1. Introduction

In linguistic texts, there is often the request to mark the reading or interpretation of a specific character (or the sound it denotes, or any other property of it which is discussed in the text) as doubtful, either if the author doubts the reading or interpretation itself, or to mark that the author is uncertain regarding the discussed property (e.g. the pronunciation of a given grapheme). The characters marked doubtful may be letters as well as e.g. metrical symbols.

This is often expressed by a combining question mark. This can be placed above or below the affected letter, depending of typographical considerations (which in turn may depend whether the text within it occurs contains high or low modifier letters to be marked).

Therefore, both versions of the combining question mark (above and below) are proposed here. In theory, they could be considered as glyph variant of the same underlying character. However, there is no precedent of a combining character which has no fixed placement relative to the base letter, and especially there is no combining class indicating such a placement variation. Introducing such a combining class also would mean to extend the combining rules specified in Unicode as such.

Therefore, it is appropriate to simply propose two characters here, with existing and proven combining classes.

In addition, also a free-standing superscript question mark is used.

This character especially is used in critical apparatuses. There, it is used in contrast with the common question mark, to mark different doubts on the reading (see also the detailed explanations in the figure legends of fig.1992a-VIII and 2001a-50):

- The ordinary question mark indicates doubt on the reading on the character, while the fact that the character was corrected in the source is not subject of the doubt.
- The superscript question mark indicates doubt on the properties of the indicated correction, while the reading of the character itself is not subject of the doubt.
Also, the superscript question mark is found within texts or sequences of metrical symbols themselves in contrast to the ordinary question mark, to denote similarly different scopes of doubt.

2. Proposed Characters

Annotations in parentheses address special issues for a character, or reference to figures where such special issues are discussed. (These annotations are not intended to be retained in the character list when copied into the standard.)

**Block: Combining Diacritical Mark Supplement**

**Combining Marks for linguistic use**

- **U+1DF5** COMBINING QUESTION MARK ABOVE
  = combining doubt mark (linguistic and metrical)
  (see fig. 1982a-75, 1989a-35, 1989a-122)

- **U+1DF6** COMBINING QUESTION MARK BELOW
  = alternative combining doubt mark
  (see fig. 1896a-109)

**Block: Superscripts and Subscripts**

- **U+209D** SUPERSCRIPT QUESTION MARK
  = \(<\text{super}>\) 003F
  = doubt mark (linguistic and metrical)
  (see fig. 1989a-122, 1992a-VIII, 1992a-42)

**Properties:**

<table>
<thead>
<tr>
<th>Character</th>
<th>Description</th>
<th>Isolation</th>
<th>Weight</th>
<th>Script</th>
<th>Visibility</th>
<th>Combining</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>U+1DF5</td>
<td>COMBINING QUESTION MARK ABOVE</td>
<td>Mn;230;NSM;;;;;;N;;;;;;;;</td>
<td>Mn;230;NSM;;;;;;N;;;;;;;;</td>
<td>Mn;230;NSM;;;;;;N;;;;;;;;</td>
<td>Mn;230;NSM;;;;;;N;;;;;;;;</td>
<td>Mn;230;NSM;;;;;;N;;;;;;;;</td>
<td>Mn;230;NSM;;;;;;N;;;;;;;;</td>
</tr>
<tr>
<td>U+1DF6</td>
<td>COMBINING QUESTION MARK BELOW</td>
<td>Mn;220;NSM;;;;;;N;;;;;;;;</td>
<td>Mn;220;NSM;;;;;;N;;;;;;;;</td>
<td>Mn;220;NSM;;;;;;N;;;;;;;;</td>
<td>Mn;220;NSM;;;;;;N;;;;;;;;</td>
<td>Mn;220;NSM;;;;;;N;;;;;;;;</td>
<td>Mn;220;NSM;;;;;;N;;;;;;;;</td>
</tr>
<tr>
<td>U+209D</td>
<td>SUPERSCRIPT QUESTION MARK</td>
<td>So;0;ON;&lt;super&gt; 003F;;;;N;;;;;</td>
<td>So;0;ON;&lt;super&gt; 003F;;;;N;;;;;</td>
<td>So;0;ON;&lt;super&gt; 003F;;;;N;;;;;</td>
<td>So;0;ON;&lt;super&gt; 003F;;;;N;;;;;</td>
<td>So;0;ON;&lt;super&gt; 003F;;;;N;;;;;</td>
<td>So;0;ON;&lt;super&gt; 003F;;;;N;;;;;</td>
</tr>
</tbody>
</table>

3. References

- [1896a] Thomsen, Vilh. – Inscriptions de l’Orkhon – Helsinki 1896
4. Examples and Figures

The figures are numbered by the referenced work (consisting of the year of edition and the letter, as in the "references" list, followed by a hyphen the page number, and following by a second letter if more than one figure is taken from a page. E.g.: "Fig. 1896a-109" means "See ref. [1896a], p.109"). References to already encoded characters are usually given in parentheses.

