

Report on progress with African scripts in Unicode and their implementations**February 5, 2020****Charles L. Riley*****Loma***

At the 5th International Conference on Mande Languages and Linguistics in 2018 in Villejuif, France, I presented on a Loma text from the 1940's that Balla Koevogui had translated, based on my transliteration of the text into Latin script. Maria Khachaturyan shared with me a set of 1970's Loma texts in Latin script on USB drive that she had scanned in while in Liberia.

In April of 2019, I made contact online with Foromo Guilavogui. He introduced me in May to his brother Fougo Guilavogui, Martin Guilavogui, and Grégoire Guilavogui. Of the four, Grégoire has the best knowledge of the Loma script, but is only reachable directly by telephone, not by WhatsApp or e-mail. Indirectly through Martin's e-mail, I sent Grégoire a copy of [L2/17-233](#), Michael Everson's cumulative chart of the Loma script, which he has reported to me that he has understood in translation. Grégoire is reported by Fougo to have given a lecture about the Loma script at a festival of Loma culture which took place in Macenta, Guinea, in late December of 2019. Fougo had also mentioned the existence of texts in his village of Deleou. He has sent me some photos of Grégoire at the festival and of a couple of examples of Loma writing.

On November 11, 2018, I met with Pat Hall and Balla Koevogui to go over transcription of two Loma texts that had been published in the 1940's, including the one I had presented on in Paris. We were interested in glossing the parts of speech for the transcription that had been done for one of them, and in moving on to the next one. Time only permitted for us to review the first text, and results were mixed.

I met again on July 25, 2019 and January 9, 2020 with Balla Koevogui in New Haven to discuss his availability, with Péma Toupou, in translating a UN document, the Universal Declaration of Human Rights, into the Loma language in Latin script as a reference document that could then be a source of conversion of the text into Loma script, pending the availability of alternative fonts to test with in the Private Use Area.

Current Status of Encoding: There remain a number of slightly different repertoires for Loma, which are documented in the latest proposal. A consensus from the user community is needed on which should be the agreed-upon repertoire, in order for a Unicode proposal to move forward.

Latest proposal: [L2/17-233](#)

Garay

Between January and July 2019, a series of Skype and other conference calls was held with the participants ranging from myself to Andrij Rovenchak, Souleiman Faye, Mamadou Diaw, Tim Brookes, and Michael Everson to discuss followup technical comments on the 2016 proposal, [L2/16-069](#). Some early conclusions were reached in response to the comments, but a more thorough line-by-line review of the proposal has yet to be completed. I was introduced by Beth Restrick, head of the library's African Collection at Boston University, and Jenny Parker, director of the West African Research Center, in January of 2020 to Mamaramé Seck, a linguist of Wolof who is based in Dakar, who has expressed interest in assisting with the completion of the project to encode the Garay script. Andrij has shown a

willingness to work on a test font encoded in the Private Use Area. Tim has carved a proverb in Wolof using the Garay script into a piece of wood that he is offering for sale to help raise money for the effort.

Current Status of Encoding: A number of outstanding encoding questions were posed by the Unicode Script Ad Hoc group in [L2/18-168](#), with some responses in [L2/19-163](#), but discussion continues on some points. A thorough review by Souleiman Faye and other users of the latest proposal is needed.

Latest proposal: [L2/16-069](#)

Bété

Between late 2018 and October of 2019, Athinkra, LLC commissioned from Bai Dodo of Zépréguhé, Côte d'Ivoire a few pages of primary lessons and a translation of the Universal Declaration of Human Rights, both into handwritten, scanned Bété script. In both cases, we worked closely with Hermann Gnepa, who facilitated communication and made a field visit to Bai Dodo's farm. These will be important sources as the script is analyzed further in light of the preliminary code chart that was produced by Andrij Rovenchak in 2017, [L2/17-323](#) (and later document with additional information on the characters, [L2/19-044](#))

Current Status of Encoding: Additional information on the script is needed, in particular addressing questions posed in the Script Ad Hoc recommendations ([L2/19-047](#)).

Latest proposal: (code charts) [L2/19-044](#), [L2/17-323](#)

Kpelle

In April of 2018 at the conference on Mande languages in Villejuif, Maria Konoshenko let me know that members of the Kpelle community were interested in the steps required to have their script become encoded. I noted to her that it would be a matter of revitalization, and that the community would have to be involved in producing its own texts. There is archival audio material to be found that should supplement the documentary record, if transcriptions of the audio material can be achieved. I have asked librarian colleagues at Indiana University and UCLA to prioritize these materials for digitization, and they have responded positively. I am hopeful that they can be made available online shortly for further research and use by the community.

Current Status of Encoding: Additional material is needed in order to have a better understanding of the script.

Latest proposal: [L2/10-063](#)

Mandombe

There has been no progress on Mandombe encoding since [L2/16-019](#) by Andrij Rovenchak, Helma Pasch, Nandefo Robert Wazi, and myself in 2016. Plans to restart the encoding process have been suggested, but no clear path has been outlined.

Current Status of Encoding: Collaboration with CENA is needed in order to proceed.

Latest proposal: [L2/16-019](#)

Scripts already encoded

- The British Library's Eyob Derillo worked on this guide to Ethiopian manuscripts: https://en.wikipedia.org/wiki/Ethiopian_manuscript_collections, to which could be added a set of about thirty manuscripts at Yale's Beinecke Rare Books and Manuscripts Library, that have been cataloged by myself with the assistance of Steve Delamarter of George Fox University and His Grace Ewestateos Gebre Kristos, making ample use of Ethiopic script in Unicode.
- The N'ko script has seen the promotion of its Wikipedia from an incubated project to a full-fledged [Wikipedia](#) in its own right. I was involved in getting it off the ground, and worked with colleagues at Harvard University to catalog a set of about sixty titles using N'ko script in the bibliographic metadata.
- There is an [incubated Vai Wikipedia](#) with articles on all of the Liberian presidents, Ebola, and the theory of relativity, edited largely by Tombekai Sherman with the support of Athinkra, LLC.
- Fonts, Keyboards, etc. for other Scripts
 - Ethiopic, N'Ko, Vai, Bassa Vah, Tifinagh, Coptic, ADLaM, and Ajami all have character apps, or "[character pickers](#)", produced by Richard Ishida with some built-in [transliteration tools](#).
 - There are [Noto fonts](#) covering each of these plus Bamum and Mende Kikakui.
 - [Gboard](#), the Google keyboard for mobile devices, covers all of the above ranges with the exception of Mende Kikakui, although the Bamum layout will be getting revised soon.
 - Keymanweb support for some of the above will be forthcoming with support from Athinkra, especially for Bamum and ADLaM.
 - Meroitic has received little attention to date to the best of my knowledge, although it has been encoded.
 - Key gaps that remain are that Coptic lacks a native Windows keyboard layout, and Bamum, Bassa Vah, and Mende Kikakui coverage is absent from common system fonts.
 - Egyptian hieroglyphics have been encoded and layout is being worked on; hieratic and demotic have yet to be encoded.