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:

<table>
<thead>
<tr>
<th>Code</th>
<th>Character</th>
</tr>
</thead>
<tbody>
<tr>
<td>2E4C</td>
<td>PARAGRAPHUS MARK</td>
</tr>
<tr>
<td>2E4D</td>
<td>POSITURA MARK</td>
</tr>
<tr>
<td>2E4E</td>
<td>COLON WITH SIDEWAYS REVERSED RAISED COMMA</td>
</tr>
<tr>
<td>2E4F</td>
<td>COLON WITH RAISED POSITURA MARK</td>
</tr>
<tr>
<td>2E40</td>
<td>TWO DOTS OVER COMMA</td>
</tr>
<tr>
<td>2E41</td>
<td>PUNCTUS ELEVATUS MARK</td>
</tr>
<tr>
<td>2E42</td>
<td>SIDEWAYS REVERSED MIDDLE COMMA</td>
</tr>
<tr>
<td>2E43</td>
<td>PUNCTUS FLEXUS MARK</td>
</tr>
<tr>
<td>2E44</td>
<td>PUNCTUS VERSUS MARK</td>
</tr>
<tr>
<td>2E45</td>
<td>LOW PUNCTUS VERSUS MARK</td>
</tr>
<tr>
<td>2E46</td>
<td>PUNCTUS INTERROGATIVUS MARK</td>
</tr>
</tbody>
</table>
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.

The quintessential study 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.

Encoding punctuation characters can be difficult. Sometimes things look as though they could be sequenced, and similar atomic characters can be found in the UCS. A sequence of characters in one dimension, like the common three-dot ellipsis ... is easy to identify with a character like U+2026 … *HORIZONTAL ELLIPSIS*. That particular character was doubtless encoded as a unique character due to its presence in 8-bit character sets like MacRoman. As well be seen in 2.2 and 2.4 below, similar
characters—sequences along a one-dimensional axis—are not proposed to be encoded atomically, because they are essentially indistinguishable from the sequence. Characters which might be encoded as sequences but which are drawn across two dimensions can be much more difficult to resolve depending on font. The character TWO DOTS OVER COMMA, for instance—·;—might be hard to recognize in a monowidth font as—·,·—two MIDDLE DOTS and COMMA. Since the UCS already includes similar punctuation characters (:\: ;:: ;;), the same approach has been taken here. One of the positura marks, on the other hand, ;; can easily be composed with FULL STOPS and SEMICOLON as it has a strong horizontal element, here on the baseline. This approach ensures good typography for important medieval punctuation characters, and is of the greatest benefit to the user community.

2.1. *Paragraphus.* The beginning of a paragraph, a section, a stanza, or proposition was marked with a symbol such as γ, Γ, §, or ¶. Later this function was replaced by the paraph, where //, ¶, or § were typical marks. One character is proposed for encoding here, U+2E4C ⹌ PARAGRAPHUS MARK. Note that U+204B ⸪ REVERSED PILCROW SIGN was derived from the typographic U+00B6 ¶ PILCROW SIGN (itself a descendant of U+2E3F ⹿ CAPITULUM), and is not a glyph variant of ⹌ 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 fulfil this function: γ, Γ, ⹌, ⸹; and ¶. Some of these can be sequenced with existing characters and some cannot. Four characters are proposed for encoding here: U+2E4D ⹍ POSITURA MARK, U+24E ⹎ COLON WITH SIDEWAYS REVERSED RAISED COMMA, U+2E4F ⹏ COLON WITH RAISED POSITURA MARK, and U+2E50 ⹐ TWO DOTS OVER COMMA. Note that colon with sideways reversed raised comma has glyph variants which look like :– and ⹐ (see figures 6, 8). Scholars typically represent this with a sideways reversed raised comma (raised is not the same thing as a middle), which does not appear to be used on its own. As noted above, ;; and ; are encoded atomically on the pattern of U+10FB ⸓ GEORGIAN PARAGRAPH SEPARATOR. The others can be composed, as they follow a strong one-dimensional horizontal line:

\[
\cdotγ = \cdot + γ \quad ·γ = · + , \quad ·;γ = · + ; + , \quad ·; = · + ; + · ; = · + ·
\]

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 (:); 5, 7 (;); 5, 9 (;).

