

Linguistic awareness in Rome: Contrasting theory and practice in light of the use of *h* in inscriptions

DÁNIEL SERES^{1,2*} 

¹ European Center of Political Philosophy, Mathias Corvinus Collegium, Hungary

² Department of History and Social Studies, Ludovika University of Public Service, Hungary

Received: January 16, 2025 • Revised manuscript received: August 29, 2025 • Accepted: September 16, 2025



© 2025 The Author(s)

ABSTRACT

This paper takes a closer look at the changes in Roman spelling standards through the Roman inscriptions of the Imperial period based on the data of the LLDB-Database. The phoneme marked with *h* disappeared from the Latin language system at the end of the Republican era, but its use in spelling remained, and thus became a prime indicator of literacy. The two most important objectives of the current study are the following: to provide evidence that the proportion of aspiration errors more than doubled from the pagan to the Christian era; to argue that the comparison of correct and incorrect forms is a valid measure of the level of contemporary spelling, thus also of literacy, and their change.

KEYWORDS

Latin linguistics, sociolinguistics, aspiration, linguistic awareness, Roman inscriptions

1. INTRODUCTION

The aim of the present paper is twofold.¹ On the one hand, it provides a summary of the history of the Latin *h* phoneme. On the other hand, the analysis examines possible changes in Roman spelling standards through the Roman inscriptions of the imperial period.

* Corresponding author. E-mail: seres.daniel@mcc.hu

¹I would like to thank the two anonymous reviewers for their careful reading of the manuscript and their constructive comments and suggestions. Any remaining mistakes are my own.

The research outlined is based on data extracted from the LLDB-Database (henceforth: LLDB or Database [<https://lldb.elte.hu/>]) – the data analysed here reflect the state of the database on 30 November 2024. LLDB has been developed primarily for dialectological research, but since it records not only linguistic errors but also non-linguistic ones, such as those of orthographical origin, it is also suitable for sociolinguistic research. The study of spelling standards can be seen as a narrow segment of sociolinguistics, since it reflects the linguistic-cultural dimensions and the language proficiency and literacy level of the people erecting the inscriptions.

The study of the use of *h* is a perfect way of measuring and tracking this phenomenon, since the /h/ phoneme disappeared from the Latin language system by the end of the Republican era or the beginning of the Imperial period, but its use in spelling remained, and thus became a prime indicator of literacy or illiteracy. Since the /h/ phoneme no longer existed in the period under study, the conventional distinction between the grapheme and the phoneme *h* is unnecessary – ⟨h⟩ and /h/ –, I consistently use a single *h* in this paper.

The inscriptional corpus of the city of Rome seems also an obvious choice since, the standard language from Varro onwards was Rome-centred, which contrasted the sophisticated, literate – *urbanus* – and the rural, illiterate – *rusticus* – usage of the Latin language. Thus, the theory of Roman linguistic awareness, known from the works of grammarians and rhetoricians, which saw Rome as an abstract ideal to copy, can be contrasted with the practice of the inscriptional corpus, which reflects the everyday language usage of the lower classes. Therefore, also what was transmitted ‘from above’, from the sublime theory of *urbanitas*, into the language of the lower classes in Rome, which were thoroughly exposed to *rusticitas*.

2. THE LETTER *H* AND ROMAN LINGUISTIC AWARENESS

2.1. History of the *h* phoneme in Latin

The continuation of the Indo-European **g^h* in the Italic languages, with the exception of Samnite and Faliscan, is uniformly *h*, although in Latin it began to disappear quite early. On the relationship and interactions of Italic languages and Latin, see Coleman (1986). Assumptions about the Archaic period are uncertain. The standard, recurring example in the previous research is the verb *diribeo*, which was unanimously derived from **dishabeo* by the disappearance of the *h* phoneme after the *s*, and thus the process predates the fourth century BC rhotacism, as mentioned by several scholars (Sommer 1902, 216; Porzio Gernia 1974, 58; Leumann 1977, 174; Allen 1978, 43; Väänänen 1981, 55; Sihler 1995, 159; Meiser 2002, 105; Traina 2002, 49). However, Adamik’s (2009, 137–138) suggestion shakes the foundations of this relative chronology, as it is possible that *diribeo* is a late word formation based on the verb *dirimo*. A similar uncertainty arises with the noun *anser*, which certainly derives from the Indo-European **g^hans*, therefore **hanser* would be its regular continuation in Latin, as described by Sommer (1902, 218), Porzio Gernia (1974, 58), Leumann (1977, 173), Allen (1978, 43), Traina & Perni (1998, 55–56), Meiser (2002, 105), Traina (2002, 52, n. 1), de Vaan (2008, s.v. *anser*), Adamik (2009, 137). However, the word is only recorded since the time of Plautus, so it is completely uncertain when it lost its initial aspiration.

We have more certain information of the Old Latin period, as described by several authors (Sommer 1902, 216; Porzio Gernia 1974, 58–59; Leumann 1977, 174–175; Allen 1978, 43;

Sihler 1995, 159; Traina & Perni 1998, 56; Traina 2002, 49). I have followed Adamik (2015) in the periodisation of Latin. From metrical examples we see that the *h* at the beginning of a word does not affect the elision, nor does it affect the shortening before another vowel (*prehendo* < **prae*hendo or *dēhinc* < **dē*hinc), nor does it form a position after a consonant. On shortening see Sihler (1995, 80). A more complex development is suggested by de Vaan (2008, s.v. *prehendo*). In some intervocalic positions it did not prevent the vowels from merging, cf. *nīl* from *nihil*, *mī* from *mihi*, *nēmo* from **ne*-hemo or *bīmus* from **dui*-himos, see de Vaan (2008 s.v. *hiems*, *hilum*, *homo*). In other words, several examples demonstrate that the pronunciation of *h* became inert in many cases during the Old Latin period. In his research on the history of writing, Adamik (2009, 139, 150–154) has pointed out that the phoneme could not have been completely inert at that time. If this had been the case, it is logical to assume that the reform of the Latin alphabet in the early third century BC would have left the letter *h* to the fate of the letter *z*, which had played a marginal role in Latin – replaced by the newly appearing *g*. Thus, he places the end of pronouncing *h* in the first half of the third century BC.

In the light of recent research, it is possible that the complete inertness of *h* can be dated back to the Classical period (from the second half / end of the second century BC to the end of the second century AD), or even to the beginning of the Classical period. Cser's (2016, 10–11, 2020, 15) suggestion that the imperfect conjugation of the verbs *veho* and *traho* may prove that the phoneme *h* must have been present in some form in Latin, since both verbs are conjugated as consonant-stems, and if the *h* sound had been completely inert, *traho* would have been inflected as *a*-stem and *veho* as *e*-stem. Cser's observation is also innovative because research has shown that the survival of the initial (e.g. *hoc*, *hodie*, *homo*, *hiems* etc.) and intervocalic (e.g. *mihi*, *nihil*, *prehendo*, *cohors* etc.) *h* is clearly due to spelling conventions. It is likely that the various archaizing laws may be behind this. For more on this, see Adamik (2009, 138–139), who also reflects Rudolf Wachter's findings. The verbs *veho* and/or *traho* are also repeatedly mentioned in such lists, cf. Leumann (1977, 173), Väänänen (1981, 55), Adamik (2009, 138). Cser's hypothesis shows, however, that *h* (cf. I consistently use a single *h* in this paper) was not only used as a phonological symbol for hiatus. For this phenomenon a typical example in linguistic history is *ahēnus*, where the presence of the phoneme *h* cannot be etymologically verified and is presumably an Umbrian loanword, cf. de Vaan (2008, s.v. *aes*). As conjugation is also involved in the cases cited, there is a legitimate question about the effect of this surplus on the spelling of the word in question, and whether the letter *h* is used more confidently in such a position. Based on LLDB data, this seems to be the case. Only three data forms refer to the verb *traho* (LLDB-31766 [TRAENS = trahens], LLDB-136209 [TRAATVR = trahatur], LLDB-48619 [TRAIT = trahit]) and these are also very late ones, while we have no data for *veho* and only two examples of its derivatives (LLDB-19623 [VEEMENS = vehemens], LLDB-76471 [VEEMENTE = vehementer]).

2.2. The usage of *h* as an indicator of being cultured

The accelerated weakening of the *h* phoneme was certainly delayed by the Hellenistic influence, which became increasingly strong in the growing Roman Empire from the second half of the second century BC. Three aspirated letters – θ (th), φ (ph), χ (ch) were established in Latin – which were completely alien to the Latin pronunciation system, but which after a relatively short time overturned the Latin rules of pronunciation and spelling, as described by several authors

(Coleman 1963, 11–12; Gignac 1975, 99. n. 1; Leumann 1977, 159–163; Allen 1987, 26–27; Biville 1990, 137–142; Mancini 1990, 9–19; Sihler 1995, 142; Adams 2003a, 109).

The frequent PH ~ F alternation in the Roman epigraphic corpus will not be analysed. The reason for this is that the phenomenon was caused by the Greek aspirated consonants becoming fricatives: /ph/ > /f/, /th/ > /θ/, /kh/ > /x/. These types of errors in fact reflected the exact contemporary Greek pronunciation, and since this sound was pronounced in the same way as the Latin sound /f/, it is quite logical that the letter *f* was used to transcribe these Greek loanwords. This error is so logical that in some cases it even fooled the Roman grammarians of the late imperial period, who gave the correct form with *f* instead of *ph*, e.g. *homfagium non monofagium*, *strofa non stropa*, *amfora non ampora*, cf. *Appendix Probi*, 47, 192, 227; for the new edition see Powell (2007). For this reason, I have included the PH ~ F alternation in all figures as an additional orthographic error – *orthographica alia*, which is not a separate LLDB category –, since it has no direct linguistic connection with the evolution of the Latin *h* phoneme. Cicero and Quintilian recall the appearance of the aspirated consonants as follows:

“Quin ego ipse, cum scirem ita maiores locutos esse, ut nusquam nisi in vocali aspiratione uterentur, loquebar sic, ut pulcros, Cetegos, triumphos, Cartaginem dicerem; aliquando, idque sero, convicio aurium cum extorta mihi veritas esset, usum loquendi populo concessi, scientiam mihi reservavi. Orcivios tamen et Matones, Otones, Caepiones, sepulcra, coronas, lacrimas dicimus, quia per aurium iudicium licet. Burrum semper Ennius, numquam Pyrrhum; vi patefecerunt Bruges, non Phryges, ipsius antiqui declarant libri. Nec enim Graecam litteram adhibebant, nunc autem etiam duas.” (Cic. *Orat.* 160.)

