Talking stones. 
Phonology in Latin inscriptions?

Giovanna Marotta

Abstract
In spite of the long-standing debate on the value of epigraphic data, especially in the studies aiming at reconstructing the sociolinguistic framework of Latin, scholars still disagree on the value to be assigned to abnormal (i.e. non classical) spellings occurring in inscriptions. Are they clues suggesting pronunciations reflecting the social class of the reader/speaker? Are they simple mistakes in writing? Are they a sign of the archaizing style typical of the epigraphic register?

The paper focuses on the graphemic alternations <i>~<e> and <u>~<o> occurring within CLASSES I, a corpus of inscriptions of the Archaic and Early periods of the Latin language. The distribution of vowel alternations in spelling is not casual, but rather suggests a plausible correspondence in phonological variation.

The fine-grained comparison carried out on the lemmata occurring in CLASSES I reveals a quite complex orthographic picture, where vowel alternations cannot be ascribable to archaism pure and simple but rather may be interpreted as evidence for a sociophonetic process sensitive to both lexical and prosodic constraints.

Keywords: Latin inscriptions, phonological variation, historical sociolinguistics.

1. The Latin language and sociolinguistic variation

Latin is one of the most studied languages in the world, for two thousand years at least. However, the large amount of material available for this language does not prevent scholars from finding something new.

For many scholars, Latin is essentially a literary object. The very label ‘Classical Latin’ crucially makes reference to the literary canon (vide infra). However, if someone chooses to imagine Latin as a natural language, i.e. as a culturally and socially constrained symbolic structure, the dimensions of speech and sociolinguistic variation cannot be ignored. Consequently, Latin should be conceived in oral terms in addition to its written form, with obvious methodological caveats.

Recent studies on Latin linguistics have two essential and complementary foci: sociolinguistics and pragmatics. Such an extension of perspective is due, at least in part, to the collection and analysis of data pertaining to
non-literary texts that provide evidence of linguistic forms not compatible with the normative tradition of Classical Latin. In particular, in the last few decades the discovery of new materials (tablets, ostraca, inscriptions) has opened new prospects for the multi-hued reconstruction of this dead language. The challenge is the reconstruction of a linguistic system resembling that of the living languages of our contemporary times, *mutatis mutandis*, of course.

To look at Latin as a spoken language necessarily implies acknowledging the occurrence of linguistic variation relative to all levels of the ‘grammar’. Language is a cognitive structure revolving around a set of general principles which are universal and partly independent from its historical and cultural contexts. At the same time, every human society, even the most archaic, shows linguistic variation according to the social class of the speaker, his education, and the communicative context in which speech acts occur.

Variation is a keyword introducing sociolinguistic analysis. The application of notions and methods of modern sociolinguistics to a closed-corpus language like Latin is not an easy task, for the following reasons:

– there are no longer any speakers;
– evidence is written only;
– sociolinguistic cues available in the sources are relatively scarce;
– metalinguistic comments about linguistic variation are marginal, because in the ancient world variation was considered essentially deviation from the linguistic norm, i.e. from the classical literary language (Winter, 1998; Giacalone Ramat, 2000; Cuzzolin and Haverling, 2009; Clackson, 2011b).

Notwithstanding these constraints, the historical sociolinguistics approach seems to be suitable to apply to classical languages. This is indeed an exciting perspective, inasmuch as it permits the discovery of new scenarios and offers new insights into dead languages.

According to a widely shared line of research, a sufficiently plausible frame of sociolinguistic variation for Latin may be approached, especially with regard to the stylistic (or diaphasic) dimension (Müller, 2000; Ferri and Probert, 2010; Clackson, 2011b; Hernández Campoy and Conde Silvestre, 2012, among others). Bright examples of such a drift are represented by the recent volumes by James Adams (2003; 2007; 2013) devoted to the Latin language.
However, the quest to reconstruct the sociolinguistic dynamics of the ancient world may be inspired only marginally by the application of both quantitative and qualitative methods of modern sociolinguistics as developed for the urban areas of contemporary societies.

The distinction between more and less privileged classes is a constant element of Roman society. For instance, in Republican times, the primary social contrast is between nobles and plebeii, although the framework of the social classes is already complex and not reducible to this simple dichotomy (Müller, 2000: 274 ff.; Clemente, 1977; Parkin and Pomeroy, 2007). Even if we admit that Roman society «was inherently more hierarchical than American society in the 1960s» (Clackson, 2011b: 514), having been grounded in slavery, the available written evidence does not allow us to clearly define the class-based variation. Despite the discovery of new materials recording the actual speech of Latin speakers of (presumably) lower social status, the data available for the reconstruction of a general sociolinguistic framework are still cursory and scattered, as underlined by Sornicola (2013: 171) with reference to other historical contexts. Therefore, evidence concerning the diastratic dimension remains sporadic and occasional. According to Müller (2000: 276), linguistic variation in Latin is always diatopic first, then diaphasic and diastratic. To extend theoretical models devised for living modern languages to ancient languages is neither simple nor direct, indeed.

