu*, *púttu*, *púttsu* '(water) well' < Lat. *puteus* (Wagner 1941a: 109).

On the basis of Sardinian-internal evidence, it is difficult to establish with certainty what the original realization of *ϑ was. Wagner (1941a: 108) follows Meyer-Lübke (1903: 21) in positing either an interdental fricative [θ] or an interdental affricate [tθ], both of which are found in the Nuorese dialects, as the most likely original value of *-ϑ-. It is not clear to me how an original realization [θ] is preferable over [ts], as found in Campidanese. It is in any case plausible that its realizations were once more uniform across Sardinia (but cf. Paulis 1984a: lxxiv–lxxvii). Logudorese *t* and Ogliastran *s* are secondary developments from ϑ and *ts* respectively (Wagner 1941a: 107–108). The remaining question is whether the

attested variation is the result of a shift [θ]/[tθ] > [ts] in Campidanese, or of a shift [ts] > [θ] in Nuorese and Logudorese. Based on the geminate outcomes of *ϑ in all Sardinian dialects and its correspondence to [ts] and [tʃ] in other Romance varieties (cf. § 8.1.1.1), an original affricate value [tθ] or [ts] seems more plausible than [θ]. For the purpose of reconstructing the earliest stages of Sardinian and of representing the attested phonetic diversity of this phoneme, the notation *-ϑ- is best understood as a graphical device for capturing the various outcomes of this single phoneme, rather than a certain phonetic reconstruction.

8.1.1.1 Non-inherited sources of *ϑ

Besides being a reflex of inherited sequences, *ϑ has several other sources. It occurs in historical loans from Greek; e.g. Srd. *ϑíu* etc. 'uncle' < late Lat. *thius* < Gr. θεῖος. It also occurs in various formations that are attested in Latin with *-z-*, or that correspond to [ts] or [tʃ] in other Romance languages; e.g. Srd. *ϑirriáre* 'to scream' vs. Sic. *zirriari* 'to squeak' etc. (Wagner 1941a: 110–111). Additionally, *-ϑ- sometimes occurs secondarily in inherited words: cf. Nuor. *aϑϑéϑϑu* 'wormwood' << **asséϑϑu* < **absentiu* << Lat. *absinthium* (DES I: 152); Log. *tilíbba* (Osilo, Ploaghe) 'bean pod' < **ϑilíbba* << **silibba* < Lat. *siliqua* (DES II: 417); Bitti *ϑèrpes*, Log. *tèrpe* 'snake' < **ϑerpe*(*s*) << Lat. *serpes* (DES II: 408). These occurrences of *ϑ are most easily understood as the result of assimilatory or dissimilatory processes, or through analogy with phonetically or semantically similar words containing *ϑ (Wagner 1941a: 117–120).

More relevant for the present study is the fact that *ϑ occurs in a large number of Sardinian words that lack a clear etymology. According to Wagner (1941a: 112–116), this subset mainly contains plant and animal names, although he mentions some counter-examples. It is not remarkable per se that a native phoneme occurring in the inherited lexicon also occurs in words without an etymology. However, *ϑ in non-inherited words tends to occur in phonological contexts where we do not expect inherited *ϑ to exist. One of these contexts is in word-initial position. The only regular sources for Sardinian *ϑ are Lat. *cį* and *tį*, which are absent from Latin onsets.[245] This means that the substantial number of words starting in *ϑ-, discussed in § 3 – 7, cannot regularly continue a Latin form. The same is valid for the Sardinian sequence **ϑi*, as it would require Lat. **cį̄i* or **tį̄i*, which did not exist. As a consequence, any instance of **ϑi* is either secondary or hails from an external source.

[245] The only potential instance is Lat. *cieō* 'to stir up', which is replaced by *citāre* already in Classical Latin and is thus not continued in Sardinian (de Vaan 2008: 113–114). Other instances are loans or names of foreign entities.

As mentioned, the phoneme **ϑ* is present in many un-etymologized Sardinian lexical items referring to native flora and fauna. In addition, it is frequent in toponyms, e.g. *Orϑullè* (It. Urzulei). This suggests that part of the instances of Srd. **ϑ* represent non-native phonemes in words that Sardinian Latin borrowed from local languages. The following is an overview of non-inherited lexical items that contain **-ϑ-*. For every instance, the most relevant dialectal forms are selected, and it is indicated when **ϑ* occurs in a position that cannot regularly continue any Latin sequence, i.e. initial *ϑ-* or *ϑi*. The order adhered to is that of § 3 – 7.

	Meaning	**Section**	**Forms**	**Reconst.**	**Initial *ϑ-**	****ϑi***
1	Arum	§ 3.1.2	*ϑoϑoróyu, sattsarói̯*	**ϑaϑarói̯*	✓	
2	Bristly ox-tongue	§ 3.1.4	*ϑiòkkoro*	**ϑiókkor-*	✓	✓
3	Celery, watercress	§ 3.1.6	*ϑurgúsa, tseliɣúsa*	**ϑur-/ϑeli-gúsa*	✓	
4	False fennel	§ 3.1.7	*tsikkiría*	**ϑikkiría*	✓	✓
5	Poppy	§ 3.1.15	*aϑáɳɖa, tϑáɳɖa*	**ϑánda-*	✓	
6	Esparto grass	§ 3.2.2	*ϑinníɣa, tsónni*	**ϑinníga, *ϑónni*	✓	✓
7	Knotgrass, horsetail	§ 3.2.4	*síntsiri, sintsurru*	**ϑ/sínϑur-*	?	?
8	Holly/Butcher's broom	§ 3.3.6	*alaϑúli, latϑòri*	**alaϑúli, *laϑòre*		
9	Osier	§ 3.3.10	*ϑòa, ϑòɣa, ϑòβa*	**ϑó(C)a*	✓	
10	Rosemary	§ 3.3.14	*ϑíppiri, tsíppiri*	**ϑíppiri*	✓	✓
11	Strawberry tree flower	§ 3.3.17	*aíϑϑo*	**aíϑϑo*		
12	Traveller's joy	§ 3.3.19	*a(u)ttsára, luttsára*	**(a)uϑϑára*		
13	Resin	§ 3.4.5	*tsònkine*	**ϑónkine*	✓	
14	Sprout, bud	§ 3.4.6	*tséurra*	**ϑérula*	✓	
15	Backswimmer	§ 4.1.2	*tsirifrúči*	**ϑirifúrki*	✓	✓
16	Dragonfly, grasshopper	§ 4.1.9	*tsimpilíɣe*	**ϑimpilíke*	✓	✓
17	Earthworm	§ 4.1.10	*ϑilingròne, tsilíngu*	**ϑiling(ul)-one*	✓	✓
18	Grasshopper	§ 4.1.11	*piβiϑϑíu, piβintsíri*	**pipíϑϑ-*		✓
19	Grasshopper	§ 4.1.12	*ϑilipírke*	**ϑilipírke*	✓	✓
20	Slug, snail	§ 4.1.15	*sittsiɣórru, tiϑiɣórru*	**ϑiϑikórnu*	✓	✓
21	Worm	§ 4.1.17	*ϑorroíɣu*	**ϑVrro(C)íku/a*	✓	
22	Eel	§ 4.2.1	*tsingòrra*	**ϑingórra*	✓	✓
23	Gecko	§ 4.3.1	*tilíbbu*	**ϑilíkwu*	✓	✓
24	Lizard	§ 4.3.2	*ϑilikèrta*	**ϑilikérta*	✓	✓

	Meaning	Section	Forms	Reconst.	Initial *ϑ-	*ϑi
25	Lizard, snake	§ 4.3.3	*tsorómpis*	**ϑor(r)ómpi(s)*	✓	
26	Skink, slug	§ 4.3.4	*ϑilikúkku*, *ϑaϑϑalúkka*	**ϑilikúku*, **ϑaϑalúk(k)a*	✓	✓
27	Kestrel	§ 4.4.4	*tserpeḍḍéri*	**ϑerpelléri*	✓	
28	Kestrel	§ 4.4.5	*attiliβríu*, *tilibríkku*	**ϑiliprí(kk)u*	✓	✓
30	Kite	§ 4.4.6	*ϑurulía*, *sittsulía* etc.	**ϑurolía*, **ϑiϑϑolía*	✓	✓
31	Bat	§ 4.5.2	*ϑuϑuréḍḍu*, *ϑiϑirṛiólu*, *tsirṛiólu*	**(ϑi-)ϑurr-*	✓	✓
32	Rocky plain	§ 5.1.22	*tsèppara*	**ϑèppara*	✓	
33	Water basin, source, stream	§ 5.1.27	*ϑúrgalu*	**ϑúrgalu*	✓	
34	Bran	§ 6.1.1	*ϑálau*	**ϑála(C)u*	✓	
35	Stack of wheat	§ 6.1.7	*tsèrga*	**ϑérga*	✓	
36	Blind	§ 7.1.1	*ϑúrpu*, *túrpu*, *tsúrpu*	**ϑúrpu*	✓	
37	Bristle	§ 7.1.2	*ϑúḍḍa*, *túḍḍa*, *tsúḍḍa*	**ϑúlla*	✓	
38	Neck	§ 7.1.4	*ϑúkru*, *ϑúɣulu*, *tsúɣu*	**ϑúk(u)lu*	✓	
39	Film, layer	§ 7.2.2	*ϑídza*, *ϑòdza*	**ϑília*, *ϑòlia*	✓	✓

Table 8.1.1: Pre-Roman words containing **ϑ*

Thirty-four out of the thirty-nine potential pre-Roman words containing **-ϑ-* start in **ϑ-*, and fourteen contain the sequence **ϑi*. Neither of these can regularly continue any sequence conforming to Latin phonotactics. It is therefore likely that the pre-Roman languages of Sardinia contained a phoneme whose phonetic realization was close to the original phonetic realization of Sardinian **ϑ* < Vulg. Lat. **cį*, **tį*. Part of the evidence shown in Table 8.1.1 has been explained as loans from Punic or Berber, the latter likely in the same Punic colonial context (see § 10.2). The remainder of the etyma may have originated in some other, unidentified, pre-Roman language. These two groups of words will be discussed in § 8.1.2 and § 8.1.3 respectively.

8.1.2 **ϑ* in Punic and Berber loans

Among the words containing **ϑ*, several have been attributed with varying degrees of confidence to a Punic stratum, having potential cognates in Semitic or

Berber.[246] These words are *tsikkiría* 'false fennel (4), *tsónni* 'esparto grass' (6), *síntsiri* 'knotgrass, horsetail' (7), *ϑíppiri* 'rosemary' (10), and *tséurra* 'sprout, bud' (14). A Punic origin has also been considered for *tsingòrra* 'eel' (22), but this is less certain.[247] We can compare the Sardinian forms as well as their reconstructions to their proposed Phoenico-Punic, other Semitic, or Berber cognates:

Srd. form	**Section**	**Srd. reconst.**	**Pun./Sem./Bb. comparandum**
tsikkiría 'false fennel'	§ 3.1.7	**ϑikkiría*	Pun. σικκιριά 'dill' (Dioscorides)
tsónni 'esparto grass'	§ 3.2.2	**ϑónni*	PBb. **sănV̆y/*sūnV̆y/*sănī/*sūnī* 'basket (made of esparto grass)'
síntsiri, *sintsurru* 'knotgrass, horsetail'	§ 3.2.4	**ϑinϑur-/*sinϑur-*	Pun. *sensur*, *zenzur* 'knotgras'
ϑíppiri, *tsíppiri* 'rosemary'	§ 3.3.14	**ϑíbbir*	Pun. *zibbir* 'rosemary'
tséurra 'sprout, bud'	§ 3.4.6	**ϑérula*	Pun. ζερα, Hebr. *zéra'* 'seed'
tsingòrra 'eel'	§ 4.2.1	**ϑingórra*	Akk. *singurru* 'kind of fish'

Table 8.1.2: Punic loans containing **ϑ*

It follows from the forms in table 2 that "Punic" *-s-* and *-z-* were borrowed as **-ϑ-* in Sardinian. It is important to stress that the lines of evidence for Punic loans into Sardinian consist of different types. These include words attributed by Greek and Roman authors to Punic (σικκιριά, *zibbir*, *zenzur*), attested Punic words with Semitic cognates (ζερα), and words with Berber comparanda (**sănV̆y/*sūnV̆y/*sănī/*sūnī*). Regardless of the ultimate provenance of these words, the substitution of *-s-* and *-z-* with **-ϑ-* is striking: the local variety of Latin could have employed its native sibilant *-s-* to represent sibilants in borrowed words, but it clearly did not.

8.1.2.1 Sardinian **ϑ* and Semitic sibilants

This treatment of sibilants in borrowings from Punic can be accounted for by assuming that the proto-Semitic phoneme **-z-* had retained its original affricated value [dz] (or [zd]) in the variety of Phoenician spoken in Sardinia. An affricated realization of **-z-*, as well as of voiceless **-s-* and emphatic **-ṣ-*, is in any case reconstructed for Proto-Semitic, Northwest-Semitic and Canaanite, to which Phoenician belongs (Kogan 2012b: 65–68). For Phoenician and the earlier stages

[246] For a discussion of the corpus of proposed Punic loans in Sardinian, see § 10.1. For a discussion of the proposed Berber loans in Sardinian and the context in which they were borrowed, see § 10.2.

[247] The Punic proposed for Camp. *míttsa* 'source' is probably no longer necessary (cf. § 5.2.5).

of Punic, **-z-* is taken to represent [dz], whereas it was written as <s> in Neo-Punic and occasionally as <z, ζ> by non-native speakers (Krahmalkov 2001: 21–22, 24–25). Among the proposed Punic loans into Sardinian Latin, the only one with a certain Semitic etymology, i.e. Pun. ζερα 'seed', contains **z-*. The three words labeled "Punic" in classical sources are σικκιριά, *zibbir*, and *sensur*/*zenzur*. Whatever the ultimate origins of these words, it is possible that they contained the same phoneme **z-*.

If the presence of **ϑ* in Sardinian words of Punic origin is indeed taken as evidence for an affricated realization of Semitic **z*, this implies that the variety of Punic, that was the source of the Sardinian punicisms, did not exhibit the merger of Semitic **s, *z, *ṣ, *ś* into /s/.[248] This merger is observed in Neo-Punic (Krahmalkov 2001: 14), and thus post-dates the Roman conquest of Sardinia by almost a century. One possibility is that Punic loans entered Sardinian Latin in the earliest stages of the Roman presence in Sardinia. Alternatively, the Sardinian variety of Punic preserved some of the contrasts between affricates and sibilants, after these had been lost in other varieties of Punic. The likelihood of the first scenario depends on one's view of the speed with which Sardinia was linguistically romanized (see § 1.2).

On the basis of the Sardinian data, it is difficult to determine what the local treatment of Punic **s*, **ṣ*, and **ś* was.[249] It cannot be ruled out that the Sardinian punicisms without a Semitic etymology continue one of these other sibilants. Although the Sardinian evidence is slim, an independent development of Sardinian Punic cannot be excluded and would be unsurprising. The relative isolation of the Punic speech community in Sardinia from the rest of the Punic-speaking world is demonstrated by a monolingual Punic inscription from the late 2nd or early 3rd c. CE in Bitia, which still uses the Punic script rather than the Neo-Punic script that had emerged elsewhere (Adams 2003: 210).

8.1.2.2 **ϑ* and Berber sibilants

One aspect that remains to be explored within this context is whether the Sardinian outcome **-ϑ-* in Punic borrowings containing sibilants can also be elucidated from a Berber perspective. Given that various of the proposed punicisms

[248] That is, at least not in the same way with /s/ as the final outcome. It is still possible that two or more of these phonemes merged into an affricate /ts/ or /dz/.

[249] Emphatic **ṣ* does occur in the root **yṣ'*, proposed by Wagner (1997: 159; 1957: 105–106; DES II: 121) to be the origin of Srd. *míttsa* 'source' (§ 5.2.5). If Srd. *míttsa* is from **mītia* < Lat. *mītis* 'ripe, mellow' however (Guarnerio 1906: 245; Mensching 1997), we have no evidence of the treatment of Punic *ṣ* in Sardinian.

in Sardinian lack a Semitic etymology, but are potentially continued in Berber (e.g. Pun. σικκιριά 'dill', Tsh. *taskra* 'teasel' [§ 3.1.7]; Pun. *zibbir*, Rif., Tmz. *azir* 'rosemary' [§ 3.3.14]), there is a possibility that these Punic words are ultimately not of Phoenician (i.e. Semitic), but rather of indigenous North African origin.[250]

On the basis of comparison between the Berber languages, Kossmann (1999: 249; 2020a: 20; 2020b: 879) reconstructs three proto-Berber sibilants, **s*, **z*, **ẓ*, that could be either singleton or geminate. For the singleton series there is no evidence that they were ever affricates. Their geminate counterparts are generally distinguished by quantity only, although some varieties show an affricate realization of geminate **s*, **z*, **ẓ* (Kossmann 2020b: 884). Regardless of whether this affrication in geminates in certain Berber varieties is an archaism, none of the proposed Berber comparanda contain a geminate sibilant. It thus seems that the Proto-Berber phonological system, at least according to our current understanding, cannot easily account for the realization of Punic sibilants in Sardinian as **ϑ* rather than as **s*. Punic is therefore a more suitable candidate for being the direct donor of the words in Table 8.1.2 than Proto-Berber.

In Libyco-Berber inscriptions from Dougga however, there is evidence for seven graphemes representing seven phonemically distinct sibilants (Kossmann 2020b: 879–883). Whether this is to be interpreted in a similar way to the Proto-Berber system — e.g. four sibilants, three of which are differentiated for length — or whether the language recorded in these inscriptions knew more sibilant contrasts than those reconstructed for Proto-Berber, is hard to establish based on the available evidence (Kossmann 2020b: 885–886). Thus, whereas it is unlikely that the Sardinian punicisms containing **-ϑ-* were borrowed into Sardinian Latin directly from Proto-Berber or its descendants, the possibility of a North African language related to Proto-Berber being spoken in Sardinia by Punic settlers cannot be ruled out on the basis of the Sardinian treatment of borrowed sibilants. See § 10.2 for a discussion of the proposed Berber loans in Sardinian.

---

[250] For such an early period, it is probably rather anachronistic to speak of "Berber". The ancestor of the modern Berber languages was undoubtedly present in the region at the time, but it cannot be excluded that any hypothetical North-African language in contact with Carthaginian Phoenician was a relative of Proto-Berber rather than descendent from it (cf. Louali & Philippson 2004: 111–114). The question whether the "Punic" population in Sardinia was predominantly Phoenician speaking or whether speakers of indigenous languages of North Africa were represented too, is addressed in § 10.1.

8.1.3 Pre-Roman *ϑ

The remainder of the non-inherited words containing *ϑ cannot be explained in terms of Punic loans. Their origin must be sought in the other language(s) that were spoken in Sardinia at least up until its romanization. The Sardinian dialectal developments of *ϑ in non-inherited, non-Punic words are identical to those of *-ϑ- in inherited and Punic words, where it may have represented an affricate (cf. § 8.1.2). It therefore stands to reason that pre-Roman *ϑ represents a phoneme of similar nature in the indigenous Sardinian pre-Roman language(s) that Latin was in contact with.

Words containing *ϑ occur across Sardinia. It occurs both word-initially (e.g. *ϑilipírke*, § 4.1.12) and word-medially (e.g. *aíϑϑo*, § 3.3.17), and can be followed by any vowel (e.g. *aϑáṇḍa*, § 3.1.15; *tsèppara*, 5.1.22; *ϑilipírke*, 4.1.12; *ϑòa*, 3.3.10; *ϑúkru* § 7.1.4). Whereas there is some evidence that pre-Roman Sardinian words could end in a consonant (§ 8.4.1), there is no evidence for *ϑ in word-final position. The only instance of pre-Roman *ϑ occurring in a cluster is *ϑráṇḍa* 'poppy', but its variants *tϑáṇḍa* and *aϑáṇḍa* suggest that *-r-* in this context is secondary. The distribution of *ϑ in the non-inherited lexicon gives the impression that the sound(s) it substituted was higher in frequency in its pre-Roman source language(s) than the clusters *-tį̇-*, *-cį̇-* were in Latin.

On the basis of these data, we can tentatively posit the existence of an affricate phoneme in language varieties throughout Sardinia, including in those spoken in the Punic colonial context (i.e. Phoenician and potentially an indigenous North-African language). Whatever the exact phonetic realization, it eventually merged with the regular reflex of Latin *-tį̇-*, *-cį̇-*. Finally, multiple Sardinian words starting in *ϑ- have been hypothesized to contain a prefix *ϑ(*i*)-. This element would have been part of the morphology of the pre-Roman source language of the words in question (Guarnerio 1904a: 68; Wagner 1932: 223–225; Wagner 1941a: 113; Wagner 1997: 263). The evidence for this putative prefix is discussed in § 9.1.2.

8.1.4 Non-inherited *ts* and *č*

In addition to words exhibiting the Sardinian-internal regular triad *ϑ – t – ts*, i.e. the reflexes of Proto-Sardinian *ϑ, there is a smaller set of words exhibiting different non-inherited affricates. Specifically, we find various words containing *ts* and *č* in the Nuorese and Logudorese dialect areas, where *ϑ regularly yields *ϑ* and *t* respectively. Strikingly, all of these words refer to animals, thus falling into a sharply delimited semantic category. The following is an overview of words containing irregular *ts* or *č*:

	Meaning	Section	Forms	*ts*	*č*
1	Trout fry	§ 4.2.3	*tsúrra*	✓	
2	Viperine snake	§ 4.3.5	*lírčis, lúrtsi, sulurtsi*	✓	✓
3	Viperine snake	§ 4.3.6	*tártsa*	✓	
4	Robin, yellow wagtail; hare	§ 4.4.9	*príńčiri, bríńčis(i), bríši, brinkíḍḍi*		✓
5	Old sheep	§ 6.1.5	*tsúrra, čúrra*	✓	✓

Table 8.1.3: Irregular occurrences of *ts* and *č*.

Wagner (1941a: 111) notes that there are several forms with *ts* in the Logudorese and Nuorese dialects that are of sound-symbolic nature. This offers an explanation for the above set of forms as well, especially since sound-symbolism is frequent in animal names. A further indication is the variation between *ts* and *č* in clearly related forms; e.g. *lírčis* ~ *lúrtsi*; *príńčiri* ~ *brísi* ~ *brinkíḍḍi* etc. However, DES explicitly claims a pre-Roman origin for some of the words listed here, i.e. *tsúrra* 'trout fry', *bríńčis* etc. 'robin, yellow wagtail; hare', and *tsúrra*, *čúrra* 'old sheep'. For *lírčis*, *lúrtsi* too, a pre-Roman origin is plausible because of the presence of other non-inherited features, such as the apparent suffix *-is* (§ 9.2.8).

In any case, *ts* and *č* in the aforementioned Logudorese and Nuorese forms cannot be equated with the phoneme **ϑ* in pre-Roman Sardinian. If *ts* and *č* are of pre-Roman origin, they must therefore represent another pronunciation. In light of the attested forms, an affricate would be the most likely source. However, it is unclear in what way it was distinct from **ϑ*. The paucity of the evidence does not permit any conclusions to be drawn as to a pre-Roman origin of these sounds. Some instances are likely the result of sound-symbolism. Others, such as *tsúrra*, *čúrra* 'old sheep' may be more recent loans (cf. § 6.1.5).

8.1.5 Unexpected final vowels

As a Romance language, the endings of the various parts of speech in Sardinian are produced by their ancestral Latin forms. In the nominal domain, Sardinian generally continues the Latin accusative forms. The most common Latin sources for nominal endings are *-am* (sg.), *-ās* (pl.) for the first declension, *-um* (sg.), *-ōs* (pl.) for the second declension, and *-em* (sg.), *-ēs* (pl.) for the third declension. These are regularly reflected in Sardinian as *-a* (sg.), *-as* (pl.); *-u* (sg.), *-os* (pl.), and *-e* (sg.), *-es* (pl.). With singular *-a*, *-u* and *-e* being by far the most common nominal endings, Sardinian nouns ending in *-i* and *-o* deserve some special attention.

8.1.5.1 Final *-i*

The third nominal declension of Latin usually had accusative singular *-em*, but it contains a small number of original *i*-stems that have *-im* as their acc. sg. end-

ing.[251] This is occasionally continued into Sardinian; e.g. *síti* (Bitti) 'thirst' < Lat. *sitis* 'id.' (DES II: 421). In other cases, original *-im* has been replaced by more common *-em* > *-e*; e.g. *frèbe* (Nuoro) 'fever' << Lat. *febrim* 'id.' (DES I: 544). Thus, the only regular inherited source of word-final *-i* is a small, closed group of Latin *i*-stems. This means that any Sardinian word ending in *-i* is suspect of not being inherited from Latin.[252] Many of these are clearly recent loans; e.g. most nouns ending in *-éri* can be explained from the Catalan agentive suffix *-er* (Wagner 1997: 224).

Nevertheless, there are quite a few etymologically obscure nouns in the Sardinian lexicon that end in *-i*. Many of these have been argued to be of pre-Roman origin in previous scholarship. The following is an overview of the words discussed in § 3 – 7 that end in an unexplained *-i*. Any relevant word-final vowel alternations are mentioned too.

	Meaning	**Section**	**Forms**	**Reconst.**	**Alt.**
1	Hemlock	§ 3.1.11	*biḍḍúri, buḍḍúri, meḍḍure, muḍḍuru*	**vullúri*	*-e, -u*
2	Mullein	§ 3.1.14	*lókri*	**lókri*	
3	Wild teasel	§ 3.1.22	*billótti, billóttiri*	**biḷḷótti*	
4	Butcher's broom, holly	§ 3.2.1	*aláši, alášḭu, aláše*	**aláse/-i*	*-e, -i̯a, -i̯u*
5	Esparto grass	§ 3.2.2	*tsónni, tsónni̯a*	**ϑónni*	*-i̯a*
6	Ropegrass	§ 3.2.5	*karkúri, krukkúri, krukkúri̯u*	**kalkúri*	*-i̯u*
7	Dog rose	§ 3.3.3	*orrolári, orrolári̯u*	**rolári*	*-i̯u*
8	Holly	§ 3.3.6	*alaϑúli, latϑòri*	**alaϑúli*	*-e*
9	Holly	§ 3.3.7	*olósti, golóstḭu, golóstri, golóstru*	**golósti*	*-i̯u, -u*
10	Hop-hornbeam	§ 3.3.8	*aúrri, árrui*	**á(C)urri/ *áruli*	
11	Maple	§ 3.3.9	*kósti, kóstiu*	**kósti*	*-i̯u*
12	Rosemary	§ 3.3.14	*ϑíppiri, tsíppiri*	**ϑíbbir*	

[251] Some of these are already attested with *-em* in the Classical period (Weiss 2009: 244 fn. 15).

[252] Note that in Campidanese dialects, word-final *-e* and *-o* merge with *-i* and *-u* respectively. The original quality of the final vowel is revealed by any *-e-* or *-o-* in the preceding syllable. These are open *-è-* [ɛ], *-ò-* [ɔ] when originally followed by *-e* or *-o* (as well as *-a*), and closed *-é-* [e], *-ó-* [o] when originally followed by *-i* or *-u*; e.g. Srd. *bónu* 'good (sg.)' < Lat. *bonum*, Srd. *bònus* 'good (pl.)' < Lat. *bonōs* (Wagner 1941a: 11).

	Meaning	Section	Forms	Reconst.	Alt.
13	Smilax	§ 3.3.15	*tétti, téttị̈u, tèttu*	**tétti*	*-ị̈u, -u*
14	Yew	§ 3.3.22	*éni*	**éni*	
15	Grasshopper	§ 4.1.11	*piβittsíri, piβiϑϑíu*	**pipiϑϑ-*	*-ị̈a, -ị̈u*
16	Trout fry	§ 4.2.2	*gróri, gróli*	**gróri*	
17	Lizard, snake	§ 4.3.3	*zorrómpi, tsorómpis*	**ϑor(r)ómpi(s)*	
18	Viperine snake	§ 4.3.5	*lúrtsi*	**lúrtsi*	
19	Robin, yellow wagtail; hare	§ 4.4.9	*príncˇiri, brínčiṡi, brínči, brinkíḍḍi* etc.	**brink/či-*	
20	Robin, wren	§ 4.4.10	*gríṡi, kiríṡi, kiríu, kríḍḍi, kérri*	**kir-/kr-*	*-u*
21	Wren	§ 4.4.12	*ḍḍóḍḍi, dóḍḍi*	**loll-*	*-o*
22	Bat	§ 4.5.1	*čížini*	**kikin-*	
23	Fawn	§ 4.5.5	*bítti, bíta, bètte, bétu*	**bitt-*	*-a, -e, -u*
24	Hare	§ 4.5.8	*léppuri, lèpparu, lèppere, lèporo*	**léppVr-*	*-e, -o, -u*
25	Rock outcropping	§ 5.1.21	*tónneri*	**tónneri*	
26	Acorn bread	§ 6.2.1	[*pan'*] *ispéli*	**ispéli*	
27	Forked juniper trunk	§ 7.2.3	*dúri*	**dúri*	

Table 8.1.4: Sardinian words ending in **-i*

8.1.5.1.1 Paragogic *-i*

All of the forms listed above are attested with non-inherited final *-i*, although many exhibit an alternation in the final vowel. The connection between Srd. *ϑíppiri* 'rosemary' and Pun. *zibbir* 'id.', as well as the existence of *brínčis* 'hare' next to *brínčiṡi* 'robin',[253] shows that one possible source of final *-i* is paragoge after an originally word-final consonant. This is a well-known feature of Sardinian; e.g. *béssiði* < *béssit* < Lat. *exīt* '((s)he/it) exits' (cf. Wagner 1952: 57–62). If a word-final *-i* occurs with *-i-* in the preceding syllable, we may therefore consider a paragogic origin. Based on this criterion, other candidates for paragoge are:[254]

- *billóttiri* 'wild teasel'
- *piβittsíri* 'grasshopper'
- *príncˇiri, brínči* 'robin, yellow wagtail; hare'
- *gríṡi, kiríṡi, kríḍḍi* 'robin, wren'

[253] For the issue of the words for 'hare' and 'robin' being homonyms, cf. § 4.4.9.

[254] One potential further candidate is *éni* 'yew' (14), but only if one accepts its reconstruction as **ágini* and its comparison to Bq. (*h*)*agin* 'id.' (Bertoldi apud Wagner 1941a: 45). As discussed in § 3.3.22, this connection is rather tenuous.

- *čížini* 'bat'
- *bítti* 'fawn'

Of these, *billóttiri* (3), *piβittsíri* (15) and *príńčiri* (19) occur next to forms without final -(*i*)*ri*; *billótti*, *piβiʒʒíu*. Possibly, this is the same element as the "collective" suffix *-́Vr* (§ 9.2.6.2) first identified by Terracini (1927: 139, 142). For *gríśi*, *kiríśi* (20) and *bítti* (23) we find the variants *gríśu*, *kiríśu* and *bíta*, *bètte*, *bètta*. This alternation could be due to replacement of original *-i* with the more common endings *-a*, *-e*, *-u*; or perhaps they represent independent strategies to resolve a word-final consonant (cf. § 8.4.1). If the abovementioned words did indeed receive their final *-i* paragogically, we would have to posit the following original forms:

- **brínč* > *brínči* 'robin; hare'
- **grís* > *gríśi/u* 'robin, wren'
- **kirís* > *kiríśi/u* 'id.'
- **kríll* > *kríḍḍi* 'id.'
- **kíkin* > *čížini* 'bat'
- **bítt* > *bítti* 'fawn'

It goes without saying that some of these (viz. **brínč*, **kríll*, **bítt*) grossly violate the constraints on Sardinian syllable structure, and for *brínči* and *kríḍḍi* there is no evidence that favors positing **brínč*, *krill* over **brínči*, *krílli* with word-final *-i*. For *bítti* 'fawn', on the other hand, we find the variants *bítta*, *bètte*, *bètta*, *bétu* (§ 4.5.5). Neither the alternation of the stressed vowel, nor that of the final vowel is easily explained by means of Sardinian internal developments. The existence of *bítti* and *bètte*, but not ***bítte* or ***bétti*, suggests that *-i* and *-e* are both paragogic vowels, regardless of the origin of the *e*/*i*-alternation in the stem. The *-a* and *-u* in *bítta*, *bètta*, *bétu* may be secondary additions due to their relative frequency in the nominal lexicon.

8.1.5.1.2 Replacement of *-i*

For none of the other words ending in *-i* is there any reason to assume that it is not original. This indicates that the pre-Roman source language(s) allowed word-final *-i*. Many of these words have variants ending in *-i̯a*/*-i̯u*; e.g.:

- *aláśi* ~ *aláśi̯u* ~ *aláśe* 'butcher's broom, holly'
- *tsónni* ~ *tsónni̯a* 'esparto grass'
- *krukkúri* ~ *krukkúri̯u* 'ropegrass'
- *orrolári* ~ *orrolári̯u* 'dog rose'

- *olósti ~ golóstɟu ~ golóstru* 'holly'
- *tétti ~ téttɟu, tittɟòne ~ tèttu* 'smilax'

Interestingly, most of these endings in *-ɟa/-ɟu* cannot be particularly old, as inherited consonants followed by *-ɟ-* undergo various sound changes. For the forms listed above containing *-ɟa/-ɟu*, we would, according to the develoments specific to their places of attestation, rather expect: ***alášu* < *alásɟu*; ***tsònğa* < *tsónnɟa*; ***krukkúrğu* < *krukkúrɟu*; ***orrolárğu* < *orrolárɟu* (cf. Wagner 1941a: 107–109, 140–152). This means that the addition of *-a*, *-u* to these nouns originally ending in *-i* is secondary and of relatively recent date.[255] Similarly, the cases in which we find *-a*, *-e* or *-u* instead of *-i* (e.g. *alášu*, *golóstru*), are probably the result of the same tendency to provide words ending in "unusual" *-i* with more common endings, but with complete replacement of *-i* rather than adding the new vowel after it.

Other words ending in *-i* have variants with productive Sardinian suffixes that obscure the original final vowel; e.g. *čížini ~ čižinéḍḍu* 'bat'.[256] Finally, *ḍḍóḍḍi* 'wren' occurs next to *gròḍḍo* 'id.'. This may be due to contamination with the Sardinian personal name *Lòḍḍo*, which is also invoked by Wagner (1933: 503–508; DES I: 593) to explain *gròḍḍo*, *gròḍḍe* 'fox' (§ 4.5.6, § 8.1.5.2).

8.1.5.2 Final *-o*

A similar case are words ending in *-o*. In the Sardinian nominal domain, the known sources for word-final *-o* are few (Wagner 1941a: 37–38). They are largely restricted to:

- loans from Catalan, Spanish or Italian; e.g. Log. *òro*, Camp. *òru* 'gold' << It. *oro* 'id.', Log. *tesòro*, Camp. *tesòru* 'treasure' << It. *tesoro* 'id.' (Wagner 1941a: 37).[257]
- words going back to Latin ablatives or adverbs in *-ō*; e.g. Log. *dòmo*, Camp. *dòmu* 'house' < Lat. *domō* (abl.) 'id.' (DES I: 476-77), Log. *sèro* 'evening', Camp. [*eri*]*sèru* < Lat. *sērō* 'late (adv.)' (DES II: 408).
- words with a paragogic *-o*; e.g. Log. *kòro* 'heart' < Lat. *cor* 'id.' (DES I: 381).

---

[255] The realizations *-rɟ-* and *-nɟ-* are continued to this day in some limited areas (Wagner 1941a: 144). These can however not account for the irregular *-ɟa/-ɟu* endings found in other dialects.

[256] Since *čížini* is regular from **kíkin* with paragogic *-i*, the diminutive ending *-éḍḍu* may have been added directly to the base.

[257] In Campidanese, word-final *-o* and *-u* merge into *-u*, with the original quality preserved in the degree of aperture of any *-e-* or *-o-* in the preceding syllable.

Like in the case of *-i*, this means that any Sardinian noun ending in *-o* is suspect of not being inherited. The following is a list of etymologically obscure words ending in *-o*:

	Meaning	**Section**	**Forms**	**Reconst.**	**Alt.**
1	Bristly ox-tongue	§ 3.1.4	*ʒiòkkoro, sóčči̱ri*	**ʒiokkoro*	*-i*
2	Strawberry-tree flower	§ 3.3.17	*aíʒʒo, aíʒʒu*	**a(C)íʒʒo*	*-u*
3	Tree heather	§ 3.3.20	*iḍḍòstro, giḍḍóstru, ʔiḍḍòstra, giḍḍòstre*	**gillostr-*	*-a, -e, -u*
4	Juniper berry	§ 3.4.3	*bòḍḍoro, bòḍḍero*	**bòllVro*	
5	Walnut	§ 3.4.8	*kòkkoro, kòkkolo*	**kòkkoro*	
6	Wren	§ 4.4.12	*gròḍḍo, ḍḍóḍḍi*	**loll-*	*-i*
7	Fox	§ 4.5.6	*gròḍḍo, gròḍḍe, lòḍḍe, gróḍḍu*	**loll-*	*-e, -u*
8	Hare	§ 4.5.8	*lèporo, lèpparu, lèppore, léppuri* etc.	**leppVr-*	*-e, -i, -u*
9	Hillock	§ 5.1.6	*mòɣoro, mòɣora*	**mòkor-*	*-a*
10	Clay used to make acorn bread	§ 6.2.2	*tòrko, tròkko, trókku*	**tòrko*	*-u*
11	Evil spirit	§ 7.2.1	*trullío, turrío*	**turlío*	

Table 8.1.5: Words ending in **-o*

Most of these words exhibit alternations in their final vowel. In several cases, final *-o* has been replaced with *-u* (*aíʒʒu, giḍḍóstru, gróḍḍu, móɣuru, trókku*), which is the most common ending of masculine words.[258] The other alternations need to be accounted for individually. The word for 'fox' (*gròḍḍo, gròḍḍe, lòḍḍe* etc.) is explained by Wagner (1933: 503–508; DES I: 593) from the Sardinian personal name *Lòḍḍo* < *Lollo*. The *-e* in *gròḍḍe* 'fox' is probably from the vocative case of this name. One of the words for 'wren' (*gròḍḍo*) is identical to that for 'fox'. It is possible that *gròḍḍo* 'wren' somehow arose due to influence from either *gròḍḍo* 'fox', or directly from the personal name *Lòḍḍo*.

8.1.5.2.1 Paragogic *-o*

As discussed for final *-i*, some of the instances of word-final *-o* may be due to the addition of a paragogic vowel to an originally consonant-final word (cf. § 8.1.5.1.1). Candidates for this are *ʒiòkkoro*, *iḍḍòstro*, *bòḍḍoro*, *kòkkoro*, *lèporo*, *mòɣoro*, and *tòrko*, since their penultimate vowels are also *-o-*. In the case of

[258] This can perhaps also explain Fonni [*sa*] *ʔiḍḍòstra* 'tree heather' (4), with secondary introduction of feminine *-a*.

ϑiòkkoro and *lèporo*, assuming a paragogic origin for word-final *-o* is in line with the fact that the final vowel seems to follow the attested vowel alternations in the preceding syllable. We find *ϑiòkkoro* next to *sóčči* (§ 3.1.4), and *lèporo* next to *lèppere* (§ 4.5.8).[259] This suggests that we need to posit **ϑi̯okkVr* 'bristly ox-tongue', **leppVr* 'hare'. For the other forms, the evidence for final *-o* being paragogic is more sparse. However, Wagner (1932: 227 fn. 3) does propose a reconstruction **mókor* for *mòɣoro* 'hillock'.

For *iḍḍòstro* 'tree heather', forms ending in all vowels except *-i* are attested.[260] This could indicate that it originally ended in a consonant, and that various different strategies were applied to resolve this — i.e. *-o* due to paragoge, *-e* by analogy to inherited consonant stems, and *-a* and *-u* due to their general frequency in the nominal domain. The resulting reconstruction **gillostr* would phonotactically be aberrant from a modern Sardinian viewpoint. The same is true for a reconstruction **tòrk* > *tòrko*. While it cannot *a priori* be assumed that the presumed pre-Roman source language of these words disallowed such heavy codas, the evidence for *-o* being paragogic origin in *iḍḍòstro* and *tòrko* is circumstantial at best (cf. § 8.4.1).

In the forms *aíϑϑo* and *trullío*, *turrío*, *-o* cannot be of paragogic origin, and is thus likely original. From the discussed evidence, we can conclude that some pre-Roman contact language had a phoneme *-o-* that could occur in final position. No words of suspected Punic origin end in *-o*, but this may be due to chance, given their small number.

8.2 Consonant alternations

8.2.1 Quantity alternations

One of the types of consonantal alternations found in the presumed pre-Roman Sardinian lexicon, are quantitative alternations. Variants involving Sardinian voiceless and voiced plosives are included here too, since the Latin opposition between geminate and singleton stops is continued as a voicing opposition in most Sardinian varieties; e.g. intervocalic Lat. *-cc-*, *-c-*, *-g-* > Camp., Log. *-k-*, *-ɣ-*, *-Ø-* (cf. Wagner 1941a). Quantitative alternation is demonstrated by forms whose respective reconstructions show a discrepancy with regard to the voicing or

[259] There is also *lèppore*, *lèppuri* 'hare', which adhere more closely to the Latin vocalism in *lepus*, *leporem* 'hare', though still not corresponding regularly to it.

[260] This is all the more interesting since this word is one of two that contain the supposed pre-Roman suffix *-òst(r)-* (§ 9.2.9), the other being *olósti*, *golóstri* 'holly' (§ 3.3.7), which is predominantly attested with *-i*.

length of a consonant. Thus, one form of a certain word might require an original voiced stop (e.g. **g*), while another needs a voiceless stop (**k*), or a voiceless geminate stop (**kk*). The following is an overview of the material from § 3 – 7 whose consonantism exhibits an alternation in voicing, length, or both. The collection includes Sardinian words whose forms require conflicting reconstructions, as well as words whose reconstructions diverge from ostensible cognate forms in Latin. Cases in which the attested alternation can be explained by folk etymology or interdialectal borrowings, have been left out.

	Meaning	**Section**	**Form**	**Reconst.**	**Alt.**
1	Celery, fool's watercress	§ 3.1.6	*ϑurgúṡa, kuɣúṡa, tseliɣúṡa*	**-gúsa, *-kúsa*	*k* ~ *g*
2	Bermuda grass	§ 3.2.1	*alášu, ollášu, (n)oroḍḍáṡile*	**ol(l)ás-*	*ll* ~ *l*
3	Esparto grass	§ 3.2.2	*tsinníɣa, tinnía*	**ϑinník/ga*	*k* ~ *g*
4	Spider	§ 4.1.14	*solóiɣa, soloíɣa*	**solo(b)ika* (vs. Lat. *solifūga, solipūga*)	*p* ~ *(b/v)* *k* ~ *g*
5	Kestrel	§ 4.4.5	*tilibríkku, tilibríu* etc.	**ϑilipríkk/(g)u*	*kk* ~ *(g)*
6	Kite	§ 4.4.6	*tirolía, tsurulía, tsioḍḍía* etc.	**ϑi/urol(l)ía*	*ll* ~ *l*
7	Hare	§ 4.5.8	*lèpparu, lèporo, lèppore, léppuri*	**léppVr-* (vs. Lat. *lepor-*)	*pp* ~ *p*

Table 8.2.1: Words with voicing/length alternations in consonants

8.2.1.1 Quantity alternation between Sardinian and Latin

Of the words listed above, two agree in consonant voicing and length within Sardinian, but show discrepancy with regard to forms attested in Latin.

Sardinian	**Latin**	**Alternation**
Srd. *solóiɣa, soloíɣa* etc. < **solo(C)ika* 'spider'	Lat. *solifūga, solipūga, sālpūga* 'some venomous animal'	Srd. **(b/v)* ~ Lat. *f/p* Srd. **k* ~ Lat. *g*
Srd. lèppere, *lèporo, lèppore, léppuri* < **léppVr-* 'hare'	Lat. *lepus, leporis* 'hare'	Srd. **pp* ~ Lat. *p*

Table 8.2.2: Quantity alternations between Sardinian and Latin

Srd. *soloíɣa* is likely cognate to Lat. *solifūga, solipūga, sālpūga*, which are evident loans from a foreign language (cf. § 4.1.14). The modern Sardinian manifestations, as well as the identification of the Latin word as referring to a Sardinian animal by Isidore and Solinus (cf. Walde & Hofmann 1954: 470; DES II: 424-425),

suggest that a native Sardinian language was the source of the Latin forms. The modern Sardinian representatives require a reconstruction **solo(C)ika*, in which *-C-* could be any voiced stop (**b*, **d*, **g*) or **v*. Since the Latin forms show a labial consonant, it is reasonable to reconstruct either **solobika* or **solovika*. Although the Sardinian reconstruction is close to the Latin attestations, the latter clearly cannot be the direct source of the Sardinian forms. For the labial consonant we find *f* and *p* in Latin vs. **b* or **v* in Sardinian. For the velar consonant, we find Lat. *g* vs. Srd. **k*. At least two explanations can be offered for this. One is that there was confusion on the Roman side as to how to represent the phonology of the borrowed word (cf. Wigman 2023: 300). Another is that the pre-Roman word in question was impacted by sound change between the borrowing into continental Latin and the borrowing into local Sardinian Latin.

The case of the Sardinian word for 'hare' is complex (§ 4.5.8). The attested forms show variation in the unstressed vowels (Srd. *lèppere*, *lèporo*, *lèppore*, *léppuri* etc.) but agree on the labial geminate. We can thus reconstruct **leppVr-*. The word is undoubtedly related to Lat. *lepus*, *leporis* 'hare', but the irregular correspondence makes direct inheritance problematic.[261] In addition, several irregular correspondences are found in Massiliot Gr. λεβηρίς 'rabbit' < **leber-*, Siculian λέπορι̣ν 'hare' < **lepor-*, Pg. *láparo*, Fr. *lapéreau* 'young rabbit' < **lappar-* (Bertoldi 1937a: 146; Hubschmid 1943). These forms have been explained by Bertoldi (1937a: 146) and Wagner (DES II: 22) as independent borrowings of a pre-Roman form **lappar-* into the languages of the Mediterranean region, including Latin and its Romance daughter languages: the mentioned Romance forms require original geminate **pp*, while Latin has singleton *p*.[262] This suggests that a word **le/appVr* existed in some language(s) of the Mediterranean region. It must have entered the Romance languages substantially later than it did Latin and colonial Greek.

[261] Lat. *lepus* 'hare' itself is of uncertain provenance as well (De Vaan 2008: 335).

[262] Lat. *lepus* was inherited regularly in Sp. *liebre*, Fr. *lièvre*, It. *lepre* etc. (FEW V: 261).

8.2.1.2 Quantity alternation within Sardinian

The remaining words listed in Table 8.2.1 also show variation in consonant quantity within Sardinian, but they lack cognates in Latin.

8.2.1.2.1 **kk* ~ **k* ~ **g*

Of the five remaining cases, three show voicing/length alternation in plosive consonants. These are:

Variant 1	**Variant 2**	**Alternation**
Nuor. *ϑurgúśa*, *kuɣúśa*, Log. *turguśòne* < **-gúśa* 'celery; fool's watercress'	South Log. *tseliɣúśa* < **-kúśa* 'celery, fool's watercress'	Nuor., Log. **k* ~ Arborense **g*
Nuor. *ϑinníɣa*, Log. *tinnía* 'esparto grass'	Camp. *tsinníɣa* < **ϑinníka* 'esparto grass'	Camp. **k* ~ Nuor., Log. **g*
Log. *tilibríkku* < **ϑilipríkku* 'kestrel'	Log. *tilibríu* < **ϑilibrí(g)u* 'kestrel'	Log. **kk* ~ Log. **g*/Ø

Table 8.2.3: Plosive quantity alternations within Sardinian

There is no discernable pattern in the irregular alternations observed in the words listed above, apart from the fact that they all involve velars. The words for 'celery, fool's watercress' and 'esparto grass' show variation between **k* ~ **g*. The word for kestrel exhibits **kk* ~ **g*/Ø. Furthermore, the geographic distribution of the forms does not aid our understanding of this alternation. For instance, the variants of the word for 'kestrel' cooccur in the same dialect area. The evidence is insufficient to attribute this alternation to a pre-Roman substrate language.

8.2.1.2.2 **ll* ~ **l*

The last two instances of possible consonantal quantity alternation concern variants alternating between *l* and *ḍḍ*, the latter of which goes back to **ll*:

Variant 1	**Variant 2**	**Alternation**
Ogl. *alášu*, Camp. *olláśu* < **olás-* 'Bermuda grass'	Barb. *(n)oroḍḍáśile*, *oroḍḍáśu* < **ollás-* 'Bermuda grass'	Barb. **ll* ~ Ogl. **l*
Nuor. *ϑurulía*, Log. *tirolía*, *turulía*, Camp. *tsurulía* etc. < **ϑirolía* 'kite'	Camp. *tsioḍḍía*, *tsueḍḍía* etc. < **ϑi(r)ollía* 'kite'	Camp. **ll* ~ Nuor., Log., Camp. **l*

Table 8.2.4: Quantity alternation involving *l*

As in the case of the alternating stops discussed in § 8.2.1.2.1, there is no obvious pattern in the geographical distribution of this alternation. In the word for 'Bermuda grass', geminate **ll* is found in the Barbaricine dialects of Olzai, Gavoi and

Ovodda, while **l* is found farther to the south in the Ogliastra and the Sarrabus. Conversely, in the word for 'kite', geminate **ll* is restricted to some Campidanese dialects in the south, while **l* is found across the island.

On the basis of such scant and disparate evidence, we cannot draw any conclusions on the specificities of consonant quantity and voicing in the hypothetical pre-Roman source language of these words. It is possible speculate that the observed alternations result from incidental inconsistencies in the phonological adaptation to Sardinian Latin during the borrowing process. It is nevertheless interesting that consonantal voice/length alternation seems to play a much smaller role in the Sardinian pre-Roman lexicon than it does in other proposed linguistic substrate contexts, such as the pre-Greek lexicon (Furnée 1972: 115–200). This might suggest that the phonological system of the local pre-Roman language(s) was somewhat comparable to that of Latin in this specific regard, causing only minor phonological problems when pre-Roman loans were incorporated into the Sardinian Latin lexicon.

8.2.2 Quality alternations

The other parameter in terms of consonant alternations is that of consonant quality. When evident cognates of suspected pre-Roman origin cannot be derived from a single proto-form, at least according to the known diachronic developments of Sardinian, this potentially offers clues on the source language. For instance, when the reconstructed consonantism exhibits variants, this could be indicative of a phoneme that was alien to Latin, reflecting different strategies to adapt it to Latin phonology.

The following is an overview of the Sardinian etyma exhibiting a consonantal alternation involving other aspects than voicing or length. Alternations that can be explained as a folk etymology, or by common tendencies of Sardinian phonology, such as irregular nasal or *-r-*insertion (Wagner 1941a: 219–227), have been left out of consideration.

	Meaning	Section	Form	Reconst.	Alt.
1	Bristly ox-tongue	§ 3.1.4	*ϑiòkkoro, kòkkoro, čòkkoro, čóčči̯ri* etc.	**ϑi̯òkkor-/*kòkkor-/*ϑòkkir-*	**ϑ(i̯) ~ *k*
2	Heather	§ 3.3.5	*túvera, úv(v)ara* etc.	**(t)úfVr-*	**t ~ Ø*

	Meaning	Section	Form	Reconst.	Alt.
3	Traveller's joy	§ 3.3.19	*auttsára, alússara, luttsára,* etc.	**a(l)uϑara*	**l ~ Ø*
4	Lizard	§ 4.3.2	*aliɣèrta, (a)liğèstra* etc.	**(a)lVkèrta/ *(a)lVkèstra*	**rt ~ *str*
5	Viperine snake	§ 4.3.2	*lírčis, lúrtsi, sulurzi*	**lírčis/*lúrϑ/tsi*	**č ~ *ϑ/ts*
6	Kite	§ 4.4.6	*tirolía, sittsulía*	**ϑirolía/*ϑiϑϑulía*	**r ~ *ϑ*
7	Robin, yellow wagtail; hare	§ 4.4.9	*bríṅčis, brinkíḍḍi, bríśi* etc.	**bríṅč-/*brink-/*bris-*	**č ~ *k ~ *s*
8	Scops owl	§ 4.4.11	*ϑònka, ϑòğğe, ϑòrža* etc.	**ϑónka/*ϑólı̯e/*ϑóru̯a*	**nk ~ *lı̯ ~ *ru̯*
9	Wren	§ 4.4.12	*ḍḍóḍḍi, gròḍḍo*	**llólli/*glóllo*	**ll ~ *gl*
10	Fox	§ 4.5.6	*lòḍḍe, gròḍḍe,* etc.	**(g)loll-*	**l ~ *gl*
11	Mouflon	§ 4.5.10	*mufròne, muɣròne* etc.	**mufròne/*mugròne*	**f ~ *g*
12	Old skinny sheep	§ 6.1.5	*čúrra, giurrítta, tsúrra*	**čúrra/*ğúrra/ *tsúrra*	**č ~ *ğ ~ *ts*
13	Tilled land	§ 6.1.8	*tèle, tevèle*	**te(b/v)éle*	**b/v ~ Ø*
14	Forked juniper trunk	§ 7.2.3	*dúle, dúri*	**dúle, *dúri*	**l ~ *r*

Table 8.2.5: Words showing consonant quality alternations

As with the quantity alternations discussed in § 8.2.1, it is challenging to discern patterns common to the quality alternations. No alternation occurs more than once in the examined data, with the possible exception of *č ~ ts* in the words for 'viperine snake' (5) and 'old sheep' (12), which are discussed in § 8.1.4. Nevertheless, a few observations can be made.

One observation is that *č* often occurs in alternation with other phonemes (cf. 1: *čòkkoro*, 6: *lírčis*, 7: *bríṅčis*, 13: *čúrra*). As discussed in § 8.1.4, the alternation between non-inherited *ts* and *č* is most easily explained as the result of sound-symbolism. The alternation of *-č-* with other consonants (e.g. 1: *čòkkoro ~ ϑiòkkoro*, 7: *bríṅčis ~ brinkíḍḍi*) is perhaps of sound-symbolic nature too, but the semantic motivations for this need to be evaluated for each case.

Another observation is that in some words a consonant alternates with the absence of a consonant (2: *túvera ~ úvara*, 3: *alússara ~ auttsára*, 13: *tevèle ~ tèle*). However, as the relevant consonants are of divergent nature (respectively *-t-*, *-l-* and *-v-*), this can hardly constitute a single phenomenon.

The same can in fact be said for the entire set of alternations in Table 8.2.5. Since none of the alternations can be shown to recur in multiple lexical items, it is impossible to make inferences about the phonological dynamics of pre-Roman words entering Sardinian Latin. Some alternations could theoretically reflect dialectal differences in the pre-Roman language varieties of Sardinia, or reflect other variables related to language contact, such as difference in the time of borrowing. Other alternations may be of more recent origin, and due to sound-symbolic or folk-etymological nature within Sardinian. On the basis of the current data it is at any rate not possible to identify any originally pre-Roman alternations.

8.3 Vowel alternations

Taken at face value, the vowel alternations attested in the material discussed in § 3 – 7 do not show a clearer pattern than the qualitative consonant alternations discussed in § 8.2.2. Besides the plethora of forms showing vowel alternations that can be explained by internal Sardinian developments, there are a number of words whose vowel alternations are less transparent:

	Meaning	**Section**	**Form**	**Reconst.**	**Alt.**
1	Esparto grass	§ 3.2.2	*ϑinníya, tsónni*	**ϑinníga/*ϑónni*	*i-í ~ ó-i*
2	Bed bug	§ 4.1.3	*koròsta, kurústa*	**koròsta/*kurústa*	*o-ò ~ u-ú*
3	Cockroach, darkling beetle	§ 4.1.7	*siśáya, saśáya, sèśè, seiśèi* etc.	**sisái̯a/*sei̯sèi̯*	*á ~ è*
4	Grasshopper	§ 4.1.12	*ϑilipírke, ϑelapòrka* etc.	**ϑilipírke/*ϑelapòrka*	*i-i-í ~ e-a-ò*
5	Spider	§ 4.1.14	*soloíya, solóiya, suíya* etc.	**solo(b)ika*; cf. Lat. *solifūga, solipūga, sālpūga*	*o-o-i ~ o-i-ū ~ ā-Ø-ū*
6	Worm	§ 4.1.17	*tserríyu, tsirríyu, tsorroíyu*	**ϑVrro(C)íku/a*	*e-í ~ i-í ~ o-oí*
7	Lizard	§ 4.3.2	*aliyèrta, kalužèrtula, ϑalaʔèrta , ϑilikèrta*	**(ka/ϑi-)lVkerta*; cf. Lat. *lacerta*	*a ~ i ~ u*
8	Skink, slug	§ 4.3.4	*tsiliyúyu, ϑalakúʔa*; *sintsilúya, sattsalúya* etc.	**ϑilikúku/*ϑalakúka* **ϑiϑilúka/ϑaϑalúka*	*i-i-ú ~ a-a-ú*
9	Viperine snake	§ 4.3.5	*lírčis, lúrtsi*	**lírčis/*lúrtsi*	*i ~ u*
10	Kite	§ 4.4.6	*tirolía, ϑurulía, tsuiḍḍía, tsueḍḍía, tsuaḍḍía* etc.	**ϑirolía/*ϑurulía/ *ϑu(r)Vllía*	*i-o ~ u-u ~ ui ~ ue ~ ua*

	Meaning	Section	Form	Reconst.	Alt.
11	Fawn, female mouflon	§ 4.5.5	*bètta*, *bètte*, *bítta*, *bítti*	**bètt-*/**bítt-*	*è* ~ *í*
12	Hare	§ 4.5.8	*lèpparu*, *lèppere*, *lèppore/o*, *léppuri*	**lèppar-*/**lèpper-*/ **lèppor-*/**léppur-*	*a* ~ *e* ~ *o* ~ *u*
13	Marten	§ 4.5.9	*assíle*, *kassíβile*, *assaíle*, *asseíle* etc.	**(k-)assíbile*/ **assabíle*/**assebíle*	*í-i* ~ *aí* ~ *eí*
14	Film, layer	§ 7.2.2	*ϑídza*, *ϑòdza*	**ϑíl̩ia*/**ϑòl̩ia*	*í* ~ *ò*

Table 8.3.1: words exhibiting vowel alternations

The bulk of the alternations listed above are found in unstressed syllables. In light of the frequent vowel assimilations and dissimilations that take place in unstressed syllables in Sardinian (cf. Wagner 1941a: 24–36), it is possible that most of these vowel alternations are in fact of recent origin, even if the variants make it difficult to identify the original vocalism. A handful of words show variation in stressed vowels. These are:

- *koròsta* ~ *kurústa* (2) 'bed bug'
- *siśáya* ~ *seiśèi* (3) 'cockroach, darkling beetle'
- *ϑilipírke* ~ *ϑelapòrka* (4) 'grasshopper'
- *lírčis* ~ *lúrtzi* (9) 'viperine snake'
- *bètta* ~ *bítta* (11) 'fawn, female mouflon etc.'
- *ϑídza* ~ *ϑòdza* (14) 'film, layer'

There is just one alternation that occurs more than once, viz. *í* ~ *ò* in *ϑilipírke* ~ *ϑelapòrka* (4) and in *ϑídza* ~ *ϑòdza* (14). However, *ϑelapòrka* 'big grasshopper' may have been influenced by *pórku* 'pig'. The alternation *í* ~ *ú* in *lírčis* ~ *lúrtsi* (9) is also found in Srd. *soloíya* ~ Lat. *solifūga*, *solipūga* (5), but the vocalism of the latter differs with respect to the Latin form rather than in Sardinian internally, making it an imperfect comparandum. In other words, there is no sufficient evidence to attribute any of these alternations to a pre-Roman Sardinian language. One minor observation can nonetheless be made. In several words starting in **ϑili-*, we find variants with **ϑala-*/*ϑela-*: cf. *ϑilikèrta* ~ *ϑalakèrta* (7) 'lizard'; *tsiliɣúɣu*, *sintsilúya* ~ *tsaβaɣúɣu*, *sattsalúya* (8) 'skink, slug'. In Nuoro we also find *ϑilipírke* 'grasshopper' next to *ϑelapòrka* 'big grasshopper'. Perhaps this is evidence for the preexistence of different forms of the same pre-Roman affix (cf. § 9.1.2).

8.4 Observations on pre-Roman phonotaxis

On the basis of the lexical data, we may attempt to make inferences about the phonotactical properties of the pre-Roman language(s) of Sardinia. This concerns aspects including syllable structure, phonological constraints etc. This possibility has been explored by Serra (1960) and Wolf (1998a: 77–81). Serra gives an overview of the vocalism found in Sardinian toponyms of pre-Roman origin. Wolf lists the various syllable types found in the pre-Roman toponyms of the Barbagia. The toponymic evidence he discusses points to a rather limited set of permitted root structures, namely *C*(*r*)*VC*, *CV*(*C*)*C*, *V*(*C*)*C*, in which any cluster of consonants could be either *-rC-*, *-sC-*, or geminate *-CC-*.[263] There are some rare cases of roots containing *-NC-* (e.g. *Pandelai*), but Wolf prefers to explain these as secondary or ultimately inherited from Latin (Wolf 1998a: 30–31). By and large, this analysis on the basis of toponyms is in line with the lexical evidence. However, there is lexical evidence for pre-Roman **-NC-*clusters; cf. *aϑáṇḍa* 'poppy' (§ 3.1.15), *tsorómpis* 'lizard, snake' (§ 4.3.3), *ϑònka* 'scops owl' (§ 4.4.11), suggesting that these clusters might in fact be original in toponyms as well. Additionally, there are some etyma that contain intervocalic clusters of the type *-Cr-*, which do not occur in the pre-Roman toponyms examined by Wolf; cf. *giḍḍòstre* etc. 'heather' (§ 3.3.20), *mufròne*, *muɣròne* 'mouflon' etc. (§ 4.5.10). While *giḍḍòstre* may owe its *-r-* to influence from the inherited suffix *-ástru* (Wagner 1952: 120), *mufròne*, *muɣròne* is attested already in the 5th c. CE as Lat. *mufron* 'mouflon', showing that this cluster must be old.

8.4.1 Word-final consonants

One question regarding syllable structure is whether pre-Roman Sardinian allowed consonants in word-final position. The strongest evidence is constituted by forms that are recorded with a final consonant. Examples are place names attested in early Sardinian sources, such as *Lacon* (Laconi), *Nugor* (Nuoro), *Migil* (Milis) etc. The direct lexical evidence for word-final consonants is discussed in § 8.4.1.1. In words ending in a vowel, unexpected final vowels can point to an original word-final consonant as well (cf § 8.1.5). If we find word-final vowel alternation, this could be the result of the independent addition of final vowels to resolve final consonants, as a nativization strategy. Additionally, the presence

[263] To an extent, there may have been more complexity that has been obscured by Sardinian sound changes. For example, all instances of *-r-* could hypothetically also go back to etymological **-l-* (Wolf 1998a: 32). Likewise, geminate *-tt-* could in theory be the outcome of **-tt-*, **-kt-*, **-pt-*, **-dd-*, **-gd-* and perhaps **-bd-* (Wagner 1941a: 192–193; Wolf 1998a: 35).

of paragogic vowels may also be indicative of an original final consonant (cf. Wagner 1941a: 57–62). These two lines of evidence are discussed in § 8.4.1.2. The following is an overview of all etyma showing one or more of the lines of evidence described above.

