

# UC Irvine

## Unicode Project

### Title

A proposal to encode Greek editorial and punctuation characters in the UCS

### Permalink

<https://escholarship.org/uc/item/3bn592k6>

### Authors

Pantelia, Maria C.  
Peevers, Richard

### Publication Date

2004-08-23

**ISO/IEC JTC 1/SC 2/WG 2  
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS  
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646<sup>1</sup>**

**Please fill all the sections A, B and C below.**

**(Please read Principles and Procedures Document for guidelines and details before filling this form.)**

See <http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html> for latest Form.

See <http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for latest Principles and Procedures document.

See <http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest roadmaps.

**A. Administrative**

1. **Title: Proposal to encode additional Greek editorial and punctuation characters in the UCS**
2. Requester's name: **Thesaurus Linguae Graecae, (University of California, Irvine)**
3. Requester type (Member body/Liaison/Individual contribution): **Expert contribution**
4. Submission date: **Oct. 6, 2003**
5. Requester's reference (if applicable):
6. (Choose one of the following:)
- This is a complete proposal: **YES**
- or, More information will be provided later:

**B. Technical - General**

1. (Choose one of the following:)
- a. This proposal is for a new script (set of characters): \_\_\_\_\_  
Proposed name of script: \_\_\_\_\_
- b. The proposal is for addition of character(s) to an existing block: **YES**  
Name of the existing block: **1 In Combining Diacritical Marks (035C); 2 In Combining Diacritical Marks Supplemental (1DC0-1DC1); 2 In General Punctuation (2059, 205D); 9 In Supplemental Punctuation (2E0E-2E16); 3 In Greek And Coptic (03FC-03FE)**
2. Number of characters in proposal: **17**
3. Proposed category (see section II, Character Categories): **C**
4. Proposed Level of Implementation (1, 2 or 3) (see clause 14, ISO/IEC 10646-1: 2000):  
**Level 3 for Diacritical Marks; Level 1 for other characters**
- Is a rationale provided for the choice? **YES**
- If Yes, reference: **Contains combining marks**
5. 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 ISO/IEC 10646-1: 2000? **Yes**
- b. Are the character shapes attached in a legible form suitable for review? **Yes**
6. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard? **David Perry and TLG Project (True Type)**  
If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used: **TLG Project, mcpantel@uci.edu**
7. 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**
8. 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**
9. 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/UnicodeCharacterDatabase.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

<sup>1</sup> Form number: N2352-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09)

**C. Technical - Justification**

1. Has this proposal for addition of character(s) been submitted before? If YES explain _____	<b>No</b>
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.)? If YES, with whom? <b>Proposal has been reviewed by Professors Roger Bagnall, Columbia University, John Oates, Duke University, William Johnson, University of Cincinnati, Michael Haslam, University of California, Los Angeles. Earlier versions of this proposals have been posted online and received comments by members of the profession.</b>	_____
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference: _____	<b>Scholarly community</b>
4. The context of use for the proposed characters (type of use; common or rare) Reference: _____	<b>Use varies</b>
5. Are the proposed characters in current use by the user community? <b>Yes. Characters are present in various editions of Greek texts and used extensively by scholars of Greek. General references provided in attached bibliography.</b> If YES, where? Reference: _____	
6. After giving due considerations to the principles in <i>Principles and Procedures document</i> (a WG 2 standing document) must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If YES, reference: _____	<b>Yes</b> <b>Accordance with the Roadmap.</b>
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	<b>Preferably</b>
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If YES, reference: _____	<b>No</b>
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference: _____	<b>No</b>
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? <b>Two characters are similar but not identical to 1FED and IFEE</b> If YES, is a rationale for its inclusion provided? If YES, reference: _____	
11. 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)? If YES, is a rationale for such use provided? If YES, reference: _____ Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? If YES, reference: _____	<b>Yes, characters 1-3</b> <b>Are combining marks</b> <b>No</b>
12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary) _____	<b>Greek Paragraphos. See footnote 3 in the proposal</b>
13. Does the proposal contain any Ideographic compatibility character(s)? If YES, is the equivalent corresponding unified ideographic character(s) identified? _____ If YES, reference: _____	<b>No</b>

## Proposal

Ancient Greek scribes generally wrote in continuous uppercase letters without separating letters into words. On occasion the scribe added in punctuation to indicate the end of a sentence, change of speaker or, that two words which could be read as one word, were in fact separate words. Editorial and punctuation characters appear abundantly in surviving papyri and when possible have been rendered in modern typography.

