

**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 to add additional characters for Greek, Latin, and Coptic to the UCS

**Source:** Michael Everson, Stephen Emmel (Universität Münster), Siegfried G. Richter (Institut für die Neutestamentliche Textforschung, Universität Münster), Susana Pedro (Faculdade de Letras da Universidade de Lisboa), António Emiliano (Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa)

**Status:** Individual Contribution

**Action:** For consideration by JTC1/SC2/WG2 and UTC

**Date:** 2010-08-05

**0. Introduction.** This proposal requests two additional letters for Coptic, and three generic punctuation marks for use with (at least) Coptic, Greek, and Latin, and a symbol with a special function in Coptic linguistics. If this proposal is accepted, the following characters will exist:

ᲀ 2CF2 COPTIC CAPITAL LETTER BOHAIRIC KHEI  
 ᲁ 03E6 coptic capital letter kheï

ᲂ 2CF3 COPTIC SMALL LETTER BOHAIRIC KHEI  
 ᲃ 03E7 coptic small letter kheï

• 2E3C RAISED DOT  
 · 002E full stop  
 · 00B7 middle dot

ᶘ 2E3D RAISED COMMA  
 , 002C comma  
 · 00B7 middle dot

ᶙ 2E3E MIDDLE COMMA  
 , 002C comma  
 · 00B7 middle dot

† 2E3F SMALL RAISED DAGGER  
 † 2020 dagger

**1. Additional letters for Coptic.** A large number of Coptic manuscripts and printed books distinguish between two different forms of the letter KHEI. One is encoded as ᲁ ᲁ COPTIC LETTER KHEI at U+03E6–U+03E7. In order to represent the orthography of these texts explicitly, ᲀ ᲀ COPTIC LETTER BOHAIRIC KHEI is proposed for encoding here at U+2CF2–U+2CF3. (Figures 1, 2, 3, 4, 5)

**2. Raised dot.** This character is required to fill a standardized set of punctuation marks used both in Greek and in Coptic. Edward Maunde Thompson describes the usage in Greek in his *Introduction to Greek and Latin Palaeography* (1912, p. 60):

A more regular system was developed in the schools of Alexandria, its invention being ascribed to Aristophanes of Byzantium (260 B.C.). This was the use of the full point with

certain values in certain positions (θέσεις): the high point (στιγμὴ τελεία), equivalent to a full stop; the point on the line (ὑποστιγμὴ), a shorter pause, equivalent to our semicolon; and the point in a middle position (στιγμὴ μέση), an ordinary pause, equivalent to our comma.

In UCS terms, the *stigmē teleia* is equivalent to U+00B7 the MIDDLE DOT (being at x-height *vis-à-vis* Greek and Coptic uncial), the *hypostigmē* is equivalent to U+002E the FULL STOP (being on the baseline), and the equivalent to the *stigmē mesē* is RAISED DOT proposed here. Note that the specific definition of the raised dot is that its height is between the full stop and the middle dot. Note too that a dot at this position can be found in THREE DOT PUNCTUATION, FIVE DOT MARK, and FIVE DOT PUNCTUATION. Compare RAISED DOT plus COLON and THREE DOT PUNCTUATION thus — ·:·.

The three dots at different heights were used in Coptic texts as in Greek ones. Hans Quecke (1977) describes this usage using slightly different terms than Thompson did:

And here [in this particular manuscript] one can recognize principles according to which the use of the dot in one of the three possible positions – *stigmē anō* [MIDDLE DOT], *stigmē mesē* [RAISED DOT], and *stigmē katō* [FULL STOP] – seems to be regulated. The dot up above the line [MIDDLE DOT] comes closest in usage to our full stop and marks the strongest break. The dot down on the line [FULL STOP] occurs after pre-posed clauses and phrases.... Also among the many occurrences of dots at the halfway height of the line [RAISED DOT] one can recognize at least two characteristic usages, that between parallel members, and that before direct and indirect speech. (Original German text in Figure 12.)

Hans Martin Schenke (1991) also describes the Coptic usage of the three dots at different heights:

The dot as a punctuation mark ... occurs in various positions: well above the line (that is, superior, at the height of the topmost part of a normal letter) [MIDDLE DOT], on the line [FULL STOP], and at various heights in between [RAISED DOT].... Taken all in all, the various heights of the dots may be seen to indicate a difference in function.... Of these dots, the one well above the line [MIDDLE DOT] clearly has the greater force of division and marks, as a rule, the break between complete sentences or verses. In contrast, both the mid-level dot [RAISED DOT] as well as the dot on the line [FULL STOP] – even though as a practical matter they are met *also* in this function, more or less – serve in principle only to set off clauses. The dot on the line has furthermore a certain outward affinity to the comma, whose tail extending below the line can be of varying length. Probably the comma is also a mark for dividing off (smaller) parts of a sentence. (Original German text in Figure 13.)

(See Figures 10 and 11.)

