ORTHOGRAPHY: ADJUSTMENTS, REFORMS, FAILURES

Ioana COSTA
University of Bucharest
ioana.costa@lls.unibuc.ro

Abstract: The writing systems were borrowed by people speaking dissimilar languages, characterised by phonetic features that not entirely matched the original set of graphemes. Discrepancies that emerged whenever a foreign writing system was borrowed are constantly obliterated by graphic conventions inside each language. Greek, Latin, nevertheless Romanian graphemes attest the adjustments made throughout the history.

Keywords: writing, alphabet, Greek, Latin, Romanian

Writing was one of the greatest human inventions, being, nevertheless, the genuine creation of a limited number of people, living in different areas and different epochs. Just as any other major invention, the writing systems were borrowed by people speaking dissimilar languages, characterised by phonetic features that not entirely matched the original set of graphemes. Discrepancies that emerged whenever a foreign writing system was borrowed are somehow obliterated by graphic conventions inside each language, without being essentially annihilated. Adjustments were made throughout the history of individual languages and only some of the attempts actually survived.

Among huge failures in adopting a writing system is to be considered the so called “Linear B”, used for writing the oldest form of Greek language, the Mycenaean. It was a rudimentary syllabic script, allowing specific signs for each of the five vowels and all the possible combinations of one consonant followed by one vowel. This syllabic pattern is highly improper for the Greek language, characterised by frequent consonant clusters and regular final consonant. Adopting this syllabic script meant predominantly some bogus vowels, *id est* the use of signs for consonant+vowel as simply consonant signs. Consequently this script was abandoned, leaving no traces in the standard Greek writing, so that deciphering the Linear B became a key event of classical and comparative philology of the mid twentieth century, involving the outstanding skills in philology, palaeography and, chiefly, cryptography of Michael Ventris and John Chadwick.

Some other borrowed sets of graphemes were more successful, without ever being completely suitable. Both the Greek and Latin alphabets display adjustments that sometimes have unexpected results when comparing their analogous lists of letters. Striking discrepancies are connected to letters *H* and *X*, present in both alphabets, with totally different phonetic values. The relationship between the two scripts is

BDD-A24375 © 2016 Ovidius University Press
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-27 06:04:42 UTC)

obvious, either as direct descendancy of Latin alphabet from the Greek one or as common inheritance of a previous (Phoenician) writing system. The *X* letter belongs to the two major Greek writing systems, the Eastern and the Western alphabet: as signifying a cluster of consonants, a digraph (*C* and *S*), in the Western alphabet, and an aspirated consonant (*CH*), in the Eastern one. They are obviously different results of recycling a grapheme that did not belong to the original set of letters borrowed from the Semitic alphabet; a visible indication of this chronology is placing the letter by the end of the alphabet, were the adjoined graphemes are inserted. The Latin graphical inheritance maintained the digraph value of letter *X*.

The other ambivalent letter mentioned above, *H*, has a more remarkable development. Once abandoned by the Greek writing system, this letter was eventually used as a conventional sign for the new phoneme that emerged in the Ionic-Attic dialect, after the long vowel *A* gradually changed, resembling more and more to a long *E*. It primarily had the value of a consonant, *H*, just as it remained in the Latin alphabet and the subsequent writing systems. Nevertheless, it's manifold use that implied a weak connection between the original value and the assumed values in different languages, in different epochs, allowed its presence as a simple graphic sign for denoting a specific phonetic value, as *TH* (Lat. *theatrum*), an aspirated consonant that virtually existed in Latin language only as a borrowed phoneme, of explicit Greek origin. This development was possible due to the diminished phonetic value in Latin, were the *H* letter was used for denoting a weak consonant, that, for instance, did not count as a full consonant in verse (allowing the elision of the adjoined vowels), eventually becoming a mute sound in Italian (*vide* the saying *non vale un' acca*: even the name of the letter was deprived by the original sound, *H*). This usage is consistent with the Ancient Greek rough breathing, *spiritus asper*, a diacritical mark, less than half a consonant, signified by half (or less than half) a letter.