---

Fig. 1896a-109:  Showing COMBINING QUESTION MARK BELOW below ordinary letters (purple arrow) and modifier letters (green arrows)-

---


---

Proposal to encode Linguistic Doubt Marks
2011-10-21
It is interesting to note a certain parallelism between the development of the strophes in O. 12 (466?) and P. 4 (452): 

O. 12 \[ e-D || \quad ? \quad e-D \mid -d \mid \mid E^2 \mid E-d \mid \]
\[ e-D \quad -E \mid \]
\[ -D \quad -E \mid \mid \]

P. 4 \[ e-D || \]
\[ e-D: \quad \quad \quad \quad e-D \mid \]
\[ e-D \quad -E \mid \]
\[ D^2 \times e- \mid D \quad -E \mid \]
\[ E \quad -D^2 \mid E^2e \mid E \quad -\mid \mid . \]

Fig. 1982a-75: Showing a specimen for COMBINING QUESTION MARK ABOVE applied to a metrical pause symbol.

Fig. 1982a-102: Showing a specimen for COMBINING QUESTION MARK ABOVE applied to a metrical symbol.

Fig. 1989a-35: Showing a specimen for COMBINING QUESTION MARK ABOVE applied to a metrical symbol.

\[ \text{VIIa} = \text{fr. 52g (A)} \]

\[ -6 \quad \mathcal{O} \quad [ \]
\[ -4 \quad \mathcal{O} \quad [ \]
\[ -3 \quad \mathcal{O} \quad [ ? ] \]

Proposal to encode Linguistic Doubt Marks
2011-10-21
Fig. 1989a-122: Showing specimens for COMBINING QUESTION MARK ABOVE (red) and SUPERSCRIPT QUESTION MARK (green), in contrast to a common question mark (blue).
Fig. 1992a-VIII: Showing specimens for SUPERSCRIPT QUESTION MARK (red), together with an ordinary question mark, explaining the different meaning of those characters when used in a critical apparatus. Right, an enlarged excerpt containing the question marks is shown. The abbreviations read:

$A^? – A$, fortasse ante correctionem – manuscript A, the indicated reading is presumed to be the one before the correction (i.e. it is doubted that the otherwise undoubted identity of the character is in fact the one before the correction)

$A^? – fortasse A$, ante correctionem – the presumed reading in manuscript A, as it was before the correction (i.e. the identity of the corrected character itself is doubted)

VIII

SIGLA CETERA

Ω

Σ

Φ

Θ

Τ

$A^1$, $A^2$

$A^a$, $A^c$

$A^?, A^?$

$A^2$, $A^3$

$A^x$

$A^a$

$A^y$

$A^m$

$A^n$

$A^r$

$A^s$

$A^{as}$

$A^l$

$A^{aw}$

lectio omnium codicum vel utique archetypi

scholium vetus (Σ$^a$ scholii lemma; Σ$^i$ scholium in cod. l)

scholium incertae aetatis

commentarius paraphrasticus in triadem, saec. f ere xii

Thomae Magistri textus

Triclinius in scholio suo

A a prima /ab altera manu correctus

A ante/post correctionem

A, fort. a.c./fort. A a.c.

A ante correctionem ab altera manu/a scholiorum scriba factam

A ante vel post correctionem (incertum utrum)

glossema in A, vel lectio quasi glossema adscripta

varia lectio cum ψ in A adscripta

A in margine

A in rasura

A tum correctus cum scholia addita sunt

lectio in A super lineam scripta

A in textu, altera lectione inter lineas vel in margine adscripta

A ut videtur

A$^a$, A$^a$
719 ἐκφύγωις Μ: -οι fere cett.  720 ἐπφύζα Βνο, πέφυζα Η
ωλεσί- ΜΒΟ λ: ούλεσί- Η: όλεσι- cett.  721 post h.v. sch. ἀ
γάρ νύκτωρ παρεχελεύσατο καὶ γέγονε in textu habent codd.
946)  725 βλαψίφρονος (-ας ΗColumnInfo ss-.+ ) οἰδιπόδα (-ποδός V,
-pόδαο HB�能 Υ) Ω: corr. Τ  726 δ’ om. Π’V  727 κλήρους
Μλ Θξ: -οις cett.  728 σκυθών Ω: corr. Dind.12  732 καὶ] ἀν
καὶ Wε Qε λ φημένοις Ω (-σιν Τ): φημένους (Bourdolot)
<ἀν> Blomf.4, <ἐγ> κατέχειν Headlam5  57  734 αὐτοκτόνως
καὶ: αὐτόκτονοι Flb: αὐτοκτόνωσιν Μς: αὐτοκτάνωσιν Μς:1
αὐτοὶ κτάνωσι(v) cett.  736 γαῖα Dind.4 640: χθονία Ω: νεφ
τέφα Weil7: καγχχωφία Newman1 56

