2003-08-24

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

L2/03-283

Doc Type: Working Group Document

Title: Proposal to add the Coptic alphabet to the BMP of the UCS

Source: NSAI/ICTSCC/SC4, Ireland Status: Member Body Contribution

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

Date: 2003-08-24

This document requests additional characters to be added to the UCS and contains the proposal summary form.

A. Administrative

1. Title

Proposal to add the Coptic alphabet to the BMP of the UCS.

2. Requester's name

Ireland

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

Member body contribution.

4. Submission date

2003-08-24

5. Requester's reference (if applicable)

N2444 (2002-05-08)

6. Choose one of the following:

6a. This is a complete proposal

Yes.

6b. More information will be provided later

No.

B. Technical – General

1. Choose one of the following:

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

No, but it is a proposal for a new block of Coptic characters.

Proposed name of script

Coptic

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

Yes, one.

1b. Name of the existing block

General punctuation

2. Number of characters in proposal

60(59+1)

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

Category A

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?

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 below.

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 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 below.

C. Technical – Justification

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

Yes. N1658, N2444. UTC has recently acknowledged the need to disunify Coptic from Greek.

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?

Stephen Emmel representing the International Association for Coptic Studies (IACS)

2c. If YES, available relevant documents

N2444

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

Scholarly, ecclesiastical, and library communities.

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

Characters are used to write the Coptic language.

4b. Reference

See below.

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

Yes.

5b. If YES, where?

By scholars, ecclesiastical researchers, Coptic Christians, and librarians.

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?

Yes.

6b. If YES, is a rationale provided?

Yes.

6c. If YES, reference

Accordance with the Roadmap. Other Coptic characters are encoded in the Greek and Coptic block.

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?

Yes.

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

Yes.

8c. If YES, reference

Coptic characters are similar to, but not identical to, Greek characters, and Copticists need to distinguish the two in plain text.

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?

Yes.

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

Yes

10c. If YES, reference

The more generic U+2627 CHI RHO must be considered to be a logotype derived from two Greek letters; the KHI RO proposed here is derived from two Coptic letters, and is used in Coptic text.

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)?

Vec

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

Ves

11c. If YES, reference

Combining diacritical marks.

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?

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

D. Proposal

On foot of the arguments given in N2444, the UTC and WG2 have decided to disunify the Coptic script from the Greek script. Currently the UCS encodes 7 Coptic letters (14 characters with the casing pairs); this proposal adds 26 letters (52 characters with the casing pairs), as well as several other characters needed for basic Coptic text processing, including standard logotypes used in Coptic text. It is recognized that there may be further additions required by scholars for special purposes, but the set here will allow basic Coptic text processing of, for instance, the Gospel of Thomas and other important Coptic texts. The set proposed here is the *core character set* for Coptic.

Combining characters

A number of generic characters are used in Coptic. The diaeresis (U+ 0308 ") can be used over ι and \mathbf{v} in syllable-initial position; the apostrophe (U+2019 ') is used to indicate the end of phonological words and clitics; the circumflex (U+0302 ^) is used in post-seventh century texts on vowels which form independent syllables. According to Fossey 1948, the acute accent (U+0301 ') can be used to indicate stress; likely this is a didactic use in modern editions.

A special combining mark called the *jinkim* is used in Coptic. The *jinkim* indicates the syllabic pronunciation of a character, that is, a consonant pronounced with a reduced vowel: so $c\bar{n}$ is pronouced [s^cn]. Coptic scribes treated the *jinkim* differently: in Sahidic, it looks like a macron; in Akhmimic

(sometimes Fayumic), it looks like a grave; in Bohairic, it looks like a dot above. At this time we accept the proposal to unify these with macron (U+ 0304 ¯), grave (U+ 0300 ˚), and dot above (U+ 0307 ˚). It may be the case that because of monofont issues with Greek varia or Latin macron that this may prove impractical, but at this time we are only proposing to add the core character set. We expect that the UTC and WG2 will look as favourably on an eventual disunification of the *jinkims* from the other combining accents as they have for the disunification of Coptic letters from Greek letters, if Copticists find they have the same problems.

Punctuation

A number of punctuation characters are used in Coptic. Several of these are generic punctuation characters: FULL STOP, COLON, SEMICOLON (= GREEK QUESTION MARK) and MIDDLE DOT (= GREEK ANO TELEIA) are used, and in edited texts quotation marks are used. A COPTIC FULL STOP *• can be used. In some didactic texts, hyphen is distinguished from a special oblique double hyphen * as can be seen in the samples below. This character has been observed in use in other scripts. We believe it should be encoded in the General Punctuation block. It forms a part of *this* proposal; we believe the evidence here is strong enough to encode it at U+2055, and we encourage UTC and WG2 not to wait for "another proposal".