Recycling vacant graphemes for signifying specific phonemes (frequently those phonemes that emerged in the history of individual phonetic history) is one possible solution of dealing with a borrowed alphabet. Remodelling an old grapheme is another solution and the history of Latin *G* letter is eloquent. Most likely due to the Etruscan writing system, that intermediated the adoption of Latin alphabet, the latter lacked the graphic distinction between voiced and voiceless stops, employing one single letter for both *C* and *G*; this situation is still visible in the traditional abbreviation of the Roman *praenomen Gaius* as *C* or *Gnaeus* as *Cn*. (*praenomen* of Etruscan origin). The archaic epigraphic testimonies attest this unique grapheme for either *G* or *C*, e.g. *VIRCO* in the *Duenos* inscription (seventh to fifth century BC). Improper for the Latin phonetic inventory, this ambivalence was annulled by a formal distinction between the two letters. The person credited with this invention is Spurius Carvilius Ruga, a freedman that lived in the third century BC (*fl*. 230): in his elementary school (the first fee-paying school in Rome) the two letters were

BDD-A24375 © 2016 Ovidius University Press
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-27 06:04:42 UTC)

for the first time consistently used for the voiced and voiceless velar plosives. The original *C* letter is obviously the Greek *gamma*, as the epigraphic form reveals, as much as the position in the Latin alphabet, coming after *A* and *B*, similar to *gamma* that comes after *alpha* and *beta*. The formal distinction is a stroke, similar to a diacritical mark. Positioning the new letter in the alphabet had to observe the rigid succession of letters due to their Greek values as numerals: it occupied the available place created by dropping of the old letter *Z*, which seemed to have been removed by the Roman censor Appius Claudius Caecus (third century BC), who found it similar to the teeth of a corpse, as Martianus Capella notes in the third *liber* of his *De nuptiis Philologiae et Mercurii* (par. 261), the chapter granted to the first of the liberal arts, the Grammar (*Grammatike* in Greek, term coined on *grammata*, "letters", *litterae* in Latin, the starting point of the term *Litteratura*).

Phonetic development of Romanian language altered some of the Latin vowels into new phonemes: "posterior-*i*" replaced some previous *i*, *a*, *e*, *u* or *o*. A new grapheme was required. There was an intermediate stage of writing, with composite alphabet, both Latin and Cyrillic. Subsequently some diacritics (the circumflex glyph) improved the basic letters *i*, *a*, *e*, *u*: *î*, *â*, *ê* and *û* were graphemes used etymologically for a single phoneme ("posterior-*i*"). Four distinct orthographic reforms (in 1904, 1932, 1953, and 1965) finally reduced the four graphemes to a single one (*î*). This approach, concerning the phonetic principle, was calibrated in 1965, admitting an etymological (and significant) exception: *român* ("Romanian") and the connected words.

The beginning of the '90s became a wide-ranging field for change. Orthography turned out to be a target: the regular grapheme *î* was largely replaced by *â*, in a manner that combined different principles (phonetic and etymological), including the position inside the word. Alf Lombard was asked to offer a specialised opinion: he wrote a dense text regarding the history of modern Romanian orthography and concluded with a plea against the proposed reform. Nevertheless, this reform became official in 1993.

Alf Lombard's arguments spring both from a phonetic approach and the history of Romanian orthography. The historical approach is a succinct account of the rather numerous orthographic reforms (no less than 41 only between 1780 and 1880, as listed by Gheorghe Adamescu, with several other major restructurings). All of them mirrored a quest for balance between the etymological (or historical and etymological) principle and the phonetic one, which, step by step, became predominant, in a pervasive tendency toward a natural orthography. Regarding the letter or letters assigned for the *close* (or *high*) *central unrounded vowel* (as in *înot, gând, hotărî*), the reforms that occurred during the twentieth century implied different solutions. This particular phoneme does not belong to the regular Indo-European inventory, but occurs as an allophone in several Indo-

BDD-A24375 © 2016 Ovidius University Press
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-27 06:04:42 UTC)

European individual languages – some Slavic languages (as Russian and Czech), some German languages (including Swedish), some Celtic languages (Irish and Welsh), some Romance language (Romanian and Portuguese) – as well as in other non-Indo-European languages. The symbol in the International Phonetic Alphabet is a letter *i* with a horizontal bar ("barred-*i*"). The phonetic features are commonly depicted in terms of height (the tongue is positioned close to the roof of the mouth, nevertheless without creating a constriction: so it is equally described as *high* or *close*), backness (the tongue is positioned halfway between a front vowel and a back vowel: so it is *central*) and roundedness or vocalic labialization (the lips are not rounded: so it is an *unrounded* vowel). Alf Lombard describes the phoneme as *posterior-i*, which is not fundamentally different, maintaining the *genus proximum* [i] (*Limba română* 1992, p. 532 *vide* also: *La pronunciation du roumain* 1933, p. 105 and p. 122-124 (medium vowel), *La langue roumain* 1974, par. 5 and 32); letter *û* was assigned to a limited number of forms, mostly four forms of the verb "to be", *a fi*, in present tense: first person singular and plural *eu sûnt, (noi) sûntem,* second person plural *voi sûnteţi* and third person plural *(ei) sûnt*; also the nouns *adûnc, adûncime* etc. (The latter is unclear, *vide* Hasdeu *adânc*, EM, *vs sîntem*, virtually paralleled by Italian *adincus*, besides the classic form *aduncus*, *Limba română* 1992, p. 534).