“In my own case, knowing that our forefathers did not use the aspirate except with a vowel, I said *pulcer*, *Cetegus*, *triumpus*, *Cartago*; after a while, – a long while indeed – the reproof of the ear forced me to abandon the correct pronunciation; I yielded to the people in the matter of usage, and kept the knowledge for myself. However, we say *Orcivius*, *Mato*, *Oto*, *Caepio*, *sepulcrum*, *corona*, *lacrima*, because the judgement of the ear permits it. Ennius always says *Burrus*, never *Pyrrhus*; *vi* patefecerunt *Bruges*, (not *Phryges*) is the reading of the oldest manuscripts of the poet. For in his day Latin had not adopted the Greek letter, now they use two.” (Trans. Harry Mortimer Hubbell)

“Parcissime ea veteres usi etiam in vocalibus, cum ‘aedos’ ‘ircos’ que dicebant. Diu deinde servatum ne consonantibus adspirarent, ut in ‘Graccis’ et ‘triumpis’. Erupit brevi tempore nimius usus, ut ‘choronae’ ‘chenturiones’ ‘praechones’ adhuc quibusdam inscriptionibus maneat, qua de re Catulli nobile epigramma est.” (Quint. *Inst.* 1.5.20.)

“Older authors used it but rarely even before vowels, saying *aedus* or *ircus*, while its conjunction with consonants was for a long time avoided, as in words such as *Graccus* or *triumpus*. Then for a short time it broke out into excessive use, witness such spelling as *chorona*, *chenturia* or *praecho*, which may still be read in certain inscriptions: the well-known epigram of Catullus will be remembered in this connexion.” (Trans. Harold Edward Butler)

Since the *h* phoneme had either disappeared completely by the Classical period or, if András Cser is correct, was only present in a very limited context, the correct pronunciation of *h* and its spelling became a matter of literacy, which could only be learned in the schools. Take a closer look at what can be learned from Cicero, writing in the first century BC, and Quintilian, writing in the first century AD. The way both put it suggests that no aspiration had been pronounced in Latin since the beginning of time. But both Cicero’s *maiores* and Quintilian’s *veteres* terms are misleading. An important question is how far back in time they could look in their research.

Research suggests that at best they could only go back to the second half of the second century BC, i.e. the beginning of the Classical period, see [Adamik \(2009, 138\)](#). Thus, their findings do not nuance the process presented in the previous chapter.

What is new, however, and of value as a source, is that the Greek names and expressions with many aspirations have found their way into Latin spelling, since originally the Greek consonants were written and pronounced without aspiration, cf. *Aciles* = Achilles (CIL 1, 567, 3), *Pilemo* = Philemon (CIL 1, 2593, 1), *Corinto* = Corintho (CIL 1, 626, 3). In addition to the names, the loanwords (e.g. *teatrum* = *theatrum* (CIL 1, 2945, 15) *struppus* = *strophos* (Andr. *poet.*9.) should also be highlighted. Several examples are analysed in detail by [Biville \(1990, 142–178\)](#) and [Mancini \(1990, 20–82\)](#). Over time, these have proved to be not only lexical but also phonological borrowings. Most scholars agree that since correct aspiration meant literacy (e.g. Quint. *Inst.* 12.10.57.) and literacy meant status, this led to hypercorrect forms in Latin, i.e. eventually the spelling and pronunciation of words of purely Latin origin were also influenced by Greek loanwords, as mentioned by several authors ([Porzio Gernia 1974, 100](#); [Sihler 1995, 142](#); [Adams 2003a, 109](#); [Herman 2003, 38](#)). In contrast, [Allen \(1978, 27\)](#) has argued, not very convincingly, that “in nearly all the attested cases the aspiration occurs in the vicinity of a liquid consonant”, and that this could be, in his view, a natural development in Latin itself. He says that the new spelling of Greek words allowed this subtle difference in pronunciation to be recorded in writing, which he imagines to be either regional or social. This phenomenon is commemorated by Catullus in his 84. *carmen*, when he mocked Arrius, who pronounced what he could but wrongly: *chommoda* instead of *commoda* (Catul. 84.1), *insidias* as *hinsidias* (Catul. 84.2), or *Hionios* instead of *Ionios* (Catul. 84.12). [Ellis \(1876, 366\)](#) had already addressed the question of social status and erudition, although for him the great number of Greeks coming to Rome seemed a more convincing argument. [Fordyce \(1978, 374\)](#) approaches the question from the point of view of erudition. [Thomson \(1997, 524\)](#) also focuses on the issue of literacy, but adds a socio-historical explanation. In his view, the large increase in the size of the plebs after the Allied wars contributed significantly to the scale of the phenomenon, cf. “One reason for the uprush of *rusticus sermo* in Rome itself may have been the influx of rural ‘Italians’ into the city after the Social Wars.” [Adams \(2007, 174\)](#) also sees it as a question of literacy. Thus, modern interpreters of the text see Arrius as a hapless man who tried to improve his social standing by demonstrating his literacy through the use of aspiration, only to have the opposite effect on his audience because of his illiteracy.

It is worth quoting [Herman \(2003, 38\)](#), who is a little more categorical than other researchers on this issue:

“It seems that certain speakers, out of snobbery, often mimicked Greek pronunciation, i.e. pronounced *h* in many cases, even in purely Latin words.”

In the quoted passage, Herman is commenting on the *carmen* of Catullus mentioned by Quintilian. Perhaps one could nuance his observation by adding that this ‘snobbery’ may have been more characteristic of the beginning of the process, when the status of aspiration ‘burst’ into the wider public consciousness, and then many tried to follow the current trend and assert their non-existent cultural superiority, but as time passed and the novelty wore off, the matter became a simple matter of literacy.

2.3. Linguistic awareness and the city of Rome

A statement similar to that of Cicero and Quintilian was made by Nigidius Figulus (first half of the first century), which leads to the next major topic, the establishment of Latin linguistic awareness.:

“Rusticus fit sermo si adspires perperam.” (Gel. 13.6.3)

“Speech becomes rustic, if you misplace the aspirates.” (Trans. John C. Rolfe)

By the first century BC, the demand for correct, precise and accurate use of language was developed. The trigger was also Hellas. From the beginning of the second century BC, as the conflicts in the East intensified and Hellas was conquered, onward there was a closer contact with Greek culture, which led to the establishment of systematic education in grammar and rhetoric. In his *De grammaticis et rhetoribus*, Suetonius, who lived at the turn of the first and second centuries AD, explains that initially there was no such education in Rome, and that the Romans despised the profession, see Suet. *Gram.* 1. Nevertheless, this form of education spread like wildfire and within a short period there were twenty popular grammar schools in Rome, see Suet. *Gram.* 3. In addition to Greek grammatical and rhetorical education, the so-called Latin rhetorical schools appeared, which, despite their popularity, were considered by their critics to be of a lower intellectual level, focusing solely on linguistic training, see Cic. *de Orat.* 3. 93–94. These schools were closed in 92 BC, cf. Suet. *Gram.* 25; Tac. *Dial.* 34. Despite their contemporary reputation, it is worth highlighting Adamik’s (2009, 203–204) observation that the language training provided here contributed to the development of linguistic awareness.

Cornificius, who may be the author of the first systematic Latin grammar, was the first to define exactly what is meant by the correct use of language, i.e. what *Latinitas* consists of:

“Latinitas est, quae sermonem purum conservat, ab omni vitio remotum. Vitia in sermone, quo minus is Latinus sit, duo possunt esse: soloecismus at barbarismus. Soloecismus est, cum in verbis pluribus consequens verbum superiori non adcommoatur. Barbarismus est, cum verbis aliquid vitiose efferatur.” (*Rhet. ad Her.* 4.12.17.)

“It is Correct Latinity which keeps the language pure, and free of any fault. The faults in language which can mar its Latinity are two: the Solecism and the Barbarism. A solecism occurs if the concord between a word and one before it in a group of words is faulty. A barbarism occurs if the verbal expression is incorrect.” (Trans. Harry Caplan)

This source demonstrates that the Greek model of grammatical and rhetorical education had taken root and spread so rapidly in Rome that by the first third of the first century BC there was an internal need and urge to define and illustrate by example what was meant by the correct use of language. Cornificius is brief in his definition and considers the error-free use of language as the basis of *Latinitas*. He divides errors into two categories: those that affect the text, and those that occur in individual words.

Terentius Varro has nuanced and expanded Cornificius’ definition:

“Latinitas est incorrupte loquendi observatio secundum linguam Romanam, constat autem, ut adserit Varro, his quattuor, natura, analogia, consuetudine, auctoritate.” (Diomedes, GL 1.439.16–30.)

“Latinity is the observation of uncorrupted speaking according to the Roman language. As Varro asserts, it consists of these four things: nature, analogy, usage and authority.” (Trans. Minna Seppänen)

Varro’s four elements (scil. of Latinity) are irrelevant to our research, but his statement that he considers Roman usage to be the correct example to follow is important. He was the first who stressed the importance and superiority of the Eternal City and its language. Moreover, he even considers the Roman use of the language to be the one to follow, for more information see [Adamik \(2009, 205–206\)](#).

The juxtaposition of rural and urban, or Roman, vocabulary also occurs in two other fragments. The first (Var. gram. L. 5.97.) is about aspiration: “in Latio rure edus, qui in urbe ut in multis <a> addito haedus.” And the second (Var. gram. L. 7.96): “Cecilius pretor ne rusticus fiat.” For these, see [Adamik \(2009, 206\)](#). In the second fragment, a process not discussed in the present paper, the monophthongisation of *ae* > *e*, was mocked. For more recent work on this process, based on the epigraphic material, see [Papini \(2020\)](#).

The next author who had influence on the Latin linguistic awareness is Marcus Tullius Cicero. In his seven works on the theory of rhetoric he demonstrated his mastery of all rhetorical techniques, and in his innumerable speeches in legal and political debates he proved himself to be an excellent orator, not only in theory but also in practice. In his *De oratore*, Cicero discusses at length the linguistic geography introduced by Varro:

“sed hanc dico suavitatem, quae exit ex ore; quae quidem ut apud Graecos Atticorum, sic in Latino sermone huius est urbis maxime propria.” (Cic. *de Orat.* 3.12.42.)

