L2/26-029

Final Proposal for Encoding the Mwangwego Script in the UCS

Oreen Yousuf | Daniel Yacob

oreen.yousuf@gmail.com | dyacob@gmail.com

To: Script Encoding Working Group (SEWG) / Unicode Technical Committee (UTC)

From: Oreen Yousuf; Daniel Yacob

Subject: Mwangwego

Date: 2025 December 5

General Overview:

This is a revised proposal to encode the Mwangwego script into the Unicode Standard. It supersedes the following documents:

- L2/12-251: "Preliminary proposal to encode the Mwangwego script in the UCS"
- L2/12-311: "Proposal to encode the Mwangwego script in the UCS"
- L2/24-241: "Proposal for Encoding the Mwangwego Script in the UCS"
- L2/25-039: "Revised Proposal for Encoding the Mwangwego Script in the UCS"

This version differs from the most recent version, L2/25-039, due to the following reasons and updates:

- ISO 15924 code for the Mwangwego script.
- More images showing evidence of use of the script. A Google Drive link holds more than 100 images showing evidence of use, and PDFs compiling selected works.
- New formatting of images for SEWG convenience.
- Letter of endorsement to encode the Mwangwego script from a Malawian linguist is included in Section VIII.
- A discussion of the visual and logical orders of composed letters is added.

I Background

The Mwangwego script was first devised in 1979 by Nolence Moses Mwangwego of Malawi, and is designed for writing the languages of Malawi. Some of these languages include the following, along with their *ISO 639 code* and speaker count: Chewa (*nya*, 2 million), Lomwe (*ngl*, 2.5 million), Sena (*swk* in Malawi, *seh* in Mozambican, *bwg* for the divergent Barwe dialect; 2.9 million), Tonga (*tog*, 170,000), Tumbuka (*tum*, 2.3 million), Yao (*yao*, 3.1 million), Nyakyusa (*nyy*, 1.4 million). The majority of these languages are also spoken in parts of nearby countries such as Mozambique, Tanzania and Zambia.

Additional letters were created up until April 7, 1997 and launched to the public. The script has not changed significantly over time and is not related to any other script. In 2003 the script received notable publicity, including an audience with Minister Kamangadzai Kingsley Chambalo, the Malawian Minister of Youth, Sports and Culture at the time.

The users of the community are students of the script, some of which are school-aged children. There was an initial community of 200-300 adherents of Mwangwego script in 2001 (Dobrovolny 2010). The number of people who have learned the script since 2001 is between 2,500-3,000 and has been regarded positively by paleographers as "reach[ing] the rank of a small movement" (Kostelnik 2023). Many are training to be teachers themselves. The script has been continuously learned and taught since 2001 in all 3 regions of Malawi: Northern, Central and Southern Regions. The script has been taught

informally since 2001. The distribution of January 2023 cohort of students is broken down as follows: 10 in Karonga, 15 in Mzuzu, 10 in the Malawian capital of Lilongwe, and 5 in Blantyre. The script is also included in primers and learning exercise materials for students. As of mid-2024, another cohort of teachers are being trained.

While there has only been 1 full book officially printed and published in the script, by the creator, as cost is a major obstacle for the majority of past and current learners, the script has taken on a life beyond the script creator and his social circle. Individuals regularly correspond in the Mwangwego script, producing letters, grocery lists, journal entries, business memos, manifestos, and more, in languages like Chichewa and Tumbuka. Script users own mobile phones and are keen to use the script in digital environments. The proposal authors have even introduced previously unknown script users to the creator. Every image of handwritten text in Section IX showing evidence of use comes from users other than the script creator. Language policy in Malawi has the government applying a limited top-down approach. Mwangwego persists and has grown instead by grassroots efforts where it has spread beyond the creator's home city and native language. Many of the images found in Section IX and the Google Drive come from Mzuzu, Lilongwe, and Blantyre. There are at least 4 fonts: one made by Tapiwanashe Sebastian Garikayi; one by Andrij Rovenchak; one by Enzo Bicudo Pepi (MetrikEnzyme); and one by Jana Reddemann and Jenna Leich, the latter of which is used in this proposal. Under consultation with the script community, the authors have developed a keyboard that applies the character model described in this proposal. The keyboard has been developed with the well-known "Keyman" software and a Microsoft Windows executable is available here: https://bit.ly/4e11h3D .

While the Malawian government endorses the Latin script for the languages mentioned above, there are no other scripts in competition among the Mwangwego script community, and the character repertoire has been stable for decades. There is also a current project to digitize Nolence Mwangwego's Chichewa language book "[illegible]?" (Latin transliteration: "*A Malawi Tili Pati?*"; English translation: "*Malawians, where are we?*"), which was written in the Mwangwego script.

II Script Name

The proposed script name is "**MWANGWEGO**", which is named after the creator, Nolence Mwangwego. This is the name used by the script community. The proposed ISO 15924 code for the Mwangwego script is "Mwgo".

On 3 April 2024, Nolence Mwangwego stated on Facebook that the script "*in Malawian languages is called '**musitu wa Mwangwego**'. The word 'musitu' comes from three parts of the script: musisi, siri, mutuyo. Mwangwego script was invented in April 1979 and was inaugurated on April 7, 1997. These are dates to remember*" (see Reference 6 in Section VII).

III Structure

The Mwangwego script is an abugida written left-to-right in horizontal lines, from the top to the bottom of a page. The basic grapheme inventory accounts for 31 consonants and 5 vowels – /a/, /ɛ/, /i/, / ɔ/, and /u/. Each basic grapheme represents a syllable composed of a consonant and a vowel (CV), or a vowel (V) by itself. Students first learn the graphemes that have an inherent /a/ vowel (e.g., /a/, /ba/, /ʧa/, etc.). These inherent /a/ vowel graphemes, as shown in the chart below under the column labeled "/a/". Any single one of these graphemes is called a **Musisi**. All Musisi (i.e., the plural) is collectively called ***Misisi***. Misisi can be slightly altered in 1 of 4 graphically consistent ways to change the grapheme's vowel from an inherent /a/ to /ɛ/, /i/, / ɔ/, or /u/. Any single one of these 4 vowel modifiers is called a **Siri**. All Siri (i.e., the plural) is collectively called ***Masiri***. Attaching Masiri to Misisi will yield a total of 160 basic V or CV syllabic graphemes. Any single one of these new graphemes with a vocalic value of /ɛ/, /i/, / ɔ/, or /u/ is called a **Musiri**. All Musiri (i.e., the plural) is collectively called ***Misiri***.

IPA	/a/	/ɛ/	/i/	/ɔ/	/u/
/vowel/	𖬀	𖬁	𖬂	𖬃	𖬄
/b/	𖬅	𖬆	𖬇	𖬈	𖬉
/ʧ/	𖬊	𖬋	𖬌	𖬍	𖬎
/d/	𖬏	𖬐	𖬑	𖬒	𖬓
/f/	𖬔	𖬕	𖬖	𖬗	𖬘
/g/	𖬙	𖬚	𖬛	𖬜	𖬝
/ɣ/	𖬞	𖬟	𖬠	𖬡	𖬢
/h/	𖬣	𖬤	𖬥	𖬦	𖬧
/ʤ/	𖬨	𖬩	𖬪	𖬫	𖬬
/ʒ/	𖬭	𖬮	𖬯	𖬰	𖬱
/k/	𖬲	𖬳	𖬴	𖬵	𖬶
/l/	𖬷	𖬸	𖬹	𖬺	𖬻
/m/	𖬼	𖬽	𖬾	𖬿	𖭀
/n/	𖭁	𖭂	𖭃	𖭄	𖭅
/ɲ/	𖭆	𖭇	𖭈	𖭉	𖭊
/p/	𖭋	𖭌	𖭍	𖭎	𖭏
/r/	𖭐	𖭑	𖭒	𖭓	𖭔
/s/	𖭕	𖭖	𖭗	𖭘	𖭙
/ʃ/	𖭚	𖭛	𖭜	𖭝	𖭞
/t/	𖭟	𖭠	𖭡	𖭢	𖭣
/ts/	𖭤	𖭥	𖭦	𖭧	𖭨
/p’/	𖭩	𖭪	𖭫	𖭬	𖭭
/v/	𖭮	𖭯	𖭰	𖭱	𖭲
/w/	𖭳	𖭴	𖭵	𖭶	𖭷
/j/	𖭸	𖭹	𖭺	𖭻	𖭼
/z/	𖭽	𖭾	𖭿	𖮀	𖮁
/dz/	𖮂	𖮃	𖮄	𖮅	𖮆
/dɬ/	𖮇	𖮈	𖮉	𖮊	𖮋
/ɬ/	𖮌	𖮍	𖮎	𖮏	𖮐
/‖/	𖮑	𖮒	𖮓	𖮔	𖮕
/!/	𖮖	𖮗	𖮘	𖮙	𖮚
/ð/	𖮛	𖮜	𖮝	𖮞	𖮟

Table 1. Inventory of the basic V and CV syllabic graphemes in the Mwangwego script.

To change the inherent /a/ vowel of Misisi to /ɛ/, /i/, /ɔ/, or /u/, you fuse Masiri (vowel modifiers) to the bottom-right of Misisi:

Emwa: ◌̱ Example: Ʊ (ba) + ◌̱ (-ɛ) = Ʊ̱ (be)

Ima: ◌ɹ Example: Ʊ (ba) + ◌ɹ (-i) = Ʊɹ (bi)

Ota: ◌ↄ Example: Ʊ(ba) + ◌ↄ (-ɔ) = Ʊↄ (bɔ)

Uyu: ◌ɾ Example: Ʊ(ba) + ◌ɾ (-u) = Ʊɾ (bu)

Mutuyo & Mituyo:

In addition to these vowel modifiers, some consonant (or tone) modifiers appear to the left (spacing) of the basic grapheme, or above or below (non-spacing) the basic grapheme. These can also be used in combination with one another to further extend the repertoire of consonant onsets or, in one case, to mark high-tone to distinguish between homographs. Not all of these modifiers can be used with all basic graphemes. Section X show which marks are possible for the Misisi. All valid possibilities are possible for the Misiri counterparts.

Any one of these marks is called **Mutuyo**. You can attach multiple Mutuyo onto a single basic grapheme. The plural of Mutuyo is ***Mituyo***. Some Mituyo configurations produce more complex spacing characters.

