Organisation internationale de normalisation

Международная организация по стандартизации

International Organization for Standardization

Jeu universel de caractères codés sur plusieurs octets

Universal Multiple-Octet Coded Character Set

Doc Type: Working Group Document
Title: Comments on 2977 (Alleged parallel between 3 N'ko Glyphs & Latin Long S)
Source: François Yergeau, Patrick Andries
Status: Individual Contribution
Date: 2005-09-04

N2977 claims that the appropriate analogy for the old N'Ko (now jona[1]) letter forms is "that to LATIN SMALL LETTER S and LATIN SMALL LETTER LONG S." This is wrong and misleading. The example given in N2977 itself, *Congreſs*, show why the analogy doesn't hold: the long s and the regular s are routinely used together in the same text, the same word even (and may not be interchanged in the *Congreſs* example). It can also convey morphological information, as the well known Wachstube and Wachſtube example illustrates. This is not the case for the old n'ko letter forms. One author may choose to use the old forms (against the advice of Souleymane Kanté, but some still do as stated in N2977) while another author chooses the new forms, but a word containing both forms will normally never occur. Both of these authors can get their preferred forms by simply picking an appropriate font, or better yet a so-called "feature" in one of the increasingly widespread smart font technologies like OpenType and AAT.

Even if the analogy held, U+017F LATIN SMALL LETTER LONG S would be a poor justification for separate encoding of the old n'ko forms: it is a compatibility character and UTR 30 has this to say about it:

> "*Latin long s* should be folded for modern text in roman type style but, other than transiently for searches, should not be folded for texts intended to be set in Fraktur type."

N2977 stresses that these old variants don't have "the kind of 'natural' historical continuity and development that the authors of N2949 seem to believe holds for N'Ko". These authors have never believed that "natural" historical continuity" — whatever this may mean — matters[2] and have said that Unicode/10646 does not encode the history of a script through the encoding of its historical character variants but only encodes its constituent character invariant of change in their glyphs.

---

[1] No longer old letters, but suddenly renamed "jona", a N'Ko term which obscures the fact that these are obsolete, abolished letter forms as stated in N2977.

[2] We are, however, well aware of the particular nature of N'ko's rapid evolution where a recent single inventor could decree that some forms should no longer be used and should be replaced by new ones.