

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

Doc Type: Working Group Document
Title: Preliminary proposal to encode the Mwangwego script in the UCS
Source: UC Berkeley Script Encoding Initiative (Universal Scripts Project)
Author: Michael Everson
Status: Liaison Contribution
Date: 2012-07-23

0. Introduction. This proposal gives preliminary information towards the encoding of an African script known as the Mwangwego script. This script was first devised in 1971 by Nolence Moses Mwangwego of Malawi, and is designed for writing the Chewa, Lomwe, Sena, Tonga, Tumbuka, Yao, Nyakyusa, and Ngonde languages, spoken in all by 15 million people. These languages belong to the Bantu branch of the Niger-Congo languages. (Rovenchak and Glavy 2011:104).

1. Structure. The Mwangwego script is typically presented to learners as a syllabary, similar to Ethiopic, but in structure is an abugida, because of the regular way in which vowel signs are added to a base character which has an inherent vowel *-a*. In addition to these vowel-signs, some spacing consonant modifiers which precede a base letter are used; these can also be used in combination with one another to extend the repertoire of consonant onsets. Some combining diacritical marks are also used, either for modification of consonants or for marking high tone to distinguish some homographs. Mwangwego is written from left-to-right.

2. Basic syllabic repertoire. The repertoire of base consonants and vowels as typically presented is as follows:

𞄂 a	𞄃 e	𞄄 i	𞄅 o	𞄆 u	𞄇 ra	𞄈 re	𞄉 ri	𞄊 ro	𞄋 ru
𞄌 ba	𞄍 be	𞄎 bi	𞄏 bo	𞄐 bu	𞄑 sa	𞄒 se	𞄓 si	𞄔 so	𞄕 su
𞄖 cha	𞄗 che	𞄘 chi	𞄙 cho	𞄚 chu	𞄛 sha	𞄜 she	𞄝 shi	𞄞 sho	𞄟 shu
𞄠 da	𞄡 de	𞄢 di	𞄣 do	𞄤 du	𞄥 ta	𞄦 te	𞄧 ti	𞄨 to	𞄩 tu
𞄪 fa	𞄫 fe	𞄬 fi	𞄭 fo	𞄮 fu	𞄯 tsa	𞄰 tse	𞄱 tsi	𞄲 tso	𞄳 tsu
𞄴 ga	𞄵 ge	𞄶 gi	𞄷 go	𞄸 gu	𞄹 psa	𞄺 pse	𞄻 psi	𞄼 pso	𞄽 psu
𞄾 gha	𞄿 ghe	𞅀 ghi	𞅁 gho	𞅂 ghu	𞅃 va	𞅄 ve	𞅅 vi	𞅆 vo	𞅇 vu
𞅈 ha	𞅉 he	𞅊 hi	𞅋 ho	𞅌 hu	𞅍 wa	𞅎 we	𞅏 wi	𞅐 wo	𞅑 wu
𞅒 ja	𞅓 je	𞅔 ji	𞅕 jo	𞅖 ju	𞅗 ya	𞅘 ye	𞅙 yi	𞅚 yo	𞅛 yu
𞅜 za	𞅝 ze	𞅞 zi	𞅟 zo	𞅠 zu	𞅡 za	𞅢 ze	𞅣 zi	𞅤 zo	𞅥 zu
𞅦 ka	𞅧 ke	𞅨 ki	𞅩 ko	𞅪 ku	𞅫 dza	𞅬 dze	𞅭 dzi	𞅮 dzo	𞅯 dzu
𞅰 la	𞅱 le	𞅲 li	𞅳 lo	𞅴 lu	𞅵 dhla	𞅶 dhle	𞅷 dhli	𞅸 dhlo	𞅹 dhlu
𞅺 ma	𞅻 me	𞅼 mi	𞅽 mo	𞅾 mu	𞅿 hla	𞆀 hle	𞆁 hli	𞆂 hlo	𞆃 hlu
𞆄 na	𞆅 ne	𞆆 ni	𞆇 no	𞆈 nu	𞆉 xa	𞆊 xe	𞆋 xi	𞆌 xo	𞆍 xu
𞆎 nya	𞆏 nye	𞆐 nyi	𞆑 nyo	𞆒 nyu	𞆓 qa	𞆔 qe	𞆕 qi	𞆖 qo	𞆗 qu
𞆘 pa	𞆙 pe	𞆚 pi	𞆛 po	𞆜 pu	𞆝 tha	𞆞 the	𞆟 thi	𞆠 tho	𞆡 thu

" ni	+	ⱱ ja	=	"ⱱ nja
" ni	+	ᶑ sa	=	"ᶑ nsa – tongue slightly touches the palate
" ni	+	ᶑ sha	=	"ᶑ nsha
" ni	+	ᶒ fa	=	"ᶒ mfa – labiodentalization
" ni	+	ᶑ va	=	"ᶑ mva

