

Universal Multiple-Octet Coded Character Set  
International Organization for Standardization  
Organisation Internationale de Normalisation  
Международная организация по стандартизации

**Doc Type:** Working Group Document  
**Title:** Proposal for the addition of four Latin characters to the UCS  
**Source:** Michael Everson and Chris Lilley  
**Status:** Individual Contribution  
**Date:** 2019-05-26

In 2012, document N4297R (L2/12-270R) proposed ten characters, eight of which were accepted for encoding. Two of these were not accepted, likely because the evidence available to us at the time was not particularly robust. This proposal requests the encoding of four Latin characters forming a casing pair. If this proposal is accepted, the following characters will exist:

|   |      |                                    |
|---|------|------------------------------------|
| Ð | A7D0 | LATIN CAPITAL LETTER TAU GALLICUM  |
| ð | A7D1 | LATIN SMALL LETTER TAU GALLICUM    |
|   |      | • Gaulish                          |
| Œ | A7D2 | LATIN CAPITAL LETTER S WITH STROKE |
| ſ | A7D3 | LATIN SMALL LETTER S WITH STROKE   |
|   |      | • Gaulish                          |

LATIN LETTER TAU GALLICUM is a letter which was devised as an extension to the Latin alphabet by native writers of the Gaulish language who lived in Roman Gaul. (Other Gaulish inscriptions use the North Italic or the Greek Alphabets) The TAU GALLICUM represents a dental affricate; most scholars believe it was [ts] or [st], though some have wondered if it might be [ð] or [θ]. Gaulish was originally written in Greek script and when Latin script began to be used for it, a gap was felt for what had been written with Greek theta. The solution they chose was to use the horizontal bar of Greek Θ within a Latin D. In another Gaulish-speaking area, the solution they chose was to put a horizontal bar through the Latin S. Modern scholars have tried various workarounds to represent the character, because there was no encoded character available. Various writers have used Icelandic eth (Ðð) or Croatian barred d (Đđ), as well as using strikethrough styled text with Latin Ð ð and even Greek Α δ̄. None of these workarounds are correct; the TAU GALLICUM was devised many centuries before either (Ðð) or Croatian barred d (Đđ), and the shapes of Ð are not the same as that of Ð anyway. In fact, the TAU GALLICUM is explicitly named in a verse by Vergil (70 BCE–19 BCE) in his *Catalepton* 2, 4:

Corinthiorum amator iste verborum,  
iste iste rhetor, namque quatenus totus  
Thucydides, tyrannus Atticae febris:  
tau Gallicum, min et sphin et – male illi sit,  
ita omnia ista verba miscuit fratri.

That lover of Corinthian words,  
that horrible, horrible rhetorician (For perfect  
Thucydides that he is, he is the tyrant of the Attic fever)  
his Gallic tau and min and sphin, and may he pay for it  
of all these words he mixed a potion for his brother.

It might be the case that part of the fun in this verse is that the repetition of “iste iste” refers to the /ts/~st/ (or even the non-Latin /θ/ or /ð/) represented by the TAU GALLICUM.

The Moro language of Sudan distinguishes dental /ǰ/ from alveolar /d/. Several orthographies exist for Moro; one writes the voiced dental as ǰ, one writes it as ǰ̄, and one as ǰ̇. Texts exist which use the TAU GALLICUM.