The risk in taking up models developed with reference to modern societies might be an unacceptable stretching of notions and methods that are heavily grounded on a mass of empirical data collected in the field, which is inconceivable for ancient languages. Hypotheses adhering to theoretical as well as formal models therefore have to be carefully tested. At the same time, the degree of linguistic awareness in the ancient world, as well as the different attitudes of speakers in matters of linguistic behaviour, cannot be automatically considered similar to those of modern societies (Kaimio, 1979: 9-19; Mancini, 2012; 2014). Such an assumption might risk a superficial analysis.

Classical Latin, a sort of ‘standard’ language (vide infra for caveats) is based on the Roman variety of Latin. The ‘standardization’ process was a literary operation developed in connection with language policies and the codification of law (Clackson and Horrocks, 2007; Cuzzolin and Haverling, 2009; Poccelli et al., 1999).

On the basis of a well-defined cultural and political design aiming at expressing the hegemony of Rome (and then, of the Latin language) on the
other Italic peoples (and their languages), a special variety of the language was chosen as the reference model for the educated elite and the upper social classes. Over the flow of time, the speakers perceived and considered this variety to be the ‘correct form’ of the language, i.e. a kind of ‘standard language’\(^1\), although the term ‘standard’ refers to a concept conceived and developed with reference to modern societies, as Sornicola (2013: 171) has properly pointed out. Therefore, in using the term ‘standard’, we should constantly be aware of the objective limits on applying this notion to classical languages (Clackson, 2015). In parallel, modern sociolinguistic studies have shown how the same notion of ‘linguistic community’ has to be conceived as a complex entity, crucially referring to social identity recognized by the speakers (cf. Le Page and Tabouret-Keller, 1985; Berruto, 2007).

The caveats discussed so far might convince scholars that the field of historical sociolinguistics is too rough to be walked through with reference to classical languages. However, in our view, these caveats do not prevent the tentative application of at least some notions and some theoretical tenets belonging to modern sociolinguistics. We do still believe that historical sociolinguistics can and must be practiced in order to achieve a richer and more realistic picture of the ancient language and society\(^2\). However, the projection of theories and methods of modern sociolinguistics onto classical languages needs to be mediated through constant reference to historical and social coordinates specific to the ancient world, since the dynamics of social and pragmatic interactions do not remain exactly the same over the centuries, due to the Uniformitarian Paradox (Labov, 1994).

2. On epigraphic data

In a tentative study of the sociolinguistic dynamics of the classical languages, a special role may be played by non-literary texts, as they are free of the filter due to the manuscript tradition. Inscriptions, papyri, private letters, ostraca, and writing tablets, represent a primary source for the study of variation in ancient languages (Marotta, in press; Marotta, Putzu and Donati, in press; Fedriani and Ramat, in press). All the texts mentioned may (but not necessarily do) reflect the *sermo cotidianus* of educated people as

---

1 In the case of Latin, the so-called Classical Latin (*vide infra*).
2 See *Marotta* (in press) for a more detailed discussion on the topic.
well as the speech of speakers belonging to the lowest social classes of Roman society (see Clackson, 2011a; Adams, 2013: 18).

Among non-literary texts, inscriptions deserve a special role, not only because they represent a direct source of evidence, but also by virtue of their reliability for the study of linguistic variation. In particular, we believe that misspellings, i.e., spellings not congruent with the ‘standard’ language as exhibited in the literary texts of the Classical period, may become useful cues for reconstructing sociolinguistic variation in the ancient world. Misspellings are traditionally considered with reference to the diachronic dimension; in our opinion, they are clues for diaphasic and diastratic dimensions too.

Nevertheless, some caveats have to be applied once more. As Adams (2003: 84) has already correctly observed, when considering any inscription on stone we should be aware of the complexity connected with its ‘authorship’, because different participants may be involved in the process:

1. the customer, i.e. the person who commissioned the inscription;
2. the author, i.e. the person who composed the text;
3. the engraver or draftsman, i.e. the person who cut it into the stone or wrote it on a different support.