“...But I mean actual charm of utterance, a merit which as among the Greeks it is peculiar to Attica so in Latin speech is specially the attribute of this city [Rome].” (Trans. Harris Rackham)

“Qua re cum sit quaedam certa vox Romani generis urbisque propria, in qua nihil offendi, nihil displicere, nihil animadverti possit, nihil sonare aut olere peregrinum, hanc sequamur neque solum rusticam asperitatem, sed etiam peregrinam insolentiam fugere discamus.” (Cic. *de Orat.* 3.12.44.)

“Consequently, as there is a particular accent peculiar to the Roman race and to our city, involving no possibility of stumbling or causing offence or unpleasantness or objection, no note or flavour of provincialism, let us make this accent our model, and learn to avoid not only rustic roughness but also provincial solecisms.” (Trans. Harris Rackham)

Cicero’s Greek parallel is quite illustrative. Several dialects [διάλεκτος] can be distinguished in ancient Greek, see [Buck \(1955\)](#) and [Jeffery’s \(1961\)](#) now classic work on the epigraphic material, and more recently [Miller \(2014\)](#). Cicero seems to have feared that the use of a rural, unsophisticated language, or a strange, not necessarily unsophisticated, but to his Roman ear unfamiliar language, might overwhelm the Roman use of language so peculiar to him and Varro, and that a kind of schism in the Latin language might be caused or intensified. And he tries to support the prominent role of Rome with a Greek parallel. According to him, Roman *latinitas* is like Attic language use among the Greeks. In the passage between the two quotations, he goes on to praise the pleasant speech and pronunciation of the Attic and despises even the most cultured Asiatic Greek (cf. Cic. *de Orat.* 3.12.43.). Behind the praise of Athens and the condemnation of Asia, in addition to the differences in pronunciation mentioned above, one can easily detect the controversy between the two great rhetorical styles, Atticism (simpler, more to the point) and Asianism (more turgid). On this see Cic. *Brut.* 35, 289; *Opt. Gen.* 6.

The last classical author to publish a monumental work on grammar and rhetoric was Quintilian. His contemporaries rightly praised his work – Martialis, for example, called him the greatest teacher of Rome (Mart. 2.90) – and his ten-volume *Institutio oratoria* is a worthy continuation of the theoretical research of Cornificius, Varro and Cicero. For him, Cicero was the ultimate point of reference, the example to be followed (Quint. *Inst.* 10.1.105, 112), so it is not surprising that Rome’s prominence in the correct use of language was also reflected in his work:

“Quare, si fieri potest, et verba omnia et vox huius alumnus urbis oleant, ut oratio Romana plane videatur, non civitate donata.” (Quint. *Inst.* 8.1.3.)

“Therefore, if possible, our voice and all our words should be such as to reveal the native of this city, so that our speech may seem to be of genuine Roman origin, and not merely to have been presented with Roman citizenship.” (Trans. Harold Edward Butler)

Robert Coleman (2000, 30–31) summarises briefly but well the linguistic standardisation efforts of the period from Varro to Quintilian. It makes no difference whether one makes a particular mistake following an example from one’s ancestors [*veteres*] or because of rural influence [*rusticitas*] or simple linguistic negligence [*barbarismus*], it is in any case contrary to the Roman-centred view of language [*urbanitas*].

2.4. Late processes and afterlife

The Roman-centred approach to language in the first centuries BC and AD was successful, and, following Cicero’s example, the contemporary Latin language has a much more unified form than Greek. Sabine Grebe (2001, 137), in her comparison of the basic requirements of *latinitas* and Ἑλληνισμός, while noting that dialects are an important aspect in Greek and also that this concept exists only in Greek – i.e. in comparison with Latin – but does not conclude that this absence is precisely the consequence of the process of standardisation of Roman language use, cf. Adamik (2009, 235).

This rigidly Rome-centric approach is particularly interesting, because among the great classical authors, only Caesar and Varro were born in Rome. Horace, Virgil, Tibullus, Propertius, Ovid, Cicero, Sallust and Livy were born in other parts of Italy. The situation is even worse for the authors of the Silver Age: the Senecas, Martialis and Quintilian were all from Spain, Juvenal and Pliny from Italy, Tacitus at best from northern Italy, but possibly from southern Gaul. I think this dichotomy makes the subject of the present study particularly interesting, to analyse the use of aspiration as a measure of literacy in the most Latin city, Rome, which at least in theory provided the standards for *Latinitas*.

However, Latin grammarians were committed to preserve the purity of the Latin language, but this standardisation was never achieved. The limited framework of the standardisation process is summarised in Adams (2003b). Focusing only on the subject of this paper, aspiration errors, it is worth examining the use of the verb *aveo*. Etymologically, the word certainly has no aspiration, as PIE reconstruction of de Vaan (2008 s.v. *aveo*) points out. Nevertheless, the aspirated form, *have* is recorded, albeit rarely, by several classical authors. These are, in chronological order, Cicero (e.g. Cic. *Fam.* 8.16.4.), Seneca the Younger (Sen. *Ben.* 6.34.3.), Martialis (14 times, e.g. Mart. 1.68.6; 3.95.14; 5.51.7.), Suetonius (Suet. *Cl.* 21.6.), Fronto (11 times, e.g. Fron. *Amic.* 2.7.1. *Ant.* 3.4.1.) and Apuleius (Apul. *Met.* 7.5.12.). The late ancient grammarians, Charistus (Char. p. 211) and Diomedes (Diom. 348), on the other hand, still considered

only the *ave / avete* forms as normative. Were they right to do so? Is it possible to ignore early authors such as Cicero and Seneca the Younger, whose use of language was a standard in other matters? Is it not possible to imagine a kind of pluralism, just as the *Oxford Latin Dictionary* considers two spellings to be normative in many cases? The current policy of the LLDB is that the question cannot be decided with certainty. Since it is only the medieval manuscript evidence of these authors which has the initial *h*, which cannot prove that the authors themselves pronounced this consonant in antiquity. Therefore, all *have = ave* forms are recorded, but these data forms also immediately are annotated with the *fortasse recte* note, and these data cannot be considered in our research as being normative variants.

Even though the question of correct aspiration has always been on the periphery of the minds of grammarians and educated authors, aspiration has completely disappeared from the Romance languages. In the *Appendix Probi* there are seven lines dealing with aspiration: 23. *cithara non citera*; 46. *Theophilus non Isophilus*; 66. *cochlea non coclia*; 67. *coc[h]leare non cocliarium*; 190. *[h]ermeneumata non erminomata*; 225. *adhuc non aduc*; 207. *hostiae non ostiae*. St Augustine also argues for the correct *aspiratio* in several places (e.g. August. Ars. 11.1; Conf. 1.18.29). Consider the Latin word *homo* [man]: where there is a continuation of the *h* phoneme in the Romance languages, the *h* either reappeared as a result of an external linguistic influence – Germanic – or as a result of some spelling restoration process, as in French *homme*, Spanish *hombre*, but there is no continuation in Italian, *uomo*, or Romanian, *uman*, as described by Väänänen (1981, 55) and Coleman (2000, 30). Dascalu’s (2019) research on the Merovingian digital text corpus, in which she examines the context in which aspirated consonants have been preserved or disappeared, promises to be of interest. The research is in its early stages and has so far focused on the origin of the words – Greek, Hebrew, Frankish – and their function – religious, civilisational, philosophical, proper names.

3. METHODOLOGY

There are two main schools of Latin dialectological research. For a summary on the emerging debates, see Adamik (2012b). The intellectual father of the first one is Paul Gaeng (1968), who based his research on a province-by-province comparison of correct and incorrect spellings for the same linguistic phenomenon. The second is based on the work of József Herman (2000), who argues that the proportion of different errors is interrelated by area, not by the correct forms, and thus develops a measure that is independent of any cultural influence, i.e. it can be used for dialectological studies. The dialectological differences between areas are then reflected in the changes of the proportions. For a brief introduction to Herman’s system, see Adamik (2012a, 99–105).

James Noel Adams (2007, 635–636), who was a leading scholar of Latin linguistics, has developed his own methodology using a combination of Gaeng’s and Herman’s. In his view, if a particular type of error does not occur in one province but does occur in another, it is not possible to conclude that it was indeed specific to one province, because the level of literacy might have been higher in the latter. As a result, there is no written record of the linguistic difference and therefore no record of the change or linguistic error being studied. According to Adams, in cases where a significant amount of data is available from the province under study and one or more exceptionally common error types are present as well, a firm conclusion can be

drawn, and in such cases the absence of the linguistic phenomenon mentioned in the example can be interpreted as a real absence.

Although Adams' modification seems convincing at first glance, Béla Adamik (2012, 100, n. 25) rightly points out that "[h]owever, the inscriptional records are characterised by the fact that there are all types of errors everywhere, just with different frequencies and different distributions." In other words, Adams' method is functional, but it is only applicable to a limited number of corpora, one might say only in ideal cases.

In this paper I follow Herman's methodology, thus in *Conclusions* I first compare the frequency of incorrect use of *h* with data of consonant errors (consonantism) and errors indicating non-linguistic changes (*errores non grammatici*), i.e. misspelling, reversal of letters, omission, Greek letters (see the full list at https://lldb.elte.hu/admin/abbrev_codes.php).

Orthographic errors, such as the misuse of the letter *h*, which are spelling errors at the level of the texts and which could only be avoided by thorough education and practice of the standard language, are well suited to the study of language use. By studying the frequency of purely orthographic errors in the Roman corpus, century by century, we can reveal the changes over time in the spelling dynamics of the users of the language, as manifested linguistically in the inscriptions, which can provide a minor but instructive element in the history of Roman linguistic culture.

The frequency of errors related to the incorrect use of *h* is first measured against the total number of consonant errors, which are purely linguistic in nature, because this allows us to separate the dynamics of the use of *h*, and thus changes in the level of literacy, from deviations reflecting real linguistic changes. The data then were compared with non-linguistic errors, a comparison which allows us to draw conclusions about another level of literacy, the quality of the textual execution of the inscriptions. Thus, the different proportions of orthographic errors from one period to another reflect changes in technical execution.