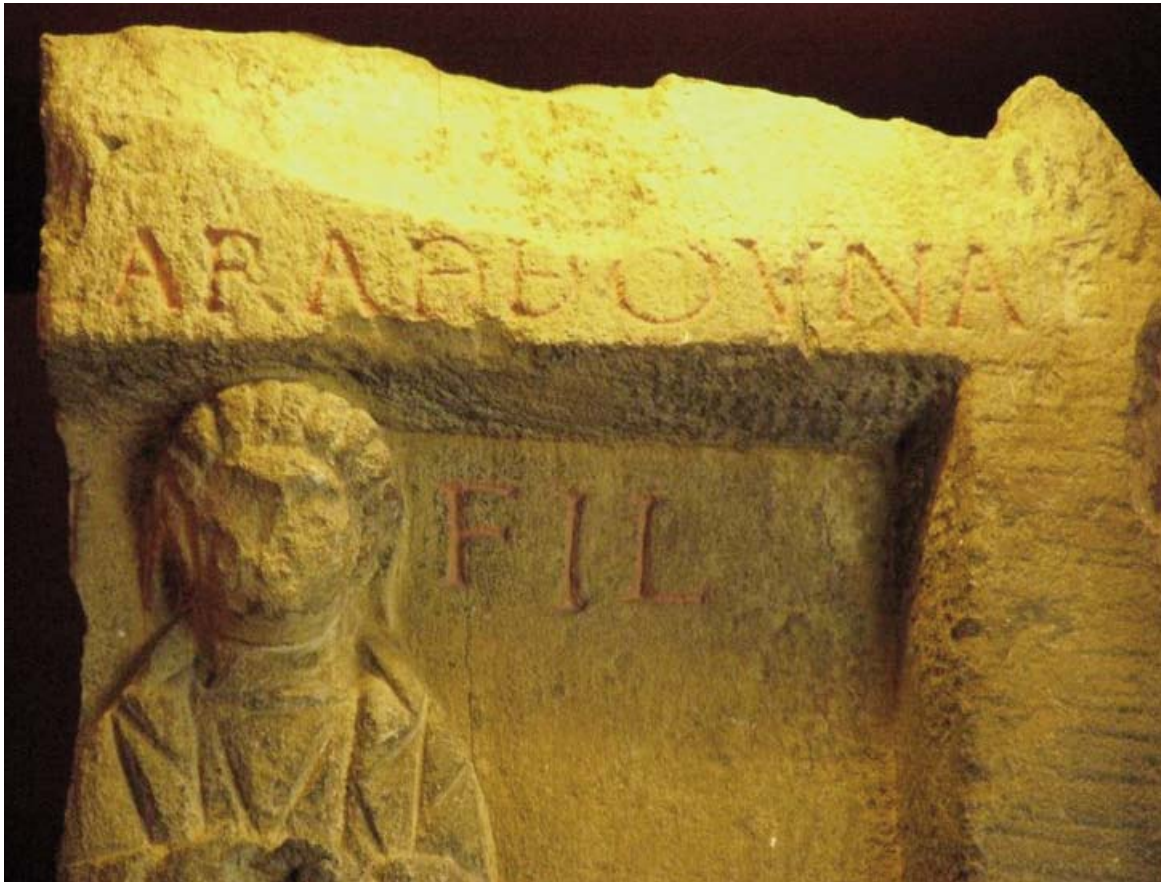
### Unicode Character Properties.

```
A7D0;LATIN CAPITAL LETTER TAU GALLICUM;Lu;0;L;;;;N;;;A7D1;
A7D1;LATIN SMALL LETTER TAU GALLICUM;Ll;0;L;;;;N;;;A7D0;;A7D0
A7D2;LATIN CAPITAL LETTER S WITH STROKE;Lu;0;L;;;;N;;;A7D3;
A7D3;LATIN SMALL LETTER S WITH STROKE;Ll;0;L;;;;N;;;A7D2;;A7D2
```

### Bibliography

- Blench, Roger. 2005. *A dictionary of the Moro language of the Nuba hills, Sudan*. Privately circulated PDF.
- Delamarre, Xavier, & Pierre-Yves Lambert. 2018. *Dictionnaire de la langue gauloise : Une approche linguistique du vieux celtique continental*. Arles: Éditions Errance. ISBN 978-2877726313
- Eska, J. F. 1998. “Tau Gallicum”, in *Studia Celtica* 32, pp 115-127. ISSN 0081-6353
- Guest, Elizabeth. 1997a. *Moro noun classes*. Privately circulated PDF.
- Guest, Elizabeth. 1997b. *Moro phonology*. Privately circulated PDF.
- Jufer, Nicole, & Thierry Luginbühl. 2001. *Répertoire des dieux gaulois*. Paris: Édition Errance. (ISBN 2-87772-200-7)
- Lambert, Pierre-Yves. 2018. *La langue gauloise: Description linguistique, commentaire d'inscriptions choisies*. Arles: Éditions Errance. ISBN 978-2877726337
- Mees, Bernard. 2002. “On Gaulish Tau”, in *Studia Celtica*, 36, pp. 21-26. ISSN 0081-6353
- Starnier, David. 2012. “Gaulish” in *Old Celtic Languages*. Spring 2012. (Evidently a course handout.)
- Tenny, Frank. 1935. “Tau Gallicum, Vergil, Catalepton II, 4”, in *The American Journal of Philology* 56, No. 3, pp. 254–256.
- von Florencourt, W. Ch.. 1851. “Unedirte Inschriften von Trier: Götterdenkmale”, in *Jahrbücher des Vereins von Altertumsfreunden im Rheinlande* 16. Pp. 63–64. Bonn: Verein von Altertumsfreunden im Rheinlande.

