

Universal Multiple-Octet Coded Character Set  
International Organization for Standardization  
Organisation Internationale de Normalisation  
Международная организация по стандартизации

L2/02-056

**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 CE 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 *Κανηρσο* '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 ϐ is taken from the small looped PHI φ without a top stem; the stem is the same width as the descender of that same PHI. It is flimsy and unattractive. The second ϐ (which we prefer) has the same ascender and descender but has the bowl of the small RHO ϱ, as do the third, fourth, and fifth. The ascender of the third ϐ is taken from that of the small ETA η. The fourth ϐ adds the ascender to the full RHO, but looks unbalanced. The fifth ϐ is a whimsical fabrication.

The next four are taken from a new font, Evertimes, which Michael Everson has been developing. The first of these ϐ is taken again from the small looped PHI φ 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 ϱ, which has an ascender which swings slightly to the left. The second ϐ is the RHO's own descender flipped up, so that the ascender also swings to the left; the ascender of the third ϐ is taken from that of the small ETA η. The fourth ϐ (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.

<i>Times Greek</i>	<i>Evertimes</i>
Δαθβο-μαρηγο ι βαβαστανο	Δαθβο-μαρηγο ι βαβαστανο
• Δαθβο-μαρηγο ι βαβαστανο	Δαθβο-μαρηγο ι βαβαστανο
Δαθβο-μαρηγο ι βαβαστανο	Δαθβο-μαρηγο ι βαβαστανο
Δαθβο-μαρηγο ι βαβαστανο	• Δαθβο-μαρηγο ι βαβαστανο
Δαθβο-μαρηγο ι βαβαστανο	

### Ordering

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

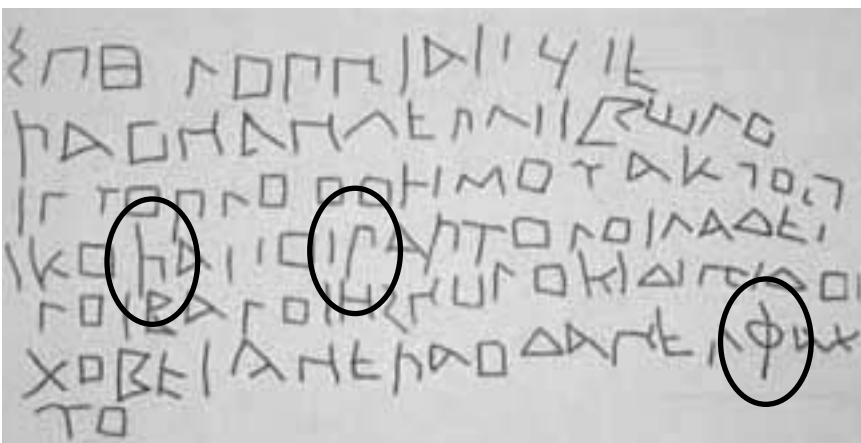


Figure 1. Bactrian inscription of Dasht-e Nawur, showing SHO, RHO, and PHI. After N. Sims-Williams and J. Cribb, "A new Bactrian inscription of Kanishka the Great", *Silk Road Art and Archaeology*, IV, [1995-1996, p. 136.