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 ‽ with a diagonal line rather than a comma can be found as a glyph variant. The offset of the top glyph element over the bottom is an integral part of the glyph design. In the Wycliffe Bible translation, the two-part character is contrasted with a similar sign lacking the lower dot, indicating a lesser pause. Clearly in such a text it is important to be able to make the distinction between the greater and lesser pause in plain text. Two characters, U+2E51 ⹊ PUNCTUS ELEVATUS MARK, and U+2E52 ⹋ SIDEWAYS REVERSED MIDDLE COMMA, have been proposed here for encoding. See Figures 7, 10, 11, 12, 13, 14, 15, 16 (;); 12, 13, (§).

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, U+2E53 ⹑ 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 (not identical to ; SEMICOLON), and the other with a sideways comma below a baseline dot. When we ask the question, "Is the placement of the glyph elements significant and needing to be preserved in a plain text encoding?", the answer for these two is “Yes”. Two characters, U+2E54 ⹔ PUNCTUS VERSUS MARK and U+2E55 ⹕ LOW PUNCTUS VERSUS MARK have been proposed here for encoding. See Figures 7, 16 (⠔); 14 (⠕).

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 /twitter/ and is typically slanted toward the right; the vertical form we know today dates to the late 15th century. One character, U+2E56 ⹖ 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; the character conventionally has two dots, unlike the modern EXCLAMATION MARK. One character, U+2E57 ⹗ 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 U+2E58 ⹘ 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, 27.

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. 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. The second set is encoded as FULL STOP, MIDDLE DOT, and U+2E59 ⹙ HIGH DOT, which is proposed here. See Figure 21.

2.5. Simplex ductus. ² This was originally a critical sign used to separate matters erroneously run together. One character, U+2E5A ⹚ 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. This character has already been accepted for encoding as U+2E4A / DOTTED SOLIDUS. 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. Proposed here as U+2E5B SIGNE DE RENVOI. See Figure 24.

2.8. **MIDDLE COMMA** (⹜) as a punctuation mark is used to indicate a variety of abbreviation. With long s, for instance, ′ is an abbreviation for sed ‘but’. It is also used sequenced alongside MIDDLE DOT ″ as an indicator of *positura*. Proposed here as U+2E5C MIDDLE COMMA. 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*. Punctuation forms based on tilde that look like math operators are systematically disunified in the UCS (compare U+223C ~ TILDE OPERATOR and U+007E ~ TILDE). The punctuation forms, in contrast to the math operators are not mirrored. Compare U+2A6A ~ TILDE OPERATOR WITH DOT ABOVE and the punctuation character U+2E1E ⩞ TILDE WITH DOT ABOVE. In addition, the range of glyph variants for this *nota* are common: ⩜. For this reason too, U+2E5D ⩝ TILDE WITH DOT ABOVE AND DOT BELOW is a different character from HOMOTHETIC. Historically a cursive form of this is the ancestor of U+A76B ꝫ 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. Proposed here as U+2E5E ⹞ VERTICAL FIVE DOTS. See Figures 26, 28.

2.11. **TRIPLE DAGGER** (⹋) is similar in use to U+2020 † DAGGER and U+2021 ‡ DOUBLE DAGGER. It indicates another level of notation. *This character has already been accepted for encoding as* U+2E4B TRIPLE DAGGER. 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+2E4D , POSITURA MARK, and U+2E5C , MIDDLE COMMA, and 2E55 ; 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+2E4E ∷ COLON WITH SIDEWAYS REVERSED RAISED COMMA, U+2E4F ∷ COLON WITH RAISED POSITURA MARK, U+2E50 ∷ TWO DOTS OVER COMMA, U+2E51 ∷ PUNCTUS ELEVATUS MARK, U+2E52 ∷ SIDEWAYS REVERSED MIDDLE COMMA,
3.3. High or tall punctuation. The characters U+003F ? QUESTION MARK, U+0021 ! EXCLAMATION MARK, U+002F / SOLIDUS, U+205A ⁏ TWO DOT PUNCTUATION, U+205B ⁛ FOUR DOT MARK, U+205D ⌹ TRICOLON, U+205E ‡ VERTICAL FOUR DOTS, and U+2E3D ⹝ VERTICAL SIX DOTS all extend from baseline or just below it to caps height, as do 2E56 ⁉ PUNCTUS INTERROGATIVUS MARK, 2E57 ⁊ PUNCTUS EXCLAMATIVUS MARK, 2E59 ⌹ HIGH DOT, 2E5A ⌷ SIMPLEX DUCTUS MARK, 2E5B ⌸ SIGNE DE RENVOI, 2E5E ⌹ VERTICAL FIVE DOTS, 2E4A ⌶ DOTTED SOLIDUS, and 2E4B ⌷ TRIPLE DAGGER.