Numerals

Numerals are indicated with letters of the alphabet, as in Greek; sometimes combining overline U+0305 is used with them. The GREEK LOWER NUMERAL SIGN U+0375 or the COMBINING DOUBLE OVERLINE U+033F can be used to indicate multiples of 1000, so $\mathbf{x} = 1$, \mathbf{x}_{i} or $\overline{\mathbf{x}} = 1000$; \mathbf{x}_{i} $\overline{\boldsymbol{\omega}}\overline{\boldsymbol{n}}\overline{\boldsymbol{\mu}} = 1888$. The GREEK NUMERAL SIGN U+0374 is used for fractions, so $\mathbf{y}' = \frac{1}{3}$, though $\frac{1}{2}$ is indicated by the symbol \mathbf{s} . See figure 11.

Abbreviations

The combining overline U+0305 is also used to indicate abbreviations: $\overline{\pi n a}$ for $\pi n \epsilon r u a$.

Additional symbols

In dates, the Diocletan era is indicated with a \mathbf{p}_{i} mi ro or \mathbf{p}_{i} khi ro (either referring to the Greek χρόνος μαρτύρων 'time of martyrs'). These are conventional symbols which cannot be algorithmically decomposed. Note the use of \mathbf{p} RO WITH STROKE as a standard abbreviation: $\mathbf{c}_{\mathbf{p}0\mathbf{c}} = \mathbf{c}_{\mathbf{T}\mathbf{a}\mathbf{v}\mathbf{p}0\mathbf{c}}$ 'cross'; it is not a ligature of * $\mathbf{T}\mathbf{a}\mathbf{v}\mathbf{p}$.

Character names.

Transliterated character names taken from Coptic dialects as given in Crum are used for the Coptic letters. A chart showing these and other variants is given below. Note that we have specified five columns for Coptic. It is known that there will be more than five *non-core* additions to this script which await further study.

