

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

Doc Type: Working Group Document

Title: Proposal to encode the Osmanya script in Plane 1 of the UCS
Source: Michael Everson
Status: Expert Contribution
Action: For consideration by JTC1/SC2/WG2
Date: 1999-01-25

A. Administrative

1. Title

Proposal to encode the Osmanya script in Plane 1 of the UCS.

2. Requester's name

Michael Everson, EGT (WG2 member for Ireland).

3. Requester type

Expert contribution.

4. Submission date

1999-01-25.

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?

Yes. Osmanya.

1b. Addition of characters to existing block? Name?

No.

2. Number of characters

27.

3. Proposed category

Category B.1.

4. Proposed level of implementation and rationale

As a simple left-to-right alphabetic script, Osmanya requires Level 1.

5a. Character names included in proposal?

Yes.

5b. Character names in accordance with guidelines?

Yes.

5c. Character shapes reviewable?

Yes (see below).

6a. Who will provide computerized font?

Michael Everson.

6b. Font currently available?

Yes.

6c. Font format?

TrueType.

7a. Are references (to other character sets, dictionaries, descriptive texts, etc.) provided?

Yes. See the bibliography below.

7b. Are published examples (such as samples from newspapers, magazines, or other sources) of use of proposed characters attached?

Yes, see below.

8. Does the proposal address other aspects of character data processing?

Yes; see Unicode properties below.

C. Technical – Justification

1. Contact with the user community?

No.

2. Information on the user community?

Probably only the scholarly and library communities, since the script was abandoned in 1973 in favour of Latin orthography for Somali.

3a. The context of use for the proposed characters?

Used to write Somali, especially from 1961–1973 (when Osmanya enjoyed official sanction); Osmanya was devised in 1920–22.

3b. Reference

See bibliography below.

4a. Proposed characters in current use?

Yes.

4b. Where?

By scholars and librarians; there could possibly still be some native use.

5a. Characters should be encoded entirely in BMP?

No.

5b. Rationale

Accordance with the Roadmap.

6. Should characters be kept in a continuous range?

Yes.

7a. Can the characters be considered a presentation form of an existing character or character sequence?

No.

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.

E. Proposal

The Osmanya script, called in Somali *far Soomaali* ‘Somali writing’ or *Ismanya*, was devised in 1920–1922 by Cismaan Yuusuf Keenadiid to represent the Somali language. It replaced an attempt by

Sheikh Uweys of the Confraternity Qadiriyyah (died 1909) to devise an Arabic-based orthography for Somali. It has in turn been replaced by the Latin orthography of Muuse Xaaji Ismaaciil Galaal (1914-1980). In 1961 both the Latin and Osmanya scripts were adopted for use in Somalia, but in 1969 there was a coup, with one of its stated aims the resolution of the debate over the country's writing system. Latin orthography was adopted finally in 1973. Gregersen said in 1977 that some 20,000 or more people use Osmanya in private correspondence and in bookkeeping and that several books and a fortnightly journal *Horseed* ('*Vanguard*') was published, all "cyclostyped" which I take to be equivalent to "cyclostyled" and to mean 'mimeographed'.

Structure


Osmanya is a simple alphabetic script, read from left to right in horizontal lines running from top to bottom. It has 22 consonants and five vowels. Long vowels are written with Arabic conventions: by the addition of ALEF to A; by the addition of WAW to U and O; by the addition of YEE to E and I. Long EE and OO can also be represented by doubling the signs.


Ordering

Alphabetical ordering follows the order of the Arabic alphabet, except that AIN (CAIN) and GHAIN (GAIN) appear to be reversed. This ordering, given in Diringer 1996, has been followed here. Diringer gives only the transliteration of the characters; in the names list below I have used the names of the Arabic letters of the alphabet, written in modern Latin Somali orthography. There may be errors in the spelling of the names.

Issues

Gregerson 1977 says the alphabet has 29 signs including 10 vowels, which doesn't quite tally with the way I'd interpret the data. It would help if I could see Cerulli 1932 and Maino 1951, but all I have is Gregersen. Gregersen says that Cerulli describes long-vowel representation as represented here in this proposal (and which was also presented in Diringer 1947). He also says that Maino describes "special symbols, which were modifications of those used for the corresponding short vowels" as being used to represent long vowels. These sound to me like ligatures, but they are probably fixed ligature-letters like æ and œ, not contextual ligatures as in Arabic. The alphabets Gregersen gives are presented below in Annex A. He shows the following orders:

Cerulli: 
 ? b t j x kh d r s sh dh g c f q k l m n w h y i u o a e

Maino: 
 b t j x kh d r s sh dh g c f q k l m n h i u o a e y w o a e

Where the two differ is that *alef*, *waw* and *yee* are considered long vowels, not consonants, and *o* and *e* have unique long forms as well. So it may be that there were two orthographies used. The sample given in Nakanishi shows its second-to-last character as what looks like an *i-alef* ligature. Gregersen says that Cismaan Yuusuf "himself denied ever using the system reported by Cerulli".