	Meaning	Section	Forms	Reconstr.	Final alt.	Possible paragoge
1	Bristly ox-tongue	§ 3.1.4	*ɟiòkkoro, sóčči̯ri* etc.	**ɟiokkor-/ *ɟókkir-*	*-i, -o*	✓
2	(Fool's) water-cress	§ 3.1.24	*ǧúɣuru, ǧúru* etc.	**i̯úguru*		✓
3	Galingale	§ 3.2.3	*sèssene, sèssini*	**séssene*		✓
4	Knotgrass, horsetail	§ 3.2.4	*síntsiri, sintsurru*	**ɟínɟi/ur-*	*-i, -u*	✓
5	Rosemary	§ 3.3.14	*ɟíppiri, tsíppiri*	**ɟíbbir*		✓
6	Tree heather	§ 3.3.20	*giḍḍòstre, iḍḍòstro, giḍḍóstru, ʔiḍḍòstra*	**gillostr-*	*-a, -e, -o, -u*	✓
7	Resin	§ 3.4.5	*tsònkine*	**ɟònkene*		✓
8	Viperine snake	§ 4.3.5	*lírčis, lúrtsi*	**lírčis, *lúrɟi*	*-is, -i*	
9	Bird of prey	§ 4.4.2	*ɟáiri*	**ɟá(C)iri*		✓
10	Robin, yellow wagtail; hare	§ 4.4.9	*príncˇiri, bríncˇis(i), bríncˇi, brinkíḍḍi*	**brink/či-*	*-is, -i*	✓
11	Robin, wren	§ 4.4.10	*gríši, kiríši, kiríu* etc.	**grísi/*kirísi/ *kiri-*	*-i, -íu, -u*	✓
12	Bat	§ 4.5.1	*čížini*	**kikin-*		✓
13	Fawn	§ 4.5.5	*bíta, bítti, bètte, bétu*	**bett-/*bitt-*	*-a, -e, -i, -u*	✓
14	Hare	§ 4.5.8	*lèpparu, lèppore, lèporo, léppuri*	**leppVr-*	*-e, -i, -o, -u*	✓
15	Hillock	§ 5.1.6	*mòɣora, mòɣoru*	**mokor-*	-a, -o	✓
16	Rocky plain	§ 5.1.23	*Theppar* (OSrd.), *tsèppara*	**ɟèppar*	*-Ø, -a*	✓
17	Border	§ 5.2.1	*lákkana*	**lákkana*		✓
18	Nuraghe	§ 5.2.2	*NURAC* (Lat.) *nuráke*	**nurak-*	*-Ø, -e*	
19	Tilled land	§ 6.1.8	*tèle, tuvèle*	**t(ev)èle*		✓

	Meaning	Section	Forms	Reconstr.	Final alt.	Possible paragoge
20	Clay used to make acorn bread	§ 6.2.2	*tòrko*, *tròkko*, *trók-ku*	**tork-*	*-o*, *-u*	✓
21	Cheek	§ 7.1.3	*tutúrru*	**tuturr-*		✓

Table 8.4.1: Words with potential evidence for a word-final consonant

8.4.1.1 Direct evidence for word-final consonants

A handful of Sardinian words occur with variants with a final consonant. In the modern dialects these are *lírčis* (8) and *bríncis* (10). In Old Sardinian we find the toponym *Theppar* (16), which probably belongs to *tsèppara* 'rocky plain'. To these direct attestations we may add two likely punicisms, whose Punic source ends in a consonant: *ϑíppiri* (5) << Pun. *zibbir* and *síntsiri*, *sintsurru* (4) << Pun. *zenzur*. Finally, there is the epigraphic attestation of *NURAC*, on the architrave of nuraghe Aidu Entos near Bortigali (Gasperini 1992: 303–306), which corresponds to modern Srd. *nuráke*, *nuráɣe* etc. 'nuraghe' (18). However, while this almost certainly native form suggests the occurrence of word-final *-c* [k], it cannot fully be excluded that the absence of a final vowel is the result of Roman epigraphic abbreviation practices.[264]

The directly attested evidence suggests that not only Punic allowed word-final consonants (cf. Friedrich 1951: 34–36), but also the pre-Roman language(s) native to Sardinia. Loanwords from both Punic and other pre-Roman language(s) would have undergone the same phonological treatment when they were borrowed into Sardinian Latin. The modern Sardinian forms provide direct evidence only for word-final *-s*. This is unsurprising, since inherited plural *-s* (unlike other consonants) is in some Sardinian dialects still realized without paragogic vowel (Wagner 1984: xxiv–xxv).

8.4.1.2 Indirect evidence for word-final consonants

For the other words listed above in Table 8.4.1, it is possible but not certain that they once ended in a consonant. Both word-final vowel alternations and possible paragoge can point to an original word-final consonant, but these arguments are not conclusive by themselves. Moreover, examples like *Laconi* < *Lacon*, *ǧáyaru* 'hound' < *gjagar*, Camp. *krášì* 'tomorrow' < Lat. *cras* show that word-final consonants could also be resolved by a paragogic vowel that did not copy the previous

[264] Cf. Gasperini's (1992: 305) interpretation, which proposes abbreviations for most forms in the inscription.

one, even when copying of the vowel is the norm (Wagner 1941a: 57–62). It is therefore possible that there are other words, not listed in Table 8.4.1, that originally also ended in a consonant.

The following words exhibit both final vowel alternation and identical vowel quality in the last two syllables:

- *ʔiòkkoro*, *sóčči̱ri* 'bristly ox-tongue'
- *giḍḍòstre*, *iḍḍòstro*, *giḍḍóstru*, *ʔiḍḍòstra* 'tree heather'
- *gríši*, *gríšu*; *kilíši*, *kilíšu* 'robin, wren'
- *bètta*, *bètte*, *bítta*, *bítti* 'fawn, roe deer, lamb'
- *lèpparu*, *lèppore*, *lèpporo*, *léppuri* 'hare'
- *tòrko*, *tròkko*, *trókku* 'clay used to make acorn bread'

All of these words show variation of the final vowel, including one variant that could be due to paragoge. Especially significant are the words whose variation in the vowel of the penultimate syllable is matched by the last syllable; cf. *ʔiòkkoro* ~ *sóčči̱ri*; *bètte* ~ *bítti*; *lèppere* ~ *lèpporo*. Rather than assuming vowel alternations in two syllables, we might propose reconstructions like **ʔiókkor*/**ʔiókkir*, **bétt*/**bítt*, **lépper*/**léppor*.

8.4.1.3 Evidence for word-final consonant clusters

The various forms for 'tree heather' contain all possible final vowels except *-i*. Among these, the *-o* of *iḍḍòstro* is the most difficult to explain in a regular way and is likely either original or paragogic. This word has been proposed to contain the same suffix *-ost*(*r*)- (§ 9.2.9) as found in *golóstri*, *golóstu̱* etc. 'holly' (§ 3.3.7), which is predominantly attested with final *-i*. If we want to explain the observed alternations in *ʔiḍḍòstra*, *giḍḍòstre*, *iḍḍòstro*, *giḍḍóstru* 'tree heather' as the result of different strategies of resolving a word-final consonant, we would have to posit a reconstruction **gillostr*, which does by no means conform to Latin or Sardinian phonotaxis (cf. § 8.1.5.2.1). The same is true for *bètte*, *bítti* etc. 'fawn' < **bétt*/**bítt* and for *tòrko*, *tórku* 'clay' < **tórk*. However, these three words do not constitute sufficient evidence to posit word-final consonant clusters for pre-Roman Sardinian. No evidence for word-final clusters is found by Wolf (1998a) either.

8.4.2 Evidence for pre-Roman vowel harmony

Serra (1960: 408–409) notes that there is a sizable number of pre-Roman Sardinian toponyms that are vocalically "monotonous", i.e. all syllables of one variant contain the same vowel. Examples are *Árzana*, *Semèstene*, *Ísili*, *Sòrgono* etc. (Ser-

ra 1960: 409–418). Serra (1960: 408–409) attributes this to some kind of vowel harmony in pre-Roman Sardinian.[265] In cases where the vocalism is not identical throughout the word, Serra (1960: 408–409) argues that the unaccented syllables nevertheless often contain the same vowel; e.g. *Gúspini*, *Tortolì*, *Tèltoro* etc. (Serra 1960: 409–418). This latter tendency is difficult to attribute to vowel harmony in a pre-Roman language. The forms that would fall under this description are too diverse in terms of vocalism and stress placement to be reliably taken as a single phenomenon. For instance, the toponyms *Gúspini*, *Tortolì*, *Tèltoro* each exhibit pairs of identical vowels, but the vowels are different and they occur in different orders.

Serra compares this putative phenomenon to North-African toponyms to support his hypothesis of a "Libyan" substrate in Sardinia. Whereas such toponymic comparisons are outside the scope of this study (cf. § 2.1.2.2.1), it is worthwhile to investigate whether forms containing three or more identical consecutive vowels are in fact found in lexical items of pre-Roman origin. The following is a list of forms of plausible pre-Roman origin discussed in § 3 – 7 that exhibit this phenomenon. Excluded are punicisms, and forms where the apparent vowel harmony involves a productive inherited suffix (e.g. *kaðúmbulu* with inherited **-ul-*). Beside the forms exhibiting potential vowel harmony and their reconstruction, cognate forms without potential vowel harmony are listed.

	Meaning	**Sect.**	**Forms with harmony**	**Reconst.**	**Vowel pattern**	**Other forms**
1	Arum	§ 3.1.2	*ϑoϑoróyu*	**ϑoϑorói̯*	*o-o-ó*	*tattaróyu*
2	Bristly ox-tongue	§ 3.1.4	*ϑiòkkoro*	**ϑi̯òkkoro*	*ó-o-o*	*sóčči̯ri*, *čóčči̯ri*
3	Hemlock	§ 3.1.11	*úḍḍuru*	**vúlluru*	*ú-u-u*	*biḍḍúri*, *búḍḍaru*, *meḍḍure*
5	Poppy	§ 3.1.15	*tzàndhara*	**ϑándara*	*á-a-a*	*ϑánda*, *ánna*
6	Watercress	§ 3.1.24	*ğúɣuru* etc.	**i̯úguru*	*ú-u-u*	
7	Galingale	§ 3.2.3	*sèssene*, *sèssini*	**sèssene*	*é-e-e*	
8	Strawberry tree fruit	§3.3.18	*liṡiriòne*	**lisiriòne*	*i-i-i-ò*	

[265] Serra (1960: 408–409) does note the difference with other languages exhibiting vowel harmony, where the harmonizing vowels match each other's timbre rather than taking over each other's quality completely.

	Meaning	Sect.	Forms with harmony	Reconst.	Vowel pattern	Other forms
9	Traveller's joy	§ 3.3.19	*attsára*	**aϑára*	*a-á-a*	*alússara*, *auttsára*
10	Mushroom	§ 3.4.3	*tuntúnnu* etc.	**tuntúnnu*	*u-ú-u*	*antúnna*, *tunníu* etc.
11	Clothes moth	§ 4.1.5	*lanásta*	**lanásta*	*a-á-a*	*lána*
12			*nástala*	**nástala*	*á-a-a*	
13	Grasshopper	§ 4.1.9	*tsimpilíɣe*	**ϑimpilíke*	*i-i-í-e*	*pilíkke*
14	Grasshopper	§ 4.1.11	*piβittsíri*, *piβiϑϑía* etc.	**pipiϑí-*	*i-i-í*	
15	Grasshopper	§ 4.1.12	*ϑilipírke* etc.	**ϑilipírke*	*i-i-í-e*	*ϑelapòrka*
16	Tick	§ 4.1.16	*kaðánka* etc.	**katánka*	*a-á-a*	
17	Kestrel	§ 4.4.4	*tserpeḍḍéri*	**ϑerpelléri*	*e-e-é-i*	*tsarpeḍḍéri*, *spraḍḍéri*
18	Kestrel	§ 4.4.5	*tilibríu*, *tilibríkku*	**ϑiliprí(kk)u*	*i-i-í-u*	
19	Robin, yellow wagtail	§ 4.4.9	*príɲčiri*	**brínciri*	*í-i-i*	*prinčóttu*
20	Robin, wren	§ 4.4.10	*kiríši, kilíši*	**kilísi*/**kirísi*	*i-í-i*	*gríšu*, *kilíšu*, *kiríu* etc.
21	Bat	§ 4.5.1	*čížini*	**kíkini*/ **kíkine*	*í-i-i*	*čižinéḍḍu* etc.
22	Hare	§ 4.5.8	*lèppere*	**lèppere*	*è-e-e*	*lèpparu*, *lèppore*, *léppuri* etc.
23	Hillock	§ 5.1.6	*mòɣoro*	**mokoro*	*ò-o-o*	*mòɣora*
24a	Rock crevice	§ 5.1.19	*kalánka*	**kalánka*	*a-á-a*	
24b			*kolonkòne*	**kalankòne*	*o-o-ò-e*	
25	Border, limit	§ 5.2.1	*lákkana*	**lákkana*	*á-a-a*	
26	Sheep that has lost its young	§ 6.1.2	*argáša*	**argása*	*a-á-a*	
27	Tilled land	§ 6.1.8	*tevèle*	**tevèle*	*e-è-e*	*tèle*
28	Cheek	§ 7.1.3	*tuttúrru*	**tuttúrru*	*u-ú-u*	

Table 8.4.2: Pre-Roman words containing three identical consecutive vowels

Although there are numerous lexical forms that exhibit vocalic "monotony" as described by Serra (1960: 408–409), many words of well-established pre-Roman origin do not (e.g. *kóstike*, *giḍḍòstre*, *òrga*, *nuráɣe* etc.). Moreover, most words

listed in Table 8.4.2 are also attested with slightly different vocalisms that contradict this phenomenon. It is likely that some of the instances of vocalic monotony are due to recent vowel assimilations, which are common in inherited forms too (cf. Wagner 1941a: 24–35). As a related observation, in various words, not all of which fit the criteria to be listed in Table 8.4.2, we find vocalic alternation by which two identical consecutive vowels change in unison.

	Meaning	Section	Form 1	Form 2
1	Arum	§ 3.1.2	*tattaróyu*	*ʔoʔʔoróyu*
2	Bristly ox-tongue	§ 3.1.4	*ʔiòkkoro*	*sóčči̯ri*
3	Bed bug	§ 4.1.3	*kurústa*	*koròsta*
4	Earthworm	§ 4.1.10	*ʔilingròne*	*ʔulungròne*
5	Lizard	§ 4.3.2	*ʔilikèrta*	*ʔalaʔèrta*
6	Skink	§ 4.3.4	*ʔilikúkku*	*ʔalakúkku*
7	Skink	§ 4.3.4	*sintsilúya*	*sattsalúya*
8	Bird species	§ 4.4.3	*korrokkáu*	*kurrukkáu*
9	Bat	§ 4.5.2	*ʔiʔʔirri̯ólu*	*ʔuʔuréḍḍu*
10	Fawn	§ 4.5.5	*bítti*	*bètte*
11	Hare	§ 4.5.8	*lèppere*	*lèporo*

Table 8.4.3: Words with two consecutive vowels alternating in unison

With the exception of the words for 'bed bug' (3) and 'fawn' (10), all of these alternations are found in unstressed syllables. This raises the suspicion that in many of these forms, one of the two unstressed alternating vowels owes its quality to assimilation (cf. Wagner 1941a: 24–35). In cases like *ʔiòkkoro* next to *sóčči̯ri* (2) 'bristly ox-tongue', the quality of the final vowel may be due to paragoge (cf. § 8.1.5).

In conclusion, it is clear that only a subsection of the potential pre-Roman lexicon exhibits a vocalism that could point to vowel harmony. Since vowel assimilation is frequent across the Sardinian dialects, the lexical evidence does not enable us to corroborate the kind of pre-Roman vowel harmony that Serra (1960) posits on the basis of vocalically "monotonous" toponyms like *Árzana*, *Semèstene*, *Ísili*, *Sòrgono*.

8.5 Conclusion

The evidence for the retrieval of phonological characteristics of the pre-Roman language(s) is discussed in this chapter. We have observed that it is rather scarce and of indirect nature. It is therefore impossible to offer a complete reconstruction of the phonological system of this language, if it was one language. However,

we can make a number of preliminary inferences on the phonological tendencies we have observed, and check whether or not they are homogenous across Sardinia. If so, they could hold clues about the linguistic diversity on the island prior to Roman conquest.

In § 8.2 and § 8.3, it has been shown that there are hardly any observable patterns in the irregular alternations that we encounter, neither in the consonantism, nor in the vocalism of the discussed material. This could be an indication that the pre-Roman language(s) of Sardinia were not strikingly different from Latin in terms of their consonant and vowel inventories. The low number of alternations in consonant quantity challenge the longstanding idea that voicing alternation in stops is a characteristic of the "Mediterranean substrate" (e.g. Bertoldi 1931: 134–137). On the other hand, there is evidence that the pre-Roman Sardinian contained a phoneme, non-native to Latin, that merged with the outcomes of Latin **kʲ* and **tʲ* in Sardinian, i.e. **ϑ* > *ϑ*, *t*, *ts* (cf. § 8.1.3). The reflexes of Sardinian **ϑ* are found in Punic loans, where it represents Pun. *z* (§ 8.1.2). This is evidence that the variety of Punic spoken in Sardinia may have preserved the affricate realization [dz] of this phoneme. In the northern part of Sardinia, some forms contain *ts* or *č*, which do not correspond to the regular outcomes of **ϑ* (§ 8.1.4). The geographical distribution of this phenomenon is rather inconsistent however, and it cannot be ruled out that most instances of it are due to sound-symbolic alterations of **ϑ* or other inherited consonants. As for the vowel inventory of the pre-Roman Sardinian, there is not much evidence that it was significantly different from that of Latin. All five inherited vowels (*a*, *e*, *i*, *o*, *u*) are attested in non-inherited words as well as in toponyms. There is no evidence for length oppositions, as Sardinian merged Latin long and short vowels. In contrast to Latin, there is no evidence for any vowel weakening process, as all five vowels appear to occur both in stressed and unstressed syllables. Word-finally, all five vowels are attested as well.

Phonotactically, a more limited set of root structures is attested in forms of likely pre-Roman origin than in inherited forms (Wolf 1998a: 28–36)(§ 8.4). There is evidence for pre-Roman language(s), including Punic, allowing word-final consonants (§ 8.4.1). Beside final *-s*, which is preserved both in inherited (e.g. plural *-s*) and in non-inherited words (e.g. *lírčis* 'viperine snake'), there is direct toponymic evidence for word-final *-l*, *-n* and *-r* (e.g. *Migil*, *Lacon*, *Mandar*). The Roman inscription containing the form *NURAC* presents potential evidence that **-k* too could be in word-final position (but cf. § 8.4.1.1). In modern Sardinian these have been resolved largely with paragogic vowels. It is possible that some non-

inherited words exhibiting word-final vowel variation originally ended in a final consonant too (cf. § 8.4.1). The strongest cases among these still only point to word-final *-*r* and *-*s* however. The evidence for word-final consonant clusters is too limited to draw any firm conclusions (§ 8.4.1.3). Finally, while there are some forms that adhere to the tendency of "vocalic monotony" described by Serra (1960), there are too many counter-examples to reliably posit vowel harmony for any pre-Roman language (§ 8.4.2). Additionally, some of the attested vocalic "monotony" can be explained by means of assimilatory processes within Sardinian (Wagner 1941a: 24–35).

In conclusion, there is no strong evidence that the pre-Roman phonological system was radically different from that of Latin. There is little to no evidence for significant differences in phonological systems between various parts of Sardinia. Non-inherited *ə is attested across the island, and Wolf's (1998a: 36) restrictions on syllable structure identified for toponyms in the Barbagia, seem to apply to lexical items attested in other regions too. This homogeneity is consistent with the preexistence of a single language or closely related language varieties across the island.

9 Morphology

Various Sardinian words of possible pre-Roman origin have been analyzed as containing pre-Roman morphemes (e.g. Guarnerio 1904a; Guarnerio 1904b; Terracini 1927; Wagner 1997: 254–280; Blasco Ferrer 1988: 173–178 etc.). In what follows, the evidence for the existence of pre-Roman affixes will be evaluated on the basis of the relevant lexical material. A pre-Roman morpheme can most confidently be posited on the basis of doublets consisting of variants with and without the affix in question (§ 2.2.2). Once the "morpheme-hood" of a certain sequence is established, it is possible to investigate its original function.

9.1 Prefixes

In the existing literature on the pre-Roman linguistic elements of Sardinia, the identification of potential pre-Roman suffixes has received significantly more attention than that of potential prefixes. This is possibly due to an Indo-European or Romance typological bias, since in these language families derivational morphology is predominantly suffix-based. For a non-attested language that was in all likelihood unrelated to Indo-European however, there is no reason to suspect *a priori* that it would have aligned typologically with Indo-European in general and with Romance in particular. In the pre-Roman Sardinian lexicon, there are in fact a number of recurring word-initial sequences that appear to exhibit the alternations indicative of substrate morphemes.

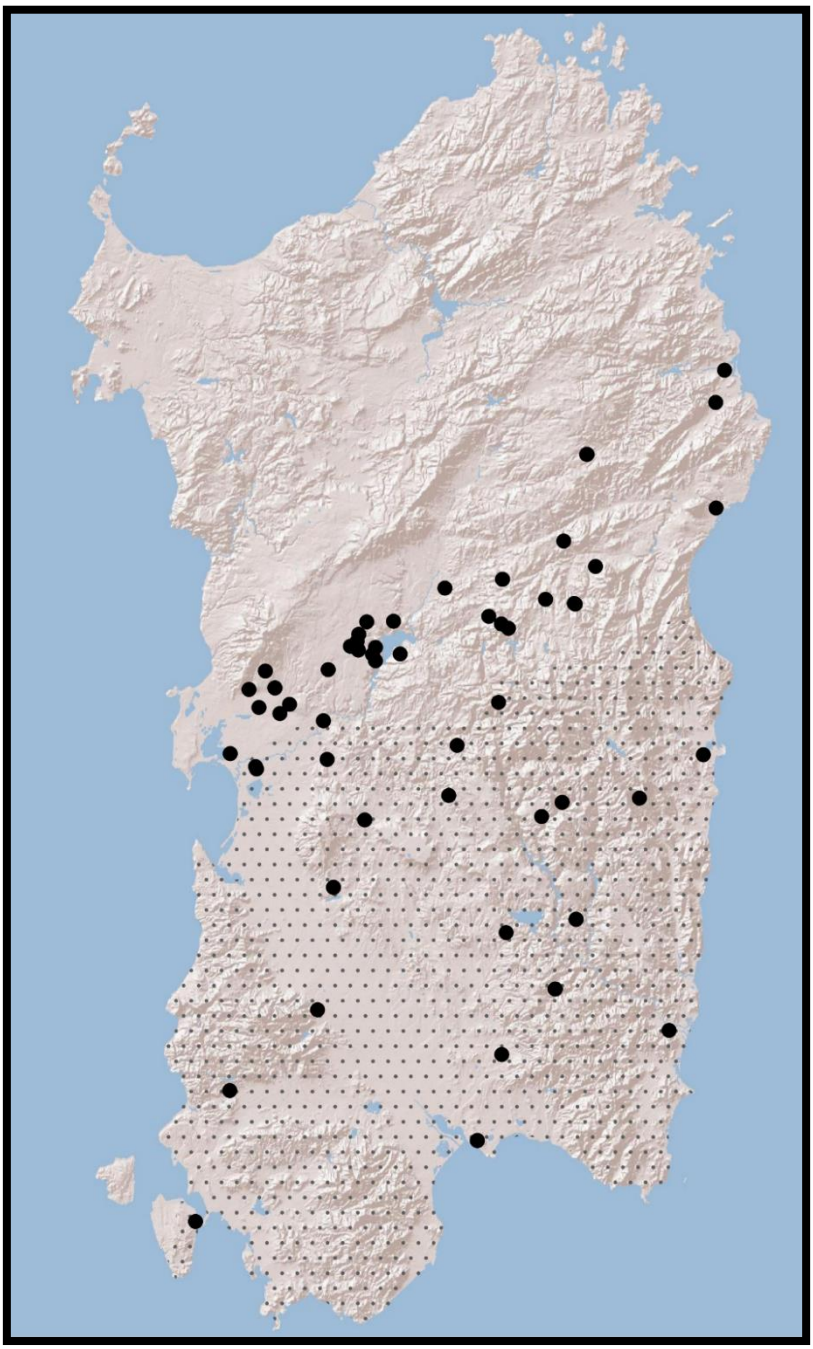

Figure 9.1.1: Distribution of forms containing **k(V)-*

9.1.1 **kV-*

First noted by Guarnerio (1904a: 58 fn. 2), there are words in the Sardinian lexicon that occur both with and without initial **k(V)-*. Some of the corresponding forms without **k(V)-* have other initial sequences, while others do not. There are six examples, all of which show an alternation.

	Meaning	Section	Form with **k(V)-*	Form without **k(V)-*	Reconstr.	Alt.
1	Celery, fool's watercress	§ 3.1.6	*kugúsa*	*ϑurgúsa, tseliɣúsa* etc.	**ku-gúsa*	✓
2	Mullein	§ 3.1.12	*kaðúmbulu* etc.	*túmbara*	**ka-túmb-*	✓
3	Pine tree	§ 3.3.12	*kampínğu* etc.	<< Lat. *pīneus*	**ka(m)-pinu̯*	?
4	Bed bug	§ 4.1.3	*kurústa*	*rústa*	**ku-rústa*	✓
5	Lizard	§ 4.3.2	*kalužèrta* etc.	*ϑilikèrta, aliɣèrta* << Lat. *lacerta*	**ka-la/ukerta*	✓
6	Marten	§ 4.5.9	*kassíβile* etc.	*assíle, síβile*	**k-assíbile*	✓

Table 9.1.1: Forms containing **kV-*

Especially notable is that two of the words that occur both with and without **k(V)-* are formed from inherited, Latin bases, i.e. Srd. *kampínğu* 'pine tree' < **kam-pínu̯* << Lat. *pīneus* 'of the pine' and Srd. *kalužèrta* < **ka-lukérta* << Lat. *lacerta* 'lizard'. Additionally, the sequence **k(V)-* alternates with **ϑ(i)-* in some forms (e.g. 1: *kugúsa* ~ *tseliɣúsa*; 5: *kalužèrta* ~ *ϑilikèrta*), which cannot be of inherited origin (cf. § 8.1.1; § 9.1.2). This is evidence that **k(V)-* was a pre-Roman prefix that could be added to Latin words (cf. Swanenvleugel 2024: 242–245). Figure 9.1.1 is an overview of the distribution of forms containing **k(V)-*. It has a clearly southern distribution, being absent from the northern portion of Sardinia (cf. also Swanenvleugel 2024: 244). This contrasts with the distribution of the pre-Roman prefix **ϑ(i)-* (§ 9.1.2).

9.1.2 **ϑ(i)-*

With regard to pre-Roman Sardinian, words starting in **ϑ-* are much discussed. This phone cannot be regularly inherited from Latin in this position (§ 8.1.1). From a semantic perspective, it has been noted that there are a substantial amount of words referring to small animals starting in **ϑi-* (Forsyth Major 1893: 154; Guarnerio 1904a: 57–58; Wagner 1932: 223–224; Wagner 1997: 263). The element has been compared to various phenomena in other languages, including the Berber feminine prefix **t-* (Wagner 1932: 223–224; 1997: 263). This is problematic because of a formal mismatch (cf. Blasco Ferrer 2011: 73). Pittau (1995:

197) compares **ϑi-* to the Etruscan demonstrative pronoun *ta/eta/ita*, which runs into similar formal problems, since Srd. **ϑi-* cannot directly go back to **t-*. Alinei (1984: 29) hypothesizes that all cases of **ϑi-* are originally from Srd. *ϑíu/tíu/tsíu* 'uncle', *ϑía/tía/tsía* 'aunt'; i.e. *ϑilikèrta* < **ϑía lakèrta* 'aunt lizard' (also Blasco Ferrer 2011: 73). This is semantically unconvincing, at the very least for the plant names like 'poppy' and 'celery, fool's watercress' that show evidence of **ϑi-* (see Table 9.1.2)(cf. Wolf 2011: 613 fn. 107). Like in the case of **kV-* (§ 9.1.1), there are some words that show an alternation between forms with and without **ϑ(i)-*.

	Meaning	Section	Forms with **ϑ(i)-*	Forms without **ϑ(i)-*	Reconst.	Alt.
1	Arum	§ 3.1.2	*ϑoϑoróyu*, *tattaróyu* etc.	Sic. *azzaru* etc.	**ϑ(-)aϑarói̯-*	?
2	Celery, fool's watercress	§ 3.1.6	*tseliɣúśa*	*ϑurgúśa*, *kugúśa* etc.	**ϑiligúsa*	✓
3	Poppy	§ 3.1.15	*ϑáṇḍa* etc.	*ánna*	**ϑ-ánda*	✓
4	Backswimmer	§ 4.1.2	*tsirifrúči*		**ϑilifurke*	×
5	Dragonfly, grasshopper	§ 4.1.9	*tsimpilíye*	[*kaḍḍ' e*] *pilíkke*	**ϑi(m)-pilíke*	✓
6	Earthworm	§ 4.1.10	*ϑilingròne* etc.		**ϑiling(u)lòne*	×
7	Grasshopper	§ 4.1.12	*ϑilipírke* etc.		**ϑilipírke*	×
8	Gecko	§ 4.3.1	*tilíbbu*		**ϑilikwu*	×
9	Lizard	§ 4.3.2	*ϑilikèrta* etc.	*kalužèrta*, *aliɣèrta* << Lat. *lacerta*	**ϑi-likerta*	✓
10	Skink, slug	§ 4.3.4	*ϑilikúkku* etc.	*alicúcu*	**ϑi-likúku*	✓
11	Kestrel	§ 4.4.5	*tilibríkku*, *at-tiliβríu* etc.		**ϑiliprí(kk)u*	×
12	Kite	§ 4.4.6	*ϑurulía*, *tirolía* etc.		**ϑurolía*	×
13	Bat	§ 4.5.2	*ϑiϑirri̯ólu*, *ϑuϑuréḍḍu* etc.	*tsirri̯ólu* etc.	**ϑi-ϑur(r)-*	✓
14	Neck	§ 7.1.4	*ϑúkru*, *ϑúɣulu* etc.		**ϑúkulu*	×

Table 9.1.2: Forms containing **ϑ(i)-*

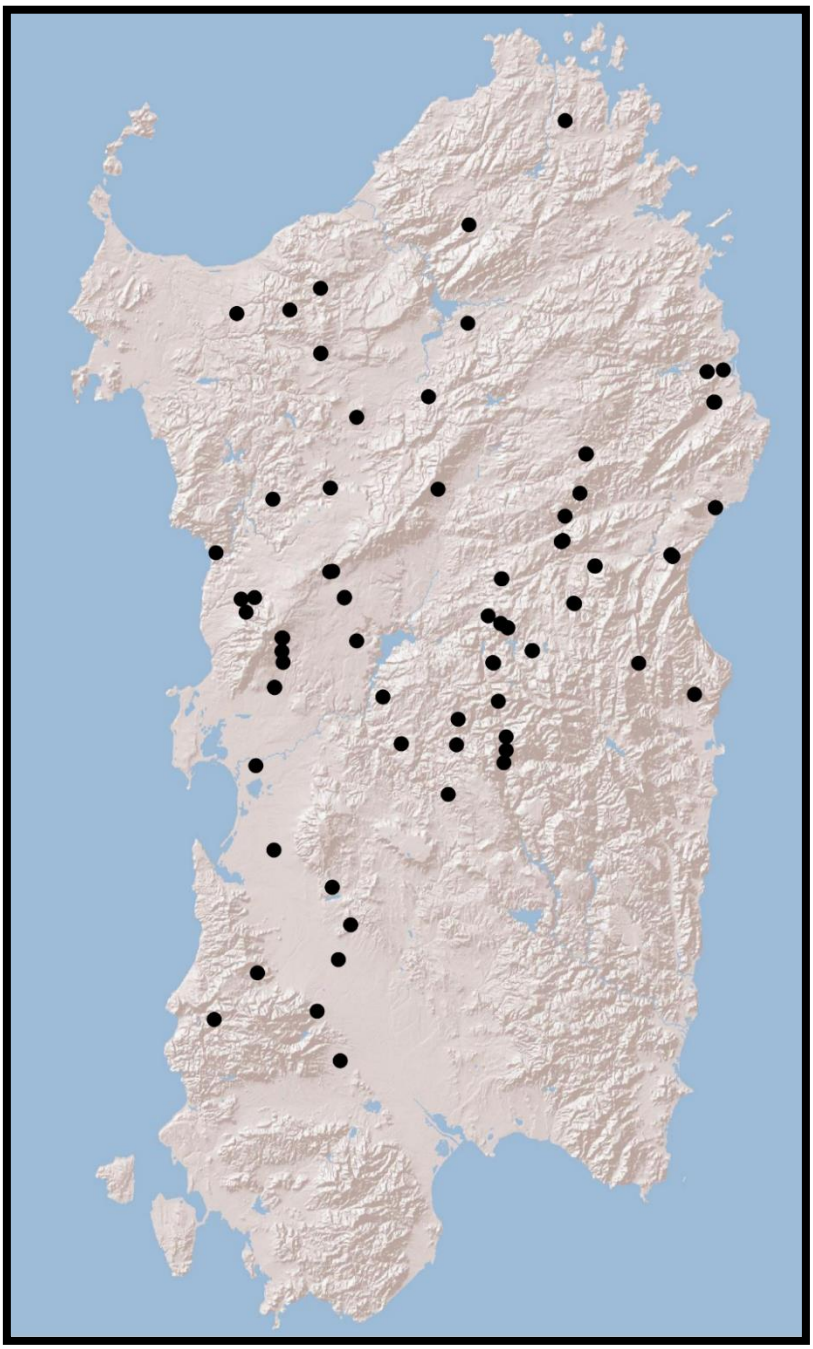

Figure 9.1.2: Distribution of forms containing **ϑi-* showing an alternation

Like **kV-* (§ 9.1.1), **ϑi-* appears to have been added at least once to a Latin stem: *lacerta* 'lizard' >> **ϑi-likérta* > Srd. *ϑilikèrta*. This, and its alternations in other words, e.g. *ϑáṇḍa* ~ *ánna* 'poppy', *ϑilikúkku* ~ *alicúcu* 'slug' etc., confirm that this was indeed a morpheme in a pre-Roman Sardinian language (cf. Swanenvleugel 2024: 256–257). Figure 9.1.2 shows the distribution of forms in which the morphemic status of **ϑ(i)-* is demonstrated by the existence of forms without **ϑ(i)-*. In contrast to the distribution of **k(V)-* (Figure 9.1.1), **ϑ(i)-* has a more northern distribution, being absent from the south-eastern portion of Sardinia (cf. also Swanenvleugel 2024: 258). The relation between **k(V)-* and **ϑ(i)-* is discussed in § 9.1.3.1.

9.1.2.1 **ϑVϑV-/*tsVtsV-*

A peculiar subset of words starting in **ϑ(i)-* show an initial sequence in **ϑVϑV-*, as already noted by Forsyth Major (1893: 154) and Guarnerio (1904a: 60). These forms all cooccur with forms that do not contain **ϑVϑV-*.

	Meaning	Section	Forms with **ϑVϑV-*	Forms without **ϑVϑV-*	Reconst.	Alt.
1	Snail, slug, beetle, cicada	§ 4.1.15	*sittsiɣórru*, *tiϑiɣórru*, *tsontsòrra* etc.	*kokòrra*, *piϑϑiɣórru* etc.	**ϑVϑV-(k)órnu*	✓
2	Skink, slug	§ 4.3.4	*sintsilúɣa*, *ϑaϑϑalúkka* etc.	*ϑilikúkku*, *ϑalakúkka*, *babbalúɣa*	**ϑVϑV-lúkka*	✓
3	Kite	§ 4.4.6	*čiččulía*, *sittsulía*	*tirolía*, *ϑurulía* etc.	**ϑiϑulía*	✓
4	Bat	§ 4.5.2	*tsintsimúrru* etc.	*tirṛiólu*, *ϑurrundéḍḍu* etc.	**ϑiϑiṛiólu*, **ϑiϑimúrru*	✓

Table 9.1.3: Forms containing **ϑVϑV-*

The examples are not uniform in nature. For Camp. *sittsiɣórru* 'slug, snail', DES (II: 422) proposes a compound of *sèttsiri* 'to put' and *kórru* 'horn'. This does not explain forms like *tiϑiɣórru* and *tsontsòrra*. It is in any case clear, however, that these words for 'snail, slug etc.' all contain the base *kórru* 'horn' < Lat. *cornu*, reshaped due to various sound-symbolic and folk-etymological processes (§ 4.1.15). Additional evidence for a more recent sound-symbolic origin for some of these forms is the fact that *-ts-* in Log. *tsintsòrra* etc. does not have a regular source (cf. § 8.1.4).

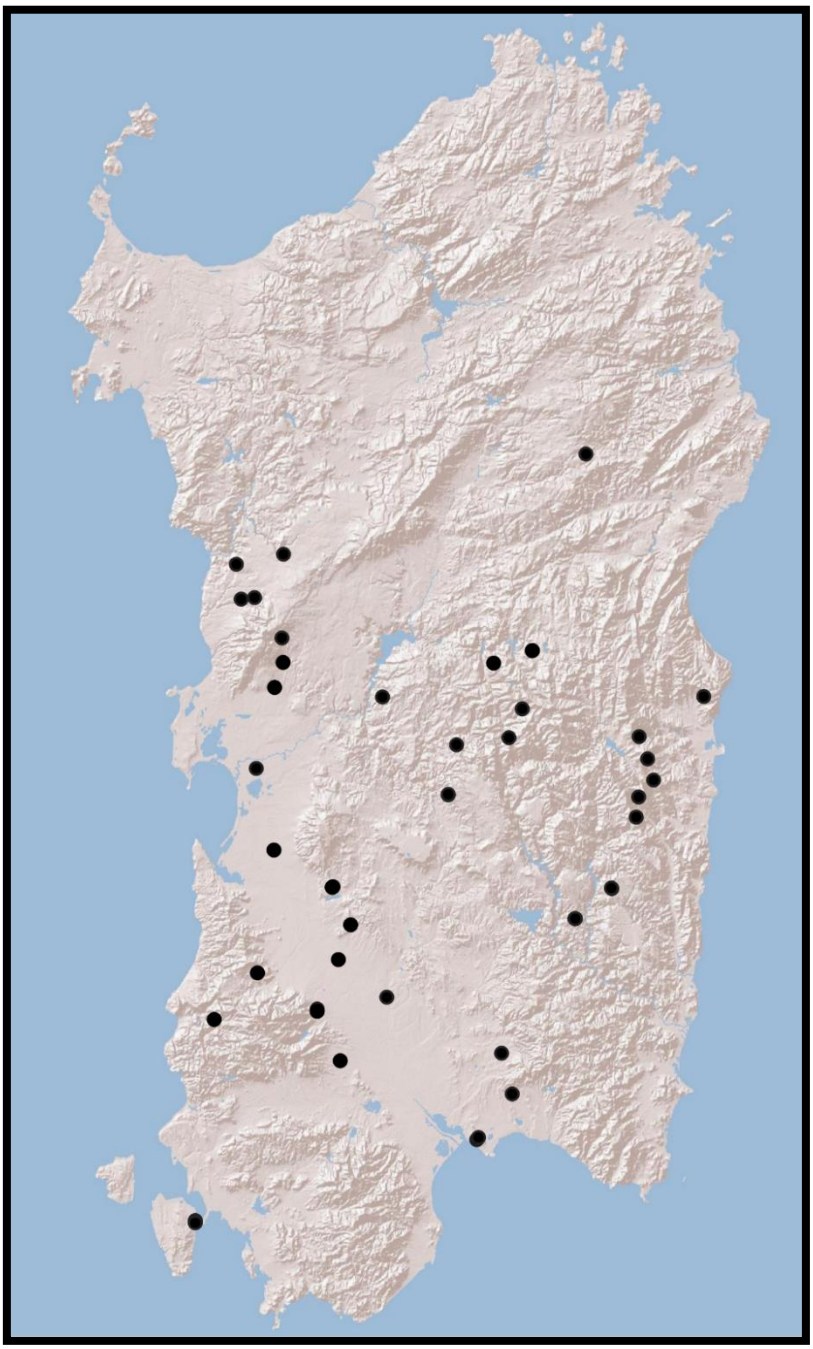

Figure 9.1.3: Distribution of forms containing **ə́Və́V-*

Moreover, we find a similar type of formation in a clearly onomatopoeic word like *ϑiϑilíkkos* 'tickles', next to *ϑakkulíttos*, *ϑeulíttos*, *tsangulíttas* etc. (DES II: 549). Similar explanations are attractive for the other listed words as well. Secondary insertion of **ϑVϑV-* could thus perhaps explain *ϑaϑϑalúkka* 'skink, slug', which corresponds irregularly to *ϑilikúkku*, *ϑalakúkka* (§ 4.3.4). Likewise, *sittsulía* 'kite' is nearly identical to more common *tirolía*, *ϑurulía* etc. < **ϑi/urolía* but with *-tts-* < **-ϑ-* instead of *-r-* (§ 4.4.6). Finally, *tsintsimúrru* 'bat' appears to be from the base **ϑi-ϑi/ur-* found in *ϑiϑiriɲólu*, *ϑuϑuréḍḍu* etc. 'bat', but with influence of *múrru* 'grey' or of homophonous *múrru* 'muzzle' (DES II: 556; cf. § 4.5.2).

Whatever the exact etymological details of these forms containing **ϑVϑV-*, they seem largely of secondary nature. As a consequence, there is no strong evidence that **ϑVϑV-* was ever a morpheme in any pre-Roman language. This is in line with the "diminutive" interpretation of this sequence proposed already by Forsyth Major (1893: 154). The distribution of forms containing **ϑVϑV-*, shown in Figure 9.1.3 suggests that the tendency of introducing **ϑVϑV-* is found predominantly in the southern half of Sardinia.

9.1.2.2 **ϑur-*

There is a single word that shows alternation between **ϑur-*, **kV-* and **ϑili-*:

	Meaning	Section	Forms with **ϑur-*	Forms without **ϑur-*	Reconst.	Alt.
1	Celery, fool's watercress	§ 3.1.6	*ϑurgúsa*, *ϑrugúsa*	*kugúsa*, *tseliɣúsa* etc.	**ϑur-gúsa*	✓

Table 9.1.4: Forms containing **ϑur-*

At face value, the alternations suggest that **ϑur-* was a distinct morphological element in the pre-Roman source language of this word. However, according to the criteria outlined in § 2.2.2, a single example does not constitute sufficient evidence to reliably identify a pre-Roman morpheme. Paulis (1992: 141) proposes a link between *ϑurgúsa* 'fool's water cress, wild celery' and *orgòsa*, *urgúsa* 'humid, marshy place', but this implies a separation of the word from all other forms denoting the same plants that also contain **-gúsa*, which is rather unattractive. Whether *ϑur-* in *ϑurgúsa* is a variant of **ϑi-*, or whether *ϑurgúsa* rose through association with *orgòsa*, *urgúsa* 'humid, shallow place', is difficult to establish on the basis of this example alone. Perhaps it is to be analyzed as **ϑ-urgúsa*, for which the form *uryúsa* (Orani) 'fool's watercress' (Blasco Ferrer 1988: 175) might be evidence.

9.1.2.3 **li-*

Many of the words that start in **ϑi-*, listed in Table 9.1.2, contain a sequence **-li-*, thus starting in **ϑili-*.

	Meaning	Section	Forms with **ϑili-*	Forms without **ϑili-*	Reconst.	Alt.
1	Celery, fool's watercress	§ 3.1.6	*tseliɣúśa*	*ϑurgúśa, kugúśa, atigúsa* etc.	**ϑeli-kúsa*	✓
2	Backswimmer	§ 4.1.2	*tsirifrúči*		**ϑilifurke*	×
3	Earthworm	§ 4.1.10	*ϑilingròne* etc.		**ϑiling(u)lòne*	×
4	Grasshopper	§ 4.1.12	*ϑilipírke* etc.		**ϑilipírke*	×
5	Gecko	§ 4.3.1	*tilíbbu*		**ϑilikwu*	×
6	Lizard	§ 4.3.2	*ϑilikèrta* etc.	*kalužèrta, aliɣèrta* << Lat. *-lacerta*	**ϑilikerta*	×
7	Skink, slug	§ 4.3.4	*ϑilikúkku* etc.	*alicúcu*	**ϑilikúku*	×
8	Kestrel	§ 4.4.5	*tilibríkku, at-tiliβríu* etc.		**ϑiliprí(kk)u*	×

Table 9.1.5: Forms containing **ϑili-*

In the case of *ϑilikèrta* 'lizard' << Lat. *lacerta*, *-li-* can be equated with the first syllable of the Latin form. The other etyma, on the other hand, do not have a Latin etymology. This raises the question whether **-li-* might in fact be a separate pre-Roman morpheme. There is a single form in which **-li-* demonstrably does not belong to the lexical stem of the form in which it is attested, namely South Log. *tseliɣúśa* 'wild celery' (Paulis 1992: 141). It seems to be related to forms like *ϑurgúśa*, *kugúśa* and *atigúsa* (Puddu 2023). The other etyma do not attest to any alternations. Whether the frequent cooccurrence of **ϑi-* and **-li-* is a matter of chance, or a petrified remain of pre-Roman morphology, is difficult to establish on the basis of the form *tseliɣúśa* alone, not least because its variant *ϑurgúśa* is not fully explained either (cf. § 9.1.2.2). Unsurprisingly, Figure 9.1.4 shows that the distribution of forms containing **ϑili-* falls within the distribution of forms with **ϑi-* (Figure 9.1.2).

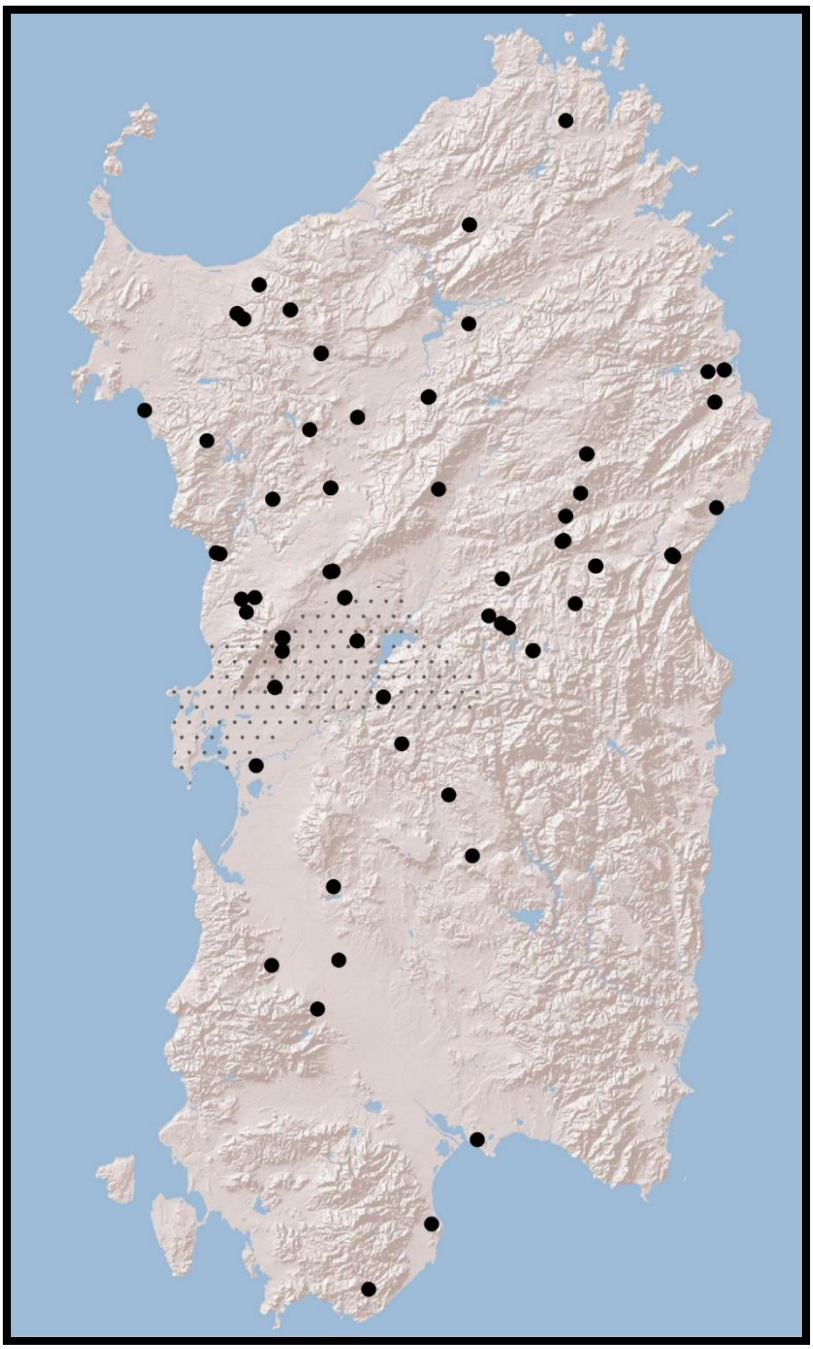

Figure 9.1.4: Distribution of forms containing **ϑili-*

9.1.3 Pre-Roman prefixes in Sardinian

In previous scholarship, there has been much attention on **ϑi-* as a potential pre-Roman prefix (§ 9.1.2). However, as we have seen, there is also evidence for the existence of a prefixed element **kV-* (§ 9.1.1). In addition, there are some indications, albeit much less strong, that there may have been a prefix **-li-* (§ 9.1.2.3). The sequence **ϑur-* does alternate with other sequences, but its single occurrence precludes any confident conclusions on its origin (§ 9.1.2.2). The identification of these apparent prefixes raises the question as to what their function was, and what they reveal about the morphology of the pre-Roman language to which they belonged.

9.1.3.1 The relation between **k(V)-* and **ϑ(i)-*

We have seen that the identified prefixes **k(V)-* and **ϑ(i)-* can occur on the same stems: e.g. *kalužèrta* ~ *ϑilikèrta* 'lizard', *kugúṡa* ~ *tseliɣúṡa* 'wild celery'. This suggests they were present in the same language and that they had a similar function. As illustrated in Figure 9.1.1 and Figure 9.1.2 **k(V)-* and **ϑ(i)-* have a different geographical distribution, with **k(V)-* occurring roughly in the southern half of Sardinia, and **ϑ(i)-* being more common in the north (cf. Swanenvleugel 2024: 259–261). Since Sardinian **ϑ-* in Latin sources goes back to palatalized *-tį̊-* and *-cį̊-*, **ϑ(i)-* may have been a palatalized variant of **k(V)-*. If one assumes this, multiple explanations are at hand for the variation between **k(V)-* and **ϑ(i)-* (Swanenvleugel 2024: 262–262). It could be conditioned by phonological factors, supported by the fact that **ϑ-* as a prefix is usually followed by the front vowel *-i-*. The variation could alternatively be due to dialectal variation in pre-Roman Sardinia, with **ϑ(i)-* being a variant of **k(V)-* in northern Sardinia. In any case, the proposed change from **k(V)-* to **ϑ(i)-* cannot have happened in Latin or Sardinian Romance, since inherited **ki-* does not palatalize in Logudorese and Nuorese, and palatalizes to *či-* in Campidanese, thus remaining distinct from **ϑi- > tsi-*.

9.1.3.2 The function of **k(V)-* and **ϑ(i)-*

By analyzing **k(V)-* and **ϑ(i)-* as variants of the same pre-Roman morpheme, the number of pre-Roman prefixes with sufficient evidence has been reduced to one (**ϑur-* and **-li-* being uncertain; § 9.1.2.2, § 9.1.2.3). In fact **ϑ(i)-* has been argued to originally not be a prefix at all, but a morphosyntactic element similar to an article (Wagner 1932: 223–224; 1997: 263; Pittau 1995: 197). The original grammatical function of **k(V)-* and **ϑ(i)-* is probably impossible to retrieve. Definiteness, noun class, number, size, proximity, pejorativity etc., all belong to the possibilities. This analysis comes closest to that of Pittau (1995: 197), although his com-

parison of **ϑi-* to the Etruscan demonstrative pronoun *ita/eta/ta* is formally problematic. Intriguingly, Etruscan had a functionally identical demonstrative pronoun *ica/eca/ca* (Wallace 2008: 59–63), which is formally a better comparison to Sardinian **k(V)-*.[266]

At some point in the history of Sardinian, **ϑi-* spread to other parts of the lexicon. It was secondarily introduced in inherited words; e.g. Log. *tilíbba* 'bean pod' < **ϑilikwa*, next to regular Camp. *siβíkkwa* < Lat. *siliqua* 'id.' (DES II: 417); *ϑilikèrta* 'lizard', next to *kalužèrta*, *aliyèrta* < Lat. *lacerta* 'id.'. It is furthermore found in such expressive nouns as *ϑiϑϑilíkkos* 'tickles' (DES II: 549). It is difficult to establish to what extent the spread of **ϑi-* was the result of analogy to pre-Roman words containing it, or of its productivity as an expressive marker. Serra (1956: 175) compares some Italian dialectal forms for 'slug', and 'snail' that seem to contain a prefixed *ca-*/*cia-* too; e.g. Garganic *calascertə* 'lizard' (Rodi Garganico) < Lat. *lucerta* 'id.', Abruzz. *cacozzə* next to *cozzə* (Città Sant'Angelo) 'slug', Campan. *ciammarruca* (Riardo, Versano) 'snail' etc. If and how these are related to Srd. **k(V)-* requires further research.

9.2 Suffixes

9.2.1 **-Ví*

Endings of the shape *-Ví* are common in Sardinian toponyms. They are found, for instance, in the toponyms *Ollolái*, *Orϑái* (It. Olzai), *Baunéi*, *Gavói* etc.[267] These were first identified as pre-Roman suffixes by Terracini (1927: 139). He argues that this pre-Roman *-Vi̯* may have had a function similar to Lat. *-ētum*, which forms collectives and is often used to denote places where specific plants grow; e.g. *pīnētum* 'pine grove' << *pīnus* 'pine', *vīnētum* 'vineyard' << *vīnum* (Weiss 2009: 293). This interpretation of *-Vi̯* is based on the observation that several toponyms containing this apparent suffix seem to have been derived from plant names of both pre-Roman and Latin origin; *Cannái* << *kánna* 'reed' (Terracini 1927: 139), *Tuvarai* << *túvara* 'heather' (Hubschmid 1963a: 151), *Colovrai* << *kolóvru* 'snake' etc. (Wagner 1997: 265).[268] Serra (1956: 189) attentively notes that various phonetic inputs could have yielded a sequence of the shape *-Vi̯*, and it is

[266] For the potential relatedness between Etruscan and a pre-Roman language of Sardinia, proposed by Pittau (1995), see § 10.4.

[267] For many more examples, see Blasco Ferrer (1984a: 10–11), Wagner (1997: 265) and Wolf (1998a: 53–55).

[268] Note that the form *Colostrais*, recorded in Terracini (1927: 139) and compared to Srd. *golóstri* etc. 'holly' by Bertoldi (1928: 231), is erroneous for *Tolostrai* (Serra 1956: 182–183).

therefore unlikely that all toponyms in *-*Vi̯* contain the same pre-Roman morpheme (cf. Wolf 1998a: 74–75 for examples).

Whereas Terracini (1927: 139) simply notes that *-ái*, *-éi* etc. could have had a function similar to Lat. *-ētum*, this has generally been taken to mean that *-Ví̯* was originally a pre-Roman collective suffix (Wagner 1997: 265; Hubschmid 1963a; Wolf 1998a: 53). Fact is however, that the pre-Roman Sardinian suffix *-Ví̯* is overwhelmingly attested in toponyms, many of which are semantically opaque (cf. also Serra 1956: 204–205). It is not fully clear to me why its identification as a collective is preferable over interpreting it simply as a suffix denoting a place. So, instead of analyzing *Cannái* and *Tuvarái* as collective 'bunch of reed' and 'bunch of heather' respectively, we may consider an analysis like 'place where reed grows', 'place where heather grows'. Such an interpretation is closer to the function of Lat. *-ētum*, and would also explain why *-Ví̯* is so frequent in toponyms. Furthermore, the fact that the few transparent placenames with *-Ví̯* are derived from plant names, does not imply that pre-Roman *-Ví̯* was attached to plant names only. The vast majority of toponyms ending in *-Ví̯* is semantically opaque, and may well have been derived from nouns in different semantic domains. Table 9.2.1 is an overview of the lexical evidence for pre-Roman *-*Vi̯*.

	Meaning	Section	Form	Reconst.	Alt.
1	Arum	§ 3.1.2	*sattsarói̯*, *tattaróyu*, *θoθθoróyu* etc.	**θaθarói̯*	×
2	Cockroach, darkling beetle	§ 4.1.7	*siśáya*, *seiśèi* etc.	**sisai̯a*, **sei̯sèi̯*	×

Table 9.2.1: Lexical forms containing *-Ví̯*

Only three lexical words discussed in § 3 – 7 exhibit an ending *-*Vi̯*, of which one (*golléi̯* 'plateau') is of Latin origin (§ 5.1.11). The two remaining examples are predominantly attested with a vowel following the sequence *-*Vi̯*-. This contrasts with what we find for toponyms in *-Ví̯*, which rather alternate with forms in final accented vowels; e.g. *Olloláis* ~ earlier *Ollala*; *Lodé* ~ earlier *Lotdé*, *Lodey* (Wolf 1988: 868–869). Because of a lack of related forms that do not contain *-*Vi̯*, there is no positive evidence that this sequence was ever a separate morpheme in these two words. Therefore, a connection with the hypothetical toponymic suffix *-*Ví̯* cannot be proven. As a consequence, the Sardinian lexicon does not provide evidence for the existence of a pre-Roman suffix *-Ví̯*, regardless of whether it had a collective or a locative function. Nevertheless, its preponderance in Sardinian toponymy still makes its morphemic status plausible.

9.2.2 *-*Vk*-

Endings of the shape *-*Vk*- are often found in toponyms; e.g. *Piaghe* (It. Ploaghe) < OSrd. *Ploauake*, *Plavaki*; *Carzeghe* (It. Cargeghe) < Osrd. *Carieke*; *Loʔomaʔe*; *Ortoiʔe*; etc. This sequence has been interpreted as a pre-Roman suffix (Wagner 1952: 17–18; Hubschmid 1963a: 165–167; Wolf 1998a: 74). In the lexicon, the same ending is notably present in the word *nuráke* 'nuraghe', which is almost certainly of pre-Roman origin.

A complicating factor for the identification of instances of pre-Roman *-*Vk*- is the fact that this sequence also has Latin sources. Primarily, it can be directly inherited for Latin words with suffixes *-āx*, *-īx*, *-ex* or *-ox*; e.g. Lat. *ferāx* 'fertile', *cīmex*, *-icem* 'bed-bug', *cornīx* 'crow', *ātrōx* 'fierce' etc. (cf. Weiss 2009: 305–307). Of these, especially Lat. *-āx* > Srd. *-áke*, *-áɣe*, *-áži* has had some degree of productivity; e.g. *preðáʔe* (Fonni) 'animal trap' << *prèða* 'stone' (Wagner 1952: 20). Other sources of Sardinian *-áke*, *-áɣe*, *-áži* are loans from Greek, Catalan, Spanish and Italian (Wagner 1952: 20). Unaccented *-́ike*, *-́iɣe*, *-́iži* has spread secondarily as well, first by accent shifts in inherited forms (e.g. Nuor. *préðike* 'partridge' << **perdíke* < Lat. *perdīx*), and later by productive suffixation (e.g. *lattúrike* (Bitti) 'spurge' << Lat. *lactoris*), particularly in the Nuorese dialects (Wagner 1952: 22–23). The productivity of these inherited suffixes raises the question whether a pre-Roman suffix *-Vk*- needs to be posited at all. After all, the semantic opacity of most pre-Roman toponyms hinders any confident identification of *-Vk*- as a suffix. Therefore, the acceptance of a pre-Roman *-Vk*- largely depends on the question whether the presence of this ending on lexical items of probable pre-Roman origin is original or due to productive extension of the same suffix from the inherited lexicon.

We will now turn to the lexical material discussed in § 3 – 7 whose stems end in *-Vk*- and which are therefore candidates for containing a suffix *-Vk*-. The material will be discussed separately according to the vowel in the ending.

9.2.2.1 *-*ák*-

All potentially pre-Roman Sardinian words ending in *-*ak*-, are stressed on the ending. This is in line with the behaviour of *-*ak*- in inherited words, which continue Lat. *-āx*. Excluding *surváke* 'horsetail' (cf. DES II: 450), which has been derived from *survare* 'to whistle' by Paulis (1992: 384), there are three candidates (Table 9.2.2).

	Meaning	Section	Form	Reconst.	Alt.
1	Oleander	§ 3.3.11	*neuláke, leonáržu, leonáǧi* etc.	**leonáke*	?
2	Cockroach, beetle	§ 4.1.6	*meláɣe, melogi* etc.	**meláke*	×
3	Nuraghe	§ 5.2.2	*nuráke, nuráɣe, nuráži, NURAC*	**nurák*(*e*)	×

Table 9.2.2: Forms containing **-ák-*

The obvious starting point for a discussion regarding hypothetical pre-Roman *-ák-* is the word *nuráke, nuráɣe, nuráži* 'nuraghe', which refers to the distinctive and typically Sardinian monuments pre-dating the Roman presence on the island. The word is actually attested in a Latin inscription as *NURAC* on a nuraghe itself. While some have analyzed **nuráke* of an **-ák-*derivation of a stem **nur-*, which is supposedly found in Old Sardinian toponyms such as *Nora, Nure* (Terracini 1927: 140; 1936: 55; Wagner 1952: 18; Hubschmid 1953: 45; Blasco Ferrer 2010a: 108–109, 115; 2011: 83), this analysis rests on formal similarity only.[269] The same goes for the few names of nuraghes starting in *nur-* (e.g. Nuridda; Hubschmid 1953: 45). As long as a semantic value of the element *nur-* cannot be established, it cannot be ruled out that **nurak-* was monomorphemic.[270]

The same issue affects the two other words containing **-ák-*, i.e. *neuláke, leonáržu, leonáǧi* etc. 'oleander' and *meláɣe* etc. 'cockroach, beetle'. The word for 'oleander' does exhibit an apparent alternation between **-áke* and **-áriu*, but *leonáržu* may have been reshaped from expected ***leonáži* < **leonáke*, rather than being formed independently from a root **leon-*. Paulis (1992: 419–420) proposes that *neuláke* 'oleander' is a derivation from inherited *néula* 'mist' by means of the supposed pre-Roman suffix *-áke*. As discussed in § 3.3.11, this does not satisfactorily explain other attested forms, including *leonáržu, leonáǧi* etc. Nevertheless, if Paulis is right, it is difficult to dismiss the possibility that *-áke* is in fact the suffix inherited from Lat. *-āx*, which was somewhat productive in Sardinian (Wagner 1952: 20). For *meláɣe* 'cockroach, beetle', there is no internal evidence for *-ák-* functioning as a suffix altogether. *Meláɣe* has been proposed by Paulis (1983: 52) to be of Byzantine origin, in which case it is irrelevant to the question whether **-ák-* existed as a pre-Roman suffix.

We are thus left with a single Sardinian word that is both undoubtedly of pre-Roman origin and in which the presence of **-ák-* is original, namely *nuráke*

[269] For a discussion of the proposed connection between Srd. *nuráke* etc. and *núrra* 'shaft; pile', cf. § 5.2.2.

[270] A connection between *-ák-* in **nuráke* and the Basque plural marker *-ak* (Blasco Ferrer 2010a: 115) is out of the question; cf. § 5.2.2.

'nuraghe'. The relevance of the next best candidate, *neuláke*, *leonáği* 'oleander', depends on one's interpretation of its etymology. Neither *nuráke* nor *neuláke*, *leonáği* offer compelling evidence for the view that **-ák-* was suffixal however. Without suffix-less bases of the type **nur-*, **neul-/*leon-*, there is no compelling reason not to treat *nuráke* 'nuraghe' and *neuláke, leonáği* 'oleander' as originally monomorphemic formations.

9.2.2.2 **-ik-*

There is some evidence for the existence of a pre-Roman ending **-ik-* (Table 9.2.3). We encounter forms with stressed *-ík-* as well as with unstressed *-́ik-*. Both of these also occur in the inherited lexicon, where they go back to Latin words in *-īx* and *-ex* respectively.

	Meaning	Section	Forms with **-ik-*	Forms without **-ik-*	Reconst.	Alt.
1	Poison hemlock	§ 3.1.11	*uḍḍúriɣu*	*buḍḍúri* etc.	**vullúri-ku*	✓
2	Holly	§ 3.3.7	*ʔolóstiʔe*, *golóstiʔe*, *olóstriɣe*	*golóstɨu* etc.	**golósti-ke*	✓
3	Maple	§ 3.3.9	*kòstike*, *kòstiɣe*	*kòsti*	**kòsti-ke*	✓
4	Spider	§ 4.1.14	*solóiɣa*, *soloíɣa*		**solo(C)ika*	×
5	Worm	§ 4.1.17	*tsarríɣa*, *tserríɣu*, *tsorroíɣu* etc.		**ϑVrroík-*	×

Table 9.2.3: Forms containing **-ik-*

The first three words in Table 9.2.3 denote plants. All three of them show an alternation between forms with and without unstressed **-́ik-*. This is good evidence that **-́ik-* in these words is of suffixal origin. Interestingly, in all three cases the suffix-less forms end in the vowel *-i*, making an analysis of this suffix as **-k-* more appropriate. All three phytonyms are generally taken to be of pre-Roman origin, which again suggests that **-(i)k-* was a pre-Roman suffix. However, Wagner (1952: 23) notes that in the Nuorese dialects, unstressed **-́ik-* is sometimes secondarily added to inherited bases; e.g. *lattúrike* (Bitti) 'spurge' < Lat. *lactoris*; Nuor. *gremédḍikes* (pl.) 'twins' << Lat. *gemellus*.

The details of this addition of **-́ik-* to inherited words are slightly different than those in *udḍḍúriɣu* (1), *ʔolóstiʔe* (2) and *kòstike* (3)(see Table 9.2.3). All instances of *-́ik-* in pre-Roman words are found in roots that ended in *-i*. The inherited examples of *lattúrike* and *gremédḍikes*, on the other hand, show that *-́ik-* could be added to inherited consonant stems as well as thematic stems. Nevertheless, it

still cannot be ruled out that the addition of *-(i)k-* to the originally pre-Roman words for 'poison hemlock', 'holly' and 'maple' is the result of the same secondary productivity of *-ík-*, perhaps aided by the fact that these words already ended in *-i*. As a consequence , they do not constitute compelling evidence for the existence of accented *-(i)k-* as a pre-Roman suffix. The remaining two words, which both refer to animals, show either stressed *-ík-* (5: *tsarríɣa*, *tsorroíɣu*), or an alternation between stressed and unstressed *-ik-* (4: *soloíɣa* ~ *solóiɣa*). This alternation has been argued to be secondary (cf. Wagner 1941a: 2–3; 1952: 23). At any rate, neither of these examples exhibits variants without *-ik-*, making the identification of *-ik-* as a suffix problematic. Note furthermore that *soloíɣa*, *solóiɣa* (4) 'spider' is probably related to Lat. *solifuga*, *solipuga* (§ 4.1.14), which clearly does not contain *-ik-*. For a discussion on the phonological discrepancies between these forms, see § 8.2.1.1.

