

# Unicode request for annotations to two mathematical characters

Neil Rees, neil.rees, biblestudies.org.uk

Kirk Miller, kirkmiller, gmail

2023 November 8

We request annotations to U+1D438 MATHEMATICAL ITALIC CAPITAL E and U+1D452 MATHEMATICAL ITALIC SMALL E so that typographers will be aware that they are used in orthography and should be supplied with anchors for diacritics and given small capital support.

Italic  $\langle E e \rangle$  is used as a distinct letter from roman  $\langle E e \rangle$  in the orthographies of two Arbëresh gospels that one of us (Rees) is attempting to digitize for modern access. Italic  $\langle E e \rangle$  stands for /ə/ and roman  $\langle E e \rangle$  for /e/. The gospels in question are translations into the Albanian dialects of Sicily and Calabria.

Modern Standard Albanian orthography uses  $\langle \ddot{E} \ddot{e} \rangle$  for /ə/. However, for these Arbëresh texts substitution with  $\langle \ddot{E} \ddot{e} \rangle$  would cause problems when the letter is combined with other diacritics, for example  $\langle \acute{e} \rangle$  in Figure 5 or  $\langle \hat{e} \rangle$  in Figure 7. In the modern standard  $\langle \ddot{E} \ddot{e} \rangle$  does not take additional diacritics, and substituting e.g.  $\langle \acute{E} \acute{e}, \hat{E} \hat{e} \rangle$  with  $\langle \acute{\ddot{E}} \acute{\ddot{e}}, \hat{\ddot{E}} \hat{\ddot{e}} \rangle$  would make them difficult to read.

Hard-coded italic  $E, e$  are available in Unicode as mathematical symbols, U+1D438 and U+1D452. There are two problems with using these characters for Arbëresh, however: (1) they do not form a casing pair, and (2) in all fonts tested, they do not accept diacritical marks, with are necessary for Arbëresh orthography.

If the Unicode characters were made into a casing pair, their properties would need to be amended as follows (additions in red):

```
1D452;MATHEMATICAL ITALIC SMALL E;Ll;0;L;<font> 0065;;;N;;;1D438;;1D438
1D438;MATHEMATICAL ITALIC CAPITAL E;Lu;0;L;<font> 0045;;;N;;;1D452;
```

This is problematic because as mathematical characters they should *not* be subject to changes in case. The Arbëresh orthography described here is not in active use, however, so a lack of casing behaviour for  $E, e$  should not cause undue difficulty for typesetting, though it will cause problems with text-searching.

Diacritic support, on the other hand, is required. Annotations similar to the following are requested so that typographers will know to provide the characters with anchor points for diacritics and to provide small-capital forms (if small capitals are provided by the font):

## Annotations

1D438;MATHEMATICAL ITALIC CAPITAL E;Lu;0;L;<font> 0045;;;N;;;;;

also used in Arbëresh orthography

1D452;MATHEMATICAL ITALIC SMALL E;Ll;0;L;<font> 0065;;;N;;;;;

also used in Arbëresh orthography

## References

Demetrio Camarda, ed. (1868). *Il Vangelo di S. Matteo, tradotto dal testo greco nel dialetto calabro-albanese di Frascineto*. London.

Vicenzo Dorsa; ed. Demetrio Camarda (1869). *Il Vangelo di S. Matteo, tradotto dal testo greco nel dialetto albanese di Piana de' Greci in Sicilia da un nativo di questo luogo*. London.

## Figures

*I segni grafici impiegati in questa versione sono i seguenti: Aa, Bb, Dd, Dz dz, Dž dž, Δ δ, Ee, Ee, Ff, Gg, Gj gj, Γ γ, Γj γj, Ii, Jj, Kk, Kj kj, Ll, Lj lj, Mm, Nn, Nj nj, Oo, Pp, Rr rr, Ss, Sš, Tt, Ts ts, Tš tš, Θ ϑ, Uu, Vv, Xx, Xj xj, Zz, Zež.*

Figure 1. Camarda (1868: i). Alphabet. A couple of the letters beside <E e> are not Unicode casing pairs, namely large <s> and <z> as capitals for <š> and <ž> (blue). Dorsa (1869) replaces those with <Š š> and <Ž ž>.

