



BRILL



brill.com/ieul

A new assessment of Messapic vocalism

Sound change, script adaptation, and synchronic phonology

Michele Bianconi | ORCID: 0000-0002-9510-0834

University of Oxford, Oxford, UK

michele.bianconi@ling-phil.ox.ac.uk

Marta Capano | ORCID: 0000-0002-0210-3551

Università per Stranieri di Siena, Siena, Italy; Rijksuniversiteit Groningen,
Groningen, The Netherlands

marta.capano@unistrasi.it

Paolo Sabbatini | ORCID: 0009-0009-9144-3723

University of California, Los Angeles, CA, USA

paolosabbatini@g.ucla.edu

Received 12 September 2024 | Accepted 2 March 2025 |

Published online 21 May 2025

Abstract

This article re-examines the vowel system of Messapic, an Indo-European language attested from the sixth to the second century BCE in Apulia. Given the limited corpus of inscriptions, written in a local script derived from Greek, the phonology of Messapic remains difficult to understand. Through a critical reassessment of the literature and a first-hand analysis of vowel alternations, this study offers a new chronology for some sound changes: namely, it challenges the widely accepted theory of a pre-historic across-the-board **o > a* shift, drawing the possible inference that this sound change was still on-going within the historical period; in addition, in exploring the synchronic values and diachronic developments of back vowels and diphthongs, particularly within key morphological classes such as the **-tor-* and **-on-* stems, it suggests that the monophthongization of certain diphthongs may have triggered other sound changes. This approach offers a more nuanced understanding of the evolution of Messapic vowels and attempts to unify previous phonological models.

Keywords

Messapic – phonology – Indo-European – alphabets – back vowels – language contact

1 Introduction¹

Messapic (or Messapian) is an Indo-European language attested between the mid sixth c. and the end of the second c. BCE across the southernmost part of Roman Apulia.² Its denomination rests on an ancient ethnonym but is largely conventional in modern scholarship.³ Nearly all of our knowledge of Messapic depends on a corpus of ca. 650 inscriptions,⁴ and in minimal part on a handful of *glossae* preserved in ancient scholarship.⁵ The inscriptions, to this day not entirely understood, are written in a local variety of the Greek script that is highly dependent on the Tarentine alphabet, and are therefore easily readable.⁶ There is a high degree of variation in the usage of the alphabet, both at

1 This paper is the outcome of joint work by the authors, but for academic purposes the final editing is to be attributed to Michele Bianconi for sections 2.1, 4.1, 4.1.1–4, and 4.1.6, Marta Capano for sections 2.2, 3.1, 3.3–5, 4.1.5, 4.6–7, and to Paolo Sabattini for sections 3.2, 3.6, 4.1.2, 4.1.7, 4.2, 4.4. All authors are completely responsible for sections 1, 4.3, 4.5, 5, and 6. We wish to thank Brent Vine, Giovanna Marotta, and an anonymous reviewer for invaluable feedback on this paper.

2 For a general bibliography on Messapic, one can refer, among others, to de Simone (1980 & 2018); Lombardo (2015); Matzinger (2019); Marchesini (2020).

3 Although some ancient sources describe at least three populations—namely the Messapians in the south of Apulia (approximately corresponding to present-day Salento), the Peucetians to the West and the Daunians to the North, in what is now the Gargano area—alternative terminology for the ancient Apulians is available (cf. Marchesini 2020: 496). This variation is somewhat visible in our epigraphic documentation, where a mostly homogenous language is attested in Salento, while two other varieties, conventionally called Peucetian and Daunian, are present in inscriptions from the areas around Bari and in the northernmost part of Apulia, respectively.

4 The vast majority of Messapic inscriptions have been collected in 2001 in the *Monumenta Linguae Messapicae* (henceforth *MLM*) by de Simone & Marchesini, while a few inscriptions have been published since. In this work, we will solely refer to the inscriptions contained in *MLM*.

5 For a list of some of the Messapic lexemes attested in ancient grammarians and lexicographers, see Lombardo (1992) and Kassel & Austin (2001 (*PCG* 1): 321).

6 In general terms, the Messapic alphabet can be classified as ‘red’, according to the traditional nomenclature (Kirchhoff 1887), but it probably does not depend on a single source (cf. §3.1 below). In this article we will adopt the following conventions: Messapic forms cited for linguistic analysis or in continuous prose will be written in the *Latin alphabet*, according to the conventions in *MLM* (e.g., *klaohi*); graphemes or sequences of letters will be cited in the

the diachronic level, as it went through several phases (cf. *MLM* i 5–18), and at the diatopic level, and one can even identify a northern “Apulian” variety (cf. § 3.1 below).

There is little consensus about the dialectal position of Messapic aside from its belonging to the Indo-European family. Although some isoglosses seem to connect Messapic with the Italic languages (e.g., *bilia*, possibly cognate of Lat. *filia*) and with Greek (e.g., the change **s > h* intervocalically and prevocalically in *Anlaut*), which might be unsurprising by virtue of their geographical proximity, another popular hypothesis connects Messapic with the rather elusive notion of ‘Illyrian’, or with Albanian and its ancestor(s).⁷ Given the lack of secure isoglosses linking this language with other clades, it is widely thought that Messapic constitutes its own branch.

Given the shortcomings of the primary data, research on the phonology of Messapic is notoriously difficult to accomplish. This is the reason why progress in this field has been less swift than it has been in other Indo-European branches. However, given the relative increase in the dataset over the last few decades, it is now time for a reassessment.

This article sets out to offer a new interpretation of the vowel system of Messapic, starting from a critical re-reading of the existing literature, with a particular focus on the implications and validity of the *vulgata* (§ 2). We are then going to offer a fine-grained analysis of spelling alternations in Messapic vowels, focusing on spelling practices and their areal and diachronic distribution (§ 3); in doing so, we will make an appropriate distinction between loanwords and inherited material. The core of the article will consist of a specific case study concerning back vowels (**(\bar{o})* and **(\bar{u})*) and diphthongs, their synchronic values, and their diachronic developments, with reference to both the segmental inventory and the implications for two well-documented morphological classes, the **-tor-* and **-on-*stems (§ 4). We shall then endeavour to uncover the assumptions behind previous accounts of the Messapic phoneme inventory and test them against our dataset and hypotheses (§ 5). We will conclude by

Greek alphabet in SMALL CAPS and between angular brackets (e.g., <KAAOH1>); phonetic and phonological transcriptions will be given in IPA, between square and slanted brackets, respectively. Reconstructed forms will be noted according to the widespread conventions used in Indo-European studies.

7 This connection has been proposed mostly on the basis of onomastics, especially the name *Tautor/Teutor* (e.g., Krahe 1955; Matzinger 2019) and, if it were correct, most features that are shared with Italic could be explained via contact. But since we have no Illyrian inscription *per se* and this term is used for any unknown language of the Balkan peninsula in the pre-Roman period, at this stage we shall refrain from engaging with questions on the dialectal classification of Messapic and from explaining *obscura per obscuriora*.

bringing all these strands together and offering a new series of considerations on Messapic phonology which we hope will generate further research (§ 6).

2 The *vulgata(e)*: implications and validity

2.1 *Messapic vocalism*

Most recent publications on Messapic phonology (de Simone 2018; Matzinger 2019) by and large depend on a seminal 1972 work by Carlo de Simone, which gives a full account of what was then known about the Messapic language. According to this article, the lack of ⟨Υ⟩ in the epichoric Greek-derived alphabet indicates that Messapic has one fewer (back) vowel than Greek, which provided the model script. However, the fact that PIE short **o* has become /a/—a change that de Simone believes has taken place ‘con assoluta certezza’ (1972: 134)—guarantees that the one back vowel is /u/. This gives us a four-vowel system.

PIE			Messapic		
/i/		/u/	/i/		/u/
/e/	/o/		/e/		
	/a/			/a/	

One would then wonder why Messapic speakers did not adopt the sign ⟨Υ⟩ for a vowel /u/. De Simone (1972: 140) offers two possible scenarios: either (i) in the Greek of Tarentum the opposition between /o/ and /u/ was being neutralized with a phonetic realization as [ɔ], and Messapic speakers chose the sign ⟨O⟩ for their back vowel, or (ii) the back vowel of Messapic was [ɔ], closer to the sound noted with ⟨O⟩ rather than ⟨Υ⟩ in the Greek of Tarentum. De Simone prefers the second solution to the first, which he believes, quite rightly, to be unverifiable.

For the long-vowel system, de Simone recognizes five distinct phonemes: crucially, he claims that original **ō* is preserved as such, that it is written ⟨O⟩, ⟨OO⟩ or ⟨OHO⟩, and that it is found both in roots (e.g., *pido*, with ⟨ΔO⟩ interpreted as the reflex of **dō*- < **deh*₃-; cf. Gk. δίδωμι, etc.) and in suffixes (e.g., -*tōr*- in *platoras* and [p]latoor[r]ih[i] and -*ōn*- in *baledonas*, *kraotedonas*, *kriθonas* and *šaillonas*). The envisaged system eventually became canonical (see Fortson 2010: 467; de Simone 2018: 1843–1844; Marchesini 2020: 506) the only difference being that, in 1988, de Simone suggested that by the fifth c. BCE a change /o:/ > /u:/ took place. This phenomenon would be triggered by a need for symmetry in the system between short and long vowels.

Before fifth c. BCE		After fifth c. BCE	
/i:/	/u:/	/i:/	/u:/
/e:/	/o:/	/e:/	
	/a:/	/a:/	

A slightly different scenario is that envisaged by Matzinger (2019: 26–27), who states that Messapic *synchronically* lacked /o/ and /o:/, and that ⟨o⟩ corresponds to /u/ and /u:/ (cf. [pl]atoor[r]ih[i] in MLM 16 Al, equivalent to the *platur* of MLM 9 Ruv, to the *platuras* of MLM 2 Gro, as well as to the Greek and Latin adaptations Πλάτωρ, Πλατουρα, and *Platura*).

Neither system is implausible *per se*, and they both hinge on the assumption of an unconditional prehistoric change *o > a. There are two main issues with Matzinger's system. First, one does not understand why Messapic speakers did not choose the sign ⟨ϣ⟩ for their phonemes /u/ and /u:/. Second, all the forms like *platur* are only found in the northern district (and in Grotta Poesia).

De Simone, instead, does tackle the problem of the adoption of ⟨o⟩ for /u(:)/ but the implications of his reasoning are not spelled out as much as they could be. One case in point: in his preferred scenario (ii, where the one back vowel of Messapic was [ɔ]; cf. above) one is effectively left with a vowel system with no /u/. In order for Messapic, whose original /o/ merged with /a/, to borrow the sign ⟨o⟩ for its back vowel, one must necessarily presuppose a prehistoric lowering of original /u/, and therefore a different *synchronic* vowel inventory, with /o/ realized as mid open [ɔ] and written ⟨o⟩.

2.2 Méndez Dosuna on the fronting of Greek /u:/

Recently, Méndez Dosuna (2021) has proposed that the fronting of the vowel /u:/ (i.e., /u(:)/ > /y(:)/) was not limited to Attic and Eastern Ionic, as traditionally believed, but affected all Greek dialects and should therefore be placed in a later period of Proto-Greek. Evidence for this hypothesis includes cases of palatalization before /u(:)/ in medieval Greek varieties of Doric origin. Following this fronting phase, many dialects would have subsequently reverted to a back pronunciation, as evidenced most clearly in Boeotian and Arcado-Cypriot. Regardless of whether this reconstruction is correct, the evidence indicates that not only are there notable cases where palatalization failed to occur in Doric (e.g., the retention of the second-person singular pronoun τύ), but also that the use of the letter *qoppa*—particularly in the sequences ρο, ρω, and ρυ—suggests the presence of a back allophone of /k/, implying that a back vowel must have been synchronically present in the dialects that employed it.⁸

⁸ See, e.g., in Magna Graecia Φυβάλας, from Lokri Epizephiri (ILocri 41, 7–6 c. BCE) and in Sicily ρύλιξς from Selinous, ca. 620–580 BCE (ISico20447).

Laconian inscriptions from later periods exhibit multiple substitutions of ⟨Υ⟩ with ⟨ΟΥ⟩, indicating that the vowel was pronounced as a back vowel. Additionally, Laconian glosses in Hesychius confirm that this phenomenon was not merely an epigraphic convention but had a wider scope.⁹ While Attic and Ionic provide strong evidence for the fronting of /u(:)/ to /y(:)/, the Ionic dialects of Magna Graecia display alternations between ⟨Υ⟩ and ⟨ΟΥ⟩,¹⁰ as well as instances where ⟨Υ⟩ replaces ⟨Ο⟩.¹¹ This clearly demonstrates that, in these dialects, no fronting of ⟨Υ⟩ took place. Furthermore, the preservation of /u(:)/ in Tsakonian and other Modern Greek varieties of the Peloponnese, as well as in Southern Italian Greek, suggests that, regardless of how ⟨Υ⟩ was realized in later stages of Proto-Greek, early Doric reverted to the original Proto-Greek *u(:) as a back vowel. Given this evidence, there is no compelling reason to assume that the Tarentine dialect, like other Laconian dialects, underwent fronting of /u(:)/. Consequently, the grapheme ⟨Υ⟩ must have represented a back vowel.

Be that as it may, one still wonders whether there is any spelling variation involving Messapic back vowels and what it actually indicates. If this is the case, are there any consequences for our understanding of Messapic phonology? In particular, did Messapic have /o/ and /o:/ at the synchronic level? Both proposed systems are based on some underlying assumptions, often going back a few generations, and the question is whether these assumptions should be revisited. At the same time, most treatments focus on a small number of *Paradebeispiele*, often obtained via the etymological method alone, without taking into account the totality of the corpus. We maintain that the increased dataset now available, along with a refined understanding of comparative Indo-European linguistics, will help us make some progress in Messapic phonology.

3 Taking a step back: Script adaptation and spelling alternations

3.1 *Messapic alphabets*

In order to answer these questions, it is worth starting from first principles, which in this case means examining further the adaptation of the Greek script for Messapic and evaluating the consequences that this has for our understand-

9 Méndez Dosuna (2021: 127) mentions “ζούγωνεϛ for *ζύγωνες ‘plowing oxen’ (Hsch. ζ 173 L.-Cunn.) besides ζύγωνεϛ (Hsch. ζ 199 L.-Cunn.)”.

10 Among the examples mentioned by Méndez Dosuna, see, e.g., Ηἔρακλειού Πῆγίνου for Ηἔρακλειού Πηγίνου (IRhregion 57, ca. 475–450).

11 See Γλαυκίω = Γλαυκίω (IRhregion 26A, ca. 450).



FIGURE 1 6–5 c. BCE Messapic alphabetic sequence
AFTER *MLM* I 7–8 & *LOMAS* 2021: 329

ing of the language, before moving on to examining relevant cases of spelling alternations.

