

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

L2/03-338

Doc Type: Working Group Document**Title:** Further discussion of the DOUBLE OBLIQUE HYPHEN**Source:** Michael Everson**Status:** Individual Contribution**Action:** For consideration by JTC1/SC2/WG2 and UTC**Date:** 2003-10-13

Continued e-mail discussion of the character DOUBLE OBLIQUE HYPHEN proposed in N2611, N2636, and N2639 has resulted in the present document. It has been urged that a document be produced which will address certain implementation concerns.

Semiticists and Cuneiformists familiar with German scholarship were aware that the glyph for U+002D HYPHEN-MINUS which is used in Fraktur typography differs from the glyph used in Roman typography. They have made use of this glyph to indicate certain morphological boundaries while continuing to use the ordinary hyphen to indicate other boundaries. Certainly they were distinguishing between the two characters. Research has shown that they also make use of the ordinary U+003D EQUALS SIGN in other contexts in their documents.

It is interesting to note that other linguists have made use of U+003D EQUALS SIGN to indicate the same kind of morphological boundary. It was suggested that the U+003D EQUALS SIGN be pressed into service with a non-plain text glyph variant in order to represent the Fraktur borrowing, but this is not satisfactory. The EQUALS SIGN has at least variant already (the longer version used in some mathematical typography), and the suggestion that it has adopted an alternate “Copto-Hittite” variant seems to many to be a misunderstanding of the underlying character identity, in addition to introducing a glyph variant for that character which would not be acceptable to anyone in a mathematical context. I suggest that it is the advent of computers which had equals signs and the abandonment of lead typefaces which had Fraktur double oblique hyphens in them might have something to do with conflation of the two *characters*.

Accordingly it is proposed to add a new character: U+xxxx DOUBLE OBLIQUE HYPHEN. This character is not intended to be used for any instance of an EQUALS SIGN which would be rendered =, even if the linguistic context might otherwise be similar or identical (in other words, there is a glyph-based distinction for what might otherwise be the same semantic function – comparable to the way modifier letters are treated). This character is also not intended to be used for instances of a Fraktur style hyphen, which should be encoded with U+002D, instead.

The repertoire of relevant glyphs is as follows:

- | | | |
|---|----|--|
| A | - | short, single-bar |
| B | ≠ | short, double-bar, uptilted |
| C | – | medium, single-bar |
| D | = | medium, double-bar |
| E | ≡ | medium, double-bar, uptilted |
| F | — | medium+, single-bar (shorter than em dash) |
| G | == | medium+, double-bar |

The repertoire of relevant characters is as follows:

1	U+002D	HYPHEN-MINUS
2	U+2010	HYPHEN
3	U+2013	EN DASH
4	U+2212	MINUS SIGN
5	U+003D	EQUALS SIGN
6	U+XXXX	DOUBLE OBLIQUE HYPHEN

Rendering rules:

1. U+002D HYPHEN-MINUS is ambiguous between use as a hyphen (A: -) and use as a minus sign (C: –), and may be rendered as either of those, depending on use and context.
2. U+2010 HYPHEN is normally rendered with glyph A (-). In Fraktur style it is rendered with glyph B (¨). In some fallback contexts it is rendered with glyph C (–).
3. U+2013 EN DASH is a fixed width dash, and should normally be rendered with glyph C (–). In some European typographic styles which emphasize the length of dashes as opposed to hyphens, it may end up rendered with glyph F (—) – but still shorter than an emdash or quotation dash.
4. U+2212 MINUS SIGN is normally rendered with glyph C (–). In fonts which exaggerate mathematical operator size, it may be rendered with glyph F (—).
5. U+003D EQUALS SIGN is normally rendered with glyph D (=). In fonts which exaggerate mathematical operator size, it may be rendered with glyph G (=), and in any case should normally have the same width as the minus sign.
6. U+XXXX DOUBLE OBLIQUE HYPHEN is normally rendered with glyph B (¨). In certain typographic contexts (e.g. the French Coptic grammar), it may be rendered with glyph E (≡) – but this is rather old-fashioned..