**Figures.**



**Figure 1.** A Gaulish tomb inscription with the name “ARADDOVNA”, illustrating the use of the TAU GALLICUM. When cased, the name would be written “Araddovna” according to modern practice. From the Museums of the Cour d’Or, Metz.



**Figure 2.** A Gaulish gold coin of the Trinovantes/Catuvellauni tribe. Although the inscription is worn, the inscription has been read as “ADDOMAROS” (“Addedomaros”). From The Celtic Coin Index, [http://www.celticcoins.ca/record.php?coin\\_id=620024](http://www.celticcoins.ca/record.php?coin_id=620024)  
Jahrbücher des Vereins von Altertumsfreunden im Rheinlande, Volumes 13-16



**Figure 3.** A Gaulish inscription dedicated to the goddess Sirona, beginning with “DEAE ðIRONAE”; compare Figure 7 below. from Saint-Avold, now in the Musée de Metz, inscription reference is CIL XIII, 04498  
 'Deae D-ironae / Maior Ma / giati filius/ vslm"

|      | a | b | c | d | e | f | g | l | m | n | o | p | r | s | t | u | x,χ | ̢ | θ |
|------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|-----|---|---|
| M.1  | λ | β | ϙ | δ | u | l | α | l | λ | μ | α | ρ | λ | σ | τ | υ | χ   | ̢ | θ |
| M.4  | λ | β | ϙ | δ | u | l | α | l | λ | μ | α | ρ | λ | σ | τ | υ | χ   | ̢ | θ |
| M.14 | λ | β | ϙ | δ | u | l | α | l | λ | μ | α | ρ | λ | σ | τ | υ | χ   | ̢ | θ |
| M.23 | λ | β | ϙ | δ | u | l | α | l | λ | μ | α | ρ | λ | σ | τ | υ | χ   | ̢ | θ |
| M.30 | λ | β | ϙ | δ | u | l | α | l | λ | μ | α | ρ | λ | σ | τ | υ | χ   | ̢ | θ |
| M.32 | λ | β | ϙ | δ | u | l | α | l | λ | μ | α | ρ | λ | σ | τ | υ | χ   | ̢ | θ |
| M.34 | λ | β | ϙ | δ | u | l | α | l | λ | μ | α | ρ | λ | σ | τ | υ | χ   | ̢ | θ |
| M.46 | λ | β | ϙ | δ | u | l | α | l | λ | μ | α | ρ | λ | σ | τ | υ | χ   | ̢ | θ |
| M.47 | λ | β | ϙ | δ | u | l | α | l | λ | μ | α | ρ | λ | σ | τ | υ | χ   | ̢ | θ |
| M.49 | λ | β | ϙ | δ | u | l | α | l | λ | μ | α | ρ | λ | σ | τ | υ | χ   | ̢ | θ |
| M.66 | λ | β | ϙ | δ | u | l | α | l | λ | μ | α | ρ | λ | σ | τ | υ | χ   | ̢ | θ |
| M.74 | λ | β | ϙ | δ | u | l | α | l | λ | μ | α | ρ | λ | σ | τ | υ | χ   | ̢ | θ |
| M.76 | λ | β | ϙ | δ | u | l | α | l | λ | μ | α | ρ | λ | σ | τ | υ | χ   | ̢ | θ |
| M.85 | λ | β | ϙ | δ | u | l | α | l | λ | μ | α | ρ | λ | σ | τ | υ | χ   | ̢ | θ |
| M.88 | λ | β | ϙ | δ | u | l | α | l | λ | μ | α | ρ | λ | σ | τ | υ | χ   | ̢ | θ |
| M.89 | λ | β | ϙ | δ | u | l | α | l | λ | μ | α | ρ | λ | σ | τ | υ | χ   | ̢ | θ |
| M.94 | λ | β | ϙ | δ | u | l | α | l | λ | μ | α | ρ | λ | σ | τ | υ | χ   | ̢ | θ |

III. 3.3.: Latin cursive script on pottery from La Graufesenque (from: RIG II-2, 370).

**Figure 4.** From Section 3.3 “Gaulish in Latin Script” of Stifter 2012. The analysis (in the rightmost column) of some of the forms as a proper rounded theta may be correct, but the analysis of the TAU GALLICUM as a barred delta must be mistaken, as clearly it is no different from the cursive Latin D and d with the addition of a horizontal stroke. (Note that a proper Latin THETA has yet to be encoded, though there is plenty of evidence for it having been borrowed into Latin for various purposes.)