The traditional notion that the Messapic alphabet (cf. Fig. 1) is directly derived from the Greek alphabet of Tarentum¹² has been challenged by a more nuanced model. According to this revised position,¹³ the presence of letter shapes (e.g., <Δ>, <Β>, <Μ>, <Α>, <Σ>) not exclusively found in the Laconian-Tarentine alphabet suggests an early experimental phase in the 6–5 centuries BCE.¹⁴ This phase was marked not only by variation in the letter shapes, but also by a lack of absolute consistency in the direction of writing.¹⁵

As is often the case in the transmission of writing systems, signs can be (i) lost, (ii) re-functionalized, or (iii) created either on the basis of other signs present in the model or *e nihilo* in order to reflect the phonetics/phonology of the target language. In the case of Messapic, <Φ> was lost (i), <Ϝ> was employed to denote [v] or [f] (ii), whereas <ϣ> and <ψ> (probably from Greek <Ψ>) were used for two different sounds (iii), presumably [h] (or [j]) and [t]/[tʰ] (or even [t͡s]/[t͡ʃ]?).¹⁶ Strategies (ii) and (iii) are not mutually exclusive, as shown by the example of Greek <χ>. This letter was the likely model for both a refunctionalized <ξ> and for a new sign, <+>, which were used for a [ʃ] that Greek did not

12 See, e.g., Jeffery (1990: 279); further bibliography in Lomas (2021: 328⁴¹).

13 Cf. Marchesini (1999: 202 with n. 63), *MLM* i 6, and Lomas (2021: 328–330).

14 Some signs are probably derived from Corinthian (e.g., <ζ>, <Α> and <ξ>/<+>), others from Euboic alphabets (e.g., <Ϝ>); see Marchesini (1999: 202); de Simone (1972, 1983, & 1984); Lejeune (1991).

15 Around 25 % of the early inscriptions are sinistrowise, and 66 % of all right-to-left inscriptions are from the early period (Lomas 2021: 330).

16 The situation is rather complex, and here we are offering a very simplified outline. The two signs seem to have a chronological distribution, as <ϣ> is mostly found in archaic inscriptions, while <ψ> is more common in later documents (*MLM* i 9–10). However, on the basis of their environments, they do not seem to have the same phonetic value. The sign <ϣ>, alongside its possible variants <ϣ̣> and <ϣ̥> (as per Lejeune 1991: 220–221) is usually found between vowels (as in, e.g., <ΒΑΛΕΥΙΑΣ> or <ΗΑΙΦΑΥΙΑΣ>) and is therefore thought to note a hiatus-breaking sound. The alternation of <ψ> with <τ> and <θ> (as in, e.g., <ΤΑΒΑΡΑ>/<ΘΑΒΑΡΑ>/<ΨΑΒΑΡΑ>) seems to suggest something phonetically not far from a dental value. On these signs, see de Simone (1972: 171–176, 1979: 111–112, & 1983: 189) and Lejeune (1991: 217–223).

possess. In some cases, such as that of the arrow-shaped sign ⟨↑⟩, it is difficult to establish whether a grapheme has been created *e nihilo* or by the modification of an already existing sign.

By the second quarter of the fifth c. BCE, variation in letter shapes diminishes and a dextroverse direction of writing becomes the standard. Whether this be due to the formation of a ‘national alphabet’ (as per Marchesini 1999: 203–204) or changing teaching practices and extent of literacy (as per Lomas 2021: 332) need not concern us here. What is relevant for our purposes is to note that this more stabilized inventory was now less prone to the changes that would independently take place in the model alphabets. If we look at vowels, in the Doric of Tarentum, besides /a/, /i/, and /u/ there were originally only two long mid vowels [ɛ:] and [ɔ:] (originally written ⟨E⟩ and ⟨O⟩ and, from the fourth c. BCE, ⟨H⟩ and gradually ⟨Ω⟩). However, with the later diffusion of the *koine* (post third c. BCE), the two new long mid vowels [e:] and [o:] were represented with ⟨EI⟩ and ⟨OY⟩, as in other parts of the Greek-speaking world, starting with Classical Athens (cf., e.g., Allen 1987: 69–79). This is hardly reflected in our Messapic corpus, which remains largely unaffected by these changes.¹⁷

However, a different story should be told for the inscriptions written in the Apulian alphabet, which are found in Daunia and Peucetia from the fourth century BCE onwards (cf. Fig. 2).¹⁸ These documents exhibit some differences from the southern (proper ‘Messapic’) alphabet, as the use of the grapheme ⟨H⟩ for a sound [ɛ:] and of ⟨Y⟩ for a back vowel. The usage of ⟨Y⟩, both in diphthongs and on its own, is otherwise only attested in the southern sanctuary of Grotta Poesia.

3.2 *The role of script variants in the analysis*

For the purposes of this paper, our dataset includes inscriptions written in both script variants, namely Messapic ‘proper’ and Apulian (as presented in *MLM*). This choice does not merely respond to needs of completeness but also has one crucial diagnostic advantage. As we have seen in § 2.1, the vowel inventory traditionally posited for Messapic hinges on the lack of ⟨Y⟩ in the Messapic (i.e.,

17 One should note the single occurrence of ⟨Ω⟩ in *MLM* Gn 26 and a small number of (mostly late) attestations of ⟨H⟩ for long /e:/ (cf. figures in n. 60).

18 Cf. de Simone (1988: 344–349); Lejeune (1991: 211–231); Lomas (2021: 332–334); de Simone rightly notes that this Apulian alphabet is not very different from the Hellenistic one, despite having preserved the Messapic trident signs and *digamma*. However, unlike the Hellenistic alphabet, it does not have a dedicated sign for /ɔ:/ (i.e., it lacks ⟨Ω⟩). Moreover, there are some morphological differences (e.g., the shape of ⟨Σ⟩, which can appear as lunate in Apulian) that are due to the later date of the introduction of the alphabet.

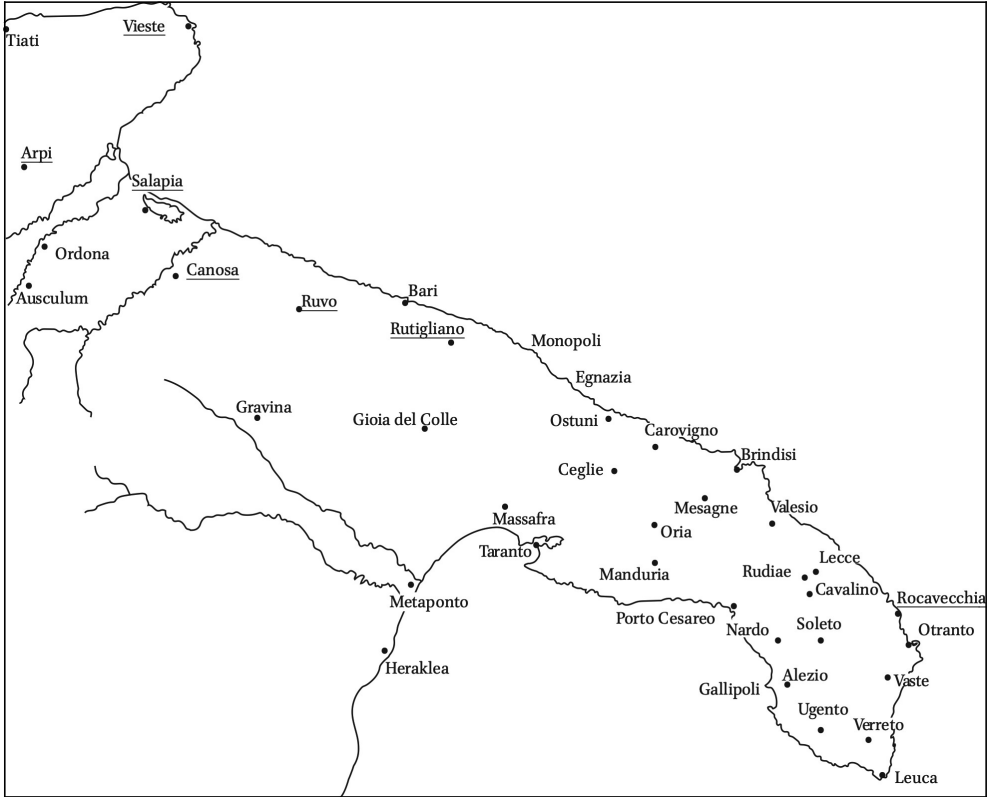


FIGURE 2 Inscriptions in the Messapic and Apulian (underlined) script
AFTER LOMAS 2021:333

non-Apulian) alphabet. Claims such as the across-the-board merger of **o* and **a* and the development **ō* > /u:/ therefore follow from a structural limitation of the evidence, and are hardly falsifiable in the absence of a writing system that can graphically represent more phonetic contrasts within back vowels. The Apulian script, as a matter of fact, can do so. Since, for all we know, the distinction between Messapic and Apulian is only graphic but not linguistic (i.e., these are *not* two different dialectal varieties), we would expect the distribution of <O> and <Y> in Apulian inscriptions to exactly match the phonological models based on the Messapic material. Namely, these two signs should be used interchangeably to notate either /o:/ and /u(:)/ (de Simone 1972) or just /u(:)/ (Matzinger 2019), making no distinction between back vowels with different historical sources and/or synchronic phonetic realizations. The sign distribution in the Apulian inscriptions does in fact point to more fine-grained phonetic contrasts on the synchronic level, which in turn call for a different

understanding of the (pre)history of the Messapic vowel system. To substantiate this point, we now present the full array of token variation involving signs for back vowels in our dataset, to then further conceptualize our findings in the remainder of the paper.

If one looks at different vowels in what *seem* to be the same lexemes, it appears that the Messapic corpus shows a number of alternations, which we report in the *Appendix*. It seems that these alternations affect all vowels, but we have already seen in § 2.1 that the analysis of back vowels has important implications for our understanding of prehistoric sound changes and, consequently, of the dialectal position of Messapic within the Indo-European family. Therefore, despite the inclusion of front vowel alternations in our count,¹⁹ the following analysis will focus on back vowels. In doing so, we are also taking into account diphthongs and long vowel notation.

3.3 *Alternations in our corpus*

Considering alphabetic variation, we notice two types of alternations within non-front vowel signs: whereas ⟨A⟩ ~ ⟨O⟩ is extremely rare—cf. *ḡobara* (1×) ~ *ḡabara* (3×)²⁰—variation between ⟨OO⟩ and either ⟨O⟩ or ⟨Y⟩, depending on the alphabet ('Apulian' ⟨Y⟩), is more common. However, it should be noted that the alternating forms are not always identical. For instance:

19 Messapic shows some token variation in the notation of front vowels, especially but not exclusively in the first syllable of the word, both in what might be inherited material (e.g., *dazomas* ~ *dezomas*, cf. Gk. δάζιμος) and in what seem to be loanwords (e.g., *artamih* ~ *artemes*, cf. Att.-Ion. ἄρτεμις, Dor. ἄρταμις, Myc. *a-te-mi-to* ~ *a-ti-mi-to* GEN.SG.). The main types of alternation we were able to identify are between ⟨A⟩ and ⟨E⟩ (documented in five lexical items, e.g., *dazomas* ~ *dezomas*), between ⟨E⟩ and ⟨I⟩ (two items, e.g., *dazehi* ~ *dazih*), and between ⟨EI⟩ and ⟨I⟩ (seven items, e.g., *deiva* ~ *diva*). Although these alternations are empirically equivalent to those involving back vowels, they have not been associated to comparable sound changes in the phonology of Messapic. Ascertaining whether they may reflect *bona fide* phonological processes or rather be the by-product of loan adaptation processes or different orthographic traditions is an endeavour we will take on in future developments of this research.

20 In addition to *ḡobara* ~ *ḡabara*, Messapic inscriptions and indirect sources point to another instance of this alternation: the Messapic personal name *Arta* (which—taking a maximalist approach—appears in six inscriptions; see *Appendix*) is also attested as *Artos* (MLM 12 Ruv, 3 c. BCE; MLM 10 Ar, second half 3 c.—end of 2 c. BCE). Since the inscriptions featuring *Arta* are roughly contemporary with those bearing *Artos*, this variation may represent another instance of the Messapic phonological shift from *o* to *a*, potentially still in progress. The same alternation occurs in literary sources: Thucydides (VII 33.3–4) names a Messapic king *Artas*, whereas Athenaeus (quoting Polemon, fr. 89 Preller, *Deipn.* III 108 F) refers to the same individual as *Artos*. However, Matzinger (2019: 114) prefers to explain the alternation by identifying *Artas* as the genuinely Messapic form, with *Artos* representing its Hellenized adaptation.

- (1) a. *poollo* (1×)²¹ ~ *polla* (1×) ~ *polonnihi* (1×)
 b. *platoorrihi* (1×) ~ *platur* (2×)

There are a few examples of forms that seem to point towards a ⟨I⟩ ~ ⟨Y⟩ and possibly a ⟨I⟩ ~ ⟨Y⟩ ~ ∅ alternation, which is traditionally explained with something equivalent to the Latin *sonus medius*:

- (2) a. *vinaihi* (1×) ~ *vunaihi* (1×), both from the sanctuary of Grotta Poesia
 b. *damatira* (1×) ~ *damatura* (1×) ~ [d]amatra (1×)²²

The alternations involving the ⟨H⟩ sign as perhaps a length marker—⟨OHO⟩ ~ ⟨AO⟩ in *bohonihi* (MLM 33 Rud, 3 c. BCE) ~ *baonnes*? (MLM 19 Cae, 3 c. BCE) and ⟨OHO⟩ ~ ⟨I⟩ in *dazohonnihi* (MLM 1 Bas 3 c. BCE) ~ *dazinnihi* (MLM 18 Bas 3 c. BCE)—seem more difficult to explain.²³

Diphthongs also display a high degree of variation and alternations:

- (3) a. ⟨EO⟩ ~ ⟨O⟩: *ϑeotor* (2×) ~ *ϑotor* (4×)
 b. ⟨AO⟩ ~ ⟨O⟩: *taotori* (4×) ~ *totor*[i] (1×)
 c. ⟨AO⟩ ~ ⟨AY⟩ ~ ⟨O⟩: *klaohi* (2×) ~ *klauhi* (2×) ~ *klohi* (3×)
 d. ⟨AY⟩ ~ ⟨EO⟩: *daus* (2×) ~ *deos* (3×)

Some of them might be graphic, such as the correspondence between ⟨AO⟩ in the Messapic alphabet and ⟨AY⟩ in the Apulian alphabet, but others seem to point to changes on the phonological level, as ⟨EO⟩ ~ ⟨O⟩ and ⟨AO/AY⟩ ~ ⟨O⟩ most probably reflect monophthongization. It would be tempting to look for a relative chronology of the ⟨AO⟩ ~ ⟨EO⟩ ~ ⟨O⟩ alternation, but the data do not provide any evidence for it.

21 It might be worth noting that the forms with the geminate ⟨OO⟩ often show gemination of the liquid consonant as well.

22 If we had not had the Apulian data, we could have noticed a reduction *damatira* ~ [d]amatra, but thanks to the inscription from Ruvo (*damatura*, MLM 10 RuV, 3 c. BCE) we can determine at least that we have a back vowel.