The script's previous proposal (Section 4, L2/12-311) proposed to encode all spacing Mituyo configurations due to the constraints of ligation at the time. For compound consonant modifiers one could use the 'rlig' or 'ccmp' OpenType features to replace a sequent of modifiers with a precomposed stacked version. 'ccmp' would be the most reliable across platforms. Therefore, we propose to only encode the standalone modifiers (i.e., Mutuyo). The set of combinations (i.e., Mituyo) is finite, as is the order in which combinations are combined. It is never correct (i.e., legible) to string them horizontally, and making them "ligatures" may easily fail in many rendering environments, which is unnecessary and undesirable. Readers do not parse the stacks per se, but read the whole cluster as a whole.

1. **MUTUYO** (spacing):

1.1 WAYA. ˗ labializes consonants. There is a unique, non-spacing counterpart of Waya that appears under a letter instead of to the left; see Mutuyo 2.1 below.

Examples of use:
˗ (waya) + Ʊ (ba) = ˗Ʊ (bwa) ˗ (waya) + ϩ (ma) = ˗ϩ (mwa)

Example words:
˗Ʊƅ (bwato) ˗ϩϥ (mwala)

1.2 NI. " is primarily used for homorganic palatalization (which may not be reflected in the transliterated romanization).

Examples of use:
" (ni) + Ɛ (da) = "Ɛ (nda) – tip of the tongue touches the front part of the palate
" (ni) + ʃ (ta) = "ʃ (nta)
" (ni) + ʃ (tsa) = "ʃ (ntsa)
" (ni) + f (dza) = "f (ndza)
" (ni) + Ƕ (a) = "Ƕ (ng'a) – rear of the tongue touches the palate
" (ni) + ɑ (ga) = "ɑ (nga)
" (ni) + 9 (cha) = "9 (ncha) – middle of the tongue touches the palate
" (ni) + ʊ (ja) = "ʊ (nja)

" (ni) + ϵ (sa) = "ϵ (nsa) – tongue slightly touches the palate
" (ni) + ϙ (sha) = "ϙ (nsha)
" (ni) + ℶ (fa) = "ℶ (mfa) – labiodentalization
" (ni) + Ꭿ (va) = "Ꭿ (mva)

Example words:
"Ʒʆ (ndati) , "ԵՍ (ng'ona) , "ϵძ১ (nsapato) , "Ϡς১ (nchalo)

1.3 HI. ¨ aspirates consonants. Its use may be limited to only a few consonants.

Examples of use:
¨ (hi) + Ͽ (cha) = ¨Ͽ (tcha) , ¨ (hi) + ʔ (ka) = ¨ʔ (kha) , ¨ (hi) + ძ (pa) = ¨ძ (pha)
¨ (hi) + ʃ (ta) = ¨ʃ (tha) ¨ (hi) + ʆ (tsa) = ¨ʆ (tsha)

Example words:
¨ʔϵ (khasu) ¨ʔϨ (khama) ¨ძς (phala) ¨১Ս (thobwa) ¨Ϡ১ (tchimo)

1.4 MI. " pre-nasalizes labial consonants. It is described as "used when both lips meet heavily." It is used with 2 Misisi (and of course their corresponding Misiri): Ս (/ba/) and ძ (/pa/).

Examples of use:
" (mi) + Ս (ba) = "Ս (mba) " (mi) + ძ (pa) = "ძ (mpa)

Example words:
"Սς (mbala) "Սᖇ (mbiri) "Սʃʃ (mbatata) "ძʔ (mpaka)

1.5 MYU. ʻ pre-nasalizes non-labial consonants. It is described as "used when pronouncing a word which involves a slight meeting of both lips." Myu occurs with many Misisi.

Examples of use:
ʻ (myu) + Ͽ (cha) = ʻϿ (mcha) ʻ (myu) + ʔ (ka) = ʻʔ (mka) ʻ (myu) + Ϩ (ma) = ʻϨ (m'ma)
ʻ (myu) + ʃ (ta) = ʻʃ (mta) ʻ (myu) + f (dza) = ʻf (mdza)

Example words:
ʻʔʔ (mkaka) ʻʔ"Ʒ (mkanda) ʻf১ (m'dziko) ʻϨf (m'mudzi)

1.6 SISA. ʼ prefixes 's-' to consonants. Very few words use Sisa as it is mainly used for foreign loan words (e.g., English; see 7:16 of reference 3 in Section VII below).

Examples of use:
ʼ(sisa) + ʔ (ka) = ʼʔ (ska) ʼ(sisa) + Ϩ (ma) = ʼϨ (sma) ʼ(sisa) + ʃ (ta) = ʼʃ (sta)

Example words:
ʼʆϨ (stima) ʼʔʆ (sketi)

1.7 TUMBU. ˈ causes prenasalization. It is only used for Yao and Nyakyusa/Ngonde.

Examples of use:
ˈ (tumbu) + Ʒ (da) = ˈƷ (n'da)
ˈ (tumbu) + Ⴘ (nya) = ˈႸ (n'nya)
ˈ (tumbu) + ϑ (ya) = ˈϑ (n'ya) – in Yao

' (tumbu) + Ɛ (wa) = 'ʘ (n'wa)
Example words:
'ɯ̈ɪ , ɥ'ⱻƔ , 'Ƴ-ⱻƨ

2. **MUTUYO** (non-spacing):

2.1 WAYA-BELOW. ◌̱ adds a following y-glide (palatalization) to consonants. This is the non-spacing counterpart of WAYA in 1.1 above.

Examples of use:
Ǝ (da) + ◌̱ (waya) = Ǝ̱ (dya)
ⱻ (ma) + ◌̱ (waya) = ⱻ̱ (mya)

Example words:
Ɫɖɖ , Ǝ̱ɩ , ୨ⱻǝ̱

2.2 MURA. ◌̏ adds a following r-glide to consonants. This Mutuyo is also used for foreign words (e.g., English).

Examples of use:
Ʊ (ba) + ◌̏ (mura) = Ʊ̏ (bra)
ʔ (ka) + ◌̏ (mura) = ʔ̏ (kra)
ʃ (ta) + ◌̏ (mura) = ʃ̏ (tra)

Example words:
ɯʃ̏ɯ , Ʊ̏ɾɯ

2.3 MULA. ◌̀ adds a following l-glide to consonants. This Mutuyo is also used for foreign words (e.g., English).

Examples of use:
Ʊ (ba) + ◌̀ (mula) = Ʊ̀ (bla)
ʔ (ka) + ◌̀ (mula) = ʔ̀ (kla)
ʃ (ta) + ◌̀ (mula) = ʃ̀ (tla)

Example word:
Ʊ̀"ʃʘ (translation: Blantyre)

2.4 PEWA. ◌̂ serves a variety of purposes, indicating affrication and other consonant modifications. It is particularly used a lot in Tumbuka, Kyangonde (a dialect of Nyakyusa; possibly other Nyakyusa dialects as well), and Yao.

Examples of use:
Ʊ (ba) + ◌̂ (pewa) = Ʊ̂ (Chichewa Latin transliteration is ŵa;
Tumbuka Latin transliteration is ba; IPA: /βa/)
Ɑ (va) + ◌̂ (pewa) = Ɑ̂ (bva)
ɼ (za) + ◌̂ (pewa) = ɼ̂ (bza)
Ɔ (fa) + ◌̂ (pewa) = Ɔ̂ (pfa)
Ր (tha) + ◌̂ (pewa) = Ր̂ (IPA: /θa/) – in Lomwe

Example words:
ÛÜ£ , Ûr·Ŀr

2.5 KWANTHU. ◌́ marks high tone, at least on words which might otherwise be misread. It is very important in Chichewa because there are 2 tones (in contrast to a language like Tumbuka having no tones).

Examples of use:
Ʒı (di) + ◌́ (kwanthu) = Ʒ́ı (ndi)

Example words:
£́"Ꮹ (mtengo; translation: tree) to distinguish from £̱́"Ꮹ (mtengo; translation: price)

3. <u>**MITUYO "*stacks*"**</u>:
As Mituyo is the plural of Mutuyo, any combination of 2 individual Mutuyo on a single grapheme can be considered a Mituyo combination. For example, "Ʉ̱ and -Û both combine a spacing Mutuyo (NI " and WAYA -, respectively) with a non-spacing Mutuyo (WAYA-BELOW ◌̱ and PEWA ◌̂, respectively). These are considered Mituyo combinations. However, it is worth specifically highlighting how the *spacing* Mutuyo (described in 1.1-1.7 above) can create spacing Mituyo "stacks" that are written on the left of a grapheme. There are 16 possible Mituyo stacks:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
MYU NI	MYU HI	MYU WAYA	MI HI	MI WAYA	NI HI	NI WAYA	HI WAYA	SISA NI	TUMBU WAYA	MYU NI HI	MYU HI WAYA	MI HI WAYA	NI HI WAYA	SISA NI HI	MYU NI HI WAYA

Table 2. Inventory of the 16 spacing Mituyo "stacks".

Similar to the (individual) Mutuyo, not all Mituyo are used for all languages. For example, the Mituyo stack TUMBU+WAYA (numbered 10 in Table 2) is only used in Nyakyusa/Ngonde.

3.1. MYU-NI.

Examples of use:
MYU-NI + Ʒ (da) = Ʒ (mnda)   MYU-NI + ʔ (ka) = ʔ (mnka)

Example words:
Ʒ"Ʒ"Ʒ , Ʒ"Ʒ

3.2. MYU-HI.

Examples of use:
MYU-HI + 9 (cha) = 9 (mchha)   MYU-HI + ʔ (ka) = ʔ (mkha)

Example words:
"d£ , "d"ᗡ

3.3. MYU-WAYA. ˙

Examples of use:

˙ MYU-WAYA + Ʊ (ba) = ˙Ʊ (mbwa) ˙ MYU-WAYA + 9 (cha) = ˙9 (mchwa)

Example words:

˙ʔ , ˙ƍɩɩ , ˙ʔɞ

3.4. MI-HI. ¨

Examples of use:

¨ MI-HI + ɗ (pa) = ¨ɗ (mpha)

Example words:

¨ɗƩ , ¨ɗ¨Ɑ

3.5. MI-WAYA. ¨

Examples of use:

¨ MI-WAYA + Ʊ (ba) = ¨Ʊ (mbwa) ¨ MI-WAYA + ɗ (pa) = ¨ɗ (mpwa)

Example words:

¨Ʊ¨Ʒ , ʕ¨Ʊ , ɾ¨Ʊ

3.6. HI-WAYA. ¨

Examples of use:

¨ HI-WAYA + ʔ (ka) = ¨ʔ (khwa) ¨ HI-WAYA + ʃ (ta) = ¨ʃ (thwa)

Example words:

¨ɗʕ , ¨ɗƩ

3.7. NI-HI. ¨

Examples of use:

¨ NI-HI + 9 (cha) = ¨9 (nchha) ¨ NI-HI + ʔ (ka) = ¨ʔ (nkha)

Example words:

¨ʔ¨Ʒ , ¨ʔƍ , ¨9Ɩ

3.8. NI-WAYA. ¨

Examples of use:

¨ NI-WAYA + Ɥ (a) = ¨Ɥ (nwa) ¨ NI-WAYA + Ʒ (da) = ¨Ʒ (ndwa)

Example words:

¨ⱭƖƖ , ¨Ɑʃ , ¨Ɑ¨Ʊ

3.9. SISA-NI. ¨

Examples of use:

¨ HI-WAYA + Ʒ (da) = ¨Ʒ (snda) ¨ HI-WAYA + Ɑ (ga) = ¨Ɑ (snga)

Example words:

ʔ٢

3.10. TUMBU-WAYA.

Examples of use:

TUMBU-WAYA + Ӈ (gha) = Ӈ (n'ghwa) TUMBU-WAYA + ℓ (ha) = ℓ (n'hwa)

Example words:

ሃ̃ᢁ

3.11. MYU-NI-HI.

Examples of use:

MYU-NI-HI + ʔ (ka) = ʔ (mnkha) MYU-NI-HI + ʃ (ta) = ʃ (mntha)

Example words:

ᢁʔ٢

3.12. MYU-HI-WAYA.

Example of use:

MYU-HI-WAYA + 9 (cha) = 9 (mchhwa)

Example words:

ʃ٢ᢁ

3.13. MI-HI-WAYA.

Examples of use:

MI-HI-WAYA + ɗ (pa) = ɗ (mphwa)

Example words:

ɗ٢ᢁ

3.14. NI-HI-WAYA.

Examples of use:

NI-HI-WAYA + 9 (cha) = 9 (nchhwa) NI-HI-WAYA + ʔ (ka) = ʔ (nkhwa)

Example words:

ʔᢁ ᢁʔ٢

3.15. SISA-NI-HI.

Examples of use:

SISA-NI-HI + ʔ (ka) = ʔ (snkha) SISA-NI-HI + ʃ (ta) = ʃ (sntha)

Example words:

ʔ٢

3.16. MYU-NI-HI-WAYA.

Examples of use:

MYU-NI-HI-WAYA + (ka) = (mnkhwa) MYU-NI-HI-WAYA + (ta) = (mnthwa)

Example words:

In summary, there are 32 Misisi (inherent /a/ vowel graphemes), 4 Masiri (vowel modifiers), 12 Mutuyo (11 consonant modifiers; 1 tone modifier), and 16 Mituyo (“stacks”) to encode.

- It is worth mentioning that the WAYA Mutuyo and the WAYA-BELOW Mutuyo *must* be encoded as two symbols and not only one. This is because if one were to process plain text (e.g., transliteration) one needs to be able to interpret the bytes as being “wa” (from WAYA) and “ya” (from WAYA-BELOW), which one couldn’t do with a single character point as position information is lost in plain text.

Punctuation and Numerals/Digits:

There are no script-specific digits in the script; Hindu-Arabic digits (i.e., 0123456789) are used. There are no script-specific punctuation marks. Words are separated using U+0020 SPACE .

IV Character Repertoire

Table 3 presents the character names for the 32 Mwanwego Misisi (singular Musisi; inherent -/a/ graphemes/letters), 4 Masiri (singular Siri; vowel marks), 12 Mutuyo (consonant/tone modifiers), and 16 Mituyo (complex Mutuyo combinations). The Chichewa language terms for the orthographic elements have been employed to clearly and distinctly identify the characters by their encoded names.

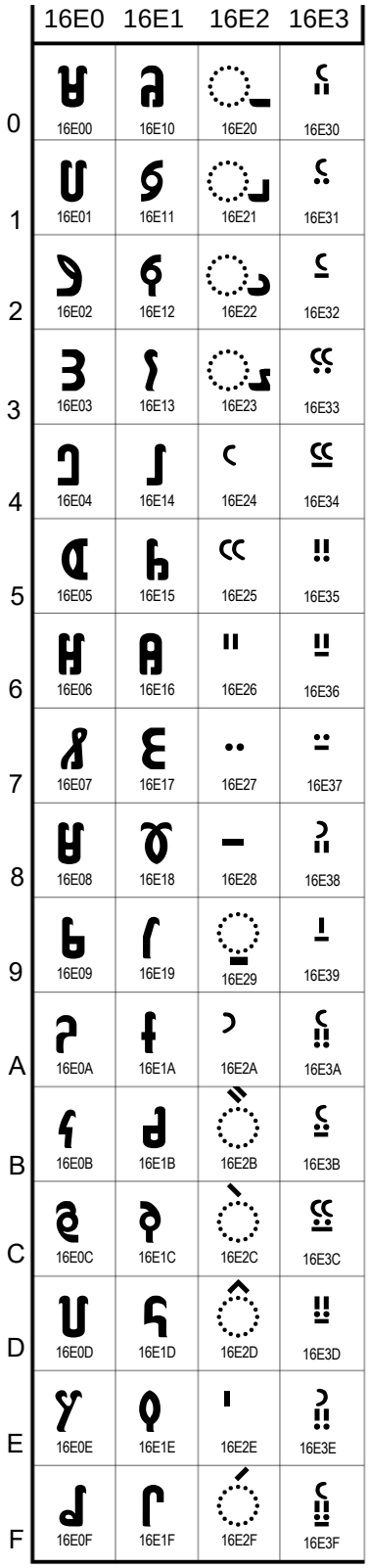
	U+16E00 MWANGWEGO LETTER A
	U+16E01 MWANGWEGO LETTER BA
	U+16E02 MWANGWEGO LETTER CHA
	U+16E03 MWANGWEGO LETTER DA
	U+16E04 MWANGWEGO LETTER FA
	U+16E05 MWANGWEGO LETTER GA
	U+16E06 MWANGWEGO LETTER GHA
	U+16E07 MWANGWEGO LETTER HA
	U+16E08 MWANGWEGO LETTER JA
	U+16E09 MWANGWEGO LETTER ZHA
	U+16E0A MWANGWEGO LETTER KA
	U+16E0B MWANGWEGO LETTER LA
	U+16E0C MWANGWEGO LETTER MA
	U+16E0D MWANGWEGO LETTER NA

𞤟	U+16E0E MWANGWEGO LETTER NYA
𞤠	U+16E0F MWANGWEGO LETTER PA
𞤡	U+16E10 MWANGWEGO LETTER RA
𞤢	U+16E11 MWANGWEGO LETTER SA
𞤣	U+16E12 MWANGWEGO LETTER SHA
𞤤	U+16E13 MWANGWEGO LETTER TA
𞤥	U+16E14 MWANGWEGO LETTER TSA
𞤦	U+16E15 MWANGWEGO LETTER PSA
𞤧	U+16E16 MWANGWEGO LETTER VA
𞤨	U+16E17 MWANGWEGO LETTER WA
𞤩	U+16E18 MWANGWEGO LETTER YA
𞤪	U+16E19 MWANGWEGO LETTER ZA
𞤫	U+16E1A MWANGWEGO LETTER DZA
𞤬	U+16E1B MWANGWEGO LETTER DHLA
𞤭	U+16E1C MWANGWEGO LETTER HLA
𞤮	U+16E1D MWANGWEGO LETTER XA
𞤯	U+16E1E MWANGWEGO LETTER QA
𞤰	U+16E1F MWANGWEGO LETTER THA
◌𞤱	U+16E20 MWANGWEGO VOWEL SIGN EMWA
◌𞤲	U+16E21 MWANGWEGO VOWEL SIGN IMA
◌𞤳	U+16E22 MWANGWEGO VOWEL SIGN OTA
◌𞤴	U+16E23 MWANGWEGO VOWEL SIGN UYU
‘	U+16E24 MWANGWEGO MODIFIER MYU
“	U+16E25 MWANGWEGO MODIFIER MI
"	U+16E26 MWANGWEGO MODIFIER NI
¨	U+16E27 MWANGWEGO MODIFIER HI
-	U+16E28 MWANGWEGO MODIFIER WAYA
◌̲	U+16E29 MWANGWEGO MODIFIER WAYA BELOW
’	U+16E2A MWANGWEGO MODIFIER SISA
◌̋	U+16E2B MWANGWEGO MODIFIER MURA
◌̀	U+16E2C MWANGWEGO MODIFIER MULA
◌̂	U+16E2D MWANGWEGO MODIFIER PEWA
'	U+16E2E MWANGWEGO MODIFIER TUMBU
◌́	U+16E2F MWANGWEGO MODIFIER KWANTHU
𞤸	U+16E30 MWANGWEGO MODIFIER MYU-NI

	U+16E31 MWANGWEGO MODIFIER MYU-HI
	U+16E32 MWANGWEGO MODIFIER MYU-WAYA
	U+16E33 MWANGWEGO MODIFIER MI-HI
	U+16E34 MWANGWEGO MODIFIER MI-WAYA
	U+16E35 MWANGWEGO MODIFIER NI-HI
	U+16E36 MWANGWEGO MODIFIER NI-WAYA
	U+16E37 MWANGWEGO MODIFIER HI-WAYA
	U+16E38 MWANGWEGO MODIFIER SISA-NI
	U+16E39 MWANGWEGO MODIFIER TUMBU-WAYA
	U+16E3A MWANGWEGO MODIFIER MYU-NI-HI
	U+16E3B MWANGWEGO MODIFIER MYU-HI-WAYA
	U+16E3C MWANGWEGO MODIFIER MI-HI-WAYA
	U+16E3D MWANGWEGO MODIFIER NI-HI-WAYA
	U+16E3E MWANGWEGO MODIFIER SISA-NI-HI
	U+16E3F MWANGWEGO MODIFIER MYU-NI-HI-WAYA

Table 3. Character Names of the Mwangwego Orthography

The Mwangwego script is used for the languages of Malawi, which include Chewa (nya), Lomwe (ngl), Sena (swk in Malawi, seh in Mozambican, bwg for the divergent Barwe dialect), Tonga (tog), Tumbuka (tum), Yao (yao), Nyakyusa (nyy).



