Doc Type: Working Group Document
Title: Proposal to add two Greek letters for Bactrian to the UCS
Source: Nicholas Sims-Williams and Michael Everson
Status: Expert Contribution
Date: 2002-01-30

A. Administrative

1. Title
Proposal to add two Greek letters for Bactrian to the UCS.

2. Requester's name
Nicholas Sims-Williams and Michael Everson.

3. Requester type
Expert contribution.

4. Submission date
2002-01-30

5. Requester’s reference

6a. Completion
This is a complete proposal.

6b. More information to be provided?
No.

B. Technical -- General

1a. New script? Name?
No.

1b. Addition of characters to existing block? Name?
Yes. Greek (proposed code position: U+03F7, U+03F8)

2. Number of characters
2.

3. Proposed category
Category A.

4. Proposed level of implementation and rationale
Level 1. Base characters with no diacritics.

5a. Character names included in proposal?
Yes.

5b. Character names in accordance with guidelines?
Yes.

5c. Character shapes reviewable?
Yes.

6a. Who will provide computerized font?
Michael Everson, Everson Typography.
6b. Font currently available?  
Yes.
6c. Font format?  
TrueType.
7a. Are references (to other character sets, dictionaries, descriptive texts, etc.) provided?  
Yes (see below).
7b. Are published examples (such as samples from newspapers, magazines, or other sources) of use of proposed characters attached?  
No.
8. Does the proposal address other aspects of character data processing?  
No.

C. Technical -- Justification
1. Contact with the user community?  
Yes. Nicholas Sims-Williams is the world’s leading expert in Bactrian.
2. Information on the user community?  
Indo-Iranian and Indo-European linguists.
3a. The context of use for the proposed characters?  
Supplementary Greek letter used to write Bactrian.
3b. Reference  
See below.
4a. Proposed characters in current use?  
Yes.
4b. Where?  
By scholars.
5a. Characters should be encoded entirely in BMP?  
Yes.
5b. Rationale  
Keeping them with other Greek characters
6. Should characters be kept in a continuous range?  
No.
7a. Can the characters be considered a presentation form of an existing character or character sequence?  
No. It is similar, but not identical, to LATIN LETTER THORN
7b. Where?  
7c. Reference
8a. Can any of the characters be considered to be similar (in appearance or function) to an existing character?  
No.
8b. Where?  
8c. Reference
9a. Combining characters or use of composite sequences included?  
No.
9b. List of composite sequences and their corresponding glyph images provided?  
No.
10. Characters with any special properties such as control function, etc. included?  
No.
D. Proposal

Bactrian, the ancient language of Bactria in northern Afghanistan, is unique among the Iranian languages in being written in the Greek script – a legacy of the conquest of Bactria by Alexander the Great in the 4th century BCE. From this period onwards the Greek language, written in the Greek script, was for a long time the exclusive language of culture and administration in Bactria. When Bactria was overrun by nomadic peoples from the north, its new rulers, the Kushans, at first continued the use of the Greek language for administrative purposes, but soon they came to use the Greek script to write the local language, Bactrian. A crucial moment in the history of this language was the decision of the Kushan ruler Kanishka to adopt Bactrian as the language of his coinage. After the first issues of Kanishka, Greek disappears from the coinage completely, replaced by Bactrian. During the first centuries CE, Bactrian could legitimately have been ranked amongst the world’s most important languages. As the language of the Kushan kings, Bactrian must have been widely known throughout a great empire, in Afghanistan, Northern India and part of Central Asia. Even after the collapse of the Kushan empire, Bactrian continued in use for at least six centuries, as is shown by the ninth-century inscriptions from the Tochi valley in Pakistan and the remnants of Buddhist and Manichean manuscripts found as far away as the Turfan oasis in western China. The career of Bactrian as a language of culture thus lasted for close to a thousand years.

Bactrian was deciphered in 1957, by W. B. Henning, after the discovery at Surkh Kotal near Baghlan of the first substantial Bactrian inscription. Previously, all that had been known of Bactrian was gleaned from the legends on the coins of the Kushans and their successors. The text Henning translated refers to the foundation of a sanctuary by the emperor Kanishka, its abandonment as a result of problems with the water-supply, and its re-establishment by a high official named Nukunzuk in the year 31 of the era of Kanishka, that is, early in the reign of his successor Huvishka.

Bactrian had a sibilant, most likely pronounced [ʃ] as in Kanishka, for which there was no letter in Greek. This letter is written, in Bactrian manuscript, very much as PHI and RHO are, with a long descender, a bowl on the right, and a tall ascender; it thus looks either like a PHI with only one bowl, or a RHO with an ascender. Bactrian manuscripts often give a unique swash terminal to PHI, RHO, and SHO alike, as can be seen in the figures below. SHO bears some superficial resemblance to LATIN LETTER THORN, and as a matter of convenience some scholars have used that letter to write it recently; but on grounds of multiscript ordering and typographic harmony, it is essential that it be encoded as a Greek letter.

Name

No traditional name is attested for this letter, but because of its similarity to RHO, the name SHO has been suggested here.

Glyphs

Since this letter passed out of use before the advent of Greek typography, it is useful to discuss the nature of its typographic design. A number of glyphs were proposed:
The capital, which is simply half a capital \( \Phi \), seems unobjectionable. The first five of the small letters were made out of the Times Greek font shipped with some versions of the Mac OS. The first of these \( \beta \) is taken from the small looped \( \Phi \) without a top stem; the stem is the same width as the descender of that same \( \Phi \). It is flimsy and unattractive. The second \( \beta \) (which we prefer) has the same ascender and descender but has the bowl of the small \( \rho \), as do the third, fourth, and fifth. The ascender of the third \( \beta \) is taken from that of the small \( \eta \). The fourth \( \beta \) adds the ascender to the full \( \rho \), but looks unbalanced. The fifth \( \xi \) is a whimsical fabrication.

The next four are taken from a new font, Evertimes, which Michael Everson has been developing. The first of these \( \beta \) is taken again from the small looped \( \Phi \) without a top stem; the stem is the same width as the descender of that same \( \Phi \). It too looks too thin in text. The remaining ones are based from the small \( \rho \), which has an ascender which swings slightly to the left. The second \( \beta \) is the \( \rho \)’s own descender flipped up, so that the ascender also swings to the left; the ascender of the third \( \beta \) is taken from that of the small \( \eta \). The fourth \( \beta \) (which we prefer) is the same, but with the ascender reversed so that it swings to the right.

Below is a sample in Bactrian (‘Dathsho-marego the eunuch’) written with the different choices in twelve-point type.

### Times Greek

\[
\Delta\rho\theta\theta\mu\alpha\rho\rho\gamma\omega \ i \ \varepsilon \beta \alpha \zeta \alpha \tau \alpha \nu \\
\Delta\rho\theta\theta\mu\alpha\rho\rho\gamma\omega \ i \ \varepsilon \beta \alpha \zeta \alpha \tau \alpha \nu \\
\Delta\rho\theta\theta\mu\alpha\rho\rho\gamma\omega \ i \ \varepsilon \beta \alpha \zeta \alpha \tau \alpha \nu \\
\Delta\rho\theta\theta\mu\alpha\rho\rho\gamma\omega \ i \ \varepsilon \beta \alpha \zeta \alpha \tau \alpha \nu
\]

### Evertimes

\[
\Delta\rho\theta\theta\mu\alpha\rho\rho\gamma\omega \ i \ \varepsilon \beta \alpha \zeta \alpha \tau \alpha \nu \\
\Delta\rho\theta\theta\mu\alpha\rho\rho\gamma\omega \ i \ \varepsilon \beta \alpha \zeta \alpha \tau \alpha \nu \\
\Delta\rho\theta\theta\mu\alpha\rho\rho\gamma\omega \ i \ \varepsilon \beta \alpha \zeta \alpha \tau \alpha \nu \\
\Delta\rho\theta\theta\mu\alpha\rho\rho\gamma\omega \ i \ \varepsilon \beta \alpha \zeta \alpha \tau \alpha \nu
\]

### Ordering

No Bactrian abecedary has been found to date. Greek letters have numeric values based on their alphabetical orders. It is conceivable that \( \text{SHO} \) could be identified with the archaic Greek letter \( \text{SAN} \), which has a value of 900; but no Bactrian text with that number has been yet discovered. \( \text{SAN} \) has not been separately encoded in the UCS, although its descendant, \( \text{SAMPI} \), has been. (\( \text{SAN} \) looks rather different from \( \text{SAMPI} \), and it is possible that it should be encoded as well for purposes of representing archaic Greek text, in the same way that \( \text{ARCHAIC KOPPA} \) and \( \text{KOPPA} \) have been disunified.) If \( \text{SHO} \) were identified with \( \text{SAN} \), it should be ordered – alphabetically – after \( \text{PI} \) (80) and before \( \text{KOPPA} \) (90). Current scholastic ordering practice for \( \text{SHO} \) does not do this: Sims-Williams orders it as a supplementary letter following \( \text{OMEGA} \), which practice we recommend for ISO/IEC 14651; some scholars have ordered it after \( \text{SIGMA} \), based on the Latin transliteration \( s \) \( \text{SIGMA} \) and \( s\acute{\ } \text{SHO} \). Since \( \text{SAN} \)’s numeric value is 900 and \( \text{OMEGA} \)’s is 800, it seems that it would do no harm to place \( \text{SHO} \) there in any case.

Figure 1. Bactrian inscription of Dasht-e Nawur, showing \( \text{SHO} \), \( \text{RHO} \), and \( \text{PHI} \). After N. Sims-Williams and J. Cribb, “A new Bactrian inscription of Kanishka the Great”, *Silk Road Art and Archaeology*, IV, [1995-1996, p. 136.]
Figure 2. Bactrian inscription of Surkh Kotal, showing SHO, PHI, and RHO. After R. Göbl, *Die Drei Versionen der Kaniska-Inschrift von Surkh Kotal*, Vienna 1965, Tafel XIII.


Figure 4. Bactrian document I 623001, showing PHI, RHO, and SHO. After N. Sims-Williams, “A Bactrian deed of manumission”, *Silk Road Art and Archaeology*, V, 1997-8 [1999], p. 210, fig. 4.

Figure 6. Sample of a modern edition of a Bactrian text. Note that LATIN LETTER THORN has been used since there was no font support for SHO. A. Maricq. “La grande inscription de Kaniska et l’étéo-tokharien”, *JA*, CCXLVI, 1958, p. 352.