#### 4. A THORNY PROBLEM OF GAULISH PHONOLOGY

The phonological system of Gaulish, possibly also of Lepontic, contains a specific sound that is traditionally called *tau Gallicum* after a passage in Vergil's *Catalepton* 2, 4:

*Corinthiorum amator iste uerborum,  
iste iste rhetor, namque quatenus totus  
Thucydides, tyrannus Atticae febris:  
tau Gallicum, min et sphin ut male illisit,  
ita omnia ista uerba miscuit fratri.*

That lover of Corinthian words,  
that... that rhetor! Even though being a complete  
Thucydides, he is a tyrant of the Attic fever:  
how he badly belched (?) the *tau Gallicum*, the *min* and *sphin*,  
thus he mixed all those words for his brother.

It is uncertain if the sound *tau Gallicum* that Vergil mentions is the same sound as the one for which the term is used today. Today it denotes a phoneme of only roughly known value that is represented by a great number of different spellings in Gaulish and possibly also Lepontic inscriptions:

Roman: **t, tt, th, tth, d, dd, ~~d~~, ~~dd~~, ts, ds, s, ss, ~~ss~~, sc, sd, st**  
Greek: **θ, θθ, σ, σσ, σθ, τ, ττ**  
Lepontic script: **san, zeta, sigma**

e.g.: ~~medu-~~, *messu-*, *μισσου-*, *medi-*, *μεθθ-*, *μεθ-*, *medsi-*, *medi-* < PIE \**medu/i-*

Wherever etymological speculations are possible, this phoneme, if it is one, goes back to earlier \**Ds*, \**st* and \**Dt* (*D* = any dental obstruent). Etymologically, it clearly corresponds to Insular Celtic *s* < \**ss* in word-initial and -internal position (against \**s* that first became \**h*, then Ø in Insular Celtic word-internally); e.g.:

PIE \**nezd-tamo-* (?) 'next': Gaul. ~~nedd~~amon, OIr. *nessam*, Cym. *nessaf*  
PIE \**med-tu-* 'judgement': Gaul. ~~medu-~~ etc., OIr. *mess*  
PIE \**melit-to-* 'sweet': Gaul. ~~melido-~~ etc., OIr. *milis*, Cym. *melys*  
PIE \**h<sub>2</sub>ster-* 'star': Gaul. ~~Dirona~~ (?), OIr. *ser*, Cym. *syr*  
PIE \**tud-to-* 'pushed': Gaul. *tuθθus* 'loads'  
PIE \**g<sup>o</sup>osti-* 'guest, stranger': Lep. *uvamokozis*, *χosioisio* (?)  
Pre-Celt. \**is-to-* 'this': Lep. *išos*

Sometimes *tau Gallicum* can also stand for strong, intensified (?) *s*:

PIE \**meh<sub>1</sub>ns* 'month', Gaul. *miš*, OIr. *mís*-  
also in acc. pl. Lep. *siTeš*, Cisalpine Gaul. *artuaš*?

and perhaps also for analogically re-introduced, strongly pronounced word-internal *s*, in opposition to regularly weakened (lenited?) intervocalic *s*:

Pre-Celt. \**b<sub>1</sub>is<sub>1</sub>je<sup>o</sup>* 'to want to become', Gaul. *bissiet* 'will be', *bissiete* 'you will be' (?)

No conclusive evidence for a comparable sound exists for Celtiberian. Inherited \**st* is retained in Celtiberian, e.g. *stena*, *boustom*. Other combinations of dental sounds and *s* may already have been simplified to mere *s(s)*. The many orthographic variants in Gaulish suggest that this sound had no direct equivalent in Latin and Greek, and that it featured a dental (because of the frequent spellings with *d, t, ...*) and a fricative component (because of *s, θ, ...*). The frequent double spellings and etymological considerations furthermore suggest gemination, i.e. prolonged pronunciation. Many phonetic suggestions for this sound have been proposed (see ESKA 1998: 116), but according to the *communis opinio* it probably was a geminate affricate [tʰ]. On the basis of a few forms where *tau Gallicum* in Gaulish cannot be derived from dental clusters or from \**st*, i.e.:

*eddic* 'and' (cp. *etic* < \**eti-k<sup>h</sup>e*, unless it is \**esti-k<sup>h</sup>e*)  
*gnatha* 'girl' (cp. *nata* < \**gn<sub>h</sub>to-* 'born')  
*madduro* (cognate with Lat. *maturus*?)  
[C]*athuboduae* (to Gaul. *catu-* 'battle, war')  
*bued* 'may be' (cp. *buet=id*, *deuorbuet=id*)

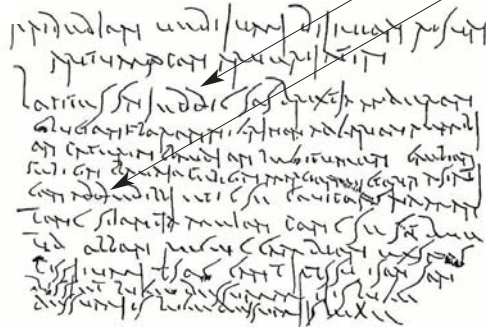
but where the sign perhaps represents lenited *t*, ESKA assumes that *tau Gallicum* may have stood for the so-called slit-*t*, a sound peculiar, for example, to Southern Hiberno-English.

**Figure 5.** From Section 4 "A thorny problem of Gaulish phonology" of Stifter 2012. In addition to the section title's fantastic pun, examples of lower-case TAU GALLICUM as well as S WITH HORIZONTAL STROKE can be seen.

6.2.8.3. L-100 (Chamalières, Puy-de-Dôme), lead tablet, magical, middle of 1<sup>st</sup> c. A.D.:

andedíon uediúúmí dííuion risun  
 artiumapon aruerífatín [or aritu?]  
 lopites sní eddic sos brixtia anderon [or lotites?]  
 clucionfloron nigrínon adgarion aemilí  
 on paterin claudíon legitumon caelion  
 pelign claudío pelign marcion uictorin asiati  
 con addedillí etic se couitoncnaman [oder poncnaman?]  
 tonc sífontío meíon ponc se sit bue  
 tid ollon reguccambion exsops  
 pissiúmí tsoe cantírtssu ison son [or rissuis onson?]  
 bissíet luge dessummiís luge  
 dessumíís luge dessumíís luxe

III. 6.55–56.: L-100 Chamalières (from MEID 1992: 41).



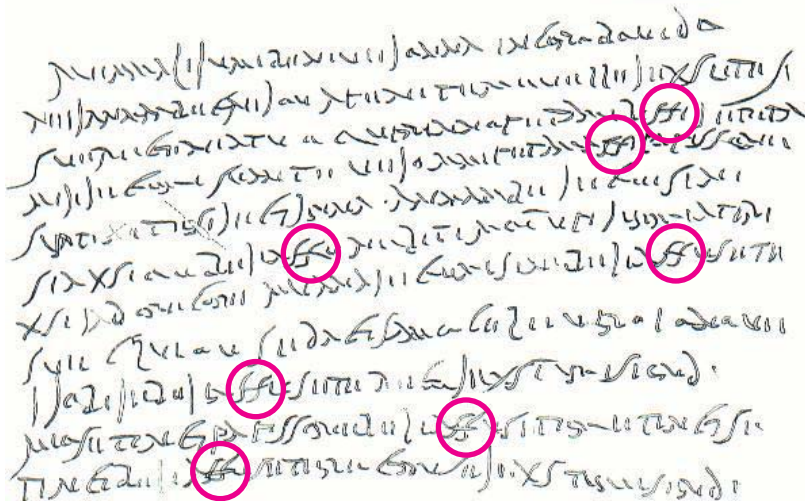
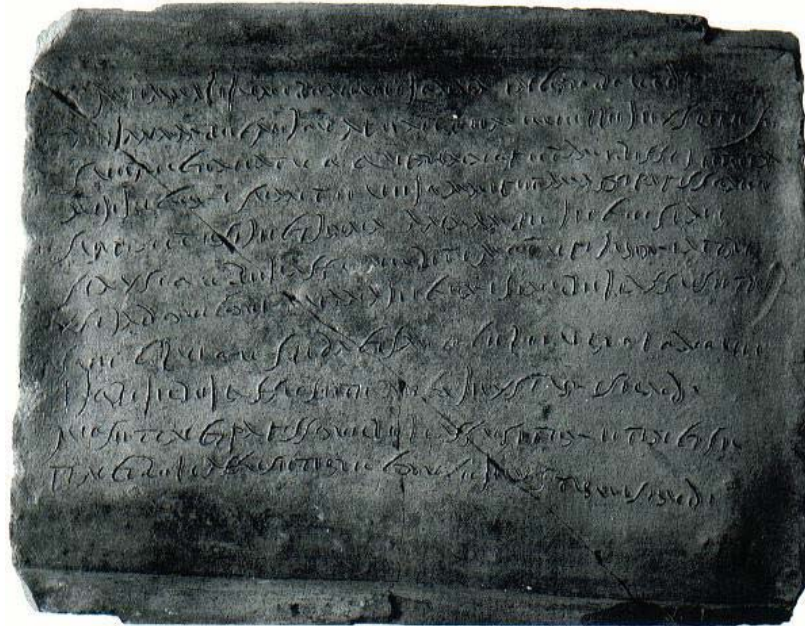
**Figure 6.** Photo, facsimile, and transcription of a Gaulish lead tablet with a magical text, including the words “eddic” and “addedillí”. Taken from Stifter 2012.