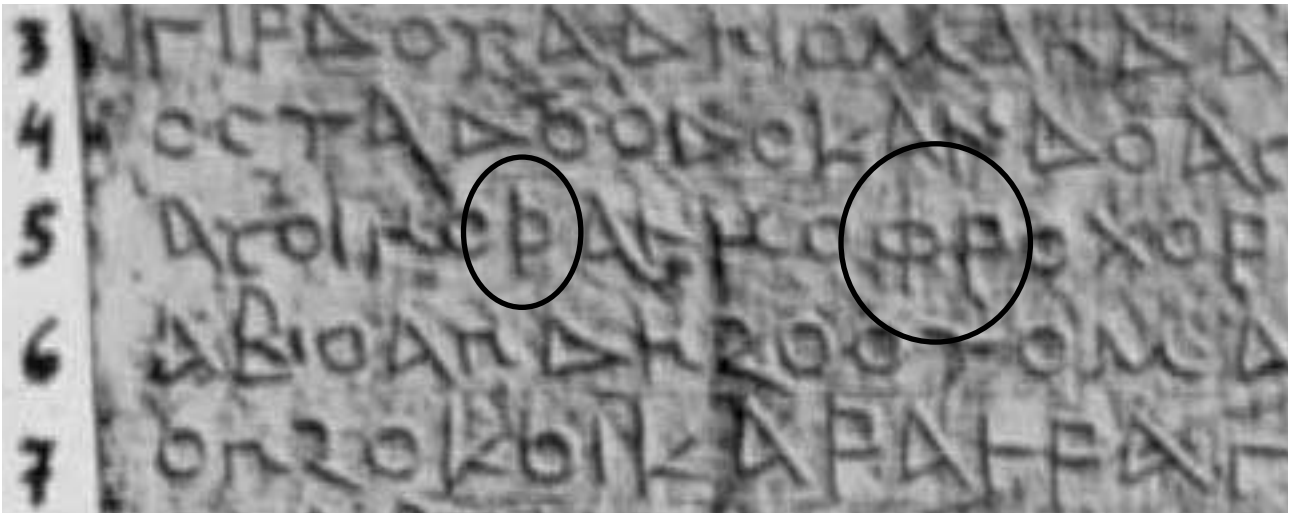


Figure 2. Bactrian inscription of Surkh Kotal, showing SHO, PHI, and RHO. After R. Göbl, *Die Drei Versionen der Kaniska-Inschrift von Surkh Kotal*, Vienna 1965, Tafel XIII.

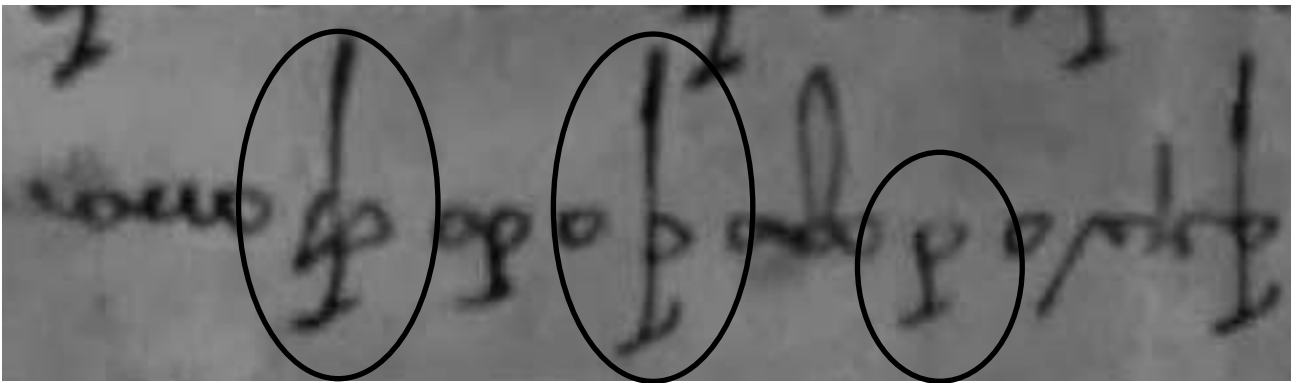


Figure 3. Bactrian letter DOC. 2, showing PHI, SHO, and RHO. After N. Sims-Williams, *Nouveaux documents sur l'histoire et la langue de la Bactriane*. *Comptes rendus de l'Académie des Inscriptions & Belles-lettres*, 1996 [1997], p. 640.