Letters (Musisi)

16E00 MWANGWEGO LETTER A
16E01 MWANGWEGO LETTER BA
16E02 MWANGWEGO LETTER CHA
16E03 MWANGWEGO LETTER DA
16E04 MWANGWEGO LETTER FA
16E05 MWANGWEGO LETTER GA
16E06 MWANGWEGO LETTER GHA
16E07 MWANGWEGO LETTER HA
16E08 MWANGWEGO LETTER JA
16E09 MWANGWEGO LETTER ZHA
16E0A MWANGWEGO LETTER KA
16E0B MWANGWEGO LETTER LA
16E0C MWANGWEGO LETTER MA
16E0D MWANGWEGO LETTER NA
16E0E MWANGWEGO LETTER NYA
16E0F MWANGWEGO LETTER PA
16E10 MWANGWEGO LETTER RA
16E11 MWANGWEGO LETTER SA
16E12 MWANGWEGO LETTER SHA
16E13 MWANGWEGO LETTER TA
16E14 MWANGWEGO LETTER TSA
16E15 MWANGWEGO LETTER PSA
16E16 MWANGWEGO LETTER VA
16E17 MWANGWEGO LETTER WA
16E18 MWANGWEGO LETTER YA
16E19 MWANGWEGO LETTER ZA
16E1A MWANGWEGO LETTER DZA
16E1B MWANGWEGO LETTER DHLA
16E1C MWANGWEGO LETTER HLA
16E1D MWANGWEGO LETTER XA
16E1E MWANGWEGO LETTER QA
16E1F MWANGWEGO LETTER THA

Vowel Signs (Siri)

16E20 MWANGWEGO VOWEL SIGN EMWA
16E21 MWANGWEGO VOWEL SIGN ITA
16E22 MWANGWEGO VOWEL SIGN OTA
16E23 MWANGWEGO VOWEL SIGN UYU

Modifiers (Mutuyo)

16E24 MWANGWEGO MODIFIER MYU
16E25 MWANGWEGO MODIFIER MI
16E26 MWANGWEGO MODIFIER NI
16E27 MWANGWEGO MODIFIER HI
16E28 MWANGWEGO MODIFIER WAYA
16E29 MWANGWEGO MODIFIER WAYA BELOW
16E2A MWANGWEGO MODIFIER SISA
16E2B MWANGWEGO MODIFIER MURA
16E2C MWANGWEGO MODIFIER MULA
16E2D MWANGWEGO MODIFIER PEWA
16E2E MWANGWEGO MODIFIER TUMBU
16E2F MWANGWEGO MODIFIER KWANTHU

Compound Modifiers (Mituyo)

16E30 MWANGWEGO MODIFIER MYU-NI
16E31 MWANGWEGO MODIFIER MYU-HI
16E32 MWANGWEGO MODIFIER MYU-WAYA
16E33 MWANGWEGO MODIFIER MI-HI
16E34 MWANGWEGO MODIFIER MI-WAYA
16E35 MWANGWEGO MODIFIER NI-HI
16E36 MWANGWEGO MODIFIER NI-WAYA
16E37 MWANGWEGO MODIFIER HI-WAYA
16E38 MWANGWEGO MODIFIER SISA-NI
16E39 MWANGWEGO MODIFIER TUMBU-WAYA
16E3A MWANGWEGO MODIFIER MYU-NI-HI
16E3B MWANGWEGO MODIFIER MYU-HI-WAYA
16E3C MWANGWEGO MODIFIER MI-HI-WAYA
16E3D MWANGWEGO MODIFIER NI-HI-WAYA
16E3E MWANGWEGO MODIFIER SISA-NI-HI
16E3F MWANGWEGO MODIFIER MYU-NI-HI-WAYA

Additional information on characters

Phonetic value

The Mutuyo TUMBU ˈ is only used for Yao, and Nyakyusa/Ngonde. The Mutuyo PEWA ◌̂ is mostly used in Tumbuka, Kyangonde (a dialect of Nyakyusa; possibly other Nyakyusa dialects as well), and Yao.

Joining information

Masiri (e.g., Vowel modifiers; singular: Siri) are attached to base letters as described above, but letters themselves do not join at all like they do in Arabic.

For the MUSISI BA (Ʊ), and the Misiri counterparts (Ʊ, Ʊ, Ʊ, Ʊ), placement of the Mutuyo PEWA ◌̂ should be above the right-side vertical line (Ʊ̂ Ʊ̂ Ʊ̂ Ʊ̂ Ʊ̂), instead of directly above the grapheme as it is for all other graphemes (e.g., Ȃ, ʃ̂). See line 3 of Figure 10.1 in the Google Drive in Section IX.

Note for typographers/font designers: the placement of the Mutuyo/Mituyo in relation to the graphemes (Misisi/Misiri) should be done the same as in Figures 3.1-3.22 and 10.1-10.9 in the Google Drive. These images should be the reference when making basic letterforms and relative position of most Mutuyo/Mituyo to the graphemes. The following are exceptions to this based on input by the user community:

- When Mutuyo/Mituyo that appear on the left of MUSISI PA ɗ, MUSISI TA ʖ, MUSISI TSA ʃ, and MUSISI DHLA ʚ, and their Misiri counterparts, they should make use of the empty space available for these graphemes. They should "fill up"/utilize the empty space. Again, this logic extends to the Misiri counterparts for these 4 graphemes.
    - See Figure 10.5 in the Google Drive for MUSISI PA ɗ. Same for Misiri ɗ, ɗ, ɗ, ɗ.
    - See Figures 10.6-10.7 in the Google Drive for MUSISI TA ʖ. Same for Misiri ʖ, ʖ, ʖ, ʖ.
    - See Figure 10.7 in the Google Drive for MUSISI TSA ʃ. Same for Misiri ʃ, ʃ, ʃ, ʃ.
    - See Figure 10.8 in the Google Drive for MUSISI DHLA ʚ. Same for Misiri ʚ, ʚ, ʚ, ʚ.

- The stacked Mituyo combinations on the left of graphemes should be perfectly symmetrical as requested by the user community:
    - Symmetrical examples: ‼ ⁝ ⁝ ‼ ‼ ‼ ‼ ‼ ‼ ‼ ‼ ‼
    - The font used in Figures 10.1-10.9 in the Google Drive does have some aspects of a Mituyo stack off-centered, such as Mutuyo MYU ʿ. However, the examples above should be the reference for this aspect of the script/future fonts.

Order Sequence

After a survey conducted with script users, we propose an encoding order of Mwangwego glyphs to follow a "Modifier;Consonant;Vowel;Modifier-Combining" (MCVMc) order. This is in contrast with other possible orders such as "Consonant;Vowel;Modifier-Combining" (CVMc), and "Consonant;Modifier-Combining;Vowel" (CMcV). In the MCVMc order, the elements appear in the lef-to-right reading order and the combining Mituyo which are always terminal. Table 4 presents two examples:

	MCVMc (proposed)	CVMc	CMcV
‼ʔɔ	‼ + ʔ + ◌ɔ	‼ + ʔ + ◌ɔ	ʔ + ‼◌ + ◌ɔ
‼ʔ̱	‼ + ʔ + ◌̱	‼ + ʔ + ◌̱	ʔ + ‼◌ + ◌̱
‼ʔ̱ɔ	‼ + ʔ + ◌̱ + ◌ɔ	‼ + ʔ + ◌ɔ + ◌̱	ʔ + ‼◌ + ◌̱ + ◌ɔ

Table 4. Comparison of the possible character encoding sequences of Mwangwego letter elements.

Punctuation

There are no script-specific punctuation marks. Latin punctuation is used.

Numbers

There are no script-specific numerals/digits. Hindu-Arabic numerals/digits (0123456789) are used.

Mutuyo glyph uniqueness:

Many Mutuyo glyphs are visually similar to existing diacritical marks presently in the Unicode inventory. The authors have considered utilization of the graphically similar marks with the Mwangwego letters and determined that doing so is not in the best interests of the user community. While the marks are graphically similar, the abstract characters that the graphs are referents for are distinctly different; which in turn would lead to cognition difficulties in their Mwangwego utilization. A discussion and comparison of the diacritic and Mutuyo marks is given in the following.

The Chichewa and other languages of Malawi use a Latin-based orthography that includes "ŵ" to denote the voiced bilabial fricative /β/. Accordingly, the user community is already familiar with the shape of the mark and its phonetical augmentation to /w/. Circumflex applied to letters of the Mwangwego script would *not* signal the *same* phonetic change. Not only might this inconsistency in the use of the symbol be confusing to users, but visually the Latin circumflex is not optimal for publishing in Mwangwego script. The visual difference between the symbols is depicted in the following tables:

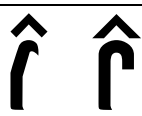
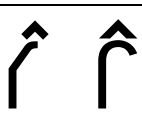
Reference Typeface	Ŵ ŵ	
Sans Serif Typeface	Ŵ ŵ	

Table 5. Comparison of Chichewa Latin Circumflex on Ŵ to Mutuyo Pewa.

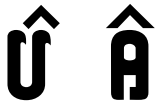
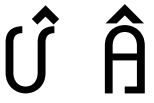
Reference Typeface	Û Â	
Sans Serif Typeface	Û Â	

Table 6. Comparison of Circumflex to Mutuyo Pewa applied to similarly shaped base letters.

Visible differences between the shapes of Circumflex and Pewa in the above examples include: stroke weight, stroke taper, inner angle, top angle, and mark positioning.

It can be expected that materials, particularly educational, will be produced that render Malawi languages in both orthographies. For these readers we believe that not conflating Circumflex and Pewa, for example, would better facilitate reading comprehension and the visual quality of publications. Maintaining the visual distinction between the symbols is also advantageous in any pain text context.

Unlike the practice with Latin diacritic, the combining Mutuyo do not change the phonetic value of a consonant letter (e.g. n + ◌̃ → ñ , phonetically /n/ → / ɲ/). Rather, they either append or prefix a phoneme to the consonant (note that with a single exception, the *non*-combining Mutuyo will prefix a phoneme).

For example:

- Ǝ (da) *vs* Ǝ̱ (dya)
- Ʊ (ba) *vs* Ʊ̏ (bra)
- ʔ (ka) *vs* ʔ̀ (kla)
- ʕ (za) *vs* ʕ̂ (bza)
- Ǝɹ (di) *vs* Ǝ́ɹ (ndi)

A noteworthy distinction between the application of the combining Mutuyo, that is unprecedented for the diacritical marks, appears in the last example above. The Mutuyo mark is the third glyph in the composition sequence. In keeping with the handwritten practices, the Mutuyo is written *after* the Masiri vowel mark while it visually appears above the consonant glyph -no different than when a Masiri is not present. Thus, the combining Mutuyo may be either the 2nd or 3rd codepoint in composed character (e.g. <C><M> or <C><V><M>). This is believed to be a different combining behavior than that exhibited by comparable diacritical marks which will always be the 2nd codepoint in a composition sequence (unless combining with another combinable mark).