Unicode Character Properties

```
2055; DOUBLE HYPHEN; Pd; 0; ET;;;; N;;; 2055;
2C80; COPTIC CAPITAL LETTER ALFA; Lu; 0; L;;;; N;;; 2CA0;
2C81; COPTIC CAPITAL LETTER VIDA; Lu; 0; L;;;; N;;; 2CA1;
2C82; COPTIC CAPITAL LETTER GAMMA; Lu; 0; L;;;; N;;; 2CA2;
2C83; COPTIC CAPITAL LETTER DALDA; Lu; 0; L;;;; N;;; 2CA3;
2C84; COPTIC CAPITAL LETTER EIE; Lu; 0; L;;;; N;;; 2CA4;
2C85; COPTIC CAPITAL LETTER SOU; Lu; 0; L;;;; N;;; 2CA5;
2C86; COPTIC CAPITAL LETTER ZATA; Lu; 0; L;;;; N;;; 2CA6;
2C87; COPTIC CAPITAL LETTER HATE; Lu; 0; L;;;; N;;; 2CA7;
2C88; COPTIC CAPITAL LETTER THETHE; Lu; 0; L;;;; N;;; 2CA9;
2C80; COPTIC CAPITAL LETTER IAUDA; Lu; 0; L;;;; N;;; 2CA9;
2C8A; COPTIC CAPITAL LETTER KAPA; Lu; 0; L;;;; N;;; 2CAB;
2C8C; COPTIC CAPITAL LETTER LAULA; Lu; 0; L;;;; N;;; 2CAB;
2C8C; COPTIC CAPITAL LETTER MI; Lu; 0; L;;;; N;;; 2CAC;
2C8D; COPTIC CAPITAL LETTER MI; Lu; 0; L;;;; N;;; 2CAD;
```

```
2C8E; COPTIC CAPITAL LETTER KSI; Lu; 0; L;;;;; N;;; 2CAE;
2C8F; COPTIC CAPITAL LETTER O; Lu; 0; L;;;;; N;;; 2CAF;
2C90; COPTIC CAPITAL LETTER PI; Lu; 0; L;;;;; N;;; 2CB0;
2C91; COPTIC CAPITAL LETTER RO; Lu; 0; L;;;;; N;;; 2CB1;
2C92; COPTIC CAPITAL LETTER SIMA; Lu; 0; L;;;;; N;;; 2CB2;
2C93; COPTIC CAPITAL LETTER TAU; Lu; 0; L;;;;; N;;; 2CB3;
2C94; COPTIC CAPITAL LETTER UA; Lu; 0; L;;;;; N;;; 2CB4;
2C95; COPTIC CAPITAL LETTER FI; Lu; 0; L;;;;; N;;; 2CB5;
2C96; COPTIC CAPITAL LETTER KHI; Lu; 0; L;;;;; N;;; 2CB6;
2C97; COPTIC CAPITAL LETTER PSI; Lu; 0; L;;;;; N;;; 2CB7;
2C98; COPTIC CAPITAL LETTER OOU; Lu; 0; L;;;;; N;;; 2CB8;
2C99; COPTIC CAPITAL LETTER AKHMIMIC KHEI; Lu; 0; L;;;;; N;;; 2CB9;
2C9A; COPTIC CAPITAL LETTER RO WITH STROKE; Lu; 0; L;;;;; N;;; 2CBA;
2CAO; COPTIC SMALL LETTER ALFA; L1; 0; L;;;;; N;;; 2C80;; 2C80
2CA1; COPTIC SMALL LETTER VIDA; L1; 0; L;;;;; N;;; 2C81;; 2C81
2CA2; COPTIC SMALL LETTER GAMMA; L1; 0; L;;;;; N;;; 2C82;; 2C82
2CA3; COPTIC SMALL LETTER DALDA; L1; 0; L;;;;; N;;; 2C83;; 2C83
2CA4; COPTIC SMALL LETTER EIE; L1; 0; L; ;; ;; N; ;; 2C84; ; 2C84
2CA5; COPTIC SMALL LETTER SOU; L1; 0; L;;;;; N;;; 2C85;; 2C85
2CA6; COPTIC SMALL LETTER ZATA; L1; 0; L;;;;; N;;; 2C86;; 2C86
2CA7; COPTIC SMALL LETTER HATE; L1; 0; L;;;;; N;;; 2C87;; 2C87
2CA8; COPTIC SMALL LETTER THETHE; L1; 0; L;;;;; N;;; 2C88;; 2C88
2CA9; COPTIC SMALL LETTER IAUDA; L1; 0; L;;;; N;;; 2C89;; 2C89
2CAA; COPTIC SMALL LETTER KAPA; L1; 0; L;;;;; N;;; 2C8A;; 2C8A
2CAB; COPTIC SMALL LETTER LAULA; L1; 0; L;;;;; N;;; 2C8B;; 2C8B
2CAC; COPTIC SMALL LETTER MI; L1; 0; L;;;;; N;;; 2C8C;; 2C8C
2CAD; COPTIC SMALL LETTER NI; L1; 0; L;;;;; N;;; 2C8D;; 2C8D
2CAE; COPTIC SMALL LETTER KSI; Ll; 0; L;;;;; N;;; 2C8E;; 2C8E
2CAF; COPTIC SMALL LETTER O; L1; 0; L;;;;; N;;; 2C8F;; 2C8F
2CB0; COPTIC SMALL LETTER PI; L1; 0; L;;;;; N;;; 2C90;; 2C90
2CB1; COPTIC SMALL LETTER RO; L1; 0; L;;;;; N;;; 2C91;; 2C91
2CB2; COPTIC SMALL LETTER SIMA; L1; 0; L;;;;; N;;; 2C92;; 2C92
2CB3;COPTIC SMALL LETTER TAU;L1;0;L;;;;;N;;;2C93;;2C93
2CB4; COPTIC SMALL LETTER UA; L1; 0; L;;;;; N;;; 2C94;; 2C94
2CB5; COPTIC SMALL LETTER FI; Ll; 0; L;;;;; N;;; 2C95;; 2C95
2CB6; COPTIC SMALL LETTER KHI; L1; 0; L;;;;; N;;; 2C96;; 2C96
2CB7;COPTIC SMALL LETTER PSI;L1;0;L;;;;N;;;2C97;;2C97
2CB8; COPTIC SMALL LETTER OOU; L1; 0; L;;;; N;;; 2C98;; 2C98
2CB9; COPTIC SMALL LETTER AKHMIMIC KHEI; L1; 0; L; ;; ;; N; ;; 2C99; ;2C99
2CBA; COPTIC SMALL LETTER RO WITH STROKE; L1; 0; L;;;;;; N;;; 2C9A;; 2C9A
2CBB; COPTIC SYMBOL MI RO; Ll; 0; L;;;;; N;;;;;
2CBC; COPTIC SYMBOL KHI RO; Ll; 0; L;;;; N;;;;
2CBD; COPTIC SYMBOL KHRISTOS; Ll; 0; L;;;;; N;;;;;
2CCE; COPTIC FULL STOP; Po; 0; CS;;;;; N;;;;;
2CCF; COPTIC FRACTION ONE HALF; No; 0; ON;;;; 1/2; N;;;;;
```

Bibliography

Crum, W. E. 2000 (1939). A Coptic dictionary. Oxford: Clarendon Press. ISBN 0-19-864404-3

Fossey, Charles. 1948. *Notices sur les caractères étrangers anciens et modernes rédigées par un groupe e savants*. Nouvelle édition míse à jour à l'occasion du 21° Congrès des Orientalistes. Paris: Imprimerie Nationale de France.

Haarmann, Harald. 1990. Die Universalgeschichte der Schrift. Frankfurt: Campus Verlag. ISBN 3-593-34346-0

Reichsdruckerei. 1924. Alphabete und Schriftzeichen des Morgen- und Abendlandes, zum allgemeinen Gebrauch mit besonderer Berücksichtigung des Buchgewerbes. Unter Mitwerkung von Fachgelehrten zusammengestellt in der Reichsdruckerei. Berlin: Reichsdruckerei.