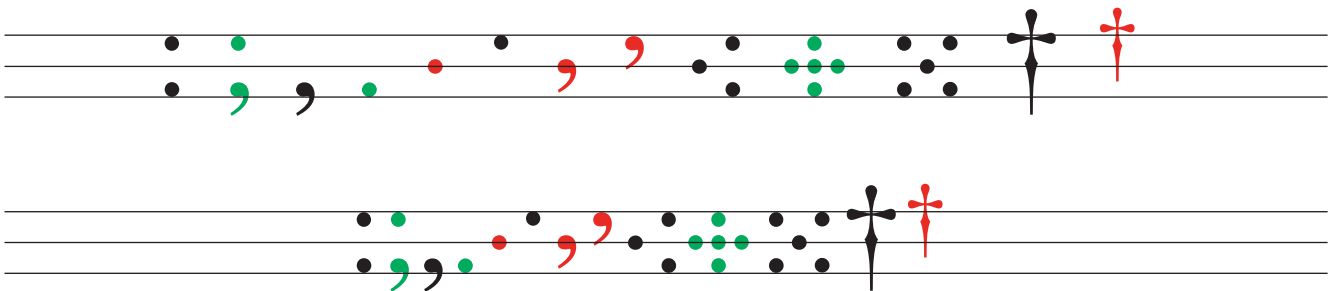
**3. Raised comma.** In many Coptic manuscripts, RAISED COMMA is a punctuation mark distinct both from COMMA and RAISED DOT. Occasionally, this feature of Coptic punctuation has been rendered very nicely in Coptic fonts. (Figures 6, 7)

**4. Middle comma.** This character was previously proposed in N3193 “Proposal to add Medievalist and Iranianist punctuation characters to the UCS” (2007-01-09). A number of characters in that document were put off for further study and this is one of them. In medieval European manuscripts, the MIDDLE COMMA is used along with MIDDLE DOT for a certain kind of positura: ‘·’. This is one of a set of positurae which can be composed with existing characters: ‘.’, and ‘.,’ are other examples. The MIDDLE COMMA is also used as an abbreviation mark; for instance, when it follows long s (ſ), the reading is *sed* ‘but’. (Figures 8, 9)

**5. Small raised dagger.** This character has a special function in Coptic linguistics to mark the “stative” (alias “qualitative”) forms of Coptic verbs, or the meaning of a stative form, in dictionaries and grammars. We do not know who first introduced this convention into Coptic linguistics, but its usage is now nearly a century old, at least, and firmly established. The existing DAGGER (a normal full-sized † on the baseline) is used by Coptologists as well, as in text editions generally, to mark (“obelize”) corrupt passages. (Figures 14, 15)

## 6. Comparison of relative glyph heights of existing punctuation and the characters proposed here.

In the illustration below, COLON, SEMICOLON, COMMA, FULL STOP, the proposed **RAISED DOT**, MIDDLE DOT, the proposed **RAISED COMMA**, the proposed **MIDDLE COMMA**, THREE DOT PUNCTUATION, FIVE DOT MARK, FIVE DOT PUNCTUATION, DAGGER, and the proposed **SMALL RAISED DAGGER** are shown. In the first line there is a space between the characters; in the second there is no space. Some characters are coloured green for contrast; the proposed characters are coloured red.



## 7. Unicode Character Properties. Character properties are proposed here.

```
2CF2;COPTIC CAPITAL LETTER BOHAIRIC KHEI;Lu;0;L;;;;N;;;2CF3;
2CF3;COPTIC SMALL LETTER BOHAIRIC KHEI;Ll;0;L;;;;N;;;2CF2;;2CF2
```

```
2E3C;RAISED DOT;Po;0;ON;;;;N;;;;;
2E3D;RAISED COMMA;Po;0;ON;;;;N;;;;;
2E3E;MIDDLE COMMA;Po;0;ON;;;;N;;;;;
2E3F;SMALL RAISED DAGGER;Po;0;ON;;;;N;;;;;
```

## 8. Bibliography.

- Chassinat, Émile. 1911. *Le quatrième livre des entretiens et épîtres de Shenouti*. (Cairo: Imprimerie de l’Institut Français d’Archéologie Orientale.
- Crum, W. E. 1939. *A Coptic Dictionary*. Oxford: Clarendon Press.
- Layton, Bentley. 2004. *A Coptic Grammar with Chrestomathy and Glossary: Sahidic Dialect*. 2d ed. Wiesbaden: Harrassowitz Verlag.
- Loew, E. A. 1914. *The Beneventan Script. A history of the South Italian minuscule*. 1999 special edition. London: Clarendon Press ISBN 0-19-924015-9.
- Quecke, Hans. 1977. *Das Lukasevangelium säidisch: Text der Handschrift PPalau Rib. Inv.-Nr. 181 mit den Varianten der Handschrift M 569*. Barcelona: Papyrologica Castrocaviana.
- Schenke, Hans-Martin, 1991. *Apostelgeschichte 1,1–15,3 im mittelägyptischen Dialekt des koptischen (Codex Glazier)*. Berlin: Akademie Verlag.
- Störk, Lothar. 1995. *Verzeichnis der orientalischen Handschriften in Deutschland*, vol. 21, *Koptische Handschriften*, vol. 3: *Die Handschriften der Staats- und Universitätsbibliothek Hamburg*, pt. 2, *Die Handschriften aus Dair Anbā Maqār*. Stuttgart: Franz Steiner Verlag.
- Störk, Lothar. 1996. *Verzeichnis der orientalischen Handschriften in Deutschland*, vol. 21, *Koptische Handschriften*, vol. 3: *Die Handschriften der Staats- und Universitätsbibliothek Hamburg*, pt. 3, *Addenda und Corrigenda zu Teil I*. Stuttgart: Franz Steiner Verlag.
- Thompson, Edward Maunde. 1912. *An Introduction to Greek and Latin Palaeography*. Oxford: Clarendon Press.



Figures

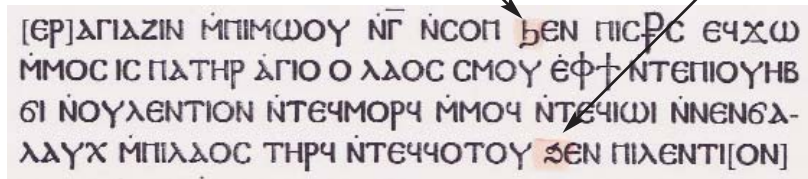


Figure 1. Sample from Störk 1996 (p. 316) of a printed Coptic text showing KHEI in line 1 and BOHAIRIC KHEI in line 4.

