ACTA CLASSICA UNIV. SCIENT. DEBRECEN.	LV.	2019.	pp. 127–146.

THE CONFUSION BETWEEN <B> AND <V> IN LATIN INSCRIPTIONS FROM SARDINIA

BY LUCIA TAMPONI

Abstract: This paper focuses on the distribution of the alternation of <B> and <V> in a corpus of Latin inscriptions from Sardinia (1st century BC – 7th century AD). The distribution of the graphemes has been related to the dating and the provenance place of the inscriptions, and the total number of occurrences has been compared with the number of corresponding forms in Classical Latin. The amount of other consonantal misspellings in the texts has been examined as well, in order to verify whether the absence of misspellings could be due to a high degree of literacy of those involved in the crafting of the inscriptions. The results of the survey show a widespread graphemic confusion between <B> and <V> in the island, especially from the third century AD. In most of the cases, Classical Latin /w/ is represented as <B>, both in initial and internal position. It will be shown that the examination of the variables considered here could shed light on the evolution of Latin /b/ and /w/ in Sardinia.

Keywords: Latin linguistics, epigraphy, phonology, corpus linguistics, historical linguistics

1. Introduction

The graphemic confusion between <B> and <V> is a widespread phenomenon across the Roman Empire. Occurrences of these alternations are particularly frequent in epigraphic texts, at least from the first century AD. For example, as illustrated by Sturtevant,[1] inscriptions display instances of *baliat* (for *valeat*), *beni* (for *veni*), *Bibius* (for *Vibius*); confusions seem to be attested in Pompeian graffiti as well, such as *Berus* (for *Verus*).[2] A few earlier examples are discussed by Campanile,[3] who illustrates a small number of instances of the use of <β> instead of <ου> to represent Lat. /w/ in Republican times, such as βίβιου (for Ουίβιου, c. 180 BC), thus proposing to backdate the phenomenon to an earlier period. As suggested by Adams,[4] the few examples quoted by Campanile might

[1] Sturtevant 1920, 43.
[2] Väänänen 1981, 50; Adams 2013, 183.
[3] Campanile 1971, 80-84.
[4] Adams 2013, 184.

be not crucial for establishing the phonetic value of Lat. <V>; however, the confusion between the two phonemes from early times, at least in some varieties or sociolects, is an interesting hypothesis to evaluate.

As far as the causes of the phenomenon are concerned, these confusions are traditionally explained as due to a change of the phonetic value of /w/ and possibly /b/. In particular, it is suggested that /w/ developed a fricative articulation[5] or, more generally, that Latin /b/ and /w/ partially merged into a bilabial fricative /β/, which was therefore alternatively represented in Latin script with <B> and <V>.[6]

Indeed, coherently with the inscriptional evidence, the testimonies of ancient grammars seem to point at least to a fricativization of Classical Latin /w/. In this respect, the testimony of Velius Longus (Keil *Gramm. Lat.* VII.58.17–19) is particularly interesting. By the second century, the author speaks of a pronunciation of Classical Latin /w/ *cum aliqua adspiratione* – at least in word-initial or intervocalic position, according to the author's examples: «*V litteram digamma esse interdum non tantum in his debemus animaduertere, in quibus sonat cum aliqua adspiratione, ut in* ualente *et* uitulo *et* primitiuo *et* genetiuo». As correctly underlined by Adams,[7] if we accept this interpretation of the term *adspiratio*, Velius Longus' statement suggests that by the second century the fricative pronunciation of Classical Lat. /w/ was the norm.

1.1. The Romance languages

The confusion between <B> and <V> is of particular interest given the outcome of Classical Latin /w/ and /b/ in the Romance languages. It is well-known that in general the two phonemes merged in intervocalic position, whereas the merger did not take place word-initially.[8]

As far as intervocalic position is concerned, /b/ became a fricative almost everywhere, though the place of articulation varies: for example, from Lat. *caballu(m)* ('horse') and *debere* ('have to'), the Romance outcomes show a labiodental fricative in e.g. Italian *cavallo* [kaˈvallo], *dovere* [doˈvere] and French *cheval* [ʃ(ə)val], *devoir* [d(ə)vwaʀ] but a bilabial fricative in Spanish *caballo*

[5] See e.g. Allen 1965, 41; Adams 2013, 183-185.

[6] Adams 2013, 183. See also Herman 2000a, 38-39 and Adamik 2017a, 15.

[7] Adams 2013, 185.

[8] From the point of view of phonology, this difference is compatible with the so-called Coda Mirror effect, illustrated e.g. in Ségéral, Scheer 2008, whereby word-initial consonants and consonants that occur after Codas experience an effect of segmental strength, while intervocalic and Coda consonants are weak. Given the complexity of the issue, will not try to deal with this topic here, reserving it for future work.

[kaβajo] and *deber* [deβer]. Similarly, /w/ became a labiodental or bilabial fricative: see e.g. Lat. *cauerna(m)* ('cavern'), *lauare* ('to wash') > It. *caverna* [kaˈvɛrna], *lavare* [laˈvare]; Fr. *caverne* [kavɛʀn], *laver* [lave]; Sp. *caverna* [kaβerna], *lavar* [laβar].[9]

As far as initial and post-consonantal positions are concerned, the merger was not equally widespread in the Romance languages. In general, as mentioned above, the two phonemes remained distinct in these environments: in word-initial position, for example, see Lat. *bucca(m)* ('mouth') and *uacca(m)* ('cow') > It. *bocca* [ˈbokka], *vacca* [ˈvakka]; Fr. *bouche* [buʃ], *vache* [vaʃ]. In post-consonantal position, see Lat. *malua(m)* ('mallow') and *herba(m)* ('grass') > It. *malva* [ˈmalva], *erba* [ˈɛrba]; Fr. *mauve* [mov], *herbe* [ɛʀb]. There are significant exceptions to this trend: in particular, in this environment, the two phonemes merged in Spanish, in Southern Italy (though with a higher degree of variation) and, significantly for the analysis presented in this paper, in Sardinia.