Proposal to encode Linguistic Doubt Marks
2011-10-21
Proposal to encode Linguistic Doubt Marks

2011-10-21

Fig. 1998a-156/157: Showing SUPERSCRIPT QUESTION MARK (red) in a critical apparatus (lower picture from p.157), in contrast to an ordinary question mark in the same apparatus (blue; upper picture from p.156). The third picture shows the same SUPERSCRIPT QUESTION MARK from p.157 by a higher resolution.

415 post 414 ferunt 1 ΖΩ, ante eum 400: post 412 trai. Cobet Misc. Crit. 369: noverat Zen quoquo loco: del. Christ 416 ἰχῶ sch 1 t B E T W: ἰχῶ R W G:, -ὁρ'Ζ F= T, -ὁρ C D F=G C ζευγὸς Ar 1 ΖΩ: ζευγὸν Zen (cf. Σ 123) 417 χείρ ττ Ω: χείρ D, χείρ' T R= 422 ἀνείσια 1:cf. ad Ι 700 423 άμα σπείραθι Plut. ΖΩ: ἀμ' ἐσπεριθαὶ 1 (αμ' ἐσπεριθαὶ) ττ= cf. Praef. ΗΠ 1 ἐπικαγιλ' ἐφιλῆς 1 400 ττ Ω: ἐκπαλικ' ἐφιλῆς 1 (cf. Ατ 415 424 εὐπέπλον 400 D Τ R, εὐπ- 1 416 ττ Ω: βαρ- θυκάλπον Plut. 425 χρυσῆι 400: χρυσῆι 1 ττ Ζ (σῆ) Ω: καταμύζατο Ar C T R=; κατεμ- (nov. Did) 1 400 ττ Ω: ἀραιήν Ar Hdn Ζ B F: ἀρ- 1 ττ Ω: βαρεάθειν W 427 χρυσῆι 1 416 416 4116 428 θεῖαν W: θεῖαν Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖαν 1 400 1017 ττ Ω: θεῖα

Ar C T R?
Proposal to encode Linguistic Doubt Marks

2011-10-21

Fig. 2001a-50: Showing SUPERSCRIPT QUESTION MARK (red).

Here (for Bacchylides 17: carmen [canto] 3, line 9 to 14 shown in the excerpt), a manuscript “A” is considered which had been corrected by two scribes.

Thus, the critical apparatus remark for line 14 following the colon reads: φαρθι in manuscript A, »²« ab altera manu correctus (corrected by second hand) »?« fortasse (perhaps) »ac« ante correctionem (before correction): φαρθιν

– i.e.: there is a correction to φαρθι in the manuscript A, possibly by the second scribe, but it is doubtful that this in fact was done by the second scribe.

θρόνιος δὲ λ[αὸς
col. vii òc, παρὰ Ζηνὸς λαχῶν
πλείσταρχον Ἑλλάνων γέρας,
oiδε πυργωθέντα πλοῦτον μὴ μελαμ-
φαρεῖ κρύπτειν εκότωι.

17. inscr. (in marg.) Ίέρων Συρακοσίων ἦπερος [Ὁλύμπια] 1 (κάρη) 2 (ίο-
ποιος) 3 (κό) Κλεοί Blasi: Κλεοί A (āc) 4. Ὀ.: prima syllaba
versus quarti strophae antistrophaeque in papyro semper tertium potius uersum
terminat (δρό) (ἰε) 5 ceou]tò e.p. (Kenyon), φέρου]tò Platt, al., ìev]tò
Edmonds, quorum uerborum nullum spatio bene conuenit (ὁχ, acc. tum deleto,
ut uid.) (καί) Νίκας scrivunt 6 Ἀγαλαῖai scribunt (ἰαί) (δύν) 7 τόδη.
Palmer (μεν) 8 τ[έκος Edmonds, στεφάνω]v e.p. (caut) 9 [ ἀπείρων |
Blass 10 (ā) [ἀνήρ | e.p. 11 (ά) 12 (πλε) γέρας Α[με] (corr.
i.).] γένος Α[με] 13 μελαμ- Α[με] (λῃ in am corr.): μελὴ Α[με] 14 φαρεῖ e.p.,
tamquam lectionem Α[με]: φαρθι, Α[με], φαρθιν Α[με] (των)}
A. Administrative