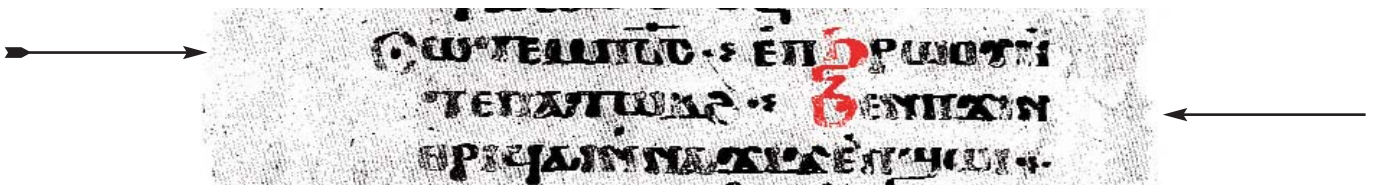


Figure 2. Sample from Störk 1995 (p. 65) of a Coptic manuscript showing KHEI in line 1 and BOHAIRIC KHEI in line 2.

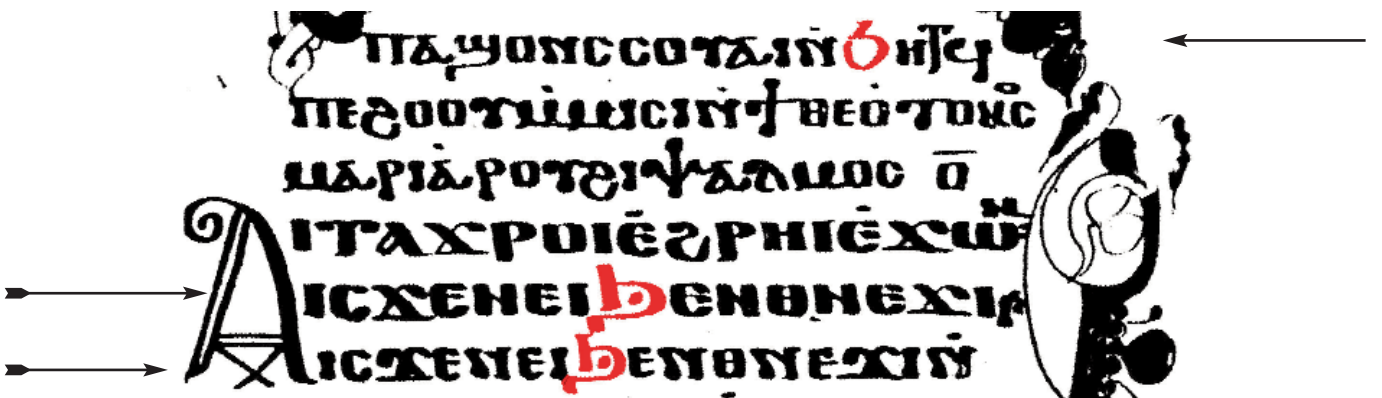


Figure 3. Sample from Störk 1995 (p. 66) of a Coptic manuscript showing BOHAIRIC KHEI in line 1 and KHEI in lines 5 and 6.

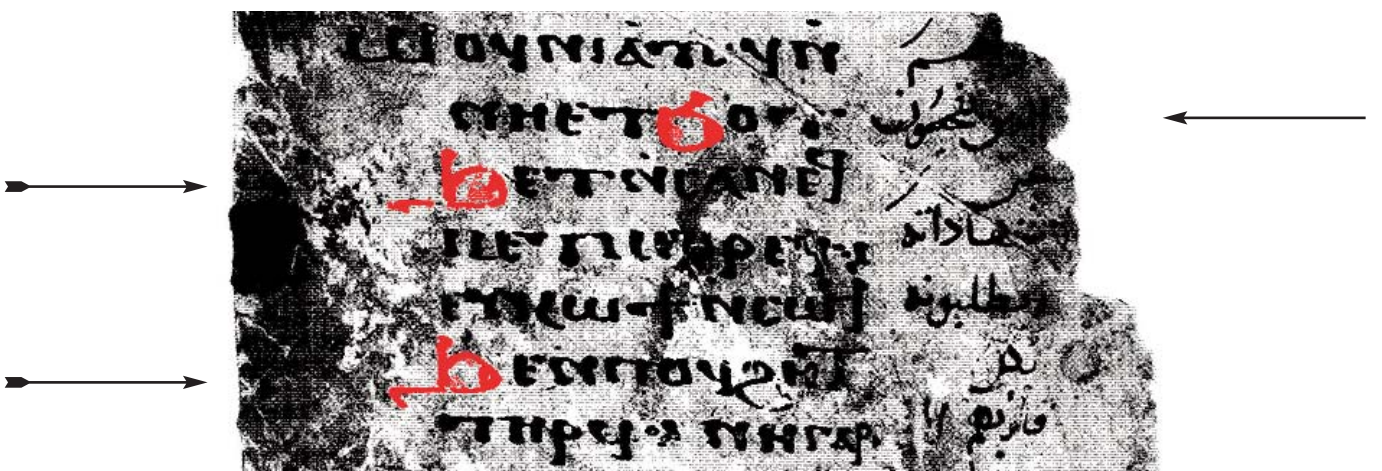


Figure 4. Sample from Störk 1995 (p. 79) of a Coptic manuscript showing BOHAIRIC KHEI in line 2 and KHEI in lines 3 and 6