The final product could be different too, according to the variable degree of education of the three participants. Draftsmen could show variable performance, with respect to their grammatical and, especially, orthographic competence. However, the stone carver could be not literate, since his skillning concerned the exact reproduction of what was lettered on the stone. The check by the author could be lacking, and the same customer could be more or less educated and careful in checking the text. Moreover, the birthplace of draftsman, author, and customer might influence the final written product. In particular, the bilingualism widespread in the ancient world could have a variable influence on the performance of the writer (cf. Adams, 2003; Mullen, 2011; 2013).

A further problematic aspect concerns archaism. Archaising forms are traditionally recognized as a typical feature of the epigraphic register, especially in the case of public and official inscriptions. Therefore, the spelling observed in inscriptions of a given age may simply reflect earlier orthographic usage, not corresponding to the current pronunciation by contemporary speakers (see § 7 below for a more detailed discussion).

However, the occurrence in inscriptions of forms belonging to past
stages of Latin as reconstructed using comparative analysis of the ancient Indo-European languages is not evidence per se against their tenacity and persistence in the spoken language for a long time.

As a general remark, we believe that to segregate epigraphic data into the restricted boundaries of archaism, and consequently to consign them to the margins of the empirical evidence, would lead to a paradox: on the one hand, we would consider Latin epigraphic data marginal, due to the availability of a very rich literary corpus; on the other, we are forced to consider inscriptions the main kind of evidence for other Italic languages, being all that is available.

The size and complexity of the debate concerning the critical evaluation of epigraphic sources reflect the difficulties intimately connected with the interpretation of data from inscriptions, as discussed elsewhere (Marotta, in press). Indeed, since different factors interact and overlap in the text of an inscription, the interpretation of non-standard spellings is not an easy matter. In principle, non-classical forms occurring in inscriptions may have multiple interpretations:

a. they may be considered archaic forms, surviving only in the epigraphic texts;
b. they may be simple scribal errors;
c. they may represent a different pronunciation from that of ‘standard’ Latin.

For instance, forms like dono or dede could be due to a desire to preserve archaising forms; conversely, they could be spellings attesting alternation between competing phonological variables, socially marked (vide infra).

Aiming at employing epigraphic data in a sociolinguistic frame, we collected a corpus of inscriptions which permitted a qualitative and quantitative analysis of the non-classical forms attested in the Corpus Inscriptionum Latinarum, starting from the Early and Archaic stages of the Latin language. Our hope is that this corpus can become a useful tool allowing fine-grained analysis of the forms attested in the epigraphic texts.

---


4 For the meaning associated with the attribute ‘non-classical’, see Donati et al. (in press), Marotta (in press).
3. Building the corpus CLASSES I

CLASSES I includes inscriptions of the Archaic and Early periods (according to Cuzzolin and Haverling, 2009). Inscriptions selected from the Corpus Inscriptionum Latinarum (CIL) for this section of our corpus are dated from ca. 350 to ca. 150 BC, with most of them falling in the 3rd century BC.

Within the large number of epigraphic texts available for these periods of Latin, the following documents have been excluded:

- inscriptions composed of single letters and initials;
- legal texts, due to their high degree of archaism;
- the inscriptions from the necropolis of Praeneste, as containing anthroponyms in nominative form almost exclusively.

CLASSES I includes 386 inscriptions, comprising 1869 words. Each epigraphic text is enriched with extra- and meta-linguistic information regarding its provenance, dating and text type. Following the traditional classification by CIL, four basic text types have been identified:

a. tituli honorarii (18 exemplars), i.e. inscriptions celebrating public people and inscriptions on public monuments;

b. tituli sepulcrales (26), i.e. epitaphs and memorial texts;

c. instrumenta domestica (246), i.e. inscriptions on domestic tools;

d. tituli sacri (96), i.e. votive inscriptions.

For the annotation of the graphic form, the classification provides for the following kinds of word:

a. complete words (e.g. CIL I² 45 diana);

b. abbreviations, i.e. every kind of shortening, including personal name initials (e.g. CIL I² 46 don for donum);

c. incomplete words, i.e. words partly completed by editors (e.g. CIL I² 448 me[nervae]);

For a detailed description of the criteria informing the selection of inscriptions, see Donati et al. (in press), De Felice et al. (in press) and Donati (2015).
It is interesting to observe that complete words represent roughly half of the entire corpus (54%), whereas the second most represented category is abbreviations (see De Felice et al., in press, for further details). The complex annotation allows the correlation of the linguistic data with the text types, thus making analyses that take into consideration all the interrelations of data in a sociolinguistic vein.

The phenomena observed (vowel alternations, syncope, final consonant deletion, archaic spelling for diphthongs, etc.) are not new, as they belong to the long tradition of studies aimed at finding evidence of sociolinguistic variation in the Latin language6.