Code positions proposed in the table are based on my Roadmap to the BMP, version 3.0, 1999-01-23.

Unicode Character Properties

Spacing letters, category "Lo", bidi category "L" (strong left to right)
 0001 2000 - 0001 201A

Bibliography

Cerulli, Enrico. 1932. "Tentativo indigeno di formare un alfabeto somalo", in *Oriente Moderno*

12:212-13.

Diringer, David. 1996 (1947). *The alphabet: a key to the history of mankind*.
 Munshiram Manoharlal. ISBN 81-215-0748-0

Gaur, Albertine. 1992. *A history of writing*. London: British Library. ISBN 0-7123-0270-0

Gregersen, Edgar A. 1977. *Language in Africa: an introductory survey*. New York: Gordon and Breach. ISBN 0-677-04380-5

Maino, Mario. 1951. "L'alfabeta 'Osmania' in Somalia", in *Rassegna di Studi Etopici* 10:108-21.

Nakanishi, Akira. 1990. *Writing systems of the world: alphabets, syllabaries, pictograms*. Rutland, VT: Charles E. Tuttle. ISBN 0-8048-1654-9

Annex A

Sample from Nakanishi 1990.

OSN GLQT NZZ
 99 SS 99X7650Z
 99952L NZZ 99
 SS OL9Z

Transcription:

ΔSN 6-627 NZZ 99 SS 99X7650Z 49452L NZZ 99 SS ΔL9Z

dal dheyr low gu aa guwrsado shishaye low gu aa deio

Sample from Gaur 1990.

6627 NZZ 99 SS 99X7650Z 952

Transcription:

6-627 NZZ 99 SS 99X7650Z 452

dheyr low gu aa guwrsadw shay

This text is suspect; it appears to be derived from the sample in Nakanishi 1990.

To the right is shown Cerulli and Maino's alphabets as given in Gregersen 1977.

	Cerulli 1932	Maino 1951
?	U	u
b	3	β
t	o	1
j	—	—
h	H	h
x	S	h
d	Δ	0
r	∇	∇
s	2	2
l	∩	∩
d	5	5
g	o	o
q	5	5
f	7	7
q	7	7
k	9	9
l	∩	∩
m	5	5
n	2	2
w	5	5
h	∩	∩
y	2	2
i	o	o
u	o	o
o	∩	∩
a	∩	∩
e	∩	∩
	y, ii	u
	w, uu	β
	oo	o
	aa	∩
	ee	u

TABLE XXX - Row 20: OSMANYA

	200	201
0	Რ	Ს
1	Ტ	Უ
2	Ფ	Ქ
3	Ღ	Ყ
4	Შ	Ჩ
5	Ც	Ძ
6	Წ	Ჭ
7	Ხ	Ჯ
8	Ჰ	Ჱ
9	Ჲ	Ჳ
A	Ჴ	Ჵ
B	Ჷ	Ჸ
C	Ჹ	Ჺ
D	᲼	
E	Ჽ	
F	Ჿ	

G = 00
P = 01

TABLE XXX - Row 20: OSMANYA

dec	hex	Name	dec	hex	Name
	00	OSMANYA LETTER ALEF			
	01	OSMANYA LETTER BEE			
	02	OSMANYA LETTER TEE			
	03	OSMANYA LETTER JIIM			
	04	OSMANYA LETTER XAA			
	05	OSMANYA LETTER KHAA			
	06	OSMANYA LETTER DAL			
	07	OSMANYA LETTER REE			
	08	OSMANYA LETTER SIIN			
	09	OSMANYA LETTER SHIIN			
	0A	OSMANYA LETTER DHAD			
	0B	OSMANYA LETTER GAIN			
	0C	OSMANYA LETTER CAIN			
	0D	OSMANYA LETTER FEE			
	0E	OSMANYA LETTER QAF			
	0F	OSMANYA LETTER KAF			
	10	OSMANYA LETTER LAM			
	11	OSMANYA LETTER MIIM			
	12	OSMANYA LETTER NUN			
	13	OSMANYA LETTER WAW			
	14	OSMANYA LETTER HEE			
	15	OSMANYA LETTER YEE			
	16	OSMANYA LETTER I			
	17	OSMANYA LETTER U			
	18	OSMANYA LETTER O			
	19	OSMANYA LETTER A			
	1A	OSMANYA LETTER E			
	1B	OSMANYA LETTER OO ?????			
	1C	OSMANYA LETTER EE ?????			
	1D	(This position shall not be used)			
	1E	(This position shall not be used)			
	1F	(This position shall not be used)			

Group 00

Plane 01

Row 20