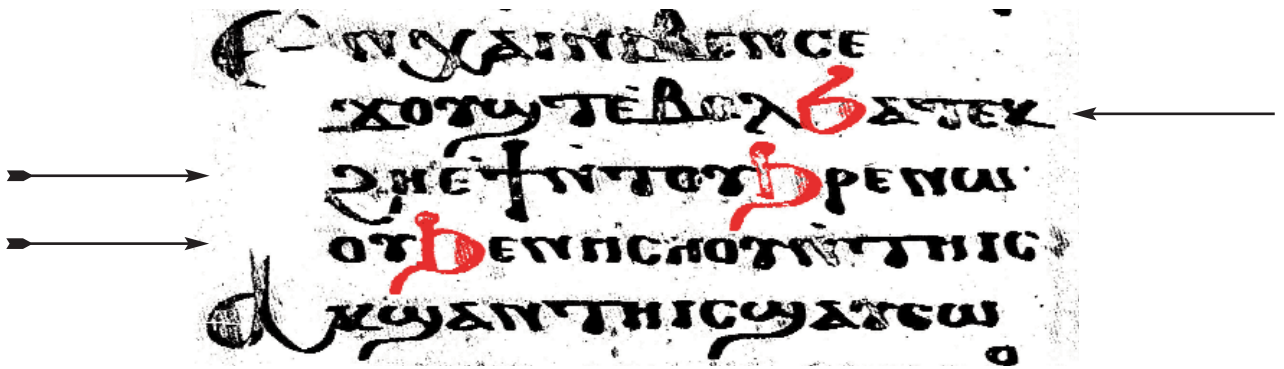


Figure 5. Sample from Störk 1995 (p. 116) of a Coptic manuscript showing BOHAIRIC KHEI in line 2 and KHEI in lines 3 and 4.



Figure 6.1. Sample from Chassinat 1911 (p. 63) of a printed Coptic text showing examples of RAISED COMMA in lines 1, 4, 7, and 20; compare the COMMA in lines 11, 15, and 19.

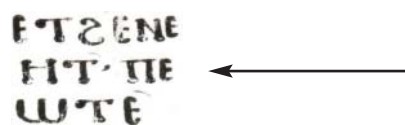


Figure 6.2. Sample of RAISED COMMA in a Coptic manuscript, in the centre of the second line of the excerpt.



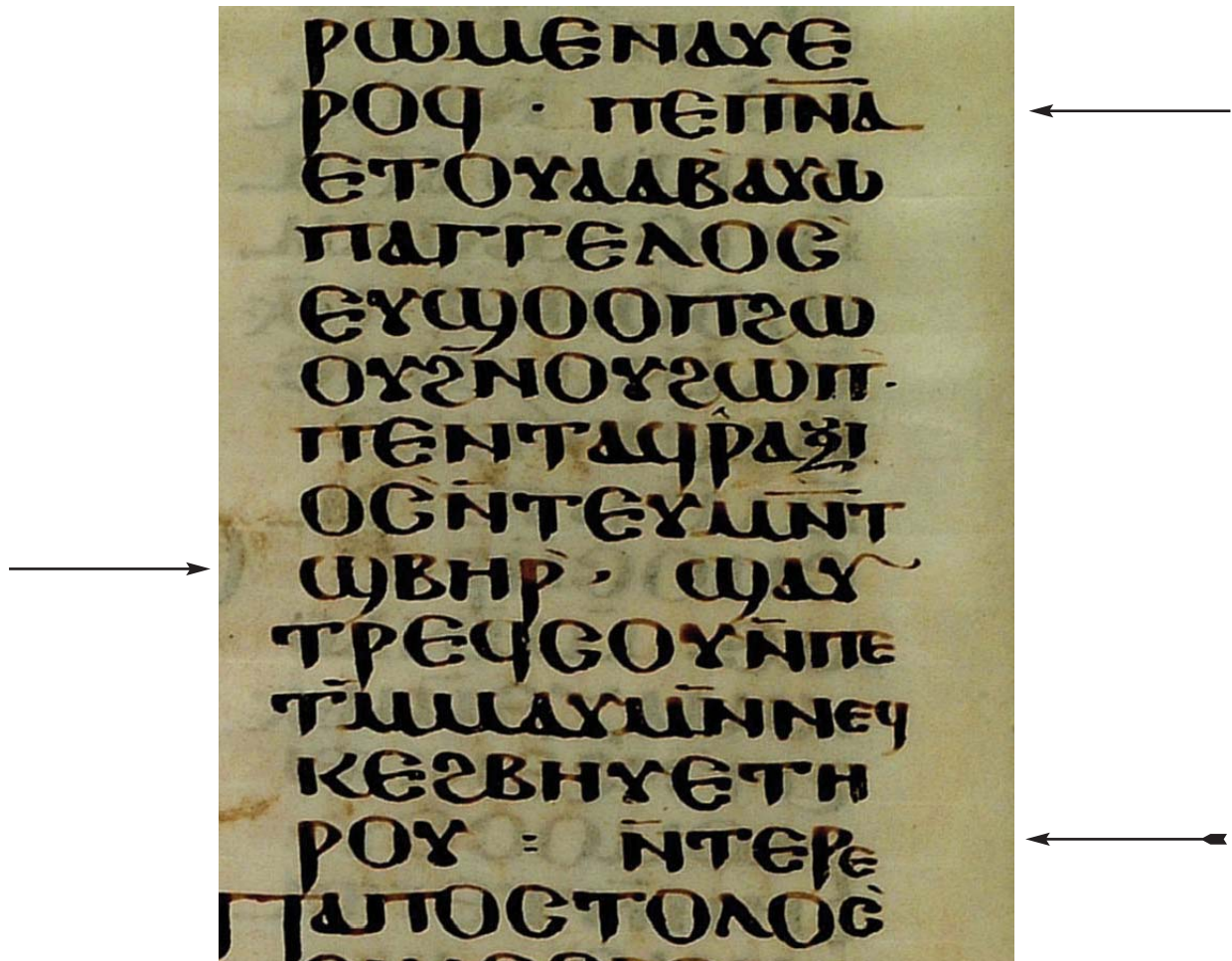


Figure 7.1. Sample of RAISED DOT (line 2) as distinct from RAISED COMMA (line 9) and COLON (line 13) in a Coptic manuscript. (The COPTIC MORPHEME DIVIDER occurs here in lines 3, 4, 6, 8, 9, 14)