Unicode Character Properties

2055;DOUBLE OBLIQUE HYPHEN;Pd;0;ON;;;;;N;;;;;

Bibliography

- Fossey, Charles. 1948. *Notices sur les caractères étrangers anciens et modernes rédigées par un groupe e savants*. Nouvelle édition mise à jour à l’occasion du 21^e Congrès des Orientalistes. Paris: Imprimerie Nationale de France.
- Güterbock, Hans G. & Harry A. Hoffner, eds. 1994. *The Hittite dictionary of the Oriental Insitute of the University of Chicago*. Chicago: Oriental Insitute of the University of Chicago.
- Haarmann, Harald. 1990. *Die Universalgeschichte der Schrift*. Frankfurt: Campus Verlag. ISBN 3-593-34346-0
- Plumley, J. Martin. 1948. *An introductory Coptic grammar (Sahidic dialect)*. London: Home & van Thal.
- Seidel, A. [s.d]. *Praktische Grammatiken der Hauptsprachen Deutsch-Südwestafrikas*. I. Nama (Sprache der Nama-Hottentotten). II. Otyiherero (Sprache der Herero). III. Oshindonga (Sprache des Ndonga-Stammes der Ovambo). Wien und Leipzig: A. Hertleben’s Verlag.

- Starke, Frank. 1977. *Die Funktionen der dimensionalen Kasus und Adverbien im Althethitischen*. (Studien zu den Boğazköy-Texten; 23) Wiesbaden: Otto Harrassowitz.
- Till, Walter C. 1966. *Koptische Grammatik: saïdischer Dialekt, mit Bibliographie, Lesestücken und Wörterverzeichnissen*. Leipzig: VEB Verlag Enzyklopädie.

Examples

Um die Erforschung des Nama hat sich besonders der Missionar Krönlein verdient gemacht. Auch Kolbe, Knudsen, Tindall, Wallmann u. a. haben nicht wenig dazu beigetragen. Wir besitzen heute ein vorzügliches Wörterbuch des Nama (J. G. Krönlein, Wortschatz der Khoi-Khoi, Berlin 1889), einige grammatische Skizzen (Wallmann, die Formenlehre der Namaquaisprache, Berlin 1857; H. Tindall, a grammar and vocabulary of the Namaqua-Hottentot language, Cape-Town 1857), sowie Übersetzungen des Neuen Testaments (von Krönlein), der Evangelien (von H. Schmelen), des Evangelium Lucae (von G. H. Knudsen), der Psalmen (von Krönlein), der Salwer biblischen Geschichte, der Agende und des luther-

Figure 1. In this example from Seidel the HYPHEN is used in both the Fraktur font (≡) and the Roman font (-). Here it is clear that the same character (U+002D) is being used with only a font distinction. It is also used as a connector in *Khoi-Khoi* and in *Cape-Town*.

Neuerdings hat Brincker auch den zweiten Band seines Wörterbuches (Deutsch-Othiherero-Oshikuanjama) veröffentlicht und der Missionär Wiehe uns mit einer neuen, vorzüglichen Darstellung der Grammatik beschenkt. Auf Grundlage dieser Vorarbeiten ist das vorliegende Büchlein verfaßt worden. Es bietet in Grammatik und Wörterverzeichnis natürlich nur das Nötigste.

Figure 2. In this example from Seidel the HYPHEN is used at a linebreak in *Wörterbuch*. Both the HYPHEN (drawn in Fraktur as a double hyphen) and a short horizontal dash with extra space on either side are used in *Deutsch-Othiherero-Oshikuanjama*. Today, in English, we encode such a distinction in Roman as *Deutsch-Otyiherero-Oshikuanjama*, where the EN DASH serves to separate the first element from the HYPHENATED compound second element; this use of EN DASH is not known in modern German.

Schmerzen an der Hand, am Fuß, Leibschmerzen, Zahnschmerzen; mba t' esura, ich habe den Schnupfen, usw. — za nu, Imperfektum von nua — azehe (vgl. E, 11, b) — ndyi kohe, Konjunktiv — k' = ku, nach — vaza, gelangen — pi, wohin? — nana (Adverb), gut — pe na, es gibt, es ist da; kape na, es gibt nicht, es ist nicht da — er hat sich Hereroart angezogen, d. h. er trägt Herero-Kleider. Ebenso otyirumbu, europäische Kleider — roye, dein (e-) — oandye, meine (oma-) — wie sagt (meint) ihr? — mua hapere, darin sind gewachsen — hana, abbrechen, abreißen — yandye, mein (o-) — omai, Plural von ei — hinga, wehen — omasa (Plural), die Kraft; n' omasa (mit Kraft), kräftig, heftig — okutyanga, Infinitiv — kutya am Anfange eines indirekten Frageatzes (vgl. oben) — rukuru, längst — zo na, sie haben (die ozo-) — karisa = ka + risa, geh' und weidel — ndyi, mir — eterera, doppelt relative Form von eta (vgl. D, 12, e) — yandye meine (o-) — tauka, zerreißen (intr.).

