L2/03-347

Doc Type: Working Group Document
Title: Background on DOUBLE OBLIQUE HYPHEN
Doc. Type: Individual Contribution
Source: Deborah Anderson
Action: For consideration by JTC1/SC2WG2, UTC
Date: October 19, 2003

I. Background

The DOUBLE OBLIQUE HYPHEN (="double hyphen" below) has been discussed in several earlier documents by Michael Everson (N2647, N2611, N2636 and N2639), as well as on email. More detailed research into the specific use of the double hyphen in certain languages is provided below (under section II), after contact with various scholars in the user communities in October 2003. A summary with specific comments on the use of this character appears in section III, and a list of modifications to the rendering rules for the glyphs and characters involved appear in section IV.

II. Information on Usage of DOUBLE OBLIQUE HYPHEN

Coptic

Based on information from Stephen Emmel, a representative of the International Association for Coptic Studies and himself a prominent Coptologist at the Institut für Ägyptologie und Koptologie, University of Münster, the double hyphen sign is standard in the grammatical analysis of Coptic. It was apparently first introduced by Ludwig Stern in his *Koptische Grammatik* (Leipzig, 1880) to mark the form of the verb before suffixes, as opposed to the (single, normal) hyphen, which was used to mark a different verb form (i.e., that used before an incorporated nominal object); the hyphen itself already had been used to separate morphs in Coptic grammar by 1830. Coptic also has also prepositional forms and other word classes that normally occur with personal suffixes, and the double hyphen was used to mark all such forms (already by Stern), whether they were verbal or not. As noted by Emmel, "some contemporary Coptic grammarians understand the double hyphen to 'stand for' any member of the paradigm of personal suffixes, but that is not exactly what Stern meant by it."

In general, Coptologists have mostly tried to imitate Stern regarding the shape of this symbol and use a sign very similar to that of a Frakturhyphen. But because such a sign was not always available to printers (and on typewriters), various substitutes have been used, including the equals sign. The example of Fossey 1948 (fig. 5 in N2647, and similar to that in figure 1 below from Mallon) with the medium double bar are not considered ideal, and "look ugly and unfortunate to the experienced Coptologist."

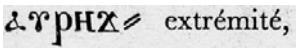

Figure 1. Example of double hyphen from Alexis Mallon, *Grammaire Copte*, 4th ed., Beyrouth: Imprimerie Catholique, 1956, p. 31.

According to S. Emmel, there is no technical standard name for this character today (other than descriptive terms like "slanted equals sign" or "slanted double hyphen") and almost no one knows its true origin (save the above information which S. Emmel has

acquired when working on the Unicode proposal). Emmel is inclined to think that a Frakturhyphen might underlie this sign, perhaps it was used in some branch of German philology in some sense of a hyphen (*Bindestrich*).

Demotic

Demotic uses the double hyphen in transcription of Demotic and other stages of Egyptian to mark a suffixed pronoun to its preceding noun (etc.). It seems to derive from Coptic, according to Janet Johnson, Professor of Egyptology in the Oriental Institute and Department of Near Eastern Languages and Civilizations

In the ongoing Chicago Demotic Dictionary project at the University of Chicago, the double hyphen is used. Two examples are shown below. Note that the equals sign is clearly distinguished from the double hyphen in figure 2.

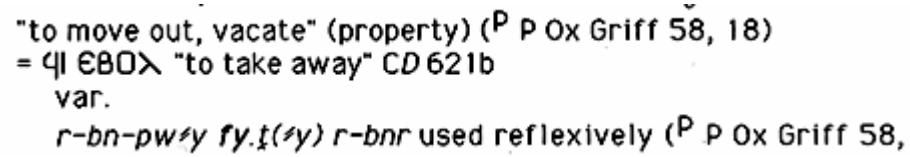

Figure 2. Example of the double hyphen (before *y* in the last line) from a sample page of the Demotic Dictionary Project, Oriental Institute, University of Chicago (Letter *F*, part 3, under the entry for *fy r-bnr* "to carry out, deliver").

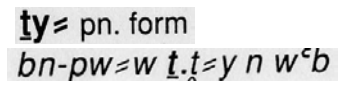

Figure 3. Two further examples of double hyphen from the Demotic Dictionary Project, showing the sign more clearly.

In answer to a question about the use of other signs in place of the double hyphen, Janet Johnson commented, "I have seen an equals sign used instead, but only when a person simply didn't have the appropriate sign in their font. One can get used to almost any convention, but the Frakturhyphen does have real functionality for many Egyptologists."

However, examples with the equals sign were found in a number of other printed sources (see figs. 4 and 5).