9.2.2.3 On the existence of pre-Roman *-*Vk*-

On the basis of the forms discussed above, there is hardly any compelling evidence for the existence of a pre-Roman suffix *-Vk-*, whatever its vocalism. Whereas it is beyond doubt that there were pre-Roman words ending in such a sequence, most notably **nurak-* 'nuraghe', we cannot conclude with confidence that this sequence was an independent morpheme rather than a part of the lexical root. This also applies to the various toponyms ending in *-Vk-* listed by Wagner (1952: 17–18) and Wolf (1998a: 74). Conversely, for those words that do exhibit alternation between forms with and without unstressed *-ík-*, it cannot be ruled out that we are dealing with secondary productivity of an inherited morpheme.

9.2.3 *-*l*-

Forms ending in *-Vl-* are well-represented among pre-Roman Sardinian toponyms (Terracini 1927: 142; Hubschmid 1963a: 163–165; Wolf 1998a: 58–59, 64–65). In addition, the ending *-Vl-* occurs in various lexical words too. As a potential pre-Roman suffix, it has received less attention than *-*Vi̯* (§ 9.2.1) and *-*Vk-* (§ 9.2.2).

9.2.3.1 *-*al*-

Wolf (1998a: 58–59) records several Barbaricine toponyms ending in unaccented *-al-*. This is unlikely to go back to Latin, as in Latin unstressed short vowels in open syllables were largely reduced to *-i-* in open syllables, or to *-u-* before *l pinguis* (Weiss 2009: 116–117). As with all affixes discussed in this study, the question is whether *-al-* was originally a suffix in fact rather part of the stem. In the lexical

material discussed in § 3 – 7, there are two words that contain this sequence (see Table 9.2.4).

	Meaning	Section	Forms with *-*al*-	Reconst.	Alt.
1	Clothes moth	§ 4.1.5	*nástala, lanásta*	**nástala* (?)	×
2	Waterbasin, stream	§ 5.1.27	*ϑúrgalu* etc.	**ϑúrgalu*	×

Table 9.2.4: Forms containing *-*al*-

In the case of the word for 'clothes moth', there is some discussion whether *nástala* or *lanásta* is the original form of the word. If *lanásta* is original, then this word does not contain a suffix *-al-*. In any case, neither *nástala* 'clothes moth' nor *ϑúrgalu* 'waterbasin, stream' exhibit any alternation that can be taken as evidence for *-al-* being a suffix. On the basis of these data there is thus no reason to posit the existence of a pre-Roman morpheme *-al-*.

9.2.3.2 *-*el*-

The sequence *-él-* occurs in various Sardinian toponyms of plausible pre-Roman origin (Serra 1956: 203–204; Wolf 1998a: 64). A pre-Roman suffix *-él-* is posited in DES (II: 481) for a single lexical word: *tevèle* 'tilled land'. However, a similar ending is found in [*pan'*] *ispéli* 'acorn bread' (see Table 9.2.5).

	Meaning	Section	Forms with *-*el*-	Reconst.	Alt.
1	Tilled land	§ 6.1.8	*tèle, tevèle*	**t(ev)èle*	×
2	Acorn bread	§ 6.2.1	[*pan'*] *ispéli*	**ispéli*	×

Table 9.2.5: Forms containing *-*el*-

Wagner (DES II: 481) reconstructs *tevèle* as **tev-èle*, with *-èle* as an original suffix. This is based on the occurrence of various toponyms ending in *-èle*, *-éli*. However, without alternating forms of this word and without any other words containing *-èle*, the evidence for **-ele/*-eli* as a pre-Roman suffix is solely based on toponyms. This renders the identification of the ending *-èle* in *tèle*, *tevèle* 'tilled land etc.' rather speculative — the lexical base may well have been monomorphemic. The same applies to *ispéli* 'acorn bread', which does not offer any evidence for *-éli* being a suffix either.

9.2.3.3 *-*il*-

The sequence *-ile*, both accented and unaccented, is found in various Sardinian toponyms and lexical words (Table 9.2.6). It has been identified as a suffix and compared to ancient North-African and Iberian toponyms (Meyer-Lübke 1925: 77; Terracini 1927: 142; Wagner 1997: 257; Hubschmid 1963a: 152–153). Serra (1956: 203) argues that *-íle* was secondarily introduced in some inherited words, by

influence of pre-Roman "Afro-Ibero-Sardinian" words. Wolf (1998a: 59, 64) presents several toponyms that end in accented *-íle* or unaccented *-́ile*, but he does not offer direct evidence that these sequences are separate morphemes.

	Meaning	Section	Forms with *-*ile*	Forms without *-*ile*	Reconst.	Alt.
1	Bermuda grass	§ 3.2.1	*oroḍḍáṡile*	*oroḍḍáṡu*	*(*nodu-*)*ollás-ile*	✓
2	Weed	§ 3.2.6	*nuríle*		**nuríle*	×
3	Marten	§ 4.5.9	*kassíβile*, *assíle* etc.		*(*k*)*assíbile*	×
4	Plateau (etc.)	§ 5.1.13	*séttile*	*séttị̊a*, *séttị̊u*	**sétt*(*i*)*-ile*	✓
5	Secluded place	§ 5.1.24	*óspile*		**óspile*	×
6	Summit	§ 5.1.26	*težíle*		**tekíle*	×

Table 9.2.6: Forms containing *-*ile*

In two examples, the ending *-ile* is accented (2: *nuríle*, 6: *težíle*). In the other four, it is unaccented. We find an alternation between suffixed and un-suffixed forms in two words only: *oroḍḍáṡile* ~ *oroḍḍáṡu* (1) 'Bermuda grass' and *séttile* ~ *séttị̊u* (4) 'plateau (etc.)'. For the other forms, the status of *-íle/-́ile* as a suffix is less certain. Moreover, both Srd. *-íle* and *-́ile* are also inherited from Lat. *-īlis* and *-ĭlis* and form productive derivations in Sardinian (Wagner 1952: 40–44). For *oroḍḍáṡile* and *séttile* it is therefore probable that *-́ile* is in fact due to productivity of a suffix inherited from Latin.

There is thus no convincing lexical evidence for *-ile* being a pre-Roman suffix. For the majority of the relevant words it is impossible to judge whether *-ile* did or did not originally belong to the lexical stem. For *óspile* 'secluded place' (5) it may be noted that it has been compared to Bq. *ospel* 'shadowy place' (Wagner 1932: 227; DES II: 198), which according Trask (2008: 320) may contain the root **bel* 'dark'.[271] If this is correct, the *-́ile* in *óspile* cannot be an original suffix.

9.2.3.4 On the existence of pre-Roman *-*Vl*-

As in the case of *-*Vk*- (§ 9.2.2.3), the lexical data do not provide sufficient evidence for the existence of any pre-Roman suffix of the shape *-*Vl*-. The only words that do show an alternation are *oroḍḍáṡile* ~ *oroḍḍáṡu* 'Bermuda grass' and *séttile* ~ *séttị̊u*, but *-íle* and *-́ile* are in fact also inherited suffixes. The fact that the

[271] See § 10.3 for a discussion of the proposed Sardinian-Basque comparanda.

lexical material does not provide evidence for a suffix (or several different suffixes) *-*Vl*- does not preclude that it is a suffix in pre-Roman toponymic forms (Terracini 1927: 142; Hubschmid 1963a: 163–165; Wolf 1998a: 58–59, 64–65). It may well be that the lexical evidence simply is too scarce. There is however also a reasonable possibility that at least some of the attested sequences -*Vl*- are originally no suffixes, but integral parts of the stem, both in lexical and toponymic forms.

9.2.4 *-*n*

9.2.4.1 *-́*Vn*-

Various pre-Roman Sardinian toponyms end in unaccented -́*Vn*- (Hubschmid 1963a: 170; e.g. Wolf 1998a: 58). The same ending is found in a small number of lexical items (Table 9.2.7).

	Meaning	Section	Forms with *-́*Vn*-	Reconst.	Alt.
1	Galingale	§ 3.2.3	*sèssene* etc.	**sèssene*	×
2	Resin	§ 3.4.5	*tsònkine* etc.	**ϑònkene*	×
3	Bat	§ 4.5.1	*čížini* etc.	**kíkin-*	×
4	Border	§ 5.2.1	*lákkana* etc.	**lákkana*	×

Table 9.2.7: Forms containing *-́*Vn*-

None of the words in Table 9.2.7 show an alternation between variants with and without final *-́*Vn*-. The existence of a pre-Roman suffix *-́*Vn*- can therefore not be established on the basis of the lexical evidence.

9.2.4.2 *-*òne*

The sequence -*òne*, including its regular dialectal correspondents, is frequent in the Sardinian lexicon. It has various sources. Mainly, it continues the suffix of Latin nouns in -*ō* (cf. Weiss 2009: 309–314), which like in other Romance languages is productive as an augmentative suffix (Wagner 1952: 57–58). Terracini (1936: 59–60) notes however that there is a significant number of evidently non-native words ending in the same suffix in Latin sources. Among these are the animal names *lalisiones*, *musmones*, *theldones* attested in Pliny. This leads Terracini to suggest that all instances of the Latin suffix -*òne* were eventually borrowed from non-Indo-European languages that Latin was in contact with. This is too bold a statement, as it is generally held that Lat. -*ō*, -*ōnem* continues the IE -*n*-stems (Weiss 2009: 309). Nevertheless, it remains plausible that the increased productivity o *-*one* in the Romance languages is the result of a conflation of the suffix -*òne* in both inherited and non-inherited words. Table 9.2.8 shows the non-inherited Sardinian etyma ending in -*òne*.

	Meaning	Section	Forms with *-*òne*	Forms without *-*òne*	Reconst.	Alt.
1	Annual mercury	§ 3.1.1	*kaðòne* etc.		**katòne*	×
2	Celery, fool's watercress	§ 3.1.6	*turgusòne*	*ϑurgúsa*	**ϑurgus-òne*	✓
3	Smilax	§ 3.3.15	*titt̰ìone* etc.	*tétti*	**tetti-òne*	✓
4	Strawberry tree	§ 3.3.16	*oliðòne* etc.		**olidòne*	×
5	Strawberry tree fruit	§ 3.3.18	*liṡòne*, *biliṡòne*, *liṡiriòne*		**lis*(*ir̰i*)*-òne*	✓
6	Earthworm	§ 4.1.10	*ϑiling*(*r*)*òne*	*tilínğa*, *tilíngu*	**ϑiling*(*ul*)*-òne*	✓
7	Plover	§ 4.4.7	*kulurdzòne*, *kulirğòni* etc.		**kulVr̰ìone*	×
8	Robin	§ 4.4.10	*griṡôi*	*gríṡi*	**gris-òne*	✓
9	Mouflon	§ 4.5.10	*murvòni*, *muɣròne* etc.	*múrva*	**mufr-òne*	✓
10	Rock crevice	§ 5.1.20	*maraɣòni*		**marakòne*	×
11	Pile of rocks	§ 5.2.3	*marganğòni*		**marganı̯ìone*	×
12	Tangle/bundle of wood	§ 7.2.4	*barraṡòne*, *bariṡòne* etc.	*barraṡólu* etc.	**bar*(*r*)*Vs-òne*	✓

Table 9.2.8: Forms containing *-*òne*

Seven of the twelve words exhibit a morphological alternation that shows that *-òne* is indeed a suffix.[272] Because of the productivity of inherited *-òne* however, it is impossible to demonstrate that *-òne* was already present in the respective pre-Roman source words. In fact, words like *titt̰ìone* and *liṡiriòne* show that *-òne* was added to *tétti* and *liṡiri-* at a late stage, since the inherited clusters **-tt̰i̯-* and **-r̰i̯-* would be expected to yield ***-tt-* and ***-ğğ-* respectively in the dialects in which these variants are attested. Possibly, the addition of *-òne* was a strategy of incorporating non-inherited nouns into the native lexicon. Some of the words in Table 9.2.8, although probably not inherited from Latin, are attested with the same suffix already in Classical times: *kaðòne* (1; cf. Lat. *catone*), *oliðòne* (4; cf. Lat. *ūnedō*), and *murvòne* (9; cf. Lat. *mufron*). With the exception of Lat. *mufron* 'mouflon', which specifically refers to a Sardinian and Corsican animal, these words are not restricted to Sardinia and are thus not indicative of a specifically pre-Roman Sardinian suffix **-òne*. Despite the uncertainty of the origin of **-òne*

[272] Although both *liṡòne* and *liṡiriòne* 'strawberry tree fruit' contain *-òne*, the fact that the element *-iri-* occurs between *liṡ-* and *-òne* in *liṡiriòne*, suggests that *-òne* is a separate morpheme.

in individual words, its frequent occurrence on Sardinian words of probable pre-Roman origin, as well as on pre-Roman toponyms of Sardinia (Terracini 1936: 59–60; Wolf 1998a: 63–64), indicates that there may have been a morpheme *-*òn*- in pre-Roman Sardinian. In the process of language shift to Latin, this pre-Roman *-*òn*- merged with the various inherited and non-inherited sources of Latin and Romance *-*one* and became a productive suffix. Because of this conflation however, it is impossible to determine with certainty which words contain pre-Roman -*òne*, and which productively added *-*òne* of inherited Latin origin.

9.2.5 *-*nk*-

Suffixes of the type *-*nk*- have often been posited for toponyms and lexical words in other regions of the Mediterranean (e.g. Hubschmid 1963a: 157). There is one instance in which we encounter an ostensible alternation of forms with and without an apparent suffix -(*a*)*nka*. The element in question is *kalánka* 'rock crevice, precipice, cave' (§ 5.1.19), which may be derived from *kála* '(animal) shelter; sheltered place' (§ 5.1.3)(Table 9.2.9).

	Meaning	Section	Forms with *-*nk*-	Forms without *-*nk*-	Reconst.	Alt.
1	Rock crevice, precipice, cave	§ 5.1.19	*kalánka*	**kála* 'animal den' (§ 5.1.3)	**kalánka*	✓

Table 9.2.9: Forms containing *-*nk*-

Both *kála* and *kalánka* have wide-spread correspondents in the western Mediterranean Romance languages. The presumed suffix-hood of *-(*a*)*nka* is predominantly based on the semantic proximity of **kála* and **kalánka*, and on the existence of other non-inherited words ostensibly containing the same suffix, such as **barrank*- 'gorge', and **palank*- 'roller, beam' (FEW II: 56). However, the presence of both **barrank*- and **palank*- in Sardinian is due to borrowing from Spanish or Catalan (cf. DES I: 181, II: 207). Moreover, Romance **palanka* hails back to Gr. φάλαγξ, -αγγος 'roller, beam' (FEW VIII: 356). Besides the coexistence of phonologically and semantically similar **kála* and **kalánka* the only other lexical evidence for *-*nk*- as a suffix is found in formations like Log. *troḍḍiánkulu*, Camp. *troḍḍiánčulu* 'someone who farts' << *troḍḍiáre* 'to fart' and Log. *piβiánkulu* 'cuddly' << *piβíu* 'caress' (Wagner 1952: 113). In this type of formation, *-ank*- clearly has an affectionate meaning. A diminutive meaning is found for *-onk*-, in Camp. *pillónku* 'skin, film' << *píllu* 'layer' (Wagner 1952: 112). These uses of -(*a*)*nk*- and -*onk*- can hardly explain *kalánka* 'rock crevice, precipice, cave' as a derivation

from *kála* '(animal) shelter; sheltered place' and are likely unrelated morphological elements.

The sequence *-*nk*- is reported to be rare in Sardinian toponyms by Hubschmid (1963a: 157). It is not mentioned at all in the toponymic studies by Wolf (1998a) and Blasco Ferrer (2011: 82–85). This reduces the evidence of *-*nk*- as a pre-Roman derivational suffix to a single lexical word, which is not sufficient to substantiate its existence in any pre-Roman Sardinian language. One possible explanation for the coexistence of *kála* and *kalánka* is that they are in fact not etymologically related, but two distinct monomorphemic words **kal*(*a*) and **kalank*(*a*) that entered Sardinian from a pre-Roman source. Alternatively, if one does not want to give up the link between these two words, it is possible to assume that Srd. *kalánka* entered Sardinian in more recent times from one of the other Romance languages in which it occurs.[273]

9.2.6 *-*r*-

Endings containing non-geminate *-*r*- can be divided into accented *-*úri* (§ 9.2.6.1) and unaccented *-*V́r*- (§ 9.2.6.2). The latter in particular has frequently been discussed as a potential pre-Roman morpheme.

9.2.6.1 *-*úri*

The ending -*úri* occurs on two plant names recorded in DES (Table 9.2.10).

	Meaning	Section	Forms with *-*úri*	Forms without *-*úri*	Reconst.	Alt.
1	Poison hemlock	§ 3.1.11	*biḍḍúri*, *buḍḍúri* etc.	*búḍḍaru*, *ḍḍúḍḍulu* etc.	**vull-úri*	✓
2	Ropegrass	§ 3.2.5	*karkúri*, *krukkúri* etc.		**kalkúri*	×

Table 9.2.10: Forms containing *-*úri*

[273] Bottiglioni (1929: 58) suggests a similar scenario for Cors. *calanca* too, making Corsican an unlikely source for the Sardinian forms.

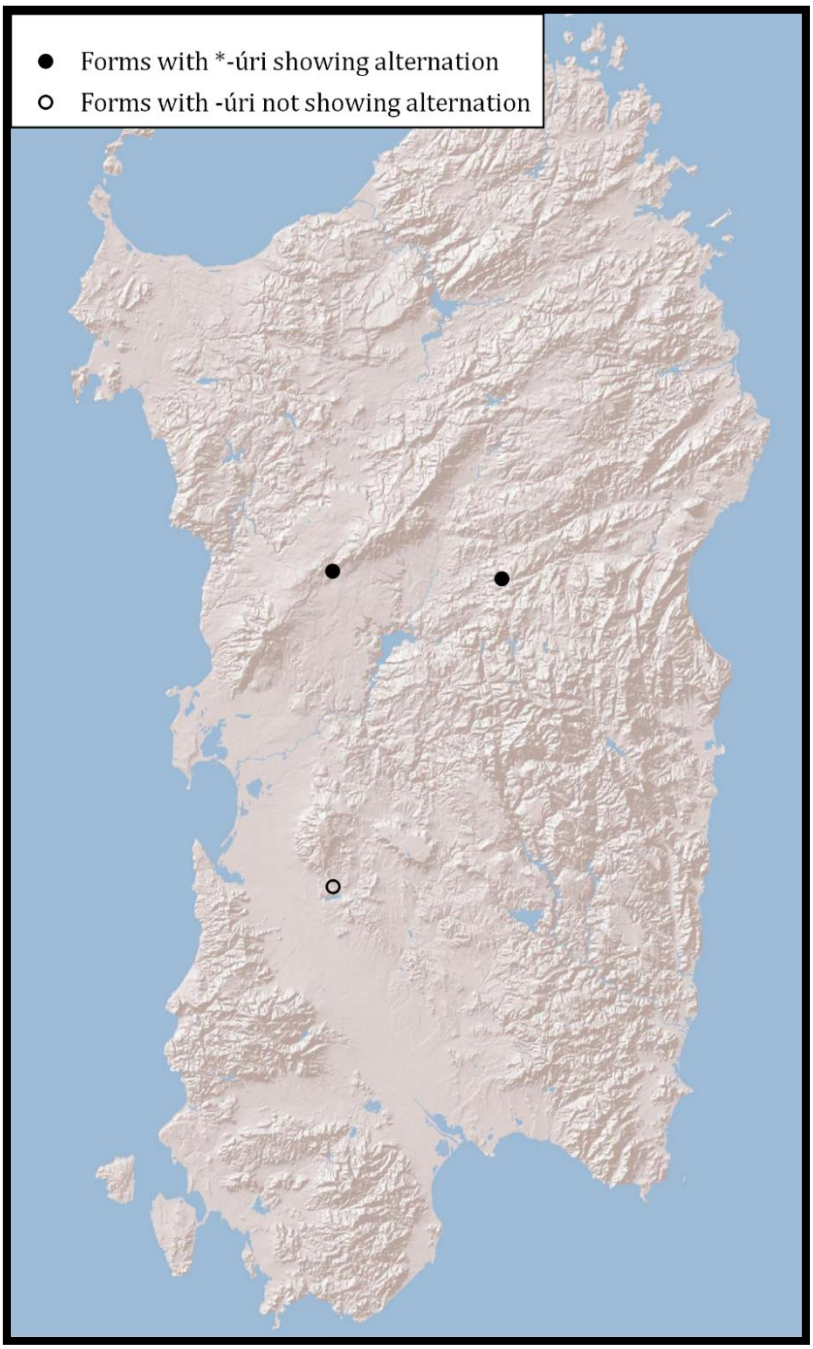

Figure 9.2.1: Distribution of forms with *-*úri*

The various forms for 'hemlock' are built from a root *buḍḍ-* < **vull-* with different suffixes: *buḍḍúri* < **b/vull-úri* ~ *búḍḍaru* < **búll-ar-* (cf. § 9.2.6) etc. This suggests that *-úri* was originally a distinct morpheme in the source language of this non-inherited word. In *karkúri* we do not find an alternation. It has been interpreted by Paulis (1992: 269) as a pre-Roman word that was folk-etymologically influenced by Camp. *krákkili*, *krákkiri* 'thick bush, dense thickets'. Paulis (1992: 269) further mentions the form *pistiḍḍúri*, a variant of *pistiḍḍòre* 'stinging nettle', of sound-symbolic origin (cf. DES II: 280).

Despite the scarcity of the relevant material, there appears to be sufficient evidence to identify *-úri* as an originally pre-Roman suffix. It is impossible to say anything about its function based on only two forms. Not much can be deduced from the geographical distribution of forms with **-úri* either (Figure 9.2.1). Wolf (1998a: 68) lists some Barbaricine toponyms in *-úri*, which may or may not contain the same element.

9.2.6.2 **-́Vr-*

Unstressed **-́Vr* has frequently been discussed with regard to pre-Roman Sardinian morphology. Terracini (1927: 139) first observed the existence of older forms of toponyms like *Mandara* for present day *Mandas*, and *Gennor* for later *Gennos*. Because of the replacement of this element *-́ar(a)/-́or* with the regular nominal plural suffix, he hypothesizes that *-́ar/or* may have been a plural ending. This interpretation is adopted by Wagner (1997: 264), Hubschmid (1963a: 175), Blasco Ferrer (1984a: 10; 2019: 39), Paulis (1985: 232a), and Pittau (1995: 196–197). Unstressed endings of the type *-́Vr* are common in other Sardinian toponyms as well, including in the longer sequence *-ÍnnVr* (Serra 1956: 184; Hubschmid 1963a: 174–175; Wolf 1998a: 59–61). In addition, many lexical words of non-Latin origin end in unaccented *-́Vr*. The question is whether these too are consistent with the analysis of *-́Vr* as a plural ending. Table 9.2.11 contains all etyma that end in unstressed **-́(V)r-*.

	Meaning	Section	Forms with **-́Vr-*	Forms without **-́Vr-*	Reconst.	Alt.
1	Bristly ox-tongue	§ 3.1.4	*ɡiòkkoro*, *sóčči̱ri* etc.		**ɡiokkVr-*	×
2	Poison hemlock	§ 3.1.11	*búḍḍaru*, *úḍḍuru*	*buḍḍúri*, *ḍḍúḍḍulu* etc.	**vúll-Vr-*	✓
3	Mullein	§ 3.1.12	*túmbara*	*kaðúmbu* etc.	**túmb-ar-*	✓
4	Poppy	§ 3.1.15	*tzàndhara*	*aɡáṇḍa*, *tɡáṇḍa* etc.	**ɡánda-r-*	✓
5	Wild teasel	§ 3.1.22	*billóttiri*	*billótti*	**bili̱ótti-r-*	✓

	Meaning	Section	Forms with **ˊ-Vr-*	Forms without **ˊ-Vr-*	Reconst.	Alt.
6	Watercress	§ 3.1.24	*ğúɣuru, ğúru* etc.		**ɟúgur-*	×
7	Knotgrass; horsetail	§ 3.2.4	*síntsiri, sintsurru*		**ϑínϑur*	×
8	Heather	§ 3.3.5	*túora, túvara, túvera* etc.		**túfVr-*	×
9	Rosemary	§ 3.3.14	*tsíppiri* etc.		**ϑíbbir*	×
10	Strawberry tree fruit	§ 3.3.18	*liṡiriòne*	*liṡòne* etc.	**lis-iri-òne*	✓
11	Yew	§ 3.3.23	*táṡaru, táṡuru* etc.	*táṡulu*	**tásVr-*	?
12	Juniper berry	§ 3.4.3	*bòḍḍoro, bòḍḍero*		**bòllVro*	×
13	Walnut	§ 3.4.8	*kòkkoro*	*kòkkolo*	**kòkkoro*	?
14	Grasshopper	§ 4.1.11	*pibittsíri* etc.	*piβiϑϑíu* etc.	**pipiϑí-r-*	✓
15	Robin, wagtail; hare	§ 4.4.9	*príncˇiri*	*brínčis, prinčóttu* etc.	**brink/č-*	✓
16	Fawn, lamb; female mouflon	§ 4.5.5	*bíttara, bíturu*	*bíta, bítti* etc.	**bitt-Vr-*	✓
17	Hare	§ 4.5.8	*lèppere, lèpporo, léppuri* etc.		**leppVr-*	×
18	Hillock	§ 5.1.6	*mòɣoro* etc.		**mòkor-*	×
19	Rock outcropping	§ 5.1.21	*tón(n)eri*		**tónneri*	×
20	Rocky plain	§ 5.1.23	*tsèppara*	*Theppage* (Top.)	**ϑéppar-*	?

Table 9.2.11: Forms containing unstressed **ˊ-Vr-*[274]

[274] The addition of *-òne* in this word must have taken place at a relatively recent date (cf. § 9.2.4.2).

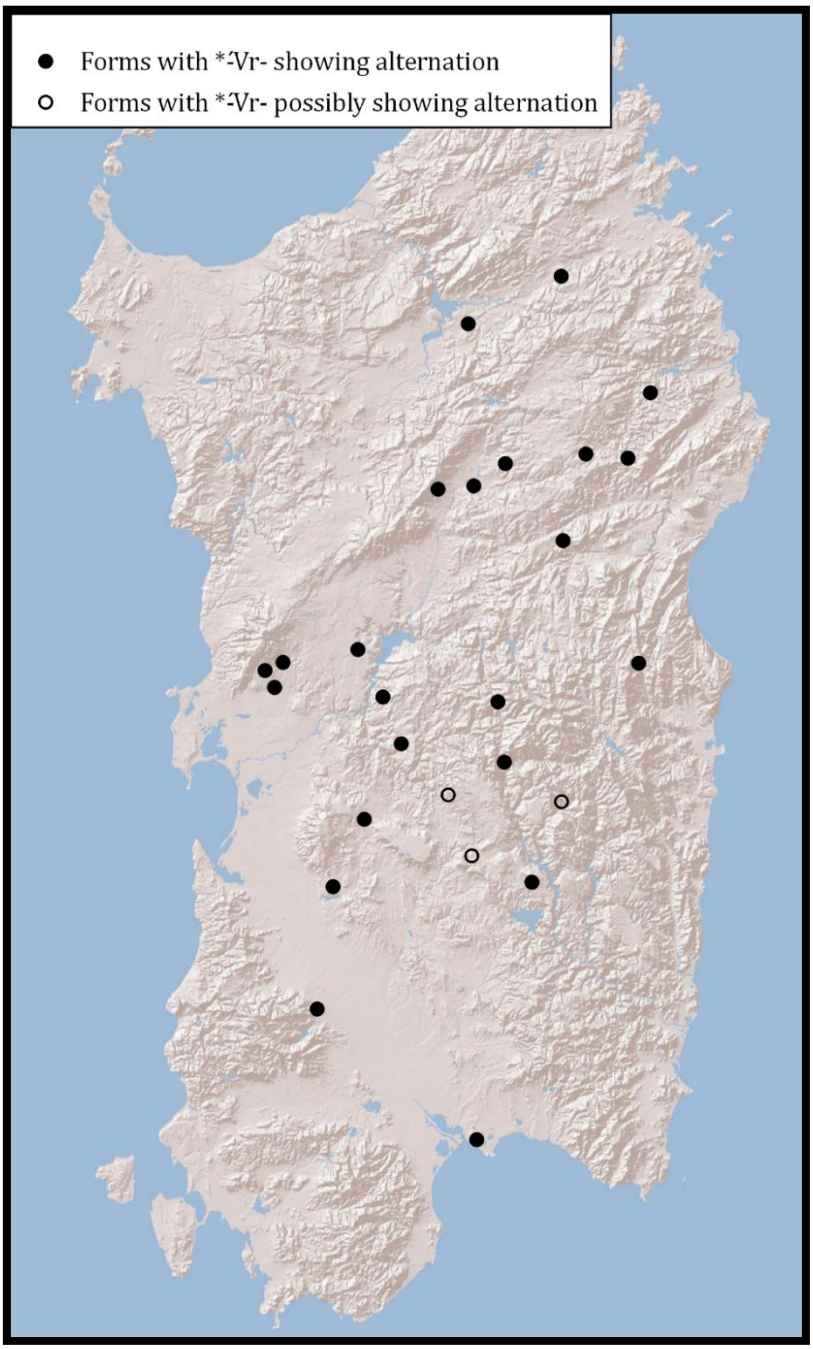

Figure 9.2.2: Distribution of forms with *-*Vr*- showing an alternation

At least two of the etyma ending in *-Vr* are of probable Punic origin: *síntsiri*, *sintsurru* (7) 'knotgrass, horsetail', cf. Pun. *senzur*, *zenzur* 'id.'; and *tsíppiri* etc. (9) 'rosemary', cf. Pun. *zibbir* 'id.' Since these forms were borrowed integrally from the language of the Punic colonists, it is evident that at least here the ending *-Vr* was not a separate morpheme and in general unrelated to the words presumably originating in the native pre-Roman languages of Sardinia.

9.2.6.2.1 Alternating *-Vr-

Of the remaining words, eight show an alternation between forms with and without *-(V)r-*. For a few others, the evidence for an alternation is less strong. The alternation *táṡaru* ~ *táṡulu* (11) 'yew' could be due to secondary replacement with *-ul-*; the same goes for *kòkkoro* ~ *kòkkolo* 'walnut'. In the alternation *tsèppara* ~ *Theppage* (20) 'rocky plain', the etymological relationship between the two forms is uncertain, due to the latter being a toponym. Figure 9.2.2 shows the distribution of forms in which the morphemic status of *-Vr-* is guaranteed by alternations; they occur across Sardinia.

9.2.6.2.2 Vocalism of *-Vr-

As for the vocalism of the suffix *-Vr*, we predominantly find *-ar-* (9x), less frequently *-or-* (6x) and *-ir-* (4x, excluding *síntsiri* and *tsíppiri* which are of Punic origin). The vocalisms *-er-* (3x) and *-ur-* (4x) cooccur mainly with other variants, suggesting an assimilatory or dissimilatory origin. In *piβiɣɣíri* and possibly in *tzàndhara*, the suffix is simply *-r-* without a vowel; cf. *piβiɣɣíu* and *ɣánda* respectively. In some cases the vowel in the suffix is identical to the preceding stressed vowel (e.g. 15: *príńčiri*), seemingly in line with the vowel harmony proposed for pre-Roman Sardinian (§ 8.4.2). In other cases, however, the vowel in the suffix is different from the stressed vowel (e.g. 1: *sóčči̱ri*, 4: *túmbara*). The available evidence does not allow for confident conclusions on the rules governing this suffix's vocalism. It must also be kept in mind that vowel changes are rather common in unaccented syllables within Sardinian, which may account for some of the observed variation (Wagner 1941a: 33–36). The word-final vowel is in most cases identical to the vowel preceding *-r-*. This points to a paragogic origin of the final vowels, and to a reconstruction of this suffix as *-Vr* (cf. § 8.4.1.2).

9.2.6.2.3 Function of *-Vr-

For all of the animal and plant names, borrowing of a plural form into Sardinian Latin is imaginable. The meaning 'rocky plain' of *tsèppara* also refers to an inherently plural or mass concept. However, none of these forms present unequivocal evidence that *-Vr* was a plural marker. The evidence for its function thus still

rests on the interpretation of the alternations found in toponyms like those discussed by Terracini (1927: 139).

9.2.6.2.4 **-́Vr-* outside Sardinia

Among the words that end in *-́Vr*, there are several that have been compared to words of non-Latin origin in other languages (Table 9.2.12).

	Meaning	Section	Sardinian	Non-Sardinian
1	Poppy	§ 3.1.15	*tzàndhara*	Lang. *ãndèr*
2	Juniper berry	§ 3.4.3	*bòḍḍoro, bòllero*	Sic. *bóḍḍaru,* Tusc. *bòllora*
3	Walnut	§ 3.4.8	*kòkkoro*	Tusc. *cóccoro*, Laz. *kùkkęrę*, Nap. *còccola*, Pg. *cócora* etc.
2	Hare	§ 4.5.8	*lèppere*, *lèpporo* etc.	Lat. *lepus*, Massiliot Gr. λεβηρίς, Pg. *láparo*, Fr. *lapéreau* etc.'
3	Hillock	§ 5.1.6	*mòɣoro* etc.	Prov. *maguelo*, Alb. *magulë*, Rom. *măgură* etc.

Table 9.2.12: Non-Sardinian comparanda to Sardinian forms with **-́Vr-*

In the cases of *bòḍḍoro* 'juniper berry' and *lèppere* 'hare', the *-́Vr*-sequence is attested in their proposed cognates. In the case of *mòɣoro* 'hillock', the *-r-* rather corresponds to an *-l-* elsewhere. For the words for 'juniper berry' and 'hare', this implies that *-́Vr-* was not a specifically Sardinian suffix. The question is whether the sequence belonged to the lexical root of the word, or whether a suffix *-́Vr-* found in pre-Roman Sardinian words was present in other pre-Roman languages of the Mediterranean region as well. An argument for the latter option are formations like Fr. *lapin* 'rabbit', Pg. *lapim*, Arag. *llapí* 'young rabbit' < **lappin-*, which contain the same root **lapp-/*lepp-* as the words for 'hare'. This is evidence in favor of a pre-Roman language affiliation between Sardinia and southern continental Europe (cf. § 11.4). A similar suffix -αρ- has also been identified in presumable pre-Greek words. Like in Sardinian it occurs most frequently on plant names, but its exact function is difficult to interpret from the lexical data (Furnée 1972: 257 fn. 36). If Srd. *kòkkoro* 'walnut' and its Romance comparanda go back to Lat. *coccum* 'scarlet oak gall' (DES : 356–357), it appears that unstressed **-Vr-* was added independently in multiple languages. Bertoldi (1937a: 167) interprets instances of **-́ar-* attested in various languages as original plurals, but because the words in question do not have clear comparanda in Sardinian, it is difficult to evaluate whether all examples are instances of one and the same pre-Roman suffix.

9.2.6.2.5 Secondary productivity of *-Vr-*

Secondary productivity of pre-Roman *-Vr-* in Sardinian words of Latin origin has been proposed for *túmbara* 'mullein' (§ 3.1.12), *ğúɣuru* 'watercress' (§ 3.1.24), *túvara* 'heather' (§ 3.3.5), and *táśaru* 'yew' (§ 3.3.23) (Paulis 1992: 245–246, 355, 410, 415). While not all of the etymologies proposed by Paulis can be accepted, the secondary productivity of *-Vr-* is demonstrated by forms like *píśiri* 'pea' << Lat. *pīsum* (Paulis 1992: 273; more examples in Serra 1960: 405). However, it is unclear whether the productivity of this suffix is a direct continuation of its hypothetical pre-Roman plural function, or whether it was added to other plant names analogically within Sardinian because of its frequent occurrence in phytonyms of pre-Roman origin.

9.2.7 *-Vrr-*

A suffix *-Vrr- has been posited for Sardinian pre-Roman linguistic elements. The evidence mainly consists of toponyms like *Ilarra*, *Motorra*, *Ludurru* etc. (Terracini 1927: 143; Wagner 1952: 107; Hubschmid 1963a: 150; Wolf 1998a: 70). It has been taken as evidence for prehistoric linguistic ties between Sardinia and the Iberian Peninsula, and to a lesser degree with Liguria and Etruria (Bertoldi 1928: 228–229; Hubschmid 1963a: 150). Terracini (1927: 143) additionally considers a North-African connection. Table 9.2.13 shows the forms discussed in § 3 – 7 that end in a sequence *-Vrr-.

	Meaning	Section	Forms with *-Vrr-	Reconst.	Alt.
1	Rush skeletonweed	§ 3.1.20	*lipòrra, limpòrra, lispòrra*	**liCpòrra*	×
2	Knotgrass; horsetail	§ 3.2.4	*sintsurru, síntsiri*	**ϑinϑur*	×
3	Caterpillar	§ 4.1.4	*kukúrra, kuɣúrra*	**kukúrra*	×
4	Eel	§ 4.2.1	*tsingòrra*	**ϑingorra*	×
5	Cheek	§ 7.1.3	*tutúrru*	**tuttúrru*	×

Table 9.2.13: Forms containing *-Vrr-*

None of the items in the Sardinian lexicon that end in *-Vrr- show an alternation of this ending. As a result, there is no convincing lexical evidence for the claim that a sequence *-Vrr- was a suffix in any pre-Roman Sardinian language (cf. Iribarren Argaiz 1997: 348–349). At face value, the words in Table 9.2.13 may have been monomorphemic. This is the case, at any rate, for *sintsurru* 'knotgrass', which was borrowed as such from Punic *zenzur* (§ 3.2.4). It should further be kept in mind that the Sardinian sequence *-rr-* can go back to both *-rr- and *-rn-, in Campidanese additionally to *-rl- when metathesized from *-rV̆l- (e.g. Camp. *méurra* < Lat. *merula* 'blackbird')(Wagner 1941a: 172–174; Wolf 1998a: 70; 1998b:

345–346). The etyma in Table 9.2.13 could thus represent three originally distinct sequences. As Wolf (1998a: 70) notes, toponyms in *-rr-* in Sardinian are not especially frequent, raising doubt about the interpretation of *-*Vrr-* as a suffix in toponyms as well. Without strong evidence for a suffix *-*Vrr-*, it cannot reliably be used as an argument for prehistoric linguistic relations between Sardinia and other Mediterranean regions, as Terracini (1927), Bertoldi (1928) and Hubschmid (1963a) do.

9.2.8 *-*s-*

A pre-Roman Sardinian suffix containing *-*s*(*s*)- has been posited on the basis of Sardinian toponyms like *Ulassai*, *Ussassai*, *Gonnesa*. These formations have been connected to ancient Iberian and Ligurian toponyms ending in *-issa*/*-essa* (Terracini 1927: 143; Bertoldi 1928: 242; 1931: 168–170; Hubschmid 1963a: 155). Hubschmid (1963a: 155) posits a possessive function for the putative *-*s*(*s*)-suffix in these toponyms. Bertoldi (1931: 167–171) cites toponyms containing a similarly shaped suffix in Basque, and interprets them as locative derivations. In the Sardinian lexicon, several words offer potential evidence for a putative suffix of the shape *-(*V*)*s*- Table 9.2.14.

	Meaning	Section	Forms with *-(*V*)*s*-	Forms without *-(*V*)*s*-	Reconst.	Alt.
1	Holly, butcher's broom	§ 3.2.1	*aláṡe*, *aláṡ̱u* etc.		**alás-*	×
2	Lizard; snake	§ 4.3.3	*tsorómpis*	*zorrómpi*	**ʒorrómpi-s*	✓
3	Viperine snake	§ 4.3.5	*lírčis*	*lúrtsi*	**lí/úrči-s*	✓
4	Robin, wagtail; hare	§ 4.4.9	*bríṅčis*, *bríṅčiṡi*	*bríṅči*, *príṅčiri* etc.	**brinči-s*	✓
5	Robin, wren	§ 4.4.10	*kiríṡi* etc.	*kiríu*, *kíri* etc.	**kiri-s-*	✓
6	Humid place	§ 5.1.7	*orgòsa*	*òrga*	**org-òs-*	✓
7	Tangle/ bundle of wood	§ 7.2.4	*barraṡòne*, *bariṡòne* etc.; cf. Cat. *barda* 'portable fence' etc.		**barr(-)Vs-*	?

Table 9.2.14: Forms containing *-*Vs-*

Among the seven etyma potentially possessing an *-*s*-suffix, five show an alternation between forms with and without *-(*V*)*s*- within Sardinian. The other two have comparanda in other languages that lack *-(*V*)*s*-. Srd. *aláṡe* < **alás-* (1) 'holly, butcher's broom' has been connected to Occitan and Italo-Romance words for 'broom' containing a base **ala-* (§ 3.2.1). Srd. *barraṡòne*, *bariṡòne* (7) 'tan-

gle/bundle of thorns etc.' has been connected to Ibero-Romance and Occitan forms going back to **barda/barra*, whose continuants denote various kinds of fences, barriers and hedges made of tangled-up plant material (§ 07.2.4). The alternation of this suffix in multiple etyma confirms the existence of a pre-Roman suffix *-(*V*)*s*-. The low number of examples hampers a functional interpretation of the suffix *-(*V*)*s*-. Bertoldi (1931: 167–171) assumed a locative function and Hubschmid (1963a: 155) a possessive, but neither seems imaginable for a suffix found mainly with words for plants and animals.

9.2.9 *-*st*(*r*)-

A suffix *-*st*(*r*)- has been posited by Bertoldi (1928: 231–233; 1930) based on substrate forms across the Mediterranean, and he argues that it formed collective nouns from plant names (Bertoldi 1930: 193). Krahe (1925: 68–71; 1929: 148) records a range of ancient place names containing *-st-*, to which he attributes an Illyrian origin. A suffix *-est*(*r*)*a*, *-ist*(*r*)*a* occurs in a number of Latin words; e.g. *genesta*, *genista* 'broom', *arista* 'awn', *lanista, lanistra* 'trainer of gladiators' etc. Herbig (1917: 165–179) argues for an Etruscan origin of *-est*(*r*)*a*, *-ist*(*r*)*a* in Latin. However, Breyer (1993: 100–102) notes that the Etruscan sequences of *-esta*, *-ista* can only be explained as the agglomeration of a genitive case ending with an enclitic demonstrative pronoun, rendering borrowing into Latin unlikely. The Sardinian continuants of Latin words ending in *-est*(*r*)*a*, *-ist*(*r*)*a* cannot inform us on the pre-Roman substrate of Sardinia and are therefore left out of consideration. There is nevertheless a small number of words in the non-inherited Sardinian lexicon that end in *-st*(*r*)- (Table 9.2.15).

	Meaning	Section	Forms with *-st(r)-	Forms without *-st(r)-	Reconst.	Alt.
1	Holly	§ 3.3.7	*golóstḍu*, *golóstri*, *olósti* etc.		**golósti*	×
2	Tree heather	§ 3.3.20	*giḍḍòstre*, *iḍḍòstro* etc.		**gillóstr-*	×
3	Clothes moth	§ 4.1.5	*lanásta*, *nástala*	*lána* (?)	**lanásta*	?
4	Lizard	§ 4.3.2	*luǧèstra*, *luǧǧèsti*	*aliɣèrta*, *kalužèrta* etc.	**luk-est*(*r*)-	✓

Table 9.2.15: Forms containing *-*st*(*r*)-

There are four words that potentially contain a suffix *-*st*(*r*)-. One of these is the word for 'lizard'. However, its ending *-èstra* etc. is likely secondary from inherited *-èrta* (< Lat. *lacerta*), which was remodeled under influence from the Latin pejo-

rative suffix *-āster* (Wagner 1952: 120; Weiss 2009: 286). *Lanásta* 'clothes moth' occurs next to *nástala*, and may have been metathesized under influence of *lána* 'wool' (§ 4.1.5).

The other two words do not show alternation. In addition, their endings do not fully correspond to each other. The forms of the word for 'holly' seem to go back to **-ósti*, with secondary introductions of *-r-* and *-u* etc (§ 3.3.7). The words for 'tree heather' (§ 3.3.20) on the other hand all contain **-óstr-* but lack the final vowel *-i* found in **golósti*.[275] What is more, none of the convincing comparanda to **golósti* 'holly' across the Mediterranean show any trace of an alternation of **-st(r)-*. The only potential evidence for the presence of a suffix in this Mediterranean word for 'holly' is the comparison to more northern European forms like PC **kolesno-*, PGm. **hulisa* 'holly', although a suffix **-t-* would suffice.

A similar issue applies to **gillostr-* 'tree heather'. It has been compared by Hubschmid (1953: 29) to Bq. *gillar*, *gilharre* etc. 'heather', which would demonstrate the suffixal nature of Srd. *-òstr-*. However, this Basque etymon has more recently been reconstructed as **inar* (Trask 1997: 306; Santano Moreno 2000: 170), which is phonologically too far removed from Srd. **gillostr-* for a plausible comparison.

In conclusion, following the criteria outlined in § 2.2.2, there is insufficient evidence to posit a pre-Roman Sardinian suffix **-st(r)-*. Within Sardinian there is no reliable case of an alternation that demonstrates its suffix-hood, nor do the proposed instances of the suffix correspond to each other. The elusive and diffuse nature of the presumed **-st(r)-*suffix in Sardinian are unexpected. Bertoldi (1928: 230–233) has shown that non-inherited formations containing **-st(r)-* are rather common in other languages across the Mediterranean. In fact, some pre-Roman Sardinian words lacking **-st(r)-* have plausible comparanda with **-st(r)-* in other languages, e.g. Srd. *aláṡe* 'butcher's broom, holly' vs. Sic. *alastra*, Lig. *arástra* 'spiny broom' etc. (cf. § 3.2.1). In light of these comparisons, a possibility is that a pre-Roman Sardinian suffix **-st(r)-* did exist, but that it did not survive by chance, with the possible exceptions of **golósti* and **gillóstr-*. On the basis of the available data however, this remains speculative.

9.3 Conclusion

In this chapter, the lexical evidence has been examined for the affixal nature of various sequences in the pre-Roman language(s) of Sardinia. Confident identifi-

[275] All other four vowels, *a, e, o, u*, are in fact attested in word-final position in this etymon.

cation of sequences as independent morphemes, rather than as parts of the lexical root, depends on the demonstrable presence of alternations (§ 2.2.2). For any non-inherited word, the pre-Roman provenance of the affix can only be confidently assumed if there is no inherited Latin affix that could have produced the same result. Although the presence of a hypothetical pre-Roman affix on toponyms cannot be used as strong evidence in favor of its morphemic status, toponyms can provide clues on the original function of the affix (e.g. § 9.2.6.2). Table 9.3.1 lists all discussed hypothetical morphemes along with the number of relevant etyma, how many of these etyma show an alternation, whether the sequence can be inherited from Latin, and whether the sequence is found in toponyms. The last column gives an evaluation of how much evidence the lexicon provides for the identification as a pre-Roman morpheme.

Affix	Section	Number of forms	Alt.	Inherited source	Toponymy	Pre-Rom. morpheme?
**kV-*	§ 9.1.1	6	5x	×	×	Likely
**ϑi-*	§ 9.1.2	14	6x	×	×	Likely
**ϑVϑV-*	§ 9.1.2.1	4	4x (?)	×	×	No evidence
**ϑur-*	§ 9.1.2.2	1	1x	×	×	Uncertain
**li-*	§ 9.1.2.3	8	1x	×	×	Uncertain
**-Vi̯*	9.2.1	2	×	×	✓	No evidence
**-ák-*	§ 9.2.2.1	3	×	✓	✓	No evidence
**-ík-/-ik-*	§ 9.2.2.2	5	3x	✓	?	No evidence
**-ʹal-*	§ 9.2.3.1	2	×	×	✓	No evidence
**-él-*	§ 9.2.3.2	2	×	×	✓	No evidence
**-íl-/-ʹil-*	§ 9.2.3.3	5	1x	✓	✓	No evidence
**-òne*	§ 9.2.4.2	12	7x	✓	✓	Likely
**-nk-*	§ 9.2.5	1	1x (?)	×	×	Uncertain
**-úri*	§ 9.2.6.1	2	1x	×	✓	Likely
**-ʹVr-*	§ 9.2.6.2	18	8x	×	✓	Likely
**-Vrr-*	§ 9.2.7	5	×	×	×	No evidence
**-(V)s-*	§ 9.2.8	6	4x	×	✓	Likely
**-st(r)-*	§ 9.2.9	4	2x (?)	×	×	No evidence

Table 9.3.1: Lexical evidence for pre-Roman morphemes

As Table 9.3.1 shows, there is good evidence for the pre-Roman origin of two of the five discussed elements preceding the lexical root: **kV-* and **ϑi-*. As discussed in § 9.1.3, these are likely originally variants of one and the same morpheme, which entered the Latin lexicon at various different times and places (Swanen-

vleugel 2024). For the elements **ϑur-* and **li-* there is some evidence, but both only alternate in a single form, and their occurrence seems to be restricted to words that are also attested with **ϑi-*. Without additional lexical evidence, or their identification in toponymy, their origin and function remains elusive.

Of the twelve suffixes discussed, only five can confidently be identified on the basis of the lexical data as actual morphemes of pre-Roman origin. For *-òne*, *-úri* and *-(V)s* we find alternations, indicating that they were at some point morphemes. All three are additionally attested in pre-Roman toponyms, strengthening this interpretation. For *-òne* it is difficult to decide which instances are of pre-Roman origin, and which from the inherited and productive Latin suffix *-ōnem*. Because the forms exhibiting an alternation generally do not show different meanings, their function is difficult to recover. This is especially the case for pre-Roman *-òne*, whose original functions were likely obfuscated by those of inherited *-òne*.

The difficulty in identifying the original grammatical functions of the morphemes *-òne*, *-úri*, and *-(V)s* demonstrates the paradox in recognizing and analyzing substrate morphology. We are only able to confidently identify an alternation of affixes in non-inherited words if the suffixed and non-suffixed variants do not differ too much in meaning. Otherwise, they may not be recognized, or the semantic distance renders the whole comparison too speculative. This means however that the affixes that can be identified are inherently difficult to assign a function to, due to their semantic "blandness". It is very much possible that the Sardinian lexicon contains forms that preserve pre-Roman derivational suffixes that significantly altered the meaning of the derivational base. However, if the suffixed and un-suffixed forms are semantically too far removed, we may not be able to convincingly identified them as cognates and to prove the morphemic status of the suffix.

The suffix **-́Vr* (§ 9.2.6.2) shows all possible lines of evidence allowing for its identification as a pre-Roman morpheme. It is found in alternation in seven cases, and cannot be from an inherited source. It is even attested in toponyms that allow us to infer its original function. Terracini (1927: 139) has shown on the basis of the replacement of this suffix in toponyms by the inherited plural ending *-s*, that *-́Vr* was likely a plural morpheme. Still, in the words ending in *-́Vr-* that do not exhibit an alternation, the sequence may have been part of the root. Moreover, there is evidence that *-́Vr-* gained some productivity within Sardinian (§ 9.2.6.2.5).

Finally, for many proposed pre-Roman suffixes, there is insufficient lexical evidence. The elements **-V́ị*, **-ák-*, **-́ik-*, **-él-*, **-nk-*, **-Vrr-* and **-st(r)-* have all been posited as pre-Roman suffixes, partly on the basis of toponyms. At some occasions, they have even been used to argue for a pre-Roman origin of certain words in which they occur. Yet, it is concluded that the available lexical evidence does not allow us to confidently identify them as suffixes.

IV
External comparisons

10 Proposed connections to known languages

Parts of the non-inherited Sardinian lexicon have been compared to various languages spoken elsewhere in the Mediterranean region. This chapter discusses the proposed connections between Sardinian and Punic, Berber, Basque, and Etruscan respectively. The comparisons of the non-inherited Sardinian lexicon to substrate phenomena outside Sardinia are discussed in § 11.

10.1 Punic

A long-lasting Punic presence on Sardinia is well-established (§ 1.3.3.1). Various non-inherited Sardinian words have been argued to be Punic loans. These can be divided into two groups. One contains the words that can be attributed to the Punic language because they are morphologically analyzable as Semitic. The other contains Sardinian words that have correspondences in classical sources for which a Punic origin is likely, even when a Semitic etymology is lacking. Although the second group was likely introduced in Sardinia during the Punic colonization, their lack of a Semitic etymology raises the question whether they did indeed originate in Phoenician, or perhaps in the language(s) of native North Africans that were involved in Carthaginian colonial enterprises. Berber comparanda could be evidence of this (cf. Argiolas 2010: 23–24)(§ 10.2).

10.1.1 Lines of evidence for Punic loans

Most of the evidence for Punic loans in Sardinian Latin does not come directly from Punic texts. In some cases Sardinian, words are compared to forms from other Semitic languages, in which case the existence of the Punic word is inferred. Stronger are the cases in which a Punic origin of Sardinian words is based on words attested in Classical Roman or Greek texts, especially when they are attributed to Punic, or to "African".[276] The main Classical sources in which Punic-Sardinian comparanda have been identified are medical works like Dioscorides' *De materia medica* and Pseudo-Apuleius' *Herbarius*. Most comparisons are therefore specific plant names, for which it is not always easy to establish a convincing Semitic etymology. In several cases, the Punic forms in question have been compared to modern Berber forms. The possible interpretations of this will be discussed in § 10.1.4.

[276] On the imprecision with which Classical authors grouped various peoples of North-Africa together under terms like *Afri* or *Punici*, see Múrcia Sànchez (2011: 362, 616–617).

10.1.2 The lexical evidence

The following is an overview of the Sardinian words that were potentially introduced to Sardinia in the context of the Phoenician and Punic colonization of Sardinia. Most of these have been proposed by previous authors, some are proposed in this study. In each case, the compared forms are printed in bold; any relevant Semitic or Berber comparanda are included as well. For each comparison, an evaluation is offered of the formal and semantic correspondence between the compared forms:

1. § 3.1.1: Srd. *kaðòne*, *kaðòni* etc. 'annual mercury; goosefoot' < ***katòne*** (DES I: 259; Paulis 1992: 374–375); cf. Lat. ***catone*** 'orache' (Dioscorides; cf. Stadler 1899: 186, 219). Further comparisons include Kb. (*w*)*aktun* 'goosefoot' < ***(ā)k(V̌)tūn***.

   Formal match: **good**
   Semantic match: **good**
   Cognacy: **compelling**
   Origin: uncertain

2. § 3.1.7: Srd. *tsikkiría* etc. 'false fennel' < ****ϑikkiría*** (DES II: 589); cf. Pun. σικκιριά 'dill' (DES II: 589; cf. Dioscorides Anazarbei 1906: III: 58 RV). Further comparisons include Hebr. שֵׁכָר *šēkār* 'alcoholic drink' < Proto-Semitic **šikar*- 'id.' (Blau 1873: 527; Kogan 2012a: 241), Gr. σίκερα 'fermented liquor, strong drink' (cf. Liddell et al. 1978: s.v. σίκερα), Tsh. *taskra* '*Echinops spinosissimus*' (van den Boogert 1998: 178, 184),[277] <*tāsakra*, *tāsikrā*, *sakrā*> (*'Umdat aṭ-ṭabīb*; cf. Tilmatine & Bustamante 2002: 433) 'black chameleon', < *taskrat*, *teskera*> '*Echinops spinosissimus*' < ****tā-sǐkrā*** < **sǐkkǐryā* (Múrcia Sànchez 2011: 433).[278]

   Formal match: **good**
   Semantic match: **good**
   Cognacy: **compelling**
   Origin: possibly Semitic, if from **sikar*- 'alcoholic drink'

[277] Defined in a medieval wordlist as 'dill (Anethum graveolens)' (van den Boogert 1998: 178, 184).

[278] If one reconstructs not two but just one short high vowel for Proto-Berber, this reconstruction would be **tā-səkrā* < **səkkəryā* (cf. Kossmann 2020a: 35–36).

3. § 3.1.19: Srd. *kúrma, kúruma, cruma* etc. 'fringed rue; cat thyme' < ***kúrma*** (Farina 1987: 363; 1989: 620; Paulis 1990: 608–612; 1992: 229–231); cf. Pun. **χουρμά** 'rue' (Dioscorides Anazarbei 1906: 57; Vattioni 1976: 527–528; Paulis 1990: 608–612; 1992: 231). Further comparisons include Akk. *ḫurmu* 'a plant species' (CAD VI: 266), Syr. (i.e. Dioscorides' Σύροι) ἁρμαλά 'rue (*Ruta chalepensis*)' (Dioscorides Anazarbei 1906: 57), Bb. <*āwurmī*> (*ʿUmdaṭ aṭ ṭabīb*; Tilmatine & Bustamante 2002: 428), Tsh. *iwrmi* 'rue'. For the Berber forms, Múrcia Sànchez (2011: 440) proposes a derivation from **urme* < ***(h)urmay*** < Pun. **ḥurmal*.

   Formal match: **good**
   Semantic match: **good**
   Cognacy: **compelling**
   Origin: possibly Semitic

4. § 3.1.23: Srd. *grúspinu, gúspinu, grúspi̯u, óspinu* 'watercress'[279] < ****gúspinu*** (DES I: 595 Paulis 1992: 241–242); cf. Pun. *cusmin* 'watercress' (Pseudo-Apuleius 20; Howald & Sigerist 1927: 59; Paulis 1992: 242).

   Formal match: **good**
   Semantic match: **good**
   Cognacy: **compelling**
   Origin: uncertain

5. § 3.2.4: Srd. *síntsiri, sintsurru* 'knotgrass; horsetail' < ****ϑinϑi/ur*** (Paulis 1992: 386–387); cf. Pun. ***sensur, zenzur, zunzur, sumsur*** 'knotgrass' (Howald & Sigerist 1927: 55; Paulis 1992: 386–387; cf. André 1985: 279).

   Formal match: **good**
   Semantic match: **good**
   Cognacy: **compelling**
   Origin: uncertain

6. § 3.3.14: Srd. *ϑíppiri, tsíppiri* etc. 'rosemary' < ****ϑibbir***- (DES II: 593); cf. Pun. ***zibbir*** 'rosemary' (Pseudo-Apuleius 80: 31; Howald & Sigerist 1927: 145; Bertoldi 1947;[280] DES II: 594). Further comparisons include Bb.

[279] Also in the meaning 'butcher's broom, but this is probably secondary (cf. § 3.1.23).
[280] *Non vidi.*

<*āzīr*> 'rosemary' (*ʿUmdat aṭ-ṭabīb*; Tilmatine & Bustamante 2002: 429), <*iazir, aziir, ouzbir*> 'rosemary' (Trabut 1935: 223), Rif., Tmz. *azir* 'rosemary, lavender; wormwood' etc. < ***ā-zV̌βīr*** (Múrcia Sànchez 2011: 428–429).

Formal match: **good**
Semantic match: **good**
Cognacy: **compelling**
Origin: possibly Berber

7. § 3.4.6: Srd. *tséurra* etc. 'sprout, bud, shoot' < ***ϑérula*** << *ϑèra (DES II: 589); cf. Pun. ζερα[φοίστ] (Dioscorides Anazarbei 1907: 177), Hebr. זֶרַע *zéraʿ*, Akk. *zēru, zarʾu* 'seed' (Paulis 1990: 616), reconstructed by Kogan (2012a: 200) as PSem. *ẓarʿ-.

Formal match: **good**
Semantic match: **good**
Cognacy:**compelling**
Origin: Semitic

8. § 4.2.1: Srd. *tsingòrra* etc. 'eel' < ***ϑingòrra*** (DES II: 592); cf. Akk. ***singurru, zingurru*** etc. 'a kind of fish), Mandaic *singūrā* (of unknown meaning)(Wagner 1997: 160; cf. Zimmern 1917: 52; CAD XV: 284-285).[281]

Formal match: **good**
Semantic match: **uncertain**
Cognacy: **uncertain**

9. § 5.1.5: Srd. *ǧára* 'basaltic and granitic highlands' < ***i̯ára*** (DES I: 603); cf. Hebr. יער ***yáʿar*** 'wood, bush' (Paulis 1994; cf. Putzu 2019: 6).[282]

Formal match: **good**
Semantic match: **good**

[281] This would be a Punic loan, with the Akkadian form as evidence for its existence in Semitic (Wagner 1997: 160). Paulis (1990: 621–623) proposes a different etymology (cf. § 4.2.1).
[282] I have unfortunately not been able to consult Paulis (1994).

Cognacy: **compelling**
Origin: Semitic

10. § 5.2.5: Srd. *míttsa, míntsa, míϑϑa* 'source' < ****míϑϑa*** (AIS 852; DES II: 120–121); cf. Pun. ***mysehi***, interpreted as 'my exit' (Plautus, Poenulus V: 921; Wagner 1957: 105–106; Friedrich 1957: 223), Hebr. מוֹצָא *mōṣā* 'exit' < PSem. **yṣ'* 'to exit' (Friedrich 1957: 223; DES: 120-121; Paulis 1990: 600). Pun. *mysehi* is however read as *mysethi* 'I have come' by Krahmalkov (2000: 303; 2001: 159). Srd. **míϑϑa* is explained by Guarnerio (1906: 245), REW (5614) and Mensching (1997: 298–301) as a derivation from Lat. *mītis* 'ripe, mellow', which is equally possible.

    Formal match: **good**
    Semantic match: **good**
    Cognacy: **unnecessary**, due to a plausible Latin etymology and possible different interpretations of the Punic form.

11. Srd. ***kému, kémbu, kèmos*** 'four, five, a handful, small quantity' (DES I: 327-328);[283] cf. Hebr. חָמֵשׁ ***ḥāmḗš*** 'five' (Spano 1863: 91 fn. 3; Wagner 1997: 160; 1957: 106–108; DES I: 327-328; but cf. Friedrich 1957: 323–324; Paulis 1990: 617–620, who are skeptical).

    Formal match: **problematic**
    Semantic match: **possible**
    Cognacy: **unlikely**

10.1.3 Discussion

Most of the proposed connections between Punic and Sardinian words can be accepted on formal and semantic grounds. The comparison between Srd. **ϑingòrra* 'eel' (8) and Akk. *singurru* 'a fish species' and Mandaic *singūrā* is intriguing, but since no Punic form is attested and because of the vague semantic identification of the Akkadian and Mandaic forms, a Punic loan in Sardinian rests on too many assumptions to be convincing (cf. Paulis 1990: 621–623). A Punic origin for Srd. **míϑϑa* 'source' (10) is phonologically and semantically possible if one assumes a Punic cognate **mūṣō* 'exit of water' on the basis of Hebr.

[283] Wagner (DES I: 327-328) suggests *kèmos* to be original, metathesized from **kòmes*, but Paulis (1990: 618–620) shows that both **kòmes* and **kèmos* are formally impossible reconstructions. He proposes the Srd. *kèmos* to be from Old Cat. *qüern* 'number of four'.

mōṣā 'exit'. This is rather speculative, especially since Wagner (1957: 105–106) and Friedrich's (1957: 223) interpretation of Pun. *mysehi* as meaning 'my exit' is disputed (cf. Krahmalkov 2000: 303; 2001: 159). Mensching (1997) has shown that a derivation from Lat. *mītis* 'ripe, mellow' for a body of water has parallels elsewhere in Romance, rendering a Punic etymology unnecessary. The connection of Srd. *kému, kèmos* 'four, five, handful' to Hebr. חָמֵשׁ *ḥāméš* (Spano 1863: 91 fn. 3; Wagner 1997: 160) must be rejected on formal grounds (Friedrich 1957: 323–324; Paulis 1990: 617–620). We are left with eight likely loans that entered Sardinian Latin from the language of the Punic population in Sardinia. Six of these are words for specific plant species (1–6), one is the word for 'sprout' (from the Punic word for 'seed'; 7), and one denotes a type of terrain (9). A further candidate could be Srd. *fríśa, fríssa*, which is cognate to *afrissa* 'dragon lily' (§ 3.1.8) found in Pseudo-Apuleius (cf. André 1985: 7). However, Pseudo-Apuleius does not identify this word as specific to Africa, and the only indication favoring a possible Punic origin is that the work as a whole was composed there (Pradel-Baquerre 2013: 20–21). Table 10.1.1 is an overview of the evaluation of the proposed Punic loans in Sardinian.

	Srd.	Comp.	Form.	Semant.	Cogn.	Origin	In Bb.
1	**katòne* 'annual mercury; goosefoot'	Lat. *catone* Bb. **(ā)k(V̌)tūn*	✓	✓	✓	Uncertain	✓
2	**ϑikkiría* 'false fennel'	Pun. *σικκιριά* Bb. **tā-sǐkrā*	✓	✓	✓	Possibly Sem.	✓
3	**kúrma* 'fringed rue; cat thyme'	Pun. χουρμά Bb. **urme*	✓	✓	✓	Likely Sem.	✓
4	**gúspinu* 'watercress'	Pun. *cusmin*	✓	✓	✓	Uncertain	×
5	**ϑinϑi/ur* 'knotgrass; horsetail'	Pun. *sensur, zenzur, zunzur*	✓	✓	✓	Uncertain	×
6	**ϑibbir* 'rosemary'	Pun. *zibbir* Bb. **ā-zV̌βīr*	✓	✓	✓	Maybe (Libyco-)Bb.	✓
7	**ϑérula* 'sprout, bud'	Pun. ζερα 'seed'	✓	✓	✓	Sem.	×
8	**ϑingòrra* 'eel'	Akk. *singurru*, Mandaic *singūrā*	✓	?	?	Uncertain	×
9	**ɟára*	Hebr. יער *yáʿar*	✓	✓	✓	Sem.	×

	Srd.	Comp.	Form.	Semant.	Cogn.	Origin	In Bb.
10	**míϑϑa* 'source'	Pun. **mūṣō*, Hebr. מוֹצָא *mōṣā*	✓	✓	×	Lat.	×
11	*kému*, *kèmos* 'handful, low number'	Hebr. חָמֵשׁ *ḥāmḗš*	×	?	×	Uncertain	×

Table 10.1.1: Proposed punicisms in Sardinian.