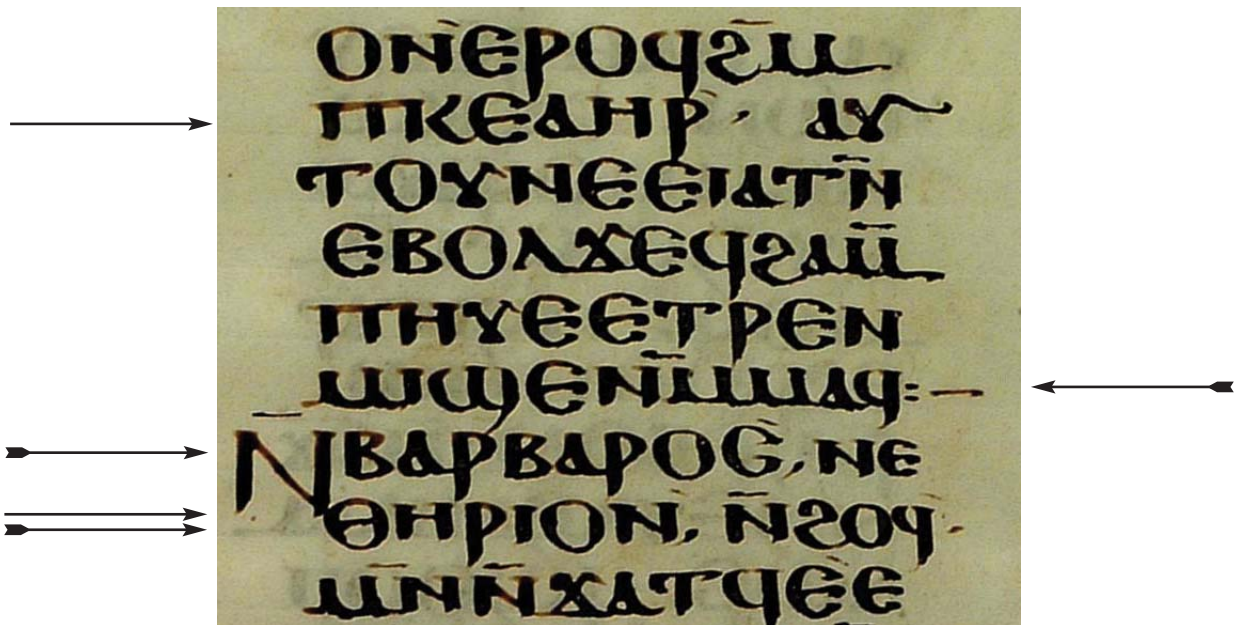


Figure 7.2. Sample of RAISED COMMA (lines 2, 8) as distinct from COMMA (lines 7, 8) and COLON (line 6) in a Coptic manuscript. (The COPTIC MORPHEME DIVIDER occurs here in lines 1, 2, 7, 8, 9).

