

TO: Unicode Technical Committee
FROM: Stephen Emmel, University of Muenster and Jost Gippert, Goethe Universität Frankfurt am Main (via Deborah Anderson, SEI, UC Berkeley)
SUBJECT: Expert Feedback on L2/13-140 SLAVONIC PARAGRAPHOS
DATE: 1 November 2013

The Unicode Technical Committee requested feedback on the use of the Slavonic Paragraphos character in the wider Byzantine milieu. The comments below provide feedback from Stephen Emmel, a specialist in Coptic, and Jost Gippert, a specialist in Indo-European and non-Indo-European languages (including Georgian and Caucasian Albanian).

Feedback from: Stephen Emmel, University of Muenster (stephen.emmel@t-online.de)

First, the use of the term "paragraphos" for the character proposed in L2/13-140R seems from my perspective (Coptic and Greek) to be infelicitous because the "paragraphos" as traditionally defined (above all in Greek paleography) occurs only in the left margin (perhaps sometimes in the right margin, but I do not know that for a fact), appearing to be between two lines of text. By definition, a paragraphos does not occur in mid-line, nor at the end of a line, as in the examples in the proposal. I would have called the proposed character a kind of "dash"; for Greek, see, e.g., E.G. Turner, *Greek Manuscripts of the Ancient World* (Oxford 1971), p. 10, last full paragraph: "The double dot : , often called dicolon, . . . combined with a dash (:—)" in mid-text.

Second, I would also not call the proposed character specifically "Slavonic." Certainly this character (including its upward flourish at the left, or "up-turn on the left side") occurs frequently also in Coptic manuscripts, in effectively identical usage, and also in combination with (i.e. following) other punctuation (raised dot, colon, three- and four-dot punctuation, etc.). I would find it hard to believe that this character is not widespread across various language/manuscript cultures. I do not know of any particular name for the character.

But third, by calling this character a "paragraphos," this proposal makes the interesting observation that the function of this particular dash is very similar (even identical?) to the function of the (Greek etc.) paragraphos, namely to mark the end/beginning of a "paragraph" (however defined). But its position (mid-line and raised) is distinct, for which reason I would say that it cannot be unified with any of the existing paragraphoi in Unicode.

Now to answer your specific question to me. In Coptic, we find a variety of shapes for the paragraphos (I mean the paragraphos proper, "the marginal paragraphus" [Turner, p. 15], whose "proper place is below the line in which the pause occurs[, thus marking] the end, not the beginning, of a sentence [or paragraph etc.]" [F.G. Kenyon, *The Palaeography of Greek Papyri* (Oxford 1899), p. 27]). The basic distinction is between a simple straight line, on the one hand, and a straight line with various more or less ornate additions, the "forked paragraphos" being the simplest of the more ornate forms. Because I understand that Unicode cannot be expected to provide a fixed code point for every variation that we find in our manuscripts (or printed editions), it has seemed to me that the three existing paragraphoi must (and can) suffice for Coptic.

But I want also to emphasize that the "Slavonic Paragraphos" is something different, even if functionally related. One might call it an "In-line paragraphos" (if "in-line" may be used in this sense, opposite to "marginal"). As sensitive as I am to specialists in a manuscript-languages wanting to have Unicode code points for each and every phenomenon in their manuscripts, in this case I cannot help but think that one should unify the "Slavonic Paragraphos" with one of the existing dashes in Unicode. For Coptic, I would say that a left-hand upward flourish on this "in-line paragraphos" is not an essential feature of the character, even though it occurs very often. Often the character is straight, or it might have a flourish on the right, or it might be slanted or curved, and so on. Therefore, in a document that I have been working on (very slowly) to present "best practices in using Unicode to encode Coptic text," I will recommend using U+2015 HORIZONTAL BAR for this purpose (with U+2E3A and U+2E3B available if one wants to make distinctions). As in Michael's proposal regarding the Caucasian Albanian paragraphos, over a Unicode-based Coptic font would use a Coptic-appropriate glyph for U+2015. I would think that one could do the same for Slavonic. (If I were to need a "real" horizontal bar in my document apart from the Coptic in-line paragraphos, then I would use a series of en- or em-dashes, as I have always done for such a purpose.)

Feedback from: Jost Gippert, Goethe Universität Frankfurt am Main (gippert@em.uni-frankfurt.de)

I have not yet finished my survey of paragraphoi across Greek-based scripts - as a matter of fact, this would be a good topic for a major investigation. Anyway, I feel uncomfortable with Stephen [Emmel]'s recommendation to use "U+2015 HORIZONTAL BAR for this purpose (with U+2E3A and U+2E3B available if one wants to make distinctions) " as this would lead us back into the times of mere font mapping, which we all hoped Unicode would put an end to. If you consider the examples [...] from the Gothic palimpsest at Wolfenbuettel, you will see that this is far from matching a mere horizontal bar, and so are the paragraphoi of other traditions I intend to document one day. So replacing this by U+2015 can be, in my view, only a temporary expedient (by the way, I have rendered the Gothic one by a sequence of U+21BC, U+203A, U+21C1, which is awkward and not less temporary). I would propose to postpone the paragraphos discussion to a day when we have enough materials at hand.