10.1.3.1 Distribution of Punic words in Sardinia

Four out of seven Sardinian words (viz. 2: **ϑikkiría*, 5: **ϑi/unϑur*, 6: **ϑíppiri*, 7: **ϑérula*) with an accepted Punic etymology are attested predominantly in the Campidanese dialects, which is roughly in line with Wagner's (1997: 158) prediction (see § 1.3.3.1). Representatives of **katòne* (1) are found not only in Campidanese, but also in Logudorese, while **kúrma* (3) and **gúspinu* (4) occur in the Nuorese dialects (cf. Figure 10.1.1). Additionally, Paulis (1990: 602) reports toponymic attestations of *típpiri*, the expected Logudorese outcome of **ϑíppiri* 'rosemary', to the northwest of the area where *tsíppiri* 'rosemary' is attested as a lexical word. As Paulis (1990) concludes, it is thus clear that Punic linguistic influence in Sardinia reached farther than assumed by Wagner (1997: 158), although it is most clearly visible in the Campidanese dialects.

10.1.3.2 Punic loans with a Semitic etymology

Among the eight accepted Punic loans in Sardinian, only two are easily etymologizable within Semitic:

- Srd. **ϑérula* 'sprout' (7) << Pun. ζερα < PSem. **zrʿ* 'seed'
- Srd. **ɟára* 'basaltic and granitic highlands' (9), cf. Hebr. יער *yáʿar* 'wood, bush'

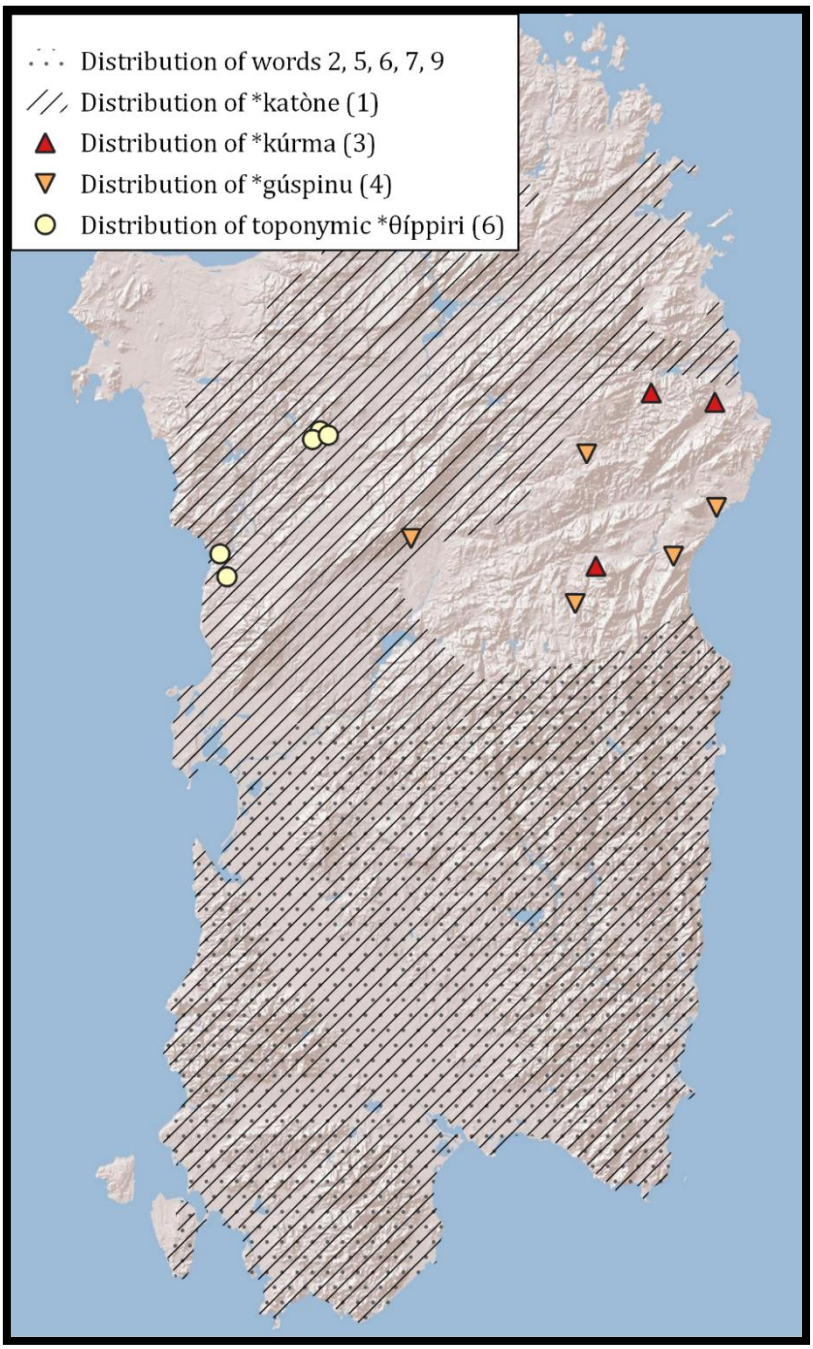

Figure 10.1.1: Distribution of punicisms in Sardinian

A Semitic etymology has also been proposed for "African" σικκιριά 'dill' (2) by Blau (1873: 527), who compares Hebr. שֵׁכָר *šēkār* 'alcoholic drink'. Reflexes of Proto-Semitic **šikar-* 'alcoholic drink' and the verbal root **škr* 'to become inebriated' are well-attested across Semitic (Kogan 2012a: 241). In order for Punic σικκιριά to be of Semitic origin, we would thus need to accept the semantic shift 'alcoholic drink' > 'dill'. The other Punic word with a proposed Semitic etymology is "African" χουρμά 'rue' (3), which has been compared to Akk. *ḫurmu* 'a plant species' (CAD VI: 266), Syr. (i.e. Dioscorides' Σύροι) ἁρμαλά 'rue (*Ruta chalepensis*)' (Dioscorides Anazarbei 1906: 57). Vattioni (1976: 527–528) and Múrcia Sànchez (2011: 440) reconstruct **ḥrml* and **ḥurmal* respectively. It is unclear to me how the *-l* would have been lost in Pun. χουρμα >> Srd. *kúrma*.

10.1.3.3 Punic loans with comparanda in Berber

Four of the eight accepted Punic loans have been compared to modern Berber forms:

- Srd. **katòne* 'annual mercury; goosefoot', Lat. *catone* 'id.', Bb. **(ā)k(V̌)tūn* 'id.' (1)
- Srd. **ϑikkiría* 'false fennel', "African" σικκιριά 'dill', Bb. **tā-sǐkrā* '*Echinops spinosissimus*' (2)
- Srd. **kúrma* 'fringed rue; cat thyme', "African" χουρμά 'rue', Bb. **urme* 'id.' (<< Pun. **ḥurmal*) (3)
- Srd. **ϑibbir* 'rosemary', Pun. *zibbir* 'id.', Bb. **ā-zV̌βīr* 'id.' (6)

In the case of Bb. **urme* 'rue' (3), a connection to Srd. **kúrma* is only plausible by assuming that both were borrowed from Pun. **ḥurma(l)* (Múrcia Sànchez 2011: 440). For Lat. *catone* 'orache' (1), a North African origin is in fact only supported by its exclusive occurrence in the Latin translation of Dioscorides (Stadler 1899: 186, 219), which may be of African origin (Adams 2007: 533). In addition, there is a probable Berber cognate **(ā)k(V̌)tūn* 'rue'. The fact that the Sardinian and Berber meanings are closer to each other than to the Latin meaning, is evidence against a Latin loan into Berber. Even if we assume that a similar form existed in Punic, there is no Semitic etymology that could account for these forms. The same goes for Pun. *zibbir* 'rosemary' (6), which has a potential Berber cognate **ā-zV̌βīr* 'id.'. The "African" form σικκιριά (2) corresponds formally to Bb. **tā-sǐkrā* '*Echinops spinosissimus*'. The semantic discrepancy between 'dill' and a species of thistle is not easy to explain, although Tsh. <*taskrā*> is defined as 'dill' in some anonymous medieval Arabic-Berber word lists (van den Boogert 1998: 178, 184).

10.1.4 Conclusion

It is beyond doubt that Phoenician was spoken as a native language in Sardinia at the time of the Roman conquest. This is evidenced by the fact that Phoenician texts continued to be produced centuries after Carthaginian influence on the island had ceased (Paulis 1990; Adams 2003: 209). The distribution of the punicisms in Sardinia suggests that Punic linguistic influence extended to most of Sardinia, although it was most pronounced in the south (§ 10.1.3.1). However, only a few of the accepted punicisms in Sardinian can satisfactorily be etymologized as Semitic, and several of the words that do not have a good Semitic etymology, are attested in Berber. There are two scenarios that could explain this:

1. The Punic language borrowed words from native North-African languages ancestral or related to Berber, after which these words entered Sardinia along with Carthaginian colonists.
2. Speakers of native North-African languages were involved in Carthaginian colonialism on Sardinia, and maintained their native language at least up to the Roman conquest.

In favor of the first scenario is the fact that the Sardinian words with both Punic and Berber comparanda are formally more similar to the compared Punic forms,[284] than to the reconstructed Proto-Berber forms. Admittedly, this could be an artefact of Proto-Berber not being identical to the North-African language that Punic was in contact with and rather a sister or descendant language *vel sim.* Unless positive evidence in favor of the second scenario can be adduced from the field of Berber linguistics, it is more parsimonious to assume the first scenario. The fact that six out of the seven accepted punicisms in Sardinia refer to specific plant species, may indicate that the Punic population of Sardinia had specific cultural or practical (e.g. culinary) uses for these plants, although this is a matter of speculation (cf. Paulis 1992: 133).

On the basis of the available evidence, it is difficult to conclude much about the linguistic contacts between Sardinian Punic and the native pre-Roman Sardinian languages. The overlap in geographical distributions of words of Punic origin and of other pre-Roman words suggests that the languages coexisted at least until Sardinia's romanization.

[284] And in the case of Srd. **katòne* to the (African) Latin form *catone*.

10.2 Berber

A "Libyan" or Berber substrate in Sardinia has often been proposed (see § 1.3.3.2). As seen in § 10.1.4, a number of accepted punicisms in Sardinian are also found in Berber, likely as the result of the intense and prolonged language contact between native North-African languages and Carthaginian Punic. The question is whether an earlier presence of Berber (or some affiliated North-African language) should be posited, independently from the Punic colonial context.

10.2.1 The lexical evidence

The following is an overview of the Sardinian words that have been proposed to have a cognate in Berber, and that would thus potentially point to a North-African (or "Libyan") presence in Sardinia at an earlier stage than the Punic colonization. The forms that are to be compared are printed in bold. For each comparison, an evaluation is given for the formal and semantic correspondence between the compared forms:

1. § 3.2.2: Srd. *thinnigas*, *tsinníɣa*, *tinnía* etc. < ****ϑinníga***; *tsónni*, *tsónnɨa* etc. < ****ϑónni*** 'esparto grass' (DES II: 547-548); cf. Nef. *tsənnít* 'esparto grass', *tsunít* 'basket woven from rushes', Kab. *isni* 'big basket made from esparto grass', Tuar. *tesănit* 'big basket with two handles' (Bertoldi 1950: 16; Ritter 2009: 434; Múrcia Sànchez 2011: 468; Argiolas 2020: 41) < ****sănV̌y/*sūnV̌y/*sănī/*sūnī.***[285]

   Formal match: **good**, at least for Srd. **ϑónni*
   Semantic match: **good**
   Cognacy: **compelling**

2. § 3.2.3: Srd. *sèssini*, *sèssene* etc. < ****sèssen(e)*** 'galingale' (DES II: 412; Paulis 1992: 449). Four Berber comparisons have been proposed for this word:
   a. cf. Kb. ***azəzzu*** 'broom' (Hubschmid 1953: 27; cf. Dallet 1982: 927).

   Formal match: **problematic**
   Semantic match: **problematic**
   Cognacy: **unlikely**

[285] Maarten Kossmann, p.c.

b. cf. Beni Snous ***tiyzzi*** 'dry esparto grass' (DES II: 412; cf. Destaing 1914: 10).[286]

Formal match: **problematic**
Semantic match: **good**
Cognacy: **unlikely**

c. Beni Snouss ***ẓẓiw*** 'strand' (DES II: 412; cf. Destaing 1914: 50).[287]

Formal match: **problematic**
Semantic match: **problematic**
Cognacy: **unlikely**

d. cf. Kb. ***ẓẓəmẓəm*** 'to be slim', ***aẓəmẓum*** 'twig', Tmz. ***azmu*** 'very fine rush used to make mats', etc. (Argiolas 2020: 40–41; cf. Dallet 1982: 945; Taïfi 1991: 807).

Formal match: **possible**, for *ẓẓemẓem* and *aẓəmẓum*
Semantic match: **problematic**
Cognacy: **unlikely**

3. § 3.3.22: Srd. *éni* 'yew' < ****éni**/***áCini*** (DES I: 489); cf. Ihahan ***taga***, Ouarzazate ***tiqqi***, Igliwa ***tiqi*** etc. 'juniper' (Hubschmid 1953: 25–26; Alessio 1957; cf. Laoust 1920: 490).

Formal match: **problematic**
Semantic match: **problematic**
Cognacy: **unlikely**

4. § 3.4.7: Srd. *mátta* 'tree, plant, shrub' etc. < ****mátta*** (DES II: 87); cf. Snh., Rif. ***tamaṭṭa*** 'stack of sheaves that are to be threshed' (Hubschmid 1953: 33; cf. Renisio 1932: 385; Serhoual 2002: 329); many other Romance comparisons in Hubschmid (1953: 33–35). Derived from Lat. *matta* 'mat' by REW (5424) and DCEC (III: 874-875).

[286] In Destaing (1914) and DES as *ϑiizzi*.
[287] In Destaing (1914) as *ẓẓẹu*, and in DES as *ž'zeu*.

Formal match: **good**
Semantic match: **possible**
Cognacy: **unlikely**. The Sardinian forms can simply be inherited from Latin, and the Berber forms are likely from Moroccan Ar. *məṭṭa*.

5. § 4.4.2: Srd. *ϑáiri* 'bird of prey' < ***ϑá(C)iri***; cf. Chenoua ***isɣi*** 'id.' (Paulis 2008: 24; cf. Laoust 1912: 136). Also compared is Bq. *sai, sari* 'vulture' (Paulis 2008: 24; cf. Löpelmann 1968: 1138).

   Formal match: **problematic**
   Semantic match: **good**
   Cognacy: **unlikely**

6. § 4.4.8: Srd. *tiváni* 'raven' < ****tifáni*** (DES II: 492); cf. Rif. *asiwan, tasiwant* 'bird of prey' (DES II: 492; cf. Serhoual 2002: 550). Paulis (2017: 111) derives the Sardinian forms from the Greek personal name Θεοφάνης.

   Formal match: **doubtful**
   Semantic match: **possible**
   Cognacy: **unlikely**. Paulis' (2017: 111) etymology is more plausible.

7. § 6.1.6: Srd. *bòḍḍa* 'old sheep' < ****bòlla**/***vòlla*** (DES I: 215). There are two possible Berber comparanda for this word:[288]
   a. cf. Bb. *ullī, wəllī* 'small livestock (i.e. sheep and goats)' < ****wV̌llī**/***wV̌llē***.[289]

   Formal match: **possible**
   Semantic match: **good**
   Cognacy: **possible**

   b. cf. Ghadames *taβali*, Mali Tuar. *tehăle*, Nef. *tili* etc. 'sheep' < ****tē-ḇălē*** (Kossmann 1999: 97).[290]

[288] The alternative Calabrese comparison to this word is discussed in § 11.1.
[289] Maarten Kossmann, p.c.
[290] Maarten Kossmann, p.c.

Formal match: **possible**
Semantic match: **good**
Cognacy: **possible**

8. § 6.2.5: Srd. *čèrda* 'mat woven from straw or rushes that serves to spread out fruits on and that is also put in carts, so that the transported straw or fruits cannot fall out' < ***kètra*** (DES I: 446-447). Two Berber comparisons have been proposed for this word:
   a. cf. Bb. *ažerʕil* 'mat (made of esparto grass or rushes)' (Bertoldi 1947b: 201; Hubschmid 1953: 22–23; cf. Haddadou 2007: 290) < ****gʷǍrtil***.[291] Further comparisons include Lat. *caetra*, *cētra* 'type of Spanish and/or African shield' (Walde & Hofmann 1938: 135), Ast. *sardu*, *sarda*, *zarda* 'wickerwork placed over the kitchen stove, to shelter the kitchen and in which hazelnuts are put to roast'[292] (Wagner 1943: 196; but cf. Hubschmid 1953: 22–23; DCEC VI: 102a).

   Formal match: **problematic**
   Semantic match: **good**
   Cognacy: **unlikely**

   b. cf. Tmz. *ɣərd* (Ayt Izdeg) 'to lie down on the floor' (cf. Taïfi 1991: 202), Kb. *əɣṛuṛəd* 'to sleep without being able to wake up' (cf. Dallet 1982: 625) < ****ɣrd*** (Argiolas 2012: 2737–2741). Further comparisons include Lat. *cherda* 'a plant species (probably eryngo, cf. André 1985: 62)', Lat. *caetra*, *cētra* 'type of Spanish and/or African shield' (Argiolas 2012: 2737–2741).

   Formal match: **problematic**
   Semantic match: **problematic**
   Cognacy: **unlikely**. Camp. *čèrda* 'mat' is most easily interpreted as a regular development of Lat. *caetra* 'type of shield', whatever its origin.[293]

[291] Maarten Kossmann, p.c.
[292] "Tejido de mimbres que se coloca sobre el llar, para abrigar la cocina y donde se ponen las avellanas para turrar".
[293] Cf. Múrcia Sànchez (2011: 423).

9. § 7.1.2: Srd. *ʒúḍḍa, tsúḍḍa* etc. 'bristle hair (of pigs and wild boars)' < ***ʒúlla*** (DES II: 553); cf. Snh. ***əzzuṭ*** 'pubic and armpit hairs' (DES II: 553; cf. Renisio 1932: 320).

   Formal match: **problematic**
   Semantic match: **possible**
   Cognacy: **unlikely**

10.2.2 Discussion

Only two of the six Sardinian words with a proposed Berber origin can be accepted on formal and semantic grounds: Srd. **ʒónni* 'esparto grass' (1), Bb. **sănV̆y/*sūnV̆y/*sănī/*sūnī* 'id.'; Srd. **bòlla/*vòlla* 'old sheep' (7), Bb. **wV̆llī/*wV̆llē* 'small lifestock (i.e. sheep and goats)' or Bb. **tē-ḇălē* 'sheep'. The others must be rejected because the compared forms are too dissimilar (2, 3, 5, 8a, 8b, 9), or because their meanings are too different (2a, 2c, 2d, 8b), or simply because the Sardinian form has an alternative, more convincing, etymology (4, 6, 8a, 8b). The following table summarizes the formal and semantic criteria of the comparisons, and whether cognacy is likely or not.

	Srd.	Comp.	Form.	Semant.	Cogn.
1	**ʒinníga, *ʒónni* 'esparto grass'	Bb. **sănV̆y/*sūnV̆y/*sănī/*sūnī* '(basket made of) esparto grass'	✓	✓	✓
2a	**sèssen(e)* 'galingale'	Kb. *azəzzu* 'broom'	×	×	×
2b		Beni Snous *tiyzzi* 'dry esparto grass'	×	✓	×
2c		Beni Snous *ẓẓiw* 'strand'	×	×	×
2d		Bb. **ẓmy* 'to sew'	×	×	×
3	**éni/*áCini* 'yew'	Ihahan *taga*, Ouarzazate *tiqqi*, Igliwa *tiqi* etc. 'juniper'	×	×	×
4	**mátta* 'tree, plant shrub'	Snh., Rif. *tamaṭṭa* 'stack of sheaves that are to be threshed'	✓	?	×
5	**ʒá(C)iri* 'bird of prey'	Chenoua *isyi* 'bird of prey'	×	✓	×
6	**tifáni* 'raven'	Rif. *asiwan, tasiwant* 'bird of prey'	?	?	×
7	**bòlla/*vòlla* 'old sheep'	Bb. **wVllī/*wVllē* 'small lifestock'	✓	✓	✓ (?)
		Bb. **tē-bălē* 'sheep'	✓	✓	✓ (?)

	Srd.	Comp.	Form.	Semant.	Cogn.
8a	**kètra* 'type of mat'	Bb. **gʲVrtil* 'mat made of esparto grass'	×	✓	×
8b		Tmz. *ɣərd* 'to spread out on the floor', Kb. *əɣruṛəd* 'to sleep without being able to wake up' < **ɣrd*	×	×	×
9	**ϑúlla* 'bristle hair (of pigs and boars')	Snh. *əzzuṭ* 'pubic and armpit hairs'	×	?	×

Table 10.2.1: Proposed Berberisms in Sardinian

10.2.3 Conclusion

On the basis of only two acceptable Berber-Sardinian comparisons that are not attested in Punic, it is unattractive to posit a pre-Roman presence of Berber in Sardinia, independently from the Punic colonial context (as was already suspected by Paulis 2008: 42). The two stronger cases (1, 7) are rather to be seen as part of the Punic loans in Sardinia, among which some potential Berber etyma have been identified (cf. § 10.1.3.3). Semantically, at least the word for 'esparto grass' (1) fits in well with the accepted punicisms (§ 10.1.4), which largely refer to culturally or economically relevant plants. The two words with accepted Berber cognates show a southern geographical distribution (Figure 10.2.1), like the majority of the accepted punicisms (Figure 10.1.1).

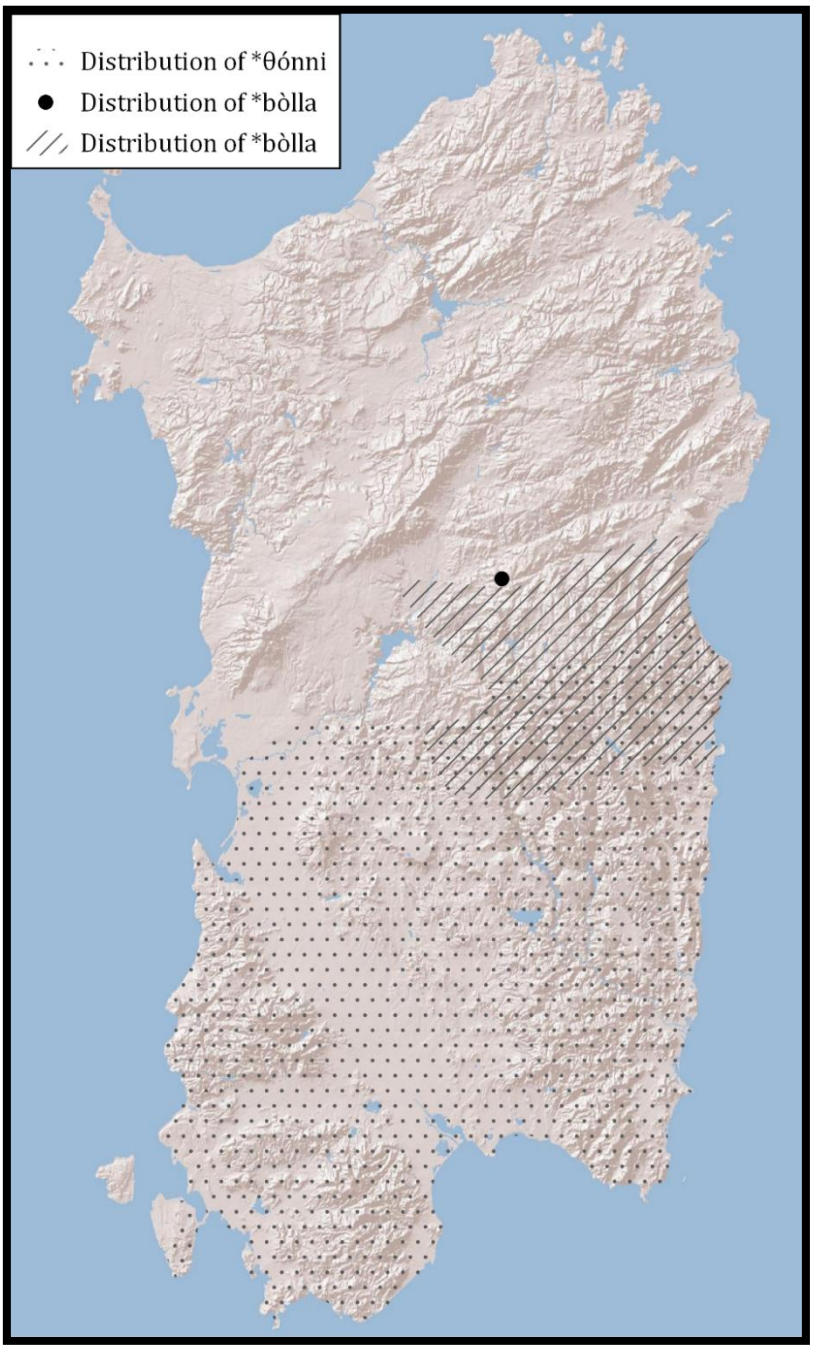

Figure 10.2.1: Distribution of Berberisms in Sardinian

10.3 Basque

A connection between Basque and pre-Roman Sardinian has been proposed by various scholars, often as part of a wider "Mediterranean substrate" (§ 1.3.2). Blasco Ferrer (2010a; 2010b; 2011; 2017) argues that the language of pre-Roman Sardinia was a sister language of Basque (but see § 1.3.3.3).

10.3.1 The lexical evidence

The following is an overview of the pre-Roman words that have been proposed to be shared by Sardinian and Basque, by various authors. In each case, the form that is to be compared is printed in bold (both for Sardinian and Basque), and a brief overview is given of any comparanda in other languages, and of any further relevant etymological considerations. For each comparison, an evaluation is given for the formal correspondence between the Sardinian and Basque comparanda, as well as for their semantic correspondence. These are used to evaluate whether cognacy between the Sardinian and Basque forms in question is at all possible and, if so, whether it is plausible.

1. § 3.1.15: Srd. *tθáṇḍa, aθáṇḍa, ánna* etc. 'poppy' < ***(ϑ-)anda** (DES I: 151); cf. Bq. ***andere*** 'lady' (Bertoldi apud Wagner 1932: 237). Further comparisons include: Egyptian ναντί 'poppy'; Lang. *ãndèr, õndèr* (dep. Lot) 'id.' (Bertoldi 1937a: 151 fn. 1), Gr. ἄνθος 'flower' (Pittau *apud* Paulis 1992: 319). Cf. also Lakarra et al. (2019: 48–50).

   Formal match: **possible**
   Semantic match: **problematic**
   Cognacy: **unlikely**. Only the Lengadocian comparison is formally and semantically acceptable.

2. § 3.3.7: Srd. *golóstḭu, golóstri, olósti, golósti?e* etc. 'holly' < ***golósti** (DES I: 584); cf. Bq. *gorosti* 'holly' < ***golosti*** (Bertoldi 1929: 261 fn. 3). Further comparisons include: Gr. κήλαστρος holly' (Hofmann 1949: 142); Navarr. *colostia, golostia* etc. (Hubschmid 1953: 29 fn. 1); Lang. *agaloussé* 'id.' < **agalústo-* (Jud 1931: 439; cf. FEW XXI: 106); Bret. *kelenn*, OIr. *cuilenn* 'id.' < PC **kolinno-* < **kolis-n-* (Hubschmid 1953: 29; van Sluis, Jørgensen & Kroonen 2023: 200); OHG *hulis*, Du. *hulst* 'id.' < PGm. **hulisa* < **kulis-* (Kroonen 2013: 253). The Mediterranean comparanda (i.e. Basque, Greek and Lengadocian) offer a match that is strong both formally and semantically (cf. § 11.1), with an exact correspondence between Sardini-

an and Basque. The Celtic and Germanic forms are formally further removed.

Formal match: **good**
Semantic match: **good**
Cognacy: **compelling**

3. § 3.3.8: Srd. *aúrri* 'hop-hornbeam' (DES I: 153) < ***a**(**C**)**urri**/***aruli**; cf. Bq. ***aurri*** 'plant with long roots and wide leaves' (Wagner 1932: 231). Cf. also Michelena (1958: 405–406).

   Formal match: **uncertain**
   Semantic match: **uncertain**
   Cognacy: **uncertain**

4. § 3.3.9: Srd. *kòstike*, *kòstiɣe*, *kòsti* 'maple' < ***kòsti*** (DES I: 392-93); cf. Bq. *astigar*, *gastigar*, *gaztigar* 'maple' < ***(g)azt**- + *ihar* 'maple' (Bertoldi 1931: 131). Further comparisons include: Tusc. *còstolo* 'ash tree'; Gr. Hes. ἄκαστος (Hesychius Alexandrinus 2018: 109); OProv. *agast* 'maple' (Alessio 1949: 132; FEW XXIV: 276). All of these are formally and semantically attractive comparanda to the Sardinian and Basque forms. Cf. also Hubschmid (1953: 82) and Michelena (1953: 48; 1990: 253 fn. 49).

   Formal match: **possible**
   Semantic match: **good**
   Cognacy: **possible**: the Tuscan, Greek and Provençal comparisons are good, but the internal structure of the Basque form is not well-understood.

5. § 3.3.20: Srd. *giḍḍòstre*, *giḍḍóstru*, [*sa*] *ʔiḍḍòstra*, [*su*] *iḍḍòstro* 'tree heather' < ***gill**(-)**óstr**- (DES I: 578); cf. Bq. *gillar*, *gilharre*, *ilharri* 'heather' < ***gil-ar(r)**- (Hubschmid 1953: 29; 1958: 214) or ***inar*** (Trask 1997: 306; Santano Moreno 2000: 170). Hubschmid (1958: 214) doubts his own earlier proposal. Note further that the evidence for *-*ósta*(*r*)- as a suffix in Sardinian is limited (§ 9.2.9).

Formal match: **problematic**
Semantic match: **good**
Cognacy: **unlikely**, because of the formal discrepancy.

6. § 3.3.22: Srd. *éni* 'yew' < ***éni*** (< ***á(g)ini*** [?]) (DES I: 489); cf. Bq. ***(h)agin*** 'yew' (Bertoldi apud Wagner 1941a: 45).

   Formal match: **possible**
   Semantic match: **good**
   Cognacy: **possible**, but the formal base of comparison is slim.

7. § 3.4.2: Srd. *kárva* 'tree branch' < ***kárva/kárba*** (DES I: 311); cf. Bq. *karbasta*, *garbazta* etc. 'stick with branches', *karbaza* 'stem, trunk' < ***karba-*** (Wagner 1932: 231; Michelena 1990: 251), or derived from Bq. *arba* < ***abar*** (Lakarra et al. 2019: 3–8). Other comparisons include: Ast. *garbu* 'small firewood', Gasc. *garbalhe* 'pine needles', Prov. *garbo* 'hollow tree trunk' etc. (Wagner 1932: 231; DCEC I: 856); Lat. *carbō*, *-ōnis* 'piece of charcoal (de Vaan 2008: 91–92).

   Formal match: **good**
   Semantic match: **good**
   Cognacy: **possible**, but the compared forms may not be old in Basque (Lakarra et al. 2019: 3–8).

8. § 4.4.2: Srd. *ϑáiri* 'bird(s) of prey; harmful bird(s)' < ***ϑá(C)iri*** (Paulis 2008: 24); cf. Bq. *sai*, *sari* 'vulture' (Löpelmann 1968: 1138; EH: 558) < ***sai*** (Trask 2008: 389). Also compared is Chenoua Bb. *isɣi* 'bird of prey' (Paulis 2008: 24; cf. Laoust 1912: 136).

   Formal match: **problematic**
   Semantic match: **good**
   Cognacy: **unlikely**

9. § 4.5.3: OSrd. *gjagar, jagaru* Srd. *ǧáɣaru* 'hunting dog' < ***ɟákaru*** (DES I: 602); cf. Bq. ***txakur*** 'dog' (Schuchardt 1905: 227; Wagner 1932: 232). Further comparisons include: Cors. *ghjácaru* etc. 'shepherd dog' (Wagner 1932: 232); Byz. Gr. ζαγάριον, MGr. ζαγάρι, Turkish *zağar* 'hunting dog', Georg. ძაღლი *dzaghli* 'dog' (Schuchardt 1905: 227). Schuchardt (1905)

proposes the Sardinian forms to be a Byzantine loan, but Wagner (1932) rejects this because of the Corsican and Basque comparanda. Nevertheless, in light of the well-established presence of Byzantine loans in Sardinian (Paulis 1983; 2017), Schuchardt's proposal remains attractive.

Formal match: **possible**
Semantic match: **good**
Cognacy: **possible**, but probably as the result of a more recent loan origin of this word in both Sardinian and Basque.

10. § 4.5.5: OSrd. *bita*, Srd. *bètte*, *bítti* etc. 'fawn, lamb, small animal' < ***bett-**, ***bitt-** (DES I: 211-12); cf. Bq. ***pitika*** 'kid (young goat)' (Wagner 1932: 233). Further comparisons include: Cat. *bita*, *bitè*, Gasc. *pito*, For. *bette*, *bitte* 'goat' < ***bitt-**, ***bett-**, ***pitt-** (Hubschmid 1953: 70; Hubschmid 1955: 189–190; cf. FEW XXII/I:288). Wagner (DES I: 212) and Michelena (1959: 525) prefer to explain these forms as independent sound-symbolic formations. This is supported by the fact that these forms are also well-attested as animals calls (FEW XXII/I: 288).

Formal match: **possible**, although there are various incongruencies within and between the different languages.
Semantic match: **good**
Cognacy: **uncertain**; there is a possibility that these forms are independent formations originating in animal calls.

11. § 5.1.4: Srd. *líttu* 'dense forest, forested terrain' < ***líttu**/***líCtu**/***líddu**/***líCdu*** (DES II: 33); cf. Bq. *lito*, *lütho* 'deep place' (Wagner 1932: 242) < ***lu**(**r**)* 'earth' + ***it**(**h**)**o*** 'drowned, choked' (Michelena 1990: 81 fn. 17).

Formal match: **problematic**
Semantic match: **problematic**
Cognacy: **unlikely**

12. § 5.1.6: Srd. *mòɣoro*, *mòɣuru* 'hillock, low hill' < ***mòkor**- (DES II: 122; Paulis 1987: 514); cf. Bq. *mokor* 'tree trunk, crust of bread, lump' < ***m-okor*** (Wagner 1932: 227; Hubschmid 1953: 49). Michelena (1990: 272) analyses Bq. *mokor* as a recently *m*-prefixed variant of *sokor* 'lump' and *okor* 'slice'. Wagner (1932: 227) and Hubschmid (1953: 49) also compare Bq.

mukurru 'summit', but this is likely metathesized from a loan from Lat. ***cumulus*** 'heap, pile' (Michelena 1990: 331).[294] Finally, Trask (2008: 276) mentions Bq. ***mokoru*** (Gipuzkoa) 'height, summit'. Other comparisons include: Old Cat. *mocorró*, Cat. *mogró* etc. 'nipple' < Lat. *mūcrō*, *-ōnis* 'sharp point' (Hubschmid 1953: 49; DCEC IV: 184); Prov. *maguelo*, Lang. *maghêlo*, *maghiêlo* etc. 'steep hill, steep place' < Pre-Rom. **maguella* (FEW XXI: 15b); Alb. *magulë* 'small pile, heap', Rom. *măgură* 'mountain, hill', MGr. μαγούλα, PSl. **mogyla* 'burial mound' (Wagner 1932: 228; Ciorănescu 1958: *măgură*; Orel 1998: 109).

Formal match: **possible**, but only for Bq. *mokoru* 'height, summit', as *mokor* 'lump' and *mukurru* 'height, summit' have alternative etymologies.
Semantic match: **good**
Cognacy: **possible**, for Bq. *mokoru* 'height, summit'. Ibero-Romance forms like Cat. *mugró* 'nipple' are inherited and thus unrelated. The forms from Southern France and the Balkans may still be compared (cf. § 11.1).

13. § 5.1.9: OSrd. *bega*, Srd. *bèɣa* 'marshy valley' < ****bèka*** < *(*i*)*baika* (DES I: 191-92; Wagner 1997: 254–255); cf. Bq. *ibai* 'river'; Old Leon. *vajka*, Sp. *vega* etc. 'meadow, fertile plain' < Pre-Rom. ****ibai-ka*** (Bertoldi 1937: 148; DCEC V: 753–754). There is a possibility, however, that the earliest attestation of OSrd. *bega* does not predate Ibero-Romance linguistic influence on Sardinia.

    Formal match: **problematic**; an unattested derivation of the Basque form must be assumed, as well irregular vocalic apheresis (Trask 1997: 420–421).
    Semantic match: **good**
    Cognacy: **unlikely**. It may be an early loan from Spanish, through Catalan.

14. § 5.1.17: Srd. *trókku* 'ravine, gorge' < ****tórku***/****trókku*** (DES II: 522); cf. Bq. *troka* 'ravine' < ****torka*** (Hubschmid 1953: 42; DCEC V: 521b–522). Other comparisons include: dial. Sp. *torco*, *torca* 'chasm; unstable terrain; pot-

[294] Another treatment of the same loanword is found in Bizkaian *bonburu*, *gonburu* 'summit' (Michelena 1990: 331).

hole etc.' (Hubschmid 1953: 42). The Basque and Ibero-Romance forms are certainly related, but Corominas (DCEC V: 521b–522) considers a Latin etymology in *torques* 'collar', and argues that the Basque word is of recent origin.

Formal match: **good**
Semantic match: **good**
Cognacy: **possible**, but rather to the Ibero-Romance forms. An inherited origin in Sardinian is not excluded.

15. § 5.1.24: Srd. *óspile* 'enclosure; gorge, natural cavern; secluded place' < ***óspile*** (DES II: 198); cf. Bq. *ospel* 'shady place' < ***ospel*** (Wagner 1932: 227). Trask (2008: 320) suggests that Bq. *ospel* contains **bel* 'dark', but has no explanation for the first element. Cf. also Michelena (1990: 272–273) on variants of the Basque word.

    Formal match: **good**
    Semantic match: **good**
    Cognacy: **compelling**

16. § 5.2.4: Srd. *moreðína*, *moretína* etc. 'pile of rocks, dry stone wall' < ***moretína*** << ***mora***- (DES II: 127-28); cf. Bq. *more*(*a*), *mora* 'cairn, pile of rocks; dolmen' < ***mora*** (DCEC IV: 146). Other comparisons include: Cors. *mora*, *mola* 'pile of rocks or wood', Tusc. *mora* 'pile of rocks', Laz. *mora* (Castelmadama) 'boulder, very big rock' (DES II: 128); Sp. *morena*, Pg. *moreia* ´pile of plant material', Frpr. *morêna* (Savoy) 'glacier moraine' (DCEC IV: 146). Cf. Hubschmid (1943: 150–151), Michelena (1990: 23–25), FEW (VI/III: 276b).

    Formal match: **good**
    Semantic match: **good**
    Cognacy: **compelling**

17. § 6.2.1: Srd. [*pan'*] *ispéli* 'bread made of acorn flour and clay' < ***ispéli*** (Wagner 1921: 60); cf. Bq. *ezpel* 'box tree', through a presumed shared original meaning 'pestle' (Paulis 2011; cf. Putzu 2019: 4).[295] However, Bq.

[295] I have unfortunately not been able to consult Paulis (2011) myself.

ezpel likely goes back to ***ez***-, a morph found in Basque tree names, and *-***bel** 'dark' (Trask 1997: 256; 2008: 188).

Formal match: **problematic**
Semantic match: **problematic**
Cognacy: **unlikely**

18. § 7.1.3: Srd. *tuttúrru* 'cheek', *tutturròne* 'mumps' < ***tuttúrru** (DES II: 537); cf. Bq. *tutur* 'crest, bump on the head', *tuturru* 'point, summit' etc. (Hubschmid 1953: 75–76) < expressive ***Cu(n)Cur**; cf. Bq. *kunkur*, *konkor* 'bump on the head', *kukur* 'bird's crest' and *kokor*, *kukutz* 'peak, summit' (Trask 1995: 70; 2008: 344).

    Formal match: **problematic**; the Basque forms do not appear to be old.
    Semantic match: **problematic**; 'cheeks' are not straightforwardly interpreted as 'crests' or 'summits' of the head.
    Cognacy: **unlikely**

19. § 7.2.4: Srd. *barraśòne*, *bariśòne*, *barraśólu* etc. 'tangle/bundle of wood (generally to block passages)' < ***barrVs**- (DES I: 181-82); cf. Bq. ***barra*** 'bar' (Rohlfs 1927b: 402; DCEC I: 510b, DECLC I: 645b). Other comparisons include: Pg. *barda* 'fence made of intertwined branches and brambles', Cat. *barda* 'portable barrier to lock up lifestock', *bardissa* 'thicket; hedge' (DCEC I: 510; DECLC I: 644–645); OProv. *barta* 'bush, brushwood' etc. (FEW I: 262). However, Bq. *barra* may be of Romance origin (Michelena 1990: 269).

    Formal match: **possible**
    Semantic match: **good**
    Cognacy: **possible**, but the Romance comparisons are semantically stronger, and the Basque form may originally be from Romance.

10.3.2 Discussion

From the overview of proposed cognates between Sardinian and Basque, it is clear that not all comparisons can be maintained. Some must be rejected on formal grounds (5, 8, 11, 13, 17, 18), while for others the meanings of the compared forms are too divergent or uncertain (1, 3, 11, 17, 18). In some other cases there does seem to be a reliable connection, but with the possibility that it is of recent origin either in Sardinian (9, 15), or in Basque (7, 9, 11, 1). The following table

summarizes the formal and semantic criteria of the comparisons, and whether cognacy is likely or not.

	Srd.	Bq.	Form.	Semant.	Cogn.
1	**(ϑ-)ánda* 'poppy'	*andere* 'lady'	✓	×	×
2	**golósti* 'holly'	*gorosti* 'holly'	✓	✓	✓
3	**a(C)urri*/**aruli* 'hop-hornbeam'	*aurri* 'unspecified plant'	?	?	?
4	**kòsti* 'maple'	**(g)azt-* 'maple (?)'	✓	✓ (?)	✓ (?)
5	**gillóstr-* 'tree heather'	**gilar(r)-*/**inar-* 'heather'	×	✓	×
6	*éni* < **éni*/**á(C)ini* 'yew'	*(h)agin* 'yew'	✓ (?)	✓	✓ (?)
7	**kárva* 'branch'	**(k)arba-* 'branch, trunk etc.'	?	✓	✓ (?)
8	**ϑá(C)ir(i)* 'bird of prey'	**sai*	×	✓	×
9	**ʒákaru* 'hunting dog'	*txakur* 'dog'	✓	✓	✓ (?)
10	**bett-*, **bitt-* 'fawn, lamb'	*pitika* 'kid'	✓ (?)	✓	?
11	**líCtu* 'dense forest'	**lu(r)-it(h)o* 'deep place'	×	×	×
12	**mòkor-* 'hillock, low hill'	*mokoru* 'height, summit'	✓ (?)	✓	✓ (?)
13	**béka* 'marshy valley'	*ibai* 'river'	×	✓ (?)	×
14	**tórku* 'ravine'	**torka* 'ravine'	✓ (?)	✓	✓ (?)
15	**óspile* 'secluded place'	**ospel* 'shady place'	✓	✓	✓
16	**moretína* << **móra-* 'pile of rocks'	**mora* 'pile of rocks, dolmen'	✓	✓	✓
17	**ispéli* 'bread made of acorns and clay'	**ez-bel* 'box tree'	×	×	×
18	**tuttúrru* 'cheek'	**Cu(n)Cur* 'point, summit; crest, bump'	×	×	×
19	**barrVs-* 'tangle of wood'	*barra* 'bar'	✓ (?)	✓	✓ (?)

Table 10.3.1: Proposed Sardinian-Basque comparisons

For ten of the nineteen discussed proposals for "Sardo-Basque" cognates, the respective formal reconstructions as well as semantics are similar enough to consider a shared origin. However, for some there are indications for a more recent origin of the compared forms in either Basque or Sardinian. The forms in question are:

- Bq. *karbaza*, *karbasta* 'etc.' (7) (<< **arba* < **abar*, according to Lakarra et al. 2019: 3–8; cf. also the various Ibero- and Gallo-Romance forms of similar form and meaning)

- Srd. *ğáyaru* 'hunting dog', Bq. *txakur* 'dog' (9) (cf. Byz. Gr. ζαγάρος; Schuchardt 1905: 227)
- Bq. *pitika* 'kid' (10) (maybe of sound-symbolic origin; Michelena 1959: 525; DES I: 212)
- Bq. *troka* < **torka* 'ravine' (14) (<< dial. Sp. *torco*, *torca* 'chasm, pit etc.'; DCEC V: 551b–522).

Because of the possibility of a more recent origin of these forms in either Basque or Sardinian, they cannot be used as evidence in favor of a prehistoric linguistic connection between Basque and the pre-Roman language(s) of Sardinia.

There are two sets of comparanda in Basque and Sardinian that potentially match both formally and semantically, but for which some caution is needed either on the formal or on the semantic side. One is the connection between Srd. *éni* 'yew' and Bq. (*h*)*agin* 'id.' While Srd. *éni* could in principle go back to **agin*, this is only one of the options. Other possible reconstructions include **éni*, **égin*(*i*), **ą́i̯n*(*i*). Because of the multitude of possible Sardinian reconstructions, the formal comparison to the Basque word loses its strength, despite the excellent semantic match. Similarly, the comparison between Srd. *kòsti* 'maple' and Bq. *gaztigar* (*g*)*astigar* 'id.' is based on the idea that the Basque forms are a compound of **gazt-* and *ihar* 'maple'. However, **gazt-* is not attested on its own (Michelena 1953: 48). Therefore, while there is formal similarity between it and Srd. *kòsti* 'maple', the meaning of Bq. **gazt-* cannot reliably be verified.

As a result, we are left with four more or less convincing sets of words that show comparable form and meaning in Sardinian and Basque, and for which a more recent (loan) origin is not evident. These words are:

- Srd. **golósti* 'holly' — Bq. *gorosti* 'id.'
- Srd. **mòkor-* 'hillock' — Bq. *mokoru* 'summit'
- Srd. **óspile* 'secluded place etc.' — Bq. *ospel* 'shady place'
- Srd. **moretína* << **móra* 'pile of rocks' — Bq. *mora* 'id.; dolmen'

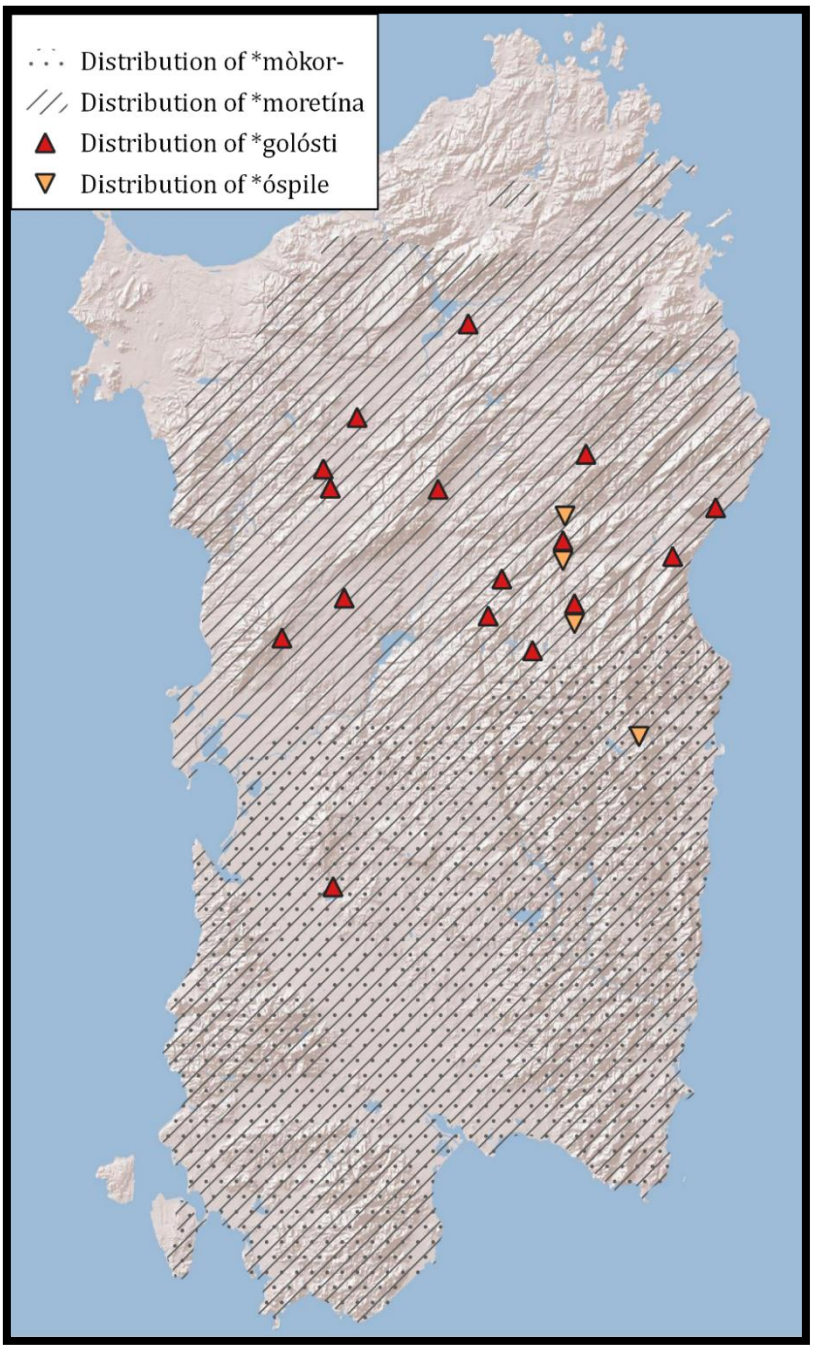

Figure 10.3.1: Distribution of possible Sardinian-Basque cognates

The small number of four Sardo-Basque comparisons that do not entail immediate formal, semantic or etymological difficulties, does not favor the hypothesis that the language(s) spoken in Sardinia prior to the arrival of the Phoenicians and Romans was a relative of Basque. Yet, it is unattractive to explain all of these forms as mere chance resemblances, especially in view of the perfect correspondences of phonologically complex forms, such as Srd. *golóstḭu* and Bq. *gorosti* 'holly'.

The low number of compelling Sardo-Basque comparisons is also unlikely to be the result of a geographically restricted Vasconic presence in Sardinia, as most of these words have a wide distribution across the Sardinian dialects (Figure 10.3.1). An alternative explanation is that these words did not originate in a language related to Basque, but that both Basque and Sardinian acquired them as loans from a third language or language family. Closer examination of the remaining four comparisons reveals that they are not restricted to Sardinian and Basque, with the exception of Srd. *óspile* 'secluded place', Bq. *ospel* 'shady place'. The word for 'holly' has comparanda in Occitan and Greek; the word for 'hillock' has been compared to forms in Occitan, Greek, Albanian and Romanian; and the word denoting 'pile of rocks' has potential cognates in Ibero-Romance, Gallo-Romance and Italo-Romance. The same is true for the less certain cases listed in Table 10.3.1: e.g. for Srd. *kòsti* 'maple', Bq. (*g*)*astigar* 'id.', cf. OProv. *agast*, Gr. Hes. ἄκαστος 'id.'; for Srd. **karva* 'branch', Bq. *(*k*)*arba-* 'branch, trunk', cf. Ast. *garba* 'small firewood', Lang. *garbasso* 'young oak', Gr. κάρφος 'arid stalk, twig, halm, hay'; for Srd. **bar*(*r*)*Vs-* 'tangle/bundle of wood used to block passages, Bq. *barra* 'bar', cf. Cat. *barda* 'portable barrier to lock up lifestock', *bardissa* 'thicket; hedge', OProv. *barta* 'brushwood'.

An additional indication that the Sardo-Basque comparisons need not originate in a language related to Basque, is the fact that the labial nasal **m* is not reconstructed for the Proto-Basque phonological system (Michelena 1990: 267). The modern occurrences of *m* in Basque are explained as the result of loans or by nasalization of original **b*, by assimilation to the dental nasal *n* (Michelena 1990: 267–271; Trask 1995: 70). Therefore, Bq. *mokoru* 'height, summit' and *mora* 'cairn, pile of rocks; dolmen' are not easily reconstructible to Proto-Basque. Michelena's reconstruction of the Proto-Basque phonological system, lacking **m*, is estimated to have been in place approximately two thousand years ago (Trask 1995: 67), which postdates the Roman conquest of Sardinia. This makes the assumption that *m-* in Bq. *mokoru* and *mora* developed secondarily and was then borrowed into the Latin of Sardinia chronologically impossible.

10.3.3 Conclusion

To conclude, in light of the small number of compelling Sardo-Basque comparanda, the Sardinian lexicon does not provide convincing evidence for the presence of a language related to Basque in pre-Roman Sardinia (cf. Paulis 2008: 27). In fact, the fact that the bulk of suspected pre-Roman Sardinian words cannot be accounted for with an appeal to Basque is counter-evidence to a Vasconic presence. Most of the small number of pre-Roman words found both in Sardinian and Basque are more widely distributed in the Mediterranean region. This cannot be accounted for by our current understanding of the historical extent of the Basque and Aquitanian languages. The Vasconic language family did occupy a larger territory on both sides of the Pyrenees, but was decidedly oriented toward the Atlantic coast. On the Mediterranean coast different, unrelated, languages were spoken (Gorrochategui 1995: 57; Trask 1995: 87). Thus, the distribution of words like Srd. *golóstị̣u*, Bq. *gorosti*, Lang. *agaloûssës*, Gr. κήλαστρος 'holly' require a different explanation that does not involve Basque as the ultimate donor language. This will be discussed in § 11.

10.4 Etruscan

The Etruscan evidence that has been proposed to point toward a connection with pre-Roman Sardinian, consists of lexical comparisons (§ 10.4.1) and structural comparisons (§ 10.4.2). Additionally, comparisons between modern Tuscan and Sardinian have been taken as evidence in favor of a Tyrrhenian presence in Sardinia (cf. § 10.4.3). For a discussion of Pittau's hypothesis involving a linguistic relationship between pre-Roman Sardinian, Etruscan, and Lydian, see § 1.3.3.4.

10.4.1 Etruscan lexical comparisons

Etruscan lexical evidence is more difficult to come by than Basque (for which see § 10.3.1). The main reason for this is that Basque is a living and well-documented language. In the case of Etruscan, we run into several issues. Firstly, the number of Etruscan lexical items whose meaning has been determined with confidence, and which can therefore be compared to Sardinian words of pre-Roman origin, is rather low (Wallace 2008: 123). Secondly, due to the nature of the attested inscriptions, many of the Etruscan words for which a meaning has been identified concern commonly used verbs, kinship terms, and specific culture terms relating to Etruscan society and religion (Wallace 2008: 123–124). We do *a priori* not expect to find substrate words in the semantic domains of common verbs and kinship terms, as these are largely regularly inherited from Latin. On the other hand, the Etruscan cultural and religious lexicon is highly specific to the Etruscan context, e.g. various types of magistrates. It is questionable whether these concepts

ever existed in pre-Roman Sardinia. At the same time, the semantic domains that are most prone to be represented in linguistic substrates, such as flora, fauna, geographical terminology etc., are largely unknown from Etruscan. As a result, it is not surprising that only a single Etruscan word has convincingly been compared to Sardinian: OSrd. *muteclu* 'rockrose', cf. the Etruscan gloss μούτουκα 'thyme' (Dioscorides Anazarbei 1906: 48; Bonfante & Bonfante 1983: 190). This does not mean that there is no lexical evidence at all however; other lines of lexical evidence are discussed in § 10.4.3 – 10.4.4.

10.4.2 Etruscan structural comparisons

In the structural domain, there are several proposals for morphological parallels between Etruscan and the pre-Roman Sardinian material.

One parallel is the sequence *-en(n)-* attested on various Sardinian toponyms like *Azzen*, *Aratena*, *Ortobène*, *Salvènnero* etc. (Serra 1952: 409; Hubschmid 1963a: 173–174; Terracini 1927: 143; Pittau 1995: 196; cf. Wolf 1998a: 63). Terracini (1927: 143) compares this to African toponyms, but Hubschmid (1963a: 173–174) and Pittau (1995: 196) prefer a Tyrrhenian explanation. Presumably, this preference is based on Latin formations referring to persons containing *-enna/us* (e.g. Lat. *sociennus* 'partner'), hypothesized to be of Etruscan origin (cf. Weiss 2009: 485). The comparison between the Sardinian toponyms and the potentially Etruscan suffix found in Latin is problematic for several reasons. On the Sardinian side, *-en(n)-* is exclusively found on toponyms, whose semantic opacity renders any analysis of its function difficult (cf. § 2.1.2.2.1).[296] On the Etruscan side, *-en(n)-* is actually not attested as such. There is a suffix *-na-*, which forms possessive adjectives from nouns (Wallace 2008: 53), but it is not evident that it is related to *-en(n)-*. Formally, the Sardinian toponyms with *-en(n)-* are more similar to Lat. *-enna/us* (Wigman 2023: 28). Due to our poor understanding of the element *-en(n)-* in Sardinian toponyms, as well as its discrepancy with the actual attested Etruscan forms, it is impossible to confidently attribute *-en(n)-* in Sardinian toponyms to a Tyrrhenian source.

Hubschmid's (1963a: 175) suggestion that the pre-Roman suffix *-ír-* (§ 9.2.6.2) is somehow related to Etruscan is more substantiated. This suffix has been interpreted by Terracini (1927: 139) as an original plural marker. Hubschmid compares it to the Etruscan plural ending *-r*, *-r(a)-* (1963a: 175). If Terracini's analysis

[296] Hubschmid (1963a: 174) does mention cases like the toponym *Tettène*, which might be derived from *tétti* 'smilax' (DES II: 480). But, here too the connection between *Tettène* and *tétti* rests solely on formal similarity, without any verifiable shared semantics.

of Srd. *-Vr-* as an original plural is correct, there is indeed a good formal and functional match with the Etruscan morpheme. It should be noted however that Etr. *-r, -r(a)-* is used to form plurals of animate nouns only (Wallace 2008: 51), which is not the case for the pre-Roman Sardinian suffix. It is however imaginable that a hypothetical Tyrrhenian language spoken in Sardinia generalized this suffix to inanimate nouns.[297]

Finally, Pittau (1995: 197) proposes a connection between the prefix *ϑi-* found on numerous pre-Roman Sardinian words (§ 9.1.2), and the Etruscan demonstrative pronoun *eta, ita, ta*. Functionally, this is imaginable, as **ϑi-* was probably a morphosyntactic element. However, if Srd. **ϑi-* is originally identical to **kV-*, as proposed in § 9.1.3.1, the connection to Etr. *ita, eta, ta* becomes less likely. Interestingly, Etruscan also had a demonstrative pronoun *ika, eca, ca*, however, apparently with identical meaning and function to Etr. *eta, ita, ta* (Wallace 2008: 59–60). Pittau (1995: 197) could thus still be right in arguing that Srd. **kV-* (>> **ϑi-*) has a match in the Etruscan demonstrative pronoun, with the adjustment that Etr. *(e/i)ka* is a more suitable comparandum than *(e/i)ta*.

For two features attested in the Pre-Roman part of the Sardinian lexicon (i.e. **-Vr-* and **kV-/*ϑi-*), a connection to Etruscan morphemes is possible on formal and functional grounds. In both cases however, the behavior of these morphemes in the Sardinian lexicon is not identical to their presumed Etruscan counterparts. Plural *-r, -r(a)-* is restricted to animate nouns in Etruscan, unlike Srd. **-Vr*. Etruscan *(e/i)ka* functioned as a demonstrative pronoun that could be enclitic to the noun, while Srd. **kV-* and **ϑi-* preceded the lexical stem. It is therefore unlikely that the pre-Roman donor language of these Sardinian features — if they are indeed related to the Etruscan forms — was Etruscan itself. Rather, we would perhaps be dealing with a related Tyrrhenian language spoken in Sardinia, which is in line with what Pittau (1995) assumes.

10.4.3 Comparing Tuscan

One other avenue for drawing lexical parallels between Etruscan and the pre-Roman Sardinian language(s), is through Tuscan (cf. Pittau 1995: 197). As the distribution of the modern Tuscan varieties lines up roughly with the ancient extent of Etruscan, it stands to reason to ascribe etymologically obscure and seemingly non-inherited words in Tuscan to an Etruscan substrate. However, comparisons between etymologically obscure Sardinian words on the one hand,

[297] Pittau (1995: 196) claims that he is the first to make this connection, but it is in fact presented in Hubschmid (1963a: 197).

and etymologically obscure Tuscan words on the other, entail a double uncertainty. In such cases an etymological connection is possible, i.e. when if the formal and semantic criteria are met, but an unequivocal attribution of the forms in question to an Etruscan origin is problematic. Our understanding of the prehistoric linguistic diversity of the region is simply not advanced enough to rule out the existence of other pre-Roman languages from which both Sardinian and Tuscan could have incorporated words.

10.4.4 The lexical evidence

The following is an overview of the linguistic comparisons that can be adduced in favor of the view that the pre-Roman language(s) of Sardinia are related to Etruscan, according to Pittau (1995: 197–233). I have left out comparisons that are exclusively based on person names or toponyms, as well as some forms that have received more convincing etymologies in works in previous scholarship. All of the comparisons are evaluated in terms of their formal as well as their semantic agreement.[298]

1. § 3.1.6: Srd. *kuɣúśa, ϑurgúśa* etc. 'Celery, fool's watercress; water hemlock' < ***-gusa*** (DES I: 422, II: 555), Srd. *aússa* 'poison hemlock' < ***a(C)ussa*** (DES I: 153); cf. Lat. ***cicūta*** 'poison hemlock' (Pittau 1995: 209).[299]

   Formal match: **problematic**
   Semantic match: **good**
   Cognacy: **unlikely**

2. § 3.1.11: Srd. *buḍḍúri, búḍḍaru,* etc. 'poison hemlock' < ***bull-úri/-Vr-*** (DES I: 204; Paulis 1992: 151); cf. Tusc. ***bìlleri*** 'mayflower' (Paulis 1992: 151; Pittau 1995: 203). Other comparisons include Lat. *berula* 'lesser water parsnip (*Berula erecta*)' (DEI I: 520b; André 1985: 35).

   Formal match: **possible**
   Semantic match: **possible**, although not straightforward
   Cognacy: **possible**

[298] For the sake of consistency, the order adhered to here is the same as in chapters § 3 – 7, and may deviate from Pittau (1995).
[299] Lat *cicūta* has been compared to Celtic forms (Pedersen 1909: 209; Wigman 2023: 193).

3. § 3.3.13: OSrd. *muteclu* 'rockrose' < ***muték(u)lu*** (DES II: 150-51); cf. Etr. μούτουκα 'thyme; rockrose' (Dioscorides Anazarbei 1906: 48; Bertoldi 1936: 312 fn. 1; Bonfante & Bonfante 1983: 190). Other comparisons include: Calab. *mútaca*, *múndaca*, *amùndaci*, *amùndici* 'etc. 'narrow-leaved cistus (*Cistus monspeliensis*)' (Bertoldi 1936: 312; Alessio 1949: 114; Rohlfs 1977: 439, 447), Lang. *mouche*, *mouge* etc., Cat. *módega* 'id.' (Hubschmid 1953: 36; FEW VI/III: 306; DECLC V: 715-716) < ***mútVka***; Cat. *motxera* (reg. Roussillon) 'id.' (FEW VI/III: 306) < ***mutukária***.

   Formal match: **good**
   Semantic match: **good**
   Cognacy: **compelling**

4. § 3.3.21: Srd. *sárpa* etc. 'willow species' (DES II: 384) < ***sárpa***; cf. Lat. *sirpus*, *scirpus* 'bulrush' < ***skerpo***- (Pittau 1995: 227; cf. de Vaan 2008: 546). However, Srd. *sárpa* may be a derivation of Lat. *sarpō* 'to cut off, to prune' (cf. Weiss 2009: 299, 300).

   Formal match: **problematic**
   Semantic match: **possible**
   Cognacy: **unlikely**.

5. § 3.4.3: Srd. *bòḍḍoro*, *bòḍḍero* 'juniper berry' < ***bóllVr***- (DES I: 215; Pittau 1995: 202); cf. Tusc. ***bòllora*** (Lucca) 'gall (on plants)'. Other comparisons include Sic. *bóḍḍaru* 'grain, kernel; bump' (DES I: 215).

   Formal match: **good**
   Semantic match: **good**
   Cognacy: **compelling**, although a sound-symbolic origin (either shared or independent) is also imaginable (cf. Paulis 1992: 444).