Since this paper focuses on spelling issues, the Gaeng method may also be appropriate, i.e. the orthographic errors of each century can be compared to other data, such as the correct spelling of the same inscriptional material. That is, how many times *h* is used correctly or incorrectly in all the inscriptions of a given period. The conclusions will point out that the comparison with consonantal and linguistic errors did not sufficiently clarify the dynamics of the aspiration processes. These analyses pointed to other linguistic and cultural processes. Therefore, a promising comparison has also been made with the data on the correct *h*-s in the *Epigraphik-Datenbank Clausz / Slaby* (https://db.edcs.eu/epigr/epi.php?s_sprache=en).

Seven centuries of data were analysed, from the first century BC to the sixth century AD. I have included the first century BC merely as a starting point in the research period to see the linguistic state of the Republic in its final days. This dataset can only serve as a simple illustration, as the LLDB, due to its scope, only records data from the symbolic beginning of the Imperial era, 27 BC. Therefore, it cannot be considered a representative sample of the first century BC. The Database provides data for the city of Rome up to the ninth century AD: 59 entries from the seventh century AD (25 inscriptions), 59 from the eighth century AD (24 inscriptions) and 17 from the ninth century AD (8 inscriptions). These do not meet the criteria of measurability due to the lack of data, therefore these centuries are not considered.

It should be noted that our data for the pagan period – first to third century AD – may be distorted, because in many cases neither the *Corpus Inscriptionum Latinarum* nor other online databases specialised in epigraphic research such as the *Epigraphik-Datenbank Clausz / Slaby* or *Epigraphic Database Roma* give an exact date (only 1–300 AD), or even less, i.e. they do not date

the inscriptions. Inscriptions undated by the sixth volume of the *Corpus Inscriptionum Latinarum* are recorded in the database with a date of 27 BC – 313 AD. This method of working was made possible by the ‘division of labour’ that once existed between the *Corpus Inscriptionum Latinarum* and the *Inscriptiones Christianae Urbis Romae*. The former published the pagan material and the latter the Christian material. However, these dates, using the “Period [A]” dating in the LLDB, all fall into the second century AD. This method means that “those data forms will be displayed where the arithmetic mean of the indicated period falls within the period specified in the search query”. This inaccuracy cannot be eliminated at this stage, but in the longer-term new publications or existing online databases may provide more accurate dates, and in the light of these, these partial results of the research can be revised.

4. ASPIRATION ERRORS IN THE ROMAN INSCRIPTIONS

4.1. First century BC

The data of the first century BC are on the borderline of evaluability – cf. LLDB time span –, but it is worth making some observations about the distribution of errors (Figure 1).

It is noticeable that 26 of the 28 data entries are onomastic in origin, which illustrates that the new spelling rules used in the transcription of Greek names caused the problem. To mention just a few examples: CELIDO = Chelido (LLDB-63846), PAMPILVS = Pamphilus (LLDB-66933) or ATENAIS = Athenais (LLDB-66781). The incorrect aspiration in the form ORCHI = Orci (LLDB-66772) is also worth pointing out, as it can be proved by a primary, epigraphic evidence – not only by Catullus’ 84 *carmen* – that the uncertainty of the aspiration affected not only Greek words but also words of Latin origin – in this example a nomen gentile – already in the first century BC.

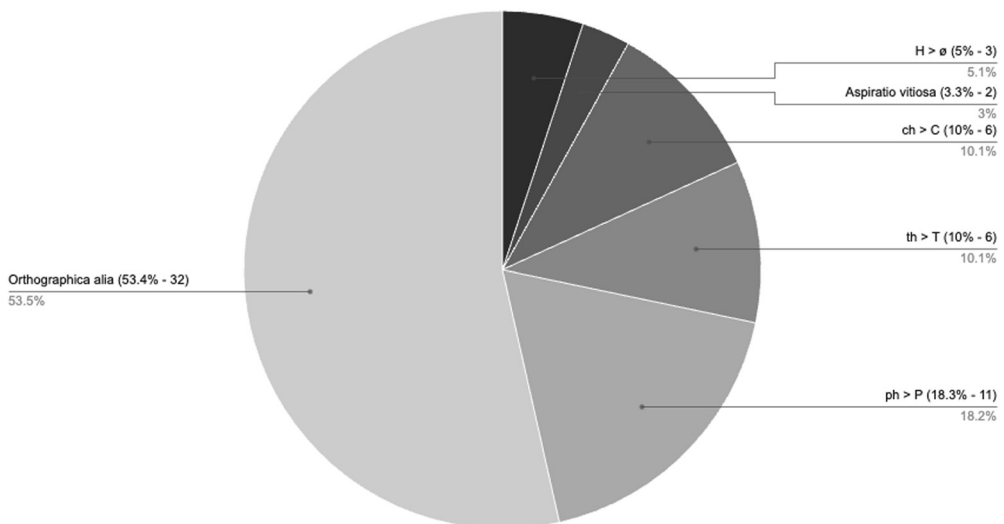


Figure 1. Aspirational errors in the I. century BC

4.2. First century AD

A total of 376 data sheets recorded aspiration errors (Figure 2). Most errors (150 records, 28.6%) fall into the *aspiratio vitiosa* category. Most of the data (139) are onomastic in origin and only 11 refer to common nouns. By far the most frequent error with 58 occurrences is the Greek name Elpis with its derivatives and declined forms, e.g. HELPIS = Elpis (LLDB-53382), HELP-IDEPHORO = Elpidophoro (LLDB-109269). Also notable with 12 data is the Greek name Tyche with its derivatives. These examples illustrate that the composer of the inscription knew that there was an aspirated consonant in the word, but not exactly which one, and whether it was one or two: e.g. CALETHVCE = Caletyche (LLDB-46232), EVTHYCHVS = Eutyclus (LLDB-108590). Among Latin proper nouns, Orcius and its derivatives are erroneously given with a *ch* seven times, e.g. ORCHIA = Orcia (LLDB-66768). Among common nouns, the Latin words *aeneus* and *lacrima* occur twice with an unnecessary aspiration at the beginning of the word and after the letter *c* (LLDB-73668, LLDB-78603; LLDB-39044, LLDB-143227).

The omission of the *h* (43 records, 8.2%) also occurs most frequently among the names, with the Euhodus, Histumennius/Histimennius forms being highlighted, e.g. EVODVS = Euhodos (LLDB-56518), ISTIMENNIVS = Histimennius (LLDB-52263). Among the six common noun forms, however, only Latin words are used, e.g. ORA = hora (LLDB-71654) or IC = hic (LLDB-57641).

Except for ARCITECT = *architectus* (LLDB-72095), there are only onomastic examples of the *ch* > *c* change (71 records, 13.5%). The Greek name Tyche and its derivatives are the most frequent, with 25 occurrences without aspiration, e.g. EVTYCIAE = Eutychiae (LLDB-90437). Then, with 9 occurrences, the name Antiochus and its derivatives occur most frequently (e.g. LLDB-169360).

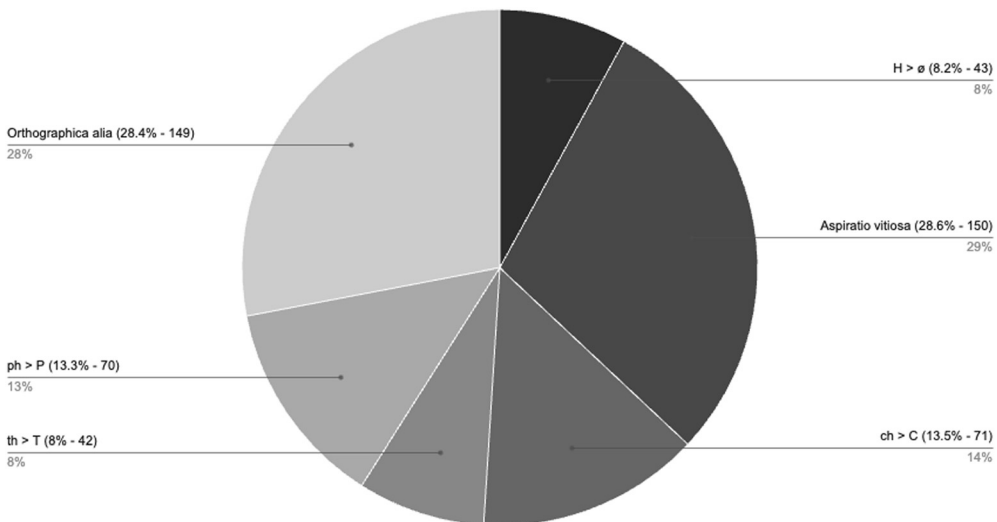


Figure 2. Aspirational errors in the I. century AD

Almost the same proportions for $ph > p$ change (70 records, 13.3%). Only two common names, a Greek one AMPITHEAT = *amphitheatro* (LLDB-143140) and a Latin one TRIVMPVS = *triumphus* (LLDB-71649) are included in the database. In the onomastic material, APRODISIVS = Aphrodisius (LLDB-143356) and its derivatives are the most frequent with 12 records, followed by NICEPORUS = Nicephorus six times (e.g. LLDB-143158).

For the $th > t$ change (42 records, 8%) onomastic examples also dominate. Most frequent, with ten occurrences, are Agathos and its compounds, e.g. AGATANGELI = Agathangeli (LLDB-57640) AGATOPODIS = Agathopodis (LLDB-38693), followed by AMARANTVS = Amaranthus (LLDB-129628) with three-to-three occurrences, and compound words formed with the Greek word for 'god', e.g. TIOPHANES = Theophanes (LLDB-39982). The only two common nouns are BIBLIOTECA = *bibliotheca* (LLDB-43324, LLDB-58064) and AMPHI-TEATRO = *amphitheatrum* (LLDB-143141, LLDB-146563).

