Title
Proposal to add Medievalist punctuation characters to the UCS

Permalink
https://escholarship.org/uc/item/6p88k38r

Authors
Everson, Michael
Baker, Peter
Grammel, Florian
[et al]

Publication Date
2016-01-25

Peer reviewed
1. Introduction. A set of characters used by specialists in medieval European philology, palaeography, and linguistics has long been absent from the Universal Character Set. This proposal requests 21 punctuation characters be added for support of medieval European linguistic and literary research and publication. If this proposal is accepted, the following characters will be encoded:

- \texttt{2E45} PARAGRAPHUS MARK
- \texttt{2E46} POSITURA MARK
- \texttt{2E47} COLON WITH SIDEWAYS REVERSED RAISED COMMA
- \texttt{2E48} COLON WITH RAISED POSITURA MARK
- \texttt{2E49} TWO DOTS OVER COMMA
- \texttt{2E4A} PUNCTUS ELEVATUS MARK
- \texttt{2E4B} SIDEWAYS REVERSED MIDDLE COMMA
- \texttt{2E4C} PUNCTUS FLEXUS MARK
- \texttt{2E4D} PUNCTUS VERSUS MARK
- \texttt{2E4E} LOW PUNCTUS VERSUS MARK
- \texttt{2E4F} PUNCTUS INTERROGATIVUS MARK
2. Functions of Medieval punctuation. Modern European punctuation comprises a set of named marks which are used with relatively well-established usages. Medieval punctuation was based on discursive functions; in some areas and at some times different configurations of dots were used to express those functions. Unification of those configurations according to the functions would not really satisfy medievalists, who need to make use of the configurations in use in the documents they study in terms of the time and place of those documents. Nevertheless, the right way to establish what configurations are needed, and what configurations can be built up out of sequences of existing or new characters, is to look first at the functions the punctuation marks serve.

One of the best studies of the history of punctuation is M. B. Parkes’ 1993 Pause and effect: an introduction to the history of punctuation in the West. He describes the development from the very neutral scriptio continua, in which neither punctuation nor word spacing was used, towards what we would consider more legible and meaningful text. The development of vernacular writing in Ireland was an important stage in this process: Latin letters having different phonemic values in Irish, for instance, were marked to distinguish them from their Latin uses. Syntactic punctuation innovation followed.
2.1. **Paragraphus.** The beginning of a paragraph, a section, a stanza, or a proposition was marked with a symbol such as γ, Γ, =Value, or §. Later this function was replaced by the **paraph**, where //, ¶, or ⸿ were typical marks. One character is proposed for encoding here, ⹅ PARAGRAPHUS MARK. See Figures 1, 2, 21.

2.2. **Positura.** The end of a section of text was marked by a **positura**—the opposite of the paragraphus. A wide variety of marks came to fulfill this function: , , .:, ;, and . Some of these can be sequenced with existing characters and some cannot. Four characters are proposed for encoding here: , POSITURA MARK, . COLON WITH SIDEWAYS REVERSED RAISED COMMA, . COLON WITH RAISED POSITURA MARK, and . TWO DOTS OVER COMMA. The others can be composed:

\[ .\gamma = . + \gamma \]
\[ .\varepsilon = . + \varepsilon \]
\[ ;\varepsilon = . + ; + . \]

Note that the angular POSITURA MARK is not identical to the modern comma. The origin of the modern comma was the **MEDIEVAL COMMA**, for which see 2.3 below. A number of other indicators of **positura** came to acquire specific meanings and are discussed in 2.2.1–2.2.5 below. See Figures 3, 4, 5 (\γ); 5, 6, 8 (\μ); 5, 7 (\ν); 5, 9 (\κ). Note that COLON WITH SIDEWAYS REVERSED RAISED COMMA looks as though it could be composed; there are handwritten glyph variants which look like :- (really :–) and :: (see figures 6, 8). But scholars typically represent this with the - SIDEWAYS REVERSED RAISED COMMA, which does not appear to be used on its own. If :: must be composed, - SIDEWAYS REVERSED RAISED COMMA must be added to the UCS.