6. § 3.4.8: Srd. *kòkkoro*, *kòkkolo* 'walnut' < ***kòkkoro*** (DES I: 356); cf. Tusc. ***cóccoro*** 'big, round hazelnut' (Castelnuovo di Garfagnana) (DES I: 356; Pittau 1995: 207). Other comparisons include: Nap. *còccola* 'walnut husk', Laz. *kùkkęrę* 'half of a walnut husk' (Castro dei Volsci), Romagnol *cocla* 'walnut'; Arag. *coca* 'walnut'; Pg. *cócora* 'cooked chestnut' etc. (DES I: 356). Wagner (DES I: 357) explains these forms from Lat. *coccum*, presumably with secondary addition of pre-Roman *-́*Vr*- (cf. § 9.2.6.2).

Formal match: **good**
Semantic match: **good**
Cognacy: **possible**, but perhaps from Lat. *coccum*.

7. § 4.1.9: Srd. [*káḍḍ' e*] *pilíkke* 'dragonfly', *tsimpilíɣe* 'grasshopper' < ****pilíke*** (DES II: 265, 547); cf. Tusc. (*s*)*piluccare* 'eating little bits at a time' (Pittau 1995: 223) << Lat. ***pilāre*** 'to depilate'.

   Formal match: **problematic**
   Semantic match: **problematic**
   Cognacy: **unlikely**

8. § 4.5.10: Srd. *mufròne*, *muɣròne* etc. 'mouflon' < ****mufròne*** (DES II: 134-135); cf. Cors. *mufrone*, *muvrone* etc. 'mouflon' < ****mufròne*** (DES II: 134-135; Pittau 1995: 218), Lat. *mufron* (Polemius Silvius, 5th c. CE) 'mouflon' (TLL VIII: 1557; Ernout & Meillet 2001: 417).

   Formal match: **good**
   Semantic match: **good**
   Cognacy: **compelling**, for the Sardinian and Corsican forms. The Latin attestation likely represents a direct loan from either Corsican or Sardinian. There is no evidence for an Etruscan origin.

9. Srd. *barrókku*, *borrókku* 'cliff, ravine, gully, gorge' < ****barrokk***- (Paulis 1987: 470; Pittau 1995: 202; cf. 5.1.28); cf. Tusc. *borro*, *borròne*, burrone 'channel, ditch, precipice' (Pittau 1995: 202) < ***bótro*** 'id.' << Gr. βόθρος 'hole, trench' (DEI I: 568b, 574b).

   Formal match: **problematic**
   Semantic match: **good**
   Cognacy: **unlikely**

10. § 5.1.15: Srd. *péntuma*, *ispéntumu* 'precipice, cliff, ravine' < ****péntum***- (DES II: 244); cf. Etr. ***penϑuna***, ***penϑna*** 'stele, cippus' (Breyer 1993: 347 fn. 515). Further comparisons include: Laz. *pèndima* (Nemi) 'sloping rock face', Campan. *péndəm*^{a} (Formicola), *péŋ́ima* (Ausonia) 'steep rock', Salent. *péntima* (Manduria, Taranto), *pèntuma* (Calimera) 'isolated

boulder; cliff' etc. (Parodi 1907: 772; Rohlfs 1927a: 163 fn. 1; Ribezzo 1933; Aebischer 1934). The Sardinian form may however be a descendant of a regional Latin formation (Parodi 1907: 772; Rohlfs 1927a: 163 fn. 1; Aebischer 1934).

Formal match: **uncertain**
Semantic match: **problematic**
Cognacy: **unlikely**

11. § 5.1.28: Srd. *karróppu*, *garróppu*, *korróppu*, *gorróppu* etc. 'whirlpool; crevice etc.' < ***garrópp**- (DES I: 308); cf. Tusc. ***karravone***, Cors. ***karravóni*** 'ravine' (Hubschmid 1953: 58; Pittau 1995: 206). Other comparisons include: Irpino *sgarrupá*, Abruzz. *scarrupà* 'to fall down, to plunge', Laz. *carapone* (Castelmadama) 'whirlpool, swamp', Nap. *scarrupo*, *sgarrupo* 'cliff, precipice', Calab. *garroppa* (Petronà) 'trough' etc. (Rohlfs 1932: 336; Wagner 1932: 243; DES I:308); Calab. *karafúnu* 'puddle' (Hubschmid 1953: 58).

    Formal match: **good**
    Semantic match: **good**
    Cognacy: **compelling**, but cf. § 5.1.28.

12. § 5.2.4: Srd. *moretína* etc. 'pile of rocks; dry-stone wall' < ***moretína*** << ***mora* (?)** (DES II: 127-28);[300] cf. Tusc. ***mòra, móra*** etc. 'pile of rocks; dry-stone wall' (Pittau 1995: 219).[301] Other comparisons include: Cors. *mora*, *mola* 'pile of rocks or wood', Laz. *mora* (Castelmadama) 'boulder, very big rock' (DES II: 128); Bq. *more*(*a*), *mora* 'cairn, pile of rocks; dolmen' < **mora* (DCEC IV: 146); Sp. *morena*, Pg. *moreia* ´pile of plant material', Frpr. *morêna* (Savoy) 'glacier moraine' (DCEC IV: 146; cf. Hubschmid 1943: 150–151; Michelena 1969: 23–25; FEW VI/III: 276b).

[300] I do not include Pittau's (1995: 219) connection to Srd. *múra*, *núrra* and *nuráke*. The occurrences of the first form (i.e. *múra*) in the Condaghe di Santa Maria di Bonarcado (Virdis 2003: sec. 1.18, 53.1, 176.2, 207.6) are misidentified by Pittau as meaning "'pile of rocks' and 'nuraghe'", but it actually denotes a 'mulberry tree' (DES II: 137-38; Virdis 2003: 263). The other forms, *núrra* 'pile of wood or rocks; chasm' and *nuráke* 'nuraghe', cannot be compared to *moretína* (or to hypothetical **mora*) because of the phonetic distance between the forms.

[301] Pittau (1995: 220) compares Lat. *mūrus* 'defensive wall', but this goes back to PIE *(*H*)*moi-ro*-, and thus does not belong here (de Vaan 2008: 396).

Formal match: **good**
Semantic match: **good**
Cognacy: **compelling**

13. Srd. *garuléu*, *galuréu*, *galiléu* 'pollen deposited in honey' (Pittau 1995: 211–212) < ***galuléu***;[302] cf. Etr. γαρουλέον 'chrysanthemum' (Bonfante & Bonfante 1983: 189; Pittau 1995: 212).

    Formal match: **good**, although the Sardinian form is similar to other productively derived words.
    Semantic match: **problematic**
    Cognacy: **unlikely**

14. Srd. *gurúle* (Bitti) 'pass, ravine between mountains' < ***gurúle*** (DES I: 597, II: 609); cf. Lat. *cūria* 'division of the Roman people; meeting place' < PIt. ***ko-wir-ija*** (de Vaan 2008: 157), Lat. [***sella***] ***curūlis*** 'type of official chair' (Pittau 1995: 213). Lat. *curūlis* has been proposed to be an Etruscan loan, but it could equally well be a derivation of *currus* 'chariot' (de Vaan 2008: 157).

    Formal match: **possible**, but only in the case of Lat. *curūlis*, which moreover has an alternative etymology entailing a productive derivation.
    Semantic match: **problematic**. Pittau (1995: 213) presumably envisages a semantic shift from 'chair' to 'mountain pass' (cf. It. *sella* 'saddle; mountain pass'). However, it is Lat. *sella* that refers to a chair, not *curūlis*, rendering the comparison unconvincing.
    Cognacy: **unlikely**

15. Srd. *nárvas*, *návras*, *náβras*, *náiriši* etc. 'earrings; earlobes' (DES II: 157-58) < ****nápr-/*náb(e/i)r-***; cf. Lat. ***napurae*** 'straw ropes', proposed to be from Etr. ***naper*** 'some surface measure' (Ettmayer 1926: 29; Breyer 1993: 265–266; TLL IX/I: 49; Bonfante & Bonfante 1983: 217). Wagner

[302] The ending of these words is reminiscent of that found *murguléu* 'cat thyme; cotton lavender' etc. (§ 3.1.5), which Paulis (1992: 119) derives from **véru* 'true' — cf. also *karduréu* 'thistle' < Lat. *carduum verum* (DES I: 300). It is not clear to me what the element *gali/u-* could be.

(DES II: 157-58) doubts the Latin etymology on formal and semantic grounds.

Formal match: **possible**
Semantic match: **problematic**, at least for the Etruscan comparandum.
Cognacy: **unlikely**; if any of the listed words are related, inheritance of the Sardinian forms from the Latin forms is semantically far more likely than a connection between the Sardinian and Etruscan forms.

10.4.5 Discussion

Of the fifteen comparisons that could potentially be indicative of the presence of a Tyrrhenian language in pre-Roman Sardinia, only five are formally and semantically close enough to be considered as evidence. These are the words for 'rockrose' (3), 'juniper berry' (5), 'mouflon' (8), 'whirlpool, crevice' (11), and 'pile of rocks; dry-stone wall' (12). For the ten remaining cases, either the semantics are too distant to confidently posit a common origin, or the formal details speak against a connection. The words for 'poison hemlock' (2) and 'walnut' (6) are uncertain cases. Table 10.4.1 summarizes the formal and semantic criteria for each discussed word.

	Srd.	**Comp.**	**Form.**	**Semant.**	**Cogn.**
1	**-gusa* 'watercress, celery; water hemlock'	Lat. *cicūta* 'poison hemlock'	×	✓	×
2	**bull-úri/-Vr-* 'poison hemlock'	Tusc. *bìlleri* (pl.) 'mayflower'	✓ (?)	?	?
3	**muték*(*u*)*lu* 'rockrose'	Etr. μούτουκα 'rockrose; thyme'	✓	✓	✓
4	**sárpa* 'willow'	Lat. *sirpus*, *scirpus* 'bulrush'	×	?	×
5	**bóllVr-* 'juniper berry'	Tusc. *bòllora* 'gall (on plants)'	✓	✓	✓
6	**kókkVr-* 'walnut'	Tusc. *cóccoro* 'big hazelnut'	✓	✓	?
7	**pilíke* 'grasshopper/dragonfly'	Tusc. (*s*)*piluccare* 'to eat little bits at a time'	×	×	×
8	**mufróne* 'mouflon'	Cors. *mufròne*, Lat. *muflon* 'mouflon'	✓	✓	✓
9	**barrokk-* 'cliff, ravine, gorge'	Tusc. *borro* 'channel, ditch, precipice	×	✓	×

	Srd.	Comp.	Form.	Semant.	Cogn.
10	**péntum-* 'precipice, cliff, ravine'	Etr. *penθuna* 'stele, cippus'	?	×	×
11	**garropp-* 'whirlpool; crevice'	Tusc. *karravone*, Cors. *karravóni* 'ravine'	✓	✓	✓
12	**moretína* << **móra* 'pile of rocks; dry-stone wall'	Tusc. *móra*, *mòra*, Cors. *mola*, *mora* 'id.'	✓	✓	✓
13	**galuléu* 'pollen deposited in honey'	Etr. γαρουλέον 'chrysanthemum'	✓	×	×
14	**gurúle* 'mountain pass, ravine'	Lat. [*sella*] *curūlis* 'official chair'	?	×	×
15	**nap*(*V*)*r-*/**nab*(*V*)*r-* 'earrings'	Lat. *napurae* 'straw ropes'; Etr. *naper* 'surface measure'	?	×	×

Table 10.4.1: Proposed lexcial evidence for Tyrrhenian Sardinian contact

10.4.5.1 Etruscan and Sardinian

As noted before, the presented material, including the convincing cases, consists of three categories: comparisons between Sardinian and Etruscan, comparisons between Sardinian and Latin, and comparisons between Sardinian and Tuscan. Of the first group, there is only a single case in which the Sardinian and Etruscan are formally and semantically close enough to consider a shared origin.[303] This is the word for 'rockrose' (3): Srd. **mutéklu*, Etr. μούτουκα. While this can potentially be invoked as an argument in favor of an Etruscan presence in Sardinia, it should be noted that continuants of **mutaka* are also found in the Iberian Peninsula (e.g. Cat. *módega*), southern France (e.g. Lang. *mouche*), and southern Italy (e.g. Calab. *mútaca* 'rockrose'). Given that Etruscan is not known to have been spoken in such a wide territory, it is better to interpret the Sardinian form in a wider Western Mediterranean context rather than as a direct loan from Etruscan (§ 11.3).

10.4.5.2 Latin and Sardinian

As far as un-etymologized Latin words with correspondents in Sardinian are concerned, there is once more just one example: Srd., Cors. **mufróne*, **múfra* 'mouflon', Lat. *mufron* 'id.'. However, this does not constitute evidence for an Etruscan origin of the Latin form. In light of the late attestation of this word (Polemius Silvius, 5th c. CE), and the fact that it refers to specifically Sardinian

[303] This is perhaps already more than we would expect, in light of our limited knowledge of the Etruscan lexicon.

and Corsican animals, it is more likely that Lat. *mufron* is a direct loan from a pre-Roman Sardinian or Corsican language. The Latin forms *musmō* and *imbrōs* (acc. pl.), both plausibly referring to the mouflon, are attested earlier and similar albeit formally less close to Srd., Cors. **mufròne* (cf. TLL VII/I: 140; VIII: 1705). If these forms are related to **mufròne*, the assumption of a loan from a native Sardinian or Corsican source stands to reason. In short, there is no evidence that any of these words for 'mouflon' is of Tyrrhenian origin, as claimed by Pittau (1995: 218).

10.4.5.3 Tuscan and Sardinian

The third category is somewhat more substantial, with a total amount of three compelling cases found in Pittau (1995).

- Srd. **bòllVro* 'juniper berry' Tusc. *bòllora* 'gall (on plants)'
- Srd. **moretína* << **móra* 'pile of rocks' Tusc. *mòra*, *móra* 'id.'
- Srd. **garróppu* 'whirlpool; crevice' Tusc. *karravone* 'ravine'

In all these cases, the Sardinian forms correspond well semantically and formally to Tuscan words that are of unknown origin. This makes them candidates for Etruscan loans. However, as with the words referring to 'rockrose' (3), all three of these Sardinian-Tuscan connections are distributed beyond Sardinian and Tuscan, beyond the known extent of Etruscan. The word for 'juniper berry' (5) has a likely cognate in Sic. *bóḍḍaru* 'grain, kernel; bump'; that for 'pile of rocks; dry-stone wall' (9) corresponds to Bq. **mora* 'cairn, pile of rocks; dolmen', Sp. *morena* 'pile of plant material' etc.; and that for 'whirlpool, crevice' (12) is attested further south as Nap. *scarrupo* 'cliff, precipice', Calab. *garroppa* 'trough' etc. In view of this geographical distribution, it is not evident that the Tuscan forms go back to Etruscan. This weakens the argument for a Tyrrhenian origin of the corresponding Sardinian forms.

10.4.6 Conclusion

As we have seen, there is limited evidence for the hypothesis that the pre-Roman Sardinian was related to Etruscan. On the structural side, two morphological elements in the non-inherited Sardinian lexicon, i.e. pluralizing **-́Vr* and morphosyntactic **kV-*/**ϑi-*, could correspond to known Etruscan morphemes, i.e. animate plural *-r*, *-ra-* and demonstrative (*e/i*)*ka*. On the lexical side however, there is but a single Sardinian word with an direct counterpart in Etruscan, viz. Srd. **muteklu*, Etr. μούτουκα 'rockrose'. The accepted Etruscan and Tuscan comparanda show a wide distribution across Sardinia (Figure 10.4.1).

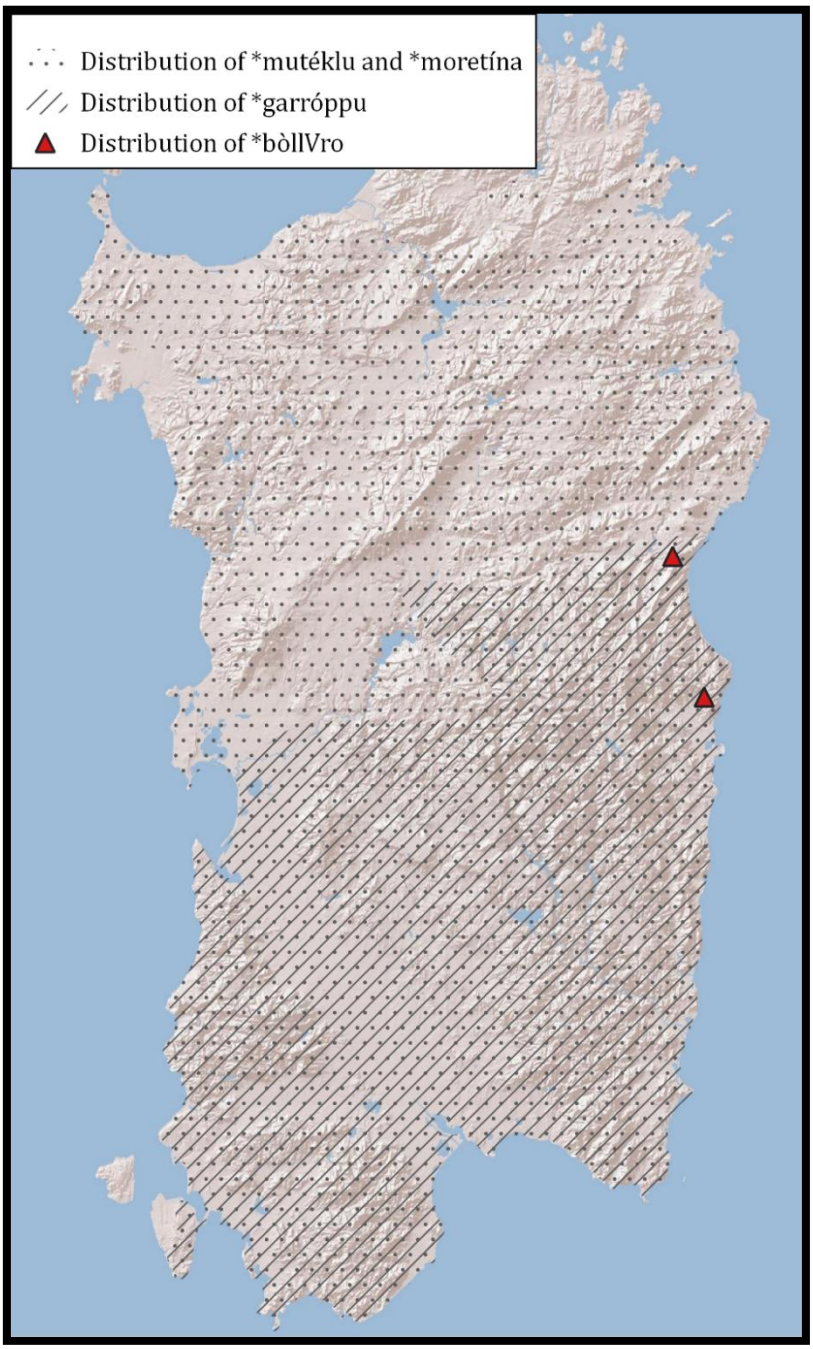

Figure 10.4.1: Distribution of accepted Etruscan and Tuscan comparisons

Because of our limited knowledge of the Etruscan lexicon, and because of the semantic nature of the parts that are known, a low number of identifiable lexical correspondents is to be expected in any scenario of contact. The dearth of clear lexical correspondences does as such not necessarily contradict the hypothetical presence of a Tyrrhenian language in Sardinia. However, even in the case of the Latin and Tuscan comparisons discussed in § 10.4.5.2 and § 10.4.5.3 it is hard to unequivocally attribute any to an Etruscan source. This is mainly due to the fact that the compared forms are distributed over a geographical area that greatly exceeds the current consensus on the historical extent of Etruscan and the Tyrrhenian language family in general. This problem in fact also applies to the single convincing Sardinian-Etruscan correspondence, i.e. the word for 'rockrose', which has compelling cognates in Catalan, Occitan and Calabrian. The distributions of the words, discussed in § 10.4.5, fit into a larger pattern of apparent pre-Roman word families found across portions of the Mediterranean (cf. § 11). On the basis of the scarce linguistic evidence presented here, it is not possible to either confirm, or reject the presence of a Tyrrhenian language in pre-Roman Sardinia, despite the two morphological correspondences discussed in § 10.4.2 being tantalizing clues.

11 Mediterranean comparisons

Numerous Sardinian words of suspected pre-Roman origin have been compared to equally obscure etyma in other languages of the Mediterranean region. This has happened predominantly in the context of the "Mediterranean substratum" hypothesis, which claims that one or more hypothetical pre-Indo-European languages were once spoken across a significant portion of the Mediterranean region. The details of this hypothesis, such as the number of different substrata and their relative chronology, differ from scholar to scholar (see § 1.3.2). In this chapter, the proposed comparisons involving Sardinian will be reviewed in the same way as in § 10. This will allow for the evaluation of the question whether there is evidence for a pre-Roman, i.e. pre-Indo-European, substrate reaching beyond Sardinia itself. In § 11.2.1 the geographical distributions of the accepted cases are discussed. Interpretations of these distributions are offered in § 11.3.

11.1 The lexical evidence

The following is an overview of non-inherited Sardinian words and their proposed comparanda in various other languages of the Mediterranean. In each case, the forms that are compared are printed in bold. In addition, an evaluation is offered for the formal correspondence between the Sardinian and the other forms, as well as for their semantic correspondence. Any comparisons discussed in § 10 that have been proposed to have wider Mediterranean comparanda, are included too:

1. § 3.1.2: Srd. *tattaróyu*, *ϑoϑϑóṛ̥u* 'arum' etc. < ***ϑaϑarój̥*** (DES II: 552; Paulis 1992: 341); cf. Sic. *anzara/u* 'arum', *azzaru* 'friar's cowl', Calab. (*an*)*zara* etc. 'arum', Pugl. *azara* 'arum' < ****a(n)zar-*** (Rohlfs 1932: 99; 1977: 84b, 802b; Alessio 1949: 110–111; Hubschmid 1953: 27); Vulg. Lat. ***asaru*** (Dioscorides) 'hazelwort/friar's cowl' (Hofmann & Auracher 1883: 60; Stadler 1899: 237, 238); Gr. ἄσαρον 'hazelwort' (Dioscorides Anazarbei 1907: 14).

   Formal match: **possible**
   Semantic match: **good**
   Cognacy: **uncertain**
   Distribution: Sardinia, southern Italy, Sicily

2. § 3.1.11: Srd. *buḍḍúri, búḍḍaru,* etc. 'poison hemlock' < ****bull-úri/-Vr-*** (DES : 204; Paulis 1992: 151); cf. Tusc. ***bìlleri*** 'mayflower' (Paulis 1992: 151;

Pittau 1995: 203); Lat. ***berula*** 'lesser water parsnip (*Berula erecta*)' (DEI I: 520b; André 1985: 35). Cf. § 10.4.4.

Formal match: **possible**
Semantic match: **possible**, although not straightforward
Cognacy: **uncertain**
Distribution: Sardinia, central Italy

3. § 3.1.15: Srd. *ϑáṇḍa, aϑáṇḍa, ánna* etc. 'poppy' < ***(ϑ-)anda*** (DES I: 151); cf. Egyptian ναντί 'poppy'; Lang. ***ãndèr, õndèr*** (dep. Lot) 'id.' (FEW XXIV: 550 fn. 21; Bertoldi 1937a: 151 fn. 1), Gr. ἄνθος 'flower' < PIE ****h₂end^{h}-*** (Pittau apud Paulis 1992: 319; cf. Beekes 2010: 104–105).[304]

Formal match: **good**, but only for the Lengadocian forms if they are not from **derbita* 'ulcer' (cf. FEW XXIV: 550 fn. 21).
Semantic match: **good**
Cognacy: **compelling**, depending on the etymology of Lang. *ãndèr, õndèr*.
Distribution: Sardinia, southern France

4. § 3.2.1: Srd. *aláṡe, aláṡi̯u* etc. 'butcher's broom, holly' < ****alás(i)*** < ****al-ás-*** (DES I: 67-68; Paulis 1992: 417–418); cf. Bearn. *aroû*, Prov. *ářę* 'broom' < ****ala***, Auv. *alète* < ****alista***, Lang. *orousto* < ****alusta*** 'common broom' (FEW XXI: 107; Paulis 1992: 418); Lig. [*la*] *lastra, lástroga, arástra* 'spiny broom; spiny shrub', Sic. *alastra* 'spiny broom' (Bertoldi 1928: 233 fn. 3; Battisti 1960: 373; Paulis 1992: 418) < ****alástra***; Lat. ***alaternus*** 'buckthorn' (Battisti 1960: 374).

Formal match: **possible**, under the assumption of **-as-* as a suffix in the Sardinian form
Semantic match: **possible**, although all compared forms refer to 'broom' while the Sardinian forms refer to 'butcher's broom'
Cognacy: **possible**; the non-Sardinian forms are a better match to each other than to the Sardinian forms, however.
Distribution: (possibly) Sardinia, southern France, northern Italy, Sicily

[304] The comparison to Bq. *andere* 'lady' (cf. Bertoldi apud Wagner 1932: 237) has been rejected in § 10.3.1.

5. § 3.3.7: Srd. *golóstḭu*, *golóstri*, *olósti*, *golóstiʔe* etc. 'holly' < ***golósti*** (DES I: 584); cf. Bq. *gorosti* 'holly' < ***golosti*** (Bertoldi 1929: 261 fn. 3); Navarr. ***colostia, golostia*** etc. (Hubschmid 1953: 29 fn. 1); Lang. *agaloussé* 'id.' < ***agalústo***- (Jud 1931: 439; cf. FEW XXI: 106); Lucanian *cèlastru*, Calab. *cègghiastru*, *cìstriddazzu*, Bov. *azzìddastro* 'holly' < ***celastru***- (DEI II: 842); Gr. **κήλαστρος** 'id.' (Hofmann 1949: 142); Bret. *kelenn*, OIr. *cuilenn* 'id.' < PC **kolinno*- < ***kolis-n***- (Hubschmid 1953: 29; van Sluis, Jørgensen & Kroonen 2023: 200); OHG *hulis*, Du. *hulst* 'id.' < PGm. **hulisa* < ***kulis***- (Kroonen 2013: 253). The Mediterranean comparanda are a very close semantic and formal match. Whether the Celtic and Germanic forms are related is less evident (cf. § 10.3.1).

   Formal match: **good**, for the Mediterranean forms.
   Semantic match: **good**
   Cognacy: **compelling**, for the Mediterranean forms.
   Distribution: Sardinia, Iberia, southern France, southern Italy, Greece

6. § 3.3.9: Srd. *kòstike*, *kòstiɣe*, *kòsti* 'maple' < ***kòsti*** (DES I: 392-93); cf. Bq. *astigar*, *gastigar*, *gaztigar* 'maple' < ***(g)azt***- + *ihar* 'maple' (Bertoldi 1931: 131; cf. Hubschmid 1953: 82; Michelena 1953: 48; Michelena 1990: 253 fn. 49); Tusc. *còstolo* 'ash tree' < ***kost***-, Gr. Hes. **ἄκαστος** (Hesychius Alexandrinus 2018: 109), OProv. *agast* etc. 'maple' < ***akást***- (Alessio 1949: 132; FEW XXIV: 276). The Occitan, Tuscan and Greek forms are formally and semantically attractive comparanda to the Sardinian forms; the formal details of the Basque comparanda are not quite clear (cf. § 10.3.1).

   Formal match: **good**, especially for the Occitan, Tuscan and Greek forms
   Semantic match: **good**
   Cognacy: **compelling**: the Occitan, Tuscan, and Greek comparisons are good, but the internal structure of the Basque forms is not well-understood.
   Distribution: Sardinia, (possibly) Iberia, southern France, Greece (?)[305]

7. § 3.3.13: OSrd. *muteclu* 'rockrose' < ***muték(u)lu*** (DES II: 150-51); cf. Etr. **μούτουκα** 'thyme; rockrose' (Dioscorides Anazarbei 1906: 48; Bertoldi

[305] The provenance of this Hesychian gloss is not known.

1936: 312 fn. 1; Bonfante & Bonfante 1983: 190); Calab. *mútaca*, *múndaca*, *amùndaci*, *amùndici* 'etc. 'narrow-leaved cistus (*Cistus monspeliensis*)' (Bertoldi 1936: 312; Alessio 1949: 114; Rohlfs 1977: 439, 447), Lang. *mouche*, *mouge* etc., Cat. *módega* 'id.' (Hubschmid 1953: 36; FEW VI/III: 306; DECLC V: 715-716) < ***mútVka***; Cat. *motxera* (reg. Roussillon) 'id.' (FEW VI/III: 306) < ***mutukária***. See § 3.3.13 for formally further removed forms with the same meaning.

Formal match: **good**
Semantic match: **good**
Cognacy: **compelling**
Distribution: Sardinia, Iberia, southern France, central Italy, southern Italy

8. § 3.3.16: Srd. *oliðòne*, *oliòni* etc. 'strawberry tree' < ***olidóne*** 'strawberry tree' (DES II: 185); cf. Gasc. *auledun*, Saintongeais *olon*, Lang. *leduno*, Pied. (*l*)*ürion* 'id.' < ***olidòne*** (Schuchardt 1904: 194; REW 9068; FEW XIV: 35). Lat *ūnedō* 'strawberry tree', recorded only in Pliny (1857: chap. 15:28), is possibly folk-etymological after a form **olidō* of more restricted distribution than Lat. *arbutus* 'strawberry tree'.

   Formal match: **good**
   Semantic match: **good**
   Cognacy: **compelling**. The possibility that an unattested form **olidō* existed in Latin and was only marginally continued in Romance cannot fully be excluded however.
   Distribution: Sardinia, southern France, northern Italy, (possibly) central Italy

9. § 3.4.2: Srd. *kárva* 'tree branch' < ***kárva**/**kárba*** (DES I: 311); cf. Bq. *karbasta*, *garbazta* etc. 'stick with branches', *karbaza* 'stem, trunk' < ***karba***- (Wagner 1932: 231; Michelena 1990: 251), or << *arba* < Bq. ***arba*** (Lakarra et al. 2019: 3–8; cf. § 10.3.1); Ast. *garbu* 'small firewood', Pg. *gravulha* 'dry pine needles', Gal. *garabullos* 'dry tips of tree branches used as small firewood', Gasc. *garbalhe* 'pine needles', Prov. *garbo* 'hollow tree trunk' etc. < ***garb***- (Wagner 1932: 231; DCEC I: 856); Lat. *carbō*, *-ōnis* 'piece of charcoal' < ***karb^{h}***- (de Vaan 2008: 91–92).

Formal match: **good**, although the Basque forms may be of more recent origin.
Semantic match: **good**
Cognacy: **compelling**
Distribution: Sardinia, Iberia, southern France, central Italy

10. § 3.4.3: Srd. *bòḍḍoro*, *bòḍḍero* 'juniper berry' < ***bóllVro*** (DES I: 215; Pittau 1995: 202); cf. Tusc. ***bòllora*** (Lucca) 'gall (on plants)', Sic. ***bóḍḍaru*** 'grain, kernel; bump' (DES I: 215).

    Formal match: **good**
    Semantic match: **good**
    Cognacy: **compelling**, although a sound-symbolic origin (either shared or independent) is also imaginable (cf. Paulis 1992: 444).
    Distribution: Sardinia, central Italy, Sicily

11. § 3.4.8: Srd. *kòkkoro*, *kòkkolo* 'walnut' < ***kòkkoro*** (DES I: 356); cf. Tusc. ***cóccoro*** 'big, round hazelnut' (Castelnuovo di Garfagnana) (DES I: 356; Pittau 1995: 207), Nap. *còccola* 'walnut husk', Laz. *kùkkẹrẹ* 'half of a walnut husk' (Castro dei Volsci), Romagnol *cocla* 'walnut'; Arag. *coca* 'walnut'; Pg. *cócora* 'cooked chestnut' etc. (DES I: 356) < ***kókk-***. Wagner (DES I: 357) explains these forms from Lat. *coccum*, presumably with secondary addition of pre-Roman **-́Vr-* (cf. § 9.2.6.2).

    Formal match: **good**
    Semantic match: **good**
    Cognacy: **possible**, but perhaps from Lat. *coccum*.
    Distribution: Sardinia, Iberia, Italy

12. § 4.5.3: OSrd. *gjagar, jagaru* Srd. *ǧáɣaru* 'hunting dog' < ****i̯ákaru*** (DES I: 602); cf. Bq. ***txakur*** 'dog' (Schuchardt 1905: 227; Wagner 1932: 232); Cors. *ghjácaru* etc. 'shepherd dog' < ****i̯ákaru*** (Wagner 1932: 232; BDLC: s.v. dog); Byz. Gr. ζαγάριον, MGr. ζαγάρι, Turkish ***zağar*** 'hunting dog', Georg. ძაღლი *dzaghli* 'dog' < Proto-Kartvelian ****ʒ₁aɣl-*** (Schuchardt 1905: 227). Schuchardt (1905) proposes the Sardinian forms to be a Byzantine loan, but Wagner (1932) rejects this because of the Corsican and Basque comparanda. I tend to agree with Schuchardt.

Formal match: **possible**
Semantic match: **good**
Cognacy: **possible**, but probably a more recent loan from Byzantine Greek in Sardinian.
Distribution: Sardinia, Iberia, Corsica, Balkans, Greece, Asia Minor

13. § 4.5.4: Srd. *kallédḍu*, *kallúttsu* etc. 'doggy, puppy' < ***kalḭ***-; cf. Alb. *këlýsh*, *këlúsh* 'young animal' << ****kul***-, Gr. **σκύλαξ** 'puppy', Gr. Hes. **κύλλας** 'id.' (Elis) (DES I:320; Demiraj 1997: 218–219; Hesychius Alexandrinus 2020: 689); Welsh *colwyn* Bret. *kolen* etc. 'puppy' < PC ****koligno*** (Meyer 1891: 186; Matasović 2008: 213).[306] The comparison revolves around a sequence **kVl*-, with different vocalisms. This is hardly enough to confidently posit an etymological connection.

    Formal match: **uncertain**, due to the shortness of the compared sequence.
    Semantic match: **good**
    Cognacy: **uncertain**; we could be dealing with a case of chance similarity.

14. § 4.5.5: OSrd. *bita*, Srd. *bètte*, *bítti* etc. 'fawn, lamb, small animal' < ***bett***-, ***bitt***- (DES I: 211-12); cf. Cat. *bita*, *bitè*, Gasc. *pito*, For. *bette*, *bitte* 'goat' < ***bitt*-/*bett*-/*pitt***- (Hubschmid 1953: 70; Hubschmid 1955: 189–190; cf. FEW XXII/I:288). Wagner (DES I: 212) prefers to explain these forms as independent onomatopoeic formations. This is supported by the fact that these forms are also well-attested as animals calls (FEW XXII/I: 288).

    Formal match: **possible**, although there are various incongruencies within and between the different languages.
    Semantic match: **good**
    Cognacy: **uncertain**; there is a possibility that these forms are independent formations originating in animal calls.

15. § 4.5.8: Srd. *lèppere*, *lèppore*, *léppuri* etc. 'hare' < ***leppVr***- (DES II: 22-23); cf. Lat. *lepus*, *-oris* 'hare' < ***lepos***- (Wigman 2023: 173), Siculian **λέποριν**

[306] The connection to Lith. *kalẽ* 'bitch' was rejected by Beekes (2010: 1361).

'hare', Massiliot Gr. (and possibly Turdetanian) λεβηρίς 'rabbit' (Bertoldi 1937a: 146); Pg. *láparo*, OProv. *laperel* 'young rabbit', OFr. *lapriel* 'rabbit' etc. < ***lappar***- (Hubschmid 1943: 246–250). Further removed are Pg. *lapim*, Arag. *llapí*, Fr. *lapin* < ***lappīnu***- (ibid.).

Formal match: **good**
Semantic match: **good**
Cognacy: **compelling**
Distribution: Sardinia, Iberia, France, Italy, Sicily

16. § 4.5.10: Srd. *mufròne*, *muɣròne* etc. 'mouflon' < ***mufròne*** (DES II: 134-135); cf. Cors. *mufrone*, *muvrone* etc. 'mouflon' < ***mufrone*** (DES II: 134-135; Pittau 1995: 218; BDLC: s.v. mouflon), Lat. *mufron* (Polemius Silvius, 5th c. CE) 'mouflon' (TLL VIII: 1557; Ernout & Meillet 2001: 417).

Formal match: **good**
Semantic match: **good**
Cognacy: **compelling**, for the Sardinian and Corsican forms. The Latin attestation likely represents a direct loan from either Corsican or Sardinian (cf. § 10.4.4).
Distribution: Sardinia, Corsica

17. § 5.1.1: Srd. *núrra* 'chasm, shaft; pile' < ***núrra*** (Hubschmid 1953: 45; DES II:176-79); cf. Ast. *noriu* 'pile of rocks made in an estate', *noiro* 'bump of earth with undergrowth' < ***noṛiu***-, Galic. ***nórrago*** 'big knot in the root of a tree' (Hubschmid 1953: 47; Hubschmid 1958: 219).

Formal match: **possible**, but with some problems.
Semantic: **problematic**
Cognacy: **unlikely**

18. § 5.1.3: Srd. *kála* 'den (of animals)' < ***kála*** (DES I: 264-65);[307] cf. Sp., Cat., OProv., Cors., It., Sic. *cala* 'cove, inlet, creek' (Wagner 1932: 242; BDLC: s.v. Cors. *cala*), Lang. *câlo* 'shelter (from the wind)', Sav. *ϑalę* 'chalet (Alpine shepherd's hut)' etc. < ***kála*** (FEW II: 50-51).

Formal match: **good**

[307] I exclude Srd. *kála* in the meaning 'cove, inlet', which is likely a loan from Sp., Cat., or It. *cala* 'id.' (Wagner 1932: 242).

Semantic match: **good**
Cognacy: **compelling**
Distribution: Sardinia, Iberia, southern France, Corsica, Italy, Sicily

19. § 5.1.6: Srd. *mòɣoro*, *mòɣuru* 'hillock, low hill' < ***mókor***- (DES II: 122; Paulis 1987: 514); cf. Bq. *mokor* 'tree trunk, crust of bread, lump' < ***m-okor*** (Wagner 1932: 227; Hubschmid 1953: 49; cf. § 10.3.1.); Bq. ***mokoru*** (Gipuzkoa) 'height, summit' (Trask 2008: 276); Prov. *maguelo*, Lang. *maghêlo*, *maghiêlo* etc. 'steep hill, steep place' < ***maguella*** (FEW XXI: 15b); Alb. ***magulë*** 'small pile, heap', Rom. *măgură* 'mountain, hill', MGr. μαγούλα, PSl. **mogyla* 'burial mound' < ***magula*** (Wagner 1932: 228; Ciorănescu 1958: s.v. *măgură*; Orel 1998: 109).[308]

    Formal match: **good**, but excluding Bq. *mokor* 'tree trunk' and Cat. *mogró* 'nipple' (cf. § 10.3.1).
    Semantic match: **good**
    Cognacy: **compelling**
    Distribution: Sardinia, Iberia, southern France, Balkans, Greece

20. § 5.1.8: Srd. *mára* 'marsh, swamp, bog; sewer, drain' < ***mára**/*mágar*** (DES II: 71); cf. Lat. *mare* 'sea' < PIE ****mori***- (de Vaan 2008: 365); Cors. ***mara*** 'deviation of water to irrigate gardens' (Terracini 1927: 144; Wagner 1932: 241), Lad. *mar* 'swampy place', *marena* 'landslide caused by a spring of water', Ven. *marecia* 'marshy terrain', It. *marazzo* 'puddle, marsh' < ***mar***- (Terracini 1927; Hubschmid 1953: 63–65; cf. REW 5349); Laz. ***marana*** 'artificial waterway' (Wagner 1932: 241); Gr. ἀμάρα 'trench, channel' (Alessio 1944: 117–118; Hubschmid 1958: 223).

    Formal match: **possible**, but only if the Sardinian forms are not from **mágar*. The similarity to Lat. *mare* 'sea' < PIE **mori* is likely coincidental.
    Semantic match: **good**, for the various forms referring to swamp-like places and artificial waterways; thus excluding Lat. *mare* 'sea'.
    Cognacy: **possible**
    Distribution: Sardinia, Corsica, northern Italy, central Italy, Greece

[308] Old Cat. *mocorró*, Cat. *mogró* etc. 'nipple' have been derived by DCEC (IV: 184) from Lat. *mūcrō*, *-ōnis* 'sharp point'.

21. § 5.1.10: OSrd. *keia* 'pit (to store grain)', Srd. *kèya*, *kèa*, *čèa* 'valley; ditch, pit' < ***kèi̯a*** (DES I: 326); cf. Arag. *cea*, *citja* 'silo; underground cavity used for grain storage and to imprison slaves', Cat. *cía* 'chasm', *θía*, *síya*, *sía* 'grain silo' (DCEC II: 58-59), Cat. *sitja* 'pit covered with earth for storing grain', *sitges* (Mallorca) 'silos, hewn out in the rock inside houses', OProv. *setgia* (Narbonne), Lang. *siejo* (Montréal-de-l'Aude) 'grain silo' (DCEC II: 59; Beekes 2010: 615), Elban *čiálę* 'shelter under a protruding rock', Dauphinois *sialet* 'chasm; land collapse in high plateaus' (cf. FEW XXI: 18), Friul. *céole, cévole* 'chasm' < ***kĕi̯a***; Gr. Hes. καίατα 'pits (split by earthquakes)', Gr. καιάδᾱς 'pit or cavern in Sparta, into which convicted persons were thrown' etc. < ***kai̯u̯a-t/d-*** or ***kau̯i̯a-t/d-*** (DCEC II: 59; Beekes 2010: 615).

    Formal match: **good**, at least for the Romance comparanda. In the Greek forms, the preservation of intervocalic -ι- is unexpected.
    Semantic match: **good**
    Cognacy: **compelling**; the connection to the Greek forms is less certain.
    Distribution: Sardinia, Iberia, southern France, northern and central Italy, (maybe) Greece

22. § 5.1.15: Srd. *péntuma*, *ispéntumu* 'precipice, cliff, ravine' < ***péntum-*** (DES II: 244); cf. Etr. ***penθuna***, ***penθna*** 'stele, cippus' (Breyer 1993: 347 fn. 515); Laz. *pèndima* (Nemi) 'sloping rock face', Campan. *péndəm*[a] (Formicola), *péńǵima* (Ausonia) 'steep rock', Salent. *péntima* (Manduria, Taranto), *pèntuma* (Calimera) 'isolated boulder; cliff' etc. < ***péntVma*** (Parodi 1907: 772; Rohlfs 1927a: 163 fn. 1; Ribezzo 1933; Aebischer 1934). The Sardinian form may however be a descendant of a regional Latin formation (Parodi 1907: 772; Rohlfs 1927a: 163 fn. 1; Aebischer 1934).

    Formal match: **good**, for the Romance forms. The correspondence to the Etruscan form is uncertain; cf. § 10.4.4.
    Semantic match: **good**, excluding the Etruscan form.
    Cognacy: **possible**, for the Romance forms. We are possibly dealing with a regionally inherited Latin formation.
    Distribution: Sardinia, central Italy, southern Italy

23. § 5.1.17: Srd. *trókku* 'ravine, gorge' < ***tórku***/***trókku*** (DES II: 522); cf. Bq. *troka* 'ravine' < ***torka***, Sp. *torco*, *torca* (Navarre) 'chasm; unstable ter-

rain; pothole etc.' (Hubschmid 1953: 42; DCEC V: 521b–522). The Basque and Ibero-Romance forms are certainly related, but Corominas (DCEC V: 521b–522) considers a Latin etymology in *torques* 'collar', and argues that the Basque word is of recent origin.

Formal match: **good**
Semantic match: **good**
Cognacy: **possible**, but only to the Ibero-Romance forms. An inherited origin is not excluded; cf. § 10.3.1.
Distribution: Sardinia, Iberia

24. § 5.1.18: Srd. *marrárdzu* etc. 'rock, boulder', Sass. *marrágğu* 'big rock' < ***marr-árịu*** (DES II: 78); cf. Gasc. *marràlh* 'big stone', Lang. *marroc* 'stone block', Prov. *marro* 'tuff', Lomb. *maròɡn* 'boulder', Trent. *maròc* 'rock, boulder', Lad. *maròk* (Val di Fassa) 'scree; formless, crumbled rock matter etc.' < ***marr-*** (Hubschmid 1953: 52–57).

Formal match: **good**
Semantic match: **good**; i.e. the forms referring to various types of rocks. I am not fully convinced by Hubschmid's (1953: 52–57) comparisons to toponyms and words for hay stacks etc.
Cognacy: **compelling**
Distribution: Sardinia, southern France, northern Italy

25. § 5.1.19: Srd. *kalánka* 'rock crevice, precipice, cave' < ***kalánka*** (DES I: 266); cf. OProv. *calanca* 'narrow alley', Prov. *caranco* 'narrow bay where ships take shelter; deep spot along the sea or a river, where one can find fish', *kalãnko* 'rock crevice', *kanãnko* 'cavern', Cors. *calanca*, *calônca* 'sheltered place between mountains or hills; rocky inlet", Tusc. *calanco* 'rockslide', Calab. *kalankə* 'landslide' < ***kalánka*** (Rohlfs 1932: 139; FEW II: 56; DES I: 266; BDLC: s.v. Cors. *calanca*).

Formal match: **good**
Semantic match: **good**
Cognacy: **compelling**[309]
Distribution: Sardinia, southern France, Corsica, central Italy, southern Italy

[309] See § 5.1.19 however for the question whether this word is old in Sardinia.

26. § 5.1.25: Srd. *arròya* etc. 'downfall of water, stream, torrent, watery ravine etc.' < ***(ar)rói̯a*** (DES I: 127); cf. Lat. ***arrugia*** (only in Pliny) 'shaft in a Spanish gold mine'; Sp. *arroyo*, Pg. *arroio* 'stream, brook, creek' < ****arrŭgiu*** (Wagner 1940: 152; DCEC I: 359); Cat., Val. *roll* 'irrigation ditch', Cat. *roi* (Mallorca) 'jet, stream', Gasc. *arrouio*, *arrouil* 'gutter, rivulet', Bearn. *arroulho* 'irrigation ditch' etc. < ****arrugula*** (DCEC I: 359; FEW XXV: 343-44); Piem. *ruğa* (Valsesia), Lomb. *róğa* (Como), Ven. *roza*, Lad. *roia* (Val Gardena), *roa* (Val di Fassa), Friul. *roye* 'millrace' < ****rugi̯a*** (REW 678; Bertoldi 1932: 400; Hubschmid 1950: 34–35; DCEC I: 359); Gr. ὀρυγή 'digging' << PIE ****h₃reu-k-*** (Hubschmid 1953: 68–69; cf. Beekes 2010: 1113).

    Formal match: **good**, excluding the Greek form.
    Semantic match: **good**
    Cognacy: **compelling**
    Distribution: Sardinia, Iberia, southern France, northern Italy

27. § 5.1.28: Srd. *karróppu*, *garróppu*, *korróppu*, *gorróppu* etc. 'whirlpool; crevice' < ****garróppu*** (DES I: 308); cf. Irpino *sgarrupá*, Abruzz. *scarrupà* 'to fall down, to plunge', Laz. *carapone* (Castelmadama) 'whirlpool, swamp', Nap. *scarrupo*, *sgarupo* 'cliff, precipice', Calab. *garroppa* (Petronà) 'trough' < ****k/garropp-*** (Rohlfs 1932: 336; Wagner 1932: 243; DES I:308); Lang. *caraven* 'chasm, precipice', Bearn. *garimbáut* 'precipice, ravine, dangerous pass', South Tusc. *karravone*, Cors. *karravóni* 'ravine', Calab. *karafúnu* 'puddle' < ****karraf/v/b-*** (FEW XXI: 20; Hubschmid 1953: 58).

    Formal match: **good**
    Semantic match: **good**
    Cognacy: **compelling**, but cf. § 5.1.28.
    Distribution: Sardinia, Southern France, Corsica, central Italy, southern Italy

28. § 5.2.4: Srd. *moreðína*, *moretína* etc. 'pile of rocks, dry stone wall' < ****moretína*** << ****mora-*** (DES II: 127-28); cf. Bq. *more*(*a*), *mora* 'cairn, pile of rocks; dolmen' < ****mora*** (DCEC IV: 146); Cors. ***mora, mola*** 'pile (of hay, rocks, wood etc.)', Tusc. ***mora*** 'pile of rocks', Laz. ***mora*** (Castelmadama)

'boulder, very big rock' (DES II: 128); Sp. *morena*, Pg. *moreia* ´pile of plant material', Frpr. *morêna* (Savoy) 'glacier moraine' (DCEC IV: 146) < ***morena*** (cf. Hubschmid 1943: 150–151; Michelena 1969: 23–25; FEW VI/III: 276b).

Formal match: **good**
Semantic match: **good**
Cognacy: **compelling**
Distribution: Sardinia, Iberia, southern France, Corsica, central Italy

29. § 6.1.4: Srd. *paβáða*, *paβáði* etc. 'hornless (of goats)' (DES II: 200) < ***papáti***; cf. Calab. ***papáttsa***, Greco-Calab. ***popáttso***, Abruzz. ***kəkǫ́ćća***, ***pəkǫ́ttsa*** 'hornless' (Rohlfs 1932: 120; AIS 1082).

Formal match: **possible**, but not quite identical. Especially the Abruzzan forms are further removed.
Semantic match: **good**
Cognacy: **uncertain**; could these be independent sound-symbolic formations?

30. § 6.1.5: Srd. *tsúrra*, *čúrra* 'old, skinny sheep' < ***čúrra*** (DES II: 602); cf. Sp. ***churra*** 'type of sheep; type of wool' (DCEC II: 409-12); Lomb. ***ciorla*** (Como) 'cow of little value', Trent. ***ciorla*** 'skinny cow', Lad. ***čorla*** 'skinny cow', (modern) Umbrian ***ciorro*** 'ram', Abruzz., Nap., North Pugl. ***tsurrə*** 'he-goat', Calab. ***zurra*** 'hermaphrodite goat', MGr. τζοῦρος (Cyprus) 'he-goat' (Rohlfs 1925: 673; Jud 1937: 312; Hubschmid 1950c: 14–15; AIS 1069, 1080).

Formal match: **good**
Semantic match: **good**
Cognacy: **unlikely**. The Sardinian forms can be Spanish loans

31. § 6.1.6: Srd. *bòḍḍa* 'old sheep' < ***bólla/vólla*** (DES I: 215); cf. Calab. [*crapa*] ***uoḍḍa*** 'sheep whose horns curve behind its ears' (DES I:215; Rohlfs 1932: 351).[310]

[310] There is also the possible comparison of Bb. **wV̌llī*/**wV̌llē* 'small lifestock (i.e. sheep and goats)' and Bb. **tē-ḇălē* 'sheep', for which cf. § 10.2.1.

Formal match: **possible**
Semantic match: **possible**; both words denote sheep, but with different attributes.
Cognacy: **possible**, but difficult to base on just two geographically very limited forms.
Distribution: Sardinia, southern Italy

32. § 6.1.8: Srd. *tèle*, *tevèle* etc. 'tilled land' < ***t(ev)èle*** (DES II: 481); cf. It. *debbio* 'slash and burn' < ***deb(V)lo***, according to Paulis both from ancient Ligurian **debelo* < PIE **d*^{h}*eg*^{wh}- 'to burn', but this language is not attested.

    Formal match: **good**
    Semantic match: **good**
    Cognacy: **compelling**, but not the ancient Ligurian etymology.

33. § 07.2.4: Srd. *barraśòne*, *barišòne*, *barraśólu* etc. 'tangle/bundle of wood (generally to block passages)' < ***barrVs***- (DES I: 181-82); cf. Bq. ***barra*** 'bar' (Rohlfs 1927b: 402; DCEC I: 510b, DECLC I: 645b); Sp., Pg. *barda* 'fence made of intertwined branches and brambles', Cat. *barda* 'portable barrier to lock up lifestock', *bardissa* 'thicket; hedge' < ***barda*** (DCEC I: 510; DECLC I: 644-45); OProv. *barta* 'bush, brushwood' etc. < ***barră/ĭt***- (FEW I: 262). However, Bq. *barra* may be of Romance origin (Michelena 1990: 269).

    Formal match: **possible**
    Semantic match: **good**
    Cognacy: **compelling**, but the Basque form may be a Romance loan (Michelena 1990: 269).
    Distribution: Sardinia, Iberia, southern France

11.2 Discussion

On the basis of the proposed pre-Roman comparanda, we can now evaluate whether the geographical distribution of the evidence can shed light on the pre-Roman linguistic configuration of the Mediterranean. If multiple cognates of pre-Roman Sardinian words are consistently found in a geographically coherent pattern, this could be evidence of the existence of a single language or language

family in that area, or of prolonged regional language contact. Table 11.2.1 provides an overview of the Sardinian words with proposed pre-Roman comparanda in the Mediterranean. In each case, the likelihood of the comparison is indicated, along with the distribution of the accepted comparanda.[311]

	Sardinian	Cogn.	Ib.	SFr.	Cors.	NIt.	CIt.	SIt., Sic.	Aeg., Balk.
1	**ϑaϑarói̯* 'arum'	?						?	?
2	**bull-úri/-Vr-* 'poison hemlock'	?					?		
3	**(ϑ-)ánda* 'poppy'	✓		✓					
4	**al-as-* 'butcher's broom, holly'	✓(?)		✓		✓		✓	
5	**golósti* 'holly'	✓	✓	✓					✓
6	**kòsti* 'maple'	✓	?	✓			✓		?
7	**mutéklu* 'rockrose'	✓	✓	✓			✓	✓	
8	**olidóne* 'strawberry tree'	✓		✓		✓	✓ (?)		
9	**kárva/kárba* 'tree branch'	✓	✓	✓			✓		
10	**bòllVro* 'juniper berry'	✓					✓	✓	
11	**kòkkoro* 'walnut'	?	✓			✓	✓	✓	
12	**i̯ákaru* 'hunting dog'	?	✓		✓				✓
13	**kali̯-* 'puppy'	?							?
14	**bett-*, **bitt-* 'fawn, lamb'	?	✓	✓					
15	**leppVr-* 'hare'	✓	✓	✓			✓	✓	
16	**mufróne* 'mouflon'	✓			✓		?		
17	**núrra* 'chasm, shaft; pile'	×							
18	**kála* 'animal den'	✓	✓	✓	✓		✓	✓	
19	**mókor-* 'hillock'	✓	✓ (?)	✓					✓
20	**mára* 'marsh; drain'	✓ (?)			✓	✓	✓		✓
21	**kéi̯a* 'valley; ditch, pit'	✓	✓	✓		✓	✓		?
22	**péntum-* 'cliff, ravine'	?					✓	✓	
23	**tórku/trókku* 'ravine, gorge'	?	✓						
24	**marr-ári̯u* 'rock'	✓		✓		✓			
25	**kalánka* 'rock crevice, preci-	✓		✓	✓		✓	✓	

[311] The abbreviations in Table 11.2.1 refer to geographical rather than linguistic distribution. For instance, CIt. (Central Italy) includes not only the Central Italian branch of Romance, but also Tuscan varieties, Latin, and Etruscan. Similarly, the single Siculian comparandum is subsumed under Sic. (Sicily). The extents of these areas are specified in § 11.2.1.

	Sardinian	Cogn.	Ib.	SFr.	Cors.	NIt.	CIt.	SIt., Sic.	Aeg., Balk.
	pice'								
26	**arróia* 'stream, watery ravine'	✓	✓	✓		✓			
27	**garróppu* 'whirlpool; crevice'	✓		✓	✓		✓	✓	
28	**moretína* 'pile of rocks; dry stone wall'	✓	✓	✓	✓		✓		
29	**papáti* 'hornless'	?						?	
30	**čúrra* 'old sheep'	×							
31	**bólla/vólla* 'old sheep'	?						?	
32	**t(ev)èle* 'tilled land'	✓					✓		
33	**barrVs-* 'bundle of wood'	✓	✓	✓					

Table 11.2.1 Distribution of proposed pre-Roman comparanda

Twenty-one of the thirty-three discussed sets of words are sufficiently close, formally and semantically, to posit a shared pre-Roman origin. Some items will be excluded from the following discussion: 'arum' (1), 'poison hemlock' (2), 'walnut' (11), 'dog' (12), 'puppy' (13), 'fawn, lamb' (14), 'chasm, shaft; pile' (17), 'cliff' (22), and 'old sheep' (30, 31). These comparisons suffer from considerable formal or semantic difficulties, rendering their connection either implausible, or difficult to verify.

11.2.1 Geographic distributions

In the following, the accepted comparanda to Sardinian non-inherited words are discussed according to the region in which they are attested. From west to east, these are the Iberian Peninsula, Southern France, Corsica, Northern Italy, Central Italy, Southern Italy and Sicily, and the Aegean and the Balkans.

11.2.1.1 Iberian Peninsula

Table 11.2.1 contains no fewer than eleven lexical comparisons of plausible pre-Roman origin (or twelve, if Bq. *gastigar* is included) between Sardinian and the languages of the Iberian Peninsula. Interestingly, Basque participates in only five of these comparisons, viz. the words for 'holly' (5), 'branch/stick etc.' (9), 'hillock, low hill' (19), 'pile of rocks' (28), and perhaps 'bundle of wood etc.' (33). Of these, the Basque forms *karbasta*, *garbazta* 'stick with branches' (9) have been argued to be a more recent formation in Basque (Lakarra et al. 2019: 3–8), and *barra* 'bar' (33) may be a loan from Romance (Michelena 1990: 269). The Basque forms *mokoru* 'height, summit' (19) and *mora* 'pile of rocks, dolmen' (28) cannot be native

either, as the phoneme **m* is not reconstructed for Proto-Basque (cf. Trask 1995: 67). As already mentioned in § 10.3.3, this suggests that the pre-Roman language(s) of Sardinia were not related to Basque.

In absence of a direct link with Basque, and given the more extensive number of ostensibly pre-Roman lexical correspondences between Sardinian and the Ibero-Romance languages, the conclusion must be that both Sardinian and Basque acquired lexical material from a third language. The distribution of the Ibero-Romance forms is not uniform. Some words are attested in the north and extending west to Portuguese, viz. 'branch etc.' (9), 'hare' (15), 'pile of rocks' (28). Others are concentrated around the Pyrenees and the east coast, viz. 'holly' (5), 'rockrose' (7), 'cove, creek' (18), 'ditch; pit' (21). Yet others have a more general distributio, viz. 'stream, torrent, channel' (26), 'tangled wood used as a barrier' (33)'. Several processes may have governed this distribution. A single group of languages may have spread across a wide area in the Iberian Peninsula, whose individual lexical items were preserved in some of the modern Romance varieties, but not in others. Another possibility is that originally pre-Roman words entered specific Romance varieties and from there diffused into neighboring Romance varieties (cf. Dworkin 2012: 27–28). Whatever the case may be, an important observation is that the majority of Iberian-Sardinian correspondences are represented in the northeastern languages of the peninsula, be it in Basque, Aragonese or Catalan. In terms of pre-Roman languages, such a distribution closely fits the extent of the extinct non-Indo-European Iberian language, as delineated by inscriptional evidence (Untermann 1983: Map 1; Dworkin 2012: 19). In conclusion, if both Sardinian and Basque borrowed from a shared donor language, a plausible candidate for such a language would be Iberian (cf. § 11.3.1).

The following is an overview of the Sardinian forms with accepted comparanda on the Iberian Peninsula, discussed in § 11.1, § 11.2. The numbering is that of § 11.1.

	Srd.	**Ib.-Rom.**	**Bq./other**
5	**golósti* 'holly'	Navarr. *colostia, golostia* 'id.'	Bq. *gorosti* 'id.'
6	**kòsti* 'maple'		Bq. *gastigar* etc. 'id.' (?)
7	**mutéklu* 'rockrose'	Cat. *módega, motxera* 'narrow-leaved cistus'	
9	**kárva*/*kárba* 'tree branch'	Ast. *garbu* 'small firewood', Pg. *gravulha* 'dry pine-needles', Gal. *garabullos* 'dry tips of branches used as firewood' etc.	Bq. *karbasta, garbazta* (?)
15	**léppVr-* 'hare'	Pg. *láparo* 'young rabbit'	Turdetanian

	Srd.	Ib.-Rom.	Bq./other
			λεβηρίς 'hare'
18	**kála* 'animal den'	Sp., Cat. *cala* 'cove, inlet, creek'	
19	**mókor-* 'hillock, low hill'		Bq. *mokoru* 'height, summit'
21	**kéi̯a* 'valley; ditch, pit (to store grain)'	Arag. *cea*, *citja* 'silo; underground cavity used for grain storage and to imprison slaves'; Cat. *cía* 'chasm', *síya*, *sía* 'grain silo', *sitja* 'pit covered with earth for storing grain, *sitges* (Mallorca) 'silos, hewn out in the rock inside houses'	
23	**tórku*/**trókku* 'ravine'	Sp. *torco*, *torca* (Navarre) 'chasm; unstable terrain; pothole etc.'	
26	**(ar)rói̯a* 'downfall of water, stream, torrent, watery ravine etc.'	Sp. *arroyo*, Pg. *orroio* 'stream, brook, creek'	Lat. *arrugia* 'shaft in a Spanish goldmine'
28	**moretína* << **mora-* 'pile of rocks; dry stone wall'	Sp. *morena*, Pg. *moreia* 'pile of plant material'	Bq. *more(a)*, *mora* 'cairn, pile of rocks; dolmen'
33	**barrVs-* 'tangle of wood (to block passages)'	Pg. *barda* 'fence made of intertwined branches and brambles', Cat. *barda* 'portable barrier to lock up lifestock', *bardissa* 'thicket, hedge'	Bq. *barra* 'bar' (?)

Table 11.2.2: Pre-Roman Sardinian comparanda in the Iberian Peninsula

11.2.1.2 Southern France

Intriguingly, the number of acceptable external comparanda to pre-Roman Sardinian words peaks in the area that is now southern France. Seventeen of the comparisons between Sardinian and Occitan forms discussed in § 11.1, i.e. approximately half of the total number, withstand formal and semantic scrutiny. They are therefore likely to represent cognate forms. This raises the question as to what the source language of these lexical items might have been. Pre-Roman Gaul is usually taken to have been predominantly Celtic-speaking (e.g. Polomé 1983: 527–528). None of the forms in question have a plausible Celtic etymology however. Nor is there evidence for Celtic ever having reached Sardinia. Non-Indo-European languages attested in (southern) Gaul are Aquitanian, Iberian, and Etruscan. Aquitanian, the ancestor to modern day Basque, was native to the Atlantic coast in the southwest of modern day France (Gorrochategui 1995: 57; Trask 1995: 87). Iberian was spoken in what is now the Languedoc region, up

to modern day Lattes (Untermann 1980: 36, 383–384). Finally, Etruscan was also spoken in Lattes, as well as at Pech Maho (Wallace 2008: 3; Dietler 2010: 97; Mullen 2013: 30). Whereas the presence of Etruscan was likely the result of Etruscan traders settling in the area relatively recently (cf. Bats 1998; Wallace 2008: 3), Iberian appears to have been the prevalent language, at least in the southwesternmost parts of what is now the French Mediterranean coast (Untermann 1980: 36).[312]

For purposes of this study, it is important to note that there is an evident overlap between Sardinian-Iberian correspondences and Sardinian-French correspondences. Eleven words of pre-Roman origin are shared between Sardinia, the Iberian Peninsula, and southern France.[313] This sizeable number suggests that whatever language or language group was the source of the shared pre-Roman vocabulary of Sardinia and the Iberian Peninsula, was also present in southern France. Looking at the geographical distribution of the Gallo-Romance comparanda, we can further observe that most are attested in Gascon and Lengadocian, north of the Pyrenees. As in the case of the comparanda in the Iberian Peninsula, this distribution is consistent with ancient Iberian being the donor language of these forms in Gallo-Romance.

The following is an overview of the Sardinian forms with likely comparanda in southern France, discussed in § 11.1, 11.2. The numbering is that of § 11.1.

	Srd.	**Gall.-Rom.**	**Other**
3	**(ϑ-)ánda-* 'poppy'	Lang. *ãndèr*, *õndèr* (dep. Lot) 'id.'	
4	**al-ás-* 'butcher's broom, holly'	Bearn. *aroû*, Prov. *ářę* 'broom', Auv. *alète* < **alista*. Lang. *orousto* < **alusta* 'common broom'	
5	**golósti* 'holly'	Lang. *agaloussé* 'id.'	
6	**kòsti* 'maple'	OProv. *agast* 'id.'	
7	**muték(u)lu* 'rockrose'	Lang. *mouche*, *mouge* 'id.'	
8	**olidóne* 'strawberry tree'	Gasc. *auledun*, Saintongeais *olon*, Lang. *leduno* 'id.'	
9	**kárva*/**kárba* 'tree	Gasc. *garbalhe* 'pine needles', Prov. *garbo*	

[312] Cf. Mullen (2013: 26–27) on the potential presence of a "Ligurian" language southeast of Massalia. This language is not attested, and its existence is largely based on mentions by classical authors. The label "Ligurian" has been used to account for various proposed substrate phenomena (§ 1.3.2.1).

[313] That is, excluding the comparison of Srd. *kóstike*, OProv. *agast*, Bq. *gaztigar* 'maple', due to uncertainties regarding the Basque form (cf. § 10.3.1).