4. Vowel alternations in CLASSES I

Vowel alternations represent the first phonological topic we addressed within the corpus CLASSES I. In particular, we focused on the graphemic alternations <i>~<e> and <u>~<o> that occur in the 386 texts such as MERITO~MERETO, DEDE~DEDIT; DONUM~DONO(m).

If epigraphic documents can be regarded as a reliable source for studying sociolinguistic variation in ancient languages (cf. Consani, in press, for Ancient Greek; Kruschwitz, 2015; De Angelis, in press; Marotta, in press; Rovai 2015), albeit with caution (cf. § 2), spelling variants may therefore be considered relevant clues for reconstructing the complex phonological frame of Latin in Republican times. In our opinion, the actual pronunciation of the language could vary with respect to speech context or according to the social strata of speakers, just as happens in today’s societies, mutatis mutandis, of course.

On the other hand, the longstanding debate in Latin linguistics about the meaning of labels like *vulgaris lingua, plebeius* or *rusticus sermo* is a clear instance of the sensitivity to sociolinguistic matters *avant la lettre* (see recent studies by Mancini, 2006 and references therein). Since the seminal work by Campanile (1971), which was then continued by other Italian scholars (e.g. Vineis, 1984; Giacalone Ramat, 2000; Molinelli, 2006), the attention to variation has been constant, seeking evidence concerning a diasystemic view of the Latin language.

The present study aims at continuing such a promising line of research by adding an integrated and quantitative approach to the epigraphic data.

The variables considered relevant for the current statistical analysis are the following:

1. letter alternations;
2. morphophonological constraints;
3. lexical stress.

With reference to the first, in *CLASSES I* the number of types (i.e. lemmata) showing <e> / <o> instead of <i> / <u> is 25 for <i> vs. <e> and 138 for <u> vs. <o>, whereas the number of tokens showing allographs is 105 for <e> and 312 for <o>\(^7\).

These quantitative data show that vowel alternations affect the back series more than the front series, in terms of both types and tokens. The same asymmetric pattern emerges in the relative distribution of the letters. The results summarized in Table 1 show that the mid allograph of the back series is much more frequent than that of the front series (<o> = 64% versus <e> = 13%).

<table>
<thead>
<tr>
<th>Letters</th>
<th>Vowel /ɪ/</th>
<th></th>
<th>Letters</th>
<th>Vowel /ʊ/</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tokens</td>
<td>%</td>
<td></td>
<td>Tokens</td>
</tr>
<tr>
<td>&lt;i&gt;</td>
<td>472</td>
<td>87</td>
<td>&lt;u&gt;</td>
<td>166</td>
</tr>
<tr>
<td>&lt;e&gt;</td>
<td>70</td>
<td>13</td>
<td>&lt;o&gt;</td>
<td>298</td>
</tr>
<tr>
<td>Total</td>
<td>542</td>
<td>100%</td>
<td>Total</td>
<td>464</td>
</tr>
</tbody>
</table>

Table 1. Spellings of Latin short high vowels in *CLASSES I*.

\(^7\) With the exclusion of a few instances of hiatus.
A Pearson’s χ-square test⁸ conducted on the frequency data relative to the distribution of the mid vs. high allographs of the front and back series (see Table 1) yields a highly significant result ($\chi^2 (1, N=1006)= 283,683, p<0.001$). This means that the distribution presented in Table 1 is not due to chance, thus confirming that the front and back series of vowels are different from each other with relation to the frequency of mid or high allographs.

5. Morphophonology and the feature [BACK]

So far, the analysis of vowel alternations occurring in CLASSES I has revealed heterogeneous behaviour. The relative proportion of the letters <e> and <o> that alternate with <i> and <u> respectively is statistically relevant (vide supra, § 4). In short, vowel alternations appear to be sensitive to the phonological feature [BACK], which crucially discriminates between the two vowel series in Latin phonology (Marotta, 1999).

The asymmetric pattern emerging from the data can quite easily be interpreted with reference to morphophonological constraints. Indeed, in 87% of the items showing <o> instead of Classical <u>, the target vowel occurs in the final syllable of the word. In particular, the ending <-o> instead of the Classical -us and -um is very frequent in nouns belonging to the second declension of Latin grammar; e.g. Cornelio (Cl. Latin Cornelius), dono (Cl. Latin donum). In considering this specific morphophonological class, a special role must be recognized for names. High-frequency scores associated with names are entirely predictable in inscriptions. And this word class is very frequent in CLASSES I too.