This last point helps highlight that Latin based combining marks may be inherently unsuitable for Mwangwego script which, being an Abugida script, is not sufficiently *Latin-like*. Review of the Unicode Character Database's "NormalizationTest.txt" finds the comparable marks applied only to the more closely related scripts Latin, Greek, and Cyrillic. Applying the marks outside of this family may also be problematic for existing text processors that make the assumption that the marks are erroneously applied when found elsewhere.

A degree of internal similarity to the Mutuyo mark is readily observed. For example: ᶜ (MYU) with ᶜᶜ (MI), ˈ (TUMBU) with ˮ (NI), and ◌̀ (MULA) with ◌̏ (MURA). Discussion with the user community and educators we find that users embody a phonological understanding of the glyphs and do not apply the visual similarity in any practical way. For example, MI (ᶜᶜ) is not perceived as a "Double MYU" (ᶜ) which is not logical phonetically. Education of the marks also does not follow what would logically be the "stroke order" of the glyphs (ˮ NI precedes ˈ TUMBU, ◌̏ MURA precedes ◌̀ MULA). Thus, we have proposed encoding the marks distinctly in keeping with the user community's perception and practices. Doing so also avoids entirely the inevitable problem of text editing where during a "copy & paste" operation, one mark is selected and pasted but not the other, or a deletion is incomplete, or a second keystroke missed in their composition.

V Properties

General Category and other properties

The Mwangwego inventory in this proposal includes the Misisi syllable bases as type Letter Other, their combining Masiri and Mutuyo diacritic marks.

```
16E00;MWANGWEGO LETTER A;Lo;0;L;;;;;N;;;;;
16E01;MWANGWEGO LETTER BA;Lo;0;L;;;;;N;;;;;
16E02;MWANGWEGO LETTER CHA;Lo;0;L;;;;;N;;;;;
16E03;MWANGWEGO LETTER DA;Lo;0;L;;;;;N;;;;;
16E04;MWANGWEGO LETTER FA;Lo;0;L;;;;;N;;;;;
16E05;MWANGWEGO LETTER GA;Lo;0;L;;;;;N;;;;;
16E06;MWANGWEGO LETTER GHA;Lo;0;L;;;;;N;;;;;
16E07;MWANGWEGO LETTER HA;Lo;0;L;;;;;N;;;;;
16E08;MWANGWEGO LETTER JA;Lo;0;L;;;;;N;;;;;
16E09;MWANGWEGO LETTER ZHA;Lo;0;L;;;;;N;;;;;
16E0A;MWANGWEGO LETTER KA;Lo;0;L;;;;;N;;;;;
16E0B;MWANGWEGO LETTER LA;Lo;0;L;;;;;N;;;;;
16E0C;MWANGWEGO LETTER MA;Lo;0;L;;;;;N;;;;;
16E0D;MWANGWEGO LETTER NA;Lo;0;L;;;;;N;;;;;
16E0E;MWANGWEGO LETTER NYA;Lo;0;L;;;;;N;;;;;
16E0F;MWANGWEGO LETTER PA;Lo;0;L;;;;;N;;;;;
16E10;MWANGWEGO LETTER RA;Lo;0;L;;;;;N;;;;;
16E11;MWANGWEGO LETTER SA;Lo;0;L;;;;;N;;;;;
16E12;MWANGWEGO LETTER SHA;Lo;0;L;;;;;N;;;;;
16E13;MWANGWEGO LETTER TA;Lo;0;L;;;;;N;;;;;
16E14;MWANGWEGO LETTER TSA;Lo;0;L;;;;;N;;;;;
16E15;MWANGWEGO LETTER PSA;Lo;0;L;;;;;N;;;;;
16E16;MWANGWEGO LETTER VA;Lo;0;L;;;;;N;;;;;
16E17;MWANGWEGO LETTER WA;Lo;0;L;;;;;N;;;;;
16E18;MWANGWEGO LETTER YA;Lo;0;L;;;;;N;;;;;
16E19;MWANGWEGO LETTER ZA;Lo;0;L;;;;;N;;;;;
16E1A;MWANGWEGO LETTER DZA;Lo;0;L;;;;;N;;;;;
16E1B;MWANGWEGO LETTER DHLA;Lo;0;L;;;;;N;;;;;
16E1C;MWANGWEGO LETTER HLA;Lo;0;L;;;;;N;;;;;
16E1D;MWANGWEGO LETTER XA;Lo;0;L;;;;;N;;;;;
16E1E;MWANGWEGO LETTER QA;Lo;0;L;;;;;N;;;;;
16E1F;MWANGWEGO LETTER THA;Lo;0;L;;;;;N;;;;;
16E20;MWANGWEGO VOWEL SIGN EMWA;Mc;204;L;;;;;N;;;;;
16E21;MWANGWEGO VOWEL SIGN IMA;Mc;204;L;;;;;N;;;;;
16E22;MWANGWEGO VOWEL SIGN OTA;Mc;204;L;;;;;N;;;;;
16E23;MWANGWEGO VOWEL SIGN UYU;Mc;204;L;;;;;N;;;;;
16E24;MWANGWEGO MODIFIER MYU;Lm;0;L;;;;;N;;;;;
16E25;MWANGWEGO MODIFIER MI;Lm;0;L;;;;;N;;;;;
16E26;MWANGWEGO MODIFIER NI;Lm;0;L;;;;;N;;;;;
16E27;MWANGWEGO MODIFIER HI;Lm;0;L;;;;;N;;;;;
16E28;MWANGWEGO MODIFIER WAYA;Lm;0;L;;;;;N;;;;;
16E29;MWANGWEGO MODIFIER WAYA BELOW;Mn;220;NSM;;;;;N;;;;;
16E2A;MWANGWEGO MODIFIER SISA;Lm;0;L;;;;;N;;;;;
16E2B;MWANGWEGO MODIFIER MURA;Mn;230;NSM;;;;;N;;;;;
16E2C;MWANGWEGO MODIFIER MULA;Mn;230;NSM;;;;;N;;;;;
16E2D;MWANGWEGO MODIFIER PEWA;Mn;230;NSM;;;;;N;;;;;
16E2E;MWANGWEGO MODIFIER TUMBU;Lm;0;L;;;;;N;;;;;
```

```
16E2F;MWANGWEGO MODIFIER KWANTHU;Mn;230;NSM;;;;;N;;;;;
16E30;MWANGWEGO MODIFIER MYU-NI;Lm;0;L;;;;;N;;;;;
16E31;MWANGWEGO MODIFIER MYU-HI;Lm;0;L;;;;;N;;;;;
16E32;MWANGWEGO MODIFIER MYU-WAYA;Lm;0;L;;;;;N;;;;;
16E33;MWANGWEGO MODIFIER MI-HI;Lm;0;L;;;;;N;;;;;
16E34;MWANGWEGO MODIFIER MI-WAYA;Lm;0;L;;;;;N;;;;;
16E35;MWANGWEGO MODIFIER NI-HI;Lm;0;L;;;;;N;;;;;
16E36;MWANGWEGO MODIFIER NI-WAYA;Lm;0;L;;;;;N;;;;;
16E37;MWANGWEGO MODIFIER HI-WAYA;Lm;0;L;;;;;N;;;;;
16E38;MWANGWEGO MODIFIER SISA-NI;Lm;0;L;;;;;N;;;;;
16E39;MWANGWEGO MODIFIER TUMBU-WAYA;Lm;0;L;;;;;N;;;;;
16E3A;MWANGWEGO MODIFIER MYU-NI-HI;Lm;0;L;;;;;N;;;;;
16E3B;MWANGWEGO MODIFIER MYU-HI-WAYA;Lm;0;L;;;;;N;;;;;
16E3C;MWANGWEGO MODIFIER MI-HI-WAYA;Lm;0;L;;;;;N;;;;;
16E3D;MWANGWEGO MODIFIER NI-HI-WAYA;Lm;0;L;;;;;N;;;;;
16E3E;MWANGWEGO MODIFIER SISA-NI-HI;Lm;0;L;;;;;N;;;;;
16E3F;MWANGWEGO MODIFIER MYU-NI-HI-WAYA;Lm;0;L;;;;;N;;;;;
```

Line breaking information

Line Breaking rules for the Mwangwego script are as follows:

- Words are separated with U+0020 SPACE. There are no other word separators in historical documents.
- Line breaks only occur at word boundaries.
- There is not a special mode found that allows line breaks within words at select positions, such as using a hyphen sign (U+002D HYPEHN-MINUS) that other scripts apply.
- Line breaks cannot occur within numbers. Numbers must always be kept together.
- There are no restrictions explicitly stated on line breaking before or after certain punctuation characters.

The Mwangwego script is likely to occur with the Latin script.

VI Collation

Misisi/Musiri

The sorting order of the basic syllabic graphemes of the script follows the ordering used in education and is shown in Table 1, starting with 𖼀, 𖼁, 𖼂, 𖼃, 𖼄, then 𖼅, 𖼆, 𖼇, 𖼈, 𖼉, …, all the way to 𖼏, 𖼐, 𖼑, 𖼒, 𖼓.

However, attaching Mutuyo/Mituyo to Misisi/Musiri significantly influences the ordering of said Misisi/Musiri; see directly below.

Mutuyo

When Mutuyo (i.e., a single modifying mark as defined in Section III) are attached to Misisi/Musiri the collation adheres to the following order:

MYU ʻ < MI ‟ < NI " < HI ¨ < WAYA - < WAYA BELOW ◌̱ < SISA ʼ < MURA ◌̏ < MULA ◌̀ < PEWA ◌̂ < TUMBU ˈ < KWANTHU ◌́.

As previously mentioned, not all Mutuyo/Mituyo attach to all Misisi/Musiri as some letter+modifier combinations are not linguistically attested. For example, the valid Mutuyo for the Musisi ʔ /ka/ are MYU ʻ, NI ", HI ¨, WAYA -, WAYA BELOW ◌̱, SISA ʼ, and KWANTHU ◌̂. See Section X for all valid Misisi and Mutuyo/Mituyo combinations.

As an example, if one is given the Musisi ʔ /ka/ and the Musiri ꝛ /kɛ/, and all valid Mutuyo attachments for ʔ, which are -ʔ, "ʔ, ʔ̱, ¨ʔ, ʻʔ, ʼʔ, and ʔ́ the order of these characters would be as follows:

ʔ, ʻʔ, "ʔ, ¨ʔ, -ʔ, ʔ̱, ʼʔ, ʔ̏, ʔ́, ꝛ

The basic Musisi ʔ /ka/ is ordered first, then each base-ʔ letter is ordered according to the Mutuyo order mentioned above. Only after all letters of the same base (ʔ in this example) are ordered, can the next vocalized letter (i.e., ꝛ /kɛ/) be ordered.