4.5 MITUYO TUMBU. The mark ^ʼ TUMBU seems to cause some sort of prenasalization.

^ʼ tumbu +	ᶑ da	=	ᶑ n'da
^ʼ tumbu +	ᶑ nya	=	ᶑ n'nya
^ʼ tumbu +	ᶑ wa	=	ᶑ n'wa
^ʼ tumbu +	ᶑ ya	=	ᶑ n'ya

4.6 MITUYO HI. The mark ^ʰ HI seems to aspirate consonants. Its use may be limited to only a few consonants.

^ʰ hi +	ᶑ cha	=	^ʰ ᶑ tcha (chha?)
^ʰ hi +	ᶒ ka	=	^ʰ ᶒ kha
^ʰ hi +	ᶑ pa	=	^ʰ ᶑ pha
^ʰ hi +	ᶑ ta	=	^ʰ ᶑ tha
^ʰ hi +	ᶑ tsa	=	^ʰ ᶑ thsa (tsha?)

4.7 MITUYO WAYA. The mark ^ʷ WAYA seems to labialize consonants.

^ʷ waya +	ᶑ ba	=	^ʷ ᶑ bwa
^ʷ waya +	ᶑ ma	=	^ʷ ᶑ mwa

5. Non-spacing modifiers. A variety of changes are also made with the use of non-spacing diacritical marks.

5.1 MITUYO WAYA BELOW. The mark ◌_{WAYA BELOW} adds a following y-glide to consonants.

◌ _{WAYA BELOW} waya +	ᶑ da	=	ᶑ _{WAYA BELOW} dya
◌ _{WAYA BELOW} waya +	ᶑ ma	=	ᶑ _{WAYA BELOW} mya

5.2 MITUYO MURA. The mark ◌[̣] MURA adds a following r-glide to consonants.

◌ [̣] mura +	ᶑ ba	=	ᶑ [̣] bra
◌ [̣] mura +	ᶒ ka	=	ᶒ [̣] kra
◌ [̣] mura +	ᶑ ta	=	ᶑ [̣] tra

5.3 MITUYO MULA. The mark ◌[̤] MULA adds a following l-glide to consonants.

◌ [̤] mula +	ᶑ ba	=	ᶑ [̤] bla
◌ [̤] mula +	ᶒ ka	=	ᶒ [̤] kla
◌ [̤] mula +	ᶑ ta	=	ᶑ [̤] tla

5.4 MITUYO PEWA. The mark $\hat{\circ}$ PEWA seems to serve a variety of purposes, indicating affrication and other consonant modifications.

$\hat{\circ}$ pewa +	U ba	=	\hat{U} wa
$\hat{\circ}$ pewa +	ɔ fa	=	$\hat{\text{ɔ}}$ pfa
$\hat{\circ}$ pewa +	ʌ va	=	$\hat{\text{ʌ}}$ bva
$\hat{\circ}$ pewa +	ɹ za	=	$\hat{\text{ɹ}}$ bza
$\hat{\circ}$ pewa +	ɽ ða	=	$\hat{\text{ɽ}}$ θa

5.5 MITUYO KWANTHU. The mark $\acute{\circ}$ KWANTHU serves to mark the high tone, at least on words which might otherwise be mis-read.

ɹ ta + \circ e + $\acute{\circ}$ kwanthu	=	$\acute{\text{ɹ}}$ " ɔ mténgo 'tree'
ɹ ta + \circ e	=	$\acute{\text{ɹ}}$ " ɔ mtengo 'price'

6. Combinations of modifiers. The following combinations of modifiers have been observed. The order in which these appear is the expected sorting order

6.1 MITUYO MI-HI.

" mi-hi	+	ɖ pa	=	"ɖ mpha
---------	---	------	---	---------

6.2 MITUYO MI-HI-WAYA.

" mi-hi-waya	+	ɖ pa	=	"ɖ mphwa
--------------	---	------	---	----------

6.3 MITUYO MI-WAYA.

" mi-waya	+	U ba	=	"U mbwa
" mi-waya	+	ɖ pa	=	"ɖ mpwa

6.4 MITUYO MYU-NI.

! myu-ni	+	ɽ ka	=	!ɽ mnka
----------	---	------	---	---------

6.5 MITUYO MYU-NI-HI.

!! myu-ni-hi	+	ɽ ka	=	!!ɽ mnkha (<i>not yet attested</i>)
--------------	---	------	---	---------------------------------------

6.6 MITUYO MYU-NI-HI-WAYA.

!!! myu-ni-hi-waya	+	ɽ ka	=	!!!ɽ mnkhwa
--------------------	---	------	---	-------------