It is worth observing that the great majority of names belonging to the second declension do not show any form ending with <-u>; rather, only <-o> is found, with occasional <-s>; for instance, atilio, claudio, decimio, fourio(s), novios, sextio, tolonio, valerio.

The spelling <-o> for the phoneme /ŭ/ affects word classes with dif-

---

⁸ The chi-square ($\chi^2$) test is a statistical test commonly used to compare the observed data (in our case, those reported in Table 1) with the expected data, i.e. those that we would expect under the null hypothesis (the hypothesis that the data are independent: in our case, that there is no significant difference between the front and the back series, with relation to the number of mid/high allographs), and to determine whether the difference between the expected frequencies and the observed frequencies is significant. Roughly speaking, the more the observed values differ from the expected ones, the less the observed distribution is likely to be due to chance.
ferent ratios, directly dependent on their frequency in the corpus. Thus, the most affected word class is names (56%), followed by common nouns (30%), whereas the other morphological classes show frequencies that are low (verbs: 7%; adjectives: 6%), and even negligible (e.g. pronouns: 0.3%).

In the front series, the spelling <e> for /i/ also shows the highest percentage of occurrence in the case of names (45%), whereas lower values are associated with common nouns (23%)\(^9\) and verbs (23%).

Dealing with the asymmetric pattern observed in the data from CLASSES I, other relevant aspects come from the study by Rovai (2015) devoted to the Greek inscriptions of Delos containing Latin names referring to Roman negociatarios active in the island. Whereas Latin /u:/ and /ũ/ are systematically spelled <ου> and <ο> respectively in Greek, /i:/ may be spelled <ει, ι, η>, and /ĩ/ <ι> or <υ>. Rovai claims that for the back series only, there is clear evidence for a phonetic distinction between Latin short and long high vowels.

In rejecting the traditional hypothesis based on alleged Sabellic features, Rovai (2015) instead considers the asymmetry to be due to general phonetic constraints. As is well-known, the articulatory space available for the production of back vowels is narrower than that for front vowels. Phonological developments involving changes in timbre are therefore expected to occur in the back series before they occur in the front series (within the Romance domain, cf. already Martinet, 1955: 98-99; Loporcaro, 2011b).

Most of the inscriptions from Delos are dated to the second century BC; that is, they are a bit later than those included in CLASSES I. Is it therefore absurd to assume that in the Latin spoken in Delos at that time Roman speakers would pronounce Atilio, dono, pocolo, Cornelio, mereto, etc., and then write atilio, dono, pocolo, cornelio, mereto, etc.? Evidence from the transliteration of Latin names into Greek supports a negative answer to that question, i.e. they probably perceived and then pronounced an ο-like sound.

\(^9\) Within the class of common nouns, the most frequent items in CLASSES I are dono(m) and poco(lo)(m), attested respectively 27 and 20 times; quite frequent also are debe(t) (11 times), mereto(d) (10), and menerva (10).
6. Vowel alternations and lexical stress

Lexical stress has also been taken as a statistical variable in the analysis of vowel alternation. Despite the skeptical opinion expressed by Adams (2013: 42), the prosodic element seems to be not irrelevant. Indeed, the evidence from the corpus CLASSES I shows that the great majority of the letters <e o> standing for high short vowels occur in unstressed syllables, more precisely in post-stress position. Table 2 presents the quantitative data concerning the interaction between vowel alternations and their prosodic context.

<table>
<thead>
<tr>
<th>prosodic context</th>
<th>&lt;e&gt; = /ĭ/</th>
<th>&lt;o&gt; = /ŭ/</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>tokens</td>
<td>%</td>
</tr>
<tr>
<td>stressed syllable</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>pre-stress syllable</td>
<td>13</td>
<td>20</td>
</tr>
<tr>
<td>post-stress syllable</td>
<td>42</td>
<td>66</td>
</tr>
<tr>
<td>monosyllables</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>64(10)</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 2. Frequency of <e o> (standing for high short vowels not in hiatus) with regard to stress position.

The primary target of vowel lowering is the post-stress syllable (<e> = /ĭ/ = 66%; <o> = /ŭ/ = 95%), suggesting that vowel lowerings of Archaic Latin are not different from those occurring in Late Latin, since both are systematic and concern post-stress syllables.

Prosodic context may become a relevant variable in the interpretation of vowel alternations insofar as we admit that [i, u], or even better: [ɪ, ʊ], co-exist with [e, o] \textit{ab antiquo}, especially in the context of unstressed syllables. Indeed, stress favours the production of tense and long vowels, whereas post-stress syllables are the best target for laxing (Labov, 1994). Moreover, an alleged Sabellic influence may also have been a factor favouring vowel lowering in Latin\(^{11}\). That short /ĭ/ and /ŭ/ had a more open pronunciation than their long counterparts is also demonstrated by the Greek transcriptions of Latin words with short <i> with Greek <ε> (e.g. Τεβέριος for Tiberius, cf. Allen, 1986).

