L2/05-323

20 October 2005

Lisa Moore
c/o The Unicode Consortium
P.O. Box 391476
Mountain View, CA 94039-1476

Dear colleagues,

This is to apprise you of our interest at Dumbarton Oaks in developing a proposal to the Unicode Consortium covering characters used in Byzantine Greek inscriptions.

"Byzantine Greek inscriptions" may need clarification. *Byzantine* here refers to the eastern Roman Empire (later called the Byzantine Empire, but only a century after its demise), from the fourth through the fifteenth century. Although a number of languages were spoken in the boundaries of the Empire at various times, *Greek* was the principal language. *Inscriptions* refers to texts that are inscribed into or painted on solid objects such as stone, mosaic, jewelry, amulets, coins, seals, weights, bricks, ivory, and steatites. These inscriptions are the main subjects of Byzantine epigraphy. Most of our publications have focused on coins and seals, so my comments below generally draw from these two subdisciplines.

The current version of Unicode (4.1) already provides for most of the glyphs used in Byzantine epigraphy. The blocks Greek and Coptic, Greek Extended, Ancient Greek Musical Notation, and Ancient Greek Numbers were built primarily with manuscript text traditions in mind, but there are a number of characters in these blocks that are known mostly through inscriptions. Many of these, however, since they represent archaic Greek symbols, are inadequate for Byzantine inscriptions, which have characters that currently have no representation in the Unicode standard. The characters can be divided into four different categories: letters, accents, punctuation, and decorations.

LETTERS

Letters may be subdivided into single letters and ligatures.

Single letters. Unicode provides adequately for the basic Greek alphabet. Seals and coins often carry unusual letterforms that change across the centuries and help scholars date the material. For example, the letter alpha (U+0391) can appear in a number of forms: ΑΑΔ∂Α. The letter may appear upside down, turned, or flipped. Semantically, all of these variants mean exactly the same thing, so it is understandable that these would not qualify for special treatment in the Unicode standard.

It is worth noting, however, that individual letters may appear in combining forms, set above a base letter. In an inscription we are to publish soon, a theta lays above an eta, shown below, in the word ἐκοιμήθη. The abbreviation for *monk*, μοναχός, is oftentimes rendered Αͯ. U+0363..U+036F has been reserved for a few combining Latin letters. Would this epigraphic habit justify a new block of combining Greek letters? There are a number of these combining letterforms, and in future communications we can provide more specific information.

Θ
Η

Ligatures. There are hundreds of Byzantine ligatures, some more common than others. If not candidates for inclusion in the Unicode standard, these ligatures may be ideal for Technical Notes, to indicate what letters may often combine with which. There tend to be fewer ligatures in Byzantine

inscriptions than in Byzantine manuscripts. Here are a few examples of Byzantine ligatures in inscriptions. The first four examples would be ideal for Notes; the last two, for consideration as new Unicode codepoints.

1 The letters mu, nu, eta, and nu sometimes forms the ligature ΜΝΗΝ.

2 The letters omicron and upsilon sometimes forms the ligatures ȣ or Ȣ. The particular form of the ligature oftentimes helps date the object.

3 The letters sigma and tau sometimes form the ligature ϛ.

4 The genitive masculine Greek definite article, tau plus omicron plus upsilon, is sometimes rendered ȣ.

5 The letters alpha and omicron can be juxtaposed to represent the title "saint" (ὁ ἅγιος): Ⓐ or Ⓐ. In this case, the symbol could not be represented by U+0391 and U+20DD since the bottom of the alpha should be set above the baseline, and its top should be set below the ascender line. U+24B6 may appear a convenient solution, but this codepoint is assigned to a character with a semantic function that differs from this abbreviation.

6 The term for *indiction* (the fifteen-year tax season) may be abbreviated ϯ.

ACCENTS

Letters are given accents in Byzantine Greek inscriptions much the same way they are in other texts. The appearance and placement of the accent often differs, but it seems that all the necessary codepoints are already in the Unicode standard.

PUNCTUATION

Punctuation marks that govern sentence structure—the semicolon, the comma, and so forth—are already adequately represented in the Unicode standard. What is especially lacking are the abbreviation marks that indicate elided letters. There are several types of these abbreviation marks, and their distinct shapes help scholars date the material. Abbreviation marks may be spacing or combining characters. They may be attached or unattached to the letter they govern. They may appear above a letter, above and to the right, or below and to the right. The glyphs below this paragraph show some of these types. A dotted circle indicates a base character with which the abbreviation mark combines. A solid circle indicates a character immediately preceding (but not combining with) an abbreviation mark.

◌⸍ ◌ϛ ◌ʾ ◌̄

○⸍ ○⊢ ○ʾ ○⁻ ○⁻ ○⸝

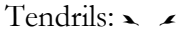
For example:

The prayer Θεοτόκε βοήθει ("Mother of God, help!") is often rendered: Θ̄K̄Ē R,Θ, Note, the overbar over three letters is an abbreviation, as are the two comma-looking marks. The capital R is a variant Byzantine letterform for a capital letter beta. In another example, a seal preserves the name of the owner, said to be Nicholas the imperial kommerkiarios, as Νικόλαος βασιλικός, rendered: ΝΙΚΟΛ́R·. Note, here two different abbreviation marks are used.

DECORATIONS

Decorations are common in inscriptions, and they often fall into certain types or categories, three of which are presented below. Tendrils mark the beginning or end of text. The globus cruciger is a decoration, but also appears frequently in coin and seal publications as its own character. Signa and

bullets/bullet variations appear on coins and represent control marks over particular batches of dies or workshop identifications. For instance:

Tendrils: 

Globus crucigers: 

Signa: ∴ ⊙ 

The glyph ∴ looks like U+2056, ∵, which also appears on coins, but its orientation is distinct. Other variations also appear: ∴ ∵. We can work on developing explanations and typologies for these various signa.

Naturally, there is more to say about each of three types of most concern to the Unicode Consortium: letters, abbreviations, and decorations. But hopefully this is enough to begin discussions on what kinds of characters would be suitable candidates for developing Unicode proposals. I look forward to your reply.

Sincerely,

Joel Kalvesmaki
Assistant Editor
Dumbarton Oaks
1703 32nd St. NW
Washington, DC 20007
(202) 777-0095
kalvesmakij@doaks.org

Developed with the assistance of John Nesbitt and Cécile Morrisson