4. Linebreaking. Line-breaking properties for these are suggested as follows. UTC recommendations are gladly accepted, if similarity to the ASCII characters ! / is a problem.

Like † U+2020 DAGGER (AI (Ambiguous)): 2E4B
Like ? U+204B REVERSED PILCROW SIGN (AL (Alphabetic)): 2E4C
Like ∴ U+2058 TURNED SEMICOLON (BA (Break After)): 2E4A, 2E4D..2E55, 2E58..2E5E
Like ! U+0021 EXCLAMATION MARK (EX (Exclamation)): 2E56..2E57

5. Unicode Character Properties. Character properties are proposed here.


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.](image1)

![Figure 2. Sample from Virgile 1509, showing PARAGRAPHUS MARK.](image2)

![Figure 3. Sample from Parkes 1993:202, showing POSITURA MARK sequenced after FULL STOP.](image3)
Figure 4. Sample from Parkes 1993:242, showing POSITURA MARK written at the ends of the discursive periods; the middle dot is used at the ends of the discursive cola.
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:

> Venite benedicti patris mei possidete regnum alleluia quod vobis paratum est ab origine 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.’

**Signos de puntuación en textos españoles**

Los signos de puntuación en códices y documentos españoles de la Edad Media es muy irregular. Agustín Millares Carlo, uno de los más rigurosos especialistas en escritura visigótica, afirmó que «los signos en códices y documentos de escritura visigótica pertenecen a más de un sistema y su equivalencia respecto a los actuales no se deja precisar...». Los signos son éstos:

```
.: ; , . /.
```

A partir del siglo xii y en textos en escritura carolingia se empezó a regularizar el sistema de puntuación con notable mengua de los signos: casi en exclusiva se utilizaron para indicar las pausas el punto (.) en línea de escritura, y el punto con virgula sobrepuést (;), la pausa final se indicó con un signo formado así: «· ·». En los documentos escritos en castellano en la baja Edad Media, la puntuación es muy irregular, adoptándose rayas y curvas acompañadas de puntos a veces para señalar y no con exactitud los periodos.

Figure 7. Sample from Núñez Contreras 1994 showing *punctus versus mark*, *punctus elenatus 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 inflectitas in uis 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 '!

... Ego temet et, alios medicos / obestgo et rogo / respondete michi! proge! quid | sibi udt permixa medicamentum multitudo / mihi quiaque te scitis / quod magis proprie prodessae possit. / et quanto! ubiis occurrat! multa quandoque perlicere / quod in uno fortute nequeas! repe-rere! / Sed quid multidum etc arume! / quid respondebis! / cum medicamina ustra de cutis in frigida! et denique de contraria in contraria communia! Nonne satis pate/ usus non | uti scientia / sed potius coniectura / Et et ad prognostica ueni-

... I earnestly entreat you and other doctors/ please reply to me/ What's the purpose of this scrambled multitude of medicines/ but that you are ignorant/ of what is most specifically efficacious/ and that you find/ that many medicines often achieve/ what you cannot hit upon in the properties of one./ But why should I assail your large numbers./ What will you reply/ when you alter your treatment from hot to cold/ and from one thing to its opposite./ Isn't it sufficiently clear/ that you employ not knowledge/ but rather conjecture./ Now

gar. Ganga xv. folc vp alandho er sogn giit vi. svfyn
dr. linda her igriform tyr gnpa lvnâ brym dyb blâ fuboc oc bvn
gvli, bar er miclo meft mengi þeira mvna nv helgi hioþung
ðvun? ren ravn bitlvp ul regin þunga en spôi vitnar at
ípar bitiði. Melnu oc mylnur ul. myrc vipar latip engi Y in

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

of Seville. An examination of the abbreviations of Monte Cassino 205 saec. xi suffices to convince us that many of them are not of the traditional Benevente stock. The constant use of symbols like # = enim, ‡ = est, ã = hoc, · = id est, q = quod, † = sed, ‡ = secundum and of the form nēe = nomine speaks for Insular influence. By a fortunate coincidence the Insular origin of the archetype can be attested by other means.

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.
Figure 24. Sample from Parkes 1993, showing SIGN DE RENVOI.

Figure 25. Sample from Thompson 1912:408, showing MIDDLE COMMA alongside MIDDLE DOT. 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 26. Sample from Owen and McKinnell showing VERTICAL FOUR DOTS and VERTICAL FIVE DOTS.
Figure 27. Samples from MS Biblioteca Nacional de España, VITR/5/9, 1st half of the 13th century, showing examples of MEDIEVAL COMMA. To the right in each picture is an example of the angular PARAGRAPHUS MARK.