2.2.1. **Punctus elevatus.** This was in origin an indicator of **positura** (ending a section) but which came to be used to indicate a major medial pause “where the sense is complete but the meaning is not” (Parkes p. 306). It is the ancestor of our modern colon. The typical shape for this is formed of a dot with a sideways reversed middle comma above and slightly to the right of it, though a form \γ with a diagonal line rather than a comma can be found as a glyph variant. In the Wycliffe Bible translation, the two-part character is contrasted with a similar sign lacking the lower dot, indicating a lesser pause. Two characters, ⹊ PUNCTUS ELEVATUS MARK, and ⹊ SIDEWAYS REVERSED MIDDLE COMMA, have been proposed here for encoding. See Figures 7, 10, 11, 12, 13, 14, 15, 16 (\γ); 12, 13, (\κ).

2.2.2. **Punctus flexus.** This was in origin an indicator of **positura** but which came to be used to indicate a minor medial pause where the sense is incomplete. It is used after a comma (not our punctuation mark “,”; see 2.3 below) or at the flexa (“the inflexion of the voice at a minor medial pause which deviates to a pitch below that adopted as the recitation tone”—Parkes p. 304). One character, ⹊ PUNCTUS FLEXUS MARK, has been proposed here for encoding. See Figures 7, 14, 15.

2.2.3. **Punctus versus.** In origin this indicated **positura** but came to indicate the terminatio of a psalm verse, or the completion of a sententia or periodus. Two characters are commonly seen to represent this function, one with the angular POSITURA MARK with a dot over it, and the other with a sideways comma below a baseline dot. Two characters, ⹞ PUNCTUS VERSUS MARK and ⹞ LOW PUNCTUS VERSUS MARK have been proposed here for encoding. See Figures 7, 16 (\gamma); 12, 13, (\gamma).

2.2.4. **Punctus interrogativus.** This originally indicated **positura** but came to indicate a question requiring an answer. This is the origin of the modern ? QUESTION MARK and ¿ INVERTED QUESTION MARK. The glyph for this character is sometimes angular with one to three hooks \µ and is typically slanted toward the right; the vertical form we know today dates to the late 15th century. One character, \µ PUNCTUS INTERROGATIVUS MARK, has been proposed here for encoding. A related function, the **punctus percontativus**, indicated a rhetorical question and is represented by U+2E2E ‽ REVERSED QUESTION MARK. See Figures 7, 13, 15, 16, 18, 21.

2.2.5. **Punctus exclamativus.** Yet another original indicator of **positura** is also known as punctus admirativus. This is the origin of the modern exclamation mark and reversed exclamation mark. A glyph variant ‿ stands upright. One character, ‿ PUNCTUS EXCLAMATIVUS MARK has been proposed here for encoding. See Figures 17, 18.
2.3. **Comma.** Functionally, the *comma* is "a division of a *colon*... usually short and rhythmically incomplete, followed by a minor disjunction of the sense where it may be necessary to pause" (Parkes p. 302). The punctuation sign which came to bear the same name was employed to show a disjunction of sense, or a minor medial pause at the end of a *comma*. This was the *MEDIEVAL COMMA* proposed here. It was most commonly used by Italian scribes in the 14th century, and is the ancestor of our modern comma, whose modern form appeared in the first typefaces. See Figures 7, 19.

2.4. **Distinctiones.** This system of punctuation places marks at different heights in an ascending order of importance. A low point indicating a minor medial pause is called *subdistinctio*, a mid-height mark indicating a major medial pause is called *media distinctio*, and a high mark, called *distinctio*, indicates a final pause where the *sententia* or *period* (the thought or opinion) is finished. Distinctiones were not only represented by the height of the marks:

```
, , ,, , , , , : : : : : : ,
```

Note that ,, ,,, and ,, should be encoded as sequences. Care should be taken in implementations not to linebreak these. Note too that : is the Georgian paragraph separator; no "generic" punctuation mark for that has been encoded (perhaps one should be). Distinctiones which are encoded according to characters at different heights are:

```
. . . . . .
```

The first set here is encoded as FULL STOP, RAISED DOT, and MIDDLE DOT (drawn a little low in this font; it should be the height of the high dot in the colon :). The second set is encoded as FULL STOP, MIDDLE DOT, and HIGH DOT. This last character is proposed here. See Figure 21.