Figure 3. In this example from Seidel the HYPHEN (drawn in Fraktur as a double hyphen) is uses in word-breaking, and is also used (drawn in Roman as a single hyphen) to show morphological boundaries. The EM DASH is used to link translations, and the EQUALS SIGN is used to show equivalence.

ε- ερ- *Richtung: zu, in, an, gegen; bezeichnet den Schuldner, den Dativ* 237; je 109 B; *Objekt* 263; *beim Vergleich* 152, 153; *mit Inf.* 337, 338, 342, 345, 346 B, 367 (50 a)
 ε- *Umstandssatz* 328;
 II. *Tempora* 334 (52 a)

ε- *Umstandssatz* 329
 ε- *Präsens* II 303; *mit ὡς- Konditionalis* 447, 448; *mit να Futurum* II 306; *mit ε Futurum* III 308
 ε- *Relativpronomen* 462
 -ε *Suffix der 2. sg. f.* 185, 190

Figure 4. Sample text from Till's grammar showing the use of the DOUBLE OBLIQUE HYPHEN, distinguished from the ordinary single hyphen.

ā = 1	k̄ = 20	t̄ = 300
b̄ = 2	l̄ = 30	γ̄ = 400
ḡ = 3	ū = 40	φ̄ = 500
ā = 4	n̄ = 50	χ̄ = 600
ē = 5	z̄ = 60	ψ̄ = 700
ē = 6	ō = 70	ω̄ = 800
z̄ = 7	p̄ = 80	† = 900
h̄ = 8	q̄ = 90	ā = 1.000
ē = 9	p̄ = 100	b̄ = 2.000
ī = 10	c̄ = 200	ī = 10.000

7. Dans la *grammaire* : on écrit l'accent aigu ' pour indiquer la voyelle formative qui porte l'accent tonique : ρῶμι, κόλσα, νοῦ†.

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, νοα=, 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.

ΦΔΙ ΠΙΒΙΟC^a ΝΕΥ ΝΙΠΟΛΙΤΕΙΑ^a ΝΤΕ ΠΕΝΙΩΤ ΕΘΟΥΑΒ ΑΠΝΑΤΟΦΟΡΟC⁽¹⁾
 ΑΒΒΑ ΒΕΝΟΦΕΡ ΠΑΝΑΧΩΡΗΤΗΣ^a ΕΤΑΦΧΑΚ ΕΒΟΛ ΝCΟΥ ΙΕ⁽²⁾ ΑΠΙΑΒΟΤ
 ΠΑΩΝΙ ΞΕΝ ΟΥΖΕΙΡΗΝΗ^a ΝΤΕ Φ† ΑΜΗΝ. Λ ΟΥΟΥΟΝ ΝΑΙΝΑΧΩΡΗΤΗΣ^a

⁽¹⁾ Abrégé pour ΑΠΝΕΥΜΑΤΟΦΟΡΟC. — ⁽²⁾ 16. — ^a Mots grecs.

Figure 5. Text from Fossey 1948 describing the use of the horizontal 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 (*trait d'union* 'stroke of union') and DOUBLE OBLIQUE HYPHEN (*deux traits obliques* 'two oblique strokes').

.....lightface double hyphen at end-of-line is a hyphen
 that should be retained
fistnote (see definition)
 +plus sign in etymology joins words or word elements
 †dagger precedes a death date

Figure 6. The DOUBLE OBLIQUE HYPHEN with its use described in the Webster's Third New International Dictionary.