1. Title: Proposal to encode Linguistic Doubt Marks in the UCS

2. Requester's name: Martin Schrage; Karl Pentzlin

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

4. Submission date: 2011-10-21

5. Requester's reference (if applicable): University of Munich, Germany (M. S.)

6. Choose one of the following:
   - This is a complete proposal: Yes
   - (or) More information will be provided later: No

B. Technical – General

1. Choose one of the following:
   - a. This proposal is for a new script (set of characters): No
   - b. The proposal is for addition of character(s) to an existing block: Yes
      - Name of the existing block: Combining Diacritical Marks Supplement; Superscripts and Subscripts

2. Number of characters in proposal: 3

3. Proposed category (select one from below - see section 2.2 of P&P document):
   - A-Contemporary
   - B.1-Specialized (small collection) [X]
   - 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

4. Is a repertoire including character names provided? Yes
   - a. If YES, are the names in accordance with the “character naming guidelines” in Annex L of P&D document? Yes
   - b. Are the character shapes attached in a legible form suitable for review? Yes

5. Fonts related:
   - a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard? The authors (if requested)
   - b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.): The authors (if requested)

6. References:
   - a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? Yes
   - b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? Yes

7. Special encoding issues:
   - 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)? No

8. Additional Information:

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 http://www.unicode.org/Public/UNIDATA/UCD.html and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.
## C. Technical - Justification

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Has this proposal for addition of character(s) been submitted before?</td>
<td>Yes</td>
</tr>
<tr>
<td>If YES explain: <em>They are contained in WG2 N3913 = L2/10-358R and are separated here from its revision</em></td>
<td></td>
</tr>
<tr>
<td>2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?</td>
<td>Yes</td>
</tr>
<tr>
<td>If YES, with whom? <em>One of the authors (M. S.) is a member of the scientific community himself</em></td>
<td></td>
</tr>
<tr>
<td>If YES, available relevant documents: <em>See text</em></td>
<td></td>
</tr>
<tr>
<td>3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?</td>
<td>Yes</td>
</tr>
<tr>
<td>Reference: <em>See text</em></td>
<td></td>
</tr>
<tr>
<td>4. The context of use for the proposed characters (type of use; common or rare)</td>
<td>Common scientific</td>
</tr>
<tr>
<td>Reference: <em>See text</em></td>
<td></td>
</tr>
<tr>
<td>5. Are the proposed characters in current use by the user community?</td>
<td>Yes</td>
</tr>
<tr>
<td>If YES, where? Reference: <em>See text</em></td>
<td></td>
</tr>
<tr>
<td>6. After giving due considerations to the principles in the P&amp;P document must the proposed characters be entirely in the BMP?</td>
<td>Yes</td>
</tr>
<tr>
<td>If YES, is a rationale provided?</td>
<td>Yes</td>
</tr>
<tr>
<td>If YES, reference: <em>To keep them in line with related characters</em></td>
<td></td>
</tr>
<tr>
<td>7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?</td>
<td>Yes</td>
</tr>
<tr>
<td>8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?</td>
<td>No</td>
</tr>
<tr>
<td>If YES, is a rationale for its inclusion provided?</td>
<td>Yes</td>
</tr>
<tr>
<td>If YES, reference:</td>
<td></td>
</tr>
<tr>
<td>9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?</td>
<td>No</td>
</tr>
<tr>
<td>If YES, is a rationale for its inclusion provided?</td>
<td>Yes</td>
</tr>
<tr>
<td>If YES, reference: <em>See text</em></td>
<td></td>
</tr>
<tr>
<td>10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?</td>
<td>Yes</td>
</tr>
<tr>
<td>If YES, is a rationale for its inclusion provided?</td>
<td>Yes</td>
</tr>
<tr>
<td>If YES, reference: <em>See text</em></td>
<td></td>
</tr>
<tr>
<td>11. Does the proposal include use of combining characters and/or use of composite sequences?</td>
<td>Yes</td>
</tr>
<tr>
<td>If YES, is a rationale for such use provided?</td>
<td>Yes</td>
</tr>
<tr>
<td>If YES, reference: <em>The proposal contains combining characters but no composite sequences</em></td>
<td></td>
</tr>
<tr>
<td>12. Does the proposal contain characters with any special properties such as control function or similar semantics?</td>
<td>No</td>
</tr>
<tr>
<td>If YES, describe in detail (include attachment if necessary)</td>
<td></td>
</tr>
<tr>
<td>13. Does the proposal contain any Ideographic compatibility character(s)?</td>
<td>Yes</td>
</tr>
<tr>
<td>If YES, is the equivalent corresponding unified ideographic character(s) identified?</td>
<td>Yes</td>
</tr>
<tr>
<td>If YES, reference:</td>
<td></td>
</tr>
</tbody>
</table>