	Srd.	Gall.-Rom.	Other
	branch'	'hollow tree trunk' etc.	
15	**léppVr-* 'hare'	OProv. *laperel* 'young rabbit', OFr. *lapriel* 'rabbit'	Massiliot Gr. λεβηρίς 'rabbit'
18	**kála* 'animal den'	OProv. *cala* 'cove, inlet, creek', Lang. *câlo* 'shelter (from the wind)', Sav. *ϑalę* 'chalet' etc.	
19	**mókor-* 'hillock, low hill'	Lang. *maghêlo*, *maghiêlo* 'steep hill, steep place'	
21	**kéi̯a* 'valley; ditch, pit (to store grain)'	OProv. *setgia* (Narbonne), Lang. *siejo* (Montréal-de-l'Aude) 'grain silo'; Dauphinois *sialet* 'chasm; land collapse in high plateaus'	
24	**marr-árḭu* 'rock'	Gasc. *marràlh* 'big stone', Lang. *marroc* 'stone block', Prov. *marro* 'tuff'	
25	**kalánka* 'rock crevice, precipice, cave'	OProv. *calanca* 'narrow alley', Prov. *caranco* 'narrow bay where ships take shelter; deep spot along the sea or a river, where one can find fish', *kalãnko* 'rock crevice', *kanãnko* 'cavern'	
26	**(ar)rói̯a* 'downfall of water, stream, torrent, watery ravine etc.'	Gasc. *arrouio, arrouil* 'gutter, rivulet', Bearn. *arroulho* 'irrigation ditch'	
27	**garróppu* 'whirlpool, crevice'	Lang. *caraven* 'chasm, precipice', Bearn. *garimbáut* 'precipice, ravine, dangerous pass'	
28	**moretína* << **mora-* 'pile of rocks; dry stone wall'	Frpr. *morêna* 'glacier moraine'	
33	**barrVs-* 'bundle of wood'	OProv. *barta* 'bush, brushwood'	

Table 11.2.3: Pre-Roman Sardinian comparanda in southern France

11.2.1.3 Corsica

Given the geographical proximity of Corsica to Sardinia, and the cultural connections between the two islands in history and prehistory, it comes as somewhat of a surprise that only seven of the comparisons discussed in § 11.1 are represented in Corsican, one of which is uncertain: viz. the word for 'dog' (12). With the exception of the word for 'mouflon' (16), which refers to an animal native to

the two islands, all comparisons have additional representatives in the Iberian Peninsula, France (12, 18, 25, 27, 28), or Italy (18, 25, 27, 28). This suggests that, even though Sardinia and Corsica likely were exposed to similar pre-Roman linguistic influences, there is no evidence for a closer linguistic affiliation between the pre-Roman languages of the two islands than between either of them and other parts of the western Mediterranean. One possible explanation for this is that Corsica was more thoroughly romanized than Sardinia. Another reason could be that Corsica was heavily tuscanized in the Middle Ages (Nesi 1988: 801; Ledgeway 2016: 208), during which native vocabulary may have been lost.

The following is an overview of the Sardinian forms with likely comparanda in Corsica, discussed in § 11.1, 11.2. The numbering is that of § 11.1.[314]

	Srd.	**Cors.**
12	**ɟákaru* 'hunting dog'	Cors. *ghjácaru* 'shepherd dog' (?)
16	**mufròne* 'mouflon'	Cors. *muvròne* 'id.'
18	**kála* 'animal den'	Cors. *cala* 'cove, inlet, creek'
20	**mára* 'marsh, swamp, bog; sewer, drain'	Cors. *mara* 'deviation of water to irrigate gardens'
25	**kalánka* 'rock crevice, precipice, cave'	Cors. *calanca* 'sheltered place between mountains or hills; rocky inlet'
27	**garróppu* 'whirlpool; crevice'	Cors. *karravóni* 'ravine'
28	**moretína* << **mora-* 'pile of rocks; dry stone wall'	Cors. *mora, mola* 'pile (of hay, rocks, wood etc.)

Table 11.2.4: Pre-Roman Sardinian comparanda in Corsica

11.2.1.4 Northern Italy

In northern Italy, here defined as all Italian regions north of Tuscany and Marche, we find six correspondences to Sardinian words of probable pre-Roman origin. All six of these are also found in southern France, and four have correspondences further south in Italy (4, 8, 20, 21).

The following is an overview of the Sardinian forms with likely comparanda in northern Italy, discussed in § 11.1, 11.2. The numbering is that of § 11.1.[315]

	Srd.	**NIt.**
4	**al-ás-* 'butcher's broom; holly'	Lig. *lastra, arástra* 'spiny broom'
8	**olidóne* 'strawberry tree'	Pied. *(l)ürion* 'id.'

[314] Srd. *mára* (20) 'only belongs here if it is not ultimately from **mágar* (cf. § 5.1.8).

[315] See fn. 314.

20	**mára* 'marsh, swamp, bog; sewer, drain'	Lad. *mar* 'swampy place', *marena* 'landslide', Ven. *marecia* 'marshy terrain'
21	**kéi̯a* 'valley; ditch, pit (to store grain)'	Friul. *céole, cévole* 'chasm'
24	**marr-ari̯u* 'rock'	Lomb. *maròɡn* 'boulder', Trent. *maròc*, Lad. *maròk* 'scree'
26	**(ar)rói̯a* 'downfall of water, stream, torrent, watery ravine etc.'	Piem. *ruǧa*, Lomb. *róǧa*, Ven. *Roza*, Lad. *roia, roa*, Friul. *roye* 'millrace'

Table 11.2.5: Pre-Roman Sardinian comparanda in northern Italy

11.2.1.5 Central Italy

The lexical evidence from Central Italy, i.e. Tuscany, Lazio, Umbria and Marche, can be divided into two categories. The first comprises the evidence from the ancient attested languages of the region, Etruscan and Latin. The second is the evidence from the local modern Romance varieties spoken in this area, including the Italian standard language, which is mainly based on Tuscan. The comparanda from Etruscan and Latin (7, 8, 9, 15) are all attested in southern France and the Iberian Peninsula as well, the sole exception being "strawberry tree" (8), which is absent from the latter. Of the modern comparanda, five reappear in southern France (6, 18, 21, 25, 27) and three in the Iberian Peninsula (18, 21, 28).

The distribution of the comparanda raises crucial questions on the linguistic relationships between pre-Roman Italy and other parts of the Mediterranean. Etruscan, being the most conspicuous pre-Roman language of Italy, has been invoked to account for various non-inherited elements in Latin and in the modern linguistic varieties of the region, most notably Tuscan (cf. § 10.4). The fact that a sizeable amount of these non-inherited elements are also found elsewhere in the Mediterranean, suggests the existence of linguistic ties between Italy, southern France, the Iberian Peninsula, and Sardinia. The questions and implications that ensue from these lexical ties, are discussed in § 11.4.

The following is an overview of the Sardinian forms with likely comparanda in northern Italy, discussed in § 11.1, 11.2. The numbering is that of § 11.1.[316]

	Srd.	**Lat./Etr.**	**CIt. Romance**
6	**kòsti* 'maple'		Tusc. *còstolo* 'ash tree'
7	**muték(u)lu* 'rockrose'	Etr. μούτουκα 'thyme; rockrose'	
8	**olidóne* 'strawberry tree'	Lat. *ūnedō* 'id.' (?)	

[316] See fn. 314.

	Srd.	Lat./Etr.	CIt. Romance
9	**kárva*/*kárba* 'tree branch'	Lat. *carbō*, *-ōnis* 'piece of charcoal'	
10	**bòllVro* 'juniper berry'	Tusc. *bòllora* 'gall'	
15	**léppVr-* 'hare'	Lat. *lepus*, *-oris* 'id.'	
18	**kála* 'animal den'		It. *cala* 'cove, inlet, creek'
20	**mára* 'marsh, swamp, bog; sewer, drain'		It. *marazzo* 'puddle, marsh', Laz. *marana* 'artificial waterway'
21	**kéi̯a* 'valley; ditch, pit (to store grain)'		Elban *čiálę* 'shelter under a protruding rock'
25	**kalánka* 'rock crevice, precipice, cave'		Tusc. *calanco* 'rockslide'
27	**garróppu* 'whirlpool; crevice'		South Tusc. *karravone* 'ravine'
28	**moretína* << **mora-* 'pile of rocks; dry stone wall'		Tusc. *mora* 'pile of rocks', Laz. *mora* 'boulder, big rocks'

Table 11.2.6: Pre-Roman Sardinian comparanda in central Italy

11.2.1.6 Southern Italy and Sicily

The distribution of Sardinian comparanda with attestations in southern Italy and Sicily is largely comparable to that in northern and central Italy, with six out of nine forms having a correspondent in southern France (viz. 4, 5, 7, 15, 25, 27), and four in the Iberian Peninsula (viz. 5, 7, 15, 18). Especially noteworthy is the exact lexical match between Sic. *alastra* 'spiny broom' and Lig. *lástra, arástra* 'id.' (cf. Bertoldi 1928: 233 fn. 3; Battisti 1960: 373). The correspondences involving central Italy, especially that of Calab. *mútaca* 'rockrose' to Etr. μούτουκα 'thyme, rockrose', would suggest some connection to Etruscan, even though that language has not been attested so far to the south (Wallace 2008: 2–3).

The following is an overview of the Sardinian forms with likely comparanda in southern Italy and Sicily, discussed in § 11.1, 11.2. The numbering is that of § 11.1.

	Srd.	SIt.	Sic.
4	**al-ás-* 'butcher's broom; holly'		Sic. *alastra* 'spiny broom'
5	**golósti* 'holly'	Lucanian *cèlastru*, Calab. *cègghiastru*, *cìstriddazzu*, Bov. *azzìddastro* 'holly'	

	Srd.	SIt.	Sic.
7	**muték(u)lu* 'rockrose'	Calab. *mútaca*, *múndaca*, *amùndaci*, *amùndici* 'narrow-leaved cistus'	Sic. *bóḍḍaru* 'grain, kernel; bump'
10	**bòllVro* 'juniper berry'		Sic. *bóḍḍaru* 'grain, kernel; bump'
15	**léppVr-* 'hare'		Siculian λέποριν 'id.'
18	**kála* 'animal den'		Sic. *cala* 'cove, inlet, creek'
25	**kalánka* 'rock crevice, precipice, cave'	Calab. *kalankə* 'landslide'	
27	**garróppu* 'whirlpool; crevice'	Calab. *kalafúnu* 'puddle'	
29	**papáti*	Calab. *papáttsa*, Greco-Calab. *popáttso*, Abruzz. *kəkǫ́ćća, pəkǫ́ttsa* 'hornless'	

Table 11.2.7: Pre-Roman Sardinian comparanda in southern Italy and Sicily

11.2.1.7 Aegean and the Balkans

In Greek, we find multiple forms that correspond, both formally and semantically, to forms in the western Mediterranean. The most conspicuous among these forms is Gr. κήλαστρος 'holly' (5), which corresponds to Srd. *golóstɨu*, Bq. *gorosti*, Lang. *agaloûssës* 'id.' etc. However, Gr. κήλαστρος 'holly' is only attested in Theophrastus (1916: chap. 9:3; Hofmann 1949: 142), and in Hesychius (2020: 596) as κήλαστραι 'trees'. It does appear to be continued in southern Italy by Lucanian *cèlastru*, Calab. *cègghiastru*, Bov. *azzìḍḍastro* etc. 'holly' (DEI II: 842; Rohlfs 1977: 104, 110, 156, 694, 808). It cannot fully be ruled out that Gr. κήλαστρος 'holly' was in fact only used in Greek varieties of Magna Graecia (cf. § 3.3.7). Similarly, the form ἄκαστος is found in Hesychius, glossed as 'maple', but he offers no information on the provenance of this form. Alessio's (1949: 132) suggestion that ἄκαστος was a Massiliote Greek form borrowed from a local language seems speculative but is supported by OProv. *agast* 'maple'. It is therefore conceivable that this word was never found in the Aegean. A connection between Srd. **kèɨ̯a* 'pit (to store grain)' (21) and Gr. καιάδᾱς 'pit or cavern in Sparta, into which convicted persons were thrown' is not certain, because the latter requires earlier **kaɨ̯u̯a-t/d-* or **kau̯ɨ̯a-t/d-*. The presence of the putative correspondents to Srd. **mokor-* 'hillock' in Albanian, Romanian, Proto-Slavic and modern Greek, suggests a more recent introduction of this word into Greece, perhaps from Slavic

**mogyla* 'burial mound', which is itself of unclear origin (Vasmer 1955: 143–144).[317] Since this particular case is the only potential evidence for substrate correspondences between Sardinia and the languages of the Balkans, other than Greek, it is possible that we are dealing with a chance correspondence.

As a result, the number of Greek-Sardinian correspondences is rather limited, especially in light of the uncertain provenance of some of the Greek forms, and of the large number of Greek words of suspected pre-Indo-European origin (cf. Beekes 2014).

The following is an overview of the Sardinian forms with likely comparanda in the Aegean region or the Balkans, discussed in § 11.1, 11.2. The numbering is that of § 11.1.[318]

	Srd.	**Gr.**	**Other**
5	**golósti* 'holly'	Gr. κήλαστρος 'id.'	
6	**kòsti* 'maple'	Gr. Hes. ἄκαστος 'id.'	
19	**mókor-* 'hillock, low hill'	MGr. μαγούλα 'burial mound'	Alb. *magulë* 'small pile, heap'; Rom. *măgură* 'mountain'; PSl. **mogyla* 'burial mound'
20	**mára* 'marsh, swamp, bog; sewer, drain'	Gr. ἀμάρα 'trench, channel'	
21	Srd. **kéi̯a* 'valley; ditch, pit (to store grain)'	Gr. καιάδᾱς 'pit or cavern in Sparta, into which convicted persons were thrown'	

Table 11.2.8: Pre-Roman Sardinian comparanda in the Aegean and Balkans

11.3 Interpretation of the observed distributions

In § 11.1, the various proposals for cognates of pre-Roman Sardinian words have been analyzed. For twenty-one out of the thirty-three proposed comparisons a good case can be made that they are indeed of a shared pre-Roman origin.[319] The following is an overview of the number of accepted comparanda between Sardinia and each of the discussed regions, in descending order:

[317] Note the agreement between the Greek and Slavic forms in the meaning 'burial mound'.

[318] See fn. 314.

[319] Some less certain cases, e.g. those that have been proposed to be the result of sound-symbolic processes, are not included in this number.

- Southern France: 17
- Central Italy: 12
- Iberian Peninsula: 11 or 12
- South Italy: 9
- Corsica: 7
- North Italy: 6
- Aegean and the Balkans: 5 (but some uncertainties)

11.3.1 Iberian in Sardinia?

The most striking observation that can be made on the basis of the geographical distribution of the pre-Roman lexicon shared between Sardinia and other regions, is that the vast majority entail comparanda in Southern France. This is all the more surprising, since this region is not home to any of the attested languages that have previously been compared to the pre-Roman Sardinian vocabulary (cf. § 10). Not much lower in number are the comparisons with comparanda in the Iberian Peninsula and Central Italy. Of the twelve Iberian-Sardinian correspondences, ten are shared with Southern France, underlining the close linguistic ties between these three regions. The relevant forms in the Iberian Peninsula are concentrated in the northeast, in Basque, Aragonese and Catalan, although some are also found further west in Castilian, Asturian and Portuguese. In France, the evidence compared to Sardinian clusters in the Languedoc, northeast of the Pyrenees, although outliers are also found farther north and east in other Gallo-Romance varieties. The fact that a considerable amount of pre-Roman Sardinian words have correspondences in a coherent and well-defined geographical context outside Sardinia (Figure 11.3.1), appears crucial. It suggests that a language spoken in northeastern Spain and southern France was related to a language once spoken in Sardinia.

As mentioned in § 11.2.1.2, the distribution of the pre-Roman lexical correspondences between Sardinia, the Iberian Peninsula, and France, overlaps with the known extent of the ancient Iberian language, which is attested from the southeastern coast of Spain to the Languedoc region in France (see § 1.3.2.2). An Iberian origin of the discussed pre-Roman words could offer an explanation of the presence of the Basque correspondents of pre-Roman Sardinian words, that cannot have been the result of a Vasconic presence in Sardinia (cf. § 10.3.3). Iberian was not related to Basque, but both languages share features, probably as a result of convergence due to intensive contact (Gorrochategui 1995; Trask 1995: 79).

There are several considerations that shed some doubt on Iberian as a source language for pre-Roman words in Sardinia, Spain and France. One is that the evidence for the existence of a phoneme /m/ in Iberian is scarce (Velaza 2019: 172–173; Moncunill & Velaza 2020: 598), similar to the non-existence of Proto-Basque **m* (Michelena 1990: 267). The phoneme /m/ does however occur in a number of pre-Roman correspondences found in Sardinia, the Iberian Peninsula, and France: e.g. Srd. *muteclu*, Cat. *mòdaga*, Lang. *mouche* 'rockrose'; Srd. *mòɣoro* 'hillock', Bq. *mokoru* 'height, summit', Lang. *maghêlo* 'steep hill'. Another possible objection to Iberian as a lexical donor is that Iberian may have expanded to its observed range not too long before the first attestations (Velaza 2019: 162). De Hoz (2009; 2011: 56–57) argues that in the northeast, Iberian was only used as a lingua franca, besides vernacular languages of a different origin. Additional evidence for a non-Iberian language in northeastern Spain and southwestern France are the non-Iberian-looking personal names attested in Iberian inscriptions in precisely this area (Gorrochategui & Vallejo 2019: 360), which have been related to onomastic practices in the "Ligurian" area (Untermann 1969: 8–9).

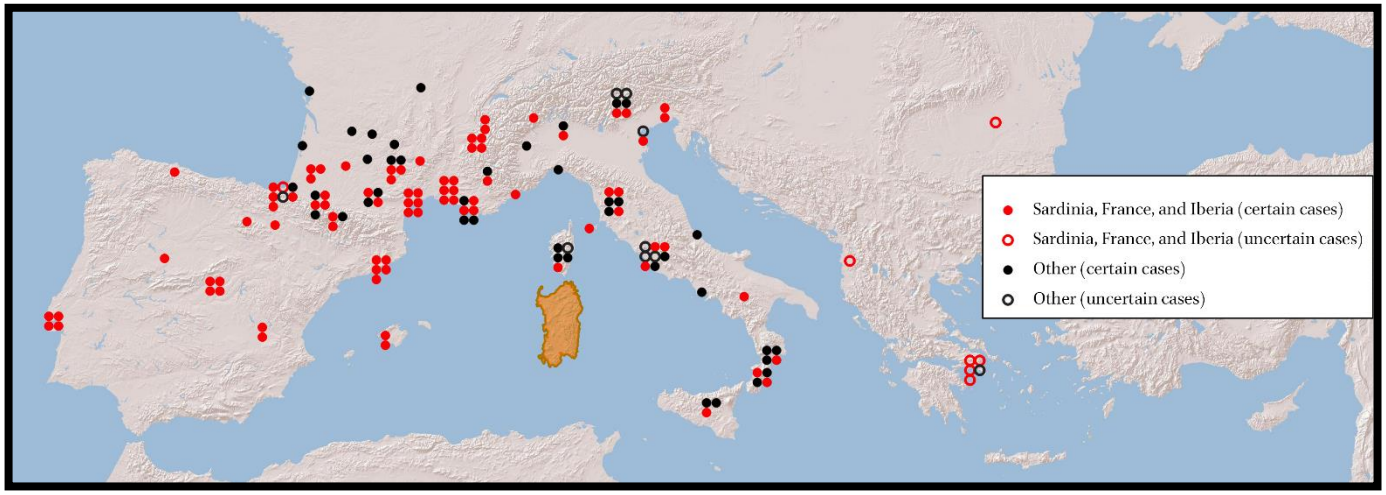

Figure 11.3.1: Distribution of pre-Roman words shared between Sardinia, Iberia, and France

The distributions of pre-Roman lexical correspondences do in any case show that some pre-Indo-European language was present both in Sardinia, and in the Iberian epigraphic area. The lexical correspondences between Sardinian, Basque, and other Romance languages in the region likely represent independent borrowings from this language. Whether this source language was Iberian itself, or some other pre-Indo-European language dominated or supplanted by Iberian, is difficult to establish based on the current evidence. In any case, these findings, based on the lexical data, corroborate the Sardinian-Iberian connection proposed by Paulis (2008) on the basis of structural and onomastic similarities.

11.3.2 Etruscan in Sardinia?

At a shared second place in order of most pre-Roman lexical correspondences to Sardinian is the region of Central Italy, with twelve accepted correspondences found in modern and ancient languages spoken in Etruria and Latium (modern day Tuscany and Lazio).[320] This may lend additional strength to Pittau's (1995) hypothesis that a language related to Etruscan was once spoken in Sardinia (cf. § 10.4), since the distribution of this rather large number of correspondences corresponds well to that of this pre-Indo-European language. It is important to note that, due to the fact that only a single Sardinian word has a convincing cognate in an actual Etruscan form (§ 10.4.1, 11.2.1.5), a possible connection between pre-Roman Sardinian and Etruscan is still based more on geographical distribution than on direct linguistic evidence — as in the case of Iberian (§ 11.3.1). Interestingly, nine of the Sardinian-Central Italian comparisons are shared with Southern France (Figure 11.3.2). This raises the suspicion that the pre-Roman linguistic connections between Sardinia and Italy cannot be seen separately from those between Sardinia, Iberia, and France (cf. § 11.4).[321]

[320] Many of these have not been discussed by Pittau (1995).

[321] Cf. Alessio (1949: 114) for a similar observation.

Figure 11.3.2: Distribution of pre-Roman words shared between Sardinia, central Italy, and France

11.3.3 Pre-Greek in Sardinia?

Whereas most accepted lexical correspondences are found in the western Mediterranean, there are five words with a more eastern distribution, all of which have a representative in Greek. The respective Sardinian and Greek forms are phonologically or semantically too dissimilar to accept direct borrowings from Byzantine Greek into Sardinian, suggesting that they must have come from a third source. The number of six correspondences is rather low in light of the large number of substrate words preserved both in Sardinian and in Greek, and the provenance of some of the Greek comparanda is unclear (e.g. ἄκαστος 'maple', κήλαστρος 'holly'; cf. § 11.2.1.7). Nor does the Sardinian pre-Roman lexicon exhibit some of the most notable morphological features present in the pre-Greek lexicon, such as the -ινθος-suffix (cf. Kroonen 2024b). Because of the uncertainty regarding the "Greek-ness" of some of the forms, there is no sufficient evidence to posit a close relationship between any pre-Roman Sardinian language and pre-Greek. The similarity between the compared forms is nevertheless striking, and will hopefully be explained in the future.

11.4 Discussion: a pre-Roman language family in the western Mediterranean?

The geographical distribution of pre-Roman comparanda shows that there were linguistic connections between Sardinia and the Iberia peninsula, southern France, and central Italy. It has been considered whether the presence of these words in Sardinia may be the result of contacts with Iberian (or some other language spoken in the Iberian epigraphic area; § 11.3.1), and a Tyrrhenian language respectively (§ 11.3.2). In this context, it is interesting to have a look at the other regions where apparent cognates of these Sardinian-Ibero-French and Sardinian-Central Italian words are found. Of the twelve Sardinian-Central Italian correspondences, seven are found in South Italy, and five are found in Corsica, accounting for most of the convincing Sardinian-South Italian and Sardinian-Corsican comparisons. If the forms listed in § 11.2.1.5 are indeed of Tyrrhenian origin, their South Italian comparanda could be due to the historical expansion of Etruscan beyond Etruria, although Etruscan has not been found as far south as Calabria (Wallace 2008: 2–3). More striking is that six out of twelve words shared between Central Italian and Sardinia are also found in both the Iberian Peninsula and Southern France, with an additional three shared between Sardinia, Central Italy and Southern France. The fact that all of these regions (Sardinia, Iberia, France, Italy) show significant overlap in their pre-Roman comparanda, suggests the presence of closely related languages across a good portion of the western

Mediterranean. This has of course been claimed numerous times by the proponents of the "Mediterranean substrate hypothesis" (cf. § 1.3.2).

Since an Iberian context has been suggested for the comparanda in § 11.3.1 and a Tyrrhenian origin is consistent with the distribution of the comparanda in § 11.3.2, one wonders if, and how, both a hypothetical Iberian and a Tyrrhenian origin of certain pre-Roman Sardinian words can be upheld. To my knowledge, no linguistic relationship between Iberian and Etruscan has ever (successfully) been proposed,[322] nor will I do so — our poor understanding of both languages makes any direct structural or lexical comparison difficult. In light of the difficulties regarding Iberian as a donor language (e.g. its apparent lack of /m/; cf. § 11.3.1), the non-Iberian language whose existence is suspected on the basis of onomastic material in Iberian inscriptions in Catalonia and Roussillon (Gorrochategui & Vallejo 2019: 360) might be a better candidate — although our understanding of this language is of course even poorer than for Iberian.

The Greek comparanda do in themselves not constitute sufficient evidence for close linguistic ties between pre-Indo-European Sardinia and Greece, but there are some striking correspondences, that are also found in other parts of the Mediterranean. All of this raises the question what linguistic reality could account for these observations, and if we can find any non-linguistic evidence supporting a presence of some language or language family across the western Mediterranean.

On the basis of the close formal resemblance of the compared forms, it appears that the donor languages of the compared pre-Roman words were rather closely related. Confidently establishing a time-depth for the dispersal of these languages is impossible. However, the date of 10 000 – 9 000 BCE, proposed by Paulis (2008: 48) for the common ancestor of pre-Roman Sardinian and its relatives in the Iberian Peninsula, seems much too early. The meanings of the accepted lexical correspondences predominantly fall in the semantic categories of wild plants (e.g. 'holly', 'maple') and natural geographical features (e.g. 'marsh', 'rock crevice'), although there are some that can be interpreted as basic agricultural concepts (e.g. 'pit [to store grain]', 'pile of rocks, dry stone wall'). The borrowing of words for concepts tied to the natural environment indicates that the observed lexical correspondences between various parts of the Mediterranean are

[322] But comparanda from both the Iberian peninsula and Italy have frequently been discussed as part of a "Mediterranean substrate" (§ 1.3.2).

not likely the result of trade contacts. At least for Sardinia, we get the impression that these words originated in the language of a settled, native population.[323]

11.5 Conclusion

In summary, on the basis of the distributions of pre-Roman lexical items shared by Sardinian and other languages in the Mediterranean, I have argued for a reinstatement of sorts of the Mediterranean substrate hypothesis (cf. § 1.3.2), though in a much more nuanced form. There is evidence for related pre-Roman languages once being spoken in Sardinia and along the northern coast of the western Mediterranean, from northeastern Spain to the Italian peninsula. However, there is no convincing evidence for a number of other pre-Roman connections proposed in the context of the Mediterranean substrate hypothesis, such as those to Basque and Berber. The lexical correspondences between Sardinian and Basque are best explained by independent contact with a third language (§ 10.3). The correspondences between Sardinian and Berber that would be evidence for a Libyan presence in Sardinia, are more likely mediated through Punic spoken in Sardinia (§ 10.2). I have not found any evidence whatsoever for a linguistic connection between Sardinia and more far-flung regions such as Anatolia and the Caucasus (cf. Bertoldi 1928: 250; 1939: 100; Hubschmid 1963b: 147–148). The only potential pre-Roman Sardinian comparanda outside the western Mediterranean are found in the Aegean, but several of these are uncertain (§ 11.2.1.7).

Regarding Sardinia specifically, the two dozen words that have comparanda elsewhere in the Mediterranean are still but a section of the entire corpus of possibly pre-Roman Sardinian words. One could wonder whether the pre-Roman Sardinian words without Mediterranean comparanda originated in another pre-Roman language native to Sardinia that was not closely related to languages elsewhere.[324] This is difficult to conclusively prove or reject, but the forms with Mediterranean comparanda are widely distributed across Sardinia, indicating that their source language too was spoken everywhere on the island. Additionally, at least one of the morphemes identified in § 9 is present on a word with Mediterranean comparanda (viz. **ϑ-ánda* 'poppy'; § 3.1.15). This suggests that the other forms containing **ϑ(i)-*, and thus **k(V)-* too, originated in the same language.

[323] *Contra* Santano Moreno (2000: 174).

[324] A similar idea is expressed by Wagner (1944: 321–322).

V
Conclusion

12 Conclusion

12.1 Summary

One of the questions addressed in this study is whether the assumption of a single pre-Roman language, besides Punic, can account for all of the non-inherited lexical material. I have found that there is no geographical patterning in the phonological features found in words of pre-Roman origin. We can, however, discern a near-complementary geographical distribution in the pre-Roman prefixes **k(V)-* and **ϑ(i)-*, which have been argued in § 9.1.3 to be variants of one and the same pre-Roman morpheme (cf. also Swanenvleugel 2024). This prefix and other accepted pre-Roman morphemes exhibit an island-wide distribution. Pre-Roman Sardinian words, excluding punicisms with accepted cognates in other languages also occur across Sardinia. All of these findings constitute evidence supporting the hypothesis of a single language, or at least closely related language varieties, having existed all across Sardinia at the time of its romanization. This language coexisted with Punic. The coexistence of other languages with a smaller distribution cannot be ruled out.

The reality of many of the previously proposed phonological and morphological features attributed to a pre-Roman language in Sardinia cannot be confirmed based on the lexical material investigated in this study. This includes the pre-Roman vowel harmony proposed by Serra (1960; cf. § 8.4.2). The same goes for a number of putative pre-Roman suffixes (§ 9.2). What can be maintained is the pre-Roman phoneme **ϑ*, the existence of word-final consonants, and various morphemes, such as **k(V)-*/**ϑ(i)-*, **-Vr*, and **-(V)s-*.

Based on the evaluation of the previously proposed extra-Sardinian comparanda, it is likely that some Sardinian words derive from a Punic substrate. This mirrors the long-lasting presence of a Punic speech community on Sardinia, which persisted for centuries after the Roman conquest of Sardinia. Previous hypotheses on a linguistic affiliation of pre-Roman Sardinian to either Berber or Basque cannot be maintained on the basis of the lexical evidence. The Sardinian-Berber correspondences are argued to be part of the Punic substrate. The few acceptable Sardinian-Basque correspondences do not point to language relatedness, but rather to independent contacts with a third pre-Roman language. The hypothesis of an affiliation between pre-Roman Sardinian and Etruscan is difficult to evaluate because of the scarce attestation of Etruscan. The positive evidence is

restricted to a single convincing Sardinian-Etruscan lexical correspondence, and two possible morphological correspondences.

The analysis of the distribution of accepted comparisons between non-inherited words in Sardinian and in other languages of the Mediterranean indicates an especially close connection between Sardinia, the Iberian Peninsula, and southern France. This is potential evidence that the pre-Roman Sardinian language was related to the ancient Iberian language. There is, moreover, a sizable number of lexical correspondences between Sardinia and the Italian peninsula. The overlap in lexical correspondences between Sardinia and the Iberian Peninsula, southern France, and Italy, is evidence that related pre-Indo-European speech varieties were spoken across the western Mediterranean region.

There is ample evidence for a scenario in which both a pre-Roman "Mediterranean" language and Punic were spoken natively in Sardinia at the time of the Roman conquest. The geographical overlap of pre-Roman Mediterranean words and Punic words shows that these languages were spoken side by side. However, the non-inherited lexicon does not provide much evidence on the sociolinguistic dynamics between these two languages. Likewise, it is impossible to establish based on the extant evidence whether or not either of these pre-Roman languages outlived the other.

12.2 Outlook

The linguistic substrate of Sardinia is far from fully clarified. Most of the details of the languages of pre-Roman Sardinia and their eventual extinction resulting from the romanization process may be forever lost to us. However, there are some avenues through which progress can be made, especially regarding our understanding of linguistic connections between pre-Roman Sardinia and other parts of the Mediterranean. First and foremost, a reappraisal of the non-inherited vocabulary of other language varieties in the western Mediterranean may help verify the reality of a western Mediterranean pre-Indo-European language or language family. Second, advances in our understanding of the Etruscan and Iberian languages should allow for better evaluation of a possible genetic relationship to the pre-Roman language of Sardinia. Finally, the interpretation of the linguistic findings presented here would significantly be furthered by comparison to insights from the fields of archeology and genetics.

Appendix: Overview of Sardinian developments

This is a simplified and non-exhaustive overview of diachronic Sardinian sound developments to facilitate easier reading of the etymological discussions in this dissertation. The developments are organized from a greater to a more restricted distribution. As such, the isoglosses setting Sardinian apart from other Romance languages are mentioned first, followed by the main isoglosses setting apart Campidanese, Logudorese and Nuorese, and finally a selection of more local developments that are nevertheless relevant. For a complete overview of Sardinian historical phonology, the reader is advised to consult Wagner (1941) or Wagner and Paulis (1984). For an overview of the phonological systems found in the modern dialects of Sardinian, see Contini (1987). Note that *b, d* and *g* are generally realized as *β, ð* and *ɣ* in non-initial position. Many of the sound laws outlined below work across word boundaries, giving rise to productive consonant mutations.

General developments from Latin to Sardinian

V̄, V̆	>	*V̆*		Loss of contrastive vowel length
ae, oe	>	*e*		
au	>	*a*		
è : ò	>	*é : ó*	*/_...i, u*	Metaphony of stressed mid vowels by a high vowel in the following syllable
-VC	>	*-VCV*		Addition of paragogic vowels after word-final consonants
bb : dd	>	*pp : tt*		Devoicing of geminate voiced stops
ct, pt	>	*tt*		
c̯i, t̯i	>	**ʒ*		
ll	>	*ḍḍ*		
s̯i	>	*s*		
v	>	*b*		Except in Bitti, where *v* is partially preserved

Isoglosses between the Sardinian macro-varieties

Logudorese and Nuorese

c /k/	=	*k*	
i̯	>	*dz, ǧ*	*/l, r, m, n_*
qu /kw/	>	*bb*	
sC-	>	*isC-*	

Logudorese

b, d, g	>	*Ø*	/V_V	
p : t : k	>	*b : d: g*	/V_	
f	>	*v*	/V_V	
l	>	*ɻ*	/C_	
**ϑ*	>	*tt*		

Nuorese

l	>	*r*	/C_	Except in Baunei, where *l* is preserved
**ϑ*	>	*ϑϑ, tϑ*		

Campidanese

e : o	>	*i : u*	/unstressed	Note that the etymological value is reflected by the metaphony of stressed mid vowels
c : g	>	*č : ǧ*	/_*e, i*	
b, d, g, ǧ	>	*Ø*	/V_V	
p : t : k : č	>	*b : d : g : ž*	/V_	
f	>	*v*	/V_V	
l	>	*r*	/C_, _C	
qu /kw/	=	*kw*		
r-	>	*arr-*		
**ϑ*	>	*tts*		

Further noteworthy developments

Logudorese

l, r, s	>	*l*	/_$C_{plosive}$	In North-Logudorese, with various subsequent dialectal developments

Nuorese

f	>	*Ø*	/V_V	In Barbaricine dialects
k	>	*ʔ*		In Barbaricine dialects

Campidanese

-Vn-	>	*-Ṽ-*	/V_V	In many parts of the Campidano
l	>	*b, w, ʀ*	/V_V	Depending on the dialect
l	>	*ll*	/V_V	In and around Cagliari

l, n	>	*ʔ*	/V_V	In the Sarrabus region; *ʔ* < *n* is preceded by a nasalized vowel
**ϑ*	>	*ϑϑ, ss*		In Ogliastran dialects
**ϑ*	>	*čč*		In Sulcis and Desulo
r	frequent metatheses and loss of *r*			

Bibliography

Abbreviations

AIS Jaberg, Karl & Jakob Jud. 1928–1940. *Sprach- und Sachatlas Italiens und der Südschweiz*. 8 vols. Zofingen: Ringier.

ALIT Bartoli, Matteo & Ugo Pellis. 1995–2018. *Atlante linguistico Italiano*. 9 vols. Roma: Istituto Poligrafico e Zecca dello Stato.

CAD Oppenheim, A. Leo, Erica Reiner & Martha T. Roth (eds.). 1964–2010. *The Assyrian dictionary of the Oriental Institute of the University of Chicago*. 21 vols. Chicago: Oriental Institute Chicago.

BDLC BDLC. 2009. Banque de données de la langue Corse. *Corsican language database*. https://bdlc.univ-corse.fr/bdlc/corse.php.

DCEC Corominas, Joan & José A. Pascual. 1980–1991. *Diccionario crítico etimológico castellano e hispánico*. 6 vols. Madrid: Gredos.

DECLC Coromines i Vigneaux, Joan, Joseph Gulsoy & Max Cahner. 1980–1991. *Diccionari etimològic i complementari de la llengua catalana*. 9 vols. Barcelona: Curial Edicions Catalanes Caixa de Pensions "La Caixa". Accessed through: https://decat.iec.cat

DEI Battisti, Carlo & Alessio, Giovanni. 1950–1957. *Dizionario etimologico italiano*. 5 vols. *Firenze*: Barbèra.

DES Wagner, Max Leopold. 1960–1964. *Dizionario etimologico sardo*. 3 vols. Heidelberg: Carl Winter Universitätsverlag.

EH Elhuyar fundazioa. 2006. *Elhuyar hiztegia*. 3rd edn. Donostia; Elhuyar edizioak.

FEW Wartburg, Walther von, Margarete Hoffert, Carl Theodor Gossen, Jean-Pierre Chambon & Jean-Paul Chauveau. 1922–2003. *Französisches etymologisches Wörterbuch : eine Darstellung des galloromanischen Sprachschatzes*. Bonn: Klopp Verlag.

REW Meyer-Lübke, Wilhelm. 1911. *Romanisches etymologisches Wörterbuch*. Heidelberg: Carl Winter's Universitätsbuchhandlung.

TLL 1900–. *Thesaurus linguae Latinae*. https://tll.degruyter.com/.

ΛΕΓ Akadimía Athinón. 1933–1989. *Leksikón tis ellinikís glóssis. Á, Istorikón leksikón tis néas ellinikís: tis te kinós omilouménis ke ton idiomáton* [Λεξικόν της ελληνικης γλώσσης. Α', Ιστορικόν λεξικόν της νέας ελληνικής: της τε κοινώς ομιλουμένης και των ιδιωμάτων]. 5 vols. Athens: Estia.

Bibliography

Adams, James Noel. 2003. *Bilingualism and the Latin language*. Cambridge: Cambridge University Press.

Adams, James Noel. 2007. *The regional diversification of Latin 200 BC - AD 600*. Cambridge: Cambridge University Press.

Aebischer, Paul. 1934. Sur l'aire occupée par le mot "*péntoma*, *péntima*" et les noms de lieu correspondants en Italie. *Archivum romanicum: nuova rivista di filologia romanza* 18. 553–559.

Aikio, Ante. 2004. An essay on substrate studies and the origin of Saami. In Irma Hyvärinen, Petri Kallio & Jarmo Korhonen (eds.), *Etymologie, Entlehnungen und Entwicklungen: Festschrift für Jorma Koivulehto zum 70. Geburtstag* (Mémoires de La Société Néophilologique de Helsinki 63), 5–34. Helsinki: Société Néophilologique.

Alcover, A.M. & Francesc de Borja Moll. 2002. *Diccionari català-valencià-balear*. Edició electrònica. Institut d'Estudis Catalans. https://dcvb.iec.cat.

Alessio, Giovanni. 1935. La base preindoeuropea **kar(r)a/gar(r)a* "pietra." *Studi Etruschi* 9. 133–151.

Alessio, Giovanni. 1937. **biblax, -āce* "rododaphne." *Studi Italiani in Filologia Classica* 14. 311–315.

Alessio, Giovanni. 1939. I nomi della cavalletta in Italia. *Archivio Glottologico Italiano* 31. 13–84.

Alessio, Giovanni. 1944. Suggerimenti e nuove indagini sul problema del sostrato "mediterraneo." *Studi Etruschi* 18. 93–158.

Alessio, Giovanni. 1949. Vestigia etrusco-mediterranee nella flora toscana. *Studi Etruschi* 20. 109–149.

Alessio, Giovanni. 1957. Stratificazione dei nomi del «*tasso* (*Taxus baccata* L.)» in Europa. *Studi Etruschi* 25. 219–264.

Alinei, Mario. 1984. *Dal totemismo al cristianesimo popolare: Sviluppi semantici nei dialetti italiani ed europei.* Alessandria: Edizioni dell'Orso.

Ambrosini, Riccardo. 1983. Orientamento storico della toponomastica dell'isola d'Elba. In Enrico Campanile (ed.), *Problemi di sostrato nelle lingue indoeuropee*, 11–26. Pisa: Giardini Editori e Stampatori.

André, Jacques. 1956. *Lexique des termes de botanique en latin*. Paris: Librairie C. Klincksieck.

André, Jacques. 1985. *Les noms de plantes dans la Rome antique*. Paris: Société d'Édition "Les Belles Lettres".

Areddu, Alberto. 2022. Sardo *thurpu* "cieco": Un grecismo tra sostrato e latinità. *Zeitschrift für Romanische Philologie* 138(1). 260–286.

Argiolas, Valeria. 2010. Gli studi sul fenicio-punico in Sardegna: Prospettive di ricerca in linguistica storica. *Annali della Facoltà di Lingue e Letterature Straniere di Sassari* 7. 21–37.

Argiolas, Valeria. 2012. Caetra: Le jonc, la tresse ou la guerre dans le travail des Sardes. In *L'Africa romana: Trasformazione dei paesaggi del potere nell'Africa settentrionale fino alla fine del mondo antico*, vol. III, 2735–2742. Roma: Carocci editore.

Argiolas, Valeria. 2020. Un état des études sur l'action du substrat/adstrat "libyque" en Sarde: Problèmes et méthodes. *Asinag* 15. 31–61.

Ascoli, Graziado Isaia. 1882. Lettere glottologiche: Prima lettera. *Rivista di Filologia e d'Istruzione Classica* 10. 1–71.

Atzori, Maria Teresa. 1982. *Sardegna* (Profilo Dei Dialetti Italiani 20). Pisa: Pacini Editore.

Bats, Michel. 1998. Marseille archaïque. Étrusques et Phocéens en Méditerranée nord-occidentale. *Mélanges de l'École Française de Rome. Antiquité* 110(2). 609–633.

Bats, Michel. 2019. Les Ligyens et les Salyens d'Hécatée à Strabon. In *D'un monde à l'autre: Contacts et acculturation en Gaule méditerranéenne*, 189–200. OpenEdition Books. https://books.openedition.org/pcjb/5386. (7 March, 2024).

Battisti, Carlo. 1928. Sui più antichi strati toponomastici dell'Alto Adige. *Studi Etruschi* 2. 647–682.

Battisti, Carlo. 1955. Review of Johannes Hubschmid, Sardische Studien. Das mediterrane Substrat des Sardischen, seine Beziehungen zum Berberischen und Baskischen, sowie zum eurafrikanischen un hispano-kaukasischen Substrat der romanischen Sprachen. *Studi Etruschi* 23. 472–484.

Battisti, Carlo. 1960. Il sostrato mediterraneo nella fitonomia greco-latina. *Studi Etruschi* 28(2). 349–384.

Beavis, Ian C. 1988. *Insects and other invertebrates in classical Antiquity*. Exeter: Exeter University Publications.

Beccaria, Gian Luigi. 1995. *I nomi del mondo: Santi, demoni, folletti e le parole perdute*. Torino: Giulio Einaudi editore.

Beekes, Robert. 1993. The position of Etruscan. In Gerhard Meiser, Jadwiga Bendahman, Jón Axel Harðarson & Christiane Schaefer (eds.), *Indogermanica et Italica. Festschrift für Helmut Rix zum 65. Geburtstag*, 12–45. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.

Beekes, Robert. 1996. Ancient European loanwords. *Historische Sprachforschung* 109. 215–236.

Beekes, Robert. 2003. The origins of the Etruscans. *Koninklijke Nederlandse Akademie van Wetenschappen: Mededelingen van de Afdeling Letterkunde, Nieuwe Reeks* 66(1). 5–59.

Beekes, Robert. 2010. *Etymological Dictionary of Greek*. (Ed.) Lucien van Beek. Leiden: Brill.

Beekes, Robert. 2014. *Pre-Greek: Phonology, morphology, lexicon*. (Ed.) Stefan Norbruis. Leiden: Brill.

Beguinot, Francesco. 1942. *Il berbero Nefûsi di Fassâṭo: Grammatica, testi raccolti dalla viva voce, vocabolarietti.* 2nd edn. Roma: Pubblicazioni dell'Istituto per l'Oriente.

Bengtson, John. 2017. The anthropological context of Euskaro-Caucasian. *Iran and the Caucasus* 21(1). 75–91.

Bertoldi, Vittorio. 1928. Antichi filoni nella toponomastica mediterranea incrociantisi nella Sardegna. *Revue de Linguistique Romane* 4. 222–250.

Bertoldi, Vittorio. 1929. *Alica* "farina bianca" e formazioni affini. *Studi Italiani in Filologia Classica* 7. 251–263.

Bertoldi, Vittorio. 1930. Gallico **cabrostos* "ligustro" (da **cabros*, il corrispondente gallico del greco κάπρος, ecc.). *Revue Celtique* 47.

Bertoldi, Vittorio. 1931. Problèmes de substrat: Essai de méthodologie dans le domaine préhistorique de la toponymie et du vocabulaire. *Bulletin de la Société de Linguistique de Paris* 32. 93–184.

Bertoldi, Vittorio. 1932. Fonema basco-guascone attestato da Plinio. *Archivum Romanicum: Nuova Rivista di Filologia Romanza* 15. 400–408.

Bertoldi, Vittorio. 1933. "Preellenico βάτος, μαντία "cespuglio, rovo" e 'preromano' **matta*, **mantia* "cespuglio, rovo." *Glotta* 21. 258–267.

Bertoldi, Vittorio. 1934. Tracce del pliniano *tibulus* "pinaster" nelle Alpi lombarde, romance e bavaro-tirolesi. *Archivum Romanicum: Nuova Rivista di Filologia Romanza* 27. 73–81.

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. 137–69.

Bertoldi, Vittorio. 1937b. Plurale mediterraneo in residui fossili. In Johannes Wils & Johan Renders (eds.), *Mélanges de linguistique et de philologie offerts à Jacq. van Ginneken: à l'occasion du soixantième anniversaire de sa naissance*, 157–169. Paris: Librairie C. Klincksiek.

Bertoldi, Vittorio. 1939. Prerogative culturali mediterranee nel lessico greco e latino. *Archivio Glottologico Italiano* 31. 85–101.

Bertoldi, Vittorio. 1947a. Sardo-Punica. Contributo alla storia della cultura punica in terra sarda. *Parola del Passato* 4. 5–38.

Bertoldi, Vittorio. 1947b. Quisquiliae ibericae. *Romance Philology* 1. 191–208.

Bertoldi, Vittorio. 1950. Metodi vecchi e nuovi nella ricerca etimologica. *Archivio Glottologico Italiano* 36. 1–48.

Bertoldi, Vittorio. 1952. Fra latino e prelatino. *Archivio Glottologico Italiano* 37. 69–81.

Bertoni, Giulio. 1913. Denominazioni del "ramarro" (Lacerta viridis) in Italia. Romania: Recueil Trimestriel Consacré à l'Étude des Langues et des Littératures Romanes 42. 161–173.

Blasco Ferrer, Eduardo. 1984a. *Storia linguistica della Sardegna*. Tübingen: Max Niemeyer Verlag.

Blasco Ferrer, Eduardo. 1984b. *Grammatica storica del catalano e dei suoi dialetti con speciale riguardo all'algherese*. Tübingen: Gunter Narr Verlag.

Blasco Ferrer, Eduardo. 1988. *Le parlate dell'Alta Ogliastra: Analisi dialettologica. Saggio di storia linguistica e culturale* (Studi Di Linguistica Sarda 1). Cagliari: Edizioni Della Torre.

Blasco Ferrer, Eduardo. 1989. Il latino e la romanizzazione della Sardegna. Vecchie e nuove ipotesi. *Archivio Glottologico Italiano* 74. 5–89.

Blasco Ferrer, Eduardo. 1993. Tracce indeuropee nella Sardegna nuragica. *Indogermanische Forschungen* 98. 177–185.

Blasco Ferrer, Eduardo. 1999. La «*s*-» latina e la fine di una crux etimologica: Sardo «*thucru*»/«*zugu*» 'collo'. *Sardegna Mediterranea* 5. 65–70.

Blasco Ferrer, Eduardo. 2010a. *Paleosardo: Le radici linguistiche della Sardegna neolitica*. Berlin: De Gruyter.

Blasco Ferrer, Eduardo. 2010b. Iberian *ordumeles*, Paleo-Sardinian *Ortumele*, *Ortarani* and *Araunele*. Cognitive semantics and substrata research. *The Journal of Indo-European Studies* 28(3/4). 374–383.

Blasco Ferrer, Eduardo. 2011. Il sostrato paleosardo: Fine d'un rebus. *Zeitschrift für Romanische Philologie* 127(1). 63–110.

Blasco Ferrer, Eduardo. 2017. Paleosardo: Sostrati e toponomastica. In Eduardo Blasco Ferrer, Peter Koch & Daniela Marzo (eds.), *Manuale di linguistica sarda*, 67–84. Berlin: De Gruyter.

Blasco Ferrer, Eduardo. 2019. Africa, Sardinia, Sicily: Consistencies and idiosyncrasies in pre-Roman, Latin, Greek, and Arabic [1992]. *Romance Philology* 73. 37–68.

Blasco Ferrer, Eduardo, Peter Koch & Daniela Marzo (eds.). 2017. *Manuale di linguistica sarda* (Manuals of Romance Linguistics 15). Berlin: De Gruyter.

Blau, Otto. 1873. Vergessene Punische Glossen. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 27. 521–532.

Böhne, Rudolf. 1950. *Zum Wortschatz der Mundart des Sárrabus*. Berlin: Deutsche Akademie der Wissenschaften.

Bonazzi, Giuliano. 1900. *Il condaghe di San Pietro di Silki. Testo logudorese ined. dei secoli XI-XIII*. Sassari: Giuseppe Dessì.

Bondì, Sandro Filippo. 1987a. La frequentazione precoloniale fenicia. In Massimo Guidetti (ed.), *Storia dei Sardi e della Sardegna*, vol. 1, 129–146. Milano: Jaca Book.

Bondì, Sandro Filippo. 1987b. La colonizzazione fenicia. In Massimo Guidetti (ed.), *Storia dei Sardi e della Sardegna*, vol. 1, 147–172. Milano: Jaca Book.

Bondì, Sandro Filippo. 1987c. La dominazione cartaginese. In Massimo Guidetti (ed.), *Storia dei Sardi e della Sardegna*, vol. 1, 173–204. Milano: Jaca Book.

Bondì, Sandro Filippo. 1987d. Le sopravivvenze puniche nella Sardegna romana. In Massimo Guidetti (ed.), *Storia dei Sardi e della Sardegna*, vol. 1, 205–212. Milano: Jaca Book.

Bonfante, Giuliano & Larissa Bonfante. 1983. *The Etruscan language: An introduction*. Manchester: Manchester University Press.

Boogert, Nico van den. 1998. *La révélation des énigmes: Lexiques arabo-berbères des XVIIe et XVIIIe siècles* (Travaux et Documents de l'Iremam 19). Aix-en-Provence: Iremam.

Bossong, Georg. 2008. *Die romanischen Sprachen: Eine vergleichende Einführung*. Hamburg: Helmut Buske Verlag.

Bossong, Georg. 2016. Classifications. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford guide to the Romane languages*, 63–72. Oxford: Oxford University Press.

Bottiglioni, Gino. 1929. *Elementi prelatini nella toponomastica corsa* (L'Italia Dialettale. Supplemento 1). Pisa: Tipografia F. Simoncini.

Bouda, Charles. 1949. Traces basques en Sardaigne? *Eusko-Jakintza* 3. 333–335.

Boutkan, Dirk H. F. 1998. On the form of North European substratum words in Germanic. *Historische Sprachforschung* 111. 102–133.

Breyer, Gertraud. 1993. *Etruskisches Sprachgut im Lateinischen unter Ausschluss des spezifisch onomastischen Bereiches*. Leuven: Uitgeverij Peeters en Departement Oriëntalistiek.

Brugnatelli, Vermondo. 1998. La morphologie des noms berbères en *w-*. Considérations diachroniques. In Mohamed El Medlaoui, Saâd Gafaiti & Fouad Saa (eds.), *Actes du Premier Congrès Chamito-Sémitique de Fès* (*12-13 mars 1997*), 51–67. Saïs-Fès: Publications de la Faculté de Lettres et des Sciences Humaines.

Caelius Aurelianus. 1757. *De morbis acutis et chronicis*. (Ed.) Jo. Conradus Amman M.D. Venice: Typis Franciscis Storti.

Calò, Carla Maria, Giuseppe Vona, Renato Robledo & Paolo Francalacci. 2021. From old markers to next generation: Reconstructing the history of the peopling of Sardinia. *Annals of Human Biology* 48(3). 203–212.

Canu, Nadia. 2016. Tra Sarcidano e Barbagia. Spunti sulla romanizzazione in una zona di transizione. In Salvatore De Vicenzo & Chiara Blasetti Fantauzzi (eds.), *Il processo di romanizzazione della provincia Sardinia et Corsica*, 275–292. Roma: Edizioni Quasar.

Caramelli, David, Cristiano Vernesi, Simona Sanna, Lourdes Sampietro, Martina Lari, Loredana Castrì, Giuseppe Vona, et al. 2007. Genetic variation in prehistoric Sardinia. *Human Genetics* 122. 327–336.

Casu, Pietro. 2011. *Vocabolario Sardo Logudorese Italiano*. (Ed.) Giulio Paulis. http://vocabolariocasu.isresardegna.it.

Cetti, Francesco. 1776. *Gli ucelli di Sardegna*. Sassari: Giuseppe Piattoli.

Chantraine, P. 1933. *La formation des noms en grec ancien*. (Ed.) Édouard Champion. Paris: Librairie ancienne Honoré Champion.

Chiang, Charleston W.K., Joseph H. Marcus, Carlo Sidore, Arjun Biddanda, Hussein Al-Asadi, Magdalena Zoledziewska, Maristella Pitzalis, et al. 2018. Genomic history of the Sardinian population. *Nature Genetics* 50. 1426–1434.

Christian, Tom. 2022. Anaygris foetida. *Trees and Shrubs Online*. treesandshrubsonline.org/articles/anagyris/anagyris-foetida/. (4 January, 2024).

Ciorănescu, Alexandru. 1958. *Dicţionarul etimologic român*. Tenerife: Universidad de la Laguna. https://dexonline.ro.

Columella, Lucius Iunius Moderatus. 1982. *Zwölf Bücher über Landwirtschaft. Buch eines Unbekannten über Baumzüchtung: Lateinisch-Deutsch*. (Trans.) Will Richter, vol. 2. München: Artemis Verlag.

Contini, M. 1987. *Etude de géographie phonétique et de phonétique instrumentale du sarde*. Alessandria: Edizioni dell'Orso.

Contini, Michel & Edward F. Tuttle. 1982. Sardinian. In Rebecca Posner & John N. Green (eds.), *Trends in Romance linguistics and philology* (Trends in Linguistics. Studies and Monographs 14), vol. 3, 171–188. Berlin: De Gruyter Mouton.

Craddock, Jerry Russell. 1969. *Latin legacy versus substratum residue: The unstressed "derivational" suffixes in the Romance vernaculars of the Western Mediterranean*. Berkeley and Los Angeles: University of California Press.

Cuno, Johann Gustav. 1878. *Vorgeschichte Roms*, vol. 1. Leipzig: Commissions-Verlag Von B.G. Teubner.

Cuny, Albert. 1910. Les mots du fonds préhellénique: En Grec, Latin et Sémitique occidental. *Revue des Études Anciennes* 12. 154–164.

Dallet, Jean-Marie. 1982. *Dictionnaire kabyle-français: Parler des At Mangellat, Algérie*. Paris: Société d'études linguistiques et anthropologiques de France.

De Azkue, Resurrección María. 1905. *Diccionario vasco-español-francés*, vol. I. Tours: Imprimerie Alfred Mame et fils.

De Decker, Filip. 2015. A new book on Pre-Greek. *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 12(1). 1–21.

De Simone, Carlo. 1998. Etrusco e "Tirreno" di Lemnos: "Urverwandschaft"? *Rivista di Filologia e d'Istruzione Classica* 126. 392–411.

Demiraj, Bardhyl. 1997. *Albanische Etymologien: Untersuchungen zum albanischen Erbwortschatz*. Amsterdam: Rodopi.

Deplano, Andrea. 2020. *Dizionario etimologico dorgalese*, vol. II. Dolianova: Grafica del Parteolla.

Dessì Schmid, Sarah. 2017. L'algherese. In Eduardo Blasco Ferrer, Peter Koch & Daniela Marzo (eds.), *Manuale di linguistica sarda*, 460–475. Berlin: De Gruyter.

Destaing, Edmond. 1914. *Dictionnaire français-berbère* (*dialecte des Beni-Snous*). Paris: Ernest Leroux.

Devoto, Giacomo. 1940. *Storia della lingua di Roma*. Bologna: Licinio Cappelli Editore.

Devoto, Giacomo. 1955. Le fasi della linguistica mediterranea. *Studi Etruschi* 23. 217–228.

Dietler, Michael. 2010. *Archaeologies of colonialism: Consumption, entanglement, and violence in ancient Mediterranean France.* Los Angeles: University of California Press.

Dioscorides Anazarbei, Pedanius. 1906. *De materia medica libri quinque*. (Ed.) Max Wellmann, vol. II. Berlin: Weidmann.

Dioscorides Anazarbei, Pedanius. 1907. *De materia medica libri quinque*. (Ed.) Max Wellmann, vol. I. Berlin: Weidmann.

Dioscorides, Pedanius. 2005. *De materia medica*. (Ed.) Lily Y. Beck. Hildesheim: Olms.

Dworkin, Steven N. 2012. *A history of the Spanish lexicon: A linguistic perspective*. Oxford: Oxford University Press.

Dyson, Stephen L. 1985. *The creation of the Roman frontier*. Princeton: Princeton University Press.

Dyson, Stephen L. & Robert J. Rowland. 2007. *Archaeology and history in Sardinia from the Stone Age to the Middle Ages: Shepherds, sailors, and conquerors*. Philadelphia: University of Pennsylvania Museum of Archaeology and Anthropology.

Ernout, Alfred & Alfred Meillet. 2001. *Dictionnaire étymologique de la langue latine: Histoire des mots*. (Ed.) Jacques André. reprint of the 4th edition (1959). Paris: Klincksieck.

Ettmayer, Karl. 1906. Zur Aussprache des lateinischen l. *Zeitschrift fur Romanische Philologie* 30. 648–659.

Ettmayer, Karl. 1926. Der Ortsname "Luzern." *Indogermanische Forschungen* 43(1). 10–39.

Farina, Luigi. 1987. *Bocabolariu sardu nugoresu-italianu*. Sassari.

Farina, Luigi. 1989. *Vocabolario italiano-sardo nuorese*. Sassari.

Farre, Claudio. 2016. Alcune considerazioni sulla Barbaria: Definizione, percezione e dinamiche di romanizzazione nella Sardegna interna. In Salvatore De Vicenzo & Chiara Blasetti Fantauzzi (eds.), *Il processo di romanizzazione della provincia Sardinia et Corsica*, 89–106. Roma: Edizioni Quasar.

Fernandes, Daniel M., Alissa Mittnik, Iñigo Olalde, Iosif Lazaridis, Olivia Cheronet, Nadin Rohland, Swapan Mallick, et al. 2020. The spread of steppe and Iranian-related ancestry in the islands of the western Mediterranean. *Nature Ecology & Evolution* 4. 334–345.

Ferraro, Giuseppe (ed.). 1891. *Canti popolari in dialetto logudorese* (Canti e Racconti Del Popolo Italiano 9). Forni Editore Bologna.

Flechia, Giovanni. 1872. *Dell'origine della voce sarda nuraghe: Congetture etimologiche*. Torino: Stamperia Reale.

Forsyth Major, C. J. 1893. Italienische Vulgärnamen der Fledermaus. *Zeitschrift für Romanische Philologie* 17. 148–160.

Forza, Francesca & Sergio Scalise. 2016. Compounding. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford guide to the Romance languages*, 524–538. Oxford: Oxford University Press.

Frazer, James George. 2012a. *Pausanias's description of Greece* [1898], vol. 1: Translation. Cambridge: Cambridge University Press.

Frazer, James George. 2012b. *Pausanias's description of Greece* [1898], vol. 5: Commentary on books IX, X. Addenda. Cambridge: Cambridge University Press.

Friedrich, Johannes. 1951. *Phönizisch - Punische Grammatik*. Roma: Pontificium Institutum Biblicum.

Friedrich, Johannes. 1957. Zur Frage punischer Lehnwörter im Sardischen. *Die Sprache* 3. 221–224.

Furnée, Edzard Johan. 1972. *Die wichtigsten konsonantischen Erscheinungen des Vorgriechischen*. Den Haag: Universiteit Leiden dissertation.

Furnée, Edzard Johan. 1979. *Vorgriechisch-Kartvelisches: Studien zum ostmediterranen Substrat nebst einem Versuch zu einer neuen pelasgischen Theorie*. Leuven: Editions Peeters.

Gasperini, L. 1992. Ricerche epigrafiche in Sardegna (I). In Bonello Lai (ed.), *Sardinia antiqua, studi in onore di Piero Meloni in occasione del suo settantesimo compleanno*, 287–323. Cagliari: Edizione della Torre.

GBIF Secretariat. 2023. Anguilla acutirostris Risso, 1827. *GBIF Backbone Taxonomy*. (12 January, 2024).

Gorrochategui, Joaquín. 1995. The Basque language and its neighbours in antiquity. In José Ignacio Hualde, Joseba A. Lakarra & Robert Lawrence Trask (eds.), *Towards a history of the Basque language*, 31–64. Amsterdam: John Benjamins Publishing Company.

Gorrochategui, Joaquín & Joseba A. Lakarra. 2001. Comparación lingüística, filología y reconstrucción del protovasco. In Francisco Villar & María Pilar Fernández Álvarez (eds.), *Religión, lengua y cultura prerromanas de Hispania*, 407–438. Salamanca: Ediciones Universidad de Salamanca.

Gorrochategui, Joaquín & José Maria Vallejo. 2019. The parts of Hispania without epigraphy. In Alejandro G. Sinner & Javier Velaza (eds.), *Palaeohispanic languages and epigraphies*, 335–364. Oxford: Oxford University Press.

Guarnerio, Pier Enea. 1904a. Postille sul lessico sardo. *Romania: Recueil Trimestriel Consacré à l'Étude des Langues et des Littératures Romanes* 33. 50–70.

Guarnerio, Pier Enea. 1904b. Ancora di *ti-*(*zi-*). *Romania: Recueil Trimestriel Consacré à l'Étude des Langues et des Littératures Romanes* 33. 258–260.

Guarnerio, Pier Enea. 1906. L'antico campidanese dei sec. XI-XIII secondo "le antiche carte volgari dell'archivio arcivescovile di Cagliari." *Studi Romanzi* 4. 189–260.

Guidetti, Massimo (ed.). 1987-1989. *Storia dei Sardi e della Sardegna*. 4 vols. Milano: Jaca Book.

Haddadou, Mohand Akli. 2007. *Dictionnaire des racines berbères communes*. Haut Commissariat à l'Amazighité.

Hamp, Eric P. 1979. The North European word for "apple." *Zeitschrift für Celtische Philologie* 37. 158–166.

Hamp, Eric P. 1990. The Pre-Indo-European language of Northern (Central) Europe. In John A. C. Greppin & Thomas L. Markey (eds.), *When worlds collide: The Indo-Europeans and the Pre-Indo-Europeans*, 291–309. Ann Arbor: Karoma Publishers.

Heizer, Robert F. 1953. Aboriginal fish poisons. *Smithsonian Insitution, Bureau of American Ethnology Bulletin* 151. 225–284.

Herbig, Gustav. 1917. Etruskisches Latein. *Indogermanische Forschungen* 37. 163–187.

Hervás, Lorenzo. 1979. *Catálogo de las lenguas de las naciones conocidas: y numeracion, division, y clases de estas segun la diversidad de sus idiomas y dialectos* [1800], vol. I. Madrid: Ediciones Atlas.

Hesychius Alexandrinus. 2018. *Hesychii Alexandrini Lexicon: Editionem post Kurt Latte continuans recensuit et emendavit*. (Ed.) Kurt Latte & Ian C. Cunningham, vol. I. Berlin: De Gruyter.

Hesychius Alexandrinus. 2020. *Hesychii Alexandrini Lexicon: Editionem post Kurt Latte continuans recensuit et emendavit*. (Ed.) Kurt Latte & Ian C. Cunningham, vol. II. Berlin: De Gruyter.

Heubeck, Alfred. 1949. Die homerische Göttersprache. *Würzburger Jahrbücher für die Altertumswissenschaft* 4(2). 197–218.

Heubeck, Alfred. 1959. *Lydiaka: Untersuchungen zu Schrift, Sprache un Göttername der Lyder* (Erlanger Forschungen Reihe A: Geisteswissenschaften 9). Erlangen: Universitätsbund Erlangen.

Hofmann, Johann Baptist. 1949. *Etymologisches Wörterbuch des Griechischen*. München: Verlag von R. Oldenbourg.

Hofmann, Konrad & T.M. Auracher. 1883. Der Longobardische Dioskorides des Marcellus Virgilius. *Romanische Forschungen* 1. 49–105.

Howald, Ernst & Henry E. Sigerist (eds.). 1927. *Antonii Musae de herba vettonica liber; Pseudoapulei herbarius; anonymi de taxone liber; sexti placiti liber medicinae ex animalibus etc.* (Corpus Medicorum Latinorum 4). Leipzig: B.G. Teubner.