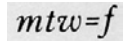

Figure 4. Example of the equals sign from R.S. Simpson (*Demotic Grammar in the Ptolemaic Sacerdotal Decrees*, Oxford, 1996, p. 27)

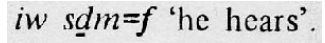

Figure 5. Example of the equals sign from a book on Middle Egyptian, from *Middle Egyptian* by John Callender (Malibu: Undena Publications, 1975, p. 5).

Hittite

a. Chicago Hittite Dictionary (CHD)

The early volume of the *Chicago Hittite Dictionary*, volume *L-N*, used a "Frakturhyphen" for two purposes:

1. to mark a clitic boundary
2. to break a word across a line (i.e., the hyphen was inserted by the printer because the word didn't fit on the line; it is not retained if the word were to be printed on one line)

The name "Frakturhyphen" for this sign appears explicitly in the Preface to the *CHD* volume *P* (1997). When asked who originally proposed using a "Frakturhyphen" in this dictionary, one of the original editors, Harry Hoffner, responded:

> "Certainly, with his background Hans Gueterbock [the other original editor of the *CHD*] was quite familiar with the Fraktur. But whether he or I first suggested it, I no longer recall. We wanted to avoid the simple hyphen for marking clitic boundaries in broad transcription, since this in some cases might lead to the mistaken impression that we were giving a word in transliteration. An equal sign might have accomplished the same end, but we were advised against it by our printers for aesthetic reasons. You may be aware that in good printer technique a longer (wider) hyphen is used to join numbers (e.g. in a date) than the ordinary hyphen. Such considerations led our printer not to favor an ordinary equal sign. And since we would have had to 'invent' a narrower equal sign to satisfy him, we decided to use the Fraktur instead. … If the Unicode Commission can oblige us with the character, that would be very convenient. If they choose not to, we will either continue its use by some other means or employ the equal sign instead." (from an email correspondence, October 9, 2003)

In the following example (figure 6), the double hyphen that appears on the first line (breaking up *uhiškanaz*) is not an example of a clitic, but rather is used as a word-break, much as a Frakturhyphen would act in a text set in Fraktur when a word needed to be broken across a line. If there had been adequate space to accommodate the full length of this word, no double hyphen would have been used. In the third line below, the double hyphen appears before the clitics *-an* and *-mu*.

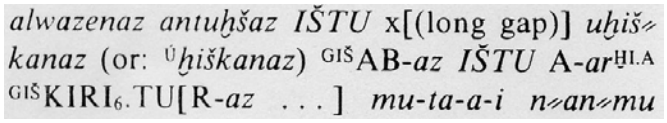

Figure 6. Example from *The Hittite Dictionary of the Oriental Institute of the University of Chicago* [=*Chicago Hittite Dictionary* or *CHD*], ed. Hans Güterbock and Harry Hoffner, volume *L-N* (Chicago: The Institute, 1980, p. 336).

In the same volume, the equals sign is clearly distinguished from the double hyphen (see fig. 7), which is used in the more general "is equivalent to" sense.

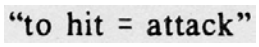

Figure 7. Example of an equals sign from *CHD*, volume *L-N*, p. 9

In later volumes of the *Chicago Hittite Dictionary*, this "word-breaking" usage of the double hyphen sign was abandoned, and a "tilde-hyphen" was adopted in order to break a Hittite word.

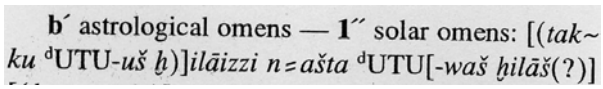

Figure 8. Example of the tilde-hyphen used to break a word, from *CHD*, volume *Š*, fascicle 1, p. 33. Note the slightly different shape of the double hyphen in the second line.

Again, the double hyphen is differentiated from the equals mark:

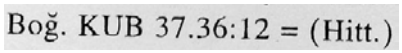

Figure 9. Example of an equals mark from *CHD*, volume *Š*, fascicle 1.

According to Harry Hoffner, one of the *CHD* editors, the online version will adopt the "tilde-hyphen" for the *L-N* set of entries, even though the print version had used the double hyphen. This usage differs from the Webster's *Third New International Dictionary* double hyphen (a "lightface double hyphen at end-of-line is a hyphen that should be retained" [cited in figure 6, document N2647, p. 4]).

The glyph's shape in the earlier *CHD* volume (*L-N*) varies from that used in the later volumes, for the lines in volume *L-N* are not lined up vertically. The shape was modified in the later volumes (*Š* fascicle 1, and *P*). The glyph shape with the two lines vertically aligned is advocated by the CHD and not the "faux" shape of the *L-N* volume.

Figure 10. Examples of the different glyphs used for the double hyphen from *CHD*, volume *L-N* and volume *Š*.