\(^{10}\) In Table 1 the number of occurrences of <e> is slightly different (70 instead of 64), since it includes six occurrences of <e> in hiatus.

\(^{11}\) Due to space constraints, we are forced to avoid entering into such a complex issue here. See Wallace (2007) for the vowel system in Sabellic languages.
1978: 49) as well as the Etruscan transcription of Tiberius as <Θefarie> in the Pyrgi tablets (cf. Clackson, 2011a).

In our opinion, graphemic alternations between <e>~<i> = /ĭ/ and <o>~<u> = /ŭ/ suggest the persistence of the same alternations in pronunciation. And ‘deviant spellings’ (with respect to the classical norm) occurring in CLASSES I may be viewed as clues for sociolinguistic variation. As is well known, spellings with <e, o> instead of <i, u> were considered by Romans to be due to archaism as well as to the so-called vox rustica, or sermo plebeius (Mancini, 2004). Even for this phonological process, antiquitas and rusticitas go hand in hand with reference to dialectal, i.e. ‘sub-standard’, features, as normally happens in the metalinguistic consciousness of Romans (speakers and grammarians).

Contra Loporcaro (2011a: 111; 2015) and Adams (2013), we believe that the phonetic process [i, u] > [ɪ, ʊ] > [e, o] occurring in spoken Latin from very early on not only is sufficient evidence that changes in quality were undergone by short high vowels, but also suggests a premature collapse of Latin vowel quantity, at least in some ‘sub-standard’ varieties of the language. Therefore, we think it is possible that vowel shifts occurred consistently and frequently in spoken Latin as early as the Archaic period and persisted without interruption through Late Latin and the Romance languages (cf. Straka, 1959; Pulgram, 1975; Vineis, 1984).

As happens in many natural languages, in spoken Latin length co-occurred with tenseness. According to the traditional picture of Latin phonology, the feature [LONG] is retained as the distinctive one, whereas [TENSE] is considered redundant. As we have discussed elsewhere (Benedetti and Marotta, 2014), these two features differ in sociophonetic status in Latin. Indeed, in vowel system, [TENSE] = [CLOSE] (Allen, 1978; Adams, 2007; 2013) was the distinctive feature in the low varieties of the language, whereas [LONG] was operative only in the highest variety, i.e. in so-called Classical Latin, which is nothing more than a register. In our view, the contrast between [TENSE] and [LAX] can be viewed as a sociophonetic feature, i.e. as a phonological property acting as a social marker able to discriminate the speech behaviour of speakers according to social parameters as well as to pragmatic contexts.

The phenomenon of vowel alternations may therefore be interpreted as another instance of ‘linguistic carsism’ applied to the history of Latin: like in a carsic river, lowerings as well as syncope appear and then disappear;
they occur in Archaic Latin and then in Late Latin and the Romance languages. In parallel, vowel quantity is lost in Late Latin and in the Romance languages. The three phonological processes belong to the same linguistic drift, aiming at confining vowel quantity to the boundaries of the phonological system, until its disappearance.

In sum, the data from the epigraphic texts of CLASSES I show systematic vowel alternations prosodically constrained. Variation in spelling may be interpreted as reflecting variation in pronunciation, and thus as a clue for sociophonetic variation in Latin.

7. On archaism in epigraphic data

Spellings divergent from the orthographic norm of Classical Latin that occur in inscriptions belonging to the Early and Archaic periods of the language are traditionally interpreted as due to archaism (e.g., Adams, 2007: 691; 2013: 42; Clackson and Horrocks, 2007: 92). The vowel alternations under scrutiny are also normally interpreted as a typical feature of the epigraphic register that resorts to archaic forms but is no longer productive in contemporary speech.

It is well known that at all times, inscriptions show a general trend towards the employment of fossilized forms belonging to the past. For instance, on tombstones in Italian cemeteries it is still usual today to find archaic and erudite words like *una prece* “a prayer” and *addì* “date”, although no Italian speaker would produce such forms nowadays. Nevertheless, to admit a general trend towards using archaism in epigraphic texts does not imply that every misspelling should be interpreted in that light. In other words, still recognizing the endurance of archaism, we do believe that alternative and parallel hypotheses are possible, inasmuch as spellings found in inscriptions can also be considered simple scribal errors or relevant clues for phonetic/phonological representation, i.e. as evidence for reconstructing the actual pronunciation of a language (cf. Montgomery, 1999).