Ritner, Robert K. 1996. "The Coptic alphabet", in Daniels, Peter T., and William Bright, eds. *The world's writing systems*. New York; Oxford: Oxford University Press. ISBN 0-19-507993-0

Taylor, Isaac. 1883. The alphabet: an account of the origin and development of letters. Vol. 1: Semitic alphabets; Vol. 2: Aryan alphabets. London: Kegan Paul.

Till, Walter C. 1966. Koptische Grammatik: saïdischer Dialekt, mit Bibliographie, Lesestücken und Wörterverzeichnissen. Leipzig: VEB Verlag Enzykopädie.

Figures

THE COPTIC ALPHABET.

			I II C			LINABL		
	Names.	Values.	Сортіс.	Greek Uncials.		Names.	Values.	Сортіс.
1	Alpha	а	2 a	入	17	Pi	p, b	Пπ
2	Vida	b, v	BB	В	18	Ro	r	Pρ
3	Gamma	g	7 5	Г	19	Sima	s	Сс
4	Dalda	d	λλ	٨	20	Tau	t, d	Ττ
5	Ei	e	Εe	е	21	He	ü	Y Y
6	So ·	6	2.8	G	22	Phi	ph	фф
7	Zita	dz	33	z	23	Khi	kh	$\infty \infty$
8	Ita	e, ī	Нн	Н	24	Psi	ps	ψψ
9	Thita	th	Ө ө	ө	25	Au	\bar{u}, \bar{o}	Шω
10	Jauta	i	Ιι	I.	26	Shei	sh	யூய
11	Kapa	k	Кк	1<	27	Fei	f	Чq
12	Laula	ı	22	入	28	Chei	ch	þр
13	Mi	m	Uu	М	29	Hori	h	SS
14	Ni	n	n H	N	30	Janjia	j.g,dzh	x x
15	Ksi	ks	33	3	31	Tshima	tsh	5 6
16	o	0	0 0	0	32	Ti	ti	44
				1				

Figure 1. The Coptic alphabet, from Taylor 1883.

Die koptischen Buchstaben und ihre Umschreibung:

λa	l i, j	Рr	(1) š, sch
B b	K k	Cs	q f
Гg	λl	Τt	ь, \$ h, ch
λd	M m	Υu, w, y	a h, ch
€ ĕ	N n	фph	2 h
Ζz	ž ks, x	X kh, ch	🗶 č, tsch
Hē, ê	Οŏ	Ψ ps	6 c, kj
O th	пр	wō,ô	十 ti

Die Namen der Buchstaben sind nur bohairisch in moderner Form erhalten und lauten nach Workell (UM 314ff.) folgendermaßen²: λ alfa, β bēda oder vēda, Γ gamma, λ dalda, ε ēje, z zāda, κ hāda, ε tútte, ι jōda, κ kabba, λ lōla, κ mēj, κ ni', z eksi, Ο ōw, π bēj, γ rōw, c sámma, τ daū, γ he, φ fīj, x kij, ψ ébsi, ω ō', ω schaj, η fāj, η chāj, 2 hōri, x dschandscha, ε schīma, † dīj, dī.

Figure 2. The Coptic alphabet, from Till 1966.