(earlier apparently they beat a stick; then) [...
^{GIŠ}*p*]*a-a-ḫi-ša UL walḫ[zi ...] nu=ššan 1* ^{GIŠ}*pa-aḫ-*
ḫi-ša x [... (x x-an-zi)] / [maḫḫan=ma=aš] (var.
 GIM-[an ...]) *parkuiš-zi* ^{GIŠ}*pa-aḫ-ḫi-ša dān wal~*
 [(ḫan)=i] “[...] he does not beat the *p.*; [...] ...
 thereupon they ... one *p.*; [but] when [he(?)] is
 pure, they beat the *p.* a second time” Bo 4951 rev. 3-5
 (birth rit.), w. dupl. KBo 17.68:2. ed. StBoT 29:124f. On the
 basis of Bo 4951, Ehelolf in OLZ 32:322f. corrected his earlier
 reading of the word as ^{GIŠ}“PA”-*aḫ-ḫi-ša* to ^{GIŠ}*pa-aḫ-ḫi-ša*.

^{GIŠ}*p.* in KBo 22.135 i 3 occurs in a rit. performed
 by a ^{LU}AZU “exorcist”; the other refs. occur in birth
 rituals. In view of the use of *walḫ-* elsewhere for
 playing percussion instruments, one should consid-
 er the possibility that *p.* was a noisemaker of some
 kind, made (partly?) of wood. Starke (StBoT 31:208)
 posits a Luwian stem *paḫḫit-* and translates “Trom-
 mel” or “Gong.”

Sommer/Ehelolf, Pap. (1924) 76; Ehelolf, OLZ 32 (1929)
 322f.; Friedrich, HW (1952) 153 (“Stock(?)”, Gerte(?)); Beck-
 man, StBoT 29 (1983) 122, 288 (tr. “stick(?)” employed only
 provisionally).

paḫḫeški- iter. v.; (a hostile action).†

Sum. [ḪAR] = Akk. [*a-r*]*a-rum* “to mill, grind”
 = Hitt. *ḫūwarzakiuwar* “repeated cursing” (tr. of
 Akk. *arāru* “to curse”) / Sum. [ḪAR] = (Akk.) [†]*zaḫ-*
a-u “resin(??)” = (Hitt.) *pa-aḫ-ḫe-eš-ki-u-wa-ar*

Figure 7. Example in Roman type from Güterbock & Hoffman 1994 showing the DOUBLE OBLIQUE HYPHEN distinguished from the HYPHEN as well as from the EQUALS SIGN.

dā- „nehmen“
iššaz-šmit „aus ihrem Mund“ (StBoT 8 I 18’); *kiššaraz-šmet* „aus ihren Hän-
 den“ (StBoT 8 II 38); *ḫalpašuiittaz* „vom Thronsessel“, *ḫaššaz* „vom Herd“
 (StBoT 8 II 39); ^{GIŠ}ŠÚ.A-*kaz-šmit* „von ihrem Sitz“, *kartaz-šmit* „aus ihrem
 Herzen“, [*tug*]*gaz-šmit* „aus ihrem Körper“ (KBo XVII 7 + = StBoT 8 Rs.IV’
 6’-7’); ^{GIŠ}BANŠUR-*az* „vom Tisch“ (KBo VI 2 II 44’ = HG I, § 47a);
 A.AB.BA-*az* „aus dem Meere“ (KBo XXII 2 Vs. 4); *ḫalenḫuaz* „aus dem
 ḫ.-Haus“ (KBo XVII 11 I 17’ = StBoT 12 I 33).

Figure 8. Example from Starke 1977 showing the DOUBLE OBLIQUE HYPHEN distinguished from the HYPHEN as well as from the EQUALS SIGN.

Note in the Dictionary and in most Grammars the hyphen is used to show at a glance the forms of verbs and prepositions which occur before a noun or pronoun.

Thus the Construct form of a verb or preposition before another noun is printed with a single hyphen eg.

$\bar{p}-$, $\pi\epsilon\chi\epsilon-$, $\epsilon-$, $\bar{\nu}-$ etc.

When the form is that used with Pronoun Suffixes (the Pronominal Form) a double hyphen is used. eg. $\alpha\alpha^*$, $\pi\epsilon\chi\alpha^*$, $\epsilon\phi\alpha^*$, $\bar{\nu}\mu\alpha^*$ etc.

Figure 9. Plumley 1948 describes the use of the HYPHEN and the DOUBLE OBLIQUE HYPHEN.