In Spanish, the result of the merger is one phoneme with two allophones: [b] post-nasally or at the beginning of a breath group, and [β] everywhere else.[10] Therefore, for example, the outcomes of the already mentioned Lat. *bucca(m)* and *uacca(m)* are *boca* [boka] and *baca* [baka], respectively; the same holds for Lat. *malua(m)* and *herba(m)* > Sp. *malva* [malβa], *herba* [jerβa].

As far as Southern Italy is concerned, it is particularly interesting to remark the sound change /w/ > /β/ in Southern Lucania[11] and the change /b/ > /v/ in e.g. *vattere* [ˈvattere] < Lat. *battere* ('to beat'), *vocca* [ˈvokka] < Lat. *bucca(m)* ('mouth').

Finally, in Sardinia /w/ evolved to *b* in most varieties, in initial position as well. Therefore, from Lat. *uacca(m)* the Sardinian outcome found already e.g. in the *Carta* of Benedetta de Lacon (1225 AD) is the plural form *baccas* ('cows'); see also Lat. *uilla(m)* > Sard. *billa* ('residential area'), found already e.g. in the *Carta* of the bishop of Suelli (c. 1215).[12] According to Wagner, the only variety where word-initial *v-* and *b-* did not merge into *b-* was apparently the one spoken in Bitti, in Central Sardinia, which is therefore traditionally thought to show a more archaic linguistic stage.[13] Just to give an example, whereas in this area the outcome of Lat. *bene* is *bène* ('well'), the outcome of Lat. *ui(gi)nti* is *vínti* ('twenty')[14].

---

[9] Adams 2013, 187; see also Herman 2000a, 46. Unless otherwise specified, the examples quoted in this paragraph are from Alkire, Rosen 2010, 32–33.

[10] Alkire, Rosen 2010, 32.

[11] Rohlfs 1966, 227.

[12] Blasco Ferrer 2017a, 129.

[13] Wagner 1984, 165–166.

[14] See also Lupinu 2000, 52, Lupinu 2003, 60-61 and Mensching, Remberger 2016, 274.

1.2. Inscriptional evidence

Given the different outcomes of the two phonemes in the Romance languages, it is interesting to note that the analyses carried out by Barbarino,[15] Herman[16] and, more recently, by Adamik[17] on the <B>/<V> confusions in Latin inscriptions show an unequal diffusion of the phenomenon through the Empire.

Firstly, Barbarino's survey is largely based on the Latin Christian inscriptions collected in Ernst Diehl's *Inscriptiones Latinae Christianae Veteres* (mainly 3rd–7th century AD) and the following areas are examined: Britain, Balkans, Dalmatia, Northern Africa, Spain (Baetica, Lusitania, Tarraconensis), Gaul (Lugudunensis, Narbonensis), Rome and Italy (Northern, Central and Southern). The rate of confusion between <B> and <V> is expressed as a percentage against the corresponding correct spellings, i.e. the number of instances of <B> for Classical Latin /b/ and <V> for Classical Latin /w/. The results of the analysis show a higher incidence of the merger in Rome (37%) and Southern Italy (30%), whereas the lowest rate of confusion is found in Lusitania, Gallia Narbonensis, Baetica and in the Balkans (<5%).[18]

The survey carried out by József Herman in 1965 yields similar results. Unlike Barbarino's approach, the rate of error is expressed as a percentage of the total number of consonantal misspellings found in each area. The inscriptions date back mostly to the 5th–6th century AD, and the regions examined are Spain, Southern Gaul (i.e. Lyon and Vienne), Northern Italy (near Bergamo and Milan, including North-Western Italy), Southern Italy (*Regio I*, *Campania*; *Regio II*, *Apulia* and *Calabria*; *Regio III*, *Lucania*; *Regio IV*, *Samnium*), Rome and Dalmatia. The highest rate of error is found in Southern Italy, with 26.7% of misspellings; the error rate is high in Rome (23.6%) and Dalmatia (17%) as well. On the contrary, the frequency of the confusions is very low in Southern Gaul (<3%) and Spain (10%).

The more recent examination by Adamik focuses on 18 provinces included in the *Computerized Historical Linguistic Database of Latin Inscriptions of the Imperial Age*.[19] The areas analyzed are the following: Lusitania, Baetica, Hispania Citerior, Narbonensis, Lugudunensis, Aquitania, Belgica, Britannia, Noricum, Dalmatia, Pannonia Inferior and Superior, Dacia, Moesia Inferior and Superior, Germania Superior, Venetia et Histria, Apulia et Calabria. The rate of error is expressed as a percentage against the total number of consonantal misspellings

[15] Barbarino 1968.
[16] Herman 1965.
[17] Adamik 2017b.
[18] Barbarino 1978, 151-154; for a critical discussion see also Adamik 2017a, 19-21.
[19] Henceforth *LLDB*.

for each region; moreover, the inscriptions are divided into two chronological periods, i.e. before and after 300 AD. The results of the survey show an unequal distribution of the phenomenon across the Empire. For the Early period (before 300 AD), the provinces which show the lowest rate of confusion are Belgica, Britannia, Baetica, Pannonia Inferior and Dacia (<2%); in contrast, the confusion is more frequent in Venetia-Histria (7%), Moesia Superior (18%) and Apulia-Calabria (34%). For the later period, the merger became more frequent in each area: the only exceptions are Belgica and Moesia Inferior, with a constant amount of confusions.