Name	Zeichen	Umschrift und Lautwert	Allgemeines. Das Koptische ist die in christlicher Zeit (seit dem 3. Jahrhundert) gesprochene und mit griechischen Buch-
416	2		staben geschriebene Sprache der Agypter. Seit dem 16. Jahr- hundert ist es vom Vulgärarabischen verdrängt und lebt heute
Alfa	A a	a	nur noch im bohairischen Dialekt als Kirchensprache fort (ähnlich dem Latein der katholischen Kirche). Der Name dieses
Vida	B&	b [b, v]	letzten Ausläufers der vorher mit Hieroglyphen (siehe Taf. 6) geschriebenen altägyptischen Sprache geht auf arabisch Qobt,
Gamma	T r	g	korrekter Qibt = τηπτιος (Αίγύπτιος) zurück. Die koptische Sprache, deren Literatur überwiegend biblisch und kirchlich ist,
Dalda	A A	d	liegt uns in mehreren Dialekten vor, den oberägyptischen des Sahidischen, Achmimischen, Fayumischen, Memphitischen und
Ei	\mathbf{e}	$e [\check{e}]$	dem unterägyptischen des Bohairischen (vgl. die unten mitge-
Sīta	3.3	z	Das Alphabet besteht aus 31 Buchstaben, die von links nach
Hîda	Ня	$ar{e}$	rechts geschrieben werden wie die griechischen, die den Grund- bestand des Alphabets bilden. Die letzten 7 Buchstaben sind der
Tīda	0 0	th[t+h]	demotischen Schrift entnommen, und zwar liegt zugrunde dem
Jōda	I ,	i [i,j]	ug demot. $\mathcal{Y}\left(\frac{\tilde{\mathbf{x}}_{-}}{\tilde{\mathbf{x}}_{-}}\right);$ q demot. $\mathcal{Y}\left(\frac{\tilde{\mathbf{x}}_{-}}{\tilde{\mathbf{x}}_{-}}\right);$
Kappa	K R	k	s demot. \mathcal{L} (\mathcal{L}); \mathcal{L} demot. \mathcal{L} (\mathcal{L});
Lōla	λλ	l	\simeq dem. \downarrow $\left(\underbrace{\downarrow} \right)$; σ dem. \frown $\left(\bigcirc \right)$; \uparrow dem. \frown $\left(\bigcirc \right)$.
Mi	II m	m	Das letzte Zeichen † ist also keine Ligatur aus T mit aufgesetztem ; es drückt den Doppellaut si aus. Der Buchstabe S
Ni	Ни	n	kommt nur im Bohairischen vor; ihm entspricht im Achmimischen 2 (aus 2 mit diakritischem Strich).
Exi	Z Z	x [ks]	Außer den vollen Vokalen und Diphthongen (die im Gegen- satz zu den Hieroglyphen mit Hilfe der griechischen Buchstaben
О	00	o [ŏ]	geschrieben werden) kennt das Koptische noch einen unbe- stimmten Murmelvokal. , den man im Sahidischen mit einem
Bi	Ππ	p	über den Konsonanten gesetzten – bezeichnet (im Bohairischen mit , im Fayumischen mit , [vgl. Textproben]): $\overline{p} \cdot r$; $\overline{cn} \cdot s^n$.
Ro	Рp	r	Die Zahlen, die im Sahidischen meist ausgeschrieben werden, bezeichnet das Bohairische mit den griechischen Buchstaben-
Sīma	C c	8	ziffern (darunter $\overline{c} = 6$; $A = 900$).
Dau	Тт	t	Abkürzungen. Für eine Anzahl häufiger griechischer Wörter (an denen die koptischen Texte reich sind) treten Abkürzungen ein:
He	ΥΥ	y [i, y]	ic für ικογς; χc für χριστος; ππα für ππεγμα u.a.m. Worttrennung und Interpunktion. Die einzelnen Wörter werden
Phi	Фф	ph[p+h]	l 1 1 1 1 1 True la
Chi	$\mathbf{x}_{\mathbf{x}}$	kh[k+h]	1.0 and in day Cabrick assessmentation was enread-
Epsi	Ψψ	p_s	dann noch durch Bindestriche getrennt werden können. Als Interpunktionszeichen kommen in Koptischen Handschriften
0	Wω		Punkt., Doppelpunkt:, satztrennendes + u ä. vor. Die neueren Drucke gebrauchen die griechischen keterpunktionen.
Schāi	Щщ	š [sch]	Sprachprobe: Jesaja 5, 18 (in drei Dialekten mit allen Lesezeichen).
Fāi	P P	f	Sahidisch: ογοι πια τοωκ πια τισολε πος πογιογε ευμοι αγω πια του πια του πογιογε πια ελευ πίλα ερευ.
Chāi	b s	b $[ch]$	Fayumisch: ordi ninetcwa nnernahi nteh nnornore
Hori	8 9	h	εμπος αγω πεγαπομια ήτρη μπμογο ήτε ογπεριές μπερου.
Dschandscha		$d\check{z}[dsch]$	Βοβαίτισος: ογοι τίπη ετεωκι πιποδι πωογ μφρη † πογπος εqιμηση πεω πιωποωια μφρη † ιπεωκ μφηας βες
Schima	6 6	$\check{c}[g,tsch]$	πογβαρει.
Di	T 4	ti	Steindorff, Abriß der koptischen Grammatik (Berlin 1921). Spiegelberg, Koptisches Handwörterbuch (Heidelberg 1921).
	<u> </u>		

Figure 3. The Coptic alphabet, from the Reichsdruckerei 1924. The three forms of the *jinkim* are shown, as is the sentence divider.

TABLE 22.10: The Sahidic Coptic Alphabet

Letter	Transcription	Phonetic Value	Name ^a	Greek/Demoi Source
λ	a	[a]	alpha	A
В	v	[v]	vēta	В
Г	k	[k]	kamma	Γ
Δ	t	[t]	talta	Δ
ε	e	[e]	ei	E
Z	S	[s]	sata	Z
н	ē	[e:]	hata	Н
θ	th	[th]	thita	Θ
I	i	[i]	iōta	I
K	k	[k]	kappa	K
Λ.	1	[1]	lauta/lole	Λ
. М	m	[m]	mē/me/mi	M
N .	n	[n]	ne	N
3	ks	[ks]	ksi	Ξ
ο	o	[o]	ou	0
π	p	[p]	pi	П
P	r	[r]	rō	P
С	s	[s]	sēmma	Σ
т	t	[t]	tau	T
Υ	u	[u]	he/ue	Y
ф	ph	[ph]	phi	Φ
x	kh	[kh]	khi	X
Ψ.	ps	[ps]	psi	Ψ
ω	ō	[o:]	ō	ω
ψ	š	[ʃ]	šai	3
q	f	[f]	fai	۶
2	h	[h]	hori	9.
X	₫	[ʤ]	dandia	٦٦
б	q	[q]	qima	L
†	ti	[ti]	ti	+

a. Examples collected in Crum 1939, and cf. Vergote 1973, vol. 1a: 7, citing Worrell 1942; and Kasser 1991.