4.3. Second century AD

Even in the second century, incorrect aspiration provides the most data, 360 out of 900 (Figure 3). The vast majority of the data are still of onomastic origin. Even more than half of these – 204 records – derive from the Greek name Elpis and its derivatives: HELPIDI = Elpidi (LLDB-108629), HELPITO = Elpisto (LLDB-46689), EVHELPISTVS = Euelpistus (LLDB-52113). On the one hand, this high figure indicates a trend in the choice of names, i.e. the name is common. On the other hand, it could easily be deduced from this frequency that the incorrect form of Helpis is copied by analogy, in the belief that it is the correct form. This increased the number of incorrect forms. The next larger group, the name Tyche and its derivatives, occurs much less frequently, 18 times: THYCE = Tyche (LLDB-56961),

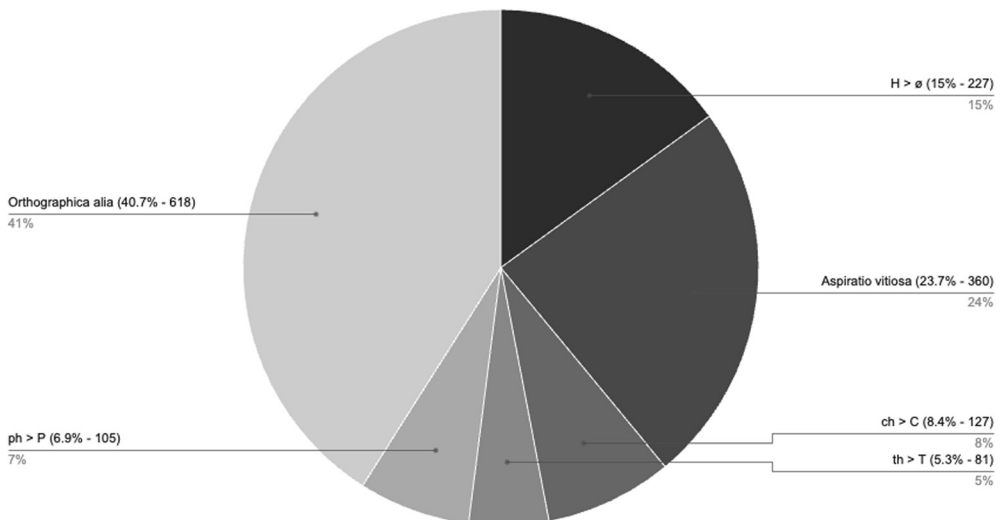


Figure 3. Aspirational errors in the II. century AD

EVTHYCHES = Eutyches (LLDB-58336), HEVTHY|CIA = Eutychia (LLDB-67251). These illustrate that the author is aware of the necessity of the aspiration but is not sure of its exact position. Errors appear in more words and positions: HAEDICVLAS = *aediculas* (LLDB-39540), HAETERNALEM = *aeternalem* (LLDB-38444), HARVNDINETVM = *arundinetum* (LLDB-110933), CHOIVGIS = *coniugis* (LLDB-41960), HEIVS = *eius* (LLDB-112109), HEORVM = *eorum* (LLDB-34062), FRATRHS = *fratres* (LLDB-40576) or SEPHVLCRV|M = *sepulcrum* (LLDB-56733). In the Latin proper names, there are isolated errors, but even in these cases there is a wide variability: HEBUR = Eburus (LLDB-40591), ORCHIVS = Orcius (LLDB-66770), VLPHIAE = Ulpiae (LLDB-110392).

The loss of *h* (227 records, 15%) was also the most significant in the onomastic data. It is most frequently recorded after names with only one or two occurrences, such as EDONE = Hedone (LLDB-37663), YACINTHVS = Hyacinthus (LLDB-110554) or YGIA = Hygia (LLDB-52152), 16 occurrences of RODIA = Rhodia and its derivatives (LLDB-39789), followed by 13-13 occurrences of ESYCHVS = Hesychus (LLDB-170617), EUODIA = Euhodia (LLDB-67200). Among the Greek common nouns YPOGAEVM = *hypogaeum* is the most frequent, with three occurrences (LLDB-52318). Of the Latin common nouns, the most common with 18 occurrences is the word ORA = *hora* (LLDB-129274), followed by ERES = *heres* with ten occurrences (LLDB-90264). None of this is surprising in the light of the source base. Since most of the data come from epitaphs, the number of hours lived or the – relatively – frequent occurrence of heirs is perfectly logical. In the Latin proper names, there are only a few examples of this phenomenon, cf. the names ORTESIA = Hortensia (LLDB-52115) or ORTORIVS = Hortorius (LLDB-107777) being the most notable.

Apart from a few sporadic examples of common nouns, the *ch* > *c* change (127 records, 8.4%) is dominated by onomastic examples, e.g. NAVARCVS = *nauarchos* (LLDB-90463), TRICLIA = *trichilam* (LLDB-111181). In the second century AD, Tyche and its derivatives also provided the most data (31), e.g. TYTYCIANO = Tychiano (LLDB-64034), EVTYCIAE = Eutychia (LLDB-56524) CALITVCE = Callityche (LLDB-34057). Second, with 17 examples, is the group of names in which the Greek word *chrysos* appears, e.g. CRYSIDI = Chrysid (LLDB-64036), CRYSOGENES = Chrysogenes (LLDB-64555) or the forms CRYSANCTVS = Chrysantus (LLDB-130621). The following names are recorded four to eight times, such as ANTIOCVS = Antiochus (LLDB-41644), CRESIMVS = Chresimus (LLDB-142440), POLYCRONIO = Polychronio (LLDB-128948) and SOTERICVS = Soterichus (LLDB-37162).

The *ph* > *p* change (105 records, 6.9%) was only reported in onomastic examples in this century. No name or group of names shows a definite increase in frequency. APRODISIA = Aphrodisia, NICHEPOR = Nicephor, POSPHORVS = Phosphorus and names containing the Greek *phil* stem were the most frequent – attested 7 to 13 times, e.g. PILONIVS = Philonius (LLDB-67356), in addition to sporadic examples such as CTESIPON = Ctesiphon (LLDB-111383), OMPALE = Omphalae (LLDB-36208) or ZEPYRI = Zephyri (LLDB-141966).

Only two non-onomastic examples are attested for *th* > *t* change out of 81 data (5.3%). There are two occurrences of the Greek noun CYATO = *cyathos* (LLDB-167666). Again, Agathos and other compounds stand out with 14 occurrences, e.g. AGATEMER = Agathemer (LLDB-57431), AGATOCLES = Agathoclés (LLDB-34588), AC|ATOTYC|HE = Agathotyche (LLDB-38237). The data is diverse, with more than 30 name or name groups for the 79 errors: e.g. ATENAIS = Athenais (LLDB-35551), CORINTIAS = Corinthias (LLDB-39479), ELEVTERRIDI = Eleutheridi (LLDB-38119), TALLVSA = Thallusa (LLDB-66910).

4.4. Third century AD

The database contains very few data forms from the third century AD, only 200 (Figure 4). This is (also) due to a difficulty in dating which cannot be overcome at this stage, so that the few data on aspiration errors may be biased and some of the data from the second century AD may in fact belong to this period. For more on this issue, see the chapter on Methodology. At the same time, the period of the barracks emperors – 235–284 – was one of the most turbulent periods in the history of Rome, so that historical events may explain the numerical decline in inscribed material.

The prevailing error in the third century AD is the loss of *h* (96 records, 21.4%). An important change is that the onomastic dominance observed so far is eliminated, with 56 of the 96 data referring to common names. Furthermore, except for two Greek words – YPOGEV = *hypogaeum* (LLDB-139334) and EPATO = *hepatos* (LLDB-72389) – only Latin words are attested, cf. with six occurrences ORA = *hora* (LLDB-123088) and ERES = *heres* (LLDB-50771), nine occurrences of the declined forms of *cohors* (LLDB-66575), and 17 occurrences of ABEO = *habeo* (LLDB-85362). Several processes may be involved in this change. On the one hand, a change in naming habits. In the LLDB the names Euhodus, Histumennius/Histimennius, Heuresis, Hermes or Rhodia and their derivatives, which were popular in the previous two centuries, are sporadically or never found on inscriptions which can be dated to the third century AD. On the other hand, terminological differences in the growing Christian material also contributed to the process. E.g. the formula *Sancti martyres Paule et Petre in mente habete / habeatis* became a recurrent element in Christian epitaphs, which increased the number of aspiration errors regarding common nouns due to the verb *habeo*. The previously mentioned dating difficulty may also distort these data, but the change in the terminology of the epitaphs allows to assume that the proportion of common names – even

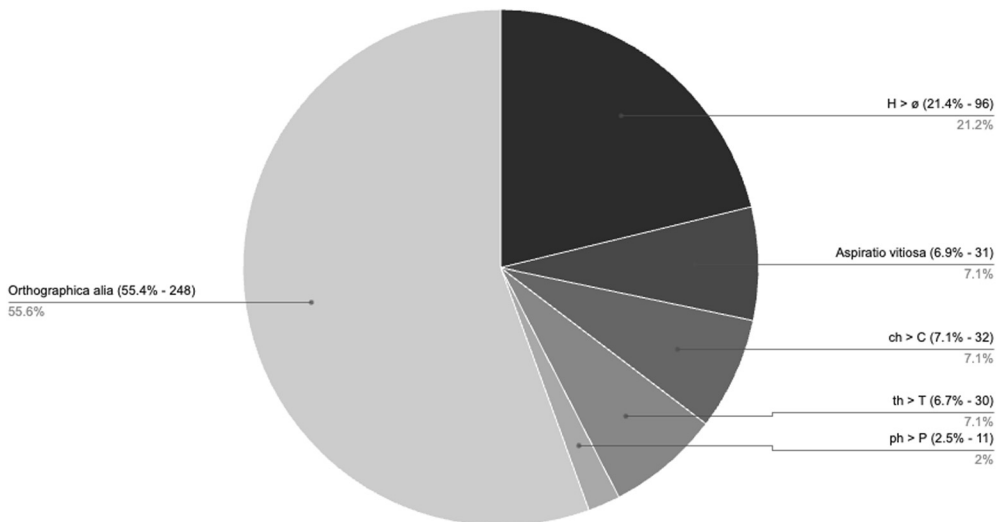


Figure 4. Aspirational errors in the III. century AD

with the somewhat distorting dating – has increased compared to proper names. There is no clearly outstanding source of error in case of proper names. The declined forms of the names ERENNIUS = Herennius (e.g. LLDB-110139) and ERMES = Hermes (e.g. LLDB-109003) appearing repeatedly, five and three times respectively.

HELPIS = Elpis and its derivatives continue to provide the largest portion of the dataset for incorrect aspiration (31 records, 6.9%), with five occurrences in this century, followed by EUTHYCHE = Eutyche (e.g. LLDB-118152) with three occurrences. Among common names there is no prominent group. Examples can be found for word-initial errors, e.g. HETERNAE = *aeternae* (LLDB-110772), consonant-related errors, cf. SACRATHIS = *sacratris* (LLDB-36987), but also indicating the separate pronunciation of vowels, see DEHE = *deae* (LLDB-98761).

