

Reply to L2/26-060 on the corruption of the basic sign list of ancient Egyptian

Mark-Jan Nederhof
University of St Andrews

March 31 2026

Abstract

Despite the occasionally dismissive tone, evading a number of central questions, **L2/26-060** addressed many of my concerns in **L2/26-059** about problems with the Unicode 16 basic sign list of Egyptian hieroglyphs. This included most of the cases of misplaced signs, which was by far the most urgent problem that needed to be resolved. However, two of the most eye-popping and harmful cases of corruption of the basic list have not received a satisfactory resolution. This document will first add further context to these two cases, and then discuss a number of other issues related to the basic list and to the syntax and semantics of the control characters.

U+131A6



v5.2



v16

Gardiner based the shape of U+131A6 on “a gold fly as an honorific decoration” [4], citing Bissing [15], which has an image with three flies hanging from a chain. The heads of the flies are of course pointing up, entirely as expected.

Gardiner further adds a reference to PSBA 22, p. 167 [1]. This shows another occurrence of the fly hieroglyph, of course with the head up.

Two further references listed by Gardiner are Urkunden IV p. 39 l. 1 and p. 893 l. 12 [14, 13]. In both cases the heads are of course pointing up again. The latter of these two is also referred to in ÄZ 48, p. 143 [2], which again confirms the head is pointing up.

These are *all* the references that Gardiner lists. One can therefore conclude, without a shadow of a doubt, that the fly hieroglyph should have its head pointing up. That is the orientation intended by Gardiner on p. 477, and the flipped orientation with the head down on p. 545 of op. cit. is clearly a printing error. This was recognized by the original creators of the basic Unicode sign list, and there was a well-justified and well-documented decision (as part of L2/07-097) to let Unicode 5.2 include the *one and only correct* orientation, consciously avoiding making the same error that was made in, for example, the Hieroglyphica [6], which blindly copied the wrong orientation from p. 545.

Of course, one can never exclude that anomalous forms may exist, and perhaps one day a fly hieroglyph with its head pointing down will be found. But no one has identified such a form since I brought this up a year ago, and the theoretical possibility of such an occurrence is moot.

Fast forward 15 years and out of the blue Unicode 16 flips the fly upside down, to *deliberately* give the fly the *wrong* orientation, and the code point no longer stands for what we thought it stood for. Existing encodings are invalidated overnight. I used the fly hieroglyph myself when encoding the Inscriptions of Ahmose Pen-Nekhet (which includes the fly mentioned above from Urkunden IV p. 39 l. 1), and the references above suggest the fly hieroglyph is by no means rare, and so it has likely been used by many others for encoding other texts in the past 15 years.

So what could possibly have been the compelling reason for turning the fly hieroglyph upside down in Unicode 16 and for thereby causing this much havoc on the user community? L2/26-060 merely has:

“That was then in 2009; later in 2023-2024, when the extension was being worked on, the following comment was made concerning that glyph as originally drawn for Unicode 5.2: Should probably be discarded, as the sign was based on iconography, not actual text, which have [the rotated form]. The current form is also used by HG and JSesh. No change.”

That JSesh has the fly upside down is not even a serious argument. JSesh [11] is an extremely valuable tool widely used in Egyptology, but it is not a standard and does not claim to be. One just needs to consider the numerous changes that are made to the font between different versions of the tool [12]. Perhaps one day the fly in JSesh will have the correct orientation. Until then, to convert from the JSesh format (i.e. Manuel de Codage) to Unicode, the fly can be easily flipped automatically; this is already done by hieropy [9].

That the Hieroglyphica has the wrong orientation is a non-argument. It was recognized in 2009 that the Hieroglyphica had it wrong, and it was decided not to make the same mistake for the Unicode 5.2 sign list. Nothing has changed since 2009. The Hieroglyphica had the wrong orientation then and still has the wrong orientation. JSesh had the wrong orientation then and still has the wrong orientation. No new inscriptions have been identified since 2009 that suggest that the fly should have its head pointing down after all.

That the particular glyph used by Gardiner was based on a piece of jewelry rather than on an inscription is similarly a non-argument. For users who are not concerned with palaeography, U+131A6 is a stylized but acceptable representation of a fly, like there are many hieroglyphs that are not realistic depictions of people, animals and objects. And even if one could argue that the glyph under the new code point U+13DEE is preferable for most occurrences of the fly, that is no reason to deliberately flip U+131A6 upside down.

This leaves absolutely no rational basis for refusing to fix what was so undeniably broken in Unicode 16.