23 The usage of ⟨H⟩ as a marker of vowel length is paralleled in Umbrian (e.g., ⟨ah⟩ and ⟨ah⟩/⟨aha⟩ for /a:/). While the document attesting the first alternation is currently available, the inscription with *dazohonnihi* is lost, and we only have a 1588 transcription that survives in two copies (MLM i 113). If ⟨OHO⟩ = /o:/ ~ ⟨I⟩ = /i/, the only similarity one could think of is that to the Latin type *origō*, *-inis*, which however owes the *-i-* in the oblique stem to Latin Vowel Weakening (< **-en-*). Alternatively, as pointed out to us by Brent Vine (p.c.), if ⟨I⟩ here can reflect a raised **ē* (as is the case in Celtic and some varieties of Sabellic), then we would have the reflexes of **-ōn-* ~ **-ēn-*, which would make sense historically. An anonymous reviewer suggests instead that a morphological account would be preferable, with both **dazōn-* and **dazīno-* being derivatives of an *o*-stem **dazo-*.

3.4 *Problems with the chronology of Messapic inscriptions*

The dating of Messapic inscriptions presents significant challenges. While de Simone and Marchesini have proposed a relative chronology of the Messapic texts based on both palaeographic and archaeological evidence (*MLM* i 5–18), their analysis also integrates comparative evidence from vase painting (e.g., ‘phase VII’, see *MLM* i 16), an approach that, while potentially informative, carries well-known methodological risks (e.g., Jeffery 1990: 63–64). The authors themselves acknowledge that certain texts cannot be reliably dated, as exemplified by those included in ‘phase IV’, which spans between the fourth and third c. BCE.²⁴ Moreover, while inscriptions can often be attributed to earlier (e.g., *MLM* 5 Cav) or later (e.g., *MLM* 24 Rud) ‘phases’ of the language, those classified under ‘phase V’ and ‘phase VI’ of *MLM* are uniformly dated to the third century BCE, which partially overlaps with ‘phase IV’. Consequently, of the ca. 500 inscriptions edited in *MLM*, 264 are assigned to the third century BCE, making it difficult to establish finer chronological distinctions.

3.5 *The role of loanwords*

To navigate these difficulties, loanwords can be used to establish a *terminus post quem*, as the presence of, e.g., Latin loanwords cannot predate the first contacts between Messapians and Romans. However, reconstructing a relative chronology of Messapic phonology based on variation in the spelling of loanwords has its own set of challenges. For example, the Latin personal name *Curtius* appears twice in Messapic texts in the genitive, possibly with the ending *-ihi* (*kor.θ.ihi* in *MLM* 1 Lup, 3 c. BCE) and then *-eih* (*korθeih* in *MLM* 10 Uz, second half 3–late 2 c. BCE). Based on the Latin evidence, we might assume that the alternation <EI> ~ <I> was ongoing around the 3–2 c. BCE, but in fact it is attested already at least one century earlier, as we can see from the probably indigenous word *blatθeihas* (*blatθei«h»as*, *MLM* 16 Gn, 4 c. BCE), which is instead attested with <I> only later (*blatθihi* *MLM* 26 Al, 3 c. BCE). These two examples, together with others in the *Appendix*, show that at the current state of research, epigraphy alone is a valuable but not conclusive tool for dating language phenomena in Messapic, and that while language contact provides a broad chronological framework, each case must be assessed individually as existing chronologies may require further refinement.

24 See, e.g., ‘La fase due è una fase di transizione, che non offre di per sé elementi datanti’ (*MLM* i 8) or, on the so-called ‘phase IV’ inscriptions: ‘Le iscrizioni attribuite a questa fase potrebbero appartenere da un punto di vista tipologico sia al periodo precedente, sia al periodo successivo’ (*MLM* i 9).

3.6 *Empirical considerations*

We now turn to the discussion of the empirical bases for Messapic sound changes affecting back vowels on the grounds of the current understanding of IE linguistics and the evidence from the published Messapic corpus. We shall assess the validity of the traditional etymologies, which often go back to the nineteenth or the early twentieth century, and still undisputedly inform most descriptions of Messapic historical phonology. In doing so, we shall attempt to distinguish inherited material from loanwords as much as possible; we will show that in many cases in which there is doubt between the two, the loanword hypothesis is the safest.²⁵ Although this inevitably reduces the dataset of inherited material, on which traditional sound change analyses rely, we suggest that a diachronic model based on synchronic alternations (cf. § 5) more effectively captures the complexity of the Messapic data.

4 The back vowels

4.1 *o > a in Messapic

One of the few certainties in Messapic phonology is that PIE *o developed into a.²⁶ This is generally maintained on the basis of a number of Messapic words supposedly going back to PIE formations with an *o in different lexical and morphological contexts:

1. Root syllables
 - i. *tabaras* < *to-b^hor-o-s (o-grade of *b^her-; semantically, cf. Umb. ařfertur 'leading priest')
 - ii. *-pandes* (in *argora-pandes*) < *argVro-pondyos (cf. Lat. *pendō* 'weigh', *pondus* 'weight')
2. Linking vowel
 - iii. *argora-pandes* < *argVro-pondyos
3. Inflectional endings
 - iv. o-stem NOM.SG. -as < *-os; e.g., *dazimas* (cf. Gk. Δάζιμος)
 - v. o-stem ACC.SG. -an < *-om; e.g., *argorian*

25 Contact hypotheses are in principle reasonable in a very multicultural and multilingual landscape such as that of Messapia; see Marchesini (2013) and the articles in D'Andria & Lombardo (1999); on Greek and Latin historical sources, cf. Lombardo (1992).

26 Cf. the strong assertion in de Simone (1972: 134): 'È un fatto assolutamente sicuro che *ō* (*o* breve) indoeuropeo è rappresentato in messapico da *ā*, o meglio, in termini funzionali, l'opposizione *ō*: *ā* è stata abbandonata ('neutralizzata') e come unico rappresentante messapico abbiamo *a*. Questo fenomeno è documentato in messapico con assoluta sicurezza.'

- vi. *o*-stem GEN.SG. *-aihi*; e.g., *dazimaihi*
 - vii. Neuter *s*-stem NOM./ACC.SG. *-as* < **-os*; e.g., *venas* < **wenos* (Lat. *Venus*, Skt. *vanas-* ‘desire’)
 - viii. *C*-stem GEN.SG. *-as* < **-os*; e.g., *kalatoras*, *baledonas*
 - ix. Athematic DAT.PL. *-bas* < **-b^h(y)os*; e.g., *logetibas*
4. Prepositions
- x. *apa-* < **apo* (cf. Gk. ἀπό ‘from’)
 - xi. *hipa-* < **supo-* (cf. Lat. *sub* ‘under’, Gk. ὑπό ‘under’, etc.)

The purpose of what follows is solely to demonstrate that the data can be interpreted in multiple ways, and that at this stage of our knowledge of Messapic diagnostic conjectures are more useful than definitive conclusions. While some traditional etymologies and explanations remain valid, these alternative interpretations highlight the complexity of the data. As we will see throughout this section, two possible scenarios concerning the reflexes of PIE **o* in Messapic emerge: we are confronted either with an on-going sound change or with one that has a different scope and chronology from what was thought until now.

4.1.1 *tabara-*

This is allegedly the most secure etymology in Messapic.²⁷ This noun appears with different initial graphemes in the earliest inscriptions (*tabara-*, *ḡabara-*, *ʰabara-*) and this already poses a problem for a solid etymology—especially if one attributes some phonological significance to ⟨θ⟩ and ⟨ʰ⟩; cf. § 3.1 above. But even if one accepts the traditional **to-b^horos* (cf. Dunkel 2014: 772–773), its relative isolation would still be an issue, as the Anatolian comparanda Hitt. *tabarna-* and CLuw. *tabariya-* are uncertain, while the Celtic preverb *to-* has a different function.²⁸ Finally, pushing the typological comparison with German *darbringer* a step further, i.e., taking *ta-* as the reflex of the demonstrative, would be even more problematic, because the initial **to-* would need a considerable number of semantic shifts (preverb < adverb (?) < demonstrative; cf. Germ. *dar-* < PGmc. **-þār* < PIE **tó-*) in order for it to be semantically transparent.²⁹

27 The idea was first developed by von Blumenthal (1936: 98–103) and then generally accepted; see, e.g., Pisani (1964: 237); Parlangèli (1965: 183–185); de Simone (1972: 135); Matzinger (2019: 12).

28 The Luwo-Hittite noun *tabarna-/labarna-* could well have developed from a personal name, while its connection with Luwian *tapariya-* ‘rule’ is debated (cf. EDHIL 520 & 830–831 with references). The item usually connected with Irish *to-* is the Old Hittite sentence-initial particle *ta* (cf. Watkins 1963: 13–14), but both are clearly different from the lexical(ized) *ta-* of Messapic.

29 The common comparison with Umb. *aifertur* ‘leading priest’ is also difficult, as this form

4.1.2 *argorapandes* and *argorian*

Pace de Simone (1972: 136¹⁹),³⁰ in this case it is important to understand whether we are dealing with a borrowing or with inherited material. The sequence *argora-* is similar to another sequence, *argorian* (*MLM* 1 Br, 3 c.), usually translated as ‘silver’, ‘coin’. This could theoretically be (i.a) a loan from Greek ἀργύριον or (i.b) from another formation based on the same root (e.g., LSJ⁹ has Ἄργουρα as a place-name); (ii) an inherited formation, although very similar to the Greek one. Now, it seems rather unlikely that this is anything other than a borrowing; it would be too exact a match with Greek, not just for the *-ro- suffix, but also for the preceding back vowel. Moreover, given that palatalization is a rather common phenomenon in Messapic (e.g., *ʰaotorra* < **ʰaotor-ya*),³¹ one would wonder why it did not happen in *argorian* and why we do not instead have a form †*argorran*. This seems to be a further hint that the form is a borrowing from ἀργύριον (i.a), as it retains the syllabification of Greek, where /i/ is syllabic. If then the form was adapted from Gk. ἀργύριον, the spelling *argorian* calls for an explanation. If we assume the *a-o* merger as prehistoric, it follows that the form was borrowed into Messapic in its prehistory, before the change had taken place—otherwise *argorian* could not display /a/ for Gk. /o/ in the ending. Under this assumption, the word could not have been borrowed after the change ceased to operate, as in that case we would expect Gk. /o/ to be mapped onto the one back vowel synchronically available in Messapic, written with <O>, rather than onto /a/. In other words, we would expect a spelling

is structurally different from *tabara-*, both in its derivational class and, more simply, in terms of *Ablaut* grade of the root. In addition, and crucially, the initial element *ař-* is synchronically a preverb in Umbrian and historically a preposition, and therefore would not require significant semantic shifts in this formation. An anonymous reviewer points out that Ved. *tád* parallels the neuter of the demonstrative becoming an adverb, and that Gk. τó-φρα ‘up or during that time’ contains **to-* and a form of **b^her-*. The problem of the morphological and lexical isolation of *tabara* would however remain.

30 È irrilevante in questa sede il problema ... se mess. *argora-* rappresenta un prestito dal greco ἄργυρος o se si tratti invece di una comune eredità indoeuropea (**arguro-*). This could in fact be crucial either for understanding the phonetic adaptations in loanwords (and thus the phonetic/phonological system of Messapic) or for refining the dialectal position of Messapic within the IE family. For instance, the formation **arguro-* is specifically Greek, although it is based on an inherited (Caland) *u*-stem **h₂érǵ-u-* (cf. Rau 2009: 131) which also gives Ved. *árjuna-*, and crucially differs from, e.g., Celt. *arg-ant-* or Lat. *arg-ent-*; if *argora-* were inherited, it would suggest that Messapic is rather close to Greek, at least given this shared **ro-* formation.

31 On palatalization, see Krahe (1928: 81–104); de Simone (1972: 156–168 & 1983: 187–191); Orioles (1972: 72–80); Lejeune (1991: 220); Ciceri (2013); Matzinger (2019: 30 & 32–33); Dupraz (2022: 172); Léger Pirus (2024).

†*argorion*, which is unattested. Our only *terminus ante quem* for the borrowing of the form, however, is the third century, which in and of itself cannot confirm this hypothesis. If, on the other hand, we do not take the prehistoric *a-o* merger as an assumption, two alternative explanations for the final *-an* are possible. Potentially, *argorian* as a borrowing could be integrated into Messapic morphology (cf. ACC.SG. *-an* in § 4.1.3). Preferably, though, the change of *o* to *a* could still be in progress at the time of the borrowing, so that a Greek /*o*/ could be mapped onto a transitional stage [a], i.e., a back vowel, and eventually noted with <A>. As for the inscription with *argorapandes* (*MLM* 1 Car, 3 c.),³² which simply reads *Θotor argorapandes*, it is currently lost and we exclusively rely on a nineteenth-century drawing. Therefore, further caution should be exercised. This word is traditionally explained as the reflex of an **arguro-pondyos*, with the second element usually connected to Lat. *pendō* ‘suspend, weigh’. In this case, one would wonder whether the change **o* > *e* (**pondyos* > **pandes*) is phonological or morphological. In addition, the absence of any signs of palatalization in the reflex of **-pondyos* (normally **θεоторιδ-ya* > *θεоторιδα*) calls for an explanation: this might have been the consequence of the post-nasal position of the **-dy-*, but it would also be compatible with the loan hypothesis. Since the number of Greek compounds in ἀργυρο^ο is very high, an alternative might be desirable: *⁰pandes* could have been borrowed from Gk. *⁰φάντης*, and *argorapandes* would be ‘silver-workman’, although Greek does not offer a perfect correspondence.³³ In this instance, we can suppose that either the voicing opposition was neutralized after nasals already in Greek (as is the case in several dialects) or one could suppose a trivial voicing assimilation in Messapic.

4.1.3 Case endings

The ending *-a(-)s* is used to mark either the thematic NOM.SG. or the athematic GEN.SG.: because of this, it is universally interpreted as the reflex of the

32 On the status of *argorapandes* within the Messapic lexicon, whether it is a loanword or genuinely Messapic, see Parlangeli (1960: 265); de Simone (1972: 134–136 & 218); and Giacomelli (1979: 49–50).