Apart from a single common noun example – DIETCI = *diaetarchi* (LLDB-55652) – only onomastic examples of the *ch* > *c* change (32 records, 7.1%) are attested. With nine examples the name Tyche and its derivatives continued to provide the most data (e.g. LLDB-138901), but with the widespread expansion of Christianity the error CRISTVS = Christus (LLDB-96831) appeared, with six occurrences, making it the second most frequent erroneous name in the modest data of the third century AD. The other occurrences are all unique with one or two examples, such as EVSTACIS = Eustachis (LLDB-68449) or NICOMACE = Nicomache (LLDB-68411).

The dramatic reduction of the *ph* > *p* change – 11 records, 2.5% compared to 105 records, 6.9% in the previous century – is a telling statistic. By this period, the pronunciation of *ph* as a fricative had become, it seems, quite common, and since it corresponded to the Latin phoneme *f*, it was used in writing as well, ignoring the Greek spelling rules.

With two exceptions – TERMAS = *thermas* (LLDB-123110) and TESAVRIS = *thesauris* (LLDB-170938) – the *th* > *t* change (30 records, 6.7%) is only reported in proper names. With five examples the most common error is AGATOPVS = Agathopus (LLDB-142045), followed by TRAX = Thrax (LLDB-135801) and BITVS = Bithus (LLDB-89658).

4.5. Fourth century AD

The loss of *h* is the leading error with 256 records (18.6%, see Figure 5). The proportion of common names (85:256) is also higher here compared to the second century AD, continuing the trend identified in the third century AD. The word ORA = *hora* (LLDB-138300) and its declined forms occur most frequently, 49 times, which continues to reflect the inherent character of the inscriptions. The second most frequent word is *habeo* (LLDB-136352) with 18 occurrences. The third one is IC = *hic* (LLDB-130232) with seven records. Some examples of rarer common nouns are ONESTA = *honesta* (LLDB-82657), *homo* (LLDB-136488) and *hypogeum* (LLDB-57226), each with one mention. Among proper names the most frequent source of errors with 28 occurrences is Hilarus and its derivatives, e.g. ILARIO = Hilario (LLDB-131919). Followed by 13 occurrences of derivatives of the name Hercules, e.g. AERCLANIVS = Herculanius (LLDB-127670), then 12 occurrences of the name Heraclia and its derivatives, e.g. ERACLIA = Heraclia (LLDB-135873). Two other groups of names stand out with nine and eight occurrences respectively. These contain the Greek word *hippos* or Helios, e.g. IPOLITVS = Hippolytus (LLDB-138461) and ELIODORA = Heliodora (LLDB-135865). The rarer forms with one occurrence each are CALLEROE = Callirhoe (LLDB-81411), EORTE = Heorte (LLDB-131518), ESPANIS = Hispanis (LLDB-119604).

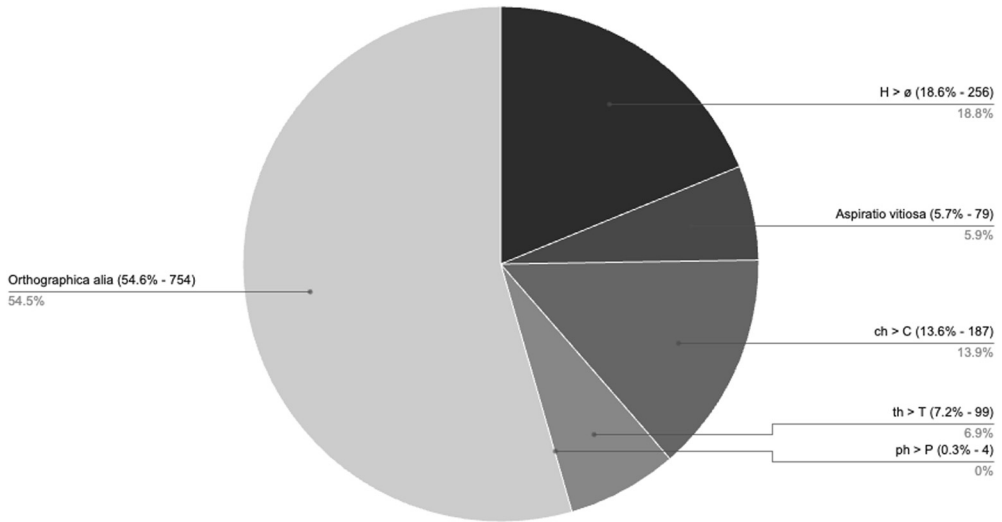


Figure 5. Aspirational errors in the IV. century AD

The Greek names Tyche (LLDB-135869) and Elpis (LLDB-67968) continue to be among the incorrect aspiration with eight and six occurrences. The common name PACHE = *pace* (LLDB-137886) was recorded seven times, again due to the Christian epitaph terminology, cf. *requiescat in pace*. There are three examples of the erroneous initial aspiration of the words *aeternus* (LLDB-56021) and *emo* (LLDB-127610). Sporadic occurrences are illustrated by HANORV = *annorum* (LLDB-131933), HVNICVS = *unicus* (LLDB-135910) and HISPIRITO = *spiritu* (LLDB-136586).

The ch > c change (187 records, 13.6%) is dominated by onomastic examples. The most frequent source of errors is the Greek name, Tyche and its derivatives, with 44 occurrences. Immediately followed by the name Christus or its derivatives with 26 occurrences. The third most common error with 16 entries is the name CLEARCO = Clearcho (LLDB-137977), followed by STILICONE = Stilichone (LLDB-130305) with 12 mentions. Among the common names, SCOLASTICVS = *scholasticus* (LLDB-88938) stands out with five examples.

The ph > p change (4 records, 0.3%) continues the trend of the third century AD. The phenomenon is recorded in only four names (LLDB-68456, LLDB-96810, LLDB-106722, LLDB-131901). The PH ~ F alternation became a common phenomenon – 289 data forms – due to the pronunciation of /ph/ as a fricative. The increase in frequency is due to terminological changes in the inscriptions. Many Greek names contain φ, as we have seen so far, which has significantly increased the possibility of PH ~ F alternation errors. However, the Edict of Milan in 313 established religious tolerance for Christianity within the Roman Empire. This allowed more explicit formulations in Christian inscriptions. Thus, it could become a recurrent phrase of fourth century AD epitaphs that the deceased person was a *neophytus* or *neophyta*, i.e. a newly converted Christian. The new formularization further increased the epigraphic appearance of the possible alternation of PH ~ F. A total of 64 errors of the type NEOFITA = *neophyta*

(LLDB-144779) were recorded in the database, while in the third century AD a total of 92 cases of PH ~ F alternation were found.

The *th > t* change (99 records, 7.2%) is still characteristic among proper nouns. The largest group, with 22 occurrences, is made up of words containing the Greek word *theos*, e.g. TEODOSIO = Theodosio (LLDB-99045) or TEOFILI = Theophili (LLDB-96561). Fifteen occurrences refer to words containing the Greek word *agathos*, e.g. AGATANGELVS = Agathangelus (LLDB-139168) or AGATOPVS = Agathopus (LLDB-95658). The third most frequent source of error is the name TIMOTEVS = Timotheus (LLDB-98977), which is recorded nine times. Only three examples of common nouns are recorded from this period: ACOLITVS = *acoluthus* (LLDB-128865) three times, followed by CATOLICE = *catholicae* (LLDB-145929) and ICTVS = *ichthys* (LLDB-129094), with one occurrence of each. It is a remarkable fact that all three words are related to Christianity.

4.6. Fifth century AD

Most of the aspiration errors are due to the loss of *h* (93 records, 16.3%, see Figure 6). Among proper names, the name Hilarus and derivatives of words containing the Greek noun *hippos* are almost as common – seven and eight occurrences – as ILARVS = Hilarus (LLDB-83816) or YPPOLITE = Hippolyte (LLDB-107904). Examples of individual cases of proper names are PYRRA = Pyrrha (LLDB-96407), OMERVS = Homerus (LLDB-96073) or ZACINIE = Hyacinthe (LLDB-107288). Among the common nouns, the most frequent occurrences of the loss of *h* were still in the declinated forms of the noun *hora* and the verb *habeo*, with 15 and 12 occurrences, e.g. LLDB-127073, LLDB-67575. Among the common nouns, the rarest occurrences were IC = *hic* (LLDB-128989) four times or VIVS = *huius* (LLDB-45074) once.

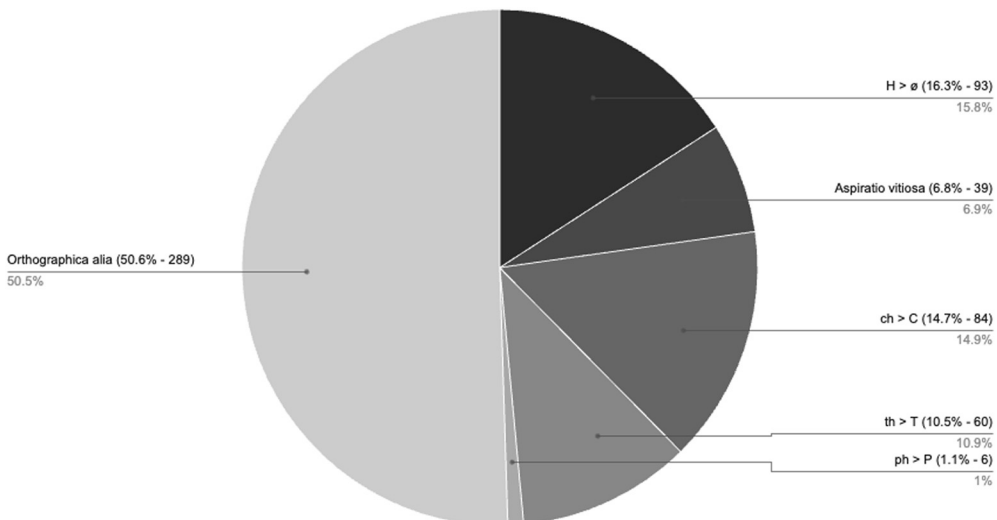


Figure 6. Aspirational errors in the V. century AD

There is no prominent source of error among the incorrect aspiration (39 records, 6.8%) in either proper or common names. There were five occurrences of the proper noun *Elpis* and four occurrences of the name *Antiochus*, e.g. *HELPIDIVS* = *Elpidius* (LLDB-36733) and *ANTHIOCO* = *Antiocho* (LLDB-128146). For proper names some sporadic examples are *CYRIACHE* = *Cyriace* (LLDB-81693), *HABENIIO* = *Aventino* (LLDB-81543) or *VICTORHIACIS* = *Victoriacius* (LLDB-67617). For common names: *HIM* = *in* (LLDB-67598), *HOC-TOBR* = *octobres* (LLDB-50131) or *PACHE* = *pace* (LLDB-138467). In other words, as in previous centuries, the incorrect aspiration can be attested in all circumstances.