U+1329F



v5.2

v16

In L2/26-059 I explained carefully why the flipping of the boat in Unicode 16 was so wrong and harmful to the user community. None of this is addressed in the 3-line response:

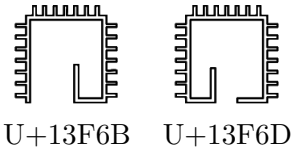
“The mirroring was specifically requested by experts; added a description based on similar signs: ‘A boat/ship with a vertical prow and a stern resembling the handle of a sickle (U001), on top of a rectangle resembling water, with an oval with rounded corners in the boat/ship.’”

I was originally told that the boat had been flipped because there was a claim from experts that the orientation in Gardiner [5] was a mistake. If this claim was indeed made, then these experts were wrong, as I explained in L2/26-059. In the latest reply, the claim of a mistake is no longer mentioned, which leaves absolutely no justification whatsoever.

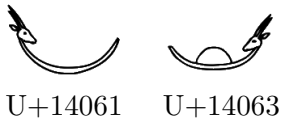
If the argument is no longer that Gardiner is claimed to have made a mistake, then I would guess that there was some kind of misguided attempt to make the sign list more uniform. But this does not hold up to scrutiny either as I will explain below.

First, I can only imagine that the mentioned ‘experts’ never used the Unicode 5.2 sign list for their own work in the past 15 years and did not suffer themselves due to this and many other forms of corruption that were introduced into the basic sign list in Unicode 16. To illustrate to them what the problem is, let me present a scenario of what may happen to *them* in 15 years from now if the flipping of U+1329F is allowed to stand. Suppose it is the year 2040, and the same ‘experts’ who wanted to have U+1329F flipped have extensively used the Unicode sign lists for their own work, they have created large databases and they have published machine-readable papers from which encodings can be copied-pasted.

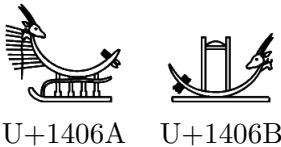
Now some other ‘expert’ comes along and starts flipping signs in the code charts for no good reason at all. For example, in the light of:



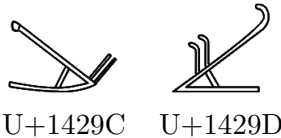
it is thought to be a good idea to flip U+13F6D in the code charts, because of the similar sign U+13F6B, without consideration for others who happen to have used U+13F6D in their work previously, and who find that their carefully compiled encodings are invalidated overnight.



Likewise, someone may decide one day to flip U+14061 or U+14063 in the code charts, because of the similarity of these two signs.



Here someone might be tempted to flip U+1406B to make it more similar to U+1406A (probably not the other way around).



Ditto.





U+142BC U+142BD

Ditto.

In the case of U+1329F and other boat signs, the flipping is particularly egregious, as the orientation of boats may carry information, as I pointed out in L2/26-059 with a reference to the paper by Fischer [3]: the choice of orientation may be influenced by the word or context in which the sign is used. For sailing north one may find the prow at the left and the stern at the right, while for sailing south one may find the reverse. This is not a steadfast rule, but there are tendencies to use one orientation or the other that are important enough for Egyptologists to care about, and encoders will want to record the exact orientation that they find in an inscription. It is then unacceptable that they wake up one morning and find that Unicode has invalidated each and every one of their carefully recorded orientations of this sign.

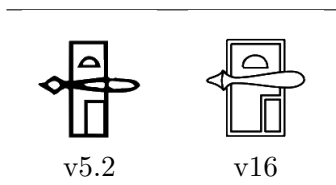
An observation along the same lines is that a number of primitives were introduced in Unicode 15 for the express purpose of rotation and mirroring. The mirroring control character mirrors left/right, and top/bottom mirroring is effectively achieved for left/right-symmetric signs by 180° rotation using a variation selector. This represented an acknowledgement that for Egyptologists, a glyph and its mirror image are not

the same thing, even in cases where the linguistic values are unaffected. For example, where a text has  it is not acceptable to encode  instead, even if this particular mirroring seems to have had no other purpose than that of graphical symmetry. It is a staggering contradiction to acknowledge in Unicode 15 that signs and their mirror images should not be confused, and to then let Unicode 16 start confusing signs and their mirror images.

Moreover, if the boat U+1329F was flipped to make the orientation of boats in the sign list more uniform, then how do we characterize the flipping of the fly U+131A6 discussed earlier? By turning the fly upside down, it now differs from the other fly U+13DEE and other insects like beetles that generally have their heads pointing up.


I would assume that the person who puts their name at the top of all the official Unicode documents carries the final if not the sole responsibility for ensuring that nothing untoward happens to the code charts that the user community relies on. In academic circles, ‘Some expert requested [...]’ carries little weight and has no explanatory power. Especially when a bad decision causes demonstrable harm to the user community, I would expect a more satisfactory response. Better still would be to reverse the harm that has been done.

