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).

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*, *ϑoϑϑóṛi̯u* 'arum' etc. < ***ϑaϑarói̯*** (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 ***$h_2$end^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. *ǎrẹ* '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. *kùkkęrę* '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 **ˊVr-* (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. *ǧáγaru* '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 ****ʒ₁ayl-*** (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' < ***norḯu-***, 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);[397] 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**

[397] 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' < ***kè̯i̯a*** (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' < ***kĕ̯i̯a***; Gr. Hes. καίατα 'pits (split by earthquakes)', Gr. καιάδᾱς 'pit or cavern in Sparta, into which convicted persons were thrown' etc. < ***ka̯i̯u̯a-t/d-*** or ***ka̯u̯i̯a-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. *maròg̀n* '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)róu̯ia*** (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. *ruǧa* (Valsesia), Lomb. *róǧa* (Como), Ven. *roza*, Lad. *roia* (Val Gardena), *roa* (Val di Fassa), Friul. *roye* 'millrace' < ***rugu̯ia*** (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. **wV̆llī*/**wV̆llē* 'small lifestock (i.e. sheep and goats)' and Bb. **tē-b̲ălē* '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. *barraṡòne*, *barisòne*, *barraṡólu* 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	**ϑaϑaróḭ* '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. *ǎrẹ* '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-áru̯u* '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)róu̯a* '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	**i̯ákaru* '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. *maròg̃n* '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	**mutéḱ(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. *golóstḭu*, 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. **kèḭa* '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 **kai̯u̯a-t/d-* or **kau̯i̯a-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).

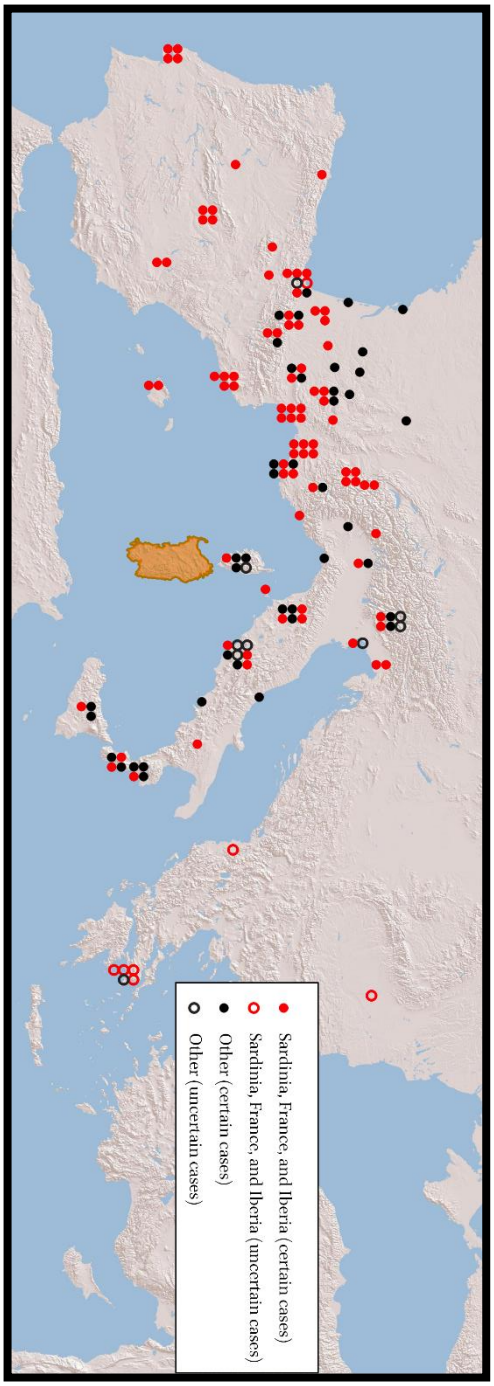

Figure 11.3.1: Distribution of pre-Roman words shared between Sardinia, Iberia, and France

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.

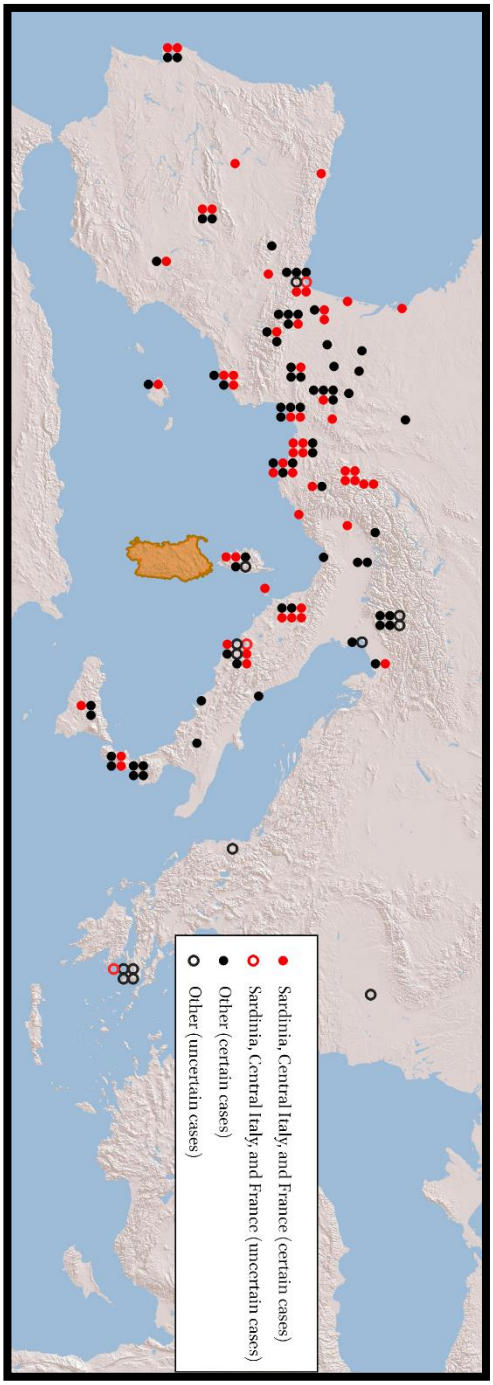

Figure 11.3.2: Distribution of pre-Roman words shared between Sardinia, central Italy, and France

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).