The most frequent example of the *ch* > *c* change (84 records, 14.7%), reported 16 times, is *Christus* or its derivatives, e.g. *CRISTO* = *Christo* (LLDB-67609) or *CRISTIANVS* = *Christianus* (LLDB-81925). Second with 14 examples are incorrect forms of the name *Stilicho*, e.g. *STILICONE* = *Stilichone* (LLDB-130323). Among the proper names, the derivatives of the Greek noun for luck are also common, with seven occurrences, e.g. *CALOETYCAE* = *Callitychae* (LLDB-67845) or *EVTICYVS* = *Euty chius* (LLDB-95926). Among the common names, the 13 occurrences of the incorrect forms of *schola* and its derivatives clearly stand out, e.g. *SCOLASTICVS* = *scholasticus* (LLDB-39067). The rarer forms of the common names are *ARCIATRI* = *archiatri* (LLDB-42735), *ARCITECTI* = *architecti* (LLDB-108275) and *BIARCVS* = *biarchus* (LLDB-42718).

The change in *th* > *t* (60 records, 10.5%) is still dominated by onomastic data. The most frequent, with 19 occurrences, is the group of names containing the Greek word *theos*, e.g. *TEODOSIO* = *Theodosio* (LLDB-108269) or *TEODOLVS* = *Theodolus* (LLDB-88892). The other prominent group with seven occurrences is the group of names containing the Greek adjective *agathos*, e.g. *AGATOPO* = *Agathopo* (LLDB-95568). A total of six common noun errors are recorded. The familiar error *ACOLITUS* = *acolithus* (LLDB-136889) appears three times, and *AMPHITEATRI* = *amphitheatri* (LLDB-146418), *CATEGRARIO* = *cathedrario* (LLDB-88220), *TERMARVM* = *thermarum* (LLDB-146276) appear once each.

The change in *ph* > *p* continues the trend known from the third to fourth centuries AD with only six data (1.1%).

4.7. Sixth century AD

By the sixth century AD, the number of epigraphic sources had significantly decreased, therefore the recordable linguistic data as well – only 47 aspiration errors are recorded in the LLDB-Database (Figure 7). The decline of the inscriptional corpus was due to the constant wars in the Apennine peninsula. After the fall of the Western Roman Empire in 476, Italy was ruled by the Ostrogoths. Justinian, in the spirit of his political programme of *renovatio Imperii Romani*, redirected considerable military force and money to the reconquest of the peninsula, which resulted in the recapture of the Eternal City in 536 by his commander Belisarius, and later of the rest of Italy. However, the success was only temporary due to Gothic revolts and later Longobard and Frankish invasions. A good indicator of the extent of the barbarian devastation is the small number of the inscriptions that were erected in Rome in the 7th and 11th centuries, cf. those written in the *Methodology* chapter.

The *th* > *t* and *ch* > *c* changes are found to be the greatest source of aspiration errors with 14–14 occurrences (21.9%–21.9%). Both records are dominated by the names of the consuls. 11 times the name *Boethius* (LLDB-37480) and five times the name *Symmachus* (LLDB-108317)

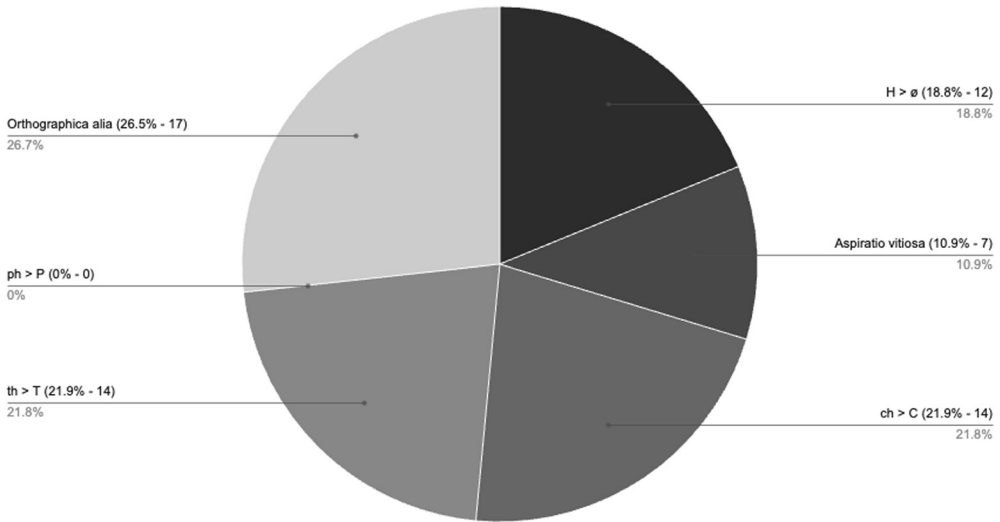


Figure 7. Aspirational errors in the VI. century AD

are attested without aspiration. For the $th > t$ change, the remaining three examples are recorded in common nouns: *ACOLT* = *acoluthi* (LLDB-82679), *AVTENTA* = *authentae* (LLDB-47703), *SPATARIVS* = *spatharius* (LLDB-37488). For the $ch > c$ process there is more diversity: *ARCIATER* = *archiatros* (LLDB-131087), *ARCIPRESBYTERO* = *archipresbytero* (LLDB-108689), *SCOLAE* = *scholae* (LLDB-108506) and *CRISTI* = *Christi* (LLDB-107591).

Examples of the loss of *h* (12 records, 18.8%) are familiar from earlier centuries, e.g. the name Hippolytus/a is used five times (LLDB-41496) or twice in the case of *habeo* (LLDB-40285).

The situation is quite similar with the incorrect aspirations (7 records, 10.9%). The forms *HELPIDIVS* = *Elpidius* (LLDB-108697), i.e. derivatives of the Greek name hope, and *STHEFANVS* = *Stephanus* (LLDB-56167) still hold. The latter is particularly revealing. Since *ph* is already transcribed as *f* in the ‘standard’ way, but it seems as if the transcriber still had some trace of a memory of the presence of aspiration within this name, and therefore mistakenly linked it to the letter *t*. This is the only period where we do not know of a single example of $ph > p$ change.

5. CONCLUSIONS

First, the frequency of the misuse of *h* has been compared with consonantism, i.e. with errors that have a real linguistic background, see Table 1 below. The data reported in this section include all aspiration errors, i.e. *aspiratio vitiosa*, $h > \emptyset$, $ch > c$, $ph > p$, $th > t$. The comparison with consonant errors seems logical – as *h* was once a consonant as well –, since the database is intended to measure linguistic processes, in accordance with its purpose and basic function. The other spelling errors category (*orthographica alia* on the previous figures) includes all errors that affect the consonant system but were excluded from the consonantism category since these

Table 1. Comparison of consonantal, aspiration and other spelling errors

Period	Consonantism	Aspiration	Other spelling
1st century BC	56.1%	20.5%	23.4%
1st century AD	52.1%	34.3%	13.6%
2nd century AD	60.3%	23.5%	16.2%
3rd century AD	67.6%	14.4%	18%
4th century AD	77.5%	10.2%	12.3%
5th century AD	76.5%	11.6%	11.9%
6th century AD	70.7%	21.5%	7.8%

errors were purely orthographic ones e.g. PH~ F, g > C, qu > CV, c > K, k > C, x > SX / CS / XS / XSS / XX.

Based on the comparison, the following conclusions can be drawn. In the pagan period – first to third centuries AD – the percentage of consonantal errors with a real linguistic background is around 60%, while in the late Christian period – fourth to sixth centuries AD – it is around 75%. This increase contradicts the logical assumption that orthographic errors should account for a higher proportion of the late centuries, when the Roman Empire was hit by barbarian invasions and migrations, followed by the collapse of the Empire, which did not serve to increase literacy.

Obviously, these proportions do not prove that spelling standards rose in the late period, but that the processes of linguistic change put such pressure on language users that the rate of linguistic error increased consequently during the centuries of the Christian era! However, comparison with consonantism has proved an inadequate standard for measuring changes in literacy levels reflected in the uncertainties in the use of *h*.

Second, the proportion of aspiration errors was compared with the proportion of non-linguistic errors, see [Table 2](#) below:

Table 2. Comparison of non-linguistic, aspiration and other spelling errors

Period	Non-linguistic	Aspiration	Other spelling
1st century BC	26.8%	34.1%	39.1%
1st century AD	41.6%	41.8%	16.6%
2nd century AD	45.7%	32.2%	22.1%
3rd century AD	59.1%	18.2%	22.7%
4th century AD	62.8%	16.8%	20.4%
5th century AD	57.7%	20.9%	21.4%
6th century AD	49.9%	36.8%	13.3%

Errors of non-linguistic origin – in fact, all typos – are themselves a kind of literacy category, like spelling, but these are realised at a different level, in the physical aspects of the textual execution of the inscription: the quality of the writing, the quality of its recording and execution. Unfortunately, this comparison did not provide a more accurate understanding of the dynamics of aspiration errors for two reasons: firstly, the drastic deterioration in the technical quality of the inscriptions is clearly visible. Thus, this analysis resulted in a similar situation to the comparison with consonantism. Secondly, it would be better to have a basis for comparison that is completely independent of the level of education, to avoid comparing the phenomenon with ‘itself’. Thus, comparisons with non-linguistic errors have also not proven effective for a better understanding of the dynamics of aspiration processes.

Based on these results, it was concluded that there was no appropriate comparison in the LLDB to infer the level of literacy of the inscriptions and their temporal changes and dynamics over time. This meant that there was no other option but to compare the data on aspiration errors with external sources. The best approach was to use epigraphical material of the *Epigraphik-Datenbank Claus / Slaby* – the data analysed reflect the state of the EDCS on 27 November 2024. We extracted the datable Latin, Latin and Greek inscriptions by century, excluding inscriptions in Greek only and inscriptions on pottery. The latter was required because the ceramic inscriptions were published in CIL volume 15, but this volume has not yet been processed by the LLDB research team. The occurrences of *h*'s were summed up in the inscriptions excluding all the *h*'s in the meta-texts, e.g. *christianae*, *honorarii*, *sarcophagus* or publication titles, as well as the *h* appearances in different corrections – {}, <> – and in fully written out forms – (). The *h*-s in additions – [] – could not be eliminated with complete certainty, therefore in the end all such *h*-s were retained and considered correct.