33 In Greek there are 16 compounds with *⁰φάντης* (a derivative from the root of *φαίνω*), e.g., *σuxoφάντης* ‘slanderer’, *τυμβοφάντης* ‘one who shows a tomb’, *ὕδροφάντης* ‘one who finds water’, *ἱεροφάντης* ‘religious congregant’. While there is no attested *†ἀργυροφάντης*, one could in theory explain *argorapandes* as a Messapic neologism made with the borrowed material *ἀργύριον* and *⁰φάντης*. If this were the case, *argorapandes* could then be interpreted either as the category of person who dedicated the inscription (e.g., ‘*Θotor* the silversmith’), or as an epithet of the god *Θotor* (e.g., ‘*Θotor* silver-bearer(?)’). With this scenario, one would obviously need to presuppose a *koine*-based formation, as we have *-της* instead of Dor. *-τᾱς*.

thematic nominative (*-o-s) in one case and the athematic genitive (*-os) in the other. Similarly, the ending *-an* is taken to be the marker of the thematic ACC.M.SG. and NOM./ACC.N.SG. (both from *-om). Despite scholars' best efforts, however, establishing clear-cut morphological classes in Messapic turns out to be a rather daunting task because of the lack of clearly established inherited material. If one were to play devil's advocate, the ending *-as* could *in principle* be a sigmatic nominative singular of an *ā*-stem, and *-an* could be its accusative singular. Incontrovertible examples of inherited neuters in *-an* < *-om are lacking: as we have just seen (cf. § 4.1.2), *argorian* is a loanword, and in theory nothing stands in the way of it being adapted to a different morphological class (e.g., a productive non-neuter declension in *-as*, *-aihi*, *-an*, etc.). As for the athematic ending *-as*, in purely IE terms the athematic GEN.SG. ending could *in principle* be *-es (cf. Lat. *-is*).³⁴ More importantly, though, de Simone himself (2018: 1844) admits the existence of a genitive in *-os*, which we find for instance in *penkeos* (cf. *MLM* 1 Sa). This creates an even greater aporia for any theory that postulates an across-the-board *o > a change happening strictly in a prehistoric period.³⁵ Once more, positing a sound change in progress with <A> reflecting a transitional pronunciation [a] can more easily reconcile the alternation *-as* ~ *-os*.³⁶ Similar considerations apply to the genitive ending *-aihi*, which *may* contain an [a] that is the outcome of prehistoric *o, but whose origin is highly debated.³⁷

34 The PIE athematic genitive singular ending *(e/o)s might appear in the *e*-grade as *-es (e.g., Lat. *canis*, with *-is* < *-es), in the *o*-grade as *-os (e.g., Gk. πῶδ-ός) or zero-grade (e.g., Ved. *agné-s*). By and large, languages tend to generalize only one ending, but on a sub-standard or dialectal level one can find instances of, e.g., an *o*-grade ending in Latin: e.g., non-urban (and possibly archaic) Lat. *VENERVS* < **wenesos*, *KASTORVS* (morphologically comparable with Mess. *kalatoras*), *NOMINVS*, *CAESARVS*. See Baldi (2002: 324) and Weiss (2020: 219) for more examples.

35 Matzinger (2019: 46) acknowledges the problem but does not call into question the chronology of the sound change.

36 The *Paradebeispiel* offered by de Simone (1972: 146) for a GEN.SG. *-as*, *kalatoras*, is a loanword from Latin *calātor*. The inscription (*MLM* 1 Tar), on a bronze *caduceus*, is complete, well preserved, and composed of three words in the genitive (*blatθihi kalatoras baletθihi* 'of *Blatθes*, the *calator* of Valesium'). The use of borrowed material to substantiate de Simone's claim could be problematic but, as an anonymous reviewer points out, the ending *-as* is attested in native personal names, such as *Molonas* (*MLM* 26 Bal; but cf. Gk. Μόλων).

37 This is one of the thorniest problems in Messapic linguistics; see contrasting opinions in Orioles (1991: 165–167), Gusmani (2006), Prosdocimi (2006), Ciceri (2012), de Simone (2018: 1845), and Matzinger (2019: 37).

4.1.4 *hipa*

This preposition (Dunkel 2014: 746–747) is only found in *hipades*, which is interpreted as a 3SG. aorist ‘(s)he dedicated’ from a preform **supo-d^hē-s-t*.³⁸ This presupposes a change **s > h / #_V*, which however leaves the many examples of the preservation of /s/ before a vowel (*sabaleidas*, *sada*, *saihikas*, *sal*, *sallih*[i], *sile*, *sitta*, *sohinnes*, *solahiaihi*, *soolles*, *souvoθoi*, etc.) without an explanation. More crucially, we find a very similar form *opades* (MLM 11 Ro; cf. de Simone 1988: 382), which renders this etymology hardly viable. Therefore, we do not regard this piece of evidence as conclusive either.

4.1.5 *logetibas*

This form appears on a funerary inscription from Alezio (MLM 10 Al), now lost, which only consists of two words, *laidehiabas logetibas*, presumably nouns in agreement with each other or a noun and an attribute. While the reading and interpretation of the first word is uncertain (see MLM ii 221), the second one is clear and has a long-standing etymology. According to de Simone (1972: 137), *logetibas*, which he believes to be a theonym, would be the dative/ablative plural of an athematic stem, with *-bas* from PIE **-b^hos*. The dative/ablative ending **-b^h(y)os* is currently not projected all the way back to PIE, but is considered a late feature (cf., e.g., Lundquist & Yates 2018: 2088–2090), developed through the fusion of the instrumental ending **-b^hi* with the original PIE dative plural ending **-os*. Specifically, **-b^hos* is reconstructed at least for the Italic and Celtic branches of the Indo-European family,³⁹ which could in theory point to a genetic connection between Messapic and the Italic languages, although this is not explicitly suggested by de Simone. However, since *logetibas* would be the only clear attestation of this dative/ablative plural ending in Messapic, much of this morphological analysis hinges on the etymological interpretation of the lexeme itself. Moreover, this very form is also a case of underapplication of the **o > a* change, which makes it somewhat problematic. Both issues are addressed in § 4.2 below.

4.1.6 *apa*

This form is thought to be a preposition cognate with Gk. ἀπό (< **h₂ep-*; cf. Dunkel 2014: 70–71); this is however most uncertain, as the sequence is only attested in three inscriptions. One of them (MLM 5 Uz) is on a lost stone

38 On *hipades*, see Krahe (1932: 120); de Simone (1962: 122); Orioles (1981: 148); Matzinger (2019: 77). This could be a verb from a purely combinatory perspective.

39 The other outcome of **-b^h(y)os*, *-bhyas*, is attested in Indo-Iranian, cf., e.g., Ved. *-bhyas*, Old Avest. *-biio*.

whose drawing has an incomplete sequence <NININKEPMAΘΙΑΠΑ> in its final line, and the other two feature the divine name *Aprodita* immediately before it: *MLM 1 Ga: klohi zis aviθos θotoridas ana aprodita apa ogrebis*; *MLM 19 Ur ...ap[r]odita apate[.].* In *MLM 5 Uz*, the context is obscure, but it is very clear that *apa* closes the inscription, therefore making its interpretation as a preposition or a preverb impossible (it could at most be a postposition or an adverb). Aside from the similar collocation, the context of *MLM 19 Ur* is obscure, so one can only really work with *MLM 1 Ga*. Here, most elements are structurally clear: we start with the *klohi zis* collocation,⁴⁰ followed by two personal names (presumably in the genitive), in turn followed by *ana* and the theonym *Aprodita*, usually thought to be both in the dative.⁴¹ The last word, *ogrebis*, is thought to be a DAT./ABL.PL. in *MLM* (ii 261) and an INSTR.PL. by Matzinger (2019: 39), and to mean ‘oaths’ (*vel sim.*). This is not impossible, but a dedication verb would perhaps be more fitting in this context, on the force of a parallel such as *MLM 24 Cae (plastas moldatθehiai bilia etθeta hipades aprod(i)ta)*, in which *hipades* fulfils such a function. If *hipades* should indeed be interpreted as a verb in the third person (regardless of its traditional etymology, cf. § 4.1.4 above), and if the vocalic alternations we find in pairs such as *mogamates-mogamatis*, *megones-mēgonis* (i.e., <E>/<I> in final syllables, especially before -s#) should be taken at face value, then one wonders whether the *apaogrebis* sequence in *MLM 1 Ga* should be considered a *verb*, perhaps with a preverb, rather than a preposition + DAT./ABL./INSTR.PL. If this were to be the case, nothing would prevent a segmentation *ap-aogrebis*, with elision of the second vowel of the preverb, which would then remain unknown. The other case where there seems to be a similar formation is *apeiūstaθi/apistaθi*, which *MLM* (ii 18–19) derives from **opi-* + **steh*₂- and compares to Greek ἀφίστημι (but note that this would be from PGk. **ap-histā-mi*), but once again the segmentation is not unambiguous (one could indeed envisage a *apeiūstaθi* < *(*h*₁)*opi-sisteh*₂-*ti*, with <EI> = [i]). While this would provide a further example of **o* > *a*, it also shows how an uncontroversial solution is hard to reach.

40 *Klohi zis*, which appears at least 6× in *MLM*, is a well-known formula in the Messapic language. Among the extensive bibliography see, e.g., Parlàngeli (1960: 323–324); Schmid (1965: 331–335); Lazzeroni (1982: 163–169); Jasanoff (1987: 105–106); Peters (2006: 329–353); Vine (2017: 774–775; 2019: 382 & 388–389); Matzinger (2019: 52–55). *Klohi* has been interpreted as an optative (Schmid 1965; Lazzeroni 1982), as an aorist subjunctive (Jasanoff 1987), or as an imperative (de Simone 1972: 151 & 1988: 368).

41 The three words on the *MLM 36 Rud* altar (*oššo hazzava θaotori*) seem to suggest that -*a* and -*o* are dative endings.

4.1.7 *venas*

Because of the striking formal similarity with the Latin theonym *Venus*, this form is generally considered its exact cognate, thus showing the change $*o > a$ in the *s*-stem suffix $*-os > -as$.⁴² The identification with *Venus* is supported by the fact that the word, found on five inscriptions (*MLM* 21 Bas, 3 Car, 1 Mo, 5 Ro, & 8 Ro), often appears in asyndetic collocations with the theonyms *Zis* (2×; ~ Gk. Ζέυς) and *Θautour* (2×; alternative spelling for *Θaotor*, see below § 4.4). Even though the identification with Lat. *Venus* is attractive on the basis of the textual evidence, the status of Mess. *venas* as its cognate, i.e., inherited material, is problematic. Etymologically, Lat. *Venus* reflects a neuter *s*-stem noun $*(h_1)wenH-os-$ ‘desire’, in turn derived from a root $*(h_1)wenH-$ ‘be fond of’ which appears to have been fairly productive in terms of verbal and nominal formations (see *EDLIL* 663; *LIV*² 682–683; *LIV*² *Add.* 89). The same *s*-stem formation as Lat. *Venus* is found in Ved. *vānas-* ‘desire’ (*EWA* ii 500–501), which preserves the original neuter gender of the inherited *s*-stem. Ostensibly, Latin has instead reanalyzed the original neuter into a feminine, given the restricted use of the noun as a theonym corresponding to the Greek *female* deity Aphrodite (see, e.g., Puhvel 1987: 151). Within Italic, the only other attestation of the stem, i.e., Osc. DAT.F.SG. $\text{F}\epsilon\nu\zeta\eta\iota$, shows feminine gender as well but, crucially, is taken by some to be a borrowing from Latin (see Untermann 2000: 837 with earlier references and *EDLIL* 663; *contra* Poccetti 2008 and Zair 2016: 107¹¹). Now, formally speaking, we could posit Messapic as a third IE language to have inherited the same *s*-stem formation that Latin and Sanskrit preserve. However, the fact that Mess. *venas* seems to pattern as a theonym and is taken by *MLM* (ii 142) as feminine NOM.SG. implies too striking a convergence with Latin, in terms of morpho-syntactic and semantic shift, to consider it a *bona fide* cognate undergoing the exact same independent innovations. Moreover, if the diffusion hypothesis is preferable to account for Osc. $\text{F}\epsilon\nu\zeta\eta\iota$, one wonders why we should not adopt it to account for Mess. *venas*. If indeed Mess. *venas* equals Lat. *Venus* in cultural terms, the cognacy hypothesis is therefore not compatible with this claim.⁴³ Not only are alternative interpretations possible,⁴⁴ but by following the

42 The first to identify *venas* as the goddess Venus was von Blumenthal (1936: 88) and his interpretation has generally been accepted (*contra*, Vetter 1943: 57 and Orioles 1981: 155). See, e.g., de Simone (1972: 136, 1988: 369–374, & 2018: 1843); Matzinger (2019: 51–52). However, the role of this *venas* in the Messapic world is more uncertain (de Simone 1991: 310; Santoro 1989: 148). For a recent re-assessment of the issue, see Dupraz (2024).

43 Dupraz (2024: esp. 279–280) independently reaches similar conclusions, from different premises.

44 We could in fact take the form as an epithet/epiclesis of *Zis* and *Θaotor* (cf. already Pagliara 1981; see Matzinger 2019: 52 with further literature; differently, see Dupraz 2024: 253),

traditional explanation one would be forced to ignore some of the evidence,⁴⁵ namely the sequence <ZIFENA>, in *MLM* 5 Vi which *could* in principle contain the same lexeme, especially given that it is found in an inscription with other theonyms.

4.2 Rules and exceptions

At a closer look, even the etymology-based *vulgata* admits some exceptions, as [o] seems to surface in some cases, which we discuss in this paragraph. From a synchronic point of view, according to de Simone, the [o] in a number of instances boils down to a combinatory variant of /a/ with fully predictable distribution—thus not as a phoneme in contrast with it. First, de Simone (1972: 137–138 & 142) notes that in some cases [o] is preserved after labial consonants, and gives a series of examples (e.g., *borrahetis*, *mogamatesi*, *morkos*; full list at 142³²). The best-looking example is *kos*, which would come from a **k^wos* comparable to the Sanskrit indefinite/interrogative *kás*. However, de Simone also observes in another note (138²⁶) that the *-pandes* discussed above, if indeed going back to **-pondyos*, would show an unexplained /a/ after /p/. He also claims (1972: 138) that the personal name (or theonym) *logeti-* features a ‘secondary labialization’. Following a hypothesis first put forward by Kretschmer (1923), *logetibas* has been compared with Gk. Λάγεις ‘fate’/‘goddess of fate’ and Sicel Λάγεις, believed to be the ‘Hellenized version’ (*Gräzisierung*) of a local goddess **Lageti*.⁴⁶ Thus, the plural *logetibas* would refer to a set of three goddesses comparable to the Greek *Moirai*. This identification, besides resting solely on late lexicographic material,⁴⁷ is also at odds with what we know about Messapic historical phonology. Crucially for our purposes, if *logetibas* is to be

rather than an independent theonym. From an etymological point of view, one could compare Av. *Vanant-* ‘the conqueror’, a minor deity to whom a Yašt (Yt. 21) is dedicated. Under this analysis, the collocation *zis venas* would correspond to ‘Lord Zeus’/‘Zeus the Conqueror’, thus implying a different trajectory of change from inherited **(h₁)wenH-os-* than Lat. *Venus*. Far from being conclusive, this alternative etymology shows how debatable the traditional interpretation is.

45 Thus, e.g., Dupraz (2024: 248³⁰): ‘elle pose des problèmes morphologiques, car *vena* ne peut pas sans arbitraire être rattaché au thème en **-s-* du nom de *Vénus* ... Nous ne discutons donc pas cette inscription ici.’