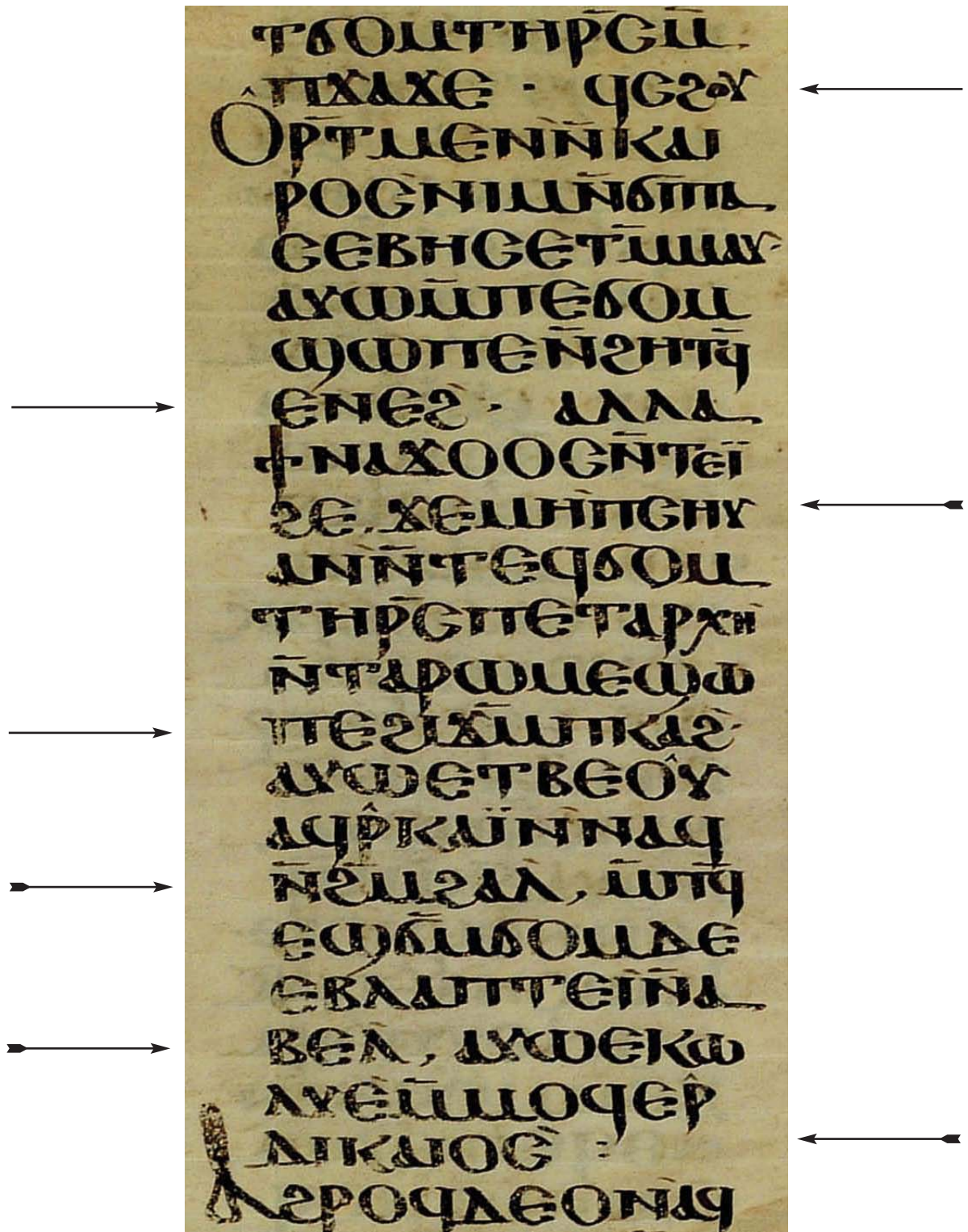


Figure 7.3. Sample of RAISED COMMA (lines 8, 14) as distinct from COMMA (lines 17, 20), FULL STOP (line 10), RAISED DOT (line 2), and COLON (line 22) in a Coptic manuscript. (The COPTIC MORPHEME DIVIDER occurs here in lines 3, 4, 8, 8, 10, 11, 12, 14, 16, 17, 18, 19, 20, 21, 22).



capti sunt. & cum illis gentilitate conuixerunt. cū  
 adhuc ad romaniam iuvenes uenerint. si commu-  
 nionem petierint quid eis obseruandum sit.  
 Si conuiuio solo gentilium. & escis immolaticis usi  
 sunt. possunt ieiunus & manus in positione pur-  
 gari. ut deinceps idolaticis abstinentes. sacramē-  
 torum xp̄i possint esse participes. Si autem  
 aut idolum adorauerunt. aut homicidiis ut forni-  
 cationibus contaminati sunt. ad communionē

THEOLOGICAL TRACTS. - A. D. 821

(capti sunt · et cum illis gentilitate conuixerunt · cum | adhuc ad romaniam ·  
 iuvenes uenerint · si communionem petierint quid eis obseruandum sit ·  
 Si conuiuio solo gentilium · et escis immolaticis usi | sunt · possunt ieiunus  
 et manus in positione purgari · ut deinceps idolaticis abstinentes · sacramē-  
 torum christi possint esse participes · Si autem | aut idolum adorauerunt · aut  
 homicidiis uel fornicationibus contaminati sunt · ad communionem)

Figure 8. Sample from Thompson 1912, showing MIDDLE COMMA alongside MIDDLE DOT.

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 Beneventan stock. The constant use  
 of symbols like  $\ddagger$  = *enim*,  $\gamma$  = *et*,  $\div$  = *est*,  $\text{h}$  = *hoc*,  $\cdot$  = *id est*,  
 $\phi$  = *quod*,  $\text{f}$  = *sed*,  $\text{f}$  = *secundum* and of the form  $\text{n}\ddot{o}\text{e}$  = *nomine*  
 speaks for Insular influence. By a fortunate coincidence the  
 Insular origin of the archetype can be attested by other means.

Figure 9. Sample from Loew 1914, showing LONG S followed by MIDDLE COMMA with the reading *sed* ‘but’.





Punkte als Satzzeichen werden im Lk-Teil weit häufiger verwendet als im Mk-Teil<sup>1</sup> und auch in etwas abweichender Weise. Es finden sich aber auch im Lk-Teil noch genügend Stellen, an denen keinerlei Satzzeichen steht, obwohl man ein solches erwartet. Im Gegensatz zum Mk-Teil, in dem fast ausschließlich der hochgestellte Punkt (in verschiedenen Stellungen) verwendet wird, macht der Lk-Teil auch ausgiebigen Gebrauch von dem Punkt unten auf der Linie. Und es lassen sich hier schon eher Prinzipien erkennen, nach denen der Gebrauch des Punktes in einer der drei möglichen Stellungen — *στιγμή ἄνω*, *στιγμή μέση* und *στιγμή κάτω* — geregelt scheint<sup>2</sup>. Der Punkt ganz oben auf der Zeile kommt unserem Punkt am nächsten und bedeutet den stärksten Einschnitt. Der Punkt unten auf der Linie steht nach voranstehenden Satzteilen und Teilsätzen. Hierher gehören alle Arten von Vordersätzen, vor allem konditionale, aber auch äquivalente Konstruktionen wie Umstandssätze oder verwandte wie Temporalsätze. Ein typischer Gebrauch ist sodann der nach satzeinleitendem Vokativ. Eine ganze Reihe von Vorkommen betrifft die Verwendung zwischen fortführenden Sätzen, teilweise vor Partikeln in entsprechender Anwendung (*ΑΛΛΑ*, *ΔΕ* oder „und“), aber auch ohne eine solche (z. B. Lk 23,18) und nach entsprechenden Fragen (z. B. Lk 7,26). Zweimal steht dieser Punkt sogar im Nominalsatz vom Typ Subjekt/Prädikat/Kopula nach dem Subjekt (124 a 21 und b 1; 125 a 20), das durch ein Attribut erweitert ist. Auch das zweimalige Vorkommen vor komparativem „als“ (Lk 15,7; 17,2) wird man als reguläre Verwendung dieses Zeichens ansehen können. Nur zweimal (von insgesamt etwa 80 Vorkommen)

<sup>1</sup> Vgl. *Mk* 16 f.