Figure 28. Query from a TEX user attempting to construct VERTICAL FOUR DOTS and VERTICAL FIVE DOTS. http://tex.stackexchange.com/questions/186634/symbols-for-four-and-five-vertical-dots
# Supplemental Punctuation

<table>
<thead>
<tr>
<th>2E00</th>
<th>2E01</th>
<th>2E02</th>
<th>2E03</th>
<th>2E04</th>
<th>2E05</th>
<th>2E06</th>
<th>2E07</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Supplemental Punctuation

<table>
<thead>
<tr>
<th>2E00</th>
<th>2E01</th>
<th>2E02</th>
<th>2E03</th>
<th>2E04</th>
<th>2E05</th>
<th>2E06</th>
<th>2E07</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Example Usage

1. **Example 1:**
   - Text: *Hello, world!*
   - Code: `Hello, world!`

2. **Example 2:**
   - Text: *This is a test.*
   - Code: `This is a test.`

## Additional Notes

- Supplemental punctuation characters can be used in various contexts, such as in mathematical expressions or in specific programming languages.
- It's important to consult the Unicode standard for the latest and most comprehensive information on these characters.
New Testament editorial symbols

2E00  RIGHT ANGLE SUBSTITUTION MARKER ➞ 231C‘ top left corner
2E01  RIGHT ANGLE DOTTED SUBSTITUTION MARKER
2E02  LEFT SUBSTITUTION BRACKET
2E03  RIGHT SUBSTITUTION BRACKET
2E04  LEFT DOTTED SUBSTITUTION BRACKET
2E05  RIGHT DOTTED SUBSTITUTION BRACKET
2E06  RAISED INTERPOLATION MARKER ➞ 22A4 T’ down tack
2E07  RAISED DOTTED INTERPOLATION MARKER
2E08  DOTTED TRANSPOSITION MARKER
2E09  LEFT TRANSPOSITION MARKER
2E0A  RIGHT TRANSPOSITION BRACKET
2E0B ☞ RAISED SQUARE • used as an opening raised omission bracket
2E0C ✯ LEFT RAISED OMISSION BRACKET • used as an opening or closing raised omission bracket
2E0D ✯ RIGHT RAISED OMISSION BRACKET • used as a closing or opening raised omission bracket

Ancient Greek textual symbols

2E0E ☠ EDITORIAL CORONIS ➞ 1FBD ‘ greek koronis
2E0F ☦ PARAGRAPHOS
2E10 ☦ FORKED PARAGRAPHOS
2E11 ☦ REVERSED FORKED PARAGRAPHOS
2E12 ☦ HYPODIASTOLE • = ypodiastoli
2E13 ☦ DOTTED OBELOS • glyph variants may look like ’ς’ or ’σ’ ➞ 2052 ⁄ commercial minus sign
2E14 ☦ DOWNDWARDS ANCORA • contrary to its formal name this symbol points upwards
2E15 ☦ UPWARDS ANCORA • contrary to its formal name this symbol points downwards
2E16 ☦ DOTTED RIGHT-POINTING ANGLE • = diple periestigmene

Ancient Near-Eastern linguistic symbol

2E17 ☖ DOUBLE OBLIQUE HYPHEN • used in ancient Near-Eastern linguistics • hyphen in Fraktur text uses 002D - or 2010 -, but with a ’ς’ glyph in Fraktur fonts ➞ 0120 - hyphen-minus ➞ 003D = equals sign ➞ 2010 - hyphen ➞ 2E40 * double hyphen

Dictionary punctuation

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

2E1A ☑ HYPHEN WITH DIAERESIS • indicates umlaut of the stem vowel of a plural form
2E1B ☑ TILDE WITH RING ABOVE • indicates change in case for derived form

Brackets

2E1C ➖ LEFT LOW PARAPHRASE BRACKET
2E1D ➖ RIGHT LOW PARAPHRASE BRACKET • used in N’Ko

Dictionary punctuation

2E1E ➖ TILDE WITH DOT ABOVE • indicates derived form changes to uppercase
2E1F ➖ TILDE WITH DOT BELOW • indicates derived form changes to lowercase

Brackets

2E20 ❋ LEFT VERTICAL BAR WITH QUILL
2E21 ❋ 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 ✓ TOP LEFT HALF BRACKET ➞ 2308 ꞅ top left corner ➞ 231C‘ top left corner ➞ 300C ❋ left corner bracket
2E23 ❋ TOP RIGHT HALF BRACKET
2E24 ❋ BOTTOM LEFT HALF BRACKET
2E25 ❋ BOTTOM RIGHT HALF BRACKET

