

Towards an Expansion of the Unicode Hieroglyph repertoire

Bob Richmond. 2016-10-28

Outline Proposal for Discussion

Work has been done during the last year to provide a basis from which to build expansions to the Egyptian Hieroglyph repertoire (see below – Background). What is missing is a practical framework to make this happen and do this in a way to best meet the requirements of the user community in a timely fashion.

I am therefore proposing we identify two separate strands for the process to be progressed in parallel.

1. An 'Expansion A' based on Classical Hieroglyphic. This is a natural development of the Unicode (2009) 1071 hieroglyph set based around Gardiner. Eversons work on Möller fits into this strand.
2. An 'Expansion P' based on Ptolemaic and later forms. Suignards work is especially useful for this strand.

This way we avoid the fragmentation of Classical and Ptolemaic that characterises Hieroglyphica (which was based on Gardiner codes only, not on an evolution of the Gardiner Middle Egyptian focussed principles).

To aid development there is an opportunity to innovate around how to go about engaging the expert user base. The **MdC analysis for Unicode Repertoire Expansions** web app (see below) illustrates how we can go beyond simply documenting lists of candidates for encoding in the traditional way to a modern approach in which experts or anyone else can evaluate their work against proposals.

It should be straightforward to refactor work done to date to use the separation model.

Background

The current Unicode (2009) set of 1071 Egyptian Hieroglyphs in Unicode is largely based on the Gardiner font and sign list. This principle was decided in 2006 in consultation with the expert user base, formally presented as [Proposal to encode Egyptian Hieroglyphs in the SMP of the UCS](#) (Everson and Richmond, April 2007) and published with Unicode 5.2 in 2009.

A [Preliminary draft of the Ptolemaic repertoire \(A: Man and his occupations\)](#) (Suignard, October 2015) introduced discussion on an extended repertoire taking Hieroglyphica (2nd Edition, 2000) as a starting point. The latest evolution of this approach is [Source analysis of an extended Egyptian Hieroglyphs repertoire \(Hieroglyphica based\)](#) (Suignard, September 2016).

A [Preliminary proposal to encode Möller's Egyptian Hieroglyphs in the SMP of the UCS](#) (Everson, September 2016) is oriented towards additional hieroglyphs useful for hieratic transcription. The Möller source was used for some elements of the 2007 proposal and regarded as a next step at that time.

Repertoire discussion featured in discussions at the I&E 2016 Cambridge meeting in July and the consensus was that a staged approach using a database of information on hieroglyphs is desirable (just as Everson and Suignard are proposing).

To facilitate better understanding of strengths and limitations of the Unicode (2009) set and potential expansions I've released a web app to analyse existing Manuel de Codage (MdC) documents - [MdC analysis for Unicode Repertoire Expansions](#). MdC accounts for a clear majority of hieroglyphic documents and databases in digital format so now all Egyptologists and others with digital data can evaluate their material in the context of the current Unicode repertoire and potential expansions. The app uses a set of 200 hieroglyphs that are candidates for a Unicode expansion. The document [Analysis of Unicode Egyptian hieroglyphs in a collection of MdC-coded transcriptions](#) provides a summary of results obtained when the app is applied to a collection of 180 MdC transcriptions containing a total of about 240,000 hieroglyphs.