---

12 Quoting Adams (2013: 42): «Some of these forms are special cases. In perfect verb endings the *e* was early (cf. e.g. Osc. *deded*; also Sihler, 1995: 461), and the spelling *-et* was very persistent, even in the imperial period. In some other cases the *e*-spelling is also old, going back for example to the time before vowel weakening had occurred (meretì)».

13 The corresponding words commonly used would be *una preghiera* and *data*, respectively.
In examining the data from *CLASSES I* concerning the phenomenon of vowel alternations (<i>~</i>e, <i>o</i>~<i>u</i>), the potentially available interpretations take on different degrees of probability with regard to single items. In particular, the resort to archaism for the occurrence of <i>e</i> and <i>o</i> (<i>/ɪ/</i> and <i>/ʊ/</i>) seems reasonable in the following cases:

- names; e.g. *atilio*, *canoleios*, *gabinio*, *hercole*, *menerva*, *apolone*, *calebus*;
- nouns referring to public positions and roles; e.g. *cosol*, *praifecto*, *tribunos*, *aidile(s)*;
- common nouns belonging to formulaic writing typical of the epigraphic register; e.g. *dono dede mereto* (e.g. *CIL I² 31*); *optumo viro* (e.g. *CIL I² 9*).

It becomes less reasonable in the case of items like *imfosos*, *dedro*, *honc*, *opos*, *oino*, *pocolo(m)*, *sacro(m)*, *tempestatebus*, although some of these forms show other spelling features compatible with the alleged archaizing fashion of the inscriptions. The resort to archaism appears even more dubious for the form *pocolo(m)*, which is particularly frequent in *CLASSES I*. Table 3 shows the allographs for this lemma with their frequencies, in both absolute and relative terms.

<table>
<thead>
<tr>
<th>Lemma</th>
<th>tokens</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>pocolo</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>pocolom&lt;sup&gt;14&lt;/sup&gt;</td>
<td>14</td>
<td>70</td>
</tr>
<tr>
<td>pococoolum</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>pocolum</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 3. *Frequency of allographs for the word poculum in CLASSES I.*

As a first note, it is worth underlining the absence of the classical form *poculum*, which should already have been in use in common orthography as well as in the speech of the third century BC, according to the traditional

<sup>14</sup> Two of these examples are likely restorations of the final letter.
hypothesis centered on archaism. The allograph POCOCOLOM, occurring once, may easily (but not for sure, of course) be interpreted as a simple material mistake of the engraver. The same interpretation may be applied to POCOLOM, despite its syncope.

In sum, the letter <u> never occurs in this word, either in a post-stress syllable or in a final syllable, even in possible material mistakes. Therefore, in the case of POCOLOM, what would the reason be for suggesting the employment of an archaic form, if contemporary people were currently pronouncing ['pokulum']?

In both its meanings of “a drinking-vessel, cup, bowl” and “the contents of a cup, a drink, draught; (pl.) social drinking, potations” (see OLD, s.v.), POCULUM can be referred to a semantic field belonging to the everyday life of Roman people. Its frequent occurrence in Plautine comedies as well as in Cicero’s writings is clear evidence of this. This word can be assigned the status of a common, ordinary word, entering into the sermo cotidianus of Latin speakers, as well as marking the possession of the cup or bowl, due to its occurrence in the instrumenta domestica of CIL. The so-called pocula deorum are a series of bowls bearing the name of deities prominent in Rome in the third century BC (e.g. Minerva, Venus, Aesculapius). As being temple souvenirs, and not dedications, they were taken away from the temple, and therefore found widely dispersed (see Beard et al., 1998).

Therefore, the epigraphic orthography with <o> instead of <u> in the word POCOLO(m) cannot be immediately justified as an archaism, especially if it is associated with artifacts such as bowls, cups and the like. The alternative hypothesis that recognizes a phonological status for the spelling used in inscriptions seems more suitable, at least for this word. The same interpretation could be extended to other items occurring in CLASSES I: nouns (e.g. CATINO, FILIOS, LOCOM), verbal forms (e.g. DEDRO, FUET, FECIE), adjectives (e.g. SACRO, MILITARE), pronouns (HEC).

Negative evidence must be taken into account too. If the writing <e> for <i> was phonetic, then we would expect to find the ending <-et> in place of <-it> over a large set of verbs, and not only for DEDET. Considering the perfect form of the verbs facio and sum, the picture drawn from our corpus is that shown in Table 4.