Hoz, Javier de. 2009. El problema de los limites de la lengua ibérica como lengua vernácula. *Palaeohispanica: revista sobre lenguas y culturas de la Hispania antigua* 9. 413–433.

Hoz, Javier de. 2011. Las funciones de la lengua ibérica como lengua vehicular. In Coline Ruiz & Eugenio Ramón Luján (eds.), *Contacts linguistiques dans l'Occident méditerranéen antique*, 27–64. Madrid: Collection de la Casa de Velázquez.

Hubschmid, Johannes. 1943. Bezeichnungen für "Kaninchen" - "Höhle" - "Steinplatte." In *Sache, Ort und Wort. Jakob Jud zum sechzigsten Geburtstag*, 150–151. Zürich.

Hubschmid, Johannes. 1949. Review of H. Vaillant-Carmanne, Mélanges de linguistique romane offerts à M. Jean Haust à l'occasion de son admission à l'éméritat. *Zeitschrift für Romanische Philologie* 65. 243–245.

Hubschmid, Johannes. 1950a. Ein etruskisch-iberischer Pflanzenname. *Museum Helveticum: Schweizerische Zeitschrift für Klassische Altertumswissenschaft* 7(4). 221–226.

Hubschmid, Johannes. 1950b. Circummediterrane Wortgruppen des westlichen Mittelmeergebietes. *Vox Romanica* 11. 125–134.

Hubschmid, Johannes. 1950c. Vorindogermanische und jüngere Wortschichten in den romanischen Mundarten der Ostalpen. *Zeitschrift für Romanische Philologie* 66. 1–94.

Hubschmid, Johannes. 1953. *Sardische Studien*. Bern: A. Francke.

Hubschmid, Johannes. 1955. Haustiernamen und Lockrufe als Zeugen vorhistorischer Sprach- und Kulturbewegungen. *Vox Romanica* 14. 184–203.

Hubschmid, Johannes. 1958. Kritische Bemerkungen zum mediterranen Substrat Sardiniens und Hispaniens. *Zeitschrift für Romanische Philologie* 74. 212–233.

Hubschmid, Johannes. 1960a. *Mediterrane Substrate: Mit besonderer Berücksichtigung des Baskischen und der west-östlichen Sprachbeziehungen* (Romanica Helvetica 70). 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. 1963a. Paläosardische Ortsnamen. In *Atti e Memorie del VII Congresso Internazionale di Scienze Onomastiche*, vol. II, 144–180. Firenze: Tipografia Francolini.

Hubschmid, Johannes. 1963b. *Thesaurus Praeromanicus: Grundlagen für ein Weitverbreitetes mediterranes Substrat, dargestellt an romanischen, baskischen und vorindogermanischen p-Suffixen*, vol. I. Bern: A. Francke.

Hubschmid, Johannes. 1965. *Thesaurus Praeromanicus: Probleme der baskischen Lautlehre und baskisch-vorromanische Etymologien*, vol. II. Bern: A. Francke.

Huld, Martin E. 1990. The linguistic typology of the Old European substrata in north central Europe. *Journal of Indo-European Studies* 18(3/4). 389–423.

ICZN. 2024. *International code of zoological nomenclature*. https://www.iczn.org/the-code/the-code-online/. (15 October, 2024).

Iribarren Argaiz, Mary Carmen. 1997. Los vocablos en *-rr-* de la lengua sarda. Conexiones con la península ibérica. *Fontes Linguae Vasconum* 76. 335–360.

Jones, Michael A. 1997. Sardinia. In Martin Maiden & Mair Parry (eds.), *The Dialects of Italy*, 376–384. London: Routledge.

Jud, Jakob. 1911. Dalla storia delle parole lombardo-ladine. *Bulletin de Dialectologie Romane* 3. 1–18, 63–86.

Jud, Jakob. 1914. Review of Carlo Salvioni, Note di lingua sarda; Bricciche sarde. *Romania: Recueil Trimestriel Consacré à l'Étude des Langues et des Littératures Romanes* 43. 451–457.

Jud, Jakob. 1931. Review of M. L. Wagner, Über die vorrömischen Bestandteile des Sardischen. *Romania: Recueil Trimestriel Consacré à l'Étude des Langues et des Littératures Romanes* 57. 438–440.

Jud, Jakob. 1937. Nachrichten. *Vox Romanica* 2. 311–313.

Jud, Jakob. 1942. Review of Max Leopold Wagner, Historische Lautlehre des Sardischen. *Vox Romanica* 6. 349–353.

Katičić, Radoslav. 1976. *Ancient languages of the Balkans*, vol. I. Den Haag: Mouton.

Kloekhorst, Alwin. 2022a. Luwians, Lydians, Etruscans, and Troy: The linguistic landscape of Northwestern Anatolia in the pre-classical period. In Ivo Hajnal, Eberhard Zangger & Jorrit Kelder (eds.), *The political geography of Western Anatolia in the Late Bronze Age. Proceedings of the EAA Conference. Bern, 7 September 2019*, 201–227. Budapest: Archaeolingua.

Kloekhorst, Alwin. 2022b. Anatolian. In Thomas Olander (ed.), *The Indo-European language family: A phylogenetic perspective*, 63–82. Cambridge: Cambridge University Press.

Kogan, Leonid. 2012a. Proto-Semitic Lexicon. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 179–258. Berlin/Boston: De Gruyter Mouton.

Kogan, Leonid. 2012b. Proto-Semitic Phonetics and Phonology. In Stefan Weninger (ed.), *The Semitic Languages: An international handbook*. Berlin: De Gruyter Mouton.

Kossmann, Maarten. 1999. *Essai sur la phonologie du proto-berbère*. Köln: Rüdiger Köppe Verlag.

Kossmann, Maarten. 2020a. Proto-Berber phonological reconstruction: An update. *Linguistique et Langues Africaines* 6. 11–42.

Kossmann, Maarten. 2020b. Sibilants in Libyco-Berber. *Journal of the American Oriental Society* 140(4). 875–888.

Krahe, Hans. 1925. *Balkanillyrischen geographischen Namen*. Heidelberg: Carl Winter's Universitätsbuchhandlung.

Krahe, Hans. 1929. *Lexikon altillyrischer Personennamen*. Heidelberg: Carl Winter's Universitätsbuchhandlung.

Krahmalkov, Charles R. 2000. *Phoenician-Punic dictionary* (Studia Phoenicia 15). Leuven: Uitgeverij Peeters & Departement Oosterse Studies.

Krahmalkov, Charles R. 2001. *A Phoenician-Punic grammar* (Handbook of Oriental Studies. Section One: The Near and Middle East 54). Leiden: Brill.

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. *Zeitschrift für Vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprachen* 38(1). 97–128.

Kretschmer, Paul. 1940. Die vorgriechischen Sprach- und Volksschichten. *Glotta* 28. 231–279.

Kronasser, Heinz. 1966. *Etymologie der hethitischen Sprache*, vol. I. Wiesbaden: Otto Harrassowitz.

Kroonen, Guus. 2013. *Etymological Dictionary of Proto-Germanic*. Leiden: Brill.

Kroonen, Guus. 2024a. A methodological introduction to sub-Indo-European Europe. In Guus Kroonen (ed.), *Sub-Indo-European Europe*, 3–26. Berlin: De Gruyter Mouton.

Kroonen, Guus. 2024b. For the *nth* time: The Pre-Greek νθ-suffix revisited. In Guus Kroonen (ed.), *Sub-Indo-European Europe*. Berlin: De Gruyter Mouton.

Kühn, Karl Gottlob. 1828. *Opuscula academica medica et philologica*, vol. 2. Lipsiae: sumtibus Leopold Vossii.

Kuiper, Franciscus B.J. 1956. The etymology of ἄνθρωπος. In Heinz Kronasser (ed.), *ΜΝΗΜΗΣ ΧΑΡΙΝ. Gedenkschrift Paul Kretschmer*, vol. I, 211–226. Wien: Wiener Sprachgesellschaft.

Kuiper, Franciscus B.J. 1995. Gothic *bagms* and Old Icelandic *ylgr*. *North-Western European Language Evolution* 25. 63–88.

Labrune, Laurence. 2021. Word-initial rhotic avoidance: A typological survey. *Glossa: A Journal of General Linguistics* 6.

Lakarra, Joseba A. 2009. Forma canónica y cambios en la forma canónica de la lengua vasca: Hacia los orígenes del bisilabismo. In *Acta Paleohispanica X* (Paleohispanica 9), 557–609. Zaragoza: Institución "Fernando el Católico."

Lakarra, Joseba A., Julen Manterola & Iñaki Segurola. 2016. Los estudios etimológicos vascos: Historia y perspectivas. In Mariano Quirós García, José Ramón Carriazo Ruiz, Emma Falque Rey & Marta Sánchez Orense (eds.), *Etimología e historia en el léxico del español: Estudios ofrecidos a José Antonio Pascual* (*Magister bonus et sapiens*), 845–871. Frankfurt am Main: Vervuert Verlagsgesellschaft.

Lakarra, Joseba A., Julen Manterola, Iñaki Segurola, Gabriel Fraile & Koro Segurola. 2019. *Euskal hiztegi historiko-etimologikoa*. Bilbo: Euskaltzaindia.

Laoust, Émile. 1912. *Étude sur le dialecte berbère du Chenoua: Comparé avec ceux des Beni-Menacer et des Beni-Salah*. Paris: Ernest Leroux.

Laoust, Émile. 1920. *Mots et choses berbères: Notes de linguistique et d'ethnographie. Dialectes du Maroc*. Paris: Augustin Challamel.

Ledgeway, Adam. 2016. Italian, Tuscan, and Corsican. In *The Oxford Guide to the Romance Languages*, 206–227. Oxford: Oxford University Press.

Lewis, Charlton T. & Charles Short. 1879. *A Latin Dictionary. Founded on Andrews' edition of Freund's Latin Dictionary*. Oxford: Clarendon Press.

Lewy, Heinrich. 1895. *Die semitischen Fremdwörter im Griechischen*. Berlin: R. Gaertners Verlagsbuchhandlung.

Liddell, H.G., Robert Scott, Henry Stuart Jones, Eric Arthur Barber, Roderick Mckenzie & Paul Maas. 1978. *A Greek-English Lexicon*. Oxford: Oxford University Press.

Lilliu, Giovanni. 1963. *La civiltà dei Sardi: Dal Paleolitico all'età dei nuraghi*. Torino: Nuova Eri.

Lilliu, Giovanni. 2002. La civiltà preistorica e nuragica in Sardegna. *Atti della Accademia Nazionale dei Lincei. Classe di Scienze Morali, Storiche e Filologiche* (9) 15(3). 219–264.

Löpelmann, Martin. 1968. *Etymologisches Wörterbuch der baskischen Sprache*. Berlin: De Gruyter.

Loporcaro, Michele. 2013. Sulla Sardegna preromana. *Romanische Forschungen* 125. 54–65.

Louali, Naima & Gérard Philippson. 2004. Berber extension into and within north-west Africa: A linguistic contribution. *Afrika und Übersee* 87. 105–130.

Lubotsky, Alexander M. 2001. The Indo-Iranian substratum. In Christian Carpelan & Petteri Koskikallio (eds.), *Early contacts between Uralic and Indo-European: Linguistic and archaeological considerations* (Mémoires de La Société Finno-Ougrienne 242), 301–317. Helsinki: Société Finno-Ougrienne.

Lupinu, Giovanni. 2005. La romanizzazione linguistica della Sardegna. In Attilio Mastino (ed.), *La Sardegna e la sua storia*, vol. 2, 193–197. Nuoro: Edizioni Il Maestrale.

Lupinu, Giovanni (ed.). 2020. *Carta de Logu – Manoscritto*. Repertorio informatizzato delle fonti documentarie e letterarie della Sardegna: https://www.reisar.eu/category/carta-de-logu-manoscritto/. (10 December, 2024).

Lupinu, Giovanni. 2023. *Manualetto di linguistica sarda*. Cagliari: UNICApress.

Maninchedda, Paolo & Antonello Murtas (eds.). 2003. *Il condaghe di San Michele di Salvennor*. Cagliari: Cooperativa Universitaria Editrice Cagliaritana.

Marcialis, Efisio. 2005. *Vocabolari*. (Ed.) Eleonora Frongia. Cagliari: Cooperativa Universitaria Editrice Cagliaritana.

Martinet, André. 1955. *Économie des changements phonétiques: Traité d,e phonologie diachronique*. Bern: A. Francke.

Mastino, Attilio. 1993. Analfabetismo e resistenza: Geografia epigrafica della Sardegna. In Alda Calbi, Angela Donati & Gabriella Poma (eds.), *L'epigrafia del villaggio*, 457–536. Faenza: Fratelli Lega Editori.

Mastrelli, Carlo Alberto. 1970. Per la latinità del sardo *littu*. In *Miscellanea in onore di Tristano Bolelli*, 227–255. Pisa: Pacini.

Matasović, Ranko. 2008. *Etymological Dictionary of Proto-Celtic*. Leiden: Brill.

Maulu, Marco. 2016. Storia, grafia e fonetica del Condaghe di San Pietro di Silki. *Zeitschrift für Romanische Philologie* 132(2). 527–556.

Maxia, Mauro. 2017. Il gallurese e il sassarese. In Eduardo Blasco Ferrer, Peter Koch & Daniela Marzo (eds.), *Manuale di linguistica sarda*, 431–445. Berlin/Boston: De Gruyter.

Mayer Modena, Maria Luisa. 1989. Zoonima ebraica e sostrato mediterraneo: Il nome della tarma. *Acme: Annali della Facoltà di Lettere e Filosofia dell'Università degli Studi di Milano* 42(2). 11–12.

Meester, Lotte. 2024. Substrate stratification: An argument against the unity of Pre-Greek. In Guus Kroonen (ed.), *Sub-Indo-European Europe*, 283–299. Berlin: De Gruyter Mouton.

Meillet, Antoine. 1908. De quelques emprunts probables en Grec et en Latin. *Mémoires de la Société de Linguistique de Paris* 15. 165–169.

Meissner, Torsten. 2013. Review of Beekes: Etymological Dictionary of Greek. *Kratylos* 58. 1–32.

Meloni, Piero. 1987. La romanizzazione. In Massimo Guidetti (ed.), *Storia dei Sardi e della Sardegna*, vol. 1, 263–295. Milano: Jaca Book.

Mensching, Guido. 1997. Campidanesisch *mitza* "Quelle" - punisch oder lateinisch? In Annegret Bollée & Johannes Kramer (eds.), *Latinitas et Romanitas: Festschrift für Hans Dieter Bork zum 65. Geburtstag*, 293–305. Bonn: Romanistischer Verlag.

Mensching, Guido. 2004. Su sardu - limba arcàica o limba moderna? In Guido Mensching & Lucia Grimaldi (eds.), *Su sardu - limba de Sardigna e limba de Europa*, 27–53. Cagliari: CUEC.

Mensching, Guido & Eva-Maria Remberger. 2016. Sardinian. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford guide to the Romance languages*, 270–291. Oxford: Oxford University Press.

Merci, Paolo (ed.). 1992. *Il Condaghe di San Nicola di Trullas*. Sassari: Carlo Delfino Editore.

Meyer, Gustav. 1891. *Etymologisches Wörterbuch der Albanischen Sprache*. Strassburg: Karl J. Trübner.

Meyer-Lübke, Wilhelm. 1903. Zur Kenntniss des Altlogudoresischen. *Sitzungsberichte der Philosophisch-Historischen Classe der Kaiserlichen Akademie der Wissenschaften* 145(5). 1–76.

Meyer-Lübke, Wilhelm. 1925. Zur Kenntnis der vorrömischen Ortsnamen der iberischen Halbinsel. In *Homenaje oferecido a Menéndez Pidal: Misce-*

lánea de estudios lingüísticos, literarios e históricos, 63–84. Madrid: Librería y casa editorial Hernando (S.A.).

Michelena, Luis. 1953. *Apellidos vascos*. San Sebastián: Biblioteca Vascongada de los Amigos de País.

Michelena, Luis. 1958. Review of Max Leopold Wagner, Dizionario etimologico sardo. *Word* 14(2–3). 405–406.

Michelena, Luis. 1959. Review of Max Leopold Wagner, Dizionario etimologico sardo. *Word: Journal of the Linguistic Circle of New York* 15. 523–527.

Michelena, Luis. 1969. Notas lingüísticas a "Collección diplomática de Irache." *Fontes Linguae Vasconum* 1. 1–59.

Michelena, Luis. 1990. *Fonética histórica vasca*. Donostia: Gipuzkoako Foru Aldundia.

Molinu, Lucia & Franck Floricic. 2017. Storia delle indagini e classificazioni. In Eduardo Blasco Ferrer, Peter Koch & Daniela Marzo (eds.), *Manuale di linguistica sarda*, 15–30. Berlin: De Gruyter.

Moncunill, Noemí & Javier Velaza. 2020. Iberian. *Palaeohispanica: Revista Sobre Lenguas y Culturas de la Hispania Antigua* 20. 591–629.

Monier-Williams, Monier. 1872. *A Sanskrit-English Dictionary*. Oxford: Clarendon Press.

Mullen, Alex. 2013. *Southern Gaul and the Mediterranean: Multilingualism and multiple identities in the Iron Age and Roman periods.* Cambridge: Cambridge University Press.

Müllenhoff, Karl. 1892. *Deutsche Altertumskunde*, vol. 3. Berlin: Weidmannische Buchhandlung.

Múrcia Sànchez, Carles. 2011. *La llengua amaziga a l'antiguitat a partir de les fonts gregues i llatines*, vol. 1. Barcelona: Promociones y Publicaciones Universitarias.

Nesi, Annalisa. 1988. Korsisch: Interne Sprachgeschichte. (Ed.) Günter Holtus, Michael Metzeltin & Christian Schmitt. *Lexikon der romanistischen Linguistik*. Tübingen: Max Niemeyer Verlag.

Oettinger, Norbert. 2010. Seevölker und Etrusker. In Yoram Cohen, Amir Gilan & Jared L. Miller (eds.), *Pax Hethitica: Studies on the Hittites and their neighbours in honour of Itamar Singer*, 233–246. Wiesbaden: Harrassowitz.

Orduña, Eduardo. 2019. The Vasco-Iberian theory. In Alejandro G. Sinner & Javier Velaza (eds.), *Palaeohispanic languages and epigraphies*, 219–239. Oxford: Oxford University Press.

Orel, Vladimir. 1998. *Albanian etymological dictionary*. Leiden: Brill.

Oštir, Karl. 1921. *Beiträge zur alarodischen Sprachwissenschaft*. Wien/Leipzig: Beyers.

Oštir, Karl. 1930. *Drei vorslavisch-etruskische Vogelnamen*. Ljubljana: Znanstveno društvo.

Pais, Ettore. 1910. Sulla civiltà dei nuraghi e sullo sviluppo sociologico della Sardegna. *Archivio Storico Sardo* 6. 85–192.

Parodi, Ernesto Giacomo. 1907. Sul raddoppiamento di consonanti postoniche negli sdruccioli italiani. *Romanische Forschungen* 23. 756–775.

Paulis, Giulio. 1983. *Lingua e cultura nella Sardegna bizantina: Testimonianze linguistiche dell'influsso greco.* Sassari: L'Asfodelo Editore.

Paulis, Giulio. 1984a. Introduzione. In Giulio Paulis (ed.), *Fonetica storica del sardo*, vii–cx. Cagliari: Gianni Trois Editore.

Paulis, Giulio. 1984b. Appendice. In Giulio Paulis (ed.), *Fonetica storica del sardo*, 475–636. Cagliari: Gianni Trois Editore.

Paulis, Giulio. 1985. Ha origini arcaiche la toponomastica montana. In *Sardegna. L'uomo e le montagne*. Sassari: Banco di Sardegna.

Paulis, Giulio. 1987. *I nomi di luogo della Sardegna*. Sassari: Carlo Delfino Editore.

Paulis, Giulio. 1990. Sopravvivenze della lingua punica in Sardegna. In *L'Africa romana. Atti del VII convegno di studio. Sassari, 15-17 dicembre 1989*, vol. 2, 599–639. Sassari: Edizioni Gallizzi.

Paulis, Giulio. 1992. *I nomi popolari delle piante in Sardegna*. Sassari: Carlo Delfino Editore.

Paulis, Giulio. 1994. L'influsso linguistico fenicio-punico in Sardegna. Nuove acquisizioni e prospettive di ricerca. In Paolo Filigheddu (ed.), *Circolazioni culturali nel Mediterraneo antico. Atti della VI giornata camito-semitica e indoeuropea,* 213–219. Cagliari: Corda.

Paulis, Giulio. 1997. *Studi sul sardo medioevale*. Nuoro: Ilisso.

Paulis, Giulio. 2008. Il paleosardo: Retrospettiva e prospettive. *ΑΙΩΝ: Annali del Dipartimento di Studi del Mondo Classico e del Mediterraneo Antico. Sezione Linguistica* 30(4). 11–61.

Paulis, Giulio. 2011. I dolci sardi nella storia della lingua e della cultura. In Anna Saderi (ed.), *Dolci in Sardegna: Storia e tradizione*, 143–185. Nuoro: Ilisso.

Paulis, Giulio. 2017. Greco e superstrati primari. In Eduardo Blasco Ferrer, Peter Koch & Daniela Marzo (eds.), *Manuale di linguistica sarda*, 104–118. Berlin: De Gruyter.

Pedersen, Holger. 1909. *Vergleichende Grammatik der keltischen Sprachen*, vol. 1. Göttingen: Vandenhoeck und Ruprecht.

Philipp, Hans. 1920. Sardinia. (Ed.) Georg Wissowa. *Paulys Real-Encyclopädie der classischen Altertumswissenschaft. Neue Bearbeitung.* (Zweite Reihe [R-Z] I). Stuttgart: J. B. Metzlersche Verlagsbuchhandlung.

Piccitto, Giorgio & Giovanni Tropea (eds.). 1977. *Vocabolario siciliano*, vol. 1. Catania: Centro di studi filologici e linguistici siciliani.

Piccitto, Giorgio & Giovanni Tropea (eds.). 2002. *Vocabolario siciliano*, vol. 5. Catania: Centro di studi filologici e linguistici siciliani.

Pinna, Claudia. 2013. Acorn bread: A traditional food of the past in Sardinia (Italy). *Journal of Cultural Heritage* 14(Supplement). S71–S74.

Pittau, Massimo. 1984. *Lessico etrusco-latino comparato col nuragico.* Sassari: Chiarella.

Pittau, Massimo. 1990. *I cognomi della Sardegna: Significato e origine di 5.000 cognomi indigeni.* Sassari: Carlo Delfino editore.

Pittau, Massimo. 1995. *Origine e parentela dei sardi e degli etruschi: Saggio storico-linguistico.* Sassari: Carlo Delfino Editore.

Pittau, Massimo. 1997. *La lingua etrusca: Grammatica e lessico*. Nuoro: Insula.

Pittau, Massimo. 2013. Toponimi della Sardegna settentrionale. http://www.pittau.it/Sardo/top_sard_settentrionale.html. (24 November, 2022).

Plinius Secundus Maior, Caius. 1857. *The Natural History of Pliny*. (Trans.) John Bostock & H.T. Riley. London: Taylor and Francis.

Polomé, Edgar C. 1983. The linguistic situation in the western provinces of the Roman empire. In Wolfgang Haase (ed.), *Teilband Sprache und Literatur* (Aufstieg Und Niedergang Der Römischen Welt 29), vol. 2, 509–553. Berlin: De Gruyter.

Polomé, Edgar C. 1986. The non-Indo-European component of the Germanic lexicon. In Annemarie Etter (ed.), *O-o-pe-ro-si, Festschrift für Ernst Risch zum 75. Geburtstag*, 661–672. Berlin/New York: De Gruyter.

Polomé, Edgar C. 1989. Substrate lexicon in Germanic. *North-Western European Language Evolution* 14. 53–73.

Polomé, Edgar C. 1990. Types of linguistic evidence for early contact: Indo-Europeans and non-Indo-Europeans. In Thomas L. Markey & John A. C. Greppin (eds.), *When worlds collide: The Indo-Europeans and the Pre-Indo-Europeans*. Ann Arbor: Karoma Publishers.

Porru, Vissentu. 1866. *Dizionariu sardu-italianu*. 2nd edn. Casteddu: Stamperia Nazionali.

Posth, Cosimo, Valentina Zaro, Maria A. Spyrou, Stefania Vai, Guido A. Gnecchi-Ruscone, Alessandra Modi, Alexander Peltzer, et al. 2021. The origin and

legacy of the Etruscans through a 2000-year archeogenomic time transect. *Science Advances* 7(39).

Pott, August Friedrich. 1853. *Die Personennamen, insbesondere die Familiennamen und ihre Entstehungsarten; auch unter Berücksichtigung der Ortsnamen.* Leipzig: F. A. Brockhaus.

Pradel-Baquerre, Mylène. 2013. *Ps.-Apulée, Herbier, introduction, traduction et commentaire*. Université Paul-Valéry Montpellier III dissertation.

Prémare, Alfred-Louis de. 1999. *Dictionnaire arabe-français: Établi sur la base de fichiers, ouvrages, enquêtes, manuscrits, études et documents divers*, vol. 11. Paris: L'Harmattan.

Puddu, Mario. 2023. *Ditzionàriu in línia de sa limba e de sa cultura sarda.* https://ditzionariu.nor-web.eu/it/.

Putzu, Ignazio. 2019. History of the Sardinian lexicon. (Ed.) Michele Loporcaro & Francesco Gardani. *The Oxford encyclopedia of Romance linguistics*. Oxford: Oxford University Press.

Renisio, Amédé. 1932. *Étude sur les dialectes berbères des Beni Iznassen, du Rif et des Senhaja de Sraïr: Grammarie, textes et lexique.* Paris: Ernest Leroux.

Ribezzo, Francesco. 1919. La regione Japigo-messapica nella tradizione e nei monumenti scritti dell'antichità. *Rivista Indo-Greco-Italica* III. 93–110.

Ribezzo, Francesco. 1920. La originaria unità tirrena dell'Italia nella toponomastica. *Rivista Indo-Greco-Italica* IV. 83–97.

Ribezzo, Francesco. 1933. Di un'altra parola di probabile origine mediterranea. Dial. ital. *péntoma* : etr. *penθuna*. *Rivista indo-greco-italica* 17. 87–88.

Ribezzo, Francesco. 1953. A che punto siamo con la interpretazione dell'etrusco. *Studi Etruschi* 22. 105–129.

Ritter, Hans. 2009. *Wörterbuch zur Sprache und Kultur der Twareg*, vol. 2. Wiesbaden: Harrassowitz.

Rix, Helmut. 1998. *Rätisch und Etruskisch* (Vorträge Und Kleinere Schriften 68). Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.

Rohlfs, Gerhard. 1925. Über Hacken und Böcke. *Zeitschrift für Romanische Philologie* 45. 662–675.

Rohlfs, Gerhard. 1927a. Die Quellen des unteritalienischen Wortschatzes. *Zeitschrift für Romanische Philologie* 46. 135–164.

Rohlfs, Gerhard. 1927b. Baskische Reliktwörter im Pyrenäengebiet. *Zeitschrift für Romanische Philologie* 47. 394–408.

Rohlfs, Gerhard. 1932. *Dizionario dialettale delle tre Calabrie: Con note etimologiche e un'introduzione sulla storia dei dialetti calabresi,* vol. I. Halle (Saale): Max Niemeyer.

Rohlfs, Gerhard. 1934. *Dizionario dialettale delle tre Calabrie: Con note etimologiche e un'introduzione sulla storia dei dialetti calabresi,* vol. II. Halle (Saale): Max Niemeyer.

Rohlfs, Gerhard. 1964. *Lexicon Graecanicum Italiae Inferioris: Etymologisches Wörterbuch der unteritalienischen Gräzität.* 2nd edn. Tübingen: Max Niemeyer Verlag.

Rohlfs, Gerhard. 1977. *Nuovo dizionario dialettale della Calabria*. Nuova edizione interamente rielaborata ampliata ed aggiornata. Ravenna: Longo editore.

Rolla, Pietro. 1895. *Fauna popolare sarda: Miscellanea di dialettologia e toponimia italiana*. Casale: Carlo Cassone.

Rolland, Eugène. 1900. *Flore populaire ou histoire naturelle des plantes dans leurs rapports avec la linguistique et le folk-lore*, vol. III. Paris: Librairie Rolland.

Rolland, Eugène. 1903. *Flore populaire ou Histoire naturelle des plantes dans leurs rapports avec la linguistique et le folk-lore*, vol. IV. Paris: Librairie F. Staude.

Rowland, Robert J. 1985. The Roman invasion of Sardinia. In Caroline Malone & Simon Stoddart (eds.), *Papers in Italian archaeology IV: The Cambridge Conference*, vol. 4, 99–117. Oxford: B.A.R.

Ruijgh, Cornelis Jord. 1967. *Études sur la grammaire et le vocabulaire du grec mycénien*. Amsterdam: Adolf M. Hakkert.

Salmons, Joe. 1991. Norhtwest Indo-European vocabulary and substrate phonology. In Roger Pearson (ed.), *Perspectives on Indo-European language, culture and religion: Studies in honor of Edgar C. Polomé*, vol. 2, 265–279. Washington DC: Institute for the Study of Man.

Salvioni, Carlo. 1909. Bricciche sarde. *Archivio Storico Sardo* 5. 211–246.

Salvioni, Carlo. 1914. Centuria di note etimologiche e lessicali. *Romania: Recueil Trimestriel Consacré à l'Étude des Langues et des Littératures Romanes* 43. 371–402, 560–586.

Santano Moreno, Julián. 2000. La esfinge sarda: La cuestión del paleosardo y sus parientes. *Nouvelle Revue d'Onomastique* 35–36. 153–186.

Schachermeyr, Fritz. 1929. *Etruskische Frühgeschichte*. Berlin: De Gruyter.

Schachermeyr, Fritz. 1964. *Die minoische Kultur des alten Kreta*. Stuttgart: W. Kohlhammer Verlag.

Schrader, Otto. 1883. *Sprachvergleichung und Urgeschichte. Linguistisch-historische Beiträge zur Erforschung des indogermanischen Altertums.* Jena: Hermann Costenoble.

Schrijver, Peter. 1997. Animal, vegetable and mineral: Some Western European substratum words. In *Sound law and analogy: Papers in honor of Robert S.P. Beekes on the occasion of his 60th birthday*, 293–316. Amsterdam: Rodopi.

Schuchardt, Hugo. 1885. *Über die Lautgesetze. Gegen die Junggrammatiker.* Berlin: Oppenheim.

Schuchardt, Hugo. 1899. *Romanische Etymologieen. II.* (Sitzungsberichte der Philosophisch-Historischen Classe der Kaiserlichen Akademie der Wissenschaften 141).

Schuchardt, Hugo. 1904. Ital. *corbezzolo*, span. *madroño*, sard. *olidone* "Erdbeerbaum"; franz. *micocoulier* "Zürgelbaum." *Zeitschrift für Romanische Philologie* 28. 192–195.

Schuchardt, Hugo. 1905. Ibero-romanisches. *Zeitschrift für Romanische Philologie* 29. 226–227.

Schuchardt, Hugo. 1907. Zur romanischen Wortgeschichte. *Zeitschrift für Romanische Philologie* 31(1). 1–35.

Schuchardt, Hugo. 1910. Sard. cugurra; anna e mele. *Zeitschrift für Romanische Philologie* 34(2). 212–215.

Serhoual, Mohammed. 2002. *Dictionnaire tarifit-français*. Tétouan: Université Abdelmalek Essaâdi dissertation.

Serra, Giandomenico. 1952. Etruschi e latini in Sardegna. In Gunnar Ahlborn, Sven Andolf & Lans Bergh (eds.), *Mélanges de philologie romane offerts à M. Karl Michaëlsson par ses amis et ses élèves*, 407–450. Göteborg: Bergendahls Boktryckeri.

Serra, Giandomenico. 1954a. Appunti su l'elemento punico e libico nell' onomastica sarda. *Vox Romanica* 13. 51–65.

Serra, Giandomenico. 1954b. Review of Rudolf Böhne, Zum Wortschatz der Mundart des Sàrrabus (Südostsardinien). *Romance Philology* 8. 211–220.

Serra, Giandomenico. 1956. Sostratos mediterraneos. *Revista de Filología Española* 40(1). 171–208.

Serra, Giandomenico. 1960. L'action du substrat lybique sur la structure des mots de la structure de la langue sarde. *Orbis: Bulletin International de Documentation Linguistique* 9(2). 404–418.

Sikkema, R., G. Caudullo & D. de Rigo. 2016. *Carpinus betulus* in Europe: Distribution, habitat, usage and threats. (Ed.) J. San-Miguel-Ayanz, D. de Rigo, G. Caudullo, T. Houston Durrant & A. Mauri. *European Atlas of Forest Tree Species*. Luxembourg: Publications Office of the European Union. https://ies-ows.jrc.ec.europa.eu/efdac/download/Atlas/pdf/Carpinus_betulus.pdf. (30 May, 2022).

Simon, Zsolt. 2021. The alleged Anatolian loanwords in Etruscan: A reconsideration. In Federico Giusfredi, Zsolt Simon & Elena Martínez Rodríguez (eds.), *Studies in the languages and language contact in Pre-Hellenistic Anatolia*, 227–254. Barcelona: Publicacions i Edicions de la Universitat de Barcelona.

Sluis, Paulus van, Anders Richardt Jørgensen & Guus Kroonen. 2023. European prehistory between Celtic and Germanic: The Celto-Germanic isoglosses revisited. In *The Indo-European puzzle revisited: Integrated archaeology, genetics, and linguistics*, 193–246. Cambridge: Cambridge University Press.

Sluis, Paulus van. 2024. Substrate alternations in Celtic. In Guus Kroonen (ed.), *Sub-Indo-European Europe*, 67–131. Berlin: De Gruyter Mouton.

Smoczyński, Wojciech. 2018. *Lithuanian etymological dictionary*. (Ed.) Axel Holvoet, Steven Young & Wayles Browne. Berlin: Peter Lang.

Solmi, Arrigo. 1905. Le carte volgari dell'archivio arcivescovile di Cagliari: Testi campidanesi dei secoli XI-XIII. *Archivio Storico Italiano, serie 5* 35. 273–330.

Spano, Giovanni. 1851. *Vocabolariu sardu-italianu et italianu-sardu*. Cagliari: Imprenta Nationale.

Spano, Giovanni. 1863. Sul titolo di Merre dato ad Esculapio nell'iscrizione trilingue di Pauli Gerrei. *Bullettino Archeologico Sardo* 9(9). 89–96.

Spano, Giovanni. 1872. *Vocabolario sardo geografico patronimico ed etimologico*. Cagliari: Antonio Alagna.

Stadler, Hermann. 1899. Dioscorides Longobardus. (Cod. Lat. Monacensis 337.) Aus T.M. Aurachers Nachlass herausgegeben und ergänzt. *Romanische Forschungen* 10. 181–247; 372–446.

Stadler, Hermann. 1901. Dioscorides Longobardus. (Cod. Lat. Monacensis 337.) Aus T.M. Aurachers Nachlass herausgegeben und ergänzt. *Romanische Forschungen* 11(1). 1–121.

Stifter, David, Martin Braun, Corinna Salomon, Michela Vignoli, Anna-Maria Adaktylos, Chiara Dezi, Eva Lettner, Corinna Scheungraber & Marcel

Schwarz (eds.). 2022. *Lexicon Leponticum* (*LexLep*). https://lexlep.univie.ac.at/wiki/Inscriptions. (5 March, 2024).

Strabo. 1877. *Geographica*. (Ed.) A. Meineke. Leipzig: Teubner.

Subak, Julius. 1909. Review of Gustav Körting, Lateinisch-Romanisches Wörterbuch (Etymologisches Wörterbuch der romanischen Hauptsprachen). *Zeitschrift fur Romanische Philologie* 33. 479–486.

Swanenvleugel, Cid. 2024. Prefixes in the Sardinian substrate. In Guus Kroonen (ed.), *Sub-Indo-European Europe*, 215–282. Berlin: De Gruyter Mouton.

Taïfi, Miloud. 1991. *Dictionnaire tamazight-français* (*Parlers du Maroc central*). Paris: L'Harmattan-Awal.

Terracini, Benvenuto. 1927. Osservazioni sugli strati più antichi della toponomastica sarda. In *Il convegno archeologico in Sardegna, giugno 1926*, 137–150. Reggio nell'Emilia: Officine grafiche reggiane.

Terracini, Benvenuto. 1936. Gli studi linguistici sulla Sardegna preromana. *Sardegna Romana* 1. 53–80.

Terracini, Benvenuto. 1959. L'ALI e i problemi della dialettologia italiana - note di geografia linguistica: Le denominazioni del pipistrello in Sardegna. *Bollettino dell'Atlante Linguistico Italiano* 3–4. 10–21.

Theophrastus. 1916. *Enquiry into plants and minor works on odours and weather signs*. (Trans.) Arthur Hort, vol. I. London: William Heinemann.

Thomason, Sarah Grey & Terrence Kaufman. 1988. *Language Contact, creolization, and genetic linguistics*. Berkely/Los Angeles/London: University of California Press.

Tilmatine, Mohand & Joaquín Bustamante. 2002. La fitonimia amazige en la ʿUmdat aṭ-ṭabīb. *Al-Andalus Magreb* 8–9(2). 461–490.

Toso, Fiorenzo. 2017. Il tabarchino. In Eduardo Blasco Ferrer, Peter Koch & Daniela Marzo (eds.), *Manuale di linguistica sarda*, 446–459. Berlin/Boston: De Gruyter.

Trabut, Louis Charles. 1935. *Flore du nord de l'Afrique: Répertoire des noms indigènes des plantes spontanées, cultivées et utilisées dans le nord de l'Afrique.* Alger: Imprimeries "La typo-Litho" et Jules Carbonel réunies.

Traina, Antonino. 1877. *Vocabolarietto delle voci siciliane dissimili dalle italiane.* Torino: G.B. Paravia e comp.

Trask, Robert Lawrence. 1995. Origins and relatives of the Basque language: Review of the evidence. In José Ignacio Hualde, Joseba A. Lakarra & Robert Lawrence Trask (eds.), *Towards a history of the Basque language*, 65–100. Amsterdam: John Benjamins Publishing Company.

Trask, Robert Lawrence. 1997. *The History of Basque*. London: Routledge.

Trask, Robert Lawrence. 2008. *Etymological Dictionary of Basque*. (Ed.) Max W. Wheeler. University of Sussex.

Trudgill, Peter. 2003. Modern Greek dialects: A preliminary classification. *Journal of Greek Linguistics* 4. 45–63.

Trudu, Enrico. 2012. Civitates, latrunculi mastrucati? Alcune note sulla romanizzazione della Barbaria. In Maria Bastiana Cocco, Alberto Gavini & Antonio Ibba (eds.), *L'africa romana: Trasformazione dei paesaggi del potere nell'Africa settentrionale fino alla fine del mondo antico*, vol. 3, 2645–2659. Roma: Carocci editore.

Untermann, Jürgen. 1969. Lengua gala y lengua ibérica en la Galia Narbonensis. *Archivo de Prehistoria Levantina* 12. 99–161.

Untermann, Jürgen. 1980. *Monumenta linguarum hispanicarum*, vol. II. Wiesbaden: Dr. Ludwig Reichert Verlag.

Untermann, Jürgen. 1983. Die althispanischen Sprachen. In Wolfgang Haase (ed.), *Teilband Sprache und Literatur* (Aufstieg und Niedergang der Römischen Welt 29), vol. 2, 791–818. Berlin: De Gruyter.

Vaan, Michiel de. 2008. *Etymological dictionary of Latin and the other Italic languages.* Leiden: Brill.

Van der Meer, Lammert Bouke. 2004. Etruscan origins: Language and archaeology. *Bulletin Antieke Beschaving: Annual Papers on Mediterranean Archaeology* 79. 51–58.

Varro, Marcus Terentius. 2015. *De lingua latina: V-VII*. (Trans.) Lucie Pultrová. Praha: Filozofická fakulta Univerzity Karlovy.

Vasmer, Max. 1955. *Russisches etymologisches Wörterbuch*, vol. 2. Heidelberg: Carl Winter Universitätsverlag.

Vattioni, Francesco. 1976. Glosse puniche. *Augustinianum* 16. 505–555.

Velaza, Javier. 2019. Iberian writing and language. In Alejandro G. Sinner & Javier Velaza (eds.), *Palaeohispanic languages and epigraphies*, 160–197. Oxford: Oxford University Press.

Verhasselt, Gertjan. 2011. The Pre-Greek linguistic substratum: A critical assessment of recent theories. *Les Études Classiques* 79. 257–283.

Virdis, Maurizio. 1988. Sardisch: Areallinguistik. In Günter Holtus, Michael Metzeltin & Christian Schmitt (eds.), *Lexikon der romanistischen Linguistik*, vol. 4, 897–912. Tübingen: Max Niemeyer Verlag.

Virdis, Maurizio (ed.). 2003. *Il Condaghe di Santa Maria di Bonarcado*. Nuoro: Ilisso.

Virdis, Maurizio. 2017. Superstrato spagnolo. In Eduardo Blasco Ferrer, Peter Koch & Daniela Marzo (eds.), *Manuale di linguistica sarda*, 168–183. Berlin: De Gruyter.

Virdis, Maurizio. 2018. Sociolinguistica storica nella Sardegna medievale. Aspetti, problemi, affioramenti: Fra dialettologia e filologia. In Giulio Paulis, Ignazio Putzu & Maurizio Virdis (eds.), *Il sardo medioevale: Tra sociolinguistica storica e ricostruzione linguistico-culturale*, 11–34. Milan: Franco Angeli.

Virdis, Maurizio. 2020. Problemi di diatopia e di diacronia della lingua sarda. Un'ipotesi di sociolinguistica storica. In Eva-Maria Remberger, Maurizio Virdis & Birgit Wagner (eds.), *Il sardo in movimento*, 31–48. Göttingen: V&R Unipress.

Vona, Giuseppe & Carla Maria Calò. 2006. History of Sardinian population (Italy, Western Mediterranean) as inferred from genetic analysis. In Carla Maria Calò & Giuseppe Vona (eds.), *Human Genetic Isolates*, 1–28. Kerala: Research Signpost.

Wagner, Max Leopold. 1907a. Gli elementi del lessico sardo. *Archivio Storico Sardo* 3. 370–419.

Wagner, Max Leopold. 1907b. *Lautlehre der südsardischen Mundarten: Mit besonderer Berücksichtigung der um den Gennargentu gesprochenen Varietäten.* Halle (Saale): Max Niemeyer Verlag.

Wagner, Max Leopold. 1921. *Das ländliche Leben Sardiniens im Spiegel der Sprache*. Heidelberg: Carl Winter's Universitätsbuchhandlung.

Wagner, Max Leopold. 1922. Los elementos español y catalán en los dialectos sardos. *Revista de Filología Española* 9(3). 221–265.

Wagner, Max Leopold. 1928. La stratificazione del lessico sardo. *Revue de Linguistique Romane* 4. 1–61.

Wagner, Max Leopold. 1930. *Studien über den sardischen Wortschatz* (Biblioteca Dell'Archivum Romanicum: Serie II vol. 16). Genève: Leo S. Olschki.

Wagner, Max Leopold. 1932. Über die vorrömische Bestandteile des Sardischen. *Archivum Romanicum: Nuova Rivista di Filologia Romanza* 15. 207–247.

Wagner, Max Leopold. 1933. Die Bezeichnungen für "Fuchs" in Sardinien. *Archivum Romanicum: Nuova Rivista di Filologia Romanza* 16. 501–514.

Wagner, Max Leopold. 1934. Weitere sardische Tiernamenstudien. *Archivum Romanicum: Nuova Rivista di Filologia Romanza* 18. 1–18; 481–494.

Wagner, Max Leopold. 1940. Über die neuen Ausgaben und die Sprache der altsardischen Urkundenbücher von S. Nicola di Trullas und S. Maria di Bonarcado. *Vox Romanica* 5. 106–164.

Wagner, Max Leopold. 1941a. *Historische Lautlehre des Sardischen*. Halle (Saale): Max Niemeyer Verlag.

Wagner, Max Leopold. 1941c. Zu den sardischen Etymologien Alessios und Pisanis. *Zeitschrift für Romanische Philologie* 61. 319–338.

Wagner, Max Leopold. 1943. Review of Giovanni Nencioni, Innovazioni africane nel lessico latino. *Zeitschrift fur Romanische Philologie* 63. 195–198.

Wagner, Max Leopold. 1944. Zum Paläosardischen. *Vox Romanica* 7. 306–323.

Wagner, Max Leopold. 1952. *Historische Wortbildungslehre des Sardischen*. Bern: A. Francke.

Wagner, Max Leopold. 1957. Die Punier und ihre Sprache in Sardinien. *Die Sprache* 3. 27–43, 78–1090.

Wagner, Max Leopold. 1984. *Fonetica storica del sardo*. (Ed.) Giulio Paulis. Cagliari: Gianni Trois Editore.

Wagner, Max Leopold. 1997. *La lingua sarda. Storia, spirito e forma* [1950]. (Ed.) Giulio Paulis. Nuoro: Ilisso.

Walde, Alois & Johann Baptist Hofmann. 1938. *Lateinisches etymologisches Wörterbuch*. 3rd edn., vol. 1. Heidelberg: Carl Winter's Universitätsbuchhandlung.

Walde, Alois & Johann Baptist Hofmann. 1954. *Lateinisches etymologisches Wörterbuch*. 3rd edn., vol. 2. Heidelberg: Carl Winter's Universitätsbuchhandlung.

Wallace, Rex E. 2008. *Zikh Rasna: A manual of the Etruscan language and inscriptions.* Ann Arbor: Beech Stave Press.

Webster, Gary S. 1996. *A prehistory of Sardinia 2300-500 BC*. Sheffield: Sheffield Academic Press.

Weiss, Michael. 2009. *Outline of the historical and comparative grammar of Latin.* Ann Arbor: Beech Stave Press.

WFO. 2021. *Helosciadium nodiflorum* (L.) W.D.J. Koch. http://www.worldfloraonline.org/taxon/wfo-0000719285. (6 April, 2023).

WFO. 2022a. *Echinops spinosissimus* subsp. *spinosissimus* Turra. https://wfoplantlist.org/plant-list/taxon/wfo-0000102217-2022-12?matched_id=wfo-0000013250&page=1. (20 April, 2023).

WFO. 2022b. *Dittrichia viscosa* subsp. viscosa (L.) Greuter. https://wfoplantlist.org/plant-list/taxon/wfo-0000101827-2022-12?matched_id=wfo-0000124293&page=1. (11 April, 2023).

WFO. 2022c. *Hirschfeldia incana* (L.) Lagr.-Foss. https://wfoplantlist.org/plant-list/taxon/wfo-0000724016-2022-12. (3 May, 2023).

WFO. 2023. *Ruscus aculeatus* L. https://wfoplantlist.org/taxon/wfo-0000732428-2023-12?matched_id=wfo-0000771488&page=1. (31 January, 2024).

WFO. 2024a. *WFO Plant List*. https://wfoplantlist.org/ (15 October, 2024).

WFO. 2024b. *Dracunculus vulgaris* Schott. https://wfoplantlist.org/taxon/wfo-0000945756-2024-06?matched_id=wfo-0000278789&page=1 (14 october 2024).

WFO. 2024c. *Micromeria graeca* (L.) Benth. ex Rchb. https://wfoplantlist.org/taxon/wfo-0000243277-2024-06?matched_id=wfo-0000304182&page=1. (15 October 2024).

Wigman, Andrew Michael. 2023. *Unde vēnistī: The prehistory of Italic through its loanword lexicon*. Leiden: Universiteit Leiden dissertation.

Wolf, Heinz Jürgen. 1988. Sardisch: Interne Sprachgeschichte III. Onomastik. (Ed.) Günter Holtus, Michael Metzeltin & Christian Schmitt. *Lexikon der romanistischen Linguistik*. Tübingen: Max Niemeyer Verlag.

Wolf, Heinz Jürgen. 1998a. *Toponomastica barbarcina: I nomi di luogo dei comuni di Fonni, Gavoi, Lodine, Mamoiada, Oliena, Ollolai, Olzai, Orgosolo, Ovodda.* Nuoro: Insula.

Wolf, Heinz Jürgen. 1998b. Sarde *pettorra*, *liporra*, le mérinos et le suffixe *-inu*. *Revue de Linguistique Romane* 62. 332–346.

Wolf, Heinz Jürgen. 2011. La question du paléosarde. *Revue de Linguistique Romane* 75. 595–622.

Zimmern, Heinrich. 1917. *Akkadische Fremdwörter als Beweis für Babylonischen Kultureinfluss*. 2nd edn. Leipzig: J.C. Hinrichs'sche Buchhandlung.

Index

Basque

agin, 161
ander, 83
andere, 484
arba, 486, 512
arroil, 324
aurrede, 133, 135
aurri, 133, 135
barra, 388, 490, 494, 521
bonburu, 288
ezpel, 360
garbazta, 486, 512
gastigar, gaztigar, astigar, 137, 485, 492, 494, 511
gilharre, 459, 485
gillar, 38, 485
gonburu, 288
gorosti, 32, 132, 484, 494, 495, 511
hagin, 161, 486, 492
hotz, 322
hozpil, 322
ibai, 296, 488
ihar, 137, 485, 492, 511
ilharri, 485
istil, 308
itil, 308
karbasta, 167, 486, 492, 512
karbaza, 167, 486, 492, 512
kokor, 376, 490
konkor, 490
kukur, 376, 490
kukutz, 376, 490
kunkur, 490
lakarra, 334
legarra, 334
lito, 283, 487
lütho, 283, 487
mehe, 139
moko, 289
mokor, 288, 487, 516
mokoru, 488, 494, 516
mora, 339, 489, 494, 501, 519
more(a), 339, 489, 501, 519
mukurru, 288, 488
okor, 288, 487
ospel, 322, 445, 489, 494
pitika, 263
sai, sari, 236, 479, 486
sokor, 288, 487
suranguilla, 227
tistil, 308
troka, 310, 488, 492, 517
tutur, 376, 490
tuturru, 376, 490
txakur, 486, 492, 513
txistil, 308
txokorro, 56
za(h)i, 345
zalauts, 345
zapalda, 322
zume, 139
zur, 139

Berber

aktun Kb., 49, 468
asiwan Rif., 244, 479, 482
āwurmī ʿUmdat aṭ ṭabīb, 90
aẓəmẓum Kb., 109, 479
ažərtil Tmz., Tsh., Rif., 368
azəzzu Kb., 109, 478, 482
aziir, 148, 470
azir Rif., Tmz., 148, 399, 470
āzīr ʿUmdat aṭ-ṭabīb, 148, 470
azmu Tmz., 109, 479
əɣruṛəd Kb., 368, 481, 482
əzzuṭ Snh., 376, 481, 482

ɣərd Tmz., 368, 481, 482
iazir, 148, 470
isɣi Chenoua, 236, 479, 482
isni Kb., 106, 478
iwrmi Tsh., 90
ouzbir, 148, 470
sakrā ʿUmdat aṭ ṭabīb, 66, 468
taga Ihahan, 161, 479, 482
tamaṭṭa Snh., Rif., 178, 479, 482
tāsakra ʿUmdat aṭ ṭabīb, 66, 468
tāsikrā ʿUmdat aṭ ṭabīb, 66, 468
tasiwant Rif., 244, 479, 482
taskra Tsh., 66, 399, 468
taskrat, 66, 468
tesănit Tuar., 106, 478
teskera, 66, 468
tiqi Igliwa, 161, 479, 482
tiqqi Ouarzazate, 161, 479, 482
tisnit, tiznit Tazerwalt, 106
tiyzzi Beni Snous, 109, 478, 482
tsənnít Nef., 106, 478
tsunít Nef., 106, 478
ullī, 353, 480
waktun Kb., 49, 468
wəllī, 353, 480
ẓẓəmẓəm Kb., 109, 479
ẓẓiw Beni Snous, 109, 478, 482

Gallo-Romance

agaloûssës Lang., 132, 484, 495, 511
agast OProv., 32, 137, 485, 494, 511
alète Auv., 105, 510
ãndèr Lang., 83, 484, 510
ắrẹ Prov., 32, 38, 105, 510
aroû Bearn., 105, 510
arrouil, arrouio Gasc., 324, 519
arroulho Bearn., 324, 519
aubredé Gasc., 135
auledun Gasc., 152
barta OProv., 388, 490, 521
bartas Lang., 388
barto Lang., 388
bette For., 263
bitte For., 263
cala OProv., 30, 283, 515
calanca OProv., 315, 518
câlo Lang., 283, 515
caranco Prov., 315, 518
casse Fr. dial., 306
coucouro Prov., 188
cuco Prov., 188
dervese Lang., 83
ẽdẹ́r Lang., 83
ender Lang., 83
enderce MFr., 83
endèrs Lang., 83
foie Fr., 32
garbalhe Gasc., 169, 486, 512
garbassat Lang., 169
garbasso Lang., 169, 494
garbo Prov., 167, 486, 512
garboucha Gasc., 169
garbout Gasc., 169
gratte-cul Fr., 122
kalãnko Prov., 315, 518
kanãnko Prov., 315, 518
katẹ́ Aostan, 306
lapéreau Fr., 410
laperel OProv., 515
lapin Fr., 271, 455, 515
lapriel OFr., 515
laurier-rose Fr., 130
leduno Lang., 152
lodissa OProv., 364

Lyon Top. Fr., 33
maghêlo, *maghiêlo* Lang., 289, 488, 516
maguelo Prov., 289, 488, 516
maguèro Gasc., 289
marràlh Gasc., 312, 518
marro Prov., 518
marroc Lang., 518
morāina Frpr., 339
morêna Frpr., 339, 489, 501, 520
mouche Lang., 146, 499, 504, 512
moudre Lang., 146
moudré Lang., 146
mouflon Fr., 275
mouge Lang., 146, 499, 512
mourre Prov., 339
olon Saintongeais, 152
onder Lang., 83
õndèr Lang., 83, 484, 510
orousto Lang., 38, 105, 510
pendo Prov., 308
pente Fr., 308
pito Gasc., 263
renard Fr., 265
sarge OFr., 364
serge Fr., 364
setgia OProv., 299, 517
sialet Dauphinois, 299, 517, 527
siejo Lang., 299, 517
ϑalę Sav., 283, 515

Greek

azzìddastro Bov., 132, 511, 531
glimbáći Bov., 140
pilimbáći Bov., 140
plembáći Bov., 140
popáttso Calab. Gr., 351
źojero Bov., 116
ἀζώγῐ̯ουρας MGr., 118
ἀζώγουρας MGr., 116, 118
ἀζώγυρος MGr., 116
ἀζώουρος MGr., 116, 118
ἀζώϋρο MGr., 116
ἀθερηίς, 120, 121
αἰθέριος, 120
ἄκαστος Hes., 137, 485, 494, 511
ἀμάρα, 293, 294, 516
ἀνάθημα, 121
ἄνθος, 83, 484, 510
ἀρίσαρον, 51
ἁρμαλά Diosc., 89, 469, 474
ἄρον, 51
ἄσαρον, 51, 52, 509
ἀττέλαβος, 204
βίβλος, 140
βόθρος, 500
γέρρον, 369
ζαγάρι MGr., 486, 513
ζαγάριον Byz., 486, 513
θεῖος, 121, 394
Θεοφάνης, 244, 479
θύλλος, θύλλα Hes., 374
θύμον, 118
θύρσιον Diosc., 146
θῶμιγξ, 108
ἰτεός, 139
κάγχαλος, 212
καιάδᾱς, 299, 517, 532, 533
καίατα, 299, 517
κάραβος, 330
καράμβας Hes., 76
κάστανον, 137
κηλάστραι Hes., 132
κήλαστρος, 132, 484, 495
κολλυρίων, 244
κύλλας Hes., 262, 514
λεβηρίς Massiliot, 271, 410, 515
μαγούλα MGr., 289, 488, 516
μέλανος, 192
μέλας, 192
μέλλαξ, 192
μούτουκα Diosc., 146
ὀνοθοῦρις, 130

ὀργάς, 290
ὀρυγή, 326
ὀρύσσω, 326
ὀρυχή, 326
σίκερα, 64, 468
σικκιριά Diosc., 64
σκύλαξ, 262, 514
τζοῦρος Cypr., 353, 520
τσέργα MGr., 362
τυφλός, 373
φάλαγξ, 448
χαλάω, 283
χλωρίων, 244
χουρμά Diosc., 90

Ibero-Romance

Almería Sp. Top., 33
arroio Pg., 324, 519
arroyo Sp., 324, 326, 519
arruoyo Leon., 326
astor Cat., 242
axebe OSp., 94
baldíssa Algh., 389
barda Cat., 457, 490, 494, 521
barda Cat., Pg., Sp., 388
barda Pg., 490, 521
bardissa Cat., 388, 490, 494, 521
bita Cat., 263
bitè Cat., 263
cabellera Sp., 94
cala Cat., Sp., 30, 281, 515
cáncano Sp., 212
carabullos Gal., 169
carba Ast., 167
carba Sp., 169
carbana Ast., 169
carvalho Pg., 169
carvallo Gal., 169
carvayo Leon., 169
cea Arag., 299
churra Sp., 353
cía Cat., 299, 517
cigarra Sp., 210
cija Arag., 299
coca Arag., 178, 499, 513
cócora Pg., 178, 499, 513
colostia, golostia, corostia, gorostia Navarr., 132, 484, 511, 525
comensar Cat., 296
cuco Sp., 188
desmoronar Sp., 340
doble Cat., 203
dóbra Algh., 203
doll Cat., 328
embraçar Cat., 296
enxebe OSp., 94
escabeche Sp., 94
escabetx Cat., 94
escarabajas Sp., 167
esparver Cat., 239
firma Sp., 124, 126
gar(r)abullos Gal., 169
gáraba Sp., 169
garabellos Gal., 169
garabizos Gal., 169
gárabu Ast., 167
garabujos Gal., 169
garamallos Gal., 169
garba Ast., 494
garbu Ast., 167, 169, 486, 512
garbucu Ast., 169
garrampallos Gal., 169
garvalha Pg., 169
gravalho Pg., 169
gravulha Pg., 169
hígado Sp., 32
jebe Sp., 94
jerga Sp., 364
jerpa Sp., 161
lána Algh., 190
láparo Pg., 271, 410, 515
lapim Pg., 271, 455, 515
laurel de flor Sp., 130
León Sp. Top., 33
llapí Arag., 271, 455, 515
lluquet Cat., 53
lluqueta Cat., 53
malecón Sp., 317
mocador Cat., 216
mocorró Old Cat., 289, 488

módega Cat., 146, 499, 504, 512
mogote Sp., 289
mogró Cat., 488
mogró, mugró Cat., 289
moguró Old Cat., 289
moreas Gal., 339
moreia Pg., 339, 489, 501, 520
morena Sp., 339, 489, 501, 505, 520
moroiço, moiroiço Pg., 339
morón Sp., 339, 340
mugaró Cat., 289
mugrón Sp., 289
muguró Old Cat., 289
múrgola Cat., 60
noiro Ast., 279, 515
noriu Ast., 279, 515
nórrago Galic., 279, 515
otoño Sp., 175
quer Cat., 319
sarda, sardu, zarda Ast., 368
sarga Cat., 161
sargantana Arag., 227
segalh Cat., 326
senill Cat., 108
serpa Sp., 161
sía Cat., 299, 517
silibrich volador Algh., 203
sitges Cat., 299, 517
sitja Cat., 299, 517
síya Cat., 299, 517
tòa Algh., 137
toll Cat., 328
torca, torco Sp., 310, 488, 492, 517, 525
vajka Old Leon., 488
vega Sp., 35, 296, 488
veiga Pg., 296
zade Sp., 139
zarpa Sp., 161
ziliblich Algh., 203
θía Cat., 299, 517

Italo-Romance

alastra Sic., 105, 510, 531
amùndaci, amùndici Calab., 146, 499, 512, 531
amùndici Calab., 499
anzana, anzara Calab., 51, 52, 509
anzara, anzaru Sic., 51, 52, 509
anzareddu, azareddu Sic., 51
arástra Lig., 105, 510, 529
autunno It., 175
azara Pugl., 51, 52, 509
azzaru Sic., 51, 429, 509
bìlleri Tusc., 74, 498, 509
bóḍḍaru Sic., 499, 505, 513
bòllora Tusc., 499, 505, 513
borro Tusc., 500
borròne Tusc., 500
bótro Tusc., 500
burrone Tusc., 500
cagliare It., 306
cala Cors., It., Sic., 30, 281, 515
calanca Cors., 315, 449
calanca, calônca Cors., 315, 518
calanco Tusc., 315, 518
canchero It., 212
carapone Laz., 330, 501, 519
cavalletta It., 196
ceca It., 216, 218
cègghiastru Calab., 132, 511, 531
cèlastru Lucanian, 132, 511, 531
céole, cévole Friul., 299, 517, 529
čiálę Elban, 299, 517, 530
cieco It., 371

ciorla Lomb., Trent., 353, 520
ciorro Umbrian, 353, 520
cìstriddazzu Calab., 132, 511, 531
còccola Nap., 178, 499, 513
cóccoro Tusc., 178, 499, 513
cocla Romagnol, 178, 499, 513
čorla Lad., 353, 520
còstolo Tusc., 32, 137, 485, 511, 530
elicriso It., 53
fegato It., 32, 85
ɣarammo̍nə, *ɣarammo̍nu* Lucanian, 317
garroppa Calab., 330, 501, 505, 519
ghiàgaru Cors., 258
ghiaia It., 285
ghiara Tusc., 285
ghjácaru Cors., 258, 486, 513
giustizia It., 366
ɟágaru, *ɟákaru* Cors., 258
iara Tusc., 285
kalankə Calab., 315, 518
karafúnu Calab., 330, 501, 519
karravone Tusc., 330, 501, 519
karravóni Cors., 330, 501, 519
kəkǫ́ćća Abruz., 351, 520
kùkkẹrẹ Laz., 178, 499, 513
kwátenə Campan., 38
lamár Lad., 293
lastra Lig., 105, 510
lástroga Lig., 105, 510
legone Tusc., 334
limbáći Calab., 140
lürion Pied., 152
mar Lad., 293, 516
mara Cors., 293, 294, 516
mara Lad., 293
maragghiuni Sic., 336
maragone Abruzz., 317
marana Laz., 293, 516
marazzo It., 516
marecia Ven., 293, 516
marena Lad., 293
maròc Trent., 518
maròɡn Lomb., 312, 518
maròk Lad., 518
mola Cors., 339, 489, 501, 519
mora Cors., Laz., Tusc., 339, 489, 501, 505, 519
mucchio Tusc., It., 146
mucchiu Calab., 146
mucchju Cors., 146
múćću Cors., 146
mufra, *muvra* Cors., 275
mufrone, *muvrone* Cors., 275, 500, 515
muk'k'u Salent., 146
múndaca Calab., 146, 499, 512, 531
mútaca Calab., 146, 499, 504, 512, 531
oleandro It., 140, 142
oro It., 406
papáttsa Calab., 351
pəkǫ́ttsa Abruz., 351, 520
péndəm^{a} Campan., 306, 500, 517
pèndima Laz., 306, 500, 517
péŋ́gima Campan., 306, 500, 517
péntica Cors., 308
péntima Salent., 306, 500, 517
pèntuma Salent., 306, 500, 517
pentúoškə Campan., 308
piluccare Tusc., 500
roa Lad., 324, 519
róğa Lomb., 324, 519, 529
roia Lad., 324, 519
roye Friul., 324, 519
roza Ven., 324, 519
ruğa Piem., 324, 519, 529
sápara Cors., 322

sargia It., 364
scarrupà Abruzz., 330, 501, 519
scarrupo, *sgarrupo* Nap., 330, 501, 505, 519
scebba Sic., 94
scerba Sic., 94
sgarrupá Irpino, 330, 501, 519
skalvắ Lombard, 169
ṣkravắ Emilian, 169
tarantola It., 221
tesoro It., 406
Tónnari Top. Cors., 319
trogolo It., 328
tsurrə Abruz., Nap., Pugl., 353, 520
túbolo It., 76
uoḍḍa Calab., 353, 520
ürion Pied., 152
vangaronu Cors., 317
varangonu Cors., 317
źanźarra Calab., 51, 52
zara Calab., 51, 52, 509
zara Sic., 51
zélle Abruz., 385
zellla Nap., 385
zirriari Sic., 394
zułę CIt., 139
zurra Calab., 353, 520