3. Unerbte Inschriften aus Erier.  
 I. Götterdenkmale.  
 1. Die Göttin Dirona (Sirona).  
 412. DEAE·BIRONA..  
 L·LVENIVS·CENSOR..  
 NV..SIGLLVM·D..  
 (Deae Dironae Lucius Lucanius Censorinus sigillum dono dedit.)  
 Inschrift eines Fussgestells von grauem Sandstein, von welchem noch Reste der Füße und der langen Gewandung des abgeschlagenen Götterbildes ersichtlich sind. Auch von einem neben demselben dargestellten Altar (oder einer ähnlichen Structur) sind Spuren vorhanden.  
 Die Inschrift ist wichtig wegen der Namensschreibung **BIRONA**, welche bisher nur auf einem einzigen Denkmale der Göttin — dem zu St. Avauld in Lothringen gefundenen Votivsteine (s. die Abbildung desselben bei Oberlin, Mus. Schoepfl. tab. I. 2.) vorgekommen war und zu irrigen Deutungen Anlass gegeben hatte. Oberlin (libr. I. pag. 15.) glaubte in der eigenthümlichen Formation des Initialzeichens (D) eine Ligatur der Buchstaben DE zu erkennen; Lehne (in der Abh. über die Göttin Sirona, Ges. Schr. III. S. 63.) hielt eine Buchstabenverwechslung des Steinhauers („Deirona statt Seirona“) für unzweifelhaft. Die Irrigkeit dieser Annahmen geht aus der gleichartigen Scriptur unsrer, mit augenscheinlicher Sorgfalt ausgeführten Inschrift hervor, in welcher überdies ein regelmässig eingezeichnetes E in dem Namen CENSORINVS sich findet.

**Figure 7.** An description from von Florencourt 1851 of text related to the goddess Sirona as BIRONA, showing a D with horizontal stroke through the whole of the letter (not just of the vertical as with Ð).

6.2.8.11. L-93 (Châteaubleau, Seine-et-Marne), tile:

1. nemnalífumi beni. uefonna incorobouido
2. nefanmanbe gniíou apeni temeuelle íexsetesi
3. sueregeniatu o quprinnopetamebissi íeteta.
4. mííí íégumi. suante uefommi petamassi papissone
5. suirexetesi íegííinna anmanbe íeguisini
6. siaxsiou. beííassunebiti moi upiíummiateri
7. xsi índore core. nuana íegumisini · beííassusete
8. sue cluio u sedagisamo cele uiro íonoue
9. ííobífe beííassusete rega íexstumisendi
10. me · setingi papissonebeííassusetemetingise
11. tingibefííassuseteregarise íexstumisendi



III. 6.61-62.: L-93 Châteaubleau (from RIG II-2, 239).

**Figure 8.** A Gaulish tile showing the use of s WITH HORIZONTAL STROKE for the Gaulish letter, in a number of words. The two s's seem to have been written first and then the horizontal stroke subsequently drawn through both, but the right way to encode this is as a single letter s, to be doubled as ss. Taken from Stifter 2012.