2.5. **Simplex ductus.** This was originally a critical sign used to separate matters erroneously run together. One character, ? SIMPLEX DUCTUS MARK has been proposed here for encoding. See Figure 22.

2.6. **Virgula suspensiva.** The *virgula suspensiva* (what we know as the / SOLIDUS) was used to mark the briefest pause or hesitation in a text. When doubled // it was an indication of paraph See 2.1 above). Humanist writers of the 14th century made a distinction whereby / indicated a break greater than that indicated by / but less than that indicated by : PUNCTUS ELEVATUS MARK. One character, / DOTTED SOLIDUS has been proposed here for encoding. See Figure 23.

2.7. **SIGNE DE RENVOI** is used to associate matter in the text with material added in the margin, and especially a passage omitted from the text by the original copyist. It was later used as a quire mark by printers. See Figure 24.

2.8. **MIDDLE COMMA** as a punctuation mark is used to indicate a variety of abbreviation. With long s, for instance, ſ is an abbreviation for *sed* 'but'. It is also used sequenced alongside MIDDLE DOT · as an indicator of *positura*. See Figure 20, 25.

2.9. **TILDE WITH DOT ABOVE AND DOT BELOW** is another *nota* or abbreviation mark, typically indicating Latin *est* 'is'. There is a math operator U+223B HOMOTHETIC, but this is not suitable for use as a *nota*. Compare U+2A6A TILDE OPERATOR WITH DOT ABOVE and the punctuation character U+2E1E TILDE WITH DOT ABOVE. There are several glyph variants for this: ≃ ≈ are common. Historically a cursive form of this is the ancestor of U+A76B LATIN SMALL LETTER ET. See Figures 6, 8, 20.

2.10. **VERTICAL FIVE DOTS** completes the set of vertical dots in the UCS. We have U+003A : COLON, U+205D : TRICOLON, U+205E : VERTICAL FOUR DOTS, and U+2E3D VERTICAL SIX DOTS. Discussion in Glasgow in July 2015 with specialists in Old Italic and North Italic indicated that five vertical dots were certainly a configuration known to be attested. They have also been attested in Runic. See Figure 26.
2.11. **TRIPLE DAGGER** ⩙ is similar in use to U+2020 † DAGGER and U+2021 ‡ DOUBLE DAGGER. It indicates another level of notation. See Figures 27, 28.

3. **Typography of medieval punctuation.** The core definition of “generic” punctuation is based on the typical shapes and sizes of punctuation as used in the Latin, Greek, and Cyrillic traditions; Georgian may be added to this list. When such punctuation is used in other scripts, such as the scripts of India or Southeast Asia, they are often modified to suit the ductus and letterforms of those scripts, but nevertheless the core definitions are based in the habits of European typography. An examination of the height of the modern : COLON U+003A, along with the size of its dots, is perhaps the best way to compare the existing set of punctuation characters and to determine how specialist needs can be met by filling out gaps or by defining characters within the context of the set of punctuation characters already encoded in the UCS.

3.1. **Low and raised punctuation.** The characters U+002C , COMMA and U+002E . FULL STOP form the basis for the system. To this are added U+2E32 , TURNED COMMA, U+2E33 . RAISED DOT, U+2E34 . RAISED COMMA, U+00B7 MIDDLE DOT, and the new characters proposed here, U+2E46 , POSITURA MARK, and U+2E56 . MIDDLE COMMA, and 2E4E ; LOW PUNCTUS VERSUS MARK:

xpb , , :, · ⩆ ⩖ ⩎

3.2. **Middle or x-height punctuation.** The characters : U+003A COLON and ; U+003B SEMICOLON define the basic height for most inline generic punctuation dots. Conforming to this configuration are U+204F ; REVERSED SEMICOLON, U+2E35 ⸵ TURNED SEMICOLON, U+10FB GEORGIAN PARAGRAPH SEPARATOR, U+2056 : THREE DOT PUNCTUATION, U+2058 ⁖ FOUR DOT PUNCTUATION, U+2059 :: FIVE DOT PUNCTUATION, U+2E2A ⸪ TWO DOTS OVER ONE DOT PUNCTUATION, U+2E2B ⸫ ONE DOT OVER DOTS PUNCTUATION, U+2E2C ⸬ SQUARED FOUR DOT PUNCTUATION, as well as the characters proposed here, U+ 2E47 ⸯ COLON WITH SIDEWAYS REVERSED RAISED COMMA, U+2E48 ⸰ COLON WITH RAISED POSITURA MARK, U+2E49 ⸱ TWO DOTS OVER COMMA, U+2E4A ⸲ PUNCTUS ELEVATUS MARK, U+2E4B ⸳ SIDEWAYS REVERSED MIDDLE COMMA, U+2E4C ⸴ PUNCTUS FLEXUS MARK, U+2E4D ; PUNCTUS VERSUS MARK, U+2E51 ⸵ MEDIEVAL COMMA, and 2E57 TILDE WITH DOT ABOVE AND DOT BELOW:

xpb :: ; · ; · ; · ; · ; · · · · · ; · · · · · · · ; · · · · · ·


Tertullian, Quintus Septimus Florens. [1493]. *Apologeticus adversus gentes*. Venetiis: B. Benalius.


7. Acknowledgements
This project was made possible in part by a grant from the U.S. National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at UC Berkeley). Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment for the Humanities.

Examples

Figure 1. Sample from Farley 1783, showing PARAGRAPHUS MARK.

Figure 2. Sample from Virgile 1509, showing PARAGRAPHUS MARK.

Figure 3. Sample from Parkes 1993:202, showing POSITURA MARK sequenced after FULL STOP.
How much those arguments lend support to us that is to universal peace which are cited as though on the authority of the blessed Cyprian on the side of Donatus against us and how much they are against those by whom they are cited: I propose to demonstrate with the help of the Lord, if herefore the necessity of responding compels me to know again those arguments which I have already set down in other books - although I shall do so as little as possible for those who have already read them and hold to them it ought not to be burdensome since not only must those things which are necessary for instruction be instilled often into those who are rather slow - but when those same things are turned over and treated in many and various ways - they also help those endowed with more capacious understanding both to learn with facility and to discourse with fluency, for I know how much it discourages a reader...
Figure 5. Description of various *positurae* from Parkes 1993, showing POSITURA MARK, COLON WITH SIDEWAYS REVERSED RAISED COMMA, COLON WITH RAISED POSITURA MARK, and TWO DOTS OVER COMMA.

Figure 6. Sample from O’Neill 1984:64, showing TILDE WITH DOT ABOVE AND DOT BELOW and COLON WITH SIDEWAYS REVERSED RAISED COMMA. The text reads:

Ueni ꞇ e benedic ꞇ i pa ꞇ ꞇ ꞇ ꞇ ꞇ ꞇ me ꞇ po ꞇ ꞇ ꞇ e
ab origin mundi · alleluia · venite sicut erat · venite: móel cáich scripsit

‘Come ye blessed of my Father take possession of the Kingdom alleluia prepared for you from the foundation of the world alleluia glory be · come As it was · Come: · Móel Cáich wrote this.’

Figure 7. Sample from Núñez Contreras 1994 showing PUNCTUS VERSUS MARK, PUNCTUS ELEVATUS MARK, MEDIEVAL COMMA, PUNCTUS FLEXUS MARK, PUNCTUS INTERROGATIVUS MARK, and COLON WITH RAISED POSITURA MARK.
Figure 8. Sample from Parkes 1993: showing colon with sideways reversed raised comma and tilde with dot above and dot below. In the manuscript the glyph variant :: is shown for ::.

Figure 9. Sample from Loew 1914:253, showing two dots over comma. Loew describes it as “the Beneventan period” (p. 269).

Figure 10. Sample from Farley 1783, showing punctus elevatus mark.
Figure 11. Sample from Wright 1960:12, showing PUNCTUS ELEVATUS MARK.

Figure 12. Sample of Wycliffe Bible text from Thompson 1912 showing PUNCTUS ELEVATUS MARK and SIDEWAYS REVERSED MIDDLE COMMA.
Figure 13. Sample of Augustinian text from Thompson 1912 showing PUNCTUS ELEVATUS MARK and SIDEWAYS REVERSED MIDDLE COMMA. Also in the manuscript is the PUNCTUS INTERROGATIVUS MARK though in the transcription the QUESTION MARK has been used.