The surveys illustrated above are remarkable and innovative. However, despite the peculiar outcome of Classical Latin /b/ and /w/, the phenomenon has not yet been systematically analyzed in one crucial area: the province of Sardinia. As shown above, Barbarino did not examine the data from the island separately from the other inscriptions from Italy; Herman excluded Sardinia from the areas under analysis; finally, Sardinia was not in the *LLDB* database, and therefore not included in Adamik's survey.

It is worth mentioning, however, the only two well-known analyses carried out so far on the Sardinian material, i.e. the ones by József Herman[20] and Giovanni Lupinu.[21]

Herman's survey takes into account Christian inscriptions from different regions of the Empire, focusing on the amount of vocalic and consonantal misspellings, among which are listed <B>~<V> confusions as well. The error rate is calculated as a percentage of the total number of misspellings found in each corpus. As far as Sardinia is concerned, the results of his analysis show a sheer number of <B>~<V> confusions in the inscriptions from the island, which corresponds to 62% of the total consonantal misspellings, mostly involving the use of <B> to represent /w/.

Finally, Lupinu's examination on Christian inscriptions from Sardinia yields similar results. According to the author, the most common type of misspellings found in Latin inscriptions from Sardinia involves the confusion between <B> and <V>, and occurrences of the phenomenon are found at least from the 1st century AD. Most of them, however, date back from the 3rd century AD onwards.

The examinations illustrated so far are noteworthy and yield interesting results. Therefore, we believe that a more complete quantitative analysis of the phenomenon in Latin inscriptions from Sardinia could shed light on the peculiarities of the change in the island. In particular, it is important to provide per-

[20] Herman 2000b.
[21] Lupinu 2000.

centages which will enable us to compare the results with those offered by Barbarino and Adamik, illustrated above, thus examining the diatopic variation across the Empire. In order to be able to perform such analysis, we built an annotated epigraphic corpus containing all the available inscriptions from Sardinia, where all the vocalic and consonantal misspellings have been manually retrieved and classified, as will be shown in the following paragraph.

2. The corpus

2.1. Selection of the inscriptions

The analysis presented in this paper has been conducted on an annotated epigraphic corpus which gathers Latin inscriptions from Sardinia; when available, the dating of the texts ranges between the 1st century BC and the 7th century AD. The corpus is provided with extra- and metalinguistic information which allows us to analyze spelling (and possibly phonetic-phonological) variants in Sardinian inscriptions and to interpret them with reference to variables such as the dating and the provenance place of the texts. It will be part of the *CLaSSES* database (Corpus for Latin Sociolinguistic Studies on Epigraphic textS),[22] developed at the Department of Philology, Literature and Linguistics of the University of Pisa, which gathers various types of non-literary Latin texts (inscriptions, writing tablets, letters) of different periods and provinces of the Roman Empire.[23]

The epigraphic texts from Sardinia have been selected through the careful examination of the main collections of Latin inscriptions from the island: *Corpus Inscriptionum Latinarum X* (*fasc. I*, section *Pars posterior inscriptiones Siciliae et Sardiniae comprehendens*);[24] *Ephemeris Epigraphica VIII* (section *Additamenta ad Corporis vol. IX et X*),[25] the two volumes by Giovanna Sotgiu, *Iscrizioni Latine della Sardegna (Supplemento al Corpus Inscriptionum Latinarum, X e all'Ephemeris Epigraphica, VIII)*,[26] and the more recent collection by Sotgiu.[27] Since the reading of the inscriptions can be problematic,[28] we

[22] The database is available online: http://classes-latin-linguistics.fileli.unipi.it/.

[23] At the moment, the database contains more than 1200 inscriptions, mainly from Rome and Central Italy, 200 ink-written tablets from Vindolanda and 219 letters from the North-African and Near-East areas. For a more detailed illustration of CLaSSES, see Marotta 2015, 2016 and De Felice, Donati, Marotta 2015.

[24] Henceforth CIL X.

[25] Henceforth Eph.Ep. VIII.

[26] Henceforth ILSard I and ILSard II.

[27] Sotgiu 1988; henceforth ANRW.

[28] On this issue see e.g. Kruschwitz 2010 and Clackson 2011.

revised and checked the texts offered by the various editions, examining the photographs of the inscriptions, when available.

Among the large number of texts available for this province (ca. 2.000), we excluded the inscriptions considered to be not relevant for linguistic analysis. In particular, the following texts have not been included in the corpus:

- inscriptions composed only of single letters and initials;
- fragmentary texts;
- inscriptions entirely in other languages (e.g. Greek).

2.2. Composition

The corpus contains 1.168 inscriptions, for a total number of 14.212 tokens. As shown in Figure 1, the areas where the highest number of inscriptions was found, indicated by bigger pins in the map, are mainly located along the coast, the so-called 'Romània costiera', where the main cities were built, such as the capital *Carales* (nowadays Cagliari) and several colonies (such as *Turris Libisonis*, nowadays Porto Torres).[29]

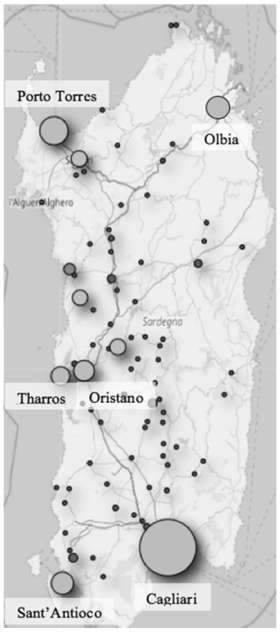

Figure 1. *Geographical distribution of the inscriptions.*

[29] Mastino 2002, 63; see also Mastino 1993, 463 ff.