Brackets

2E26 ≤ LEFT SIDEWAYS U BRACKET ➞ 2282 ⩾ subset of
2E27 ⇒ RIGHT SIDEWAYS U BRACKET ➞ 2283 ⩾ superset of
2E28 ⩾ LEFT DOUBLE PARENTHESIS ➞ 2985 ( left white parenthesis ➞ FF5F ⩾ fullwidth left white parenthesis
2E29 ⩾ RIGHT DOUBLE PARENTHESIS

Historic punctuation

2E2A ⩾ TWO DOTS OVER ONE DOT PUNCTUATION
2E2B ⩾ ONE DOT OVER TWO DOTS PUNCTUATION
2E2C ⩾ SQUARED FOUR DOT PUNCTUATION
2E2D ⩾ FIVE DOT MARK
2E2E ⩾ REVERSED QUESTION MARK ➞ punctus percontativus ➞ 003F ? question mark ➞ 00BF ⩾ inverted question mark ➞ 061F ⩾ arabic question mark
2E2F ⩾ VERTICAL TILDE • used for Cyrillic yerik ➞ 033E ⩾ combining vertical tilde ➞ A67F ⩾ cyrillic payrerek
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 glyphic 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

Church Slavonic punctuation

2E45  ₩ INVERTED LOW KAVYKA

2E46  ₪ INVERTED LOW KAVYKA WITH KAVYKA ABOVE

2E47  ₦ LOW KAVYKA

2E48  ₦ LOW KAVYKA WITH DOT

2E49  ﹦ DOUBLE STACKED COMMA

Medieval punctuation

2E4A  / DOTTED SOLIDUS
→ indicates a medial disjunction less than solidus but more than punctus elevatus
= a form of virgula suspensiva

2E4B  § TRIPLE DAGGER
→ 2020 † dagger
→ 2021 ‡ double dagger

2E4C  † PARAGRAPHUS MARK
→ indicates the beginning of a paragraph, section, stanza, or proposition
→ 006F † pilcrow sign
→ 204B † reversed pilcrow sign
→ 2E0F † paragraphos
→ 2E3F † capitulum

2E4D  ₫ POSITURA MARK
→ indicates the end of a section of text
→ 002C , comma

2E4E  • COLON WITH SIDEWAYS REVERSED RAISED COMMA
→ indicates the end of a section of text

2E4F  • COLON WITH RAISED POSITURA MARK
→ indicates the end of a section of text

2E50  • TWO DOTS OVER COMMA
→ indicates the end of a section of text

2E51  • PUNCTUS ELEVATUS MARK
→ indicates a minor medial pause where the sense is complete but the meaning is not

2E52  • SIDEWAYS REVERSED MIDDLE COMMA
→ indicates a brief medial pause

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

Date: 2016-05-04  21
2E53  ⹓ PUNCTUS FLEXUS MARK
- indicates a minor medial pause where the sense is incomplete

2E54  ⹔ PUNCTUS VERSUS MARK
- indicates the melodic formula at the end of a psalm verse
- indicates the completion of a single idea

2E55  ⹕ LOW PUNCTUS VERSUS MARK
2E56  ⹗ PUNCTUS INTERROGATIVUS MARK
- \003F \? question mark
- \00BF \¿ inverted question mark
- \2E2E \? reversed question mark
- indicates a non-rhetorical question

2E57  ⹘ MEDIEVAL COMMA
- indicates a minor medial pause or disjunction of sense

2E58  ⹙ MEDIEVAL COMMA
- indicates a minor medial pause or disjunction of sense

2E59  ⹙ HIGH DOT
- a “distinctio” which indicates a final pause in series with full stop and raised dot

2E5A  ⹛ SIMPLEX DUCTUS MARK
- originally used to separate matters erroneously run together

2E5B  ⹜ MEDIEVAL COMMA
- used as an abbreviation sign
- \00b7 \· middle dot
- \2E34 \⸴ raised comma

2E5C  ⹝ TILDE WITH DOT ABOVE AND DOT BELOW
- \223B \∻ homothetic
- used as an abbreviation sign for “est”

2E5D  ⹞ VERTICAL FIVE DOTS
- \205D \__: tricolon
- \205E \__: vertical four dots
- \2E3D \__: vertical six dots
A. Administrative
1. Title
Revised 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
2016-05-04
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
Supplementary Punctuation
2. Number of characters in proposal
19.
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.
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?
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?

No.

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?

Yes.

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

Yes.

10c. If YES, reference

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

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?

No.

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?