As shown in figure 8 of document N2647, the double hyphen is used outside the *Chicago Hittite Dictionary* (with an example in figure 8 from *Die Funktionen der dimensionalen Kasus und Adverbien im Althethitischen* by Frank Starke in the series Studien zu den Boğazköy-Texten, Wiebaden: Harrassowitz, 1977).

b. In other Hittite texts
In other Hittite texts, an equals sign is still used widely to mark the morphological boundary. Different means have been used in order to differentiate the use of an equals sign as a clitic marker versus a general "is equal to" sense, as described below.

(1) The "clitic marker" does not have surrounding whitespace, whereas the use of the equals sign to show equivalencies does have whitespace:

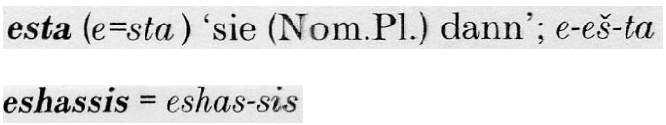

Figure 11. Examples of an equals sign used as a clitic marker without surrounding whitespace (top), versus an identical equals sign used to show equivalencies, but with surrounding whitespace (bottom), from Johann Tischler, *Hethitisches Handwörterbuch*, Innsbruck 2001, p. 33.

(2) The "clitic marker" can have a different length from the equals sign (meaning "is equivalent to"):

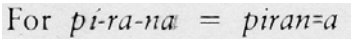

Figure 12. Example of two different lengths of equals mark to show two different uses of the equals sign, from Itamar Singer, *The Hittite KI.LAM Festival*, Part One, Studien zu den Bogazköy-Texten, Heft 27, Wiesbaden: Harrassowitz, 1983, p. 154.

Hurrian
According to Theo van den Hout, Oriental Institute, University of Chicago, and Gernot Wilhelm, Prof. für Altorientalistik, Institut für Altertumswissenschaften Universität Würzburg, the equals sign is also widely used to mark morpheme boundaries. The *Journal of the American Oriental Society*, however, uses a "double hyphen." Dr. Joost Hazenbos, Altorientalisches Institut, University of Leipzig notes that while the equals sign is the most commonly used, "[t]here are, however, other ways to mark morpheme boundaries in Hurrian …the minus sign, or the plus sign (for the last possibility see Speiser's classic *Introduction to Hurrian*, AASOR 20, 1941)."

Anatolian Languages: Cuneiform Luwian, Hieroglyphic Luwian, Lycian, and Lydian
The equals sign is used in the various related Anatolian languages, perhaps based on usage introduced by H. Craig Melchert. The following examples all show the use of the equals sign to mark morphological boundaries:

Cuneiform Luwian

tappaši=tar tapala

Figure 13. From Frank Stark, *Die Keilschrift-luwischen Texte in Umschrift*, Wiesbaden: Harrassowitz, 1985 (Studien zu den Boğazköy-Texten,

25 pa-a=tar a-ap-pa

Figure 14. From Craig Melchert's online cuneiform Luwian text corpus (http://www.unc.edu/~melchert/CLUVIAN.pdf):

(Hieroglyphic Luwian likewise uses the equals mark. No examples are provided here.)

Lycian

me=i(j)=es-eri-hhati : me=hri-qla

Figure 15. From H. Craig Melchert, in an article "Once More on the Conclusion of the Lycian Trilingual of the Létôon" *Historische Sprachforschung* 112, p. 75, 1999.

Lydian

3 laqrisa=k qela=k

Figure 16. From H. Craig Melchert's Lydian corpus, http://www.unc.edu/~melchert/lydian.pdf):

Linguistic Texts

As defined in April 2003 "The Leipzig Glossing Rules: Conventions for interlinear morpheme-by-morpheme glosses" (http://www.eva.mpg.de/lingua/files/morpheme.html), "Clitic boundaries are marked by an equals sign, both in the object language and in the gloss." An example from West Greenlandic appears in figure 17.

(3) West Greenlandic (Fortescue 1984:127)

palasi=lu niuirtur=lu

priest=and shopkeeper=and

'both the priest and the shopkeeper'

Figure 17. An example of the use of an equals sign to mark morpheme boundaries in http://www.eva.mpg.de/lingua/files/morpheme.html

This usage also found in print:

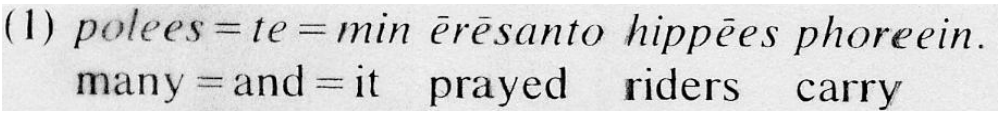

Figure 18. An example of the equals signs to separate clitics, from Stephen Anderson, "Wackernagel's Revenge: Clitics, Morphology, and the Syntax of Second Position" in *Language*, vol. 69, no. 1, 1993, p. 68)