46 Later supported, among others, by de Simone (1972: 137–138, also n. 25) and by Beekes (*EDG* 821).

47 See the Hesychian gloss λ 50 Λάγεις· θεός, Σικελοί ‘Lagesis: God(dess). Sicels’ (cf. also Phot., λ 16 Λάγεις· θεός Σικελιακή ‘Lagesis: Sicel goddess’). The association of Λάγεις with Sicels rather than with other indigenous groups of Sicily like the Sicanians is well paralleled in the lexicographers (cf. Poccetti 2012: 66). It should be noted that Λάγεις does not appear in Sicilian epigraphic texts, and it must therefore be handled with caution.

taken as a loanword, one would need to explain why an original /a/ in Greek and Sicel (Λάχ-, Λάγ-) was changed to /o/ in Messapic (*log-*), as we usually observe the opposite pattern (i.e., *o > a*). Moreover, to explain the preservation of *-ti(s)*, we would need to assume that the borrowing happened before the Greek assibilation of **-ti-*.⁴⁸ If, on the other hand, we tried to posit an inherited action noun in **-ti-* from the same root of Gk. λαγχάνω ‘obtain’, we would be faced with an obscure root etymology (see *DELG* 611–612; *EDG* 821). De Simone’s (1972: 137) solution of a ‘secondary labialization’ is not particularly informative, as he does not specify whether he envisages a situation of contact or rather thinks of inherited material. In either case, the presence of /o/ in *logetibas* would lack an overt phonological trigger, since the only comparable phenomenon is attested immediately *after* labials (cf. above).⁴⁹ Another interesting loanword is the form *kokkehias* (date unclear). This word, if borrowed from Lat. *Cocceius*, would be even more problematic for the traditional theory. If the vocalism in the first syllables, Mess. *kok-* and Lat. *Coc-*, suggests that the borrowing took place after the *a-o* merger, the ending in *-as* would seem to reflect the change. Moreover, for it to undergo the merger, we would need to push the date of borrowing back to when Lat. *-us* was still *-os*—a quite unlikely scenario. We must therefore admit some reshaping of the borrowing to match Messapic morphology.

4.3 Reassessing the **o > a* shift: A sound change in progress?

On the basis of the analyses above, there does not seem to be very strong evidence for an across-the-board prehistoric change **o > a* in Messapic, as the material that is certainly inherited is rather meagre. The evidence we are working with mostly consists of loanwords, which is only informative about the synchronic phonology of Messapic and loan adaptation processes. De Simone’s solution envisions a system with one fewer phoneme in the Messapic sound inventory (i.e., with no /o/ *vel sim.*) and still generates some surface examples

48 Alternatively, if *-ti-* was restored in Messapic, it is unclear what acted as an analogical basis for its restoration. Kretschmer (1923: 278) also entertains the possibility that Λάχεσις is a Sicilian loanword both into Greek and into Messapic, but ultimately refutes it.

49 In an attempt to explain the vocalism, Kretschmer (1923: 280) postulates the existence, next to Λάχεσις, of an unattested Arcadian or Aeolic **Λόχεσις* and claims that this alleged form **Λόχεσις* should be explained through an *o*-outcome of syllabic **ŋ* and **ŋ̥*; cf., e.g., δέκο and έκοτόν in Arcadian, δέκοτος and έντος in Lesbian (*contra*, see Buck 1955: 20 & 96–97). However, if the Messapians borrowed this theonym from Aeolic (or Arcadian) Greek speaking populations, then *logetibas* would show preservation of **o*, which could only happen if the loanword took place in a later phase, when the Messapic change **o > a* had ceased to be productive.

of [o]. He first postulates an across-the-board merger of every /o/ with /a/, to then admit a synchronic sound law producing new instances of [o]. From a diachronic point of view, this seems rather anti-economical, and therefore an alternative might be sought. Having observed that the **o > a* change did not take place across the board as universally thought, we suggest that we might be in front of a sound change in progress during the historical period. While it is difficult to establish the specific environments for the retention of **o*, we will see that, at the very least, the labial environment identified by de Simone might not be the only significant one (cf. § 5 below).

4.4 Reflexes of **ō*: **-tor-* and **-on-*stems

Turning to the reflexes of **ō*, as was mentioned above (§ 2.1) de Simone thinks that **ō* is preserved as such in Messapic and is written as ⟨o⟩, ⟨oo⟩, and ⟨oho⟩. This is noted:

- a. In root syllables: e.g., *pido* '(s)he gave' (root aorist) < **pi-dō-t*
- b. In suffixes: e.g., *-tōr-* in *platoras* (GEN.) < **platōr* (cf. Gk. Πλάτωρ); *-ōn-* in *baledonas* (GEN.)

Since most details of Messapic verbal morphology still escape us, and most forms have been identified through the etymological method alone,⁵⁰ and are thus susceptible of different interpretations,⁵¹ we are going to focus on b., which offers enough material for a falsifiable claim. In a brief detour on nominal morphology, we shall first look at **-tor-*stems and then at **-on-*stems.

Under standard Erlangen theory, **-tor-*stems are reconstructed as amphiki-netic because of the suffixal *o*-vocalism, which is peculiar to this morphological class, but one must admit some restructuring of the stress pattern: e.g., NOM.SG. **ġēnh₁-tōr-Ø >(>)* Ved. *jánitā*, Gk. γενέτωρ 'begetter', GEN.SG. **ġnh₁-tr-és/ós >>*

50 In de Simone (2018: 1845–1846), several verbs are given more or less certain IE etymologies: *apistaθi* (3SG.PRS.IND.) < **opi-steh₂-ti*, *berain* (3PL.PRS.OPT.) < **b^her-oi-nt*, *beran* (3PL.PRS.SBJV.) < **b^her-ā-nt*, *hadive* (3SG.PRF.IND.) < **sod-i-v-e*, *hipades* (3SG.AOR.IND.) < **supo-d^hē-s-t*, *klaohi/klohi* (2SG.PRS.IMPV.) < **kleus-*, *(ni)ligaves* (3SG.PRF.IND.) < **leġ^h* (?). Others are left without an etymology: *aupave*, *dareti*, *(in)kermaθi*, *kraapati*, *no* [= Lat. *sum*], *pido* (which, as we have just seen, was interpreted as from a **pi-dō-t* in 1972), *preve*, *stahan*, *stihati*. The form *epeigrave/ipigrave* is interpreted as a loanword (< Gk. ἐπιγράφω) adapted to Messapic morphology (*-āve*). To these, Matzinger (2019: 63–83) adds *atallinda* (?), *deineθi*, *haiθraaθi*, *intaθi*, *invinta/invitati*, *kroseti*, not all of which are surely verbs, and none of which have a certain etymology.

51 A case in point: the very form *pido* (cf. previous note) given by de Simone presupposes the deletion of /t/ in *Auslaut* and a phonotactic constraint similar to the one found, e.g., in Greek or Luwian; excluding uncertain segmentations, this seems to be contradicted by the forms *dazet* (MLM 19 Al, MLM 6 Bas, MLM 3 Car, MLM 16 Gn, MLM 22 Ur, MLM 24 Ur) and *bosat* (MLM 13 Gn); see MLM ii 84 and Matzinger (2019: 34).

Gk. γενέτορος; see, e.g., Ringe (2017: 75) and Weiss (2020: 281). This is why these stems have been identified by some as acrostatic, on the basis of the fixed root stress in Vedic (cf., e.g., Rix 1992: 123; Gotō 2017: 348–349; Neri 2017: 107).

PIE **-tor*-stems are reflected in the Messapic corpus by a total of 77 tokens, out of which we are excluding 18 forms on epigraphic grounds.⁵² 44 tokens out of the 59 deemed reliable belong to one type, i.e., *θaotor-*, a divine and personal name connected by de Simone to a preform **teut-* or to a ppp. of PIE **teuH-* or **teuh₂-*. The remaining **-tor*-stems in Messapic are *plator-* (12×), *kalator-* (2×), and *kantor-* (1×).

This material, as summarized in (4), points to some paradigm levelling of the suffixal vocalism. De Simone (1972: 139) interprets the suffix with a long *ō*, thus assuming a levelling based on the lengthened grade in the NOM.SG. This follows from the aforementioned claim that **ō* merged with **ā* in the prehistory of Messapic and would be notated as ⟨A⟩, never as ⟨O⟩ (or variants). Moreover, this argument finds some support in the spelling alternation ⟨O⟩/⟨OY⟩/⟨Y⟩, whereby the Messapic vowel system supposedly only had one long back vowel.

(4) Messapic *-tor*-stems

NOM.SG.	⟨-TOP⟩/⟨-TOYP⟩/⟨-TYP⟩	cf. <i>θaotor</i> , <i>θautour</i> , <i>platur</i>
ACC.SG.	—————	
GEN.SG.	⟨-TOPΑΣ⟩/⟨-TYPΑΣ⟩	cf. <i>θaotoras</i> , <i>platuras</i>
DAT.SG.	⟨-TOPI⟩/⟨-TOYPI⟩/⟨-TYPI⟩	cf. <i>θaotori</i> , <i>θautouri</i> , <i>θoturi</i>

Such analogical levelling is well paralleled in the Indo-European languages, and it is widely assumed to have happened in particular in Italic (cf. Heidermanns 1996: 135 with n. 3 and reference to von Planta 1892). As is well known, Latin has generalized the lengthened grade to all case forms. As for Sabellic, this is illustrated, e.g., by Umb. *ařfertur* (NOM.SG.), *ařferture* (DAT.SG.), *speture* (DAT.SG.), and even by the non-primary derivative *speturie* (showing **-tōr-yo-*; cf. *platoorrihi*), as well as Osc. *regaturei* among other oblique case forms in Sabellic. One could perhaps entertain the idea that Messapic instead resembles Greek and generalizes the short *o*-grade (outside of the NOM.SG.). This would however call for a different understanding of the *a-o* merger in Messapic, which in this scenario would still be in progress at the time of the attested inscrip-

52 We have excluded from our count all forms whose reading is uncertain and whose epigraphic support is not available for autopsy. For instance, the sequence *ehestorihi* on MLM 3 Car, is not singled out as a word.

tions, rather than having taken place in the prehistory of Messapic. Moreover, one would need to explain the ⟨Y⟩ and ⟨OY⟩ spellings in the GEN.SG. and DAT.SG.

As for the PIE **-on-*stems, these also seem to have displayed an amphikinetic inflection (see, e.g., Gotō 2017: 350). However, it is somewhat difficult to tell apart ‘plain’ **-on-*stems from ‘complex’ **-mon-* or ‘Hoffmann’ **-Hon-* stems, which seem to have inflected according to various ablaut templates.⁵³ Moreover, a great number of reconstructible **-on-*stems seem to be non-primary derivatives, as per (5).

(5) Examples of **-on-*stems

- a. PIE **g^wreh₂w-on-* ‘pressing stone, millstone’ > Ved. *grāv-an-*, Eng. *quern*, OIr. *brau*, *broon* (GEN.SG.), Lith. *gìrnos* (PL.); derived from the adj. **g^wrh₂-u-* ‘heavy’ > Ved. *gurú-*, Gk. βάρύς, etc.
- b. PIE **tetk-on-* ‘builder’ > Ved. *takšan-*, Av. *tašan-*, Gk. τέκτων, built on a reduplicated stem **te-tk-*
- c. The Latin ‘*Catō*-type’ (examples from Weiss 2020: 329):
 - i. *Catō*, *Catōnis* ‘sharpy’ ← *catus* ‘sharp’
 - ii. *Varrō*, *Varrōnis* ← *uārus* ‘bow-legged’
 - iii. *mandūcō* ‘glutton’ (Pompon., Apul.) ← *mandūcus* (name of masked character with champing jaws in Atellan farce) < ‘chewing’ ← *mandere* ‘to chew’

The status of the **-on-*stems, therefore, is relatively complicated from a comparative standpoint. Moreover, if Messapic underwent the same type of analogical levelling observed for the **-tor-*stems (cf. above)—as per de Simone—such an extension of a lengthened grade **-ōn-* would not find a parallel in Latin: in the most common reflex of simple **-on-*stems, the ‘*Catō*-type’ (5c), the long **ō* results from the contraction of **-o-on-*.

Some complications also arise when we consider spelling variation in Messapic. On the basis of 17 reliable tokens out of 27 in our dataset, distributed among the types given in (6), a spelling ⟨OY⟩ is the only one found in our corpus for the NOM. SG., but ⟨O⟩ is consistently used in the other case-forms—see (7). The only positive piece of evidence that de Simone puts forward for a levelling based on the lengthened grade **ō* is now limited to the form *dazohonnihi* (*MLM* 1 Bas), but this inscription is not very reliable (cf. n. 22 above).

53 For instance, Neri (2017: 135) mentions both amphikinetic and acrostic inflection for these stems.

(6) Messapic *-on*-stem types

<i>baledon-</i> (1×)	<i>kelon-</i> (1×)	<i>megon-</i> (1×, Grotta Poesia)
<i>bohon-</i> (1×)	<i>kraotedon-</i> (1×)	<i>molon-</i> (1×)
<i>dazihon-</i> (3×)	<i>lahon-</i> (2×)	<i>šaillon-</i> (2×)
<i>eredon-</i> (1×)	<i>laparedon</i> (1×)	<i>vastihon-</i> (1×)
<i>iθaredon-</i> (1×)		

(7) Messapic *-on*-stem inflection

NOM.SG.	⟨-OYN⟩	cf. <i>dazoun</i>
NOM.PL.?	⟨-ONEΣ⟩?	cf. <i>megones</i> ?
GEN.SG.	⟨-ONΑΣ⟩	cf. <i>kraotedonas</i>
DAT.SG.	⟨-ONI⟩	cf. <i>dazihoni</i>

4.5 *The a-o merger, its phonological implications, and the role of diphthongs*

We may therefore entertain the hypothesis that Messapic generalized the short *o*-grade (outside of the NOM.SG.) in the **-on*-stems, as no alternation that is comparable to that of the **-tor*-stems seems to be visible. This independently calls for a different understanding of the *a-o* merger, which may still be in-progress at the time of the attested inscriptions, rather than having taken place in the prehistory of Messapic.

In several cases, the spelling of inherited **ō* coexists with the spelling of both diphthongs ⟨AO⟩, ⟨AY⟩, and their monophthongized version ⟨O⟩. De Simone describes this change as a merger of inherited **eu* and **ou* to **ou*, then changed to *au* because of the *a-o* merger. This stage of the change corresponds to the spellings ⟨AO⟩ or ⟨AY⟩ in Messapic. Finally, monophthongization to *ō* took place, and the new vowel is represented as ⟨O⟩. By comparing the notations of inherited and secondary /o:/, we can identify some diagnostic differences that may clarify the phonetic and phonological value of these graphemes. This shows that when the alphabet was adopted the monophthongization process was not complete, as we can still observe the change in progress in our documents.

We have seen that the notation of inherited **ō* in the context of **-tor*-stems varies significantly: the suffix is noted as ⟨-TOP⟩, ⟨-TOYP⟩, or ⟨-TYP⟩ (and also ⟨-TOOP-⟩ in the **yo*-derivative [p]latoor[r]ih[i]). On the other hand, it is interesting to note that the long /o:/ yielded by monophthongization is only written with ⟨O⟩, and never with ⟨Y⟩ or ⟨OY⟩ (i.e., we never find forms such as †θoutor or †θutor). At the same time, our dataset shows that ⟨OY⟩ for what is ⟨O⟩ elsewhere (in the same lexical item) coexists with ⟨AY⟩ where elsewhere we find ⟨O⟩ (in the same lexical item). One can say that, if the forms θautour and θautouri are reliable, then ⟨OY⟩ might be another way of spelling long /o:/.