Contritio et inflicitas in suis corum: et uiam pacis non cognouerunt: non est timor det ante oculos corum:

Figure 14. Sample from Parkes 1993: showing PUNCTUS FLEXUS MARK, PUNCTUS ELEVATUS MARK, and LOW PUNCTUS VERSUS MARK.

Figure 15. Sample from Parkes 1993: showing PUNCTUS FLEXUS MARK, PUNCTUS INTERROGATIVUS MARK, and PUNCTUS ELEVATUS MARK.
Figure 16. Sample from Parkes 1993:293, showing PUNCTUS INTERROGATIVUS MARK, PUNCTUS ELEVATUS MARK, and PUNCTUS VERSUS MARK.

Figure 17. Sample from Ouy 1987, showing PUNCTUS EXCLAMATIVUS MARK with its glyph variant ! as opposed to the usual /.
Figure 18. Sample from Parkes 1993, showing PUNCTUS EXCLAMATIVUS MARK and PUNCTUS INTERROGATIVUS MARK. Although the scan is not very clear, there are definitely two dots on the !.

Figure 19. Sample from Guðvarður Már Gunnlaugsson 2001, showing MEDIEVAL COMMA.

Figure 20. Sample from Loew 1914, showing LONG S followed by MIDDLE COMMA with the reading sed ‘but’, and showing TILDE WITH DOT ABOVE AND DOT BELOW with the reading est.

Apart from the symbol for enim, the other characters here have been encoded or can be composed. For enim, it is a bit of a question what is to be represented. Forms with middle dot ·n· occur, and in some hands an N or n can look like what is shown in the sample above. See Cappelli pp. 229–230.

A character for enim is not requested in this proposal.
Figure 21. Sample from Parkes 1993 showing HIGH DOT, PUNCTUS INTERROGATIVUS MARK and PARAGRAPHUS MARK.

Figure 22. Sample from Parkes 1993:164, showing many examples of SIMPLEX DUCTUS MARK.
By contrast Petrarch used the pointed *virgula* where an interpolated statement has a different relationship to other statements in the immediate context:

```
Numquam tam iuuenis / numquamque tam glorioe cupidus / in / quod interdum me fuisse non inficior, quin maluerint bonus esse quam doctus. 
```

Never so young / and never so greedy for fame / have I been / I do not deny having been that occasionally. but that I have chosen to be good rather than learned.

**Figure 23.** Sample from Parkes 1993, showing DOTTED SOLIDUS.

---

Any sign used to associate matter in the text with material added in the margin, and especially a passage omitted from the text by the original copyist. ‘.’ was a sign used by printers.

**Figure 24.** Sample from Parkes 1993, showing SIGNE DE RENVOI.

---

The x-height here is low, but neither the dot nor comma rest on the baseline. (Compare the dots in the title THEOLOGICAL TRACTS.—A. D. 821.

**Figure 25.** Sample from Thompson 1912:408, showing MIDDLE COMMA alongside MIDDLE DOT.

---

The intended rune-forms of the inscription are interpreted as follows:

```
5 10 15 20 25 30
PORST-NEINARSSUNREISTRUNARPSAR
```

**Figure 26.** Sample from Owen and McKinnell showing VERTICAL FOUR DOTS and VERTICAL FIVE DOTS.
Figure 27. Sample showing ⹙ TRIPLE DAGGER (circled) alongside † DAGGER and ‡ DOUBLE DAGGER. Also shown in the last line is the ⼏ PUNCTUS INTERROGATIVUS MARK.

