

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 for encoding the Carian script in the UCS

Source: Michael Everson

Status: Individual Contribution

Action: For consideration by JTC1/SC2/WG2 and UTC

Date: 2005-04-27

The Carian script is used to write the Carian language, an ancient Indo-European language of southwestern Anatolia. It dates from the first millennium BCE. While a few texts have been found in Caria, most of the written evidence comes from Carian communities in Egypt. The script is related to Greek.

Processing

The primary direction of writing is right-to-left in Egyptian texts, but left-to-right in Caria. Word dividers are not regularly employed; *scriptio continua* is common. Some values of letters remain unknown or in dispute (though their distinction from other letters is not).

Names

The names used for the characters here are based on their Latin transliterations.

Bibliography

Melchert, H. Craig. 2004. "Carian", in *The Cambridge encyclopedia of the world's ancient languages*, ed. Roger Woodard, 609-613. Cambridge: Cambridge University Press.

Swiggers, Pierre, & Wolfgang Jenniges. 1996. "The Anatolian alphabets", in *The World's Writing Systems*, ed. Peter T. Daniels & William Bright. New York; Oxford: Oxford University Press. ISBN 0-19-507993-0

Unicode Character Properties

```
10B20;CARIAN LETTER A;Lo;0;R;;;;N;;;;;
10B21;CARIAN LETTER D;Lo;0;R;;;;N;;;;;
10B22;CARIAN LETTER L;Lo;0;R;;;;N;;;;;
10B23;CARIAN LETTER UUU;Lo;0;R;;;;N;;;;;
10B24;CARIAN LETTER R;Lo;0;R;;;;N;;;;;
10B25;CARIAN LETTER LY;Lo;0;R;;;;N;;;;;
10B26;CARIAN LETTER Q;Lo;0;R;;;;N;;;;;
10B27;CARIAN LETTER B;Lo;0;R;;;;N;;;;;
10B28;CARIAN LETTER M;Lo;0;R;;;;N;;;;;
10B29;CARIAN LETTER O;Lo;0;R;;;;N;;;;;
10B2A;CARIAN LETTER T;Lo;0;R;;;;N;;;;;
10B2B;CARIAN LETTER SH;Lo;0;R;;;;N;;;;;
10B2C;CARIAN LETTER S;Lo;0;R;;;;N;;;;;
10B2D;CARIAN LETTER U;Lo;0;R;;;;N;;;;;
10B2E;CARIAN LETTER X;Lo;0;R;;;;N;;;;;
10B2F;CARIAN LETTER N;Lo;0;R;;;;N;;;;;
10B30;CARIAN LETTER P;Lo;0;R;;;;N;;;;;
10B31;CARIAN LETTER SS;Lo;0;R;;;;N;;;;;
10B32;CARIAN LETTER I;Lo;0;R;;;;N;;;;;
10B33;CARIAN LETTER E;Lo;0;R;;;;N;;;;;
10B34;CARIAN LETTER W;Lo;0;R;;;;N;;;;;
10B35;CARIAN LETTER K;Lo;0;R;;;;N;;;;;
```

10B36;CARIAN LETTER UU;Lo;0;R;;;;;N;;;;;
 10B37;CARIAN LETTER II;Lo;0;R;;;;;N;;;;;
 10B38;CARIAN LETTER TT;Lo;0;R;;;;;N;;;;;
 10B39;CARIAN LETTER WW;Lo;0;R;;;;;N;;;;;

Figures

TABLE 22.9: *The Letters of the Carian Alphabet^a*

1	Α Δ Λ	a
3	ϸ	d
4	Δ	l
5	Ε Ε	ù (alternating with sign 28)
6	Ϝ Ϝ Ϝ	r
7	Ι Ι	λ
9	⊕	q
10	Γ Γ Λ	b
11	Λ Ν	m
12	Ο	o (corresponding to Greek ω)
14	ϙ	t
15	Ϡ Ϡ Ϡ	š
17	Μ Μ	s
19	Υ V	u (/u)
20	ϕ	ñ
21	Χ +	χ
22	Υ V	n
24	Δ Δ Μ	p
25	⊙ ⊙	ś
26	⊗ ⊗ ⊗	i
27	□	e (corresponding to Greek η)
28	ϙ	w (/ə) (alternating with sign 5)
29, 30	∇ ∇ ∇	k
32	Π Π	ú (or ü)
38	Η Η	í (/i)
40	↑ †	τ (/t)
41	⊖	variant of 28

a. Omitted numbers refer to signs included in decipherment lists for which no established value can yet be given.

Figure 1. Table of Carian letters from Swiggers & Jenniges 1996.