It is therefore possible to hypothesize, at this stage, not only that these were historically distinct vowels, but that they also had different phonetic qualities, which in turn caused the different spellings to arise. A remarkable example is in the form *ſoturi*, which, synchronically, displays two different back vowels (otherwise we would have expected <ΘΟΤΟΡΙ>, which we actually find in *MLM* 21 Ro and *MLM* 27 Rud, with the trident sign, or ^{**}<ΘΥΤΥΡΙ>).

4.6 Reflexes of *u

Since spelling variation and phonological changes seem to have greatly affected the Messapic back vowels, it is necessary to ascertain the outcomes of PIE *u(:). According to the *vulgata*, once again based on de Simone (1972: 139–141), PIE short *u is preserved in Messapic, but realized very open and therefore spelled with the sign <ο>, rather than with <Υ>, because the sound was closer to the Tarentine /o/ than to the Tarentine /u/. This of course implies that a lowering has taken place, although de Simone does not explicitly state it. However, revisiting his own argument a few years later, de Simone (1988: 339–351) argued that there is evidence for the neutralization of the opposition between /i/ and /u/ in Messapic, with a surface realization as [y]. This seems somewhat comparable to the Latin *sonus medius* (cf. *optimus/optumus*, etc.). The examination of the inscriptions from Grotta Poesia, together with the evidence from the rest of the Messapic-speaking area, shows a number of alternations such as:

- (8) a. *hipades* ~ *hopades* < *supo-
 b. *hipaka* | (*h*)*ipaka* ~ *hopaka* | *hopakoa*, also probably < *supo⁻⁵⁴
 c. *damatira* ~ *damatura* ~ [d]amatra⁵⁵
 d. *vinaihi* ~ *vnaihi*-

In a few cases, these alternating vowels seem to be relevant for the reflex of PIE *u, as in (8a), while in others it is impossible to establish for certain whether these forms constitute the Messapic outcome of PIE *u.⁵⁶ This evidence is best explained by postulating a five-vowel system for the short series, with two back vowels. Establishing the phonetic realization of these back sounds is less

54 One should note that the etymology of *hipades* and *hipaka* rests on the validity of assumed change *s > h / #_V; cf. § 4.1.4.

55 The evidence for *Damatura* and *vnaihi* comes respectively Ruvo and from Grotta Poesia and shows the north-Apulian conventions of using <Υ>. The puzzling form *Damatira* may be explained by postulating anaptyxis following borrowing from Doric.

56 As for the remaining forms (see *Appendix*), they either have an uncertain etymology, or can be traced back to words without a PIE *u.

straightforward, and Matzinger (2019: 27), who also bases his interpretations on the kind of alternations presented in (8), prefers a realization [ɥ] rather than [y].

4.7 Reflexes of *ū

Turning to the reflexes of PIE *ū, de Simone (1972: 148) tentatively hypothesizes that PIE *ū yielded Mess. *ū*, though admitting that we do not have any clear examples in the epigraphic record. However, as already noted by Krahe (1928: 47), ancient lexicographers might provide us with a path: in Hesychius (β 1318 L.-Cunn.) we read βύριον ὄκημα (βύριον: 'house'), and the same word is attributed to Messapic by the *Etymologicum magnum* (Lasserre & Livadaras 1976),⁵⁷ providing a good parallel for OHG and OE *būr* 'hut' and thus assigning an Indo-European etymology to βύριον. This interpretation is generally accepted, and although Matzinger (2019: 27) does not explicitly take a stance on the issue, he seems to follow the *vulgata* in hypothesizing that *ū remained as such in Messapic.

Despite the fragility of some of the etymologies listed and the limited evidence, it is probable that Messapic preserves at least some reflexes of *u(:). If we accept de Simone's proposals for *u, according to which Messapic displays a lowering and a *sonus medius*, we must postulate that PIE *u had two reflexes. Nevertheless, at this stage we are left with several questions relative to the chronology of these sound changes (whether or not we can assume that lowering preceded fronting or determine which environments are concerned; see §5 below). Furthermore, previous scholars generally assumed that these changes have taken place mostly in the prehistory of Messapic, but the evidence suggests that some of these changes are still in progress at the time of our evidence. In the next section, we shall address all these questions, and others, integrating the evidence and theories that we have outlined so far.

5 Models for Messapic vocalism

On the basis of the empirical considerations offered above, we will now evaluate how compatible the current models for the Messapic vowel system are

57 However, some details reveal that the parallel, though intriguing, is not without problems: where it is said that the Messapians call home βάριον, the word is preserved both as εὐβύριον (compared to εὐοικον) and βάριον (*Etym. magn.* 389,25). It should be noted that the variant βάρυ- 'home' (βαυριόθεν) appears also in the elegy of Cleon Curialis (*SH* 340 l. 2), who is not connected with Messapic.

with the available data. Starting from the reflexes of $*\bar{o}$, which seem to be rather uncontroversial, we outline two possible scenarios, each with its own difficulties. The crucial point of our argumentation will be that monophthongization (cf. § 4.5) is key to the understanding of the long vowel system.

In one possible scenario (i), PIE $*\bar{o} >$ Mess. /u:/, written with ⟨o⟩—and this would be Matzinger's assumption (cf. § 2.1); then the monophthongization of /au/ creates a new /o:/. The need then arises to graphically distinguish this new /o:/ from /u:/, written ⟨o⟩, and therefore ⟨γ⟩ is employed to write the original /u:/ from PIE $*\bar{o}$. For instance: ⟨ΠΛΑΤΥΡ⟩ = /platu:r/ and ⟨ΘΟΤΟΡ⟩ = /t^(h/j)o:tur/ (cf. ⟨ΘΟΤΥΡΙ⟩ = /t^(h/j)o:turi/). However, this would encounter some problems. First of all, we have to postulate a prehistoric raising because historically the vowel has to be a long /o:/. If then the change $*\bar{o} >$ /u:/ took place before the written record (i.e., the adoption of the alphabet), we would expect that Messapic speakers would have ⟨γ⟩ instead. Also, if one needs to graphically distinguish these new /o:/ from /u:/, one wonders why the *old* convention (/o:/ = ⟨o⟩) is changed, that is why the phonetic value of a given grapheme is reassigned.

In an alternative scenario (ii), PIE $*\bar{o} >$ \bar{o} , but by the fifth century BCE $\bar{o} >$ \bar{u} (as per de Simone, cf. § 2.1), and this further change was a consequence of an unbalance in the system due to the asymmetry between long and short vowels. In this scenario, the monophthongization of /au/ (as per above) does not seem to have any impact on the overall system, but—*contra* de Simone⁵⁸—we believe that the introduction of a new (monophthongized) /o:/ = [ɔ:], which is more open than the original /o:/ = [o:], *did* have some consequences. Either the old /o:/ started to be written as ⟨γ⟩ (with variations), pointing to an orthographic change, or the new /o:/ (= [ɔ:]) caused a raising of the old /o:/ to /u:/ via a push chain. The implication is that we would have an inherited (PIE) /o:/ and a secondary /o:/ which is the outcome of monophthongization of PIE $*ou$ and $*eu$.

Let us now move to more uncertain terrain, that is to say the fate of $*o$. As we have seen, the *a-o* merger in Messapic meets with some empirical problems, as it rests on somewhat unconvincing etymologies, and spelling variation does not always align with the sound change. Moreover, de Simone (1972: 137–138) himself admitted that in at least one environment, i.e., after labial consonants, the merger was not complete, and /o/ was preserved—the best example being

58 Cf. de Simone (1988: 348): 'Un altro fenomeno chiaramente esterno è l'impiego di ⟨ou⟩ per indicare / \bar{u} / (*ḡautour*, *ḡautouri*), corrente in ambiente ellenistico, che non può essere spiegato come fatto messapico (grafia inversa), perché non esiste in questa lingua un dittongo /ou/ possibilmente monottonghizzato (> / \bar{u} /)'.

kos < **k^wos* (§ 4.2). However, he also explained some other instances of synchronic [o] as secondary changes *a* > *o*, which is both theoretically and empirically unsatisfactory.

Our tentative suggestion, which builds on de Simone's, is that **o*, rather than being prehistorically changed to *a* across the board and sporadically reverted, was retained as such in more environments than initially thought—i.e., not exclusively after labial consonants. Some exploratory soundings of the corpus⁵⁹ seem to confirm this hypothesis,⁶⁰ and also reveal the curious fact that, despite the obvious abundance of *a*, the mid vowels *e* and *o* have a comparable number of tokens.⁶¹ Although certain tendencies might be identified, it is for the time being impossible to draw firm conclusions from a purely quantitative study. One should also acknowledge that variation seems to be somewhat less present in functional morphemes such as endings (with the exception of **-os* in a few environments; cf. § 4.1.3), which does not rule out that morphology might be involved in the diffusion of this sound change.

Be that as it may, a hypothesis that better accommodates the underapplication of the *o* > *a* rule is that the sound change was still in progress by the time our records were produced instead of belonging to the prehistory of Messapic. It follows that, at the moment of the adaptation of the Greek alphabet, a synchronic /*o*/ was available in the phonemic inventory of Messapic. We may then

59 We assembled a database from the lexicon in *MLM* ii. Excluding sequences of less than four letters in incomplete contexts, whatever is marked with (?) by *MLM*, duplicate forms, and deleting restorations that are not absolutely obvious, we gathered 863 items. This approach has some obvious limitations, but a sophisticated statistical study would go beyond the scope of this paper and be in any case hindered by the nature of the material.

60 E.g., the fact that 25.59% of all <O>s in /oC/ sequences precede <P> (vs. only 10.9% of all <A>s in /aC/ sequences) and that 81 out of 203 /r/ sequences (39.4%) are <OP> might point to a possible retention of /*o*/ before /*r*/. Among these words, besides a probable reflex of a *u*-stem, *morkos*, we observe the presence of *-tor*- stems (e.g., *kalator*, *kantorrihi*, *haštorryhi*) including the orthographic variations of the divine names *Θaotor* and of *Plator*. Additionally, the dataset comprises nouns with less well-established or uncertain etymologies (e.g., **andoraion*, potentially also attested as **andirah*(*h*)*o*, **balaoriaihi*, and **šoborihi*) alongside terms that likely represent loanwords (cf. *argorian*, § 4.1.2). Similarly, the fact that 161 out of 360 <O>s (44.7%) are found before <P>/<A>/<M>/<N> (vs. 271 out of 1009 <A>s, 26.85% of the total) might suggest even more environments for the preservation of /*o*/.

61 Our Python-aided count returned 2563 vowel tokens (excluding those in the sequences: <AO>, <EO>, <OHO>, <OO>, <OY>, and <AA>) for our corpus, with the following figures: <A>: 1009 occurrences, 39.37% of total; <E>: 350 occurrences, 13.66% of vowels; <I>: 816 occurrences, 31.84% of total; <O>: 360 occurrences, 14.05% of total; <Y>: 22 occurrences, 0.86% of total; <H> (= *ē*): 6 occurrences, 0.23% of total.

wonder why, in the process of script adaptation, a distinction between /o/ and /u/ was not reflected in the sign inventory and only ⟨o⟩ was selected.

Our proposed solution is that /u/ was lowered in the prehistory of Messapic so that the phonetic distinction between the reflexes of *o and those of *u was not marked enough to justify the adoption of two separate signs. What we are positing is not a complete o-u merger, but rather the existence of two *phonetically* distinct o-sounds, namely [o] < *u and [ɔ] < *o. This accommodates the fact that only the reflexes of *u appear to undergo the *sonus medius* treatment discussed above. Furthermore, hypothesizing the existence of two o-sounds provides us with a reasonable trigger for the change o > a, which is descriptively a lowering.

On the basis of these premises, one can set up the following relative chronology:

- i. Before alphabet adoption: *u > o; *o and *ō are unchanged
- ii. In-progress at the time of the extant records (stage I): *o > a in some environments; monophthongization of ⟨AO⟩ (< *ou < *eu, *ou) to [ɔ:]
- iii. In-progress at the time of the extant records (stage II): new [ɔ:] causes raising of old *ō to [u:] (*alternatively: just an orthographic change*)

The changes we described proceed in two opposite directions, as we observe the lowerings *u > o and (*o) > a in the short-vowel series vis-à-vis the raising $\bar{o} > \bar{u}$ in the long-vowel series. Far from constituting a problem, our relative chronology qualifies the raising $\bar{o} > \bar{u}$ as a late development, i.e., not taking place at the same time as the lowerings, but all changes we posit have a specific phonetic motivation and conform to the Principles I and II of chain shifts outlined in Labov (1994: 116), by which long vowels rise and short vowels fall.⁶²

This scenario could potentially find a typological parallel within the Indo-European family, namely in Hittite. This is also a language in which PIE *(\bar{o}) had lowered to /a(:)/ in most—though not all—environments. For a long time, it was thought that Hittite had no synchronic /o(:)/. This was partly due to its writing system, which did not have a series of signs containing the vowel /o(:)/, and partly to our incomplete understanding of its historical phonology. However, in the last 15–20 years, it has become reasonably clear that not all PIE *o and *ō became /a(:)/ (cf. Melchert 2020: 263). Hittite also shows monophthongizations⁶³ and prehistoric and historic (*u) > o lowerings and raisings that seem to be comparable to what happens in Messapic. Interestingly, *eu, *ou,

62 We wish to thank an anonymous reviewer for reminding us of the pertinence of Labov's principles to our case.

63 For a detailed treatment of Hittite diphthongs, cf. Kimball (1994) and especially *EDHIL* 51–59.

and **u* do not yield /o(:)/ in one environment (*C_N*) where **o* does not seem to be merged with /a/, which seems to be indicative of a push chain.

6 Conclusions

In light of all of what we have seen, we can conclude by offering the following considerations. From a synchronic perspective, we have shown that the vowel alternations we see in our corpus can help us understand more about Messapic phonology, and that it is important to distinguish between loanwords and inherited material.

From the diachronic point of view, we have challenged the commonly-held view of a complete prehistoric **o > a* change, and have entertained the possibility that this sound change might still be in progress; we also took into account monophthongization phenomena, which we suggest might have caused a push chain, or at least an orthographic shift, in the long back vowels. This hypothesis is indeed attractive, as it best explains the spelling variation and accommodates more elegantly the two models proposed in the literature (de Simone's and Matzinger's). One could agree with Matzinger in positing an /o:/ > /u:/ change in the history of Messapic, but we would suggest a different chronology and scope: what was an across-the-board change that pre-dated our extant documents is now a change limited to inherited **ō* and triggered by the monophthongization of <AO> (< **ou* < **eu*, **ou*) that appears to be in progress in our evidence. If this is the case, one could maintain with de Simone that PIE **ō* remained as such in early Messapic (though with a later chronology for the change to *ū*), and that the monophthongization of <AO> produced *ō*. Our proposal tries to unify these theories and, taking them a step further, to clarify the phonetic value of the grapheme <O>. Suggesting that, potentially, <O> could note both [ɔ:] and [u:] would also seem to be coherent with the idea that Messapic borrowed <O> for its one and only short back vowel /o/ = [ɔ].