Figure 4. The Coptic alphabet and names from Ritner 1996.

Let.	Num.		Names	
1	1	Da	алфа	alpha
2	2	Bß	вита, віда	bêta, vida
3	3	22	Vanua .	gamma
4	4	$\Delta \lambda$	<u> </u>	dlda, dalda
5	5	ϵ_{ϵ}	ei, eie	ei, eie
	6	23	C00 Y	soou
6	7	ζζ	ረዘ ፕል,	zêta, zita, zata
7	8	Нн	гнта, нта, гіта, гат€	hêta, êta, hita, hate
8	9	$\Theta \Phi$	θ нта, θ гта, θ \in θ \in	thêta, thita, thethe
9	10	lı .	ιωτα, 10τα, ιαγΔα	iôta, iota, iauda
10	20	\mathbf{K}_{K}	каппа, капа	kappa, kapa
11	30	$\lambda\lambda$	λ a γ λ a, λ a γ λ a, λ o λ ϵ	lauda, laula, lole
12	40	Uu	uh, u€, ui	mê, me, mi
13	50	Nи	ne, nnh, n	ne, nnê, n
14	60	Z .₹	₹।	ksi
15	70	Oo	0 ° , 0	ou, o
16	80	Ππ	πι	pi
17	100	Pp	pw, &po, po	rô, hro, ro
18	200	Cc	CHUUA, CYUUA, CIUA, CUUA	sêmma, summa, sima, smma
19	300	Tт	ፕ ልኖ	tau
20	400	$\Upsilon_{\mathbf{Y}}$	&e, Te, Ta	he, ue, ua
21	500	$\Phi\Phi$	Фі	phi
22	600	$\mathbf{x}_{\mathbf{x}}$	χı	khi
23	700	$\mathbf{\Phi}\Psi$	Ψι	psi
24	800	$\mathbf{w}_{\mathbf{w}}$	w, at, wot	ô, au, ôou
25		\mathfrak{M} ത	ພູຂາ, ພູຍາ	šai, šei
26	90	વુ	વઢા, વદા	fai, fei
27		ક ેશ્	<u> </u>	[unattested]
28		Ъ₫	žei, žai	hei, hai
29		ઙ ક	Sobi	hori
30		Χx	xanxia, xenxe	janjia, jenje
31		6 °6	бıна	čima
		¼ †	गा, †ा, †हा, गहा, गम, ग ह	ti, <i>ti</i> i, <i>ti</i> ei, tei, tê, te
	900	₽p	· ·	

Figure 5. The Coptic alphabet. The first column indicates the letter order (for example, PHI is the twenty-first letter of the alphabet) as stated in Crum. The second column gives the numeric value of the letters. Transliterated character names taken from Coptic dialects as given in Crum 2000 are used for the Coptic letters. Note especially that Crum distinguishes AKHMIMIC KHEI, KHEI, and HORI in different entries in his dictionary.

```
є- еро» Richtung: zu, in, an,
                              € Umstandssatz 329
                              6= Präsens II 303; mit ΦAN-
 gegen; bezeichnet
                       den
 Schuldner, den Dativ 237; je
                               Konditionalis 447, 448; mit
                                NA Futurum II 306; mit e
 109 B; Objekt 263; beim
 Vergleich 152, 153; mit
                               Futurum III 308
 Inf. 337, 338, 342, 345,
                              6- Relativpronomen 462.
 346 B, 367 (50a)
                              -6 Suffix der 2. sg. f. 185, 190
e- Umstandssatz 328;
 II. Tempora 334 (52a)
```

Figure 6. Sample text from Till's grammar showing the use of the DOUBLE HYPHEN, distinguished from the ordinary single hyphen.

(300) Das koptische Alphabet im Vergleich zur griechischen Unzialschrift

		8	8	
Zahlen- wert	Laut- wert	Name	Koptisch	Griech. Unziale
1	а	alfa	3.	a
2	b, v	vēda	В	В
3	g	gamma	r	r
4	d	dalda	Δ	S
5	ē	ēje	ϵ	ϵ .
6	-	sou	3	C,
7	z	zāda	3	Z
8	ē	hāda	Н	Н
9	t-h	tutte	θ	ө
10	j, i	jōda	ı	1
20	k	kabba	ĸ	K
30	1	lōla	λ	λ
40	m	mēj	u	M
50	n	ni	H	N
60	ks	eksi	Z	3
70	ŏ	ou	0	0
80	p	bej	Π	π
100	r	rou	Р	Р
200	s	samma	C	C
300	t	daū	T	T
400	i	he	r	Y
500	p-h	fīj	Ф	ф
600	k-h	kij	\boldsymbol{x}	×
700	ps	ebsi	क	*
800	ō	ō	ш	w
900	-	_	₽,₽	

Figure 7. Chart showing Coptic with the Greek uncial forms from which they were derived, from Haarmann 1990.