As far as the dating of the texts is concerned, we annotated this information only when available and based on non-linguistic criteria, in order to avoid the danger of circularity.[30] Since the dating of the inscriptions can be highly problematic,[31] this information was available only for 616 texts (9.379 tokens). As mentioned above, the dated inscriptions cover a broad time span, from the 1st century BC to the 7th century AD (Table 1).

Dating	Number of inscriptions
1st BC–1st AD	9
1st–3rd AD	474
4th–5th AD	101
6th–7th AD	32
Total	616

Table 1. *Dating of the inscriptions.*

In particular, the majority of the texts date back to the 1st–3rd century AD, before the end of the Roman influence and the conquest of the island by the Vandals in 455 AD[32] and, subsequently, their defeat and the start of the Byzantine domination in 533–534 AD.[33]

2.3. Annotation

All the selected texts have been digitalized; the entire corpus was tokenized and an index was created. In this way, each token (i.e. a string of contiguous characters between two spaces) is univocally associated to a token-ID containing the epigraphic collection and the number of the inscription, as well as the position of the token within the text.

[30] On this issue see Penney 2011, 221.

[31] For a detailed insight on the main issues connected to the dating of epigraphic texts, see Cooley 2012, 398 ff.

[32] Blasco Ferrer 1984, 84.

[33] Spanu 2005, 506-507; Blasco Ferrer 2017b, 86.

The innovative aspect of our corpus is the linguistic analysis of each token, which at the moment focuses on the phonetic aspects of the language. The deviant spellings, i.e. spellings which do not conform to the Classical norms, were manually retrieved. It is important to highlight that we did not annotate the misspellings presumably due to extralinguistic reasons, such as the state and the dimension of the support. For example, we did not label as deviant spellings instances of omission of final consonants due to a lack of space on the support; similarly, we did not annotate omissions of graphemes at the end of the line if the support was damaged. Subsequently, following the same criteria adopted for CLaSSES, all misspellings were classified according to the type of variation phenomena that distinguish them from the corresponding classical equivalents. The phenomena identified involve vowels, consonants and morpho-phonology (i.e. involving morphological endings). For each level, we labelled the different types of linguistic phenomena through specific labels. For example, as far as vowels are concerned, we identified the following ones: alternations, deletion, epenthesis, monophthongization.

Finally, each token was annotated with extralinguistic information regarding the place of provenance and the dating of each inscription (when available). In this way, it is possible to relate these variables to the graphic variants identified. This corpus will enable us to shed light on the characteristics of the variety of Latin spoken in Sardinia, as well as to examine several phenomena connected to the development of the Sardinian varieties. In particular, it has been possible to cast some light on the diachronic evolution of the phenomenon of <B>/<V> confusion in Sardinian Latin from the 1st century BC to the 7th century AD, as will be shown in the following paragraph.

3. <B>/<V> confusion in Sardinia

The analysis presented in this paper focuses on the instances of confusion between <B> and <V> in the 616 dated inscriptions included in the corpus (9.379 tokens, see §2.2). We decided to limit our analysis to dated inscriptions with the aim to trace the diachronic development of the sound change in the island. Moreover, as will be shown in the following paragraphs, this methodology enabled us to obtain data comparable to those illustrated in Barbarino and Adamik's surveys, examining both the diachronic and the diatopic levels of variation. However, we do not exclude the possibility of extending our analysis to undated inscriptions in the near future.

3.1. Error rate

We identified 120 instances of <B> for <V> and <V> for <B> in Sardinia, which amount to 7.4% when calculating the error rate against the corresponding correct spellings in the corpus, as shown in Table 2.

	Tokens	**%**
<B>/<V> confusions	120	7.4%
Correct spellings	1510	92.6%
Total	1630	100%

Table 2. *<B>/<V> confusion in Sardinia.*

It is interesting to note that the vast majority of the instances involves the use of the grapheme for the bilabial stop in place of the one for the semivowel. As shown in Tables 3 and 4, we identified 111 instances of <B> for <V> (11.4%) against 9 of the reverse type (1.4%).

Grapheme	**Tokens**	**%**
<B>	111	11.4%
<V>	859	88.6%
Total	970	100%

Table 3. *Tokens showing <B> for <V> in Sardinia.*

Grapheme	**Tokens**	**%**
<V>	9	1.4%
<B>	655	98.6%
Total	664	100%

Table 4. *Tokens showing <V> for <B> in Sardinia.*

Though this trend is not exclusive of Sardinia,[34] the use of <B> for <V> in the inscriptions could be interpreted as a cue for the strengthening of the phoneme in the island. These data seem to anticipate the Romance development of the Sardinian variety, where /w/ (graphically represented as <V>) evolved to *b*, especially in word-initial position (see §1.1). Moreover, our inscriptional data seem

[34] See Adams 2013, 183–190.

coherent with later written documents from Sardinia, which show that the process had already begun by the 11th–13th centuries: the grapheme representing the bilabial stop was already used in place of the one for the semivowel in the first Sardinian administrative documents, such as the *Condaghe di S. Pietro di Silki* and the *Carte Volgari dell'Archivio Arcivescovile di Cagliari*.[35]

3.2. Comparison with other areas

In order to verify whether these trends are characteristic of Sardinia, our analysis needs to be compared with inscriptional data from other regions of the Empire. In particular, our percentages are comparable with the ones given by Barbarino[36] in his examination of the phenomenon across the Roman Empire, where the same methodology for calculating the error rate is adopted (i.e. by expressing it as a percentage of the corresponding correct spellings). In order to obtain comparable data, the analysis illustrated in this paragraph is limited to the inscriptions belonging to the time frame examined by Barbarino (mainly 3rd–7th century AD). The results of our survey are presented in Table 5.