U+1325F






I’m less concerned here about the Z-order of the composite glyph in the code charts than about what L2/26-060 adds about compositional encoding of the same group (which is actually irrelevant to the Z-order, which is an issue of visual realization rather than of encoding). Three decompositions are claimed to be possible, all of which are incorrect:

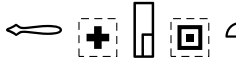


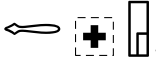

- 1.

2. 

3. “O029 is an overlay to the rest”

The first is already incorrect because insertion at the top assumes that  has an empty space at the top *from above*. It hasn't. Even apart from this, the sequence would represent something different from what is claimed (as in the case of the second sequence).

The syntax of the second sequence is allowable, but it represents something different, namely the overlay  inserted in (i.e. entirely enclosed within) . But that is not what U+1325F is supposed to represent.

The third is similarly incorrect. An overlay has two operands, a flat horizontal group (a single sign or a horizontal arrangement of signs) and a flat vertical group (a single sign or a vertical arrangement of signs), in this order. The character sequence  is in fact allowable and is one of two possible decompositions of U+1325F, but it is not an overlay of  with a group that has an insertion, which is not allowable, but rather an insertion of  within the overlay . The other possible decomposition is  (depending on which canonical ordering within simple overlays one prefers).

Whenever an insertion and an overlay interact, the insertion *must* come last as in the last two encodings above. The reason is self-evident: insertion places a smaller sign in “free space” in or around a larger sign or overlay of signs, and that free space generally does not exist or has not been delimited until *after* the overlay has happened. The syntax governing the relation between overlays and insertions was specified several times in documents in the L2 repository, including in the approved 2021 document introducing new control characters; cf. L2/18-236, L2/21-248, L2/24-079.

For background, I released my first implementation of the insertion and overlay primitives of Unicode in 2018, making use of an existing implementation of RES [7]. Since then I released three more implementations of the Unicode controls [10, 8, 9]. If the number of stars these implementations have received on GitHub is anything to go by, then these packages must have at least dozens of users, who altogether may well have encoded hundreds of texts. Naturally, all my implementations use the same syntax that was agreed in the original Unicode proposals, and over the course of the least 8 years, this syntax has proven itself. The phrase “although there is an update coming”, which L2/26-060 casually throws out, suggests that in 2026, almost 10 years after the insertion and overlay controls were introduced in Unicode, the syntax is still somehow up in the air (it isn't), and that users who have used the controls in the past can now expect that their encodings will soon be invalidated because of some revision or other of the syntax or semantics concocted behind closed doors by Unicode insiders, without involving developers and encoders like myself who were at the forefront of the introduction of the controls of ancient Egyptian.

Encoding order in overlays

“it is expected that the sequence order will use the kEH_Cat values” (p. 3)

This refers to encoding of simple overlays (involving only two signs). I already addressed this in another document, but let me reiterate unambiguously: **under no circumstance will any of my implementations of the control characters involve kEH_Cat values in any shape or form.** As one of the main developers of implementations of encodings of ancient Egyptian over the past 25 years, and those of Unicode

in particular over the past 10 years, I will not accept that such a bad idea is forced upon me and upon others in the user community that I represent. The idea to use `KEH.Cat` values for a canonical ordering in simple overlays seems to have been concocted behind closed doors by Unicode insiders and I never heard about this until reading L2/26-060. If I had been consulted, I would have rejected this idea outright.

U+13017



v5.2



v16

(See L2/26-060 for further background.)

The continued fiddling with the angle of the staff is unproductive. In Unicode 16 the staff was straight vertical, while in Unicode 17 it became slightly slanted away from the man (i.e. the point where the staff rests on the ground is further away from the man than the hand holding the staff), to make it closer to the Unicode 5.2 shape. In Unicode 18, the staff will apparently have the opposite angle, to be slanted towards the man. I don't necessarily object to tiny changes in appearance that are not significant, but at the same time, nothing is achieved in this way. A link between the exact angle of the staff and the particular use of the sign (e.g. 'old age' versus 'wisdom') is illusory. In all likelihood, there is a single grapheme covering a number of related meanings, with a large variation in the angle of the staff between individual tokens that is uncorrelated with those meanings.

U+13055

(See L2/26-060 for further background.) The glyph was updated, but Unikemet-18 still has "with long hair".

U+132BF

(See L2/26-060 for further background.) The glyph was updated, but Unikemet-18 still has "with a rounded top".

U+1315D



v5.2



v16

I have a feeling that the elaborate standard with loop and attachment from Unicode 17 onward will come back to haunt us. For one thing, the alternative sign U+13C5D has the vertical bar all the way at the right end leaving no space for a bottom-right insertion. Encoders who see an occurrence in an inscription with a plain standard as in Unicode 5.2, and who want to encode a plain standard, will have to choose between having the vertical bar in the wrong place (and then possibly not being able to encode a bottom-right insertion) or pretending there is a loop and attachment anyway.