When punctuation is present, it falls into one of two kinds: (1) marginal or semi-marginal characters which mark the end of a section of text (e.g. Coronis, Paragraphos); (2) characters mixed in with the text to mark pauses, end of sense, or separation between words (e.g. Stigme, Hypodiatole).

A number of editorial characters are attributed to and named after Aristarchos of Samothrace (c. 216-144 BCE),<sup>2</sup> fifth head of the Library at Alexandria. Aristarchos is known to have provided a major edition of the works of Homer, which forms the basis of our modern editions. For a complete list of the extant papyri containing these characters, see McNamee (1992) 28-9.

## Character Properties

The additions to Combining Diacritical Marks are “Mn”.

The additions to Combining Diacritical Marks Supplemental are “Mn”.

The additions to General Punctuation are “Po”.

The additions to Supplemental Punctuation are “Po”.

The additions to Greek And Coptic are “So”.

## Bibliography

Grenfell, B.P. & Hunt, A.S. et al, *The Oxyrhynchus Papyri*, multiple volumes (London, 1898 - onwards)

Kenyon F.G. & Bell H.I., *Greek Papyri in the British Museum III (Nos. 485—1331)* (London, 1907)

Kenyon, F.G., *The Palaeography of Greek Papyri*. (Oxford, 1899)

Kenyon, F.G., “Abbreviations and Symbols in Greek Papyri” in Oikonomides, A. N. (ed), *Abbreviations in Greek Inscriptions: Papyri, Manuscripts and Early Printed Books* (Chicago, 1974)

Lockwood, J.F., Browning, R. Wilson, N.G. “Aristarchus” in *OCD*<sup>3</sup> (1996) 159

Lockwood, J.F., Browning, R. Wilson, N.G. “Zenodotus” in *OCD*<sup>3</sup> (1996) 1653-6

McNamee, K., *Sigla and Select Marginalia in Greek Literary Papyri*. (Papyrologica Bruxellensia 26) (Brussels, 1992)

*Thesaurus Linguae Graecae, Beta Code Manual* (University of California, Irvine, 2002)

Turner, E.G. with Parsons, P.J., *Greek Manuscripts of the Ancient World* 2<sup>nd</sup> ed. (London, 1987)

Wace, A.J.B. & Stubbings, F.H., *A Companion to Homer* (London, 1962)

Wattenbach, W., *Anleitung zur Griechischen Palaeographie*. 3<sup>rd</sup> ed. (Leipzig, 1987)




Zimmermann, F., *Griechische Roman-Papyri und verwandte Texte* (Heidelberg, 1936)

---












<sup>2</sup> Lockwood, J.F., Browning, R. Wilson, N.G. “Aristarchus” in *OCD*<sup>3</sup> (1996) 159

## Ancient Greek Combining Diacritical Marks, General Punctuation and Additions to the Greek and Coptic Block

### Combining Diacritical Marks

	Name	Unicode	Comments
<b>Amended Accents</b>			
1		Combining Dotted Grave Accent	Similar, but not identical to 1FEE
2		Combining Dotted Acute Accent	Similar, but not identical to 1FED
<b>Word Joiner</b>			
3		Combining Double Breve Below	· Greek Papyrological Hyphen