Figure 28. Sample showing ⹙ TRIPLE DAGGER alongside † DAGGER and ‡ DOUBLE DAGGER. Also shown are other signs of notation, * ASTERISK and § SECTION SIGN.
<table>
<thead>
<tr>
<th>Unicode Code Points</th>
<th>Characters</th>
</tr>
</thead>
<tbody>
<tr>
<td>2E00 2E01 2E02 2E03 2E04 2E05 2E06 2E07 2E08 2E09 2E0A 2E0B 2E0C 2E0D 2E0E 2E0F</td>
<td>⸀ ⸁ ⸂ ⸃ ⸄ ⸅ ⸆ ⸇ ⸈ ⸉ ⸊ ⸋ ⸌ ⸍ ⸎ ⸏</td>
</tr>
<tr>
<td>2E10 2E11 2E12 2E13 2E14 2E15 2E16 2E17 2E18 2E19 2E1A 2E1B 2E1C 2E1D 2E1E 2E1F</td>
<td>⸐ ⸑ ⸒ ⸓ ⸔ ⸕ ⸖ ⸗ ⸘ ⸙ ⸚ ⸛ ⸜ ⸝ ⸞ ⸟</td>
</tr>
<tr>
<td>2E20 2E21 2E22 2E23 2E24 2E25 2E26 2E27 2E28 2E29 2E2A 2E2B 2E2C 2E2D 2E2E 2E2F</td>
<td>⹀ ⹁ ⹂ ⹃ ⹄ ⹅ ⹆ ⹇ ⹈ ⹉ ⹊ ⹋ ⹌ ⹍ ⹎ ⹏</td>
</tr>
<tr>
<td>2E30 2E31 2E32 2E33 2E34 2E35 2E36 2E37 2E38 2E39 2E3A 2E3B 2E3C 2E3D 2E3E 2E3F</td>
<td>⹐ ⹑ ⹒ ⹓ ⹔ ⹕ ⹖ ⹗ ⹘ ⹙ ⹚ ⹛ ⹜ ⹝ ⹞ ⹟</td>
</tr>
<tr>
<td>2E40 2E41 2E42 2E43 2E44 2E45 2E46 2E47 2E48 2E49 2E4A 2E4B 2E4C 2E4D 2E4E 2E4F</td>
<td>⹐ ⹑ ⹒ ⹓ ⹔ ⹕ ⹖ ⹗ ⹘ ⹙ ⹚ ⹛ ⹜ ⹝ ⹞ ⹟</td>
</tr>
</tbody>
</table>
New Testament editorial symbols

2E00 ꞌ RIGHT ANGLE SUBSTITUTION MARKER
02C1 ꞌ top left corner
2E01 Ɥ RIGHT ANGLE DOTTED SUBSTITUTION MARKER
2E02 Ꞓ LEFT SUBSTITUTE BRACKET
2E03 ꞓ RIGHT SUBSTITUTE BRACKET
2E04 ꞔ LEFT DOTTED SUBSTITUTE BRACKET
2E05 ꞕ RIGHT DOTTED SUBSTITUTE BRACKET
2E06 Ꞟ RAISED INTERPOLATION MARKER
02D4 Ꞟ down tack
2E07 ꞟ RAISED DOTTED INTERPOLATION MARKER
2E08 Ꞡ DOTTED TRANSPOSITION MARKER
2E09 ꞡ LEFT TRANSPOSITION BRACKET
2E0A Ꞣ RIGHT TRANSPOSITION BRACKET
2E0B ꞣ RAISED SQUARE
02D1 ꞣ used as an opening raised omission bracket
2E0C Ꞥ LEFT RAISED OMISSION BRACKET
02D0 Ꞥ used as an opening or closing raised omission bracket
2E0D ꞥ RIGHT RAISED OMISSION BRACKET
02D2 ꞥ used as a closing or opening raised omission bracket

Ancient Greek textual symbols

2E0E Ꞑ EDITORIAL CORONIS
01BD Ꞑ greek koronis
2E0F ꞑ PARAGRAPHOS
2E10 Ꞓ FORKED PARAGRAPHOS
2E11 ꞓ REVERSED FORKED PARAGRAPHOS
2E12 ꞔ HYPODIASTOLE
2E13 ꞕ DOTTED OBELOS
02D7 ꞕ glyph variants may look like ‘-’ or ‘-’
02D3 ꞕ commercial minus sign
2E14 Ꞗ DOWNWARDS ANCORA
02E4 Ꞗ contrary to its formal name this symbol points upwards
2E15 ꞗ UPWARDS ANCORA
02E5 ꞗ contrary to its formal name this symbol points downwards
2E16 Ꞙ DOTTED RIGHT-POINTING ANGLE
02D9 Ꞙ diplo peristigmene