	Tokens	**%**
<B>/<V> confusions	73	20.1%
Correct spellings	291	79.9%
Total	364	100%

Table 5. *<B>/<V> confusion in Sardinia (3rd–7th century AD).*

The percentage of tokens showing the phenomenon in Sardinia is 20.1%. It is therefore a remarkable amount of misspellings: there is much more confusion than in Northern and Central Italy, where according to Barbarino[37] 8% of the instances are misspelt; the same holds for *Hispania Tarraconensis*, with 9% of misspellings. However, it is worth noting that the error rate in Sardinia is still considerably lower than other regions, particularly Rome (35%) and Africa (32%), to which the linguistic features of Sardinia are frequently associated.[38] However, it is important to highlight that the inscriptions included in our corpus cover a broader time span (1st century BC – 7th century AD; see §2.2). Therefore, in order to examine the diachronic evolution of the change, we decided to take

[35] Wagner 1984, 163-165.
[36] Barbarino 1978.
[37] Barbarino 1978, 151-154.
[38] See e.g. Fanciullo 1992; Lupinu 2000, 53; Lupinu 2003; Lorenzetti, Schirru 2010; Loporcaro 2015, 48 ff.

into account all the dated inscriptions at our disposal. As shown in Table 6, we subdivided the inscriptions into three time frames which correspond to the different dominations in the island: 1st century BC–3rd century AD (before the weakening of the Roman influence on the island), 4th–5th centuries AD (from the Vandal domination to the Byzantine one); 6th century AD (from the Byzantine conquest onwards).

	1st BC–3rd AD	**4th–5th AD**	**6th–7th AD**
<B>/<V> confusions	47	40	33
%	3.7%	15%	34%
Correct spellings	1219	227	64
%	96.3%	85%	66%
Total	1266	267	97

Table 6. *Diachronic evolution of the confusion between <B> and <V> in the corpus.*

As shown above, we observe a growing confusion between <B> and <V>, especially from the 4th century onwards, as Adamik[39] pointed out for other areas (see §1.2). The instances of the phenomenon amount to 3.7% in the first centuries of the Empire; from the 4th century onwards, the percentage rises to 15% and finally doubles after the 6th century (34%). As mentioned above, the majority of the occurrences involve the use of <B> for <V>, except for 4 instances of *vene* (for *bene*) and 5 of *Olvie/Olviae* (for *Olbiae*), belonging to the 2nd–3rd century AD.[40]

Again, those data are coherent with the Romance evolution of /w/ in Sardinia: the phoneme tends to be strengthened, especially in later inscriptions, until the percentage of occurrences of <B> for <V> amounts to one-third in the 6th–7th century AD.

3.3. Literacy

As shown in the previous paragraphs, the methodology of calculating the error rate as a percentage against the total number of correct spellings can be useful for the comparison of corpora of different size. Therefore, following this method is particularly useful when comparing a relatively small epigraphic corpus such

[39] Adamik 2017b.

[40] The relevant inscriptions are the following: CIL X 7619, 7833, 7833, 8027, Eph. Epigr. VIII 722, 774, 776, 798 and ANRW B92.

as the Sardinian one with much bigger corpora (e.g. the corpus of inscriptions from Rome).

However, this methodology does not account for the level of literacy of those involved in the crafting of the inscriptions. The literacy level is instead an important variable when examining spelling variations in epigraphic sources: if the level of literacy of the writers is high, the lack of misspellings in the inscriptions could not be immediately taken as a reflection of their pronunciation, since the latter could be obscured by their knowledge of the classical norms.

In particular, as far as our corpus is concerned, it is worth noting that half of the inscriptions which do not show <B>/<V> confusions (i.e. where <B> and <V> stand for Classical Latin /b/ and /w/, respectively) do not show any misspelling at all (see Table 7).

	N. of inscriptions	**%**
Presence of other types of misspellings	212	45.2%
Absence of other types of misspellings	257	54.8%
Total	469	100%

Table 7. *Percentage of inscriptions showing misspellings other than <B>/<V> confusions.*

Therefore, in half of the inscriptions, we cannot exclude that the lack of graphic confusion between <B> and <V> could be attributed to a high level of education of the writers. In other words, in these cases, the spelling could not reflect the pronunciation of those involved in the crafting of the inscriptions.

In order to address this issue, we adopted the methodology proposed by Adamik,[41] which follows the one initially proposed by József Herman.[42] Therefore, we calculated the percentage of <B>/<V> confusions against the total number of consonantal misspellings occurring in the dated inscriptions in the corpus.[43]

Given the methodological approach behind our annotation (see §2.3), the deviant spellings analysed are not due to extralinguistic reasons, and could be considered to be phonetic spellings. Therefore, by taking into account only the inscriptions with at least one misspelling, we are only analysing texts produced by speakers which had uncertainties in (at least) one other point of the language. As

[41] Adamik 2016; 2017b.

[42] With reference to Sardinia, see e.g. Herman 2000b.

[43] The consonantal misspellings taken into account are the following: deletion of consonants (final *-s*, *-m*, *-t* and nasals, such as *ns*>*s*, etc.), insertion of consonants, assimilation, dissimilation, non-etymological gemination, degemination, confusion between voiced and voiceless stops, loss of aspiration.

a consequence, the absence of <B>/<V> confusions in these texts could be reasonably due to a correspondence between the classical norms and the pronunciation of the writers/inscribers.

As shown in Table 8, the relative frequency of the phenomenon is quite high, and amounts to one-third of the total consonantal misspellings (33.1%).