From the point of view of typology, while at the synchronic level all models are potentially viable, we have observed that our suggestion may find an interesting parallel in Hittite, which has similar developments in the back vowels and comparable monophthongization phenomena. Our analysis also has some implications for historical morphology, as we have argued that while one can identify an *ō*-grade analogical levelling in the **-tor*-stems, the same cannot be done for the **-on*-stems. On several occasions we have highlighted the relevance of morphological environments for different outcomes of sound change, but we have refrained from proposing an extensive morphologically sensitive analysis, which would be appropriate for a separate paper.

Another issue that remains to be settled is how the front vowels fit in the overall picture; both a traditional analysis and a preliminary quantitative survey have shown that /e/ has distributional patterns that look similar to those of /o/. The distribution and developments of front vowels in Messapic have not been addressed as thoroughly as those of back vowels, and one can only hope that future findings and further research will help fill this gap.

References

- Allen, W. Sidney. 1987. *Vox Graeca. A guide to the pronunciation of classical Greek*³. Cambridge: Cambridge University Press.
- D'Andria, Francesco, & Mario Lombardo (eds.). 1999. *I Greci in Terra d'Otranto*. Galatina: Congedo.
- Baldi, Philip. 2002. *The foundations of Latin*². Berlin: De Gruyter.
- von Blumenthal, Albrecht. 1936. Zur Interpretation der messapischen Inschriften. *Indogermanische Forschungen* 54: 81–113.
- Buck, Carl D. 1955. *The Greek dialects. Grammar, selected inscriptions, glossary*. Chicago: University of Chicago Press.
- Ciceri, Marco. 2012. Il genitivo messapico in *-ihi*. *Annali della Facoltà di Lettere e Filosofia dell'Università degli Studi di Milano* 65: 71–102.
- Ciceri, Marco. 2013. A proposito di messapico <s(s)>/<š(š)>. *Studi e saggi linguistici* 51: 129–173.
- Cunningham, Ian C. 2018. *Hesychii Alexandrini Lexicon editionem post Kurt Latte continuans recensuit et emendavit i, A–Delta*. Berlin: De Gruyter.
- Cunningham, Ian C. 2020. *Hesychii Alexandrini Lexicon editionem post Kurt Latte continuans recensuit et emendavit ii, [1] Epsilon–Iota. 2) Kappa–Omicron*. Berlin: De Gruyter.
- DELG = Chantraine, Pierre. 1999. *Dictionnaire étymologique de la langue grecque. Histoire des mots*. Paris: Klincksieck.
- Dunkel, George E. 2014. *Lexikon der indogermanischen Partikeln und Pronominalstämme*. Heidelberg: Winter.
- Dupraz, Emmanuel. 2022. Zu einigen möglichen Prohibitivformeln im Messapischen. *Historische Sprachforschung* 135: 163–181.
- Dupraz, Emmanuel. 2024. Vénus messapienne, Vénus osque, Vénus campanienne. Deux propositions. In *Petit dieux des romains et leurs voisins. Enquête comparatiste sur les hiérarchies divines dans les cultures romaines, italiques et grecques*, ed. Francesca Prescendi & Françoise Van Haepere, 241–287. Turnhout: Brepols.
- EDG = Beekes, Robert S.P. 2010. *Etymological dictionary of Greek*. Leiden: Brill.
- EDHIL = Kloekhorst, Alwin. 2008. *Etymological dictionary of the Hittite inherited lexicon*. Leiden: Brill.

- EDLIL = de Vaan, Michiel. 2008. *Etymological dictionary of Latin and the other Italic languages*. Leiden: Brill.
- EWA = Mayrhofer, Manfred. 1986–. *Etymologisches Wörterbuch des Altindoarischen*. Heidelberg: Winter.
- Fortson, Benjamin W. 2010. *Indo-European language and culture. An introduction*². Malden (MA): Wiley-Blackwell.
- Giacomelli, Roberto. 1979. *I grecismi del messapico*. Brescia: Paideia Editrice.
- Gotō, Toshifumi. 2017. The morphology of Indic (old Indo-Aryan). In *The handbook of comparative and historical Indo-European linguistics*, ed. Matthias A. Fritz, Jared Klein, Brian Joseph, & Mark Wenthe, 344–377. Berlin: De Gruyter.
- Gusmani, Roberto. 2006. Ancora sul genitivo messapico in -(A)IHI. In *Studi di antichità linguistiche in memoria di Ciro Santoro*, ed. Maria Teresa Laporta, 199–205. Bari: Cacucci Editore.
- Heidermanns, Frank. 1996. *Sabellische Nominalbildung. Untersuchungen zur nominalen Wortbildung im Oskisch-Umbrischen*. Habilitationsschrift. Köln: Universität zu Köln.
- Ilocri = Del Monaco, Lavinio. 2013. *Iscrizioni greche d'Italia. Locri*. Roma: Quasar edizioni.
- IRhegion = D'Amore, Lucia. 2007. *Iscrizioni greche d'Italia. Reggio Calabria*. Roma: Unione accademica nazionale.
- ISic = Prag, Jonathan R.W. (ed.). 2017–2024. *I.Sicily: Inscriptions of Ancient Sicily*. Available at: <http://sicily.classics.ox.ac.uk>.
- Jasanoff, Jay. 1987. Some irregular imperatives in Tocharian. In *Studies in memory of Warren Cowgill*, ed. Calvert Watkins, 92–112. Berlin: De Gruyter.
- Jeffery, Lilian H. 1990. *The local scripts of archaic Greece. A study of the origin of the Greek alphabet and its development from the eight to the fifth centuries B.C.*² Revised edition with supplement by A.W. Johnston. Oxford: Clarendon Press.
- Kimball, Sara E. 1994. The IE short diphthongs *oi, *ai, *ou and *au in Hittite. *Die Sprache* 36: 1–28.
- Kirchhoff, Adolf. 1887. *Studien zur Geschichte des Griechischen Alphabets*⁴. Berlin: Gütersloh/C. Bertelsmann.
- Krahe, Hans. 1928. Sprachliche Untersuchungen zu den messapischen Inschriften. *Glotta* 17: 81–104.
- Krahe, Hans. 1932. Illyrisches. *Glotta* 20: 188–196.
- Krahe, Hans. 1955. *Die Sprache der Illyrier 1, Die Quellen*. Wiesbaden: Otto Harrassowitz.
- Kretschmer, Paul. 1923. Messapische Göttinnen. *Glotta* 12: 278–283.
- Labov, William. 1994. *Principles of linguistic change 1, Internal factors*. Oxford/Cambridge (MA): Wiley-Blackwell.
- Lasserre, François, & Nicolaus Livadaras. 1976. *Etymologicum magnum genuinum. Symeonis etymologicum una cum magna grammatica; Etymologicum magnum auctum*. Roma: Edizioni dell'Ateneo.

- Lazzeroni, Romano. 1982. Messap. *kla(o)hi* = sscr. *śroši*. Un εἶδωλον della comparazione. *Studi e saggi linguistici* 22: 163–169.
- Léger Pirus, Viviane. 2024. La palatalisation des consonnes devant *-y- en messapien: étude de la diversité des graphies. *Les Études Classiques* 92: 51–72.
- Lejeune, Michel. 1991. Sur la translittération du messapien. *AION. Annali del Dipartimento di Studi del Mondo Classico e del Mediterraneo Antico. Sezione linguistica* 13: 211–231.
- LIV² = *Lexicon der indogermanischen Verben. Die Wurzeln und ihre Primärstambildungen*², bearbeitet von Martin Kümmel & Helmut Rix. Wiesbaden: Reichert, 2001.
- LIV² Add. = Kümmel, Martin J. 2011–2023. *Addenda und corrigenda zu LIV²*. Consulted on-line at <https://www.gw.uni-jena.de/phifakmedia/fakultaet/einrichtungen/institute/institut-fuer-orientalistik/indogermanistik/dateien/mitarbeiter/martin-kuemmel/publikationen/kuemmel-liv2-add.pdf>. Latest update: 3 February 2024.
- Lomas, Kathryn. 2021. The Greek alphabet in south-east Italy. Literacy and culture of writing between Greeks and non-Greeks. In *The early Greek alphabets. Origin, diffusion, uses*, ed. Robert Parker & Philippa M. Steele, 320–347. Oxford: Oxford University Press.
- Lombardo, Mario. 1992. *I Messapi e la Messapia nelle fonti letterarie greche e latine*. Galatina: Congedo.
- Lombardo, Mario. 2015. I Messapi. Origini, vicende, e civiltà. In *Soletto ritrovata. Ricerche archeologiche e linguaggi digitali per la fruizione*, ed. Maria Teresa Giannotta, Francesco Gabellone, Maria Federica Stifani, & Lavinia Donateo, 9–26. Galatina: Editrice Salentina.
- LSJ⁹ = Henry G. Liddell, Robert Scott, Henry S. Jones, & Roderick McKenzie. 1940 [+ Suppl. 1968 & 1996]. *A Greek-English lexicon*⁹. Oxford: Oxford University Press.
- Lundquist, Jesse, & Anthony D. Yates D. 2018. The morphology of Proto-Indo-European. In *The handbook of comparative and historical Indo-European linguistics*, ed. Matthias A. Fritz, Jared Klein, Brian Joseph, & Mark Wenthe, 2079–2195. Berlin: De Gruyter.
- Marchesini, Simona. 1999. Confini e frontiera nella grecità d'Occidente. La situazione alfabetica. In *Confini e frontiere nella grecità d'Occidente. Atti del trentasettesimo Convegno di Studi sulla Magna Grecia, Taranto, 3–6 ottobre 1997*, ed. Attilio Stazio & Stefania Ceccoli, 173–212. Napoli: Arte Tipografica.
- Marchesini, Simona. 2013. Quali lingue, quali popoli nell'Apulia di v e iv secolo a.C. In *La comunicazione verbale tra Greci e indigeni in Apulia nel v–iv secolo a.C.: quali elementi? Atti del Seminario di Studi linguistici, archeologici e storici (Bari, Università degli Studi 'Aldo Moro', 30 Ottobre 2012)*, ed. Luigi Todisco, 19–33. Napoli: Lofredo.
- Marchesini, Simona. 2020. Messapico. *Paleohispanica* 20: 495–530.

- Matzinger, Joachim. 2019. *Messapisch*. Wiesbaden: Reichert.
- Melchert, H. Craig. 2020. Hittite historical phonology after 100 years (and after 20 years). In *Hrozny' and Hittite. The first hundred years. Proceedings of the international conference held at Charles University, Prague, 11–14 November 2015*, ed. Ronald I. Kim, Jana Mynářová, & Peter Pavúk, 258–276. Leiden: Brill.
- Méndez Dosuna, Julián. 2021. The pronunciation of upsilon and related matters. A u-turn. In *The early Greek alphabets. Origins, diffusion, uses*, ed. Robert Parker & Philippa M. Steele, 119–145. Oxford: Oxford University Press.
- MLM = de Simone, Carlo, & Simona Marchesini. 1992. *Monumenta linguae Messapicae*. Wiesbaden: Reichert.
- Neri, Sergio. 2017. *Elementi di morfologia flessiva nominale indoeuropea*. Perugia: Università degli Studi di Perugia.
- Orioles, Vincenzo. 1972. Su alcuni fenomeni di palatalizzazione e di assibilazione nelle lingue dell'Italia antica. *Studi linguistici salentini* 5: 69–100.
- Orioles, Vincenzo. 1981. Il messapico. In *Nuovi materiali per la ricerca indoeuropeistica*, ed. Enrico Campanile, 139–160. Pisa: Giardini.
- Orioles, Vincenzo. 1991. Il messapico nel quadro indoeuropeo. Tra innovazione e conservazione. In *Rapporti linguistici e culturali tra i popoli dell'Italia antica. Pisa, 6–7 ott. 1989*, ed. Enrico Campanile, 157–175. Pisa: Giardini.
- Pagliara, Cosimo. 1981. Materiali epigrafici da Vaste e Muro. *Studi di antichità. Lecce* 2: 207–220.
- Parlangèli, Oronzo. 1960. *Studi messapici (iscrizioni, lessico, glosse e indici)*. Milano: Istituto Lombardo di Scienze e Lettere.
- Parlangèli, Oronzo. 1965. Nuove iscrizioni messapiche. *Indogermanische Forschungen* 70: 172–190.
- PCG i = Kassel, Rudolf, & Colin Austin. 2001. *Poetae comici Graeci i, Comoedia dorica. Mimi. Phylaces*. Berlin: De Gruyter.
- Peters, Martin. 2006. Sur morphologischen Einordnung von messapisch *klaohi*. In *Studi di antichità linguistiche in memoria di Ciro Santoro*, ed. Maria Teresa Laporta, 329–353. Bari: Cacucci Editore.
- Pisani, Vittore. 1964. *Le lingue dell'Italia antica oltre il latino²*. Torino: Rosenberg & Selier.
- von Planta, Robert. 1892. *Grammatik der oskisch-umbrischen Dialekte i, Einleitung und Lautlehre*. Strassburg: K.J. Trübner.
- Pocetti, Paolo. 2008. Notes de linguistique italique. Nouvelle série. 5. Une nouvelle inscription osque et le nom de *Venus*. *Revue des études latines* 86: 24–36.
- Pocetti, Paolo. 2012. Language relations in Sicily. Evidence for the speech of the Σικελοί, the Σικωνοί and others. In *Languages and linguistic contact in ancient Sicily*, ed. Olga Tribulato, 49–94. Cambridge: Cambridge University Press.
- Preller, Ludwig. 1964. *Polemonis Periegetae Fragmenta collegit, digessit, notis auxit*