<sup>2</sup> Die Stellung des Zeichens ist fast immer klar. Nur an wenigen Stellen kann man zweifeln, ob die *στιγμή ἄνω* oder *μέση* gemeint ist. Die *στιγμή κάτω* und *μέση* stehen bei einem *τ* unter dessen Querbalken.

16

steht die *στιγμή κάτω* nach meinem Urteil am Satzschluß (Lk 2,49; 18,21), was man ohne Schwierigkeit als Versehen betrachten kann. Auch unter den vielen Vorkommen des Punktes auf halber Höhe der Zeile lassen sich zumindest zwei charakteristische Gebrauchsweisen erkennen, die zwischen parallelen Gliedern und die vor der direkten und indirekten Redc. Im letzten Fall kann der Punkt vor oder hinter dem *ΧΕ* stehen, wobei zu Beginn des Lk-Teiles der Handschrift die beiden Stellungen ziemlich wahllos wechseln, später aber die Stellung hinter dem *ΧΕ* deutlich überwiegt<sup>1</sup>.

<sup>1</sup> Auch hierin ein gewisser Unterschied zum Mk-Teil, wo bei *ΧΕ* gewöhnlich gar kein Punkt steht, aber vor dem *ΧΕ*, wo er ausnahmsweise gesetzt ist; siehe *Mk* 17.

**Figure 12.** Quecke 1977, pp. 15–16, with a statement about the use in Coptic manuscripts of punctuating dots at three different heights; see §3 above on page 2. Footnote 2 to the first passage reads: “The placement of the mark is almost always clear. Only in a few places is it doubtful whether the *stigmē anō* [MIDDLE DOT] or *mesē* [RAISED DOT] is intended. The *stigmē katō* [FULL STOP] and *mesē* [RAISED DOT] when next to a *τ* stand under its horizontal bar.” Thus: *τ*. FULL STOP, *τ*. RAISED DOT, *τ*. MIDDLE DOT. The Greek terms used here are literally *stigmē anō* ‘dot above’, *stigmē mesē* ‘middle dot’, *stigmē katō* ‘dot below’.

Mit dem Punkt als Satzzeichen verhält es sich im Codex Glazier wie im Codex Scheide. Er kommt in verschiedenen Stellungen vor: auf der Oberzeile (also hochgestellt, auf der Höhe des oberen Endes der Normalbuchstaben), auf der Unterzeile und in verschiedenen Höhen dazwischen. Die Position des Punktes kann äußere Gründe haben. Er wird gern da gesetzt, wo am meisten Platz ist; d. h., hinter T und Y wird er – unabhängig von seiner Funktion – häufig niedrig gesetzt. Gleichwohl dürfte – aufs Ganze gesehen – die verschiedene Höhe der Punkte eine Differenz in der Funktion anzeigen. Die konkrete Weise, wie in der Handschrift die Punkte wirklich gesetzt sind, legt es nahe, die fast oben und die fast unten gesetzten Punkte als bloß graphische Varianten des oberzeiligen bzw. des unterzeiligen Punktes zu deuten und in allen übrigen verschiedene Ausführungen eines Punktes auf halber Höhe zu sehen – und entsprechend im Druck wiederzugeben.<sup>85</sup> Von diesen Punkten hat der oberzeilige deutlich die größere Trennungskraft und bildet in der Regel die Zäsur zwischen ganzen Sätzen bzw. Versen. Demgegenüber dienen sowohl der halbhohe als auch der unterzeilige Punkt, obgleich sie praktisch, mehr oder weniger, auch in dieser Funktion begegnen, im Prinzip nur der Abgrenzung von Satzteilen. Der unterzeilige Punkt hat wiederum eine gewisse äußerliche wie funktionelle Affinität zum Komma, dessen von der Unterzeile ausgehender Schwanz verschieden lang ausfallen kann. Auch das Komma ist wohl eine Zäsur zur Abgrenzung von (kleineren) Satzteilen. Freilich könnte man an der realen Existenz des Kommas in unserer Handschrift fast zweifeln. In den meisten Fällen könnte man jedenfalls die betreffenden, in der Handschrift ja wirklich vorhandenen Schwänze auch als etwas nach rechts verrutschte Bogenschlußelemente der Buchstaben € und C deuten.<sup>86</sup> Aber einerseits erscheint der Schwanz auch, wenn das Bogenschlußelement, normalerweise ein aufgesetzter Punkt, schon da ist, andererseits erscheint das gleiche Zeichen auch nach dem Buchstaben X. Gleichwohl bleibt die Wiedergabe des Schwanzes nach € und C eine Interpretationsfrage. Wir interpretieren und drucken dieses Zeichen nach € und C als Satzteiltrenner „Komma“, wenn es deutlich von dem Buchstaben abgerückt ist, groß genug ausfällt und an einer Stelle erscheint, wo eine Zäsur Sinn gibt. Man könnte dann in bezug auf die Kommata unserer Handschrift sagen: Das Komma erfüllt nach Wörtern, die auf €, C und X enden, in etwa bzw. weithin die gleiche Funktion wie der Apostroph bei den Wörtern mit den Endkonsonanten β, κ, λ, μ, ν, π, ρ, τ, ϑ (Γ, Λ, Ο). Nur in dieser Optik dürfte übrigens verständlich sein, daß das Komma auch am Ende eines Satzes erscheint mit dem satztrennenden Punkt darüber.