	Tokens	**%**
<B>/<V> confusions	120	33.1%
Other consonantal misspellings	243	66.9%
Total	363	100%

Table 8. *Percentage of <B>/<V> confusions against the total number of consonantal misspellings.*

This approach is particularly useful since our data can be compared with the extensive examination made by Adamik,[44] who analysed the frequencies of <B>/<V> confusions in proportion to all other consonantal mistakes for 18 provinces of the Empire,[45] both for the 'early period' (before 300 AD) and for the 'late' one (after 300 AD). In this way, we can examine both the diatopic and diachronic dimensions of the variation.

In order to obtain comparable data, we divided the inscriptions from Sardinia into two broad periods, i.e. before and after the beginning of the 4th century AD, as shown in Table 9 and Figure 2.

	Before 300 AD		**After 300 AD**	
	Tokens	%	Tokens	%
<B>/<V> confusions	47	19.9%	73	57.5%
Other consonantal misspellings	189	80.1%	54	42.5%
Total	236	100%	127	100%

Table 9. *Percentage of <B>/<V> confusions against the total number of consonantal misspellings (before and after 300 AD).*

[44] Adamik 2017b.

[45] As mentioned in §1.2, Sardinia was not examined in this study.

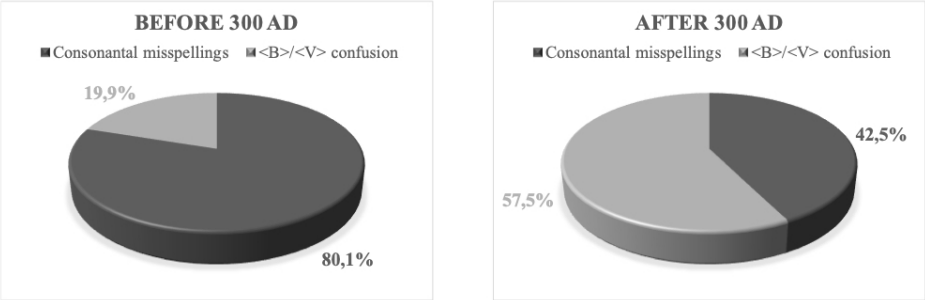

Figure 2. *Percentage of <B>/<V> confusions against the total number of consonantal misspellings (before and after 300 AD).*

The results of the examination show a widespread graphemic confusion between <B> and <V> in Sardinia, especially from the 4th century AD, confirming the general trend observed by Adamik. In the Early period, the relative frequency of the phenomenon amounts to 19.9% and then rises dramatically in the Later period, reaching 57.5%. It is also interesting to note that <B>/<V> confusions are particularly frequent in the island in comparison with the regions analyzed by Adamik. For the Early period, only Apulia-Calabria shows a higher rate of confusion than Sardinia, with 34% of misspellings. For the Later period, Sardinia surpasses even Apulia-Calabria, where the percentage of misspellings amounts to 51%.

This picture is even clearer if we adopt our periodization based on the different dominations on the island, as shown in Table 10 and Figure 3.

	1st BC–3rd AD		4th–5th AD		6th–7th AD	
	Tokens	**%**	**Tokens**	**%**	**Tokens**	**%**
<B>/<V> confusions	47	19.9%	40	50.6%	33	68.8%
Other consonantal misspellings	189	80.1%	39	49.4%	15	31.2%
Total	236	100%	79	100%	48	100%

Table 10. *Percentage of <B>/<V> confusions against the total number of consonantal misspellings.*

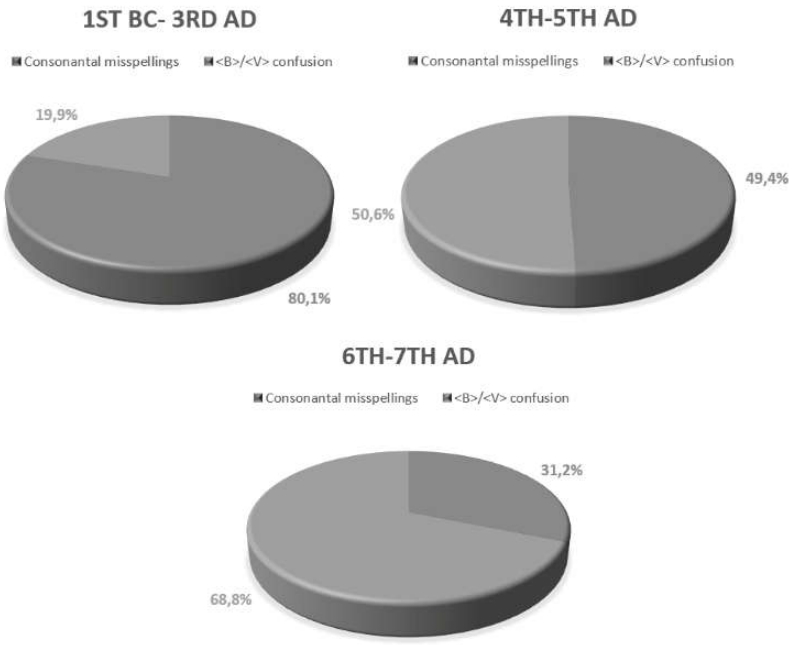

Figure 3. *Percentage of <B>/<V> confusions against the total number of consonantal misspellings.*

During the first time span analysed, i.e. between the 1st century BC and the 3rd c. AD, the relative frequency of the phenomenon is 19.9%. The amount of confusions then increases significantly with the beginning of the 4th century, when it involves the majority of the occurrences of <B> and <V> (50.6%); finally, the phenomenon reaches two-thirds of the instances after the 6th century (68.8%).