Ancient Near-Eastern linguistic symbol

2E17 ꞙ DOUBLE OBLIQUE HYPHEN
02D6 ꞙ used in ancient Near-Eastern linguistics
02D5 ꞙ hyphen in Fraktur text uses 02D0 - or 2010 -, but with a ‘-’ glyph in Fraktur fonts
02D2 ꞙ hyphen-minus
02D3 ꞙ equals sign
02D0 ꞙ hyphen
02D2 ꞙ double hyphen

General punctuation

2E18 Ꞧ INVERTED INTERROBANG
02E7 Ꞧ gnrbrtnri
02D0 Ꞧ interrobang
2E19 Ꞟ PALM BRANCH
02D2 Ꞟ used as a separator

Dictionary punctuation

These punctuation marks are used mostly in German dictionaries, to indicate umlaut or case changes with abbreviated stems.

2E1A ꞕ HYPHEN WITH DIAERESIS
00F5 ꞕ indicates umlaut of the stem vowel of a plural form
2E1B Ꞗ TILDE WITH RING ABOVE
0268 Ꞗ indicates change in case for derived form

Brackets

2E1C Ꞑ LEFT LOW PARAPHRASE BRACKET
2E1D ꞑ RIGHT LOW PARAPHRASE BRACKET
02B6 ꞑ used in N’Ko

Dictionary punctuation

2E1E Ꞓ TILDE WITH DOT ABOVE
00F7 Ꞓ indicates derived form changes to uppercase
2E1F ꞓ TILDE WITH DOT BELOW
00F8 ꞓ indicates derived form changes to lowercase

Brackets

2E20 ꞛ LEFT VERTICAL BAR WITH QUILL
2E21 Ꞝ RIGHT VERTICAL BAR WITH QUILL

Half brackets

These form a set of four corner brackets and are used editorially. They are distinguished from mathematical floor and ceiling characters. Occasionally quine corners are substituted for half brackets.

2E22 ꞡ TOP LEFT HALF BRACKET
02C8 ꞡ top left corner
02C1 ꞡ top left corner
02C0 Ꞓ left bracket corner
2E23 Ꞣ TOP RIGHT HALF BRACKET
2E24 ꞣ BOTTOM LEFT HALF BRACKET
2E25 Ꞥ BOTTOM RIGHT HALF BRACKET

Brackets

2E26 ꞥ LEFT SIDEWAYS U BRACKET
02C6 ꞥ subset of
2E27 Ꞧ RIGHT SIDEWAYS U BRACKET
02C7 Ꞧ superset of
2E28 ꞧ LEFT DOUBLE PARENTHESIS
0296 ꞧ left white parenthesis
0295 ꞧ FF5F fullwidth left white parenthesis
2E29 Ꞩ RIGHT DOUBLE PARENTHESIS

Historic punctuation

2E2A ꞩ TWO DOTS OVER ONE DOT PUNCTUATION
2E2B Ɦ ONE DOT OVER TWO DOTS PUNCTUATION
2E2C Ɜ SQUARED FOUR DOT PUNCTUATION
2E2D Ɡ FIVE DOT MARK
2E2E Ɬ REVERSED QUESTION MARK
003F Ɬ question mark
000F Ɬ inverted question mark
061F Ɬ arabic question mark
2E2F Ɪ VERTICAL TILDE
0261 Ɪ used for Cyrillic yerik
033E Ɪ combining vertical tilde
0A67 Ɪ combining vertical tilde

Printed using UniBook™ (http://www.unicode.org/unibook/)

Date: 2015-12-19
Supplemental Punctuation

Double hyphen

The double hyphen is used in transcription of old German manuscripts, and occasionally as a non-standard punctuation mark. It is not intended for the representation of normal hyphens, whose doubled forms in Fraktur text are considered glicphic variants.