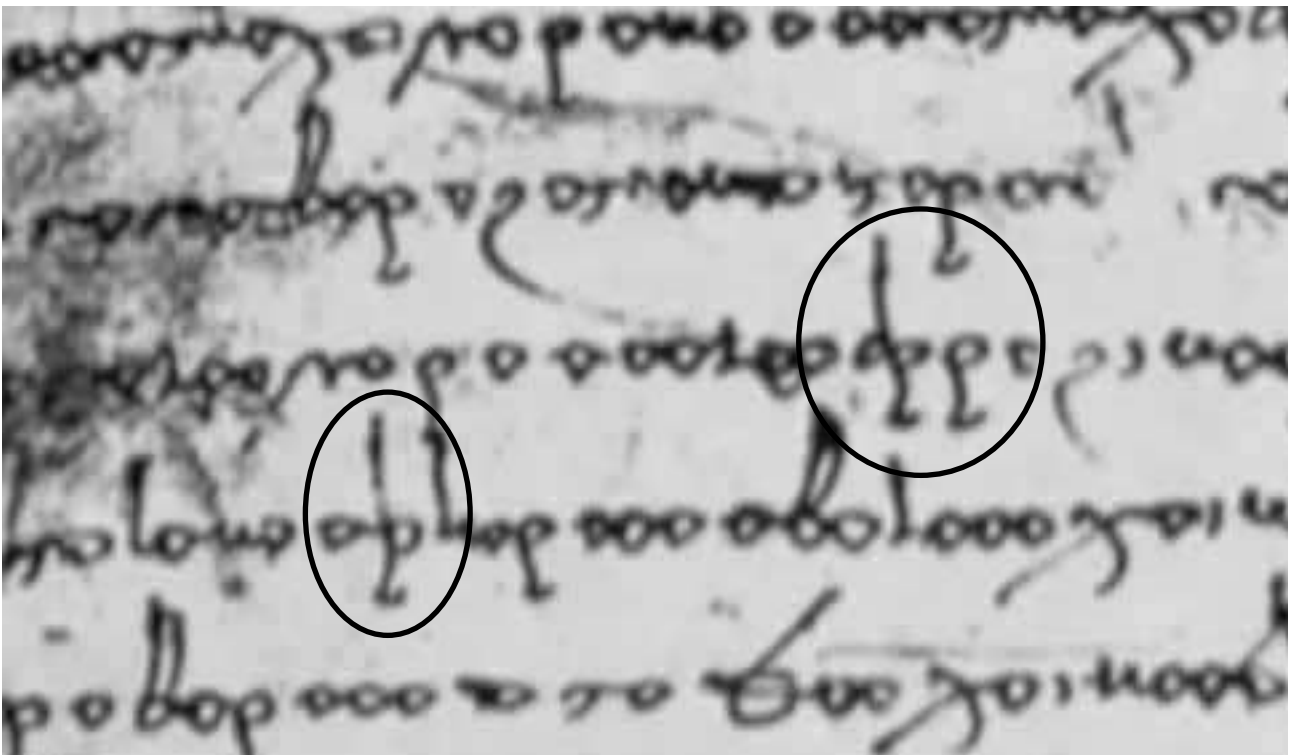


Figure 4. Bactrian document I 623001, showing PHI, RHO, and SHO. After N. Sims-Williams, "A Bactrian deed of manumission", *Silk Road Art and Archaeology*, V, 1997-8 [1999], p. 210, fig. 4.