3.4. Position in the word

The last variable we took into account is the position in the word of the phoneme showing the confusion. This analysis is particularly interesting given the Sardinian outcome of *b-* and *v-* in most varieties, which both merged into *b-* (as shown in §1.1). For this reason, we decided to calculate the relative frequency of the phenomenon for each position (word-initial, intervocalic, and post-consonantal). Our results are summarized in Table 11.

Position	Tokens	%
#_	60	50%
V_V	40	33.3%
C_V	20	16.7%
Total	120	100%

Table 11. *Position of the phonemes showing <B>/<V> confusion.*

As shown in the Table, our corpus displays a variety of positions targeted by the phenomenon: initial (e.g. *bixit* for *vixit*), intervocalic (e.g. *requiebit* for *requievit*) and post-consonantal (e.g. *Silbius* for *Silvius*). However, the confusion seem to be particularly frequent in word-initial position (60 tokens out of 120, i.e. 50%). Since most of the instances point to a strengthening of the phoneme (see §3.1), our data seem to anticipate the Romance outcome of the Sardinian variety. However, these preliminary results need to be completed by a finer-grained analysis which takes into account external sandhi rules. Therefore, in the near future we plan to verify whether the instances of <B> for <V> in initial position are actually in intervocalic position in phonosyntax (e.g. in the sequence *qui bixit*), or in word-initial and postconsonantal position, following the methodology proposed by Adamik.[46]

4. Conclusions

The analysis illustrated so far allows us to draw some preliminary conclusions about the confusion between <B> and <V> in Latin inscriptions from Sardinia. First of all, the creation of a complete corpus which gathers the available inscriptions from this area enabled us to observe a widespread graphemic confusion between <B> and <V> in the island from early times. This trend is evident through the comparison with Barbarino's data, and it is even clearer when expressing the error rate as a percentage of the total number of consonantal misspellings. Following this methodology, thanks to the wide chronological range of the inscriptions included in the corpus, we found a constant increase of the phenomenon, especially from the 4th century AD onwards.

The incidence of the confusion in the island is even more evident when comparing the data from Sardinia with the areas examined by Adamik.[47] Among those regions, only Apulia-Calabria shows a higher rate of confusion for the Early period, whereas Sardinia shows the highest rate for the Later period.

[46] Adamik 2017a, 24 ff.
[47] Adamik 2017b.

Finally, most of the instances involve the use of the grapheme for the bilabial stop in place of the one for the semivowel. Though this distribution is not a unique characteristic of Sardinian inscriptions, it might point to a strengthening of the phoneme, especially in initial position, where the use of <B> for <V> seems to be particularly frequent, coherently with the development of most Sardinian varieties.

To conclude, by taking into account variables such as dating, place of provenance and literacy level we were able to examine both the diatopic and diachronic levels of variation. This approach enabled us to cast some light on the evolution of Latin /b/ and /w/ in the island and on the Romance evolution of the Sardinian varieties.

Bibliography

Adamik 2016 = Adamik, B.: The frequency of syncope in the Latin of the Empire: A statistical and dialectological study based on the analysis of inscriptions. In: Poccetti, P. (ed.): *Latinitatis Rationes, Descriptive and Historical Accounts for the Latin Language*. Berlin–Boston, 3-21.

— 2017a = Adamik, B.: Potential Greek influence on the Vulgar Latin sound change [b]>[β]: Dialectological evidence from inscriptions. *Acta Antiqua Academiae Scientiarum Hungaricae* 57, 11-33.

— 2017b = Adamik, B.: On the Vulgar Latin merger of /b/ and /w/ and its correlation with the loss of intervocalic /w/: Dialectological evidence from inscriptions. *PALLAS* 103, 25-36.

Adams 2013 = Adams, J. N.: *Social Variation and the Latin Language*. Cambridge.

Alkire, Rosen 2010 = Alkire, T., Rosen, C.: *Romance Languages: a Historical Introduction*. Cambridge.

Allen 1965 = Allen, W. S.: *Vox Latina: A Guide to the Pronunciation of Classical Latin*. Cambridge.

Barbarino 1978 = Barbarino, J.: *The evolution of the Latin /b/-/u/ merger: A quantitative and comparative analysis of the B-V alternation in Latin inscriptions*. Chapel Hill.

Blasco Ferrer 1984 = Blasco Ferrer, E.: *Storia linguistica della Sardegna*. Tübingen.

Blasco Ferrer 2017a = Blasco Ferrer, E.: Sardo antico. In: Blasco Ferrer, E., Koch, P., Marzo, D. (*eds.*): *Manuale di linguistica sarda*. Berlin–Boston, 119-136.

Blasco Ferrer 2017b = Blasco Ferrer, E.: Il latino e la romanizzaizone. In: Blasco Ferrer, E., Koch, P., Marzo, D. (*eds.*): *Manuale di linguistica sarda*. Berlin–Boston, 85-103.

Campanile 1971 = Campanile, E.: Due studi sul latino volgare. *L'Italia Dialettale* 34, 1-64.

Clackson 2011 = Clackson, J.: Latin Inscriptions and Documents. In: Clackson, J. (2011, *ed.*): *A Companion to the Latin Language*. Malden, Mass., 29-39.

Cooley 2012 = Cooley, A. E.: *The Cambridge Manual of Latin Epigraphy*. Cambridge.

De Felice, Donati, Marotta 2015 = De Felice, I., Donati, M., Marotta, G.: CLaSSES: A new digital resource for Latin epigraphy. *Italian Journal of Computational Linguistics* 1(1), 119-130.