- L. Preller; accedunt de Polemonis vita et scriptis et de historia atque arte periegetarum commentationes*². Amsterdam: Adolf M. Hakkert.
- Prosdocimi, Aldo Luigi. 2006. Il genitivo messapico in *-ihi*. In *Studi linguistici in onore di R. Gusmani*, ed. Raffaella Bombi, Guido Cifoletti, Fabiana Fusco, Lucia Innocente, & Vincenzo Orioles, 1421–1432. Alessandria: Edizioni dell'Orso.
- Puhvel, Jaan. 1987. *Comparative mythology*. Baltimore: Johns Hopkins University Press.
- Rau, Jeremy. 2009. *Indo-European nominal morphology. The decads and the Caland system*. Innsbruck: Institut für Sprachen und Literaturen der Universität Innsbruck.
- Ringe, Don. 2017. *From Proto-Indo-European to Proto-Germanic*². Oxford: Oxford University Press.
- Rix, Helmut. 1992. *Historische Grammatik des Griechischen. Laut- und Formenlehre*². Darmstadt: Wissenschaftliche Buchgesellschaft.
- Santoro, Ciro. 1989. Il lessico del 'divino' e della religione messapica. In *Atti del IX Convegno dei comuni messapici, peuceti e dauni (Oria, 24–25 novembre 1984)*, 139–180. Bari: Editrice Tipografica.
- Schmid, Wolfgang P. 1965. Messapisch *klaohi, klohi*. *Studi salentini* 14: 331–335.
- SH* = Lloyd-Jones, Hugh, Peter J. Parsons, Heinz-Günther Nesselrath, & John U. Powell. 1983. *Supplementum Hellenisticum. Texte und Kommentare*. Berlin: De Gruyter.
- de Simone, Carlo. 1962. Die messapische Sprache seit 1939. *Kratylos* 7: 113–125.
- de Simone, Carlo. 1972. La lingua messapica. Tentativo di una sintesi. In *Le genti non greche della Magna Grecia. Atti dell' XI Convegno di Studi sulla Magna Grecia. Taranto, 10–15 ott. 1971*, ed. Attilio Stazio, 125–201. Napoli: Arte Tipografica.
- de Simone, Carlo. 1979. Il Messapico. In *Le iscrizioni pre-latine in Italia (Roma, 14–15 marzo 1977)*, 171–182. Roma: Accademia Nazionale dei Lincei.
- de Simone, Carlo. 1980. Italien. In *Die Sprachen im römischen Reich der Kaiserzeit. Kolloquium vom bis 10. April 1974*, ed. Günter Neumann & Jürgen Untermann, 65–81. Köln: Rheinland Verlag / Bonn: Habelt Verlag.
- de Simone, Carlo. 1983. L'evidenza messapica, tra grafematica e fonologia. *AION. Annali del Dipartimento di Studi del Mondo Classico e del Mediterraneo Antico. Sezione linguistica* 5: 183–195.
- de Simone, Carlo. 1984. La posizione linguistica della Daunia. In *La civiltà dei Dauni nel quadro del mondo italico. Atti del XIII Convegno di studi etruschi e italici Manfredonia 12–27 giugno 1980*, ed. Aldo Neppi Modona, 113–127. Firenze: Olschki.
- de Simone, Carlo. 1988. Iscrizioni messapiche della Grotta della Poesia (Melendugno, Lecce). *Annali della Scuola Normale Superiore di Pisa* 18: 325–415.
- de Simone, Carlo. 1991. La lingua messapica oggi. Un bilancio critico. In *I Messapi, Atti del trentesimo convegno di studi sulla Magna Grecia, Taranto-Lecce, 4–9 ottobre 1990*, ed. Giovanni Pugliese Carratelli, 297–322. Napoli: Arte Tipografica.
- de Simone, Carlo. 2018. Messapic. In *Comparative Indo-European linguistics. An international handbook of language comparison and the reconstruction of Indo-European*,

- ed. Jared Klein, Matthias A. Fritz, Brian Joseph, & Mark Wenthe, 1839–1850. Berlin: De Gruyter.
- Theodoridis, Christos. 1998. *Photii Patriarchae lexicon*. Volumen II: E–M. Berlin: De Gruyter.
- Untermann, Jürgen. 2000. *Wörterbuch des Oskisch-Umbrischen*. Heidelberg: Winter.
- Vetter, Emil. 1943. Literaturbericht 1934–1938. *Glotta* 30: 15–84.
- Vine, Brent. 2017. Armenian *lsem* ‘to hear’. In *Usque ad radices. Indo-European studies in honour of Birgit Anette Olsen*, ed. Bjarne Simmelkjær, Sandgaard Hansen, Adam Hyllested, Anders Richardt Jørgensen, Guus Kroonen, Jenny Helena Larsson, Benedicte Nielsen Whitehead, Thomas Olander, & Tobias Mosbæk Søborg, 767–781. Copenhagen: Museum Tusulanum Press.
- Vine, Brent. 2019. Faliscan *foied*, Latin *hodiē* ‘today’, and Italic **d(i)jē-*. In *QAZZU warrai. Anatolian and Indo-European Studies in Honor of Kazuhiko Yoshida*, ed. Adam Alvah Catt, Ronald I. Kim, & Brent Vine, 380–392. Ann Arbor: Beech Stave Press.
- Watkins, Calvert. 1963. Preliminaries to a historical and comparative analysis of the syntax of the Old Irish verb. *Celtica* 6: 1–49.
- Weiss, Michael L. 2020. *Outline of the historical and comparative grammar of Latin²*. Ann Arbor: Beech Stave Press.
- Zair, Nicholas. 2016. *Oscan in the Greek alphabet*. Cambridge: Cambridge University Press.

Appendix: List of vowel alternations in Messapic

(Potential loanwords marked with following *)

Alternation	Attestations
1 a ~ e	<ul style="list-style-type: none"> – <i>artamihi</i>* (MLM 21 Bas, first half 6–first half 5 c. BCE) ~ <i>artemes</i>* (MLM 7 Cae, 4–3 c. BCE) – <i>dazomas</i> (MLM 3 Ro, MLM 23 Ro, MLM 27 Ro; MLM 21 Gn, 4 c. BCE; MLM 4 Cae, 4 c. BCE) ~ <i>dezomas</i> (4 Ro) – <i>kazareihī</i>* (MLM 1 Bas, 3 c. BCE) ~ <i>kezareihē</i>* (MLM 1 Di, 3 c. BCE) – <i>šarnaiath</i>[i] (MLM 1 Hy, 4 c. BCE) ~ <i>šernaihi</i> (MLM 22 Cae, 4 c. BCE) – <i>dazzes</i> (MLM 16 Gn, 4 c. BCE) ~ <i>dezes</i> (MLM 21 Lup, second half 3–late 2 c. BCE)
2 a ~ o	<ul style="list-style-type: none"> – <i>artos</i> (MLM 10 Ar, second half 3 c.–end of 2 c. BCE; MLM 12 Ruv, 3 c. BCE) ~ <i>arta</i> (MLM 9 Os, 4 c. BCE); cf. also <i>arθas</i> (MLM 37 Cae, 3 c. BCE; MLM 4 Man, 3 c. BCE) and <i>ar^ϕaihi</i> (MLM 17 Al, second half 3 c.–end of 2 c. BCE; MLM 38 Al, 3 c. BCE; MLM 31 Ca, date uncertain) – <i>θabara</i> (MLM 3 Bas, second half 5 c. BCE; MLM 2 Man, 3 c. BCE; MLM 6 Ur, 3 c. BCE) ~ <i>θobara</i> (MLM 22 Rud, 3 c. BCE)

(cont.)

	Alternation	Attestations
3	ao ~ o	– <i>baolihi</i> (MLM 26 Bas, first half 6–first half 5 c. BCE; MLM 22 Ne, first half 6–first half 5 c. BCE) ~ <i>bollihi</i> (MLM 7 Gn, 3 c. BCE) – <i>taotori</i> (MLM 9 Bal, 4 c. BCE; MLM 16 Ro, MLM 19 Ro, MLM 23 Ro) ~ <i>totor</i> [<i>i</i>] (MLM 3 Car, 3 c. BCE) – <i>traohanθihi</i> (MLM 5 Bal, 3 c. BCE) ~ <i>trohanθes</i> (MLM 35 Cae, 3 c. BCE; MLM 5 Os, 3 c. BCE)
4	ao ~ oho	– <i>baonnes</i> ? (MLM 19 Cae, 3 c. BCE) ~ <i>bohonihi</i> (MLM 33 Rud, 3 c. BCE)
5	ao ~ au	– <i>daos</i> (MLM 3 Ro, MLM 12 Ro, MLM 18 Ro) ~ <i>daus</i> (MLM 4 Ro, MLM 5 Ro)
6	ao ~ au ~ o	– <i>klauihi</i> (MLM 4 Ro; MLM 5 Ro) ~ <i>klaohi</i> (MLM 8 Ro; MLM 1 Br, 3 c. BCE) ~ <i>klohi</i> (MLM 1 Bas, 3 c. BCE; MLM 1 Ga, 3 c. BCE; MLM 2 OI, 4 c. BCE) – <i>θautour</i> (MLM 4 Ro; MLM 5 Ro) ~ <i>θaotor</i> (MLM 35 Rud, second half 3–late 2 c. BCE; MLM 1 Br, 3 c. BCE; MLM 3 Ro; MLM 8 Ro; MLM 15 Ro) ~ <i>θotor</i> (MLM 21 Ro)
7	au ~ eo	– <i>daus</i> (MLM 4 Ro; MLM 5 Ro) ~ <i>deos</i> (MLM 13 Ro; MLM 14 Ro; MLM 19 Ro) – <i>aviθaos</i> (MLM 32 Al, second half 5 c. BCE; MLM 25 Ur, 3 c. BCE) ~ <i>aviθeos</i> (MLM 30 Al, 3 c. BCE; [avi]θeos MLM 15 Al, 3 c. BCE)
8	e ~ i	– <i>dazehi</i> (MLM 16 Ro) ~ <i>dazihi</i> (MLM 3 Car, 3 c. BCE) – <i>mogelihe</i> «os»? (MLM 6 Rud, 3 c. BCE) ~ <i>mogilles</i> (MLM 25 Gn, 4–3 c. BCE)
9	ei ~ i	– <i>deiva</i> (MLM 1 Sa, 3 c. BCE; MLM 5 Vi, 3 c. BCE) ~ <i>diva</i> (MLM 3 Ar, 4 c. BCE; MLM 7 Ar, 3 c. BCE; MLM 1 Vi, 4–3 c. BCE) – <i>veinan</i> (MLM 1 Bas, 3 c. BCE) ~ <i>vinas</i> (MLM 35 Ur, late 5 c. BCE) – <i>epeigrave</i> (MLM 1 Di, 3 c. BCE) ~ <i>epig[ra]van</i> (MLM 3 Car, 3 c. BCE) – <i>blatθei</i> « <i>h</i> »as (MLM 16 Gn, 4 c. BCE) ~ <i>blatθihi</i> (MLM 26 Al, 3 c. BCE) – <i>beileia</i> (MLM 7 Rud, second half 3 c. BCE) ~ <i>bilias</i> (MLM 24 Rud, second half 3–late 2 c. BCE; MLM 13 Ur, 4 c. BCE; MLM 24 Cae, second half 3–late 2 c. BCE; MLM 17 Ro) – <i>korθeithi</i> (MLM 10 Uz, second half 3–late 2 c. BCE) ~ <i>kor.θ.ithi</i> (MLM 1 Lup, 3 c. BCE)
10	eo ~ o	– <i>θeotor</i> (MLM 42 Lup, 3 c. BCE; MLM 12, 13 Ro) ~ <i>θotor</i> (MLM 1 Car, 3 c. BCE; MLM 31 Gn, 3 c. BCE; MLM 11, 12 Lup, 3 c. BCE; MLM 17 Lup, 4 c. BCE) – <i>θeotoras</i> (MLM 15 Bas, 3 c. BCE; MLM 2 Os, 4–3 c. BCE; MLM 32 Cae, 3 c. BCE) ~ <i>θotoras</i> (MLM 32 Cae, 3 c. BCE) – <i>θeotori</i> (MLM 6 Ro; MLM 7 Ro; MLM 13 Ro) ~ <i>θotori</i> (MLM 21 Ro) – <i>aviθeos</i> (MLM 15 Al, 3 c. BCE; MLM 30 Al, 3 c. BCE) ~ <i>aviθos</i> (MLM 1 Ga, 3 c. BCE)
11	i ~ o	– <i>andirahho</i> (MLM 4, 16 Ro) ~ <i>andoraion</i> (MLM 20 Ro; <i>andoraion</i> MLM 21 Ro) – <i>dazimas</i> (MLM 6 Cae, 3 c. BCE) ~ <i>dazomas</i> (MLM 21 Gn, 4 c. BCE; MLM 4 Cae, 3 c. BCE; MLM 3, 7, 27 Ro) – <i>hipaka</i> (19 Ur, 3 c. BCE) ~ [h]opaka (4 Vi, 4–3 c. BCE) – <i>hipades</i> (24 Cae, second half 3–late 2 c. BCE) ~ <i>opades</i> (11 Ro)
12	i ~ oho?	– <i>dazinnih</i> ? (MLM 18 Bas, 3 c. BCE) ~ <i>dazohonnih</i> (MLM 1 Bas, 3 c. BCE)
13	i ~ u	– <i>vinaihi</i> (MLM 18 Ro) ~ <i>vunaihi</i> (MLM 5 Ro)
14	i ~ u ~ ∅	– <i>damatira</i> * (MLM 13 Me, first half 6–first half 5 c. BCE) ~ <i>damatura</i> * (MLM 10 Ruv, 3 c. BCE) ~ [d]amatra* (MLM 14 Gn, 4 c. BCE)

(cont.)

Alternation	Attestations
15 o ~ u	<ul style="list-style-type: none"> – <i>ǵotori</i> (MLM 21 Ro) ~ <i>ǵoturi</i> (MLM 21 Ro) – <i>koletos</i> (MLM 3 Ruv, 3 c. BCE) ~ <i>kuloi</i> (MLM 2 Ruv, 3 c. BCE) – <i>plator</i> (MLM 23 Cae, 3 c. BCE; MLM 17 Ro) ~ <i>platur</i> (MLM 2 Gro, second half 5 c. BCE; MLM 3 Sa, 3 c. BCE; MLM 9 Ruv, 3 c. BCE) – <i>platoras</i> (MLM 1 Ve, 3 c. BCE; MLM 2 Cae, 3 c. BCE) ~ <i>platuras</i> (MLM 2 Gro, second half 5 c. BCE)
16 o ~ oo	<ul style="list-style-type: none"> – <i>stabos</i> (MLM 36 Al, 3 c. BCE) ~ <i>staboos</i> (MLM 1 Bas, 3 c. BCE) – <i>polla</i> (MLM 20 Ve, 3 c. BCE), <i>polonnihi</i> (MLM 20 Gn, 3 c. BCE) ~ <i>pollo</i> (MLM 31 Al, 3 c. BCE)
17 o ~ ou	<ul style="list-style-type: none"> – <i>dazos</i> (MLM 8 Uz; MLM 12 Ro) ~ <i>dazoun</i> (MLM 1 Can, 3 c. BCE) – <i>taotori</i> (MLM 9 Bal, 4 c. BCE; MLM 16, 19, 23 Ro) ~ <i>ǵautouri</i> (MLM 4 Ro) – <i>ǵaotor</i> (MLM 3, 8, 15 Ro; <i>taotor</i> MLM 24 Me, 4–3 c. BCE) ~ <i>ǵautour</i> (MLM 2 Ro)
18 oo ~ oa	<ul style="list-style-type: none"> – <i>staboos</i> (MLM 1 Bas, 3 c. BCE) ~ <i>staboas</i> (MLM 49 Lup, 3 c. BCE; <i>stabovas</i> MLM 29 Al, 4 c. BCE)
19 u ~ oo	<ul style="list-style-type: none"> – <i>platur</i> (MLM 3 Sa, 3 c. BCE; MLM 9 Ruv, 3 c. BCE) ~ <i>platoorrihi</i> (MLM 7 Gn, 3 c. BCE)