---

15 E.g. Pl. St. 272; Pl. Truc. 43; Cic. Sen. 46.
verb *facio* | verb *sum*  
---|---
perfect form | tokens | perfect form | tokens  
FECE | 1 | FUET | 1
FECID | 1 | FUIT | 2
FECIT | 16 | – | –

Table 4. Allographs for perfect forms of the verbs *facio* and *sum* in CLASSES I.

The very low number of instances for perfect forms of FUET~FUIT prevents us from drawing legitimate conclusions, whereas the data concerning FECIT indicate that this form was the one commonly written and produced by Latin speakers, although the occurrence of one instance with final -e (FECE, CIL I² 416) cannot be ignored.

On the other hand, the occurrence in CLASSES I of other perfect forms that do not show the spelling alternation <-IT~-ET> (DEDICAVIT, CIL I² 395; 396, CEPIT, CIL I² 7; 9, SUBIGIT, CIL I² 7, PERIT, CIL I² 2841) suggests that the morpheme in current use was the one written <-IT>. However, we cannot exclude that <-IT> could stand not only for [it], but also for [It], or even [et].

In conclusion, the vowel alternations found in the corpus CLASSES I and discussed so far do not unequivocally speak in favor of the simple preservation of fossilized forms. How can it be demonstrated that scribes maintained the longstanding spellings, if older documents evidencing these early spellings are no longer available? From the archeological and historical points of view, archaism appears to be nothing more than a null hypothesis for the interpretation of phenomena like those analyzed here. Vowel variation can be interpreted instead as a clue for sociophonetic variation occurring in the Latin language viewed in its diasystemic dimensions.

8. Conclusions

The fine-grained comparison carried out on the lemmata occurring in CLASSES I has revealed a quite complex orthographic picture, where some vowel alternations at least are not ascribable to archaism pure and simple but

---

16 Cf. CIL I² 9.
17 The two instances of FUIT occur in the same epigraph, CIL I² 7.
rather may be interpreted as evidence for a sociophonetic process sensitive to both lexical and prosodic constraints.

Spellings divergent from what would become the orthographic norm of Classical Latin constitute a reliable corpus of data providing clues to Latin as spoken during the third century BC. In particular, we would like to state that not all vowel alternations occurring in the epigraphic corpus CLASSES I should be considered engraving errors or archaisms.

A final methodological remark concerns the contrast between innovation and preservation in the spelling of epigraphic data. Although written documents of a closed-corpus language may record sporadic innovations besides the preservation of earlier phonetic and orthographic habits\textsuperscript{18}, nevertheless the assumption that Latin inscriptions (and especially those belonging to the category of instrumentum domesticum) could represent the contemporary state of the language does not appear to be meaningless. In other words, inscriptions in Early and Archaic Latin recorded not only earlier orthographic (and hence phonetic) habits, but also the actual spelling and pronunciation of contemporary Roman speakers.

Another aspect must be taken into account in order to understand the complex sociocultural picture of Rome in the third century BC. In this period, a sort of linguistic and orthographic norm could be already in use in Rome, due to the high number of foreigners who reached Rome and needed to acquire Latin in both its oral and written form (see Wachter, 1987: § 222). This pre-classical orthographic norm could employ mid vowels instead of high vowels, especially in unstressed position and in some endings with morphophonological value.

Therefore, the forms with [e o], alternating with [ɪʊ] and written <e, o> were not absent from the Latin diasystem, but rather continued to be used in the spoken language through the centuries until Late Latin, especially at the lowest levels of Roman society. Romance developments support such a drift (see Loporcaro, 2011a; 2015).

In conclusion, we hope to have brought evidence for the value of epigraphic data as a useful source for reconstructing the sociolinguistic frame of a closed-corpus language like Latin. If so, inscriptions may be considered ‘talking stones’ \textit{stricto sensu}.

\textsuperscript{18} The topic has been recently addressed by Lorenzetti and Schirru (2013: 592 ff.) with reference to spelling with <k> and <c> in Latin inscriptions of the Tripolitanian area.
Acknowledgements

Thanks are due to James Clackson, Pierluigi Cuzzolin, Margherita Donati, Rolando Ferri, Francesca Mencacci and Francesco Rovai for their useful comments on the topic addressed in this article. The research presented is part of the PRIN project Linguistic representations of identity. Sociolinguistic models and historical linguistics (PRIN2010, prot. 2010HXPF2_001).

Bibliography


GIOVANNA MAROTTA
Dipartimento di Filologia, Letteratura e Linguistica
Università di Pisa
Via Santa Maria 36
56126 Pisa (Italy)
giovanna.marotta@unipi.it