To provide another example: if one is given the Musisi Ʊ /ba/ and the Misiri Ա /bɛ/, Ɯ /bi/, Ɓ /bɔ/, and Ƴ /bu/. The valid Mutuyo for Ʊ /ba/ (which are also ***always*** valid for the Misiri counterparts) are MYU ʻ, MI ‟, WAYA -, WAYA BELOW ◌̱, MURA ◌̏, MULA ◌̀, and PEWA ◌̂ (see Section X). If you attach all Mutuyo to Musisi Ʊ /ba/ and all Misiri, the ordering of all of these characters, including the base Musisi and Misiri, would be as follows:

Ʊ < ʻƱ < ‟Ʊ < -Ʊ < Ʊ̱ < Ʊ̏ < Ʊ̀ < Ʊ̂ < Ա < ʻԱ < ‟Ա < -Ա < Ա̱ < Ա̏ < Ա̀ < Ա̂ < Ɯ < ʻƜ < ‟Ɯ < -Ɯ < Ɯ̱ < Ɯ̏ < Ɯ̀ < Ɯ̂ < Ɓ < ʻƁ < ‟Ɓ < -Ɓ < Ɓ̱ < Ɓ̏ < Ɓ̀ < Ɓ̂ < Ƴ < ʻƳ < ‟Ƴ < -Ƴ < Ƴ̱ < Ƴ̏ < Ƴ̀ < Ƴ̂

Mituyo

Graphemes with only one Mutuyo are ordered before any instances of a grapheme with Mituyo (i.e., multiple Mutuyo).

There are only three ways to configure a Mituyo combination around a Musisi/letter:

1. **Left** and Bottom (example: "ʕ̱)
2. **Top** and Left (example: -ʕ̂)
3. **Top** and Bottom (example: ʕ̱̂)

You cannot have a Mituyo configuration where all three positions (top, left, and bottom) are occupied by a Mutuyo (i.e., it is linguistically impossible for these languages). The Mutuyo that is written first, and therefore considered first when ordering is in **bold** above for each of the three configurations.

Furthermore, if there are multiple Mutuyo stacked on the left of a grapheme (e.g., "ʕ, ¨ʕ) ordering of the left stack is decided by reading the individual Mutuyo from the top of the stack (first) to the bottom of the stack(second), and adhering to the general ordering of individual Mutuyo

2-part Mituyo

Below are all 19 linguistically possible 2-part Mituyo combinations. This is also their order (from left to right):

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
																		

Table 7. Order of all 2-part Mituyo.

For convenience, the order of individual Mutuyo is: MYU ‘ < MI " < NI " < HI ¨ < WAYA - < WAYA BELOW ◌̲ < SISA ’ < MURA ◌̰ < MULA ◌̀ < PEWA ◌̂ < TUMBU ' < KWANTHU ◌́.

Therefore, we can see in Table 7 that all 2-part Mituyo that contain a MYU ‘ are ordered first before 2-part Mituyo that “start with” a MI ", etc.

Furthermore, as we can see in Table 7, the 2-part Mituyo combinations that are labeled as 1, 2, and 3 are all “stacks” that would appear on the left side of a grapheme. Stacks 1, 2, and 3 all have a MYU ‘ at the top of their stacks. The ordering of these 3 stacks is determined by parsing the stacks’ Mutuyo components from top to bottom. They each start with MYU ‘, which can be ignored. The second Mutuyo in stacks 1, 2, and 3 are NI ", HI ¨, and WAYA -,respectively. As NI " comes before HI ¨ and WAYA - in the Mutuyo ordering,  is ordered before  and  . Then, as HI ¨ comes before WAYA - in the Mutuyo ordering, is ordered before .

Another example are the 2-part Mituyo combinations labeled under 14, 15, and 16 in Table 7 ( ,  , and ). They have top-left, top-left, and top-bottom configurations around a grapheme, respectively. They each begin with a PEWA ◌̂. Their order is also determined by looking at the second Mutuyo in the Mityuo combination. Combination 14 has a NI ", 15 has a WAYA -, and 16 has a WAYA BELOW ◌̲ as their second Mutuyo (according to the order of reading their respective configurations). As NI " comes before WAYA - in the Mutuyo ordering, and WAYA - comes before WAYA BELOW ◌̲, their order as  < < .

3-part Mituyo

Below are the 7 linguistically possible 3-part Mituyo combinations. This is also their order (from left to right):

1	2	3	4	5	6	7
						

Table 8. Order of all 3-part Mituyo

The exact same logic applies to 3-part Mituyo.

4-part Mituyo

There is only one linguistically possible 4-part Mituyo combination. It is ordered after 3-part Mituyo combinations.

1
[illegible]

Table 9. The sole 4-part Mituyo combination.

The exact same logic applies to 4-part Mituyo.

Complete collation example

1. A basic grapheme (Misisi) is ordered first.
    a. Example: [illegible] /ka/
2. Then all valid Mutuyo for that basic grapheme are ordered.
    a. Example: [illegible]
3. Then all valid 2-part Mutuyo for that basic grapheme are ordered.
    a. Example: [illegible]
4. Then all valid 3-part Mituyo for that basic grapheme are ordered.
    a. Example: [illegible]
5. Then all valid 4-part Mituo for that basic grapheme are ordered.
    a. Example: [illegible]
6. Then the next vocalization for that basic grapheme is ordered.
    a. Example: [illegible] /ke/
7. Then the ordering continues for all the same Mutuyo, and 2/3/4-part Mituyo for that new vocalized grapheme.
8. Then the next basic grapheme (Misisi) continues the entire cycle.

Section X shows every possible Mutuyo/Mituyo combination attached to the Misisi (i.e., all inherent /a/ syllabic graphemes).

VII References

1. Mwangwego, Nolence. 2011. [illegible] (A Malawi tili pati? Wolemba ndi Nolensi Mose Mwangwegho). Blantyre: [self-published; printed by Blantyre Print and Packaging].
2. Everson, Michael. “Proposal to Encode the Mwangwego Script in the UCS.” The Unicode Consortium, September 25, 2012. http://www.unicode.org/wg2/docs/n4323.pdf.
3. Mwanwego, Nolence. “Part 1 Mwangwego Script.” YouTube, September 18, 2015. https://youtu.be/ShOzBdgJGgo?si=dBzx4KGTWjaSDR8a.
4. Mwangwego, Nolence. “Part 2 of 4 Mwangwengo Script.” YouTube, September 18, 2015. https://youtu.be/UyzLbGVBgog?si=1Ur6qTy_74Yik3Uo.
5. Mwangwego, Nolence. “Part 3of 4 Mwangwego Script.” YouTube, September 19, 2015. https://youtu.be/uaiEJfvChSw?si=xErrtIiIloqjaXu6.
6. Mwangwego, Nolence. “Part 4 of 4 Mwangwego Script.” YouTube, September 19, 2015. https://youtu.be/rC5k-Y5RhJU?si=3lVCBYRjOmpDHLYz.
7. https://www.facebook.com/groups/455153627868133/permalink/7667291373320953/
8. Dobrovolny, Michelle, “The Politics of Malawi’s Alphabet”. https://web.archive.org/web/20160304000609/http:/speakjhr.com/2010/11/the-politics-of-malawis-alphabet/ November 26, 2010.
9. Kostelnik, Jannis. "Technological Adaptation and the Role of Social Media in the Proliferation of the Newly Created Minim Script from Burkina Faso." Ahyu: A Journal of Language and Literature 6 (2023): 71-87.

VIII Acknowledgements

We would like to humbly thank Nolence Mwangwego and the Mwangwego script user community that have immensely helped us understand the nuances of this script.

An endorsement letter for the encoding of the Mwangwego script written by a University of Malawi linguist is attached/included on the next page.

This project was made possible in part by a grant from the Mellon Foundation to the Script Encoding Initiative at the University of California, Berkeley.

Endorsement/Recommendation for the Mwangwego script

I, the undersigned, would like to recommend the Mangwego script to be included into the Unicode and eventually be displayed on all digital devices.

When the Mwangwego script was launched in the late 90s, I happened to be one of its first learners.

The script is so easy to learn and to use and I wholeheartedly endorse its inclusion in the Unicode.

M.Chimasula

Mazunzo Prince Chimasula

Lecturer of French at University of Malawi
princechimasula@gmail.com

The views expressed herein do not represent my employer. They are expressed in personal capacity as a student of the script.

IX Examples

Figure 1. Excerpt of Genesis 1 from a Tumbuka Bible. A Latin orthography transliteration can be found in the Google Drive.