6.7 MITUYO MYU-HI.

! myu-hi	+	ɽ ka	=	!ɽ mkha
----------	---	------	---	---------

6.8 MITUYO MYU-WAYA.

! myu-waya	+	ɽ ka	=	!ɽ mkwa
------------	---	------	---	---------

6.9 MITUYO SISA-NI.

ᵀᵀ sisa-ni + ᵀ ka = ᵀᵀ snka (*not yet attested*)

6.10 MITUYO SISA-NI-HI.

ᵀᵀᵀ sisa-ni-hi + ᵀ ka = ᵀᵀᵀ snkha

6.11 MITUYO NI-HI.

ᵀᵀ ni-hi + ᵀ ka = ᵀᵀ nkha

6.12 MITUYO NI-HI-WAYA.

ᵀᵀᵀ ni-hi-waya + ᵀ ka = ᵀᵀᵀ nkhwa

6.13 MITUYO NI-WAYA.

ᵀᵀ ni-waya + ᵀ ka = ᵀᵀ nkwa

6.14 MITUYO HI-WAYA.

ᵀ hi-waya + ᵀ ka = ᵀᵀ khwa

ᵀ hi-waya + ᵀ pa = ᵀᵀ phwa

7. Punctuation and numbers. European digits and punctuation are used and no script-specific marks are known. The characters observed in text so far are comma, full stop, parentheses, and question mark.

8. Bibliography.

Rovenchak, Andrij & Jason Glavy. 2011. “Mwangwego script”, in *African Writing Systems of the Modern Age: The Sub-Saharan Region*. New Haven, Buena Park, New Rochelle, London, Lviv, Abidjan: Athinkra. ISBN 978-0-9818294-1-8

9. Acknowledgements. This project was made possible in part by a grant from the U.S. National Endowment for the Humanities, which funded the which funded the Universal Scripts Project (part of the Script Encoding Initiative at UC Berkeley) in respect of the Mwangwego encoding Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment for the Humanities.

	16E0	16E1	16E2	16E3	16E4
0	𐌲 16E00	𐌳 16E10	𐌴 16E20	𐌵 16E30	
1	𐌶 16E01	𐌷 16E11	𐌸 16E21	𐌹 16E31	
2	𐌺 16E02	𐌻 16E12	𐌼 16E22	𐌽 16E32	
3	𐌾 16E03	𐌿 16E13	𐍀 16E23	𐍁 16E33	
4	𐍄 16E04	𐍅 16E14	𐍆 16E24	𐍇 16E34	
5	𐍈 16E05	𐍉 16E15	𐍊 16E25	𐍋 16E35	
6	𐍌 16E06	𐍍 16E16	𐍎 16E26	𐍏 16E36	
7	𐍐 16E07	𐍑 16E17	𐍒 16E27	𐍓 16E37	
8	𐍕 16E08	𐍖 16E18	𐍗 16E28	𐍘 16E38	
9	𐍙 16E09	𐍚 16E19	𐍛 16E29	𐍜 16E39	
A	𐍞 16E0A	𐍟 16E1A	𐍠 16E2A	𐍡 16E3A	
B	𐍣 16E0B	𐍤 16E1B	𐍥 16E2B	𐍦 16E3B	
C	𐍨 16E0C	𐍩 16E1C	𐍪 16E2C	𐍫 16E3C	
D	𐍮 16E0D	𐍯 16E1D	𐍰 16E2D	𐍱 16E3D	
E	𐍴 16E0E	𐍵 16E1E	𐍶 16E2E		
F	𐍸 16E0F	𐍹 16E1F	𐍺 16E2F		

Letters

16E00	𐌀	MWANGWEGO LETTER A
16E01	𐌁	MWANGWEGO LETTER BA
16E02	𐌂	MWANGWEGO LETTER CHA
16E03	𐌃	MWANGWEGO LETTER DA
16E04	𐌄	MWANGWEGO LETTER FA
16E05	𐌅	MWANGWEGO LETTER GA
16E06	𐌆	MWANGWEGO LETTER GHA
16E07	𐌇	MWANGWEGO LETTER HA
16E08	𐌈	MWANGWEGO LETTER JA
16E09	𐌉	MWANGWEGO LETTER ZHA
16E0A	𐌊	MWANGWEGO LETTER KA
16E0B	𐌋	MWANGWEGO LETTER LA
16E0C	𐌌	MWANGWEGO LETTER MA
16E0D	𐌍	MWANGWEGO LETTER NA
16E0E	𐌎	MWANGWEGO LETTER NYA
16E0F	𐌏	MWANGWEGO LETTER PA
16E10	𐌐	MWANGWEGO LETTER RA
16E11	𐌑	MWANGWEGO LETTER SA
16E12	𐌒	MWANGWEGO LETTER SHA
16E13	𐌓	MWANGWEGO LETTER TA
16E14	𐌔	MWANGWEGO LETTER TSA
16E15	𐌕	MWANGWEGO LETTER PSA
16E16	𐌖	MWANGWEGO LETTER VA
16E17	𐌗	MWANGWEGO LETTER WA
16E18	𐌘	MWANGWEGO LETTER YA
16E19	𐌙	MWANGWEGO LETTER ZA
16E1A	𐌚	MWANGWEGO LETTER DZA
16E1B	𐌛	MWANGWEGO LETTER DHLA
16E1C	𐌜	MWANGWEGO LETTER HLA
16E1D	𐌝	MWANGWEGO LETTER XA
16E1E	𐌞	MWANGWEGO LETTER QA
16E1F	𐌟	MWANGWEGO LETTER THA

