

Comments on proposals for the Universal Multiple-Octed Coded Character Set - **WG2 N2025-2**

Prof. Dr. W. Röllig

- 1.) When generating characters from Old-Semitic languages, you need to know what you are aiming for. It can be sensible to generate, for a limited field of application, standardized character forms that can be used in print. This is the case, e.g. for the Ugaritic cuneiform, the Old Persian cuneiform, for Phoenician and for the Old South Arabian alphabet. Here it is merely reasonable to use standardized forms that in a sense develop an abstraction out of the variants resulting from daily use. However, I would like to stress the fact that such standardized forms are merely useful for specific, very limited purposes, possibly for the editions of specific collections of inscriptions that are not concerned with paleographic questions.
- 2.) The very limited field of application for such generated characters can be explained within our science by the fact that unknown characters are normally transcribed into the Latin script, often with the help of diacritical signs. In the last decades this has also been the dominant practice in letter press printing. Only in specific fields of the presentation of e. g. Phoenician and Aramaic is the Hebraic character repertoire applied, namely in form of the so-called Quadratschrift. More is not needed from the scientific point of view of grammar and linguistics.
- 3.) Paleography is a central field of older Semitistics. Here the object is to reproduce inscriptions as closely to their original as possible. For this, the forms of the letters are scrupulously to be taken into consideration. As these are not standardized and scientific interest is focused especially on unusual characters, it is pointless to generate standardized characters. Each edition of inscriptions will therefore still have to be made by using drawings, under certain circumstances also with the help of a scanner. It is pointless to generate character repertoires for each of the different variants of characters that are found in different regions and appeared in different times. This is the case for Phoenecian as well as for Aramaeic, Old South Arabian, Ugarian etc. It is definitely also the case for the field of cuneiform languages, where it is also pointless to cover the very numerous character by computer. In a very limited range, possibly for introductory studies for teaching purposes a character reservoir of main characters could be defined (ca. 550 characters), that, in a standardized form, could be used for specific purposes. For scientific purposes properly speaking they would be of no use.
- 4.) It is absolutely superfluous to generate the characters of Proto-Sinatic inscriptions, as especially in this case a high variability of characters is of the essence. These can be clearly reproduced in photographs and drawings, but are not normally used in publications. Furthermore, there are only very few such inscriptions which do not justify the necessary effort.

Closing remark

The literature used by authors of character lists is mostly of a secondary nature, i.e. one that has already exploited the truly scientific publications in a popular way. Through this, of course, mistakes have been added. Some of this literature is also clearly no longer up to date. From this follows, furthermore, that sometimes character forms appear in these character lists that are not correct or at least cannot be understood in this manner anymore today, and that on the other hand the very numerous variants that have since been discovered, have, for reasons unbeknownst to me, not been taken into account. If they knew these variants, the authors should have noticed that their undertaking was not very useful.

I would be grateful if you could introduce my comments into the discussion about the relevant projects — in the hope that no superfluous work is done which is of no use to the academic world.