Fanciullo 1992 = Fanciullo, F.: Un capitolo della Romania submersa: il latino africano. In: *Actes du XVIIIe Congrès International de Linguistique et de Philologie Romanes (Trier 1986), I.* Tübingen, 162-187.

Herman 1965 = Herman, J.: Aspects de la différenciation territoriale du Latin sous l'Empire. In: Herman, J. (1990): *Du latin aux langues romanes. Études de linguistique historique*. Tübingen, 10-28.
— 2000a = Herman, J.: *Vulgar Latin. Translated by Roger Wright.* Pennsylvania.
— 2000b = Herman, J.: Differenze territoriali nel latino parlato dell'Italia: un contributo preliminare. In: Herman, J., Marinetti, A. (2000, *eds.*): *La preistoria dell'italiano. Atti della Tavola Rotonda di Linguistica Storica. Università Ca' Foscari di Venezia 11-13 giugno 1998*. Tübingen, 123-135.
Kruschwitz 2010 = Kruschwitz, P.: Romanes eunt domus. Linguistic aspects of the sub-literary Latin in Pompeian wall inscriptions. In: Evans, T.V., Obbink, D. D. (2010, *eds.*): *The Language of the Papyri*. Oxford–New York, 156-170.
Loporcaro 2015 = Loporcaro, M.: *Vowel length from Latin to Romance*. Oxford.
Lorenzetti, Schirru 2010 = Lorenzetti, L., Schirru, G.: Un indizio della conservazione di /k/ dinanzi a vocale anteriore nell'epigrafia cristiana di Tripolitania. In: Tantillo, I., Francesca, B. (*eds.*): *Leptis Magna. Una città e le sue iscrizioni in epoca tardoromana*. Cassino, 303-311.
Lupinu 2000 = Lupinu, G.: *Latino epigrafico della Sardegna. Aspetti fonetici*, Officina linguistica 3. Nuoro.
— 2003 = Lupinu, G.: Tra latino epigrafico e sardoromanzo: sulla datazione di alcuni sviluppi fonetici. *VERBUM* 5(1), 59-68.
Marotta 2015 = Marotta, G.: Talking stones. Phonology in Latin inscriptions. *Studi e Saggi Linguistici* 53(2), 39-64.
— 2016 = Marotta, G.: Sociolinguistica storica ed epigrafia latina. Il corpus CLaSSES I. *Linguarum Varietas* 5, 145-160.
Mastino 1993 = Mastino, A.: Analfabetismo e resistenza: geografia epigrafica della Sardegna. In: Calbi, A., Donati, A., Poma, G. (1993, *eds.*): *L'epigrafia del villaggio*. Faenza, 457-536.
— 2002 = Mastino, A.: La Sardegna romana. In: Brigaglia, M., Mastino, A., Ortu, G. G. (2002, *eds.*): *Storia della Sardegna. 1: dalla Preistoria all'età bizantina.* Roma–Bari, 52-92.
Mensching, Remberger 2016 = Mensching, G., Remberger, E.: Sardinian. In: Ledgeway, A., Maiden, M. (2016, *eds.*): *The Oxford guide to the Romance languages* (Oxford guides to the world's Languages). Oxford, 270-291.
Penney 2011 = Penney, J.: Archaic and old Latin. In: Clackson, J. (2011, *ed.*): *A Companion to the Latin Language*. Malden, Mass., 220-235.
Rohlfs 1966 = Rohlfs, G.: *Grammatica storica della lingua italiana et dei suoi dialetti I: Fonetica*. Torino.
Ségéral, Scheer 2008 = Ségéral, P., Scheer, T.,: The Coda Mirror, stress and positional parameters. In: Brandão de Carvalho, J., Scheer, T., Philippe, S. (2008, *eds.*): *Lenition and Fortition*. Berlin–Boston, 483-518.
Spanu 2005 = Spanu, P.G.: L'età vandalica. In: Mastino, A. (2005, *ed.*): *Storia della Sardegna Antica.* Nuoro, 499-509.
Sturtevant 1920 = Sturtevant, E. H.: *The Pronunciation of Greek and Latin*. Chicago.
Väänänen 1981 = Väänänen, V.: *Introduction au latin vulgaire*. Paris.
Wagner 1984 = Wagner, M. L.: *Fonetica storica del sardo, ristampa con traduzione, introduzione e appendice di Paulis, G.* Cagliari.

List of Abbreviations

ANRW = Sotgiu, G. (1988): L'epigrafia latina in Sardegna dopo il C.I.L. X e l'E.E. VIII. In: Temporini, H., Haase, W. (1988): *Aufstieg und Niedergang der römischen Welt (ANRW), II: Principat, 11.1*. Berlin–New York.

CIL X = Mommsen, T.: *Corpus Inscriptionum Latinarum*, vol. X *Inscriptiones Bruttiorum, Lucaniae, Campaniae, Siciliae, Sardiniae latinae*, fasc. I, section *Pars posterior inscriptiones Siciliae et Sardiniae comprehendens*.

Eph.Epigr. VIII = Ihm, M.: Additamenta ad Corporis vol. IX et X. In: *Ephemeris epigraphica. Corporis inscriptionum Latinarum supplementum, edita iussu Instituti archaeologici Romani, volume VIII*. Berolini, 1-22.

ILSard I = Sotgiu, G.: *Iscrizioni Latine della Sardegna (Supplemento al Corpus Inscriptionum Latinarum, X e all'Ephemeris Epigraphica, VIII), I*. Padova.

ILSard II = Sotgiu, G.: *Iscrizioni Latine della Sardegna, II: Instrumentum domesticum. I. Lucerne*. Padova.

Keil Gramm. Lat. = Keil, H.: *Grammatici Latini*. Lipsia.

(ISSN 0418-453X)