III. Comments

1. The DOUBLE OBLIQUE HYPHEN sign has evolved into three distinct "species":

a. The Frakturhyphen, which is a glyph used when a hyphen (U+002D or U+2010) is rendered in Latin in the Fraktur style. It is also used in Hittite, but only in a very limited way and appears not to have been further adopted by other publications (or later volumes of the *CHD*).

b. The Coptic/Demotic/Hittite "Frakturhyphen," which originates (for Hittite at least) from an equals sign; the origin of the Coptic shape is murkier. This sign has now has caught on, so that people feel that it is graphically distinct from an equals sign (even though equivalent data using an equals sign exists).

c. Webster's *Third New International Dictionary* convention of displaying a hyphen at a linebreak as a "Frakturhyphen" when it derives from a lexical "-" (U+002D) in a word, rather than being an inserted hyphen.

2. The editors of the *Chicago Hittite Dictionary* (Harry Hoffner), Demotic Dictionary Project (Janet Johnson, University of Chicago), and the publisher of the *Journal of the American Oriental Society* (Jim Eisenbrauns, publisher of *JAOS*) have all requested the inclusion of this double hyphen. For the *CHD*, rendering the glyph has been an issue, particularly as they will be going to an online version. The *CHD* requests the glyph shape reflect the form used in the *Š* and *P* volumes of the *CHD*, and not the "faux" Fraktur-form used in the earlier *L-N* volume.

3. For Hittite: While the encoding of this symbol will reflect the usage in the *Chicago Hittite Dictionary*—itself a respected reference work—the double hyphen is not yet thoroughly entrenched in publications with Hittite generally; the equals sign appears in many standard texts (including some volumes of the StBoT series noted above) and has also made significant inroads already in publications of related Indo-European Anatolian languages, Luwian and Lycian (as well as in texts in the non-Indo-European language Hurrian).

In those Hittite texts where an equals sign alone is used (and no use of the double hyphen is employed), publishers have often made an effort to distinguish the different equals signs typographically, the "is equals to" use is typically surrounded with whitespace. Some publishers make the equals sign used for morphological juncture marker shorter than the "is equals to" sign.

Hittite scholars need to be made aware that if the double hyphen is encoded this will impact their searching, since searching for the double hyphen will not result in hits in those Hittite texts which may have used the equals sign to mark clitic boundaries.

4. There is a widespread feeling that the "double oblique hyphen" is the same as a Frakturhyphen because of the glyph shape. If a Fraktur font were widely available or glyphs for Fraktur were included in a widely distributed font, users could erroneously select a glyph for the hyphen character (U+002D or U+2010) in place of the double hyphen. A "best practices" guide for Hittite, Coptic, etc., identifying the correct Unicode character can help forestall such problems, as will availability of fonts for specific fields with the correct underlying characters.

5. The equals sign has been used to different degrees in texts from various languages, but perhaps more markedly in Hittite and related Anatolian languages and Hurrian. Whether the prestige of the *CHD* and the availability of a font with the double hyphen will encourage others to adopt the double hyphen in Anatolian texts over the equals sign remains to be seen; it is possible.

Although I have not investigated it, it seems unlikely that general linguistic texts would adopt the double hyphen, because the equals mark has already been used widely to mark a morphological juncture and most linguists (at least in the U.S.) don't tend to have much exposure to the smaller, more circumscribed world of Anatolian/Coptic/Demotic texts in transcription, where they might encounter it.

IV. Suggested Modifications to Rendering Rules from N2647
(Suggested changes in **bold**)
("X" is used for any missing glyphs in the font)

The repertoire of relevant glyphs is as follows:
A - short, single-bar
B X short, double-bar, uptilted
C – medium, single-bar
D = medium, double-bar
E X 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 ("X" is used for any missing glyphs in my font):

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 (X). In some fallback contexts it is rendered with glyph C (–). **Hittite has used glyph B in a very limited way to break a word across a line; in this context it occurs within a run of Latin letters.**

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. **In some linguistic texts (i.e., Hittite transcribed texts), the two glyphs (D and G) may be used to distinguish a morphological juncture (glyph D) from the use of its general meaning, "is equivalent to" (glyph G).**

6. U+xxxx DOUBLE OBLIQUE HYPHEN is normally rendered with glyph B (X). In certain typographic contexts (e.g. the French Coptic grammar), it may be rendered with glyph E (X) – but this is rather old-fashioned..

Note: There is also a glyph of this character raised off the base line. One might compare the form in Demotic, figure 2. How widespread it is, I have not yet investigated. Below is an example from *An Introduction to Sahidic Coptic* by Thomas O. Lambdin, Macon, Ga.: Mercer University, 1983):