80 10 насы не нішаже сонп' ентаїс стоно того аты ацераїсот нізі дідтмос

12 їотдас ошмас (1) аты пежац же пе таре соєрминсіа инесіщаже уна

14 жі тпе ан мітмот (2) пеже іс мітреу до нізі пет'щіне су'щіне щантеу 16 жіне аты ротан' сущанізтне уна штртр аты сущаніштортр унар

18 (blank) ішпире аты унар рро ежм птиру (3) пеже іс же стіща 20 хоос инти нізі нет'сык рит' титти же сіс ринте ст'митеро ри тпе е 22 сіє прадит' нар щорп' єршти ите

Figure 8. Sample text from an edition of the Gospel of Thomas showing use of *jinkim*, diaeresis, apostrophe (drawn as a grave-like stroke), and middle dot.

5. Accentuation. L'écriture copte emploie surtout l'accent grave (ou un point) et des traits horizontaux suscrits au-dessus de certaines lettres; c'est pour indiquer qu'une voyelle brève ou auxiliaire є, correspondant à notre e accentué ou atone, est sous-entendue par l'écriture.

Dans ce cas, le bohaïrique emploie un accent grave ` (ou un point), au-dessus de la première consonne des mots : ùkaz (ou ikaz) pour eukaz *mkah.

Toutesois, dans le corps du mot, le bohaïrique écrit la voyelle brève ϵ .

Tous les autres dialectes remplacent l'accent ' ou ' par un trait horizontal T tracé au-dessus de la première consonne : PNOBE pour EDNOBE ernobe.

Dans le corps du mot, le sa'idique indique aussi l'e accentué ou atone par le même trait entre les deux consonnes et au-dessus d'elles : Zuue pour hémme; cwTu pour sōtém; wNZ pour ōnéh; enex pupu pour epekrémrém.

Les règles d'accentuation sont souvent hors d'usage, bon nombre de manuscrits n'étant pas accentués.

- 6. Le trait horizontal suscrit est encore employé :
- 1° Pour indiquer que certains mots très fréquents, coptes ou grecs, sont écrits en abrégé :

```
      φt, φt pour φνογτ «Dieu».
      πc pour πσοις «le Seigneur».

      σc pour θεος «Dieu».
      κε pour κγριε «Seigneur».

      ιc, iπς pour incoγς «Jésus».
      πνα pour πνεγμα «esprit».

      κε pour κγριε «Seigneur».
      πνα pour πνεγμα «esprit».
```

Figure 9. Text from Fossey 1948 describing the use of the *jinkim* and the mark of abbreviation.

Für häufige nomina sacra werden Kurzschreibungen (Zusammenziehungen) verwendet, die durch einen waagrechten Strich darüber als solche gekennzeichnet sind. Die gebräuchlichsten sind: $\overline{A}\overline{A}\overline{A} = A\lambda\gamma GIA$, \overline{IC} = IHCOYC, $\overline{IH}\overline{\lambda} = ICP\lambda H\lambda$, $\Theta\overline{I}\overline{\lambda}\overline{H}\overline{M} = \ThetaIGPOYC\lambda\lambda HM$, $\overline{IN}\overline{\lambda} = \Pi NGYM\lambda$, $CPO\overline{C} = CT\lambda\gamma POC$, $\overline{C}\overline{W}P = CWTHP$. $\overline{XC} = XPICTOC$, $\overline{XC} = XOGIC$.

Figure 10. Text from Till 1966 showing abbreviations, including the use of C

wöhnlich sind sie durch einen waagrechten Strich darüber als Zahlen gekennzeichnet. Die Zahlwerte sind: $\lambda = 1$, B = 2, $\Gamma = 3$, $\lambda = 4$, $\epsilon = 5$, $\varsigma = 6$, z = 7, H = 8, $\theta = 9$, I = 10, K = 20, $\lambda = 30$, M = 40, N = 50, z = 60, O = 70, D = 80, P = 100, C = 200, D = 300, D = 400, D = 500, D = 600, D = 600

Die Brüche, deren Zähler mit ganz wenigen Ausnahmen stets 1 ist, werden dadurch bezeichnet, daß der Zahl für den Nenner ein schräger Strich rechts oben beigefügt wird: $\Gamma' = \frac{1}{3}$. Ausnahmen: $f = \frac{1}{2}$.

Figure 11. Coptic numerals from Till 1966.

ī =	1	$\vec{K}_{\cdot} = 20$	$\bar{\mathbf{T}} = 300$
B ===		$\overline{\lambda} = 30$	7 = 400
<u>r</u> =		$\overline{\mathbf{u}} = 40$	$\overline{\phi} = 500$
$\bar{\lambda} =$	_	$\vec{N} = 50$	x = 600
$\overline{\epsilon} =$		$\overline{3} = 60$	₹ = 700
= 3		$\overline{0} = 70$	$\overline{\mathbf{w}} = 800$
z =		$\vec{n} = 80$	† = 900
H =	•	$\bar{\mathbf{q}} = 90$	<u>a</u> = 1.000
- =		$\bar{p} = 100$	$\overline{\overline{B}} = 2.000$
ī =	· ·	$\overline{c} = 200$	= 10.000

7. Dans la grammaire: on écrit l'accent aigu ' pour indiquer la voyelle formative qui porte l'accent tonique: ρώμι, κόκτε, νοήτ.