In the history of Romanian orthography, the *posterior-i* was assigned to several letters, in time or simultaneously, *e.g.* lână **lana,** *câmp* **campus,** *ânger* **angelus,** *vênt* **ventus,** *avênd* **habendo,** *tênĕr* **tener,** *rîde* **ridet,** *rîu* **rivus,** *în* **in,** *hotârî* (composite form including the Latin verb **ire**).

The status of the forms of the verb "to be", *a fi*, is peculiar. As Alf Lombard briefly mentiones (*Limba română* 1992, p. 534-535), from the corresponding Latin forms (1) **sum**, (2) **sumus**, (3) **estis**, (4) **sunt**, the XVIth century Romanian employed: (1) *sănt/sămt/sint,* (2) *sem/săm,* or (by the end of the XVIth century) *săntem/ sîntem/sintem,* (3) *seţi/set/siţi*, or (by the end of the XVIth century) *sănteţi/ sînteţi/sinteţi,* (4) *sănt/sămt/sint*. The transfer from Latin to Romanian implied several stages; it is obvious that the tradition was broken for the second and third forms mentioned here. On the other hand, the orthography is significant: the letter *u* in these forms is lowly attested in their history, outclassed by the *ă* and (mostly before the cluster *nt*) *î* forms, so that they seem to point to a *posterior-i* pronunciation. The recurrent form *sunt* is part of a cultural (not linguistic) process of re-enacting the Latin origins, highly envisaged by the Latinist trend of the XIXth century; nevertheless, the *u*-forms are attested in some subdialects of modern Romanian language. The standard Romanian attested *posterior-i* forms, basically descending from the Latin subjunctive mood: **sim**, **sis**, **sitis**, **sint**.

The plethora of letters assigned to *posterior-i* was diminished in time. The four letters (*â, ê, î, û*) were reduced: the first to be discarded were *ê* and *û* (with the exception of *u*-forms of the verb "to be", recurrent in

BDD-A24375 © 2016 Ovidius University Press
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-27 06:04:42 UTC)

cultivated speech). The 1932 rules generalised the *â*-writing, maintaining the letter *î* for the initial vowel (e.g. *împărat, îngust*) and the final vowel in verbs and the connected forms (*hotărî*, *hotărît*, *hotărîtor*), *vide* S. Puşcariu and T. A. Naum, *Îndreptar şi vocabular ortografic*, "Cartea românească", 1932. Aiming to observe the historical and etymological principle, as in *lână* for Latin **lana**, *împărat* for Latin **imperator**, it was still confusing, as *înger* for **angelus**, *îngust* for **angustus**, *sân* for **sinus**, *râu* for **riuus** were obvious deviations.

Alf Lombard praised the sound principles displayed by Candrea and Adamescu in their illustrated encyclopaedic dictionary (*Dicţionarul enciclopedic ilustrat "Cartea Românească*", 1931): *lînă, înger, rîde, român*. All these personal rules strictly observed in that dictionary finally became the norm in 1953 and 1965. Actually, September 16th 1953 was the day the phonetic rule triumphed over the (mixed) etymological rule. The exception of *român-România* and the connected words (1965) was a necessary reverence for the national identity.

The debate over the topic of *î* vs *â* occasioned a concise and steady answer of Alf Lombard: there are no solid reasons to assign two distinct letters for one and only phoneme and, more over, the etymological principle could not be observed by this simple couple of letters. The solution (*Limba română* 1992, p. 538) ought to be simple, easily put in rules, obeying the principle one sound-one letter, obeying the tradition of the language. All these requirements could not be observed simultaneously, so that one sacrifice ought to be made. The most reasonable is to sacrifice the etymology, mostly as it simply can not be always displayed, for several reasons (graphical complexity, objective incertitude, non-Latin origin). Although **lana**, **ventus** and **ridet** include three distinct Latin vowels, their Romanian outcome displays one and only sound: *-î-*.