In a database of this size, there are bound to be misspellings due to inaccurate data entry. I experienced this myself on several occasions while building the LLDB. But it is not an exaggeration to assume that there are not enough misspellings to affect the comparison in any meaningful way, i.e. the correctly spelt *h*'s were compared with the misspelled or unspelled *h*'s recorded in the LLDB-Database, thus understandably reflecting changes in literacy levels (Table 3).

Table 3. Comparison of EDCS and LLDB-Database data

Period	EDCS number of insc.	EDCS H-number	LLDB H > ø	LLDB H > 0 number as a percentage of EDCS insc. number	LLDB H > 0 number as a percentage of EDCS H-number
1st century BC	1,188	1,033	28	2.4%	2.7%
1st century AD	11,526	7,927	376	3.4%	4.7%
2nd century AD	6,510	6,000	900	13.8%	15%
3rd century AD	4,805	3,323	200	4.2%	6%
4th century AD	14,939	2,648	625	4.2%	23.6%
5th century AD	2,851	1,347	282	9.9%	20.9%
6th century AD	305	411	47	15.4%	11.4%

Table 4. Comparison of EDCS and LLDB-Database data by periods

Period	EDCS number of insc.	EDCS H-number	LLDB H > ø	LLDB H > 0 number as a percentage of EDCS insc. number	LLDB H > 0 number as a percentage of EDCS H-number
1st–3rd centuries AD	22,841	17,250	1,476	6.5%	8.6%
4th–6th centuries AD	18,095	4,406	954	5.3%	21.7%

The comparison shows a clear upward trend from the first century BC to the second century AD, cf. 2.7% < 4.7% < 15%. In the third century AD normative knowledge and normative usage of h seems to have been stabilised, possibly as a result of the distortion caused by the previously mentioned dating problems, on this see *Methodology*. From the fourth century AD, with a small increase, the decline in standards became more pronounced, continuing in the fifth century AD and then moderating slightly in the sixth century AD. But by then the data are so limited compared to previous centuries that their validity is questionable.

Due to the difficulties of dating mentioned in *Methodology* and the limited data from the sixth century AD, it is reasonable to divide the data into two periods, one pagan – first to third century AD – and one Christian – fourth to sixth century AD. In this case the change in the aspiration trend is quite impressive (Table 4).

Overall, therefore, the percentage of the EDCS H number in the LLDB H > ø increased from 8.6% in the pagan era to 21.7% in the Christian era. This comparison also shows that it is not possible to relate our aspiration data to the number of inscriptions alone.

To summarize the results: by analysing aspiration errors by century through the Roman inscriptional material, we have argued (1) that there are significantly more errors due to consonantism in the late Imperial period than in its first three centuries, since linguistic changes became more rapid, intense and widespread toward the end; (2) that the quality of the technical execution of the inscriptions steadily deteriorated over time; (3) that the proportion of aspiration errors more than doubled from the pagan to the Christian era (from 8.6% to 21.7%); (4) that the comparison of correct and incorrect forms is a valid measure of the level of spelling and thus of the level of literacy, and their change.

REFERENCES

- Adamik, Béla. 2009. A latin nyelv története. Az indoeurópai alapnyelvtől a klasszikus latinig [History of the Latin language. From the Proto-Indo-European language to classical Latin]. Budapest: Argumentum kiadó.

- Adamik, Béla. 2012a. A császárkori feliratok vulgáris latin nyelvi adatainak dialektológiai érvényessége [The dialectological validity of the Vulgar Latin data of imperial inscriptions]. *Antik Tanulmányok* 56. 91–105.
- Adamik, Béla. 2012b. In search of the regional diversification of Latin: Some methodological considerations in employing the inscriptional evidence. In Fr. Biville, M.-K. Lhommé and D. Vallat (eds.) *Latin vulgaire – latin tardif IX.: Actes du IXe colloque international sur le latin vulgaire et tardif*, Lyon, 6–9 septembre 2009. Lyon: Maison de l’Orient et de la Méditerranée. 123–139.
- Adamik, Béla. 2015. The periodization of Latin: An old question revisited. In G. V. M. Haverling (ed.) *Latin linguistics in the early 21st century: Acts of the 16th International Colloquium on Latin Linguistics*, Uppsala, June 6th–11th, 2011. Uppsala: Uppsala Universitet. 640–652.
- Adams, James Noel. 2003a. *Bilingualism and the Latin language*. Cambridge: CUP.
- Adams, James Noel. 2003b. ‘Romanitas’ and the Latin language. *The Classical Quarterly* 53. 184–205.
- Adams, James Noel. 2007. *The regional diversification of Latin, 200BC – AD 600*. Cambridge: CUP.
- Allen, W. Sidney. 1978². *Vox Latina. A guide the pronunciation of classical Latin*. Cambridge: CUP.
- Allen, W. Sidney. 1987³. *Vox Graeca. A guide the pronunciation of classical Greek*. Cambridge: CUP.
- Biville, Frédérique. 1990. *Les emprunts du latin au grec. Approche phonétique, Tome I. Introduction et consonantisme*. Louvain & Paris: Peeters.
- Buck, Carl Darling. 1955. *The Greek dialects*. Chicago, IL: UCP.
- Coleman, Robert. 1963. Two linguistic topics in Quintilian. *The Classical Quarterly* 13. 1–18.
- Coleman, Robert. 1986. The central Italic languages in the period of Roman expansion. *Transactions of the Philological Society* 84. 100–131.
- Coleman, Robert. 2000. The contribution to vulgar Latin in Italy from other languages in the peninsula; Some case studies. In J. Herman, A. Marinetti and L. Mondin (eds.) *La preistoria dell’italiano. Atti della Tavola Rotonda di Linguistica Storica. Università Ca’ Foscari di Venezia, 11–13 giugno 1998*. Tübingen: Niemeyer 23–39.
- Cser, András. 2016. *Aspects of the phonology and morphology of classical Latin. Akadémiai doktori értekezés. Doctor of the Hung. Acad. of Sci. thesis*. Budapest.
- Cser, András. 2020. *The phonology of classical Latin*. Oxford: Wiley-Blackwell.
- Dascalu, Ioana-Rucsandra. 2019. *Les consonnes aspirées dans les textes latins tardifs*. *Acta Antiqua Academiae Scientiarum Hungaricae* 59. 45–51.
- Ellis, Robinson. 1876. *A commentary on Catullus*. Oxford: OUP.
- Fordyce, Christian James. 1978^{rev}. *Catullus*. Oxford: OUP.
- Gaeng, Paul A. 1968. *An inquiry into local variations in vulgar Latin, as reflected in the vocalism of Christian inscriptions*. Chapel Hill, NC: UNCP.
- Gignac, Francis Thomas. 1975. *A grammar of the Greek papyri of the Roman and Byzantine periods. Vol. I. Phonology*. Milano: Cisalpino & La Goliardica.
- Grebe, Sabine. 2001. Views of correct speech in Varro and Quintilian. *Journal of Latin Linguistics* 6. 135–164.
- Herman, József. 2000. Differenze territoriali nel latino parlato dell’Italia tardo-imperiale: Un contributo preliminare. In J. Herman, A. Marinetti and L. Mondin (eds.) *La preistoria dell’italiano. Atti della Tavola Rotonda di Linguistica Storica. Università Ca’ Foscari di Venezia 11–13 giugno 1998*. Tübingen: Niemeyer. 123–135.
- Herman, József. 2003. *Vulgáris latin. Az újlatin nyelvek kialakulásának útja [Vulgar Latin. The evolution of the Romance languages]*. Budapest: Tinta kiadó.

- Jeffery, Lilian Hamilton. 1961. *The local scripts of archaic Greece. A study of the origin of the Greek alphabet and its development from the eighth to the fifth centuries B. C.* Oxford: OUP.
- Leumann, Manu. 1977². *Lateinische Laut- und Formenlehre.* München: C. H. Beck.
- Mancini, Marco. 1990. *Aspirata greche e geminate latine.* Viterbo: Istituto di Studi romanzi.
- Meiser, Gerhard. 2002². *Historische Laut- und Formenlehre der lateinischen Sprache.* Dramstadt: Wissenschaftliche Buchgesellschaft.
- Miller, Gary D. 2014. *Ancient Greek dialects and early authors. Introduction to the dialect mixture in Homer, with notes on lyric and Herodotus.* Boston & Berlin: De Gruyter.
- Papini, Alessandro. 2020. A preliminary investigation on the <ae>/ <e> graphemic oscillation in Latin inscriptions from Rome. The relationship between vowel alternations, lexical stress and syllabic structure. *Acta Classica Universitatis Scientiarum Debreceniensis* 56. 107–123.
- Porzio Gernia, Maria Luisa. 1974. Vicende storiche e culturali dell'aspirazione latina. *Archivio glottologico italiano* 59. 56–102.
- Powell, Jonathan G. F. 2007. A new text of the Appendix Probi. *The Classical Quarterly* 57. 687–700.
- Sihler, Andrew L. 1995. *New comparative grammar of Greek and Latin.* Oxford & New York, NY: OUP.
- Sommer, Ferdinand. 1902. *Handbuch der lateinische Laut- und Formenlehre.* Heidelberg: Winter.
- Thomson, Douglas F. S. 1997. *Catullus. Phoenix supplementary, Vol. 34.* Toronto: UTP.
- Traina, Alfonso and Giorgio Bernardi Perni. 1998⁶. *Propedeutica al latino universitario.* Bologna: Pàtron.
- Traina, Alfonso. 2002⁵. *L'alphabeto e la pronunzia del latino.* Bologna: Pàtron.
- de Vaan, Michiel. 2008. *Etymological dictionary of Latin and the other Italic languages.* Leiden & Boston, MA: Brill.
- Väänänen, Veikko. 1981³. *Introduction au latin vulgaire.* Paris: C. Klincksieck.

Open Access statement. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited, a link to the CC License is provided, and changes – if any – are indicated. (SID_1)