Dans les grammaires et les vocabulaires, on use aussi de traits d'union dans les mots composés et pour séparer préfixes et suffixes de la racine. Le mot copte à l'état construit, c'est-à-dire uni à son complément, est suivi d'un trait d'union, peu-, qui le réunira à son complément éventuel; le mot à l'état pronominal, c'est-à-dire accompagné d'un pronom, est suivi de deux traits obliques, por, qui annoncent un pronom éventuel.

8. Certains manuscrits coptes sont ponctués. La ponctuation des textes imprimés est généralement ajoutée par l'éditeur.

Consulter: A. Mallon, Grammaire copte, Beyrouth, 1907.

TEXTE COPTE.

Фаі півіос и нец ніполітеїа итє пеніют євотав шплатофорос (1) авва венофер піанахюрнтнє є єтархюк євол исот $\overline{18}$ (2) шпіавот пашні рен отвеїрнин итє ф $\overline{1}$ ашни. Д ототон наінахюрнтнє $\overline{1}$ (1) Abrégé pour шписчатофорос. — (2) 16. — « Mots grecs.

Abrege pour unnerual ropopoc. — 10. — mois gices.

Figure 12. Text from Fossey 1948 describing the use of the horizonal stroke over numerals and the double horizontal stroke over thousands. The text also describes the use of the acute accent to indicate tonic accent, and the distinction between hyphen and DOUBLE HYPHEN.

COPTE MEMPHITIQUE.

LETTRES ALPHABÉTIQUES.

MAJUSCULES. MINUSCULES. A B P T A G F M A H в н у взак т д в в і к х и н д о п тку жи хои рс **Р** С Т Ү Ф **х** Ψ **ш у** ρ ω ψ χ ψ κ κτΓ т **д х з ч** SIGNES DE NUMÉRATION. و بو و که ه د د س له ما ۱ ط و ۱ ۶ و س س د 此女子业 V= 2 € 5 3 b € LETTRES ACCENTUÉES, LIGATURE ET SIGNES DIVERS. \sim , \sim \dot{s} \dot{s} \dot{s} \dot{s} \dot{s} \dot{s} \dot{s} COPTE THÉBAIN. LETTRES ALPHABÉTIQUES. $x \psi u u u d x c +$ ABRÉVIATIONS, LIGATURES ET MONOGRAMMES. ACCENTS ET SIGNES DIVERS.

Figure 13. The Coptic alphabet, from the Imprimerie de l'Institut Français d'Archéologie Orientale.

TABLE XX - Row 2C: COPTIC

	2C8	2C9	2CA	2CB	2CC	2055
0	A	Π	a	π		=
1	В	P	В	p		
2	2	C	5	С		
3	$\mathbf{\lambda}$	T	λ	Л		
4	ϵ	r	€	r		
5	£,	Ф	٤	Ф		
6	ζ	X	ζ	×		
7	Н	4	Н	Ψ		
8	θ	w	0	w		G = 00 P = 00
9	1	ક	ı	ર		
Α	K	₽	κ	p		
В	A		y	.p.		
С	u		u	*		
D	И		N	cic		
Е	Z.		ξ		÷	
F	0		0		5	

TABLE XX - Row 2C: COPTIC

hex	Name	hex	Name
801 823 845 667 88 88 88 88 88 89 99 99 99 99 99 99 99	COPTIC CAPITAL LETTER ALFA COPTIC CAPITAL LETTER VIDA COPTIC CAPITAL LETTER BALDA COPTIC CAPITAL LETTER BALDA COPTIC CAPITAL LETTER BEIE COPTIC CAPITAL LETTER SU COPTIC CAPITAL LETTER SU COPTIC CAPITAL LETTER TATE COPTIC CAPITAL LETTER ALTE COPTIC CAPITAL LETTER HATE COPTIC CAPITAL LETTER HATE COPTIC CAPITAL LETTER HAUDA COPTIC CAPITAL LETTER MIDA COPTIC CAPITAL LETTER MIDA COPTIC CAPITAL LETTER MIDA COPTIC CAPITAL LETTER MI COPTIC CAPITAL LETTER MI COPTIC CAPITAL LETTER MI COPTIC CAPITAL LETTER NI COPTIC CAPITAL LETTER NI COPTIC CAPITAL LETTER NI COPTIC CAPITAL LETTER SI COPTIC CAPITAL LETTER PI COPTIC CAPITAL LETTER SI COPTIC SMALL LETTER SI COPTIC SMA	2055	DOUBLE HYPHEN