U+13163

“The presence of a lappet in the original Gardiner G31 (U+13163) is vestigial at best and is not corroborated by other sources.”

That we are now arguing about the difference between lappets and vestigial lappets is a clear indication that this endeavor has painted itself into a corner by trying to distinguish what is indistinguishable.

The viewpoint that Unicode should be able to distinguish between a lappet and a vestigial lappet cannot possibly be reconciled with the viewpoint that Unicode need not distinguish between a boat sailing north and a boat sailing south (see above under U+1329F).

U+1335D

“Most sources have a heart.”

I deduce from the way this is phrased that in L2/26-060 “sources” refers to normalized fonts. There still appears to be a discrepancy between two irreconcilable viewpoints: that of the philologists who often overemphasised graphical details of individual tokens in original inscriptions, and that of L2/26-060, who overemphasises graphical details of glyphs in fonts. Much of what I see in Unikemet and the code charts today tends towards either one or the other extreme. A consistent view still seems to be lacking what it is that Extended-A aimed to achieve.

U+1334D

“All other sources use a fork, not a round drill-bit, difference is minor. Description change is accepted.”

Elsewhere there are multiple code points depending on the choice of drill bit, such as for U+142C1 and U+142C2. But here the difference is minor and should be ignored?

There was a change in the description in Unikemet-18, but in a different part of the description. What I was referring to was “a horizontal line on top of” the drill bit.

U+1339E

In Unikemet-18 correct ‘overliad’ to ‘overlaid’.

Conclusions

Vast amounts of effort have been poured into the Extended-A sign list in recent years. I would hope that that was motivated by a desire to serve the interests of users in Egyptology.

A few things that do *not* serve the interests of users:

- undermining the idea of Unicode as a standard, by arbitrarily flipping the orientation of signs, invalidating encodings overnight,
- casually suggesting that the syntax is somehow still up in the air and that no texts can yet be confidently encoded.

If serving the interests of users is no longer its aim, then I would call upon the Unicode Consortium to reflect on what is.

References

- [1] *Proceedings Of The Society Of Biblical Archaeology*, volume 22, 1900. URL: <https://archive.org/details/ProceedingsOfTheSocietyOfBiblicalArchaeologyVol221900/page/n253/mode/2up>.
- [2] *Zeitschrift für ägyptische Sprache und Altertumskunde*, volume 48. 1911. URL: <https://archive.org/details/zeitschriftfr48brug/page/142/mode/2up>.
- [3] H.G. Fischer. *The Orientation of Hieroglyphs, Part I, Reversals*. The Metropolitan Museum of Art, New York, 1977. URL: <https://archive.org/details/TheOrientationOfHieroglyphs/page/n143/mode/2up>.
- [4] A. Gardiner. *Egyptian Grammar*. Griffith Institute, Ashmolean Museum, Oxford, 1957.
- [5] A.H. Gardiner. *Catalogue of the Egyptian hieroglyphic printing type, from matrices owned and controlled by Dr. Alan H. Gardiner*. University Press, Oxford, 1928. URL: <https://www.unicode.org/L2/L2005/05313-Gardiner28-57.pdf>.
- [6] N. Grimal, J. Hallof, and D. van der Plas. *Hieroglyphica*. Publications Interuniversitaires de Recherches Égyptologiques Informatisées, Utrecht, Paris, 2 edition, 2000.
- [7] Mark-Jan Nederhof. *resjs*, 2019. URL: <https://github.com/nederhof/resjs>.
- [8] Mark-Jan Nederhof. *hierojax*, 2026. URL: <https://github.com/nederhof/hierojax>.
- [9] Mark-Jan Nederhof. *hieropy*, 2026. URL: <https://pypi.org/project/hieropy/>.
- [10] Mark-Jan Nederhof. *opentypehiero*, 2026. URL: <https://github.com/nederhof/opentypehiero>.
- [11] S. Rosmorduc. *JSesh*, 2026. URL: <http://jseshdoc.qenherkhopeshef.org>.
- [12] S. Rosmorduc. *Sign Differences Across JSesh Versions*, 2026. URL: <https://jsesh.qenherkhopeshef.org/signdiffs>.
- [13] K. Sethe. *Urkunden der 18. Dynastie, Fascicle 12*. Hinrichs, Leipzig, 1907.
- [14] K. Sethe. *Urkunden der 18. Dynastie, Fascicle 1*. Hinrichs, Leipzig, 1927.
- [15] F.W. von Bissing. *Ein thebanischer Grabfund aus dem Anfang des neuen Reichs*. Verlag von Alexander Duncker, Berlin, 1900. URL: <https://meretsegerbooks.cld.bz/bissingthebanischergrabfund/48/>.