<sup>85</sup> Vgl. zu den drei Stellungen des Punktes in alten Handschriften und der damit gegebenen verschiedenen Trennkraft besonders Quecke: LKEv sa., 15f.

<sup>86</sup> Vgl. die Beschreibung dieser beiden Buchstaben oben S. 32.34.

**Figure 13.** Schenke 1991, pp. 39–40, with a statement about the use in Coptic manuscripts of punctuating dots at three different heights; see §3 above on page 2.



<p>αταρηζ, -αυρηζ SB negative adj, limitless, boundless: Job 5 9 S (B ατσοτσετ) ανεξιχνιαστος, cf Eph 3 8 B; 2 Pet 2 4 S (B Gk diff); BMis 416 S of sea, BG 81 S of light.</p> <p>αc SF, ααc S, εc AF, απαc SB, απεc F adj, old: a. Lev 25 22 SB, Jos 9 10 S, Jer 45 11 B, Mt 9 17 SB, Ro 6 6 SB παλαιός; Is 65 22 SB ειρε πα., He 8 13 SBF sim παλαιου; Si 9 12 S (only) αρχαίος; CaiEuch 460 B μετχαχι πα. χρόνιος; AM 252 B φτραφι πα. Old Test, MG</p>	<p>Wess 18 124 S I cannot strike with ουσ. or with ογσερωc, Mor 31 58 S † πογμινше нш.</p> <p>αcιnc S υ ααα.</p> <p>αcαι SF, εcιεε A, αcαι B, αcωγ† SF (αcο-ογ† PS 218, αcωγ ib), αcεωγ† S (Tor 27), εcωγ† A, αcωγ† (αcωγ 2 Kg 1 23, Cat 53 &amp;c) BO (AZ 21 100) ιβ, ι. intr, A be light, relieved. a. Jo 4 52 B (S αατοπ) κομφοτ. εχειν, Ac 27 38 SB κομφιζειν, BSM 13 B of floods εαγα., BM 1104 S of health, Tell us γεατπεα. b. with dependent προ: c. c. Ion 1 ε SA (B εβολ αα) κομφιζεσθαι</p>
--	--

Figure 14. Sample from Crum 1939 (p. 17, see col. 2, lines 4–6) showing standard Coptological usage of a SMALL RAISED DAGGER to mark a certain verb form called the “stative” (or “qualitative”).

161 Specific negation of the lexical content (as such) of an infinitive is expressed by the negator τ̄- 105(c). E.g. πτ̄-†-<sup>0</sup>βωντ δε να-α ShAmél II 233:13 “Not making Him angry”; πτ̄-οπ-κ̄ ApophPatr 105 (Chaîne 25:2 = Z. 298:7) “Not to think highly of oneself” 181(f); νανογ-<sup>0</sup>τ̄-ογем-<sup>0</sup>αα ογδε ε-<sup>0</sup>τ̄-σε-<sup>0</sup>ηρ̄ Rom 14:21 “It is right not to eat meat or drink wine”; νανογ-πτ̄-ωινε ηζογο c-<sup>0</sup>ωινc ShAmél I 38:1 “Not seeking is better than seeking.” For the compatibility of τ̄- and τpc-, cf. 359.

STATIVE

162 Some Egyptian Coptic verbs are also actualized in a ‘stative’ form, also called ‘qualitative’, which expresses being-in-a-state. The stative describes the enduring state of a subject in a condition that is associated with verbal action or verbally expressed acquisition of a quality. E.g. cοτπ ‘is chosen’ (is in the enduring state of having been chosen), is exquisite; μoογ† ‘is dead’ (is in the state of having died); ογoβω ‘is white’. In modern dictionary entries, the stative form is identified by a raised dagger: cοτπ†, μoογ†, etc. All statives, except the five listed in 168(c), can contrast in meaning one way or another with the corresponding infinitive; cf. 168(a)(b). E.g.

†-cοτπ†: †-cωτπ	I am exquisite†, chosen†: I choose
†-μoογ†: †-μoγ	I am dead†: I die
†-ογoβω†: †-ογβαω	I am white†: I become white

The formation of the stative is described in 190. The only syntactic function of the stative is predicate in the durative sentence; its presence signals nexus 248.

For a few statives, no corresponding infinitive is attested in Sahidic, e.g. κιωγ† ‘is fat’, ζοογ† ‘is bad’.

Figure 15. Sample from Layton 2004 (p. 126 § 162) showing standard Coptological usage of a SMALL RAISED DAGGER to mark a certain verb form called the “stative” (or “qualitative”), and here also to mark the meanings of such forms.

## A. Administrative

1. Title

**Proposal to encode additional characters for Greek, Latin, and Coptic in the UCS**

2. Requester's name

**Michael Everson, Stephen Emmel, Siegfried Richter, Susana Pedro, António Emiliano**

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

**Individual contribution.**

4. Submission date

**2010-08-05**

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

**Coptic and Supplementary Punctuation**

2. Number of characters in proposal

**6 (2, 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 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. 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.

**No.**

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?

**Copticists, Biblical scholars, 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 Coptic and diacritical marks.**

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

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

**See §3 above.**

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?

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?