

To: Unicode Technical Committee
From: Carlos Pallán (via Debbie Anderson, SEI, UC Berkeley)
Subject: Updated List of Characters for Mayan Codices
Date: 26 September 2020

Attached is the updated list of characters found in the Mayan Codices; it updates what was contained in [L2/19-171](#) “Final List of Characters and Quadrats for Mayan Codices,” and reflects progress since 2019

This list is instrumental in the ongoing work of creating and implementing the first font for Mayan hieroglyphs (by Pallán in collaboration with Andrew Glass, Microsoft). The list assigns preliminary code points for the characters, which will be used in a future proposal to the Unicode Consortium. The work is also integral to the development of a state-of-the-art virtual keyboard for Mayan, which will allow real-time text input of signs (e.g. glyphs) non-linearly, into realistic configurations (e.g. "quadrats"), including vertical and horizontal joins and various types of ligatures (e.g. infixation, conflation, superimposition).

List of Characters

The list of characters includes distinct signs and their variants, compiled by Pallán from the extant Mayan codices from Yucatan, preserved at libraries in Dresden, Madrid and Paris.

The headings are as follows (from left to right):

NcodeX¹_CODE: catalog codes assigned by semantic and visual categories with the sequencing defined largely by rate of occurrence for each class. The categories were based on a new taxonomy developed by Pallán that incorporates principles from Gardiner's list for Egyptian hieroglyphs. The system uses one uppercase letter for semantic/visual class (e.g. L for mammals), followed by a number and, in cases of variants, a lowercase letter (e.g. L24a, L24b, etc.)

Picture: vectorial SVG renderings by Pallán (based on codical instances with certain variants modified after Knoróзов 1955,² 1963³) for each sign found in the cross-referenced glyph-catalogs. NOTE: Since each character has been cross-referenced with several other glyph catalogs, visualizations of renderings for those catalogs can be interactively changed as part of the NcodeX Project Database “Concordance” tool.

¹ **NcodeX** is a sub-project created under the umbrella of Script Encoding Initiative with the main task of encoding Mayan hieroglyphs in a progressive fashion, with the Postclassic set of characters from the Mayan codices as a first step. Work is underway to incorporate a much larger set of characters from the Monumental inscriptions of the Classic period. Work has benefitted from funding from a Unicode Adopt-a-Character Grant, a Google Research Grant (PI: Deborah Anderson), a NEH grant to the Universal Scripts Project at UC Berkeley (PR-253360-17, part of SEI, PI: Deborah Anderson), and an NEH grant to the Unicode Consortium (PI: Dr. Gabrielle Vail). Two digital domains have been created to capture and showcase work being made by the project.

² *Sistema Pis'ma Drevnih Mayya. Moscow: Institut Etnografii, Akademiya Nauk SSSR. 1955.*

³ *Pis'mennost' Indeytsev Mayya. Moscow-Leningrad: Institut Etnografii, Akademiya Nauk SSSR. 1963.*

Numeric: unique legacy numeric ID for each character; glyph variants are indicated by lowercase letters (e.g. 11259a, 11259b, etc.)

codepoints: preliminary code points in standard hexadecimal notation (e.g. U+1522E, U+1522F, etc.)

Phonetic1 and Phonetic2 (2 columns): the two more widely used reading values for signs. LOGOGRAPHS (signs with meaning) appear in uppercase and syllables (phonetic values) are shown in lowercase

Translation1 and Translation2 (2 columns): the meaning of each logograph, matching the logograph in the corresponding field

Iconic_Value: field that allows classification of the signs based on their pictorial representations and visual features, rather than their phonetic values. This set of conventions was developed in scholarly literature to designate the signs, based on English glosses of the inferred 'real-word' or cosmological notions and referents related to the pictorial forms. While classification is difficult for the more abstract signs and scholarly consensus has yet to be reached for several forms, the values have proven especially useful in the analysis and classification of problematic/undeciphered signs, hence it was deemed relevant to include this column as well. (Note: "UNK" = Unknown)