|           |     |           |       |   |            |  |
|-----------|-----|-----------|-------|---|------------|--|
| com       |     |           | conj. | also  | أيضاً.     |  |
| coña      |     |           | v.    | to starve                                   |            |  |
| copo      |     |           | v.    | to catch (bus)                              |            |  |
| copo      |     |           | v.    | to follow s.o. in order to catch them       |            |  |
| corċo     |     |           | v.    | to poke down a hole                         |            |  |
| cum       |     |           | conj. | also  |            |  |
| cwadaċo   | pl. | cwadaċr   | v.t.  | to anoint s.t., to put a salve on a wound   | إمسح على.  |  |
| cwado     | pl. | cwadr     | v.    | to anoint                                   | إمسح.      |  |
| cwareċo   |     |           | v.    | to be tired, to be homesick                 |            |  |
| cwaro     | pl. | cwarer    | v.    | to be tired                                 | إتعب.      |  |
| cwëri     | pl. | cwërir    | v.    | to make tired, to annoy, to pester          | أجهد-أتعب. |  |
|           |     |           |       |   |            |  |
|           |     |           |       |   |            |  |
| D.        |     |           |       |   |            |  |
|           |     |           |       |   |            |  |
| daċo      |     |           | v.    | to walk beside                              |            |  |
| dëria     |     | rëre      | n.    | side of body                                |            |  |
| dia       |     | iria      | n.    | cow   |            |  |
| dundo     |     |           | v.    | to cough                                    |            |  |
| dwaċo     |     |           | v.    | to push                                     |            |  |
|           |     |           |       |   |            |  |
|           |     |           |       |   |            |  |
| Đ.        |     |           |       |   |            |  |
|           |     |           |       |   |            |  |
| đaba      |     |           | n.    | clouds                                      | سحاب.      |  |
| đaba      |     |           | n.    | chaff                                       | تبن.       |  |
| đabacelo  |     |           | n.    | raising up, promoting                       | الإرتفاع.  |  |
| đabëcia   |     |           | n.    | lifting up                                  | الإصعاد.   |  |
| đabërtula |     | yabërtula | n.    | locust sp.                                  |            | green, solitary, appears in the rainy season |
| đablaċa   |     | yablaċa   | n.    | small pestle                                |            |  |
| đaboŋa    |     |           | n.    | curse                                       |            |  |
| đabortwa  |     |           | n.    | riding                                      | الركوب     |  |
| đaborwa   |     |           | n.    | flying                                      | الطيران.   |  |
| đaboċwa   |     |           | n.    | climbing                                    | الصعود.    |  |
| đabria    |     |           | n.    | loosing, releasing, setting free, salvation | الإطلاق.   |  |
| đabunëċia |     |           | n.    | blessing                                    |            |  |
| đad       | pl. | rad       | n.    | way   | طريق.      |  |
| đad       | pl. | rad       | n.    | road  | شارع.      |  |

Figure 9. A page from Blench 2005 showing TAU GALLICUM used for the Moro language. The text here places the Đ and đ in the Latin-1 positions for Š and š. It also uses Latin-1 ñ for ŋ and ç for ċ.



## Moro Noun Classes

Elizabeth Guest 8/5/97

This document outlines the Moro noun classes. The data was provided by Angelo, Israel, Ibrahim and a few others. The noun classes are identified by their singular and plural prefixes and by their concords and are summarised in the following chart ('∅' indicates zero prefix):

| Singular prefix | singular concord | plural prefix | plural concord | semantic identification                         |
|-----------------|------------------|---------------|----------------|---|
| ∅/w-            | g-               | l-            | l-             | people  |
| l-              | l-               | ñ-            | ñ-             | animals and body parts?                         |
| ɖ-              | ɖ-               | w-            | g-             | trees   |
| ∅               | g-               | n-            | n-             | common things                                   |
| l-/ɽ-/ɽr-       | l-/ɽr-/ɽr-       | ŋ-            | ŋ-             | long things, hollow & deep things, round things |
| ɖ-              | ɖ-               | r-            | r-             | long things?                                    |
| ɖ-              | ɖ-               | ∅             | g-             | ??  |
| ɖ-              | ɖ-               | y-            | y-             | large and harmful things                        |
| ŋ-              | ŋ-               | ñ-            | ñ-             | domestic and small animals                      |
| ŋ-              | ŋ-               | —             | —              | liquids and abstract nouns                      |
| ɖ-              | ɖ-               | —             | —              | abstract nouns (emotions)                       |
|                 | r-               |               | y-             | cow, goat, other irregular nouns                |
|                 | y-               |               | y-             | foreign words                                   |