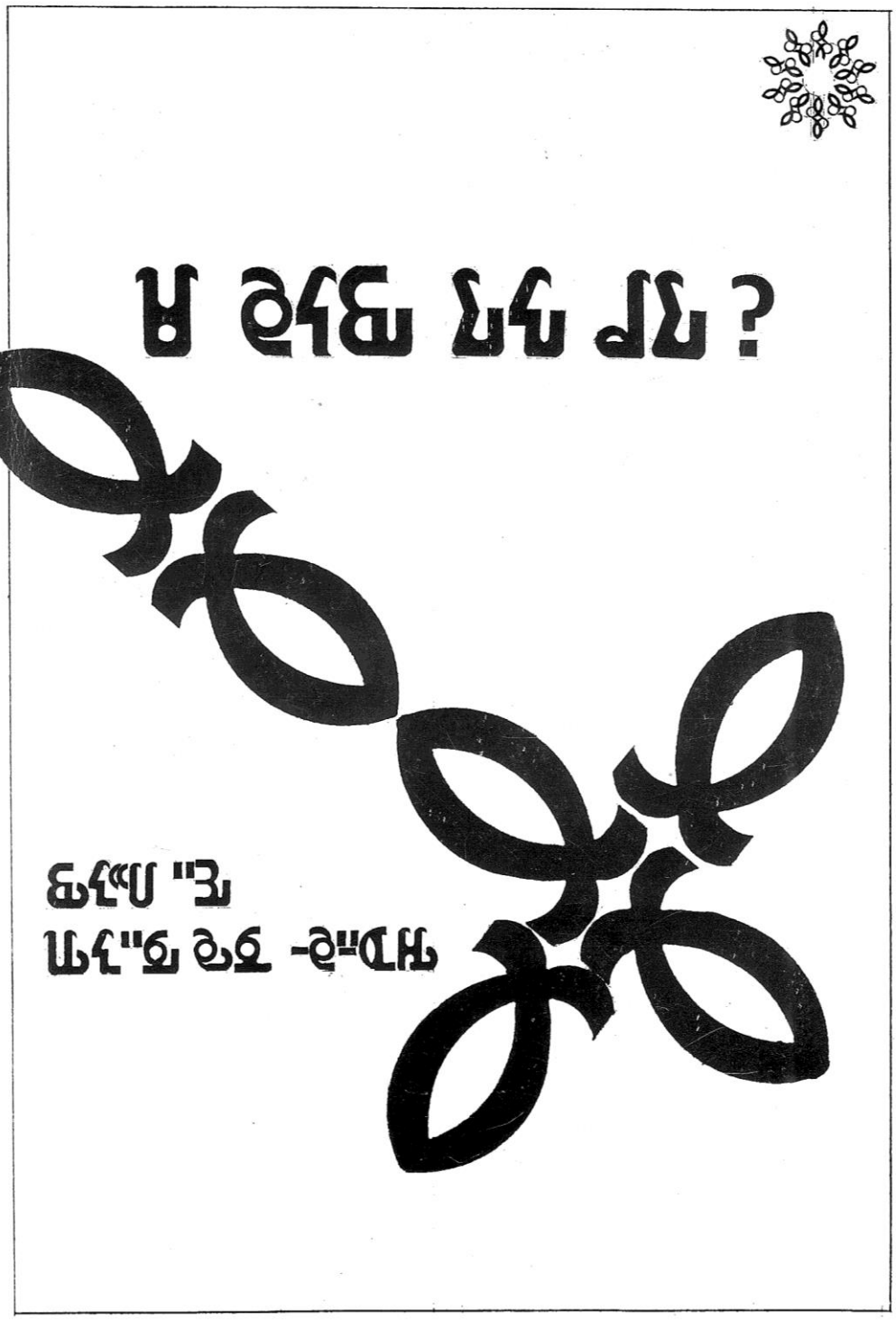
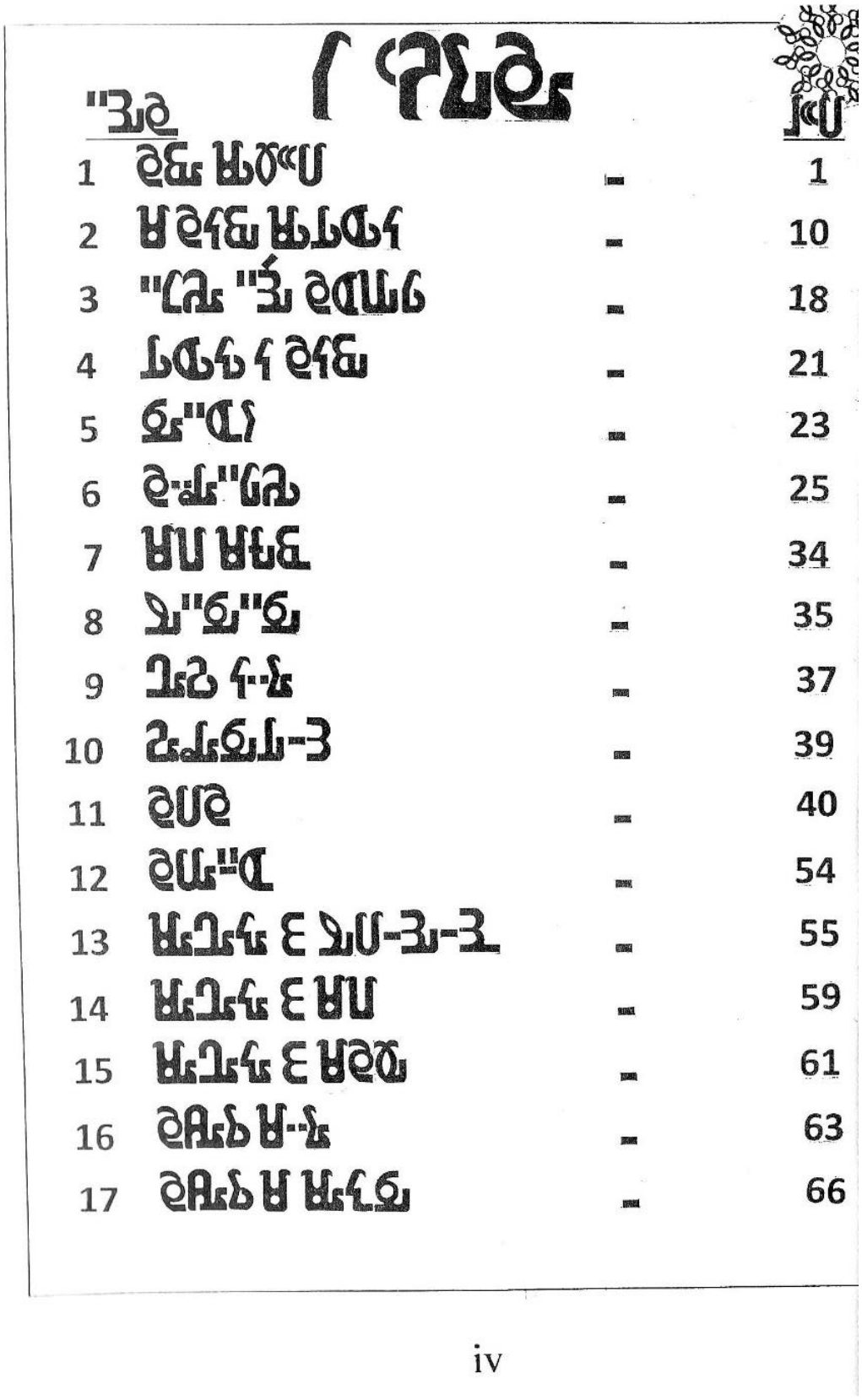


Figure 2.1. Cover page of the Chichewa language book “Ꮋ ᏣᎶᏮ ᏁᎶ ᎴᏁ?” (transliteration: “*A Malawi Tili Pati*”; translation: “*Malawians, where are we?*”). Mwangwego, Nolence. 2011. Ꮋ ᏣᎶᏮ ᏁᎶ ᎴᏁ? ᏫᎶᏬᏍ ᏋᏯ ᏓᎶᏬᏮ ᏅᏧ ᏣᏬᏍᏆᏂᎶ (A Malawi tili pati? Wolemba ndi Nolensi Mose Mwangwegho). Blantyre: [self-published; printed by Blantyre Print and Packaging].

iii

Figure 2.2. “ᎯᏳᎩᏁᏂᏗ ᎺᏯ” page iii.



ᏗᏟᏍᏗᏓᏂᏙᏗ

"ᏓᏥ		ᏥᏌ
1	ᎾᎿ ᏂᏚᎵᏍᏗ	1
2	ᏂᎦᏛᏁᎲ ᏂᏓᏆᏍᏗᏱ	10
3	"ᏣᏓᏅ "ᏱᏗ ᎾᏤᏃᎮᏍ	18
4	ᏗᏆᏍᏗᏱ ᏧᏛᏁᎲ	21
5	ᎧᏅ"ᏣᏘ	23
6	ᎾᏍᏅ"ᏱᏓᏂ	25
7	ᏂᏙᏂ ᏂᏐᏋᏢ	34
8	ᏓᏁ"ᎧᏁ"ᎧᏓ	35
9	ᏧᏂᏍ ᏱᏍᏎ	37
10	ᏕᏎᏆᏓᏁᏗᏱ	39
11	ᎾᏆᎾ	40
12	ᎾᏆᏎ"ᏣᏘ	54
13	ᏎᏗᏍᏗᏎ Ꭼ ᏓᏆᏍᏗᏱᏍᏗ	55
14	ᏎᏗᏍᏗᏎ Ꭼ ᏂᏙᏂ	59
15	ᏎᏗᏍᏗᏎ Ꭼ ᏂᏧᏣᏯ	61
16	ᎾᏨᏁᏖ ᏂᏍᏎ	63
17	ᎾᏨᏁᏖ ᏂᏧ ᏎᏣᏫᏯ	66

iv

Figure 2.3. Table of contents of “Ꮀ ᏣᏁᎩ ᏗᏂᏁ ᏧᏁᏏ?”, page iv.

Figure 2.4. Page 1 of “[illegible]”

Figure 3. Handwritten Chichewa letter from 2012. The author was from Blantyre; the recipient was in Lilongwe. The topic is on agricultural practices.

Figure 4. A letter from 2002. Written from someone in Blantyre to another person in Blantyre. The topic is on literacy.

Figure 5. Page from a very long diary that was mailed from the author (in Malawi) to other user(s) (in South Africa) for interpersonal communication/text interchange. A 40-page excerpt of the longer text is in the Google Drive.