Table 23.1 A subset of characters of the Carian alphabet

Character	Transcription
Α	a
ϸ	d
Δ	l
Ε	ù
Ϝ	r
Ι	λ
⊕	q
Γ	b
Λ	m
Ο	o
ϙ	t
ϛ	š
Μ	s
Υ	u
Χ	x
Υ	n
ΔΔ	p
Φ	ś
ϑ	i
Ϟ	e
ϙ	w
∇	k
Π	ú
Η	í
↑	τ
⊥	w

Figure 2. Table of Carian letters from Melchert 2004.

TABLE XX - Row 10B: CARIAN

	10B2	10B3
0	A	Λ
1	⋈	⊖
2	Δ	⊣
3	Ɔ	◻
4	Ɔ	⊞
5	I	∇
6	⊕	⏏
7	Γ	⏏
8	N	↑
9	⊙	⏏
A	♀	
B	♁	
C	Λ	
D	Υ	
E	×	
F	Υ	

G = 00
P = 01

TABLE XX - Row 10B: CARIAN

hex	Name
20	CARIAN LETTER A
21	CARIAN LETTER D
22	CARIAN LETTER L
23	CARIAN LETTER UUU
24	CARIAN LETTER R
25	CARIAN LETTER LY
26	CARIAN LETTER Q
27	CARIAN LETTER B
28	CARIAN LETTER M
29	CARIAN LETTER O
2A	CARIAN LETTER T
2B	CARIAN LETTER SH
2C	CARIAN LETTER S
2D	CARIAN LETTER U
2E	CARIAN LETTER X
2F	CARIAN LETTER N
30	CARIAN LETTER P
31	CARIAN LETTER SS
32	CARIAN LETTER I
33	CARIAN LETTER E
34	CARIAN LETTER W
35	CARIAN LETTER K
36	CARIAN LETTER UU
37	CARIAN LETTER II
38	CARIAN LETTER TT
39	CARIAN LETTER WW
3A	(This position shall not be used)
3B	(This position shall not be used)
3C	(This position shall not be used)
3D	(This position shall not be used)
3E	(This position shall not be used)
3F	(This position shall not be used)

hex	Name
-----	------

A. Administrative

1. Title

Proposal for encoding the Carian script in the UCS

2. Requester's name

Michael Everson

3. Requester type (Member body/Liaison/Individual contribution)

Individual contribution.

4. Submission date

2005-04-27

5. Requester's reference (if applicable)

6. Choose one of the following:

6a. This is a complete proposal

No.

6b. More information will be provided later

This is a preliminary proposal.

B. Technical – General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters)

Yes.

Proposed name of script

Carian.

1b. The proposal is for addition of character(s) to an existing block

No.

1b. Name of the existing block

2. Number of characters in proposal

27.

3. Proposed category (see section II, Character Categories)

Category C

4a. Proposed Level of Implementation (1, 2 or 3) (see clause 14, ISO/IEC 10646-1: 2000)

Level 1.

4b. Is a rationale provided for the choice?

Yes.

4c. If YES, reference

Spacing characters are proposed.

5a. Is a repertoire including character names provided?

Yes.

5b. If YES, are the names in accordance with the character naming guidelines in Annex L of ISO/IEC 10646-1: 2000?

Yes.

5c. Are the character shapes attached in a legible form suitable for review?

Yes.

6a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Michael Everson. TrueType.

6b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Michael Everson, Fontographer.

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

Yes, see bibliography above.

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

Yes.

8. 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)?

Yes, see 9 below.

9. 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 Unicode Character Database <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.

Yes, see Unicode properties above.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

No

2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

Yes.

2b. If YES, with whom?

Deborah Anderson, University of California, Berkeley

2c. If YES, available relevant documents

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

Scholarly communities researching the Carian language.

4a. The context of use for the proposed characters (type of use; common or rare)

Carian is fairly rare as these things go.

4b. Reference

5a. Are the proposed characters in current use by the user community?

Yes.

5b. If YES, where?

By scholars worldwide.

6a. After giving due considerations to the principles in Principles and Procedures document (a WG 2 standing document) must the proposed characters be entirely in the BMP?

No.

6b. If YES, is a rationale provided?

6c. If YES, reference

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

Yes, they should be encoded in a single block as presented here.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No.

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No.

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

No.

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

11a. 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)?

No.

11b. If YES, is a rationale for such use provided?

11c. If YES, reference

12a. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

12b. If YES, reference

13a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

13b. If YES, describe in detail (include attachment if necessary)

14a. Does the proposal contain any Ideographic compatibility character(s)?

No.

14b. If YES, is the equivalent corresponding unified ideographic character(s) identified?

14c. If YES, reference