**Figure 10.** A page from Guest 1997a showing TAU GALLICUM used for the Moro language in a list of noun-class prefixes.

I have asked the group for many contrasts, but I have not tried to be exhaustive. Instead I have concentrated on combinations that seem to be a problem between dialects.

|       |        |                                  |        |                           |
|-------|--------|----------------------------------|--------|---------------------------|
| b/p   | abəlo  | deny                             | apəlo  | take them                 |
| f/b   | fiðu   | find                             | biðu   | vomit                     |
| k/g   | kero   | bread                            | gero   | harvest                   |
| d/t   | ldu    | frog                             | ltu    | shelter                   |
| t/ɽ   | eta    | rock                             | eta    | lake                      |
| d/ɖ   | lwadia | tiredness                        | lwadia | cross                     |
| t/ɖ   | təro   | prevent                          | ɖəro   | look after baby           |
| ɽ/ɖ   | miɽu   | take and give to<br>someone else | miðu   | have full udder or breast |
| nd/nd | ndəŋ   | on hand                          | ndəŋ   | firm                      |
| c/j   | caca   | girl's name                      | jaja   | girl's name               |
| m/n   | mo     | take                             | no     | hear                      |
| ñ/ŋ   | ñere   | girls                            | ŋere   | girl                      |
| ɽ/r   | ɽo     | be                               | ro     | kick, to spear            |
| r/l   | raɽo   | crawl                            | laɽo   | cultivate                 |
| ɖ/s   | idia   | son                              | isia   | fire                      |
| o/u   | tu     | drink                            | to     | lose                      |
| e/ə   | erme   | in the rib                       | r-əme  | rib                       |
| ə/ë   | bəɽu   | beat                             | bəɽu   | miscarry                  |
| e/i   | gede   | it is green                      | gidi   | will                      |
| e/a   | pieło  | above                            | plalo  | under                     |

**Figure 11.** A page from Guest 1997b showing TAU GALLICUM used for the Moro language in a list of phonemic minimal pairs.

## A. Administrative

1. Title

**Proposal for the addition of four Latin characters in the UCS**

2. Requester's name

**Michael Everson, Chris Lilley**

3. Requester type (Member body/Liaison/Individual contribution)

**Individual contribution.**

4. Submission date

**2019-05-26**

5. Requester's reference (if applicable)

6. Choose one of the following:

6a. This is a complete proposal

**Yes.**

6b. More information will be provided later

**No.**

## B. Technical – General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters)

**No.**

1b. Proposed name of script

1c. The proposal is for addition of character(s) to an existing block

**Yes**

1d. Name of the existing block

**Latin Extended-D**

2. Number of characters in proposal

**4.**

3. Proposed category (A-Contemporary; B.1-Specialized (small collection); B.2-Specialized (large collection); C-Major extinct; D-Attested extinct; E-Minor extinct; F-Archaic Hieroglyphic or Ideographic; G-Obscure or questionable usage symbols)

**Category A.**

4a. Is a repertoire including character names provided?

**Yes.**

4b. If YES, are the names in accordance with the “character naming guidelines” in Annex L of P&P document?

**Yes.**

4c. Are the character shapes attached in a legible form suitable for review?

**Yes.**

5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

**Michael Everson.**

5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

**Michael Everson, Fontographer.**

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

**Yes.**

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

**Yes.**

7. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

**Yes.**

8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database <http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

**See above.**

## C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

**Yes. N4297R (L2/12-270R).**

2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

**Yes.**

2b. If YES, with whom?

**Michael Everson, Chris Lilley (Gaulish); Roger Blench, Peter Jenks, Sharon Rose, Angelo Naser (Moro).**

2c. If YES, available relevant documents

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

**Linguists, Celticists, Africanists. Also Moro speakers in one practical orthography.**

4a. The context of use for the proposed characters (type of use; common or rare)

**Used historically and in modern editions.**

4b. Reference

5a. Are the proposed characters in current use by the user community?

**Yes.**

5b. If YES, where?

**Various publications.**

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

**Yes.**

6b. If YES, is a rationale provided?

**Yes.**

6c. If YES, reference

**Accordance with the Roadmap. Keep with other Latin phonetic characters.**

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

**No.**

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

**No.**

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

**No.**

9b. If YES, is a rationale for its inclusion provided?

**No.**

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

**No.**

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?

**No.**

11b. If YES, is a rationale for such use provided?

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

**No.**

11e. If YES, reference

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

**No.**

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

**No.**

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?