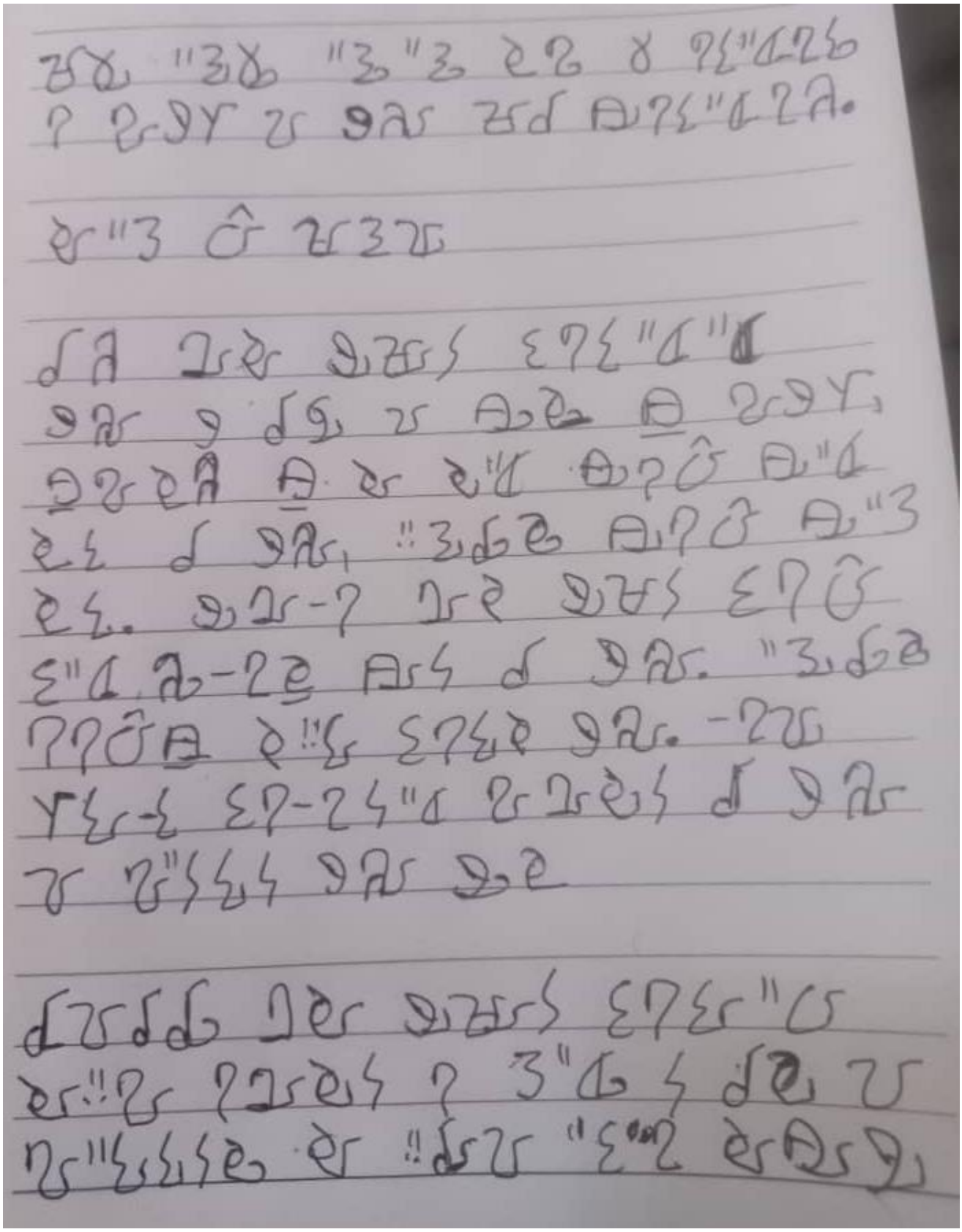


Figure 6 Handwritten journal entry from a user in Mzuzu.

Figure 7. Chichewa text on literacy, what it means to be literate, etc. Written around 2015.

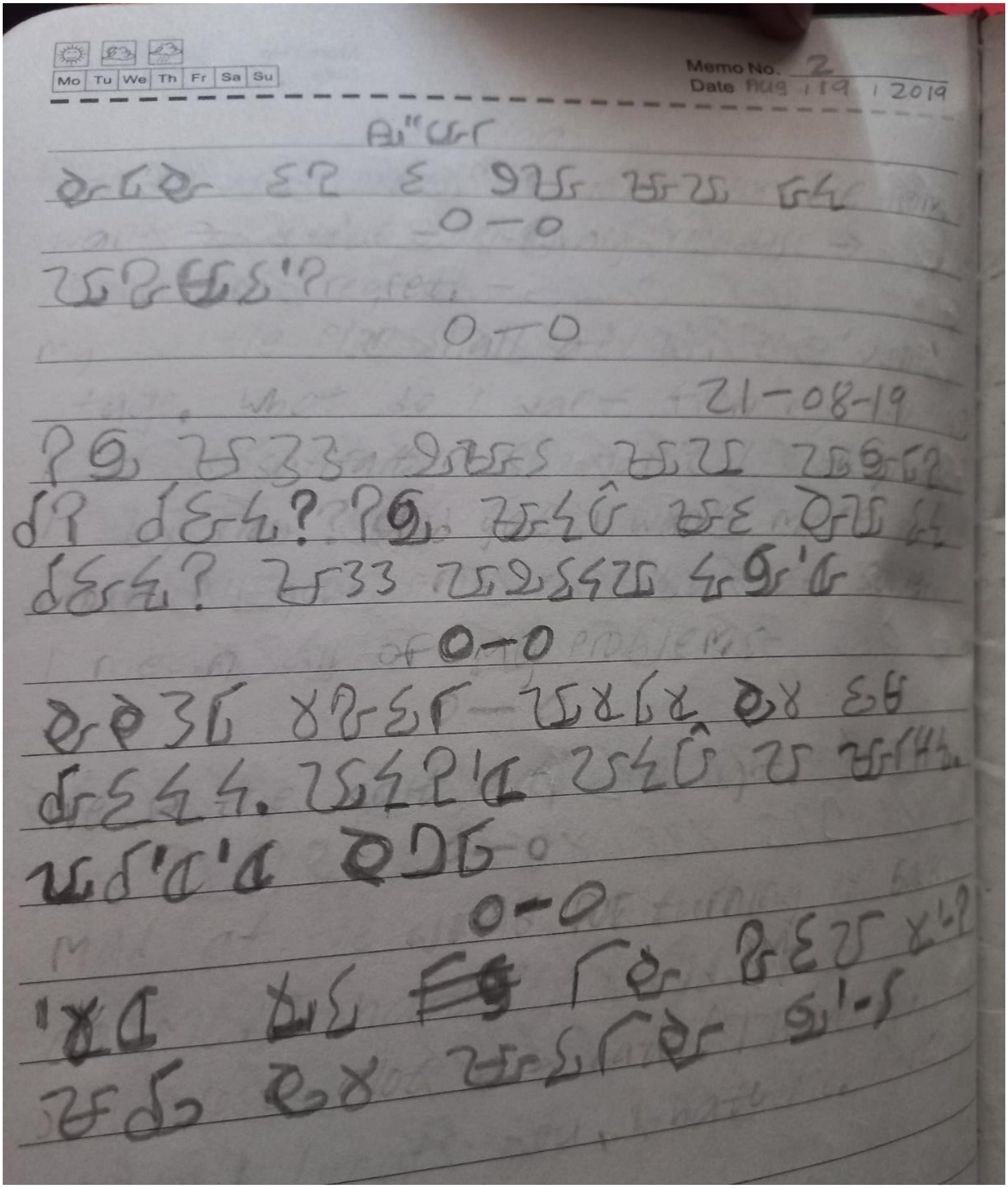


Figure 8. August 19, 2019 Tumbuka-language journal entry from a Malawian user in Mzuzu.

Please note, the last word on this image, [illegible], has the Mituyo stack deconstructed, with the Tumbu Mutuyo incorrectly written on the left of the Waya Mutuyo, instead of properly placing it on top of the latter. This is an error the student made, which is part of the learning curve. There are several instances of the user *correctly* constructing this same Mituyo in these journal entries in Figure 4.6 (twice in paragraph 3; again in paragraph 4).

Figure 9. Chichewa text about changing mindsets of Malawians. Written in early February 2025.

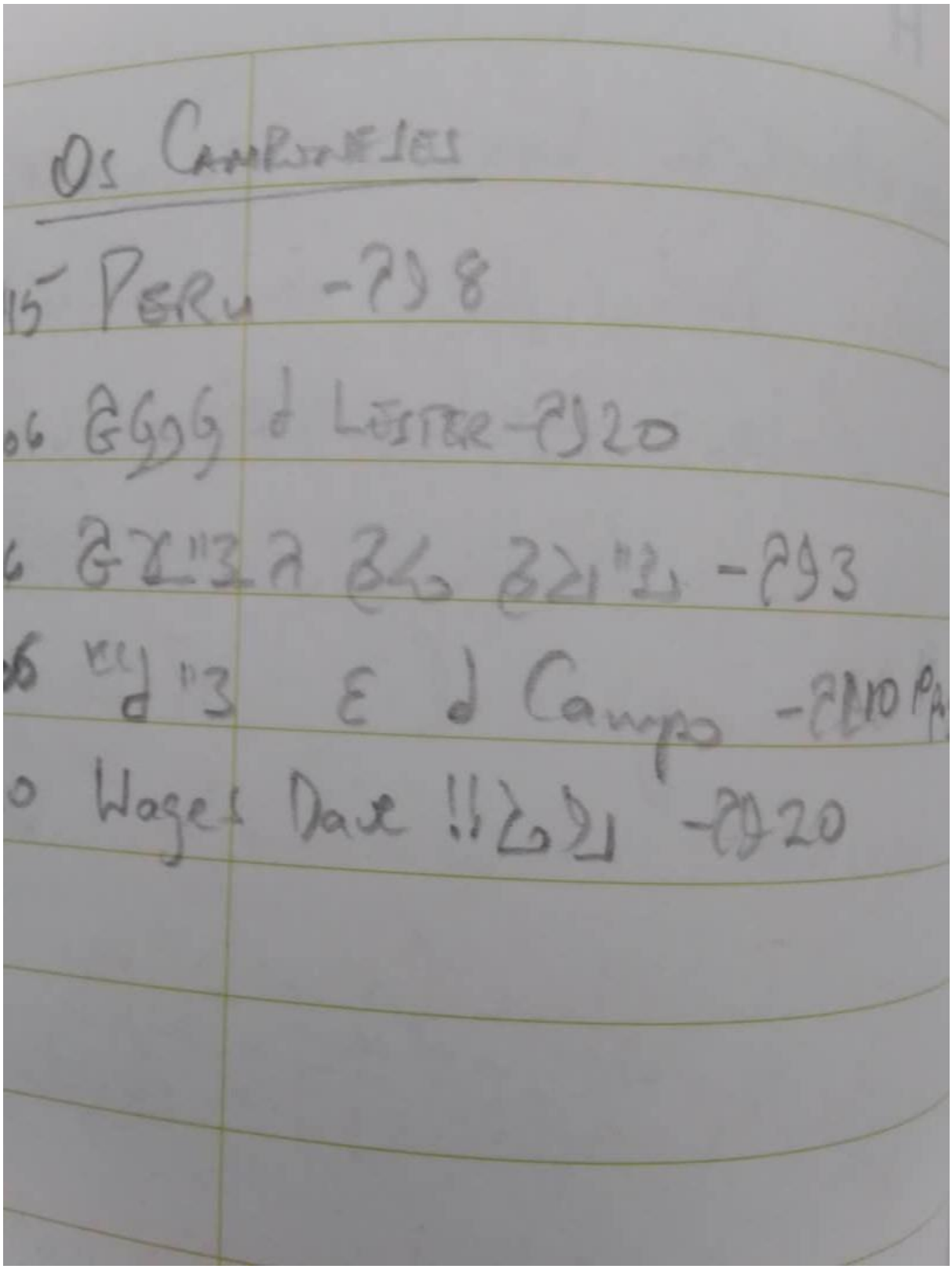


Figure 10. List in Chichewa.