The topic is to be found in various studies, *e.g*. the series of relatively short studies hosted by *România literară* in 2002 (nr. 38-42), written by linguists and cultural personalities: Dumitru Irimia, Matilda Caragiu-Marioţeanu, Nicolae Manolescu, Sorin Mărculescu, George Pruteanu, Victor Iancu. The topic is also implicitly approached by the writings that consistently display the previous orthographic system, which is not allowed by the rules of the Romanian Academy (and the publications placed under its aegis), but is accepted by some of the best Romanian publishing houses and some major periodicals.

The analysis of Professor Alf Lombard deserves being re-enacted mostly in the light of the new *Orthographic, orthoepic and morphological dictionary* (DOOM 2005): it highly welcomes the natural tendencies of the everyday Romanian language, including syllabic units in compounds. A word like *capîntortură*, outcome of three distinct words, reveals its triad

BDD-A24375 © 2016 Ovidius University Press
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-27 06:04:42 UTC)

structure by the presence of letter *-î-* in the middle of it. The recommended first two syllables (*ca-pîn-* and, only as a second recommendation, *cap-în-*) opens the gates for forgetting the etymology. All that remains is the phonetic reality, virtually admitting a ***capântortură* writing. Identically *răsînţelegere*, recommended as *ră-sîn-* and only secondly, *răs-în-*, virtually allows a ***răsânţelegere* writing. From this situation to *hotărât* gliding to ***hotărâ* seems to be only one small step – small step for some words, but a terrifying giant leap for the orthography, *i.e*. the 1993 orthography. Accepting that phonetics prevails over etymology, might vividly attest the accuracy of Alf Lombard's diagnosis, who, to summarize, discarded the presence of two distinct letters for one and only phoneme and, facing the graphic alternative of using *a* vs *i* to designate this specific phoneme, explicitly preferred the *i*-form, for several reasons, including the phonetic features that are more similar to the *close* vowel *i* than to the *open* vowel, *a*, leading to labelling it as *posterior i*.

The long history of choosing letters to match phonemes in various languages is never a straight, one-way journey. Attempts are to be made in order to find the best possible solutions and failures are to be accepted in order to observe the natural – and implicitly simple – development of language and writing.

Bibliography

ADAMESCU, Gheorghe, 1931: *Probleme actuale ale ortografiei româneşti. Cu un adaos cuprinzînd "Regulile ortografiei româneşti" şi un "Glosar"*, Bucureşti : Socec.

ADAMESCU, Gheorghe, 1932, *Ortografia românească după modificările din 1932. Reguli şi explicări urmate de un glosar*, Bucureşti : Socec.

CARDONA, Giorgio Raimondo, 1987, *Storia universale della scrittura*, Milano : Mondadori.

COSTA, Ioana, 2008[2], *Fonetică istorică latină*, Bucureşti : Editura Universităţii din Bucureşti.

COSTA, Ioana, 2011, *Papirus, pergament, hîrtie. Începuturile cărţii*, Bucureşti: Humanitas.

DOOM (*Dicţionarul ortografic, ortoepic şi morfologic al limbii Române*), 2005[2], Bucureşti: Univers Enciclopedic.

JEAN, Georges, 1986, *L'écriture, mémoire des hommes*, Paris : Gallimard.

JEAN, Georges, 1992, *La scrittura, memoria degli uomini*, Trieste : Electa/Gallimard.

IRIMIA, Dumitru, 2002, « De ce scriu şi susţin scrierea cu î din i? », in: *România literară*, 49.

LOMBARD, Alf, 1935, *La prononciation du roumain*, Uppsala: A.-B. Lundequistska bokhandeln.

LOMBARD, Alf, 1974, *La langue roumaine. Une présentation*, Paris : Klincksieck.

BDD-A24375 © 2016 Ovidius University Press
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-27 06:04:42 UTC)

LOMBARD, Alf, 1992, « Despre folosirea literelor â şi î », in *Limba română*, 10, pp. 531-542.
SIHLER, Andrew L., 1995, *New comparative grammar of Greek and Latin*, New York, Oxford: Oxford University Press.

BDD-A24375 © 2016 Ovidius University Press
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-27 06:04:42 UTC)