2E40 " DOUBLE HYPHEN
  → 003D = equals sign
  → 2010 = hyphen
  → 2E17 « double oblique hyphen
  → 30A0 = katakana-hiragana double hyphen
  → A78A = modifier letter short equals sign

Reversed punctuation

2E41 ‥ REVERSED COMMA
  → 002C , comma
  → 060C 、 arabic comma

2E42 ‡ DOUBLE LOW-REVERSED-9 QUOTATION MARK
  → 201E „, double low-9 quotation mark

Punctuation marks

2E43 ← DASH WITH LEFT UPTURN
2E44 ′ DOUBLE SUSPENSION MARK

Medieval punctuation

2E45 ′ PARAGRAPHUS MARK
  • indicates the beginning of a paragraph, section, stanza, or proposition
  → 006B ¶ pilcrow sign
  → 204B † reversed pilcrow sign
  → 2E0F ␣ paragraphos
  → 2E3F ⏵ capitulum
  → 002C , comma

2E47 Ꞌ COLON WITH SIDEWAYS REVERSED RAISED COMMA
  • indicates the end of a section of text
  → 002C , comma

2E48 ꞊ COLON WITH RAISED POSITURA MARK
  • indicates the end of a section of text

2E49 Ꞌ ꞉ COLON WITH RAISED OVER COMMA
  • indicates the end of a section of text

2E4A ꞊ PUNCTUS ELEVATUS MARK
  • indicates a minor medial pause where the sense is complete but the meaning is not

2E4B ′ SIDEWAYS REVERSED MIDDLE COMMA
  • indicates a brief medial pause

2E4C ꞊ PUNCTUS FLEXUS MARK
  • indicates a minor medial pause where the sense is incomplete

2E4D ꞊ PUNCTUS VERSUS MARK
  • indicates the melodic formula at the end of a psalm verse
  • indicates the completion of a single idea

2E4E ꞊ LOW PUNCTUS VERSUS MARK

2E4F ꞊ PUNCTUS INTERROGATIVUS MARK
  • question mark
  → 003F ? question mark
  → 00BF ¿ inverted question mark
  → 2E2E ′ reversed question mark
  • indicates a non-rhetorical question

Printed using UniBook™
(http://www.unicode.org/unibook/)

Date: 2015-12-19
<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>2E50</td>
<td>PUNCTUS EXCLAMATIVUS MARK</td>
<td>punctus admirativus, indicates an exclamation</td>
</tr>
<tr>
<td>2E51</td>
<td>MEDIEVAL COMMA</td>
<td>indicates a minor medial pause or disjunction of sense</td>
</tr>
<tr>
<td>2E52</td>
<td>HIGH DOT</td>
<td>a “distinctio” which indicates a final pause in series with full stop and raised dot</td>
</tr>
<tr>
<td>2E53</td>
<td>SIMPLEX DUCTUS MARK</td>
<td>originally used to separate matters erroneously run together</td>
</tr>
<tr>
<td>2E54</td>
<td>DOTTED SOLIDUS</td>
<td>indicates a medial disjunction less than solidus but more than punctus elevatus</td>
</tr>
<tr>
<td>2E55</td>
<td>SIGN DE RENVOI</td>
<td>associates the text with external notes</td>
</tr>
<tr>
<td>2E56</td>
<td>MIDDLE COMMA</td>
<td>used as an abbreviation sign</td>
</tr>
<tr>
<td>2E57</td>
<td>TILDE WITH DOT ABOVE AND DOT BELOW</td>
<td>used as an abbreviation sign for “est”</td>
</tr>
<tr>
<td>2E58</td>
<td>VERTICAL FIVE DOTS</td>
<td></td>
</tr>
</tbody>
</table>
A. Administrative
1. Title
Proposal to add Medievalist punctuation characters to the UCS
2. Requester’s name
Script Encoding Initiative
3. Requester type (Member body/Liaison/Individual contribution)
Liaison contribution.
4. Submission date
2016-01-25
5. Requester’s reference (if applicable)

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
Supplementary Punctuation
2. Number of characters in proposal
21.
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 B.1.
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.
See above.

C. Technical – Justification
1. Has this proposal for addition of character(s) been submitted before? If YES, explain.
Yes, some of the characters have. See N3193.
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?
The authors are members of the user community.
3c. 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?
Medievalists, Latinists, and other scholars.
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?

**Scholarly 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 punctuation 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?
9c. If YES, reference

**Some characters are ancestors of modern characters.**

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?
   Yes.
10b. If YES, is a rationale for its inclusion provided?
   Yes.
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?