

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: Working Group Document
Title: Discussion of the Mayanist Latin letter CUATRILLO WITH COMMA
Source: Michael Everson
Status: Individual Contribution
Date: 2006-04-16

On the Unicore discussion list, Karl Pentzlin asked: “Why do the LATIN capital/small CUATRILLO WITH COMMA need their own codepoints, instead of using LATIN capital/small CUATRILLO + U+0326 COMBINING COMMA BELOW?”

The mark does not behave like a combining comma below.

First note that real G and real g (not the cuatrillo) are found with U+0327 COMBINING CEDILLA in Latvian orthography, where the cedilla is conventionally drawn with a comma-shape. But with the lower-case g it is not drawn underneath the g, but rather is turned and drawn above it. This is an aside, but it’s just something to note.

Second, note that in Figures 18 and 19 of N3082 the true cedilla is drawn with a comma-like shape. This is sort of “typographic handwriting”, though since neither of those two samples shows the CUATRILLO WITH COMMA... well, let’s just say there’s room for confusion that would be better avoided.

Third, see Figure 20 of N3082, where a glyph variant of CUATRILLO WITH COMMA is shown with a horizontal bar rather than with a comma. If encoded as a unique character, font designers can more easily make the choice to draw it this way; it’s not at all reasonable to expect COMBINING COMMA BELOW to render in this fashion.

Fourth, see Figures 23-26 of N3082. Here we see CUATRILLO WITH COMMA alongside C WITH CEDILLA, though the marks look the same. In large lettering in Figure 23 and 25, it is drawn, not under the cuatrillo, but to the side, “nestled” as I describe it. It’s notable too that in “xtinu4ibah” (Figure 23) and “ço4il” (Figure 26) it is drawn to the right, as though underneath the following vowel. And in the last line of Figure 26, the text reads:

Hagari ahauł Vikuba4 kamama, citan tiłax cablał rubi rutata,

Now here the comma in Vikubag, is attached to the letter, and it is not a true comma as the other two are.

I believe that the preferred behaviour (with the “nestling” I describe in 3.2) should be the one that gets standardized. The COMBINING COMMA BELOW is not otherwise used by Mayanists and there is no particular advantage to decomposing this letter and hoping that font designers will remember to include a glyph for it. Because, of course, if special precomposed glyph attention is not given to the string CUATRILLO + COMBINING COMMA BELOW, then the comma will either render by default superimposed on the CUATRILLO 4 or by being centred below its tail 4—both unsatisfactory default rendering.

Considering the woeful typographic history these poor letters have had, it seems prudent to encode capital and small CUATRILLO WITH COMMA as unique characters 4 4, and let them do their job without burdening Mayanists with rendering issues for them.