Latin

absinthium, 394
acer, 137
acinus, 161
aetheria, 120
afrissa, 68
ala, 167
alaternus, 105, 510
amurca, *amurga*, 60
anguilla, 218
arbor, 116
arbutus, 152
arisarum, 51, 52
aron, 52
arrugia, 324, 519
asaro, *asaru*, 52, 509
asaru, 51, 52
ātrōx, 440
autumnus, 173, 175
berula, 498, 510
biblace, 140
bicornis, 210
caetra, *cētra*, 368, 480, 481
calāre, 283
calcāre, 112
cancer, 212, 214
cannabis, *cannabu*, 345
Caralis Top., 319
carbō, 169, 486, 512, 530
carduum, 57, 60, 502
casa, 29
catellus, 262
catena, 212, 214
catinus, 115
catone, 49, 447
catus, 49
cherda, 368, 481
chlōriō, 244
cicer, 56
cicer rusticum, 56
cicūta, 63, 498
cieō, 394
cīmex, 440
citāre, 394
coactilis, 306
coagulare, 306
coagulāre, 306
coccum, 180
cōchlea, 188
collēgium, 301
conus, 142
cor, 406
cornīx, 440
cornu, 432
cornū, 209
cornus, 230
corrugus, 326
cras, 185
crusta, 185
crux, 185
cucutium, 64
cumulus, 288, 488
cunnus, 68, 70
curculiō, 188
cūria, 502
currus, 502
curūlis, 502

domō (abl.), 29
domus, 406
eryngion, 120
exīt, 404
expaginare, 161
faciō, 393
febrim, 402
ferāx, 440
ferus, 89
ficatus, 85, 86
flamma, 124
flammeus, 124, 126
fringilla, 246
fūnis, 108
furca, 184
gemellus, 442
glans, 267
glārea, 285
gregārius, 348
gurguliō, 188
helicrysum, 81
ianua, 29
īliceus, 283
imbrōs, 276, 505
iteon, 139
iūgis, 102
iugulum, 379
iūs, 85, 86
iustitia, 366
lacerta, 223, 429, 432, 435
lactoris, 440, 442
lacūna, 334
lalisiones, 446
lancearius, 51
legiō, 33
lepus, 271, 407, 409, 410, 514
lepus, -oris, 269, 271
liber, 140
lōdīx, 364
lolium, 281
lucus, 122
Lugdūnum Top., 33
macula, 60
malva, 116
mare, 293, 294, 516
marga, 317
Marianus PN, 265
marum, 57, 59, 60
matta, 178
mergere, 289
merula, 456
mūcrō, 488
mūcrō, 289
mufron, 275, 276, 447, 500, 504, 515
mūrex, 317
mūrus, 340, 501
musmō, 275, 276, 505
musmones, 446
mutilus, 60
napurae, 502
nasturtium, 98
nebula, 140
nōdus, 115
nōmen, 139
NURAC, 8, 418, 424, 441
oleástru, 132
onotirum, 130
orbus, 371
parietina, 339
peditare, 115
pendere, 308
perdīx, 440
pētra, 373
pilāre, 500
pīnētum, 438
pīneus, 142
pīnus, 86, 142, 438
porta, 29
puteus, 393
quattuor, 139
rubus, 122
rūga, 324
ruta, 90
sale, 85
sālpūga, 206, 409
sanguen, 139
sappīnus, 142
sarpō, 499
scirpus, 499
sectilis, 304
senex, 204
sērica, 364
sērō, 406
serpes, 394
siliqua, 394
sirpus, 499
sitis, 402
socer, 60
solifuga, solipuga, 206, 214, 409, 443
sucula, 379
suppūrāre, 371
taxus, 164, 165
Thasius Plin., 164, 165
theldones, 446
thius, 394
tithymalus, 76
torques, 310, 312, 489, 518
trulla, 382

tuber, 128
tūber, 128
tūfa, 128
turbare, 79
turbāre, 77
turbiscus, 77, 79
turpis, 371
ūnedō, 150, 152, 447
verus, 59, 89
vervēx, 73
vindēmia, 73
vīnētum, 438
vīnum, 438
vulpēs, 167

Sardinian

aǧǧayarai, **258**
aíssu, **152**
aíϑϑo, **152**, 154, 395, 400, 406, 408
aíϑϑu, **152**, 406, 407
alakáṡu, **53**
aláṡe, 32, **102**, 105, 130, 403, 405, 457, 510
aláṡi, **102**, 403, 405
aláṡi̯u, **102**, 403, 405, 457, 510
aláṡu, **102**, 105, 406, 409, 411
alaússa, 70, **85**, 86
alaϑúli, 105, **128**, 395, 403
alicúcu, **228**, 429, 432, 435
aliɣèlta, **223**
aliɣèrta, **223**, 225, 413, 415, 428, 429, 435, 458
aliǧèstra, **223**, 413
aližèrta, **223**
allákkana, **332**
alúsara, **157**
alússara, **157**, 413, 414, 421
ambíḍḍa, 218
andzalòra, **206**
andzalóriɣa, **206**
andzèṡa, **196**
angesa, **196**
anǧèṡa, **196**
anǧulètta, **223**
angwíḍḍa, 218
ánna, **81**, 421, 429, 484, 510
annurilare, **115**
ansíle, **272**
antsònka, **250**, 252
antúnna, **173**, 421
antúnnu, **173**
argáǧǧu, 348
argárǧu, 348
argáša, **348**, 422
argašare, **348**
arǧiléstru, **223**
arišéu, 59
arresólu, **388**
arròa, **324**
arroia OSrd., 326
arròya, **324**
arróyu, **324**
arrú, 122
árrui, **133**, 403
artiòkkoro, **56**
Árzana Top., 420
asíli, **272**
assaíle, **272**, 415
asseíle, **272**, 415
assettiliɣáðu, **304**
assíle, **272**, 415, 428, 445
assíle Gallur., **272**
astòre, 242
asturólia, **242**
ateliɣèlta, **223**
atigusa, 63
atigúsa, **62**, 63, 435
atóndzu, 175
atónǧu, 175
attaliɣèlte, **223**
attaliɣúɣe, **227**
attaliβíχχe, **203**
attaltingòne, **198**
attèrga, **362**
atteriβílke, **203**
atterringòne, **198**
attídza, **385**
attilibíske, **203**
attiliɣèrta, 63, **223**
attiliɣúɣu, **227**
attilingia, 228
attilinǧòne, **198**
attiliβílke, **203**
attiliβríu, **239**, 429, 435
attiloría, **242**

attòa, **137**
attolinǧòne, **198**
attsára, **157**, 421
attsíbba, **94**
attularia, **242**
attulínǧa, **198**
atturulía, **242**
atúndzu, 175
atúnǧu, 175
Atzara Top., 157
aúrri, **133**, 135, 403
áurri, **133**
aússa, 70, 85, 498
aussára, **157**
auttsáða, 157
auttsára, **157**, 395, 413, 414, 421
aužayarai, **258**
áybburu Sass., 116
azzacauga, 228
aɣáṇḍa, **81**, 395, 400, 416, 451, 484, 510
aɣɣéɣɣu, 394
baballótti, 96
babbakòrra, **209**
babbalótti, 96
babbalúya, **228**
babbaśáu, **192**, 194
babbaśáyu, **192**, 194
babbóį, 209, 228
baɣaðíu, 351
bagaliòni Gallur., **150**
balatúle, **128**
bappakòrra, **209**
bariśòne, **388**, 447, 490, 521
barrákka, 210
barrakórru, 210
barrakórru Gallur., **209**
barraśólu, **388**, 490, 521
barraśòne, **388**, 447, 457, 490, 521
barraśòni Sass., **388**
barrittsòne, **388**
barrókku, 500
basíli Sass., **271**
bassíli Gallur., **271**
báttile, 306
báttoro, 139
Baunéi Top., 438
baβaðía, 351
baβáðia, 351
bèɣa, **294**, 296, 488
bega OSrd., 488
bèra, 89
berriśòne, **388**
berβèɣe, 73
béssiði, 404
bèta, **262**
bètta, **262**, 404, 405, 415, 419
bètte, **262**, 263, 403, 404, 405, 415, 419, 420, 423
bétu, **262**, 405
bibbiɣòrra, **209**, 210
bibillótti, 96
bibillótti surname, 96
biḍḍúri, **73**, 403, 421
biðíle, **308**
biðíli, **308**
biðílli, **308**
biðíri, **308**
biɣòrra, **209**, 210
bilindzòne, 154
biliśòne, **154**, 447
biliśòni, **154**
billótti, **96**, 403, 404, 452
billótti surname, 96
billóttiri, **96**, 246, 403, 404, 452
binnènna, 73
bíta, **262**, 403, 404, 418, 452
bítalu, **263**
bitaréḍḍu, **263**
bítaru, **263**
bíteru, **263**
bítta, **262**, 405, 415, 419
bíttara, **263**, 452
bíttera, **263**
bítti, **262**, 263, 403, 404, 405, 415, 418, 419, 420, 423, 452
bítula, **263**
bítulu, **263**
bíturu, **263**, 452
bobbóį, 228
bòḍḍa, **353**, 520
bóina, 188
borrókku, 500
brínči, **246**, 250, 403, 404, 405, 418, 457
brínčis, **246**, 401, 404, 413, 418, 452, 457
brínčiśi, **246**, 403, 404, 418, 457
brinkare, 203
brinkíḍḍi, **246**, 250, 401, 403, 413, 418

brinzis cannèḍḍa, **246**
bríši, **246**, 401, 413
búḍḍa, 74
búḍḍaru, **73**, 74, 421, 449, 451, 498, 509
buḍḍúri, **73**, 74, 403, 451, 498, 509
bukka, 209
bukkaɣòrra, **209**
bunnánaru, **68**
Bunnánnaru Top., 68
bunnánneru, **68**, 70
butúyu, **379**
cadinancia Sass., **211**
Cannái Top., 438, 439
caraíle, *caràili*, *caraíli*, **319**
Carieke Top., 440
Carzeghe Top., 440
čèrda, **366**
čerβúsku, 196
chirisi, **249**
čiččulía, **242**, 432
čiǧǧinéḍḍu, **254**
čilèstra, **223**
čiliβríu, **239**
činčiðéḍḍu, **254**
činčiréḍḍu, **254**
činčirṛ́ólu, **257**
činči?órru, **209**
cipiri OSrd., 146
čirolía, **242**
čižinéḍḍu, **254**, 406, 422
čížini, **254**, 403, 404, 405, 406, 418, 422
čížiri, 56
čóčči̱ri, **56**, 412, 421
čòkkoro, **56**, 412, 413
Colovrai Top., 438
čònka, **250**
corrocàu, **236**
crúma, **89**
Cucurra PN, 188
čúɣu, **379**
cugúrra, **188**
cugùrra, **188**
cumençat, *cumenz(z)at* OSrd., 296
čúrpu, **371**
čúrra, **351**, 401, 413, 520
currucàu, **236**
ḍḍóḍḍi, 250, **252**, 403, 406, 413
ḍḍúḍḍulu, **73**, 74, 449, 451
ðiliβírke, **203**
dóḍḍi, **252**
dòmo, 29, 406
dòmu, 29
dról'u, 328
dulče, 373
dúle, 413
Durgále Top., 328
dúri, **385**, 403, 413
dzèrda, **366**
dzeúrra, **175**
dzingòrra, **216**
èɣa, **294**
elindzòne, 154
éni, **161**, 403, 486, 492
erikína, 211
erišéru, 59
erisòne, **388**
essíle, **272**
fadzòrβa, **116**
fátto, 393
fáttsu, 393
fáϑϑo, 393
ferréli, 59
ferréri, 59
fiɣáu, 32
fíkatu, 32
framíu, **124**, 126
frámi̱u, **124**
frámma, 124, 126
frámmi̱u, **124**
frèbe, 402
frísa, **66**
fríssa, **66**
fručiḍḍáða, 188
fručiḍḍai, 184
frumíu, **124**
fruskèḍḍa, 98
frúski̱u, 98
frúskru, 98
fuisèssini, **109**
furčíḍḍa, 184
fúrka, 184
Gabutele Top., 356
ǧaɣarare, **258**
ǧáɣaru, **258**, 486, 492, 513
galiléu, 502
ɣallái, **181**
gallíciu, **267**
gallíssi, **267**
gallítsu, **267**
ɣalluččéḍḍu, **260**
galuréu, 502
ǧára, **285**

gaṛánka, **211**
garduléu, 59
garóppu, **330**
ğárra, 285
garróppu, **330**, 501, 519
garuléu, 502
ɣattsédḍu, **260**
Gavói Top., 438
ğèlda, **366**
Gennor Top., 451
Gennos Top., 451
ğèrda, **366**
ggaddenánču Sass., **211**
ggíri, **249**
ghiru, **249**
giḍḍòstre, 38, 132
giḍḍòstre, *giḍḍóstru*, *iḍḍòstro*, *ʔiḍḍòstra*, **158**, 406, 407, 408, 416, 417, 419, 420, 422, 458, 485
giḍḍóstru, 407
ğiğğinéḍḍu, **254**
ğiléstru, **223**
gilindzòne, 154
giliṡòni, **154**
ğinğiréḍḍu, **254**
ɣíri, **249**, 250
ğiṛíli, **308**
gíru, **249**
Gitil Top., 308
giurrítta, **351**, 413
giuru, **100**
gjagar OSrd., **258**, 486, 513
góḍḍu Sass., 379
golléi̯, **301**
golóstiu̯, 32, 38, **130**, 403, 405, 420, 458, 484, 494, 495, 511
golóstiʔe, **130**
golóstiʔe, 442
golóstiʔe, 484, 511
golóstri, 32, **130**, 132, 403, 408, 420, 438, 458, 484, 511
golóstru, 403, 405, 406
ğòlva, *ǵòlva*, *ğòrva*, **116**
gorróppu, **330**, 501, 519
graðanánka, **211**
graɣárğu, 348
gránde, 267
grassíle, **272**
grassíβile, **272**
grássu, 273
greméḍḍikes, 442
gríṡi, **249**, 250, 403, 404, 405, 418, 419, 447
griṡôi, **249**, 447
grisótti, **249**
grisu, **249**
gríṡu, **249**, 250, 404, 405, 419, 422
gròḍḍe, **265**, 406, 407, 413
ɣròḍḍe, **265**
gròḍḍo, **252**, **265**, 406, 407, 413
gróḍḍu, **265**, 407
gróli, **218**, 403
gróri, **218**, 221, 403
gruɣudzòne, 188
gruɣullòni, 188
gruke, 185
grúspinu, **98**
grúspi̯u, **98**
grústi̯u, **348**
grústu, **348**
grúži, 185
guḍḍútulu, **73**, 74
ğuğaréḍḍu, **100**
ğúğaru, **100**
ğúğğuru, *ǵúǵguru*, **100**
ɣuɣúrra, **188**
ǵúǵguru, **100**
ğúɣuru, **100**, 417, 421, 452
Guillèi Top., **301**
Gulai, *Gullai* Top., **301**
ɣullári, **122**
gulléi̯, **301**
Gullèi Top., **301**
gurínu, 115
ğúru, *ǵúru*, **100**, **417**, **452**
gurúle, 502
ğuruyúru Sass., **100**
Gúspini Top., 420
gúspinu, **98**
ğuspínu, **86**
ğúspinu, **86**
ğustítsi̯a, 366
gutturáta, 310
gútturu, 310
hanazzema, 121

hárva, **167**
hárβa, **167**
huhurraóina, **188**
iḍḍòstro, 420
iðíle, **308**
iðíli, **308**
Ilarra Top., 456
iliòne, **150**
implassandullu OSrd., 296
intulgaláre, 328
inturgaláre, 328
iṛí̩, **308**
iṡbírru, 29
Ísili, 420
išòkkoro, **56**
ispaị̀nare, 161
ispéli, **358**, 360, 362, 403, 444
ispéntuma, 306
ispéntumu, 306, 500, 517
issúrpu, **371**
istiòkkoro, **56**
istrugaláða, **328**
istrúgulu, **328**
itsúḍḍas, **374**
jagaru OSrd., **258**, 486, 513
kaðánka, **211**, 422
káḍḍ' e déu, 198
káḍḍ' e pilíkke, **196**, 198, 204, 500
káḍḍu 'e demóniu, 196
káḍḍu 'e déu, 196
káḍḍu 'e déus, 196
kaðèna, 211
kaðenále, 212
kaðenánča, **211**
kaðenánka, **211**
kaðínu, 115
kaðòne, **47**, 49, 447
kaðòni, **47**
kaðranánka, **211**
kaðrenánka, **211**
kaðrúmbulu, **74**
kaðúmbu, 38, **74**, 76, 77
kaðúmbulu, **74**, 77, 428
kadúmbuʀu, **74**
kaðúmbuŕu, 77
kaðúmbuʔu, **74**, 77
kadzáre, 306
kála, 30, **281**, 283, 315, 448, 449
kalaɣáṡu, 216
kalakáṡu, **53**
kalánka, 283, **315**, 422, 448, 518
kalankáṡu, 216
kalažèḍḍa, **223**
kalekáṡu, **53**
kalíttsa, **267**
kallarèḍḍa, **181**
kallédḍu, **260**, 262, 267, 514
kallelédḍu, **260**
kallelléḍḍu, **260**
kalliğğonédḍu, **260**
kallíttsa, **267**
kalluččédḍu, **260**
kallúčču, **260**
kallúttsu, **260**, 514
kalužèrta, **223**, 428, 429, 435, 437, 458
kalužèrtula, **223**, 415
kampínğu, **142**, 144, 428
kánkala, **211**, 212
kánkaru, 212
kánna, 438
kanna de billótti, **96**
kannákka, 212
kannáu, 345
kánnau, 345
kantúnna, **173**
karćúmbulu, **74**
kardánka, **211**
kardu, 77
kardúmbulu, **74**
kardúmmulu, **74**
karduréu, 57, 59, 60, 502
karğilètta, **223**
kaʀišèḍḍa, **223**
karkáre, 112
kárku, 112
karkúri, **112**, 403, 449
karróppu, **330**, 501, 519
karúmbulu, **74**
kárva, **167**, 169, 486, 512
kassí, **272**
kassíli, **272**
kassíβi, **272**
kassíβile, **272**, 415, 428, 445
káṡu, 53
katánka, 315
katèna, 211

katranákka, **211**
kattéḍḍu, 262
kaʔižèḍḍa, **223**
kaβeliˌèra, 94
kaϑϑéḍḍu, 262
kèa, **297**
keia OSrd., 299
kelleisòne, **154**
kérri, **249**, 403, 418
kèya, **297**
kilíši, **249**, 419, 422
kilíšu, **249**, 419, 422
kíri, 457
kiríši, **249**, 403, 404, 405, 418, 422, 457
kiríši e βèrka, **249**
kiríšu, 404, 405
kiríu, **249**, 250, 403, 418, 422, 457
kóḍḍu Gallur., 379
koðòne, **47**
kokkóˌi, 230
kokkóile, 230
kòkkolo, 499, 513
kokkolóḍḍe, 230
kòkkoro, **56**, 412, 499, 513
kokkoróˌi, 230
kokkoròne, 56
kokòrra, **209**, 210, 432
kolonkòne, 422
kolóstru, **130**, 132
kolóvru, 438
kompínğu, **142**, 144
kòro, 406
koròsta, **184**, 185, 414, 415, 423
kòrra, 209, 230
korrokkáu, **236**, 423
korróppu, **330**, 501, 519
korroβákka, **209**, 210
kórru, 209, 432
kòsti, 32, 38, 132, **135**, 403, 442, 485, 492, 494, 511
kòstiɣe, **135**, 442, 485, 511
kóstike, 132, 422
kòstike, 38, 132, **135**, 137, 442, 485, 511
kóstˌiu, **130**
kožuètta, **223**
kraðanákka, **211**
krákkili, 112, 451
krákkiri, 112, 451
krákku, 112
krakkúri, **112**
krás, 185
kríḍḍi, **249**, 403, 404, 405, 418
kriḍḍólu, **249**
kròɣa, 188
kròka, 188, 230
krokkèḍḍa, 230
krukkúri, **112**, 403, 405, 449
krukkúrˌiu, **112**, 403, 405
kruβižồi, **244**
kuɣúrra, **188**, 456
kugúša, **62**, 70, 428, 429, 434, 435, 437
kuɣúša, 63, 409, 411, 498
kuɣúttsula, 64
kuhúrrabóina, **188**
kukkéḍḍu, 175
kukkúi, 175
kukkumèle, 175
kukúrra, **188**, 210, 456
kukúrravóina, **188**
kukúsa, 63
kul(l)ári, **122**
kulinğòni, **244**
kulirğòni, **244**, 447
kúlu, 122, 124
kulurdzòne, **244**, 447
kúnnu, 70
kúrma, **89**, 90, 185
kurrukkáu, **236**, 423
kúruma, **89**, 185
kurústa, **184**, 185, 348, 414, 415, 423, 428
Lacon OSrd. Top., 424
lákkana, **332**, 418, 422
lákkara, **332**
lákku, 334
lákuna, **332**
lána, 190
lána Algh., **190**
lanásta, **190**, 421, 444
lánde, 267
lattúrike, 440, 442
latϑòri, **128**, 130
láu, 140, 142
láu spinóšu, 128, 130
launáži, **140**, 142
láʔana, **332**
leonáğˌi, **140**, 142, 441
leonáržu, **140**, 142, 441
lèpari Sass., **269**

lèpere, **269**
lèpore, **269**
lèporo, **269**, 403, 407, 409, 410, 423
lèpparu Gallur., **269**, 403, 407, 409, 415, 418, 419, 422
lèppere, 92, **269**, 403, 407, 409, 410, 415, 422, 423, 452, 514
lèppere de ardzòla, 269
lepperéḍḍu, **269**
léppi léppi, **269**
lèppiri, 92, **269**
lèppore, 92, **269**, 407, 409, 410, 415, 418, 419, 422, 514
lepporèḍḍa, **269**
lepporéḍḍu, **269**
lèpporo, 415, 418, 419, 452
leppuréḍḍu, **269**
léppuri, 92, **269**, 403, 409, 410, 415, 418, 419, 422, 452, 514
lèppuri, 92, **269**, 407
liðồi, **150**
liðòne, **150**
liğèstra, **223**
limpòra, **92**
limpòrra, **92**, 456
linnálβu, 373
liòni, **150**, 154
lipòrra, **92**, 456
lírčis, **230**, 233, 401, 413, 415, 417, 418, 424, 457
liširiòne, **154**, 421, 447, 452
lišòne, **154**, 447, 452
lispòrra, **92**, 456
líttu, **283**, 487
littu OSrd., 283
lítϑϑu, **283**
lóḍḍe, **265**
lòḍḍe, **265**, 407, 413
Lòḍḍo PN, 265, 406, 407
loḍḍóšu, **265**
Lodé Top., 439
Lodey Top., 439
lodissa OProv., 366
lodzána, **279**
lokáši, **53**
lokášu, **53**
lokáu, **53**
lòkri, 403
lòkri, lókri, **79**, 81
Lollo PN, 265, 407
lonáržu, 140, 142
lóssi̯a, **364**, 366, 368
Lotdé Top., 439
ložána, **279**
Loʔomaʔe Top., 440
lú(*o*), 122
Ludurru Top., 456
ludzána, **279**
luğána, **279**
luğèstra, **223**, 458
luğğèsti, **223**, 458
luiana OSrd., 281
Luianathos OSrd., 281
lukéšu, **53**
lukkíttu, **53**
lukréžu, **53**, 81
lulğètta, **223**
lurdzána, **279**
lúrtsi, **230**, 233, 401, 403, 413, 415, 417, 457
lúrtzi, 415
lúša, **364**
lússa, **364**
lússi̯a, **364**
lússias, **364**
luttsára, **157**, 395, 413
lužắa, **279**
luzana, **279**
lúži, 53, 157
máiða, 161
malangoni Gallur., **190**
mameɣeríši, **249**
mámma de ambíḍḍa, 249
mamma de séssini, **109**
mamma de sizziɣórru, 249
mammuléu, **57**, 59, 60
Mandar OSrd. Top., 424
Mandara Top., 451
Mandas Top., 451
mára, **293**, 294, 317, 516
marayáu, **233**
marayòni, **317**, 336, 447
marètta, **293**
márga, 60

Margangioni Top., 336
marganǧòni, **336**, 447
marǧáni, 265
marrayáu, **233**
marráǧǧu, **312**
marráǧǧu Sass., **312**, 518
marráɣu, **233**
marráldzu, **312**
marrangáu, **233**
marrárdzu, **312**, 518
marum, 60
mát(*t*)*a*, **178**
matiǵèḍḍas, **178**
matižèḍḍas, **178**
mattédu, **178**
mattièḍḍas, **178**
mattiɣèḍḍas, **178**
maϑrikúṩa, 118
meḍḍure, **73**, 403, 421
meláɣe, **190**, 441
melàghe, **190**
melogi, **190**, 441
meluliðòne, **150**
mènta, 132
mentástru, 132
mérula, 382
méurra, 382, 456
míća, **340**
Migil OSrd. Top., 424
míntsa, **340**
mítsa, **340**
míttsa, **340**
míϑϑa, **340**
moðerína, **339**
mòɣora, **288**, 422
mòɣoro, **288**, 289, 407, 422, 452, 487, 516
móɣuru, 407
mòɣuru, **288**, 407, 418, 487, 516
moreðína, **339**, 489, 519
moretína, **339**, 489, 501, 519
Morgongiori Top., 336
Motorra Top., 456
mráɣa, 60
mraganǧòi, **336**
mrúɣu, 60
mrúva, **275**
mruvòi, **275**
mruvòne, **275**
mruvòni, **275**
múćću Cors., Gallur., 146
múčču Sass., 146
muddéɣu, **144**
muḍḍuru, 73, 74, 403
muðéčču, **144**
muðéǧǧu, **144**
mudeglu OCamp., **144**
muðéɣu, **144**
muðerína, **339**
muðeríɲa, **339**
muðéyu, **144**
muðréɣu, **144**
muðréhu, **144**
muðréku, **144**
muðréχu, **144**
múfra, **275**
mufròne, **275**, 413, 416, 417, 500, 515
múɣra, **275**
muɣròne, **275**, 413, 416, 447, 500, 515
muɣuòes, **57**, 60
mukkaðòre, 216
mummuléu, **57**, 59, 60
munkaðòre, 216
múra, 89
muráke, **334**
muravèra, **89**
murðéǧǧu, **144**
murdéɣru, **144**
murdéɣu, **144**
murdéyu, **144**
mureðína, **339**, 340
muréɣu, **144**
muretína, **339**
murga, **57**, 60
múrga, 60, **275**
murgòne, **275**
múrgu, 60
Murgueliai Top., 57
murguéu, **57**
murguèu, **57**
murguèus, **57**, 60
murguléu, 60, 502
murgulèu, **57**
muriɣína, **339**
murmuréu, **57**, 59, 60
murồi, **275**
muròne, **275**
murtélʔu, **144**
múru, 339
múrva, **275**, 447
murvồi, **275**

murvòne, **275**
murvòni, **275**, 447
muteclu OLog., **144**, 146
muteclu OSrd., 496, 499, 511
mutréku, **144**
múvra, **275**
muvròne, **275**
muvròni, **275**
muvròni Gallur., **275**
náiriši, 502
nárvas, 502
nástala, **190**, 421, 444
nástula, **190**
návras, 502
naybbúttsa Sass., 116
náβras, 502
Neoneli Top., 356
néula, 140, 142, 441
neuláɣe, **140**
neulaǧǧi, **140**
neuláǧi, **140**
neuláke, **140**, 441
neuláχe, **140**
nódu, 103
nóðu, 115
nòmene, 139
Nora Top., 441
noroḍḍáśile, **102**, 105, 409, 411
nóu, 115
ntsaúrru, **175**
nuráɣe, **334**, 419, 422, 441
nuráke, 279, **334**, 418, 419, 440, 441
nuráži, **334**, 441
Nure Top., 441
nuríle, **115**, 445
núrra, **277**, 279, 515
odzástru, 132
oĩõi, **150**
oláśi, **102**
oliðòne, **150**, 152, 154, 447
oliòne, **150**
oliòni, 140, 142, **150**
Ollala Top., 439
ollásu, 103
olláśu, **102**, 105, **106**, **109**, 409, 411
olliòni, **150**
Ollolái Top., 438, 439
olokáśu, **53**
olósti, **130**, 403, 405, 408, 458, 484, 511
olóstiu, **130**
olóstri, **130**
olóstriɣe, **130**, 442
olóstru, **130**
oppínu, 142
òrga, 422
orgòsa, 63, **290**, 434
orgòše, **290**
Orgòsolo Top., 290
oʀiòni, **150**
òro, 406
oroḍḍáśile, **102**, 409, 411, 445, 446
oroḍḍáśu, **102**, 103, 411, 445, 446
orròa, **324**
orrolári, **122**, 403, 405
orroláriu, **122**, 403, 405
orròsa ɣullári, **122**, 124
orrú, 122
orruláriu, **122**
Ortoiʔe Top., 440
Orθái Top., 438
Orθullè Top., 395
óspile, 322, 445, 489, 494
óspinu, **98**
oβiòi, **150**
oβiòi, **150**
padzáre, 306
pasquiri OSrd., 296
pattáda, **306**
pattáða, **306**
pattháta, **306**
paβáða, **351**
paβáði, **351**
paβári, **351**
pèðra, 373
péntuma, **306**, 500, 517
pèrda, 373
Piaghe Top., 440
piðai, 115
pilíkke, 196, 429
pilindzòne, 154
piliśòne, **154**
pillónku, 449
píllu, 449
pínu, 142
píra, 132
pirai, 115
pirástru, 132
piśiyò, **209**, 210, 211
píśiri, 246
pistiḍḍòre, 451

pistiḍḍúri, 451
pittsiɣòrra, **209**, 210
píttsu, 210
piβiánkulu, 448
piβintsíri, 200, 396
piβirittsíri, 200
piβisíu, 200
piβissíu, 200
piβittséri, 200
piβittsíri, 200, 246, 403, 404, 421
piβittsíu, 200
piβíu, 448
piβiɣɣía, 200, 421
piβiɣɣíu, 200, 396, 403, 404
piɣɣiɣórru, **209**, 210, 432
píɣɣu, 210
Plavaki Top., 440
Ploauake Top., 440
pòrka, 203
pottaminkòrras, 209, 211
prèða, 373, 440
prèða e láʔna, **332**
preðáʔe, 440
préðike, 440
príncˇiri, **246**, 401, 403, 404, 418, 422, 452, 457
prinčóttu, **246**, 422, 452
púttsu, 393
púttu, 393
púɣɣu, 393
r'ákana Sass., **332**
ʀépuį, **269**
ròya, **324**
rú, 122
rúɣe, 185
rúke, 185
rullári, **122**
rúɬta, 184
runáɣe, **334**
runáke, **334**
rúša, **364**
rústa, **184**, 185, 348, 428
rustìa, **348**
rustólu, **348**
rúβu, 122
sakkáyu, 326
salaúspe, **165**
salaúspu, **165**, 167
sámbene, 139
sampunáre, 142
sapunare, 142
saraúipu Sass., **165**, 167
sárpa, **161**, 499
Sarunele Top., 356
sasaǧa, **192**
saṡáǧa, **192**, 194
saṡáya, **192**, 194, 414
sassalúɣa, **228**
ṡassalúɣa, **228**
satsaúɣa, **228**
sattsaɣúʀa, **228**
sattsalúɣa, **228**, 415, 416, 423
sattsamurréḍḍu, **257**
sattsaróį, **51**, 395, 439
sattsaúɣa, **228**
sattsemurréḍḍu, **257**
sazzaroi, **51**
sèðere, 304
seiṡèi, **192**, 194, 414, 415, 439
Semèstene Top., 420
senèɣra, **204**
sèrga, **356**
sèro, 406
sèrra, 356
seṡáya, **192**, 194
sèṡè, **192**, 194, 414
sèssene, **109**, 417, 421
sessinárǧu, **109**
sèssini, **109**, 417, 421
sèssiri, **109**
séthile, **304**
séttįa, **304**, 445
settįáre, **304**
séttiðu, 304
séttile, **304**, 445
séttįu, **304**, 445
sèttsere, 210
sèttsiri, 210, 432
silibrich volador Algh., 203
sintsiɣórru, **209**
sintsilúɣa, **227**, 415, 416, 423
sintsimurréḍḍu, **257**
síntsiri, **112**, 395, 397, 417, 418, 452, 456
sintsurru, **112**, 395, 397, 417, 418, 452, 456
síppiri, **146**
siṡáḍḍza, **192**
siṡáǧa, **192**
sisaia, **192**

siśáya, **192**, 194, 414, 415, 439
sissiriɣòša, **209**
síti, 402
sitseɣórru, **209**
sittsiɣárru, **209**
sittsiɣòlla, **209**
sittsiɣórru, **209**, 210, 230, 396, 432
sittsimurréḍḍu, **257**
sittsiringòni, **198**
sittsulía, **242**, 396, 413, 432, 434
siziɣórru, 209
síβile, **272**, 428
skaβéčču, 94
sóčči̱ri, **56**, 96, 406, 407, 417, 419, 421, 423, 451
soli̱onážu, **140**, 142
sòlla, **385**
soloíɣa, **206**, 214, 409, 414, 442
solóiɣa, **206**, 409, 414, 442
solóriɣa, **206**
sónni̱a, **106**
Sòrgono Top., 420
sórgu, 60
soṡòɣa, **228**
sossòɣa, **228**
sottsoróyu, **51**
šòva, **137**, 139
spadderi, **239**
spai̱nai, 161
splíɣu, 373
spraḍḍéri, **239**, 422
spraβéri, 239
srápa, **161**
ṡráppa, **161**
súḍḍa, **374**
súɣu, **379**
suíɣa, **206**, 414
sullungòne, **198**
sulurtsi, **230**, 401
suntsòrra, **209**, 210
tabakòrra, **209**
tákka, **301**
tákku, **301**
taláu, **345**
tártsa, **233**, 401
táṡaru, **164**, 165, 452
táṡiru, **164**
tássu, 164, 165
táṡua, **164**, 165
táṡulu, **164**, 165, 452
táṡuru, **164**, 165, 452
tattaróǧu, **51**
tattaroiu, **51**
tattaróyu, **51**, 421, 423, 429, 439, 455, 509
taxus, 165
tectinosu OLog., 148
teláu, **345**
tèle, **356**, 413, 414, 418, 444
Tèltoro Top., 420
tèrga, **362**
tería, **118**, 120, 121
tèrpe, 394
tesòro, 406
Teti Top., 150
teti̱òne, **148**
Tettène Top., 150
tétti, *téti*, **148**, 403, 405
Tettinosa Top., 150
téttiu, *tétiu*, **148**, 403, 405
tèttu Gallur., **148**, 403, 405
tevèle, **356**, 413, 414, 422, 444
težíle, **328**, 445
Theppage OSrd. Top., 452
Theppar OSrd. Top., 418
thinnigas OLog., 106
thurpu OSrd., 371
tía, 429
tídza, **385**
tilíbba, 394, 438
tilíbbu, **221**, 396, 429, 435
tilibílke, **203**
tilibricu, **239**
tilibriki, **203**
tilibrikku, **203**
tilibríkku, **239**, 409, 411, 422, 429, 435
tilibriku, **203**
tilibrirke, **203**
tilibríu, **239**, 409, 411, 422
tiliɣèlta, **223**
tiliɣèrta, **223**
tiliɣèťťa, **223**
tiliɣúɣu, **227**
tiligulu, 228
tilígulu, 228
tilikètta, **223**

tilínğa, **198**, 228, 447
tilingòne, **198**
tilinğòne, **198**, 200
tilingròne, **198**
tilíngu, 447
tilipírke, **203**
tilipische, **203**
tilipriku, **203**
tiliβíχχe, **203**
timándza, 121
timònğa, 121
tinnía, **106**
tintiồi, **148**
tintiòni, **148**
tintiɾi̯óluβèḍḍe, **257**
tintirɾi̯ólu, **257**
típpiri Top., 148
tir'ibíkku Sass., 239
tiría, **118**
tiribíkku Sass., **203**, **239**
tiriɣètta Sass., **223**
tirigúru Sass., **227**
tiringòni, **198**
tirolía, **242**, 409, 413, 415, 429, 432, 434
tirɾi̯ólu, **257**
tirɾi̯ólu 'e βèḍḍe, **257**
tirɾi̯ólumalaβèḍḍe, **257**
tirɾi̯óluβèḍḍe, **257**
titímbalu, 76
titirɾi̯ól' impeḍḍáðu, **257**
tittɨ̯òne, **148**, 405, 447
tíu, 429
tiváni, **244**
tiϑiɣórru, **209**, 210, 396, 432
tòa, **137**
tòa Algh., **137**
tòdza, **385**
tón(n)eri, **319**
tònka, **250**
tónneri, 403
Tónneri, 452
tòrko, 310, 312, 360, **362**, 407, 408, 418, 419, 420
tórku, 420
Tortolì Top., 420
tottoróyu, **51**
travòḍḍa, 77
trivòḍḍa, 77
troccale, 328
tròḍḍa, 77, 79
troḍḍi̯ánčulu, 448
troḍḍi̯ánkulu, 448
troḍḍiáre, 448
trógliu, 328
trókko, 360
tròkko, 310, 312, 362, 407, 418, 419
trókku, 310, 312, **362**, 407, 418, 419, 488, 517
tról'u, 328
trovòḍḍa, 77, 79
trúbulu, 328
trúḍḍa, 382
truèḍḍa Gallur., 77, 79
trúɣu, **379**
trullío, **382**, 407, 408
truváre, 79
truvèḍḍa, 77, 79
truvísku, 77, 79
truβísku, 79
tsangulíttas, 434
tsántsa, **81**
tsantsallúɣa, **228**
tsára, 157
tsarpeḍḍéri, **239**, 422
tsarríɣa, **214**, 442
tsattsalúɣa, **228**
tsaβaɣúɣu, **227**, 228, 415, 416
tséɣu, 371
tseliɣúsa, **62**, 63, 395, 409, 411, 428, 429, 434, 435, 437
tsèppara, 322, 418, 452
tsèrga, 356, **362**, 396, 400
tserpeḍḍéri, **239**, 396, 422
tsèrpi, 239
tserríɣa, **214**
tserríɣu/a, **214**, 414, 442
tserringòni, **198**
tséurra, **175**, 395, 397
tsía, 429
tsiaʀiɣúɣu, **227**
tsiaʀingồi, **198**
tsíbba, **94**
tsiḍḍí, 250
tsikkiría, **64**, 395, 397
tsilibríłke, **203**
tsilibrínkidi, **203**
tsiliɣèrta, **223**
tsiliɣúɣu, **227**, 415, 416

tsilikúkku, 227
tsilingòne, **198**
tsilíngu, **198**, 395
tsiliprí?e, **203**
tsiliβríkke, **203**
tsimpilíɣe, **196**, 198, 203, 204, 395, 421, 429, 500
tsingòrra, **216**, 396, 397, 456
tsinníɣa, **106**
tsintsiɣórru, **209**
tsintsikòrra, **209**
tsintsimurréḍḍu, **257**
tsintsimúrru, **257**, 432, 434
tsintsòrra, **209**, 210, 432
tsintsuréḍḍu, **257**
tsioḍḍía, **242**, 409, 415
tsíppiri, **146**, 395, 397, 403, 417, 452
tsirifrúči, **184**, 395, 429, 435
tsirikèlta Gallur., **223**
tsirikkía, **64**
tsirikúkku, **227**
tsiringǒi, **198**
tsiringòne, **198**
tsirinòni, **198**
tsiriñòni Gallur., **198**
tsirríɣu, **214**, 414
tsirringǒi, **198**
tsirringòni, **198**
tsirṛióluβèḍḍe, **257**
tsirṛioluβèḍḍe, **257**
tsirṛióra, **257**
tsirṛiòra Sass., **257**
tsitárra, **209**, 210
tsitsirṛiólu, **257**
tsittsiɣórru, **209**
tsíu, 121, 429
tsòa, **137**
tsònka, 175, **250**, 252
tsònki, **175**
tsónkine, **175**
tsònkine, **175**, 395
tsónni, **106**, 108, 395, 397, 403, 405, 414
tsónnṛa, **106**, 403, 405
tsontsikòrra, **209**, 210, 211
tsontsòrra, **209**, 210, 432
tsorómpis, **227**, 416
tsorroíɣu, **214**, 414, 442
tsottsorói, **51**
tsottsoróyu, **51**
tsrappeḍḍéri, **239**
tsrèɣa, 356, **362**
tsreppeḍḍéri, **239**
tsrúɣu, **379**
tsrúppu, **371**
tsuaḍḍía, **242**, 415
tsúḍḍa, **374**, 396
tsúḍḍa Gallur., **374**
tsúḍḍi Gallur., Sass., **374**
tsúḍḍu Sass., **374**
tsueḍḍía, **242**
tsúɣu, **379**, 396
tsuiḍḍía, **242**
tsúkkru, **379**
tsuntsuréḍḍu, **257**
tsuntsuríttu, **257**
tsuntsurréḍḍu, **257**
tsuntsurríttu, **257**
tsúrpu, **371**, 396
tsúrra, **221**, **351**, 401, 413, 520
tsurríttu, **221**
tsurrundéḍḍu, **258**
tsurulía, **242**, 409
tsutsurríttu, **257**
túḍḍa, **374**, 396
túfera, **126**
túǵu, **379**
túǧu, **379**
túm(b)u, 118
túmbara, **74**, 76, 428, 451
túmbaru, **74**, 76, 77
túnna, **173**
tunníu, **173**, 421
tuñníu, **173**
túnnṛu, **173**
tuntunníu, **173**
tuntúnnu, **173**, 421
túora, **126**, 452
turbare, 79
túrgalu, 328
turgusòne, 447
turguṡòne, **62**, 63, 411
túrpu, **371**, 396
túrra, 382
turrío, **382**, 407, 408
turulía, **242**
tutturreri, **257**
tutturròne, **376**, 490
tuttúrru, **376**, 422, 490

tutúmbaru, **74**, 76
tuturréri, **257**
tutúrru, 418, 456
túv(v)era, **126**
túvara, **126**, 128, 438, 452
Tuvarái, 438
Tuvarái Top., 439
tuvèle, 356, 418
túvera, 412, 414, 452
túvu, 128
túyu, **379**
tzándhara, 81
tzàndhara, **81**, 421, 451
tϑáṇḍa, **81**, 395, 429, 451, 484, 510
tϑutϑuréḍḍu, **257**
tϑutϑurréri, **257**
uḍḍuréḍḍu, **73**
uḍḍúri, **73**
uḍḍúriɣu, **73**, 442
úḍḍuru, **73**, 421, 451
urgúsa, 63, 434
urɣúsa, **62**
úrma, 89
utúyu, **379**
úvara, 412, 414
úvvara, **126**, 128
vákka, 211
vassíli Gallur., **271**
vega OSrd., 35
vervèke, 73
vinnínna, 73
vrattakáṡu, **53**
vríssa, 66
vrústu, **348**
vuḍḍúri, **73**
yánna, 29
yènna, 29
yuruyúru Sass., 100
zalki, 345
zaraclua, 228
žèðra, **366**
ziliblich Algh., **203**
zilibricu Gallur., **203**
zilibrínkiḍḍi, **203**
zingòrra, **216**
zinnía, **106**
zirifrúci, **184**
zirringòni, **198**
żissiɣórru, **209**
zorroiga, **214**
zorrómpi, **227**
zúḍḍa, **374**
zúɣu, **379**
ʔála, **281**
ʔassíβile, **272**
ʔèppiri, **269**
ʔiḍḍòsta, **158**
ʔiḍḍòstra, 407
ʔolóstiʔe, **130**, 442
ʔúrma, 89
ʔurústa, **184**
ϑáiri, **236**, 417, 479, 486
ϑakkulíttos, 434
ϑalakèrta, 416
ϑalakúkka, **227**, 432, 434
ϑalakúkku, **227**, 423
ϑalakúʔa, **227**
ϑálau, **345**, 396
ϑalaʔèrta, **223**, 228, 415, 423
ϑánda, 421
ϑáṇḍa, **81**
ϑaϑϑalakúkka, **227**
ϑaϑϑalúkka, **228**, 396, 432, 434
ϑelakúkka, **227**
ϑelakúkku, 203
ϑelapòrka, **203**, 416, 421
ϑèrga, **362**
ϑería, **118**
ϑèrpes, 394
ϑeulíttos, 434
ϑía, 429
ϑídza, **385**, 396, 415
ϑilibríʔe, **203**
ϑilikèrta, 204, **223**, 228, 230, 396, 415, 416, 423, 428, 429, 435, 437, 438
ϑilikètta, **223**
ϑilikúkku, **227**, 396, 423, 429, 432, 434, 435
ϑilikúku, 204
ϑilingòne, 447
ϑilingròne, **198**, 395, 423, 429, 435, 447
ϑilipírke, **203**, 396, 400, 416, 421, 429, 435
ϑilipríʔe, **203**
ϑiliʔèrta, **223**
ϑinnía, **106**
ϑinníɣa, **106**, 395, 414
ϑiòkkoro, **56**, 395, 406, 407, 412, 413, 417, 419, 421, 423, 451

ϑíppiri, **146**, 395, 397, 403, 404, 417, 418
ϑiría, **118**, 120, 121
ϑirriáre, 394
ϑíu, 121, 394, 429
ϑiϑímbalu, 76
ϑiϑirḭólu, 434
ϑiϑirrḭólu, **257**, 396, 429
ϑiϑϑilíkkos, 434, 438
ϑiϑϑirrḭólu, 423
ϑòa, **137**, 139, 395, 400
ϑòdza, **385**, 396, 415
ϑòɣa, **137**, 139, 395
ϑòğğa, **385**
ϑòğğe, **250**, 252, 413
ϑònka, **250**, 413, 416
ϑóppu, 371
ϑorroíɣu, **214**, 396
ϑòrža, **250**, 252, 413
ϑòβa, **137**, 139, 395
ϑoϑoróyu, **51**, 395, 421, 429
ϑoϑϑórḭu, **51**, 509
ϑoϑϑoróyu, 423, 439
ϑráṇḍa, **81**
ϑráppa, **161**
ϑrúɣu, **379**
ϑrugúsa, **62**
ϑrugúṡa, 434
ϑrúk(*k*)*u*, **379**
ϑrúχu, **379**
ϑúčču, **379**
ϑúḍḍa, **374**, 396
ϑúɣru, **379**
ϑúɣulu, **379**, 396, 429
ϑúkru, **379**, 396, 400, 429
ϑulungròne, **198**, 200, 423
ϑulunkòne, **198**
ϑúlʔu, **379**
ϑúrgalu, **328**, 396, 444
ϑurgúsa, **62**, 70, 395, 434, 447
ϑurgúṡa, 63, 409, 411, 428, 429, 434, 435, 498
ϑúrpu, **371**, 373, 374, 396
ϑurulía, **242**, 396, 415, 429, 432, 434
ϑurunkròne, **198**
ϑuϑuréḍḍu, 254, **257**, 396, 423, 429, 434
ϑuϑuréri, **257**
ϑuϑurréri, **257**
χattséḍḍu, **260**

Semitic

cusmin Pun., 98
faras aš-šayṭān فرس الشيطان Ar., 196
ḥhamra Tun., 90
ḫurmu Akk., 89
kharma Tun., 89
məṭṭa Moroccan Ar., 178, 479
mita מִטָּה Hebr., 178
mysehi Pun., 398
šabb شَبّ Ar., 94
šēkār שֵׁכָר Hebr., 64, 468
sensur Pun., 112, 397
senzur Pun., 454
singūrā Aramaic, 216
singurru Akk., 216, 397
yáʿar יער Hebr., 285, 470, 473
zarʾu Akk., 175
zenzur, 454
zenzur Pun., 112, 397, 398, 418, 456
zéraʿ זֶרַע Hebr., 175, 397, 470
zēru Akk., 175, 470
zibbir Pun., 146, 397, 398, 399, 404, 418, 454
ἁρμαλά Syr., 89, 469, 474
ζερα Pun., 175, 397, 398
σικκιριά Pun., 397, 398, 399
σικκιριά Pun., 64
σιχκιριάμ Pun., 64
χουρμά Pun., 89, 90

Other

Almere Du. Top., 33
amë Alb., 293
barr MW, OIr., 389
cergă Rom., 362
cergë Alb., 362
çergë Alb., 362
çerge Turk., 362
charj Bret., 364
colwyn Welsh, 262, 514
cuilenn OIr., 132, 484, 511
derbita Gaul., 83
dervoed Bret., 83
dzaghli ძაღლი Georg., 486, 513
gamule Alb., 289
gamulë Alb., 289
hulis OHG, 132, 484, 511
hulst Du., 484, 511
kaḍamba Skt., 76
kalẽ Lith., 262, 514
kelenn Bret., 132, 484, 511
këlúsh Alb., 514
këlýsh Alb., 262, 514
kolen Bret., 262, 514
magulë Alb., 289, 488, 516
măgură Rom., 289, 488, 516
maklo Georg., 276
melkdistel Du., 76
Reinhard Gm. Pers., 265
sargie Flemish, 364
Schassen Westfalian, 364
sea holly Eng., 120
serch Bret., 364
wolfsmelk Du., 76
zağar Turkish, 486, 513
λέποριν Siculian, 271, 410, 514
ναντί Egyptian, 83, 484, 510

Nederlandse samenvatting

Dit proefschrift heeft als doel een beter inzicht te krijgen in de achtergrond en aard van de talen die in Sardinië gesproken werden ten tijde van de Romeinse verovering van het eiland. Vandaag de dag wordt in Sardinië het Sardijns gesproken, dat van het Latijn afstamt. Door nauwkeurig onderzoek naar de Sardijnse woordenschat is het echter mogelijk om sporen te vinden van substraattalen, d.w.z. de verdwenen voor-Romeinse talen waar het Latijn in Sardinië mee in aanraking kwam. Het proefschrift bestaat uit vijf delen.

In het eerste deel worden de achtergrond van het onderzoek (hoofdstuk 1) en de werkwijze (hoofdstuk 2) uiteengezet. In hoofdstuk 1 wordt de omvangrijke onderzoeksgeschiedenis van het vraagstuk besproken. Vele woorden en eigenschappen van het Sardijns zijn op enig ogenblik toegeschreven aan een of meerdere substraattalen, en hiervoor zijn ook verschillende verwantschappen voorgesteld. Deze voorstellen worden later in het proefschrift aan de hand van de in hoofdstuk 2 uitgewerkte werkwijze getoetst. Deze werkwijze behelst het herkennen van Sardijnse woorden met een substraatoorsprong, het analyseren van de kenmerken van deze woorden, en het reconstrueren van de taalkundige voorgeschiedenis die aan het bewijs ten grondslag ligt.

Het tweede deel bevat de etymologische analyse van 164 Sardijnse woorden waarvoor een substraatoorsprong voorgesteld is, of waarvan een Latijnse oorsprong onzeker is. De woorden zijn onderverdeeld aan de hand van hun betekenis in planten (hoofdstuk 3), dieren (hoofdstuk 4), geografische zaken (hoofdstuk 5), landbouw (hoofdstuk 6) en overige woorden (hoofdstuk 7). Elk besproken woord is voorzien van een kaart met daarop de verspreiding van de verschillende varianten van dat woord binnen Sardinië. Per woord wordt afgewogen of een voor-Romeinse oorsprong aannemelijk is en, zo ja, of er mogelijke verwanten buiten Sardinië bekend zijn.

In het derde deel wordt onderzocht of de Sardijnse woorden waarvan een substraatoorsprong aannemelijk is inzicht kunnen bieden in de eigenschappen van hun voor-Romeinse brontaal. In hoofdstuk 8 wordt dit voor klankkenmerken gedaan. Er wordt vastgesteld dat het voor-Romeins Sardijnse klanksysteem niet opvallend veel afweek van dat van het Latijn, en dat voor een aantal eerder voorgestelde voor-Romeinse kenmerken onvoldoende bewijs is. In hoofdstuk 9 wordt het bewijs voor voor-Romeinse morfemen onder de loep genomen. In het

verleden zijn vijf voorvoegsels en twaalf achtervoegsels aan een voor-Romeinse taal toegeschreven. Hiervan kunnen twee voorvoegsels en vijf achtervoegsels inderdaad bevestigd worden. Voor de voorvoegsels **kV-* en **ϑi-* en voor het achtervoegsel **-Vr* worden mogelijke grammaticale functies besproken. Er zijn onvoldoende gegevens beschikbaar om iets te kunnen zeggen over de functie van de andere toevoegsels.

Het vierde deel betreft de verwantschap van de voor-Romeinse taal of talen van Sardinië. De gangbaarste voorstellen voor geattesteerde verwanten van een Sardijnse substraattaal, namelijk het Punisch, Berber, Baskisch en Etruskisch, worden een voor een besproken in hoofdstuk 10. Van de vier kandidaten is het bewijs voor een Punisch substraat het sterkst. Dit is begrijpelijk in het licht van de geschiedkundig bekende Fenicische en Carthaagse aanwezigheid in Sardinië. Het Punisch werd naast inheemse voor-Romeinse Sardijnse talen gesproken, tot lang na de Romeinse verovering. Het kleine aantal Sardijnse woorden met een mogelijk Berbers cognaat moet waarschijnlijk in een Carthaagse context geplaatst worden. Wat het Baskisch betreft, kunnen slechts vier van de negentien voorgestelde woordvergelijkingen gehandhaafd worden. Dit is een te gering aantal om de inheemse taal van voor-Romeins Sardinië als verwant van het Baskisch te bestempelen. Aannemelijker is dat het Baskisch onafhankelijk in contact stond met een taal die nauw verwant was aan een taal in voor-Romeins Sardinië. Voor slechts één Sardijns woord kan een Etruskische cognaat worden ontwaard, hetgeen gezien onze magere kennis van de woordenschat van deze taal niet verrassend is. Desondanks zijn er mogelijk gedeelde structurele kenmerken tussen het Etruskisch en het voor-Romeins Sardijns. In hoofdstuk 11 wordt het bewijs behandeld voor verwantschappen tussen het voor-Romeins Sardijns en andere substraattalen zonder schriftgetuigenissen buiten Sardinië. Hieruit wordt afgeleid dat er voor verscheidene Sardijnse woorden van substraatoorsprong overeenkomsten te vinden zijn in het westelijke Middellandse Zeegebied, met een zwaartepunt in hedendaags noordoost-Spanje en Zuid-Frankrijk. Dit duidt op het bestaan van een voor-Romeinse taal of taalfamilie in dit gebied. Dit komt ten dele overeen met eerdere voorstellen tot een "Mediterraan substraat", zij het in een geografisch beperktere verspreiding.

Tot slot bevat deel vijf (hoofdstuk 12) een uitgebreidere samenvatting van de bevindingen van het proefschrift en enkele vooruitblikken op mogelijk vervolgonderzoek.

Arresùmini sardu

Su scopu de custa dissertatzioni est de cumprendi mellus su contestu e s'ìndula de is lìnguas chi ddas fueddànt in Sardìnnia in sa simana de sa cunchista arromana de s'ìsula. Oindii in Sardìnnia si-nci fueddat sa lìngua sarda, chi calat dae su latinu. Intamus, po mesu de unu scrucullu atentu de su bocabolariu sardu fait a agatai arrastus de lìnguas de sostratu, est a nai, de is lìnguas preromanas sparèssidas in Sardinnia cun is chi su latinu at ténniu cuntatu. Sa dissertatzioni est partzida in cincu partis.[325]

Sa primu parti esponit su contestu (capìtulu 1) e su métodu (capìtulu 2) de sa circa. Su capìtulu 1 ndi tratat s'amprosa stòria de sa circa a custa chistioni. Medas fueddus e naturalesas de su sardu ddus ant atribuius a una o prus lìnguas de sostratu, e po issas lìnguas ddis ant propostu nàscidas diferentis. Funt a avalorai custas propostas prus ainnantis in sa dissertatzioni, imperendi su métodu cumpostu in su capìtulu 2. Custu métodu incluit s'identificatzioni de fueddus sardus chi calant dae una lìngua de sostratu, s'anàlisi de is naturalesas de custus fueddus e s'arricostrutzioni de sa preistòria linguìstica chi at nàsciu a is fatus linguìsticus.

Sa de duas partis achipit s'analisi etimològica de 164 fueddus sardus po is chi nc'ant propostu una nàscida de sostratu o po is chi una nàscida latina no est segura. Is fueddus ddus apu cumpartziu a basi de su sentidu insoru in prantas (capìtulu 3), animalis (capìtulu 4), nòminis geogràficus (capìtulu 5), bocabolàriu de massarìtzia (capìtulu 6) e àterus fueddus (capìtulu 7). Cun dònnia fueddu s'agatat una mapa chi amostat sa spartzidura de is variantis diferentis suus in Sardìnnia. Si-ndi avalorat po dònnia fueddu chi fait una nàscida ipotètica preromana e, si fait, chi nci funt fueddus foras de sa Sardìnnia chi nc'iant a podi essi arrelataus.

In sa de tres partis spriculaus si is fueddus sardus, po is chi fait una nàscida de sostratu, si-ndi podint donai un'arviu de is naturalesas de sa lìngua preromana dae sa chi nci funt caladas. Su capìtulu 8 si cuncentrat a is sonus de sa lìngua: si-nci cuncruit ca su sistema fónicu de su sardu preromanu no scistàt meda de su

[325] Cust'arresùmini est scritu in sardu campidanesu, chi est sa bariedadi de sa lìngua sarda chi connoscu mellus. Po s'ortografia, apu sighiu is *Arrègulas po ortografia, fonètica, morfologia e fueddàriu de sa Norma Campidanesa de sa Lìngua Sarda* (2009), de su Comitau Scientìficu po sa Norma Campidanesa de su Sardu Standard.

sistema latinu, e ca no nci funt proas bastantis po sa presèntzia de unas cantu naturalesas preromanas chi ddas ant propostu innantis. Su capìtulu 9 esaminat is proas po is morfemas preromanus supostus. Innantis, cincu prefissus e doxi suffissus nci-ddus ant ascritus a una lìngua preromana. De issus, ndi podeus confirmai duus prefissus e cincu suffissus comenti talis. Fintzas si-ndi chistionant is funtzionis grammaticalis de is prefissus **kV-* e **ϑi-* e de su suffissu **-Vr*. No nci funt datus bastantis po nd'apubai is funtzionis de is aterus affissus.

Sa de cuàturu partis castiat su parentiu de sa lìngua o de is lìnguas preromanas de sa Sardinnia. In su capìtulu 10 si-nd'arrexonat de is propostas de su prus importu mannu de calis lìnguas atestadas iant a podi essi arrelatadas cun su sostratu sardu, is chi incluint su pùnicu, su bèrberu, su bascu e s'etruscu. De issus, sa proa de unu sostratu pùnicu parit essi sa prus crara, fintzas ca sa presèntzia fenìcia e cartaginesa in Sardinnia est connota de s'antigóriu; su pùnicu ddu chistionànt a costau de is lìnguas preromanas sardas de su logu fintzas a meda tempus apustis de sa cunchista arromana. Po su nùmeru bàsciu de is fueddus sardus po is chi fait chi siant arrelataus a fueddus bèrberus, est capassu chi tochit a ddus poni in d'unu contestu cartaginesu. Po su chi pertocat su bascu, fait a aguantai sceti cuàturu de is 19 acaramentus lessicalis propostus. Custu est tropu pagu po chi no potzaus classificai sa lìngua originali de Sardìnnia preromana comenti unu parenti de su bascu. Est prus capassu chi su bascu fessit in contatu a solus cun d'una lìngua chi a borta sua fut arrelatada a strintu cun d'una lìngua preromana de sa Sardìnnia. Fait a agatai sceti una corrispondèntzia etrusca, su chi no si depit essi unu spantu casi ca de su bocabolàriu di custa lìngua ndi scieus pagu meda. Intamus, nci funt calincunas naturalesas struturalis chi forsis s'etruscu e su sardu preromanu condividiant. Su capìtulu 11 analisat is proas de acàpius intra de su sardu preromanu e de àteras lìnguas de sostratu no atestadas, foras de Sardìnnia. Si notat ca po calincunus fueddus sardus nàscius de sostratu ddis podeus agatai fueddus arrelataus in àterus logus in su Mediterràneu ocidentali, cun d'unu centru de graidadi in sa Spànnia orientali de susu e in sa Frantza de giossu. Istu inditat ca in s'antigóriu nci fut una lìngua o una famìlia de lìnguas preromanas in cust'arregioni. Issu apicigat giai cun is propostas chi ddas ant fatas prima, de unu "sostratu mediterràneu", mancai lacanau prus a strintu.

Sa de cincu partis (capìtulu 12), a ùrtimu, cuntenit un'arresùmini prus cumpriu de is arresurtaus de sa dissertatzioni e calincunas prospettivas de circa chi si iat a podi fai in su benidori.

Riassunto italiano

L'obiettivo di questa dissertazione è comprendere meglio il contesto e la natura delle lingue parlate in Sardegna all'epoca della conquista romana dell'isola. Oggigiorno vi si parla il sardo, che discende dal latino. Tuttavia, tramite un attento esame del vocabolario sardo è possibile trovare tracce di lingue di sostrato, cioè delle lingue preromane scomparse in contatto con il latino di Sardegna. La dissertazione è suddivisa in cinque parti.

La prima parte espone il quadro generale (capitolo 1) e la metodologia (Capitolo 2) dell'indagine. Il capitolo 1 presenta l'ampia storia della ricerca su questo tema. Molte parole e caratteristiche del sardo sono state attribuite a una o più lingue di sostrato, per le quali sono state proposte anche varie affiliazioni. Le proposte vengono valutate più avanti nella tesi, utilizzando la metodologia elaborata nel capitolo 2. Questa metodologia comprende l'identificazione di parole sarde con un'origine di sostrato, l'analisi delle caratteristiche di queste parole e la ricostruzione della preistoria linguistica che sta all'origine dei dati linguistici.

La seconda parte contiene l'analisi etimologica di 164 voci sarde per le quali è stata proposta un'origine di sostrato o per le quali è incerta l'origine latina. Le voci sono suddivise sulla base del loro significato in piante (capitolo 3), animali (capitolo 4), termini geografici (capitolo 5), lessico agricolo (capitolo 6) e restanti voci (capitolo 7). Ogni voce trattata è corredata da una mappa che mostra la distribuzione delle sue diverse varianti in Sardegna. Per ogni voce si valuta se sia plausibile un'origine preromana e, in caso affermativo, se si conoscano parole potenzialmente correlate al di fuori della Sardegna.

La terza parte è dedita ad indagare se le voci sarde per le quali è plausibile un'origine di sostrato possano offrire indizi sulle caratteristiche della lingua di partenza preromana. Il capitolo 8 si concentra sui suoni della lingua: vi si conclude che il sistema fonico del sardo preromano non differiva notevolmente da quello del latino e che non ci sono prove sufficienti per la presenza di alcune caratteristiche preromane precedentemente proposte. Il capitolo 9 esamina le prove relative ai supposti morfemi preromani. In passato, cinque prefissi e dodici suffissi sono stati attribuiti a una lingua preromana. Di questi, due prefissi e cinque suffissi possono essere confermati come tali. Inoltre, nel capitolo vengono discusse le possibili funzioni grammaticali dei prefissi **kV-* e **ϑi-* e del

suffisso *-*Vr*. Non sono disponibili dati sufficienti per pronunciarsi sulla funzione degli altri affissi.

La quarta sezione riguarda l'affiliazione della lingua o delle lingue preromane della Sardegna. Nel capitolo 10 si discutono le proposte più rilevanti di lingue attestate ipoteticamente collegate con il sostrato sardo, ossia il punico, il berbero, il basco e l'etrusco. La presenza di un sostrato punico sembra essere la più plausibile, anche alla luce della presenza fenicia e cartaginese storicamente nota in Sardegna; il punico è stato parlato fianco a fianco con le lingue indigene sarde preromane fino a molto tempo dopo la conquista romana. Il piccolo numero di voci sarde con possibili parole correlate in berbero probabilmente va collocato in un contesto cartaginese. Per quanto riguarda il basco, solo quattro dei 19 raffronti lessicali proposti possono essere mantenuti. Questo numero è troppo esiguo per classificare la lingua indigena della Sardegna preromana come correlata con il basco. Più plausibile è che il basco sia stato in contatto indipendente con una lingua a sua volta strettamente imparentata con una lingua preromana della Sardegna. Solo per un'unica voce sarda è possibile individuare una corrispondenza etrusca – il che non sorprende data la nostra scarsa conoscenza del lessico di questa lingua. Tuttavia, vi sono alcune caratteristiche strutturali possibilmente condivise tra l'etrusco e il sardo preromano. Il capitolo 11 analizza le possibili prove di legami tra il sardo preromano e altre lingue di sostrato non attestate, al di fuori di Sardegna. Si nota che per varie voci sarde originarie di sostrato si possono trovare voci correlate nel Mediterraneo occidentale, con un centro di gravità nella Spagna nord-orientale e nella Francia meridionale. Questo suggerisce l'esistenza preistorica di una lingua o una famiglia linguistica preromana in quest'area. Ciò è in parte coerente con le precedenti proposte di un "substrato mediterraneo", seppure con una distribuzione geograficamente più ristretta.

La quinta parte (capitolo 12), in conclusione, contiene un riassunto più completo dei risultati della dissertazione e alcune prospettive su possibili ulteriori ricerche.

Curriculum Vitae

Cid Swanenvleugel was born on 23 July 1998 in Delft, The Netherlands. He grew up in Meerssen, where he attended Stella Maris College from 2009 to 2015. In 2015 he began a BA in linguistics at Leiden university, specializing in Comparative Indo-European linguistics and taking additional courses in Middle Eastern Studies. During these years, he developed an interest in the linguistic history of the Romance languages and spent a semester at the NIMAR institute in Rabat, Morocco. In 2019, he started an internship with the ERC-funded EUROLITHIC research project. In the same year, he obtained the BA degree with a thesis on livestock terminology in the Romance languages.

Immediately after, he enrolled in the Research MA program in linguistics at Leiden university, receiving his degree in 2021 with a thesis on Proto-Italo-Celtic reconstruction, supervised by Dr. Tijmen Pronk. Throughout his MA, he continued contributing to the EUROLITHIC project as a student assistant. From 2021 to 2024, he was employed as a PhD candidate at Leiden University, within the same project, supervised by Prof. Dr. Guus Kroonen. During this period, he presented at international conferences such as ICHL, (co-)authored a number articles, co-organized the Comparative Indo-European Lectures (CIEL) series in 2022-2023, and spent a semester at the University of Cagliari. In 2024, he taught a BA and MA course on the Historical Grammar of Latin.