16E36	-	MWANGWEGO MITUYO HI
16E37	=	MWANGWEGO MITUYO HI-WAYA
16E38	-	MWANGWEGO MITUYO WAYA
16E39	○	MWANGWEGO MITUYO WAYA BELOW
16E3A	◌̣	MWANGWEGO MITUYO MURA
16E3B	◌̤	MWANGWEGO MITUYO MULA
16E3C	◌̥	MWANGWEGO MITUYO PEWA
16E3D	◌̦	MWANGWEGO MITUYO KWANTHU

Vowel signs

16E20	◌̇	MWANGWEGO VOWEL SIGN E
16E21	◌̈	MWANGWEGO VOWEL SIGN I
16E22	◌̉	MWANGWEGO VOWEL SIGN O
16E23	◌̊	MWANGWEGO VOWEL SIGN U

Consonant modifiers

16E24	◌̋	MWANGWEGO MITUYO MI
16E25	◌̌	MWANGWEGO MITUYO MI-HI
16E26	◌̍	MWANGWEGO MITUYO MI-HI-WAYA
16E27	◌̎	MWANGWEGO MITUYO MI-WAYA
16E28	◌̏	MWANGWEGO MITUYO MYU
16E29	◌̐	MWANGWEGO MITUYO MYU-NI
16E2A	◌̑	MWANGWEGO MITUYO MYU-NI-HI
16E2B	◌̒	MWANGWEGO MITUYO MYU-NI-HI-WAYA
16E2C	◌̓	MWANGWEGO MITUYO MYU-HI
16E2D	◌̔	MWANGWEGO MITUYO MYU-WAYA
16E2E	◌̕	MWANGWEGO MITUYO SISA
16E2F	◌̖	MWANGWEGO MITUYO SISA-NI
16E30	◌̗	MWANGWEGO MITUYO SISA-NI-HI
16E31	◌̘	MWANGWEGO MITUYO NI
16E32	◌̙	MWANGWEGO MITUYO NI-HI
16E33	◌̚	MWANGWEGO MITUYO NI-HI-WAYA
16E34	◌̛	MWANGWEGO MITUYO NI-WAYA
16E35	◌̜	MWANGWEGO MITUYO TUMBU

A. Administrative

1. Title

Preliminary proposal to encode the Mwangwego script in the UCS

2. Requester's name

UC Berkeley Script Encoding Initiative (Universal Scripts Project)

3. Requester type (Member body/Liaison/Individual contribution)

Liaison contribution.

4. Submission date

2012-07-23

5. Requester's reference (if applicable)

6. Choose one of the following:

6a. This is a complete proposal

No.

6b. More information will be provided later

Yes.

B. Technical – General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters)

Yes.

1b. Proposed name of script

Mwangwego.

1c. The proposal is for addition of character(s) to an existing block

No.

1d. Name of the existing block

2. Number of characters in proposal

62.

3. Proposed category (A-Contemporary; B.1-Specialized (small collection); B.2-Specialized (large collection); C-Major extinct; D-Attested extinct; E-Minor extinct; F-Archaic Hieroglyphic or Ideographic; G-Obscure or questionable usage symbols)

Category A.

4a. Is a repertoire including character names provided?

Yes.

4b. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?

Yes.

4c. Are the character shapes attached in a legible form suitable for review?

Yes.

5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Jana Reddemann and Jenna Leich, via Michael Everson.

5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Michael Everson, Fontographer.

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

Yes.

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

No.

7. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

No.

8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database <http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

See above.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

No.

2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

Yes.

2b. If YES, with whom?

Nolence Mwangwego.

2c. If YES, available relevant documents

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

Africanists and Bantuists.

4a. The context of use for the proposed characters (type of use; common or rare)

Used for living languages.

4b. Reference

5a. Are the proposed characters in current use by the user community?

Yes.

5b. If YES, where?

In Malawi.

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

No.

6b. If YES, is a rationale provided?

6c. If YES, reference

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

Yes.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No.

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No.

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

No.

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences?

Yes.

11b. If YES, is a rationale for such use provided?

No.

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

11e. If YES, reference

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No.

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?