### Supplemental Punctuation

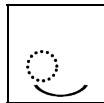
<b>Section and Pause Markers, Word Separators</b>			
4		Editorial Coronis	Note: name ‘Editorial Coronis’ used to distinguish character from 1FBD (Greek Koronis)  Many glyph variants. Three basic forms:  ,  and  .
5		Paragraphos <sup>3</sup>	
6		Forked Paragraphos	Glyph variant: 
7		Reversed Forked Paragraphos	Glyph variant: 
8	⋮	Tricolon	Similar to 22EE, but with different properties.
9	⋮⋮	Quintuple Colon	=Greek Pentoklion. Unifiable with (Glagolitic block)
10	⋸	Hypodiastole	Not the same as 002C (see example below)
<b>Text Highlighters and Markers</b>			
11	Ϸ	Reversed Greek Capital Lunate Sigma Symbol	= Greek Antisigma
12	Ϸ̇	Dotted Obelos	Many glyph variants, including Ϸ̇, Ϸ̇, Ϸ̇. See McNamee 37 for a complete list.
13	↴	Downwards Ancora	
14	↵	Upwards Ancora	
<b>Aristarchean Editorial Notation</b>			
15	Ϸ̇	Dotted Right-Pointing Angle	= Diple Periestigmene Glyph variants:  and 
16	Ϸ̇	Dotted Greek Capital Lunate Sigma Symbol	= Sigma Periestigmenon
17	Ϸ̇	Reversed Dotted Greek Capital Lunate Sigma Symbol	= Antisigma Periestigmenon

<sup>3</sup> The Greek paragraphos is placed at the beginning of the line but it may refer to a break in the text at any point in the line. Therefore, it should not be treated as a combining character--such as an underscore--and placed under the first character in the line. The Paragraphos should rather be a horizontal line generally stretching under the first one and a half letters of the line it refers to, and possibly extending into the margin. It should be given a no-space line of its own.

**TABLE 035C: COMBINING DIACRITICAL MARKS**

035

C



**TABLE 035C: COMBINING DIACRITICAL MARKS**

hex	Name
035C	COMBINING DOUBLE BREVE BELOW = GREEK PAPHROLOGICAL HYPHEN

**TABLE 1DC0-1CD1: COMBINING DIACRITICAL MARKS SUPPLEMENTAL**

	1DC
0	◌̈́
1	◌̈́̂

**TABLE 1DC0-1CD1: COMBINING DIACRITICAL MARKS SUPPLEMENTAL**

hex	Name
1DC0	COMBINING DOTTED GRAVE ACCENT
1DC1	COMBINING DOTTED ACUTE ACCENT

**TABLE 2059, 205D: GENERAL PUNCTUATION**










205	⋮
9	⋮
D	⋮

**TABLE 2059, 205D: GENERAL PUNCTUATION**

hex	Name
2059	FIVE DOT PUNCTUATION = PENTONKION
205D	TRICOLON



**TABLE 2E0E-2E16: SUPPLEMENTAL PUNCTUATION**

	2E0	2E1
0		
1		
2		
3		
4		
5		
6		
7		
8		
9		
A		
B		
C		
D		
E		
F		

**TABLE 2E0E-2E16: SUPPLEMENTAL PUNCTUATION**

hex	Name
2E0E	EDITORIAL CORONIS
2E0F	PARAGRAPHOS
2E10	FORKED PARAGRAPHOS
2E11	REVERSED FORKED PARAGRAPHOS
2E12	HYPODIASTOLE
2E13	DOTTED OBELOS
2E14	DOWNWARD ANCORA
2E15	UPWARDS ANCORA
2E16	DOTTED RIGHT-POINTING ANGLE = DIPLE PERIESTIGMENE

**TABLE 03FC-03FE: GREEK AND COPTIC**

03F

D	Ϸ
E	Ϸ̣
F	Ϸ̣̅

**TABLE 03FC-03FE: GREEK AND COPTIC**

hex	Name
03FD	REVERSED GREEK CAPITAL LUNATE SIGMA SYMBOL = ANTISIGMA
03FE	DOTTED GREEK CAPITAL LUNATE SIGMA SYMBOL = SIGMA PERIESTIGMENON
03DF	REVERSED DOTTED GREEK CAPITAL LUNATE SIGMA SYMBOL = ANTISIGMA PERIESTIGMENON