*L IBER i jenise te Džeslite Krištite birate Dávidite, birate Abrámite.*

Figure 2. *ibid.* (1868: 3). Italic <E> and <e> (red) contrasting with roman <E> and <e> (blue).

11 Aχιέρρα e **lé** diâyi: e ši **éngjejet'** i u-kjasen, e e **šerbéijen.**

12 Si gjegji prân Džesúi se Džuáni kle zenne, u-prúare nde **Galilê:**

37 Te **jête** te folet tâje, **éχj, éχj:** jo, jo: até tš' ište. mé šume se ktô fjale, ište per te like.

38 Gjékjete se u-ða: Sî per sî, ðempe per ðempe.

39 E ú ju ðome, te mos **zéχiχje** me te ligun: e ne ndonjerî te ðefte ndô šplakê mbi fakjen tende te **dréjeten,** priri edé te jetren.

Figure 3. *ibid.* (1868: 11, 17). Italic <é> and <ê> (red) contrasting with roman <é> and <ê> (blue). <é> is a short tonic vowel and <ê> a long tonic vowel.

## **ŠÉITI VANGJELJE**

2 **Émeret** e dimbediét Apóstuljevet jân kta. I pari, Simuni tše θexet Piéter, e Ndreu vlau tîje,

Figure 4. Dorsa (1869: 1, 30). Capital italic <É> (red) and <E> contrasting with roman <E> and <é>.

22 Gjiθ prana kî šurbés **u-bé,** se t' iš e **bēn e θēna** ka Inzót per profeten tše θot:

23 Šiχ virgjera kat dâlje me bârr, e kat bēnje nje bîr: e kân ti θerresen **émerin** e tîje **Manuêl, tše vién** me θēn: Inzót me nê.

24 Sgjúar prana Sepa ka gjumi, beri si i θa **éngjeli** Tinzoti, e mbióθ grúan e tîje.

Figure 5. *ibid.* (1869: 3). Italic <é>, <ê>, <ē> and <é̇> contrasting with roman <é> and <ê>. The two nasal vowels in this dialect are <e> and <i>.

8 Arđur prana mbrémanet θot i zoti vréštése atíje tše i **běneje** piéste: Θerrite jurnatiéret, e jipi rrogen, zēnur ka te lúrtemit njera tek te paret.

9 Mos mbâši ár, jo **regjént**, jo turrés nder brezet tâje:

Figure 6. *ibid.* (1869: 30, 69). Other instances of italic <ě>.

20 Aχíera zû e nemereje χôrvet, tek **u-běň** mé te šumet merákulje te tîje, se nēng u-metanósetin.

21 Mierezî tîj, Xorazín, mierezî tîj, Betsáide: se, nde Tîr, e Sidôn u-kišen **běň** merákuljit, tše **u-běň** tek ju, kâ mot tše kišen bēň peniténdzie nde θes e nde χí.

22 Frēnge kte ju θom: Tirit, e Sidonit kan te **jěň mé** te ljea penet te dita džudidzit, se juve.

8 E gjiθ kto jân te zēnit e te **đěmburavet**.

9 Aχíerna ju **věň** nde pēn, e ju vrasen: e jini te dašur ljik ka gjiθ gjindet per **émerin** tim.

Figure 7. *ibid.* (1869: 36, 86). Italic <ě, é> alongside <ê, é, ē> contrasting with <ê>.

13 Aχíerna i θot Pilati: **Ēng** gjegjen sâ θôn mbi tîj martrît?

Figure 8. *ibid.* (1869: 103). Capital italic <Ē>.

28 E kure vate te špia, i u-kjasen te vérberite. E i Đote Džesúii: Kini besse se ú munde bunje keté? E atá i Đane: **Ēχj**, Zot.

Figure 9. Camarda (1868: 31). Capital italic <Ē>.