αχβρονο τ' ν' σ' μανο μιρογανο  
 ρωσο γωβτο καλδο αγγιτο μανο τητ-  
 ο βαβορανο σωπανο οτο μανο πιο  
 σοροσανιγο σωπανο ασοφαγο μοζδο  
 παζοβορανο ιαιοιο γωρο λασο ι' οτο  
 σιλγο σπαλιγο κ' α' βαμδδιγο<sup>38</sup> οιταβο  
 σοφαρο <μανο> ταδο καλδαζο τητο οτο αζο  
 πιο λασο γωρο ιαιοιο οτο οιστο οτο σοφ-  
 αρο σπαλιγο μολο αβο ριζμο αβο ζαμα-  
 νο νατωζανο ταδο πορδογο βοοανο  
 αζο τητο οτο αζο πιο ιωγο πιδο πανζ-  
 ο οτο αζο τητο ασο πιο πιταδιμο ταδ-  
 ο καλδο πιο νατωζδο<sup>39</sup> ταδαζο τητο ασο  
 χοβο τωζα[ν]ο

Figure 5. Sample of a modern edition of a Bactrian text. N. Sims-Williams, *Bactrian documents from Northern Afghanistan, Vol. I: Legal and economic documents*, Oxford, 2000 [2001], p. 63.

(L'upsilon est transcrit par h, pour éviter toute confusion sur sa valeur phonétique, cf. p. 350.)

- 1 Ειδο μαλιζο μο Κανηρκο Οσινδο. Βαγολαγγο σιδοι βαγο ραο Κ-  
ανηρκι ναμο. Βαργο κηρδο ταδιοσκεδοφορδαμσο μαλιζο φρο-  
γυρδο ταδηιο μανδαρο ασο νιστοχοτο ασιδο μαλιζο αεαεγ-  
ο σταδο οδο καλδο ασολροχομινανοι ειρο σταδο. Ταδοι βαγε
- 5 ασοινοθαλμο φροχορτινδο — ταδο αεο λραφο ραστινδο  
αεο ανδηζο — οτο μαλιζο πιδοριγδο. Τα καλδο Νοκ-  
ουζκοι καρλαραγοι Φρει χοαδηο γο κηδο φρειστα-  
ρο αεο ραοι βαγοποηροι λοιχοεσοταροιφιζογαργοι α-  
λορχαλο κηδο φαρο οισπο ανομοσαδο βαργανο ω-
- 10 σογδο μαγγο πιδοι ιωγο οδο ηηρσο χρονο Νεισανο μ-  
αο μαλο αγαδο αμο βαγολαγγο ταδηιο μαλιζο πο-  
ρογατο ταδηιο ειρο σαδο κανδο οτηιο αεο οζσοαστο  
οτηιο πιδο ασαγγε ιθο οιληρδο ατανο αεο μαλιζο φ-  
αρο καρανο αεο μαγαοηιο οδο καλδανο ασολρο
- 15 ηομινανοι ειρο βοοηιο. Ταδανοι βαγε ασοινοθ-

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'éteo-tokharien", *JA*, CCXLVI, 1958, p. 352.

Die Endung dieses Wortes hat in letzterer Zeit Anlaß zu  
 einer Diskussion gegeben, da Baron Staël-Holstein darin immer  
 das Suffix des Plurals erkennen wollte, was von anderer Seite  
 bestritten wurde. In der Tat haben Fleet (*JRAS.* 1914, 374 ff.)  
 und Allan (ib. 408 ff.) nachgewiesen, daß es unmöglich ist, in  
 der Münzlegende ΠΑΟΝΑΝΟ ΠΑΟ ΚΑΝΗΡΚΙ ΚΟΡΑΝΟ das letzte  
 Wort, wie der russische Gelehrte vorschlug, als den Anfang  
 zu nehmen und als Genitiv Pluralis von ΠΑΟΝΑΝΟ ΠΑΟ ab-  
 hängen zu lassen: ‚der König der Könige der Kušan‘, woraus  
 implicite folgt, daß es hier nur Nominativ si. sein kann: ‚der  
 Kušan‘, als Apposition zu dem vorhergehenden Eigennamen.

Figure 6. Sample of CAPITAL LETTER BACTRIAN SHO. J. Kirste, "Orobazes", *Sitzungsberichte der Kais. Akademie der Wissenschaften in Wien, Phil.-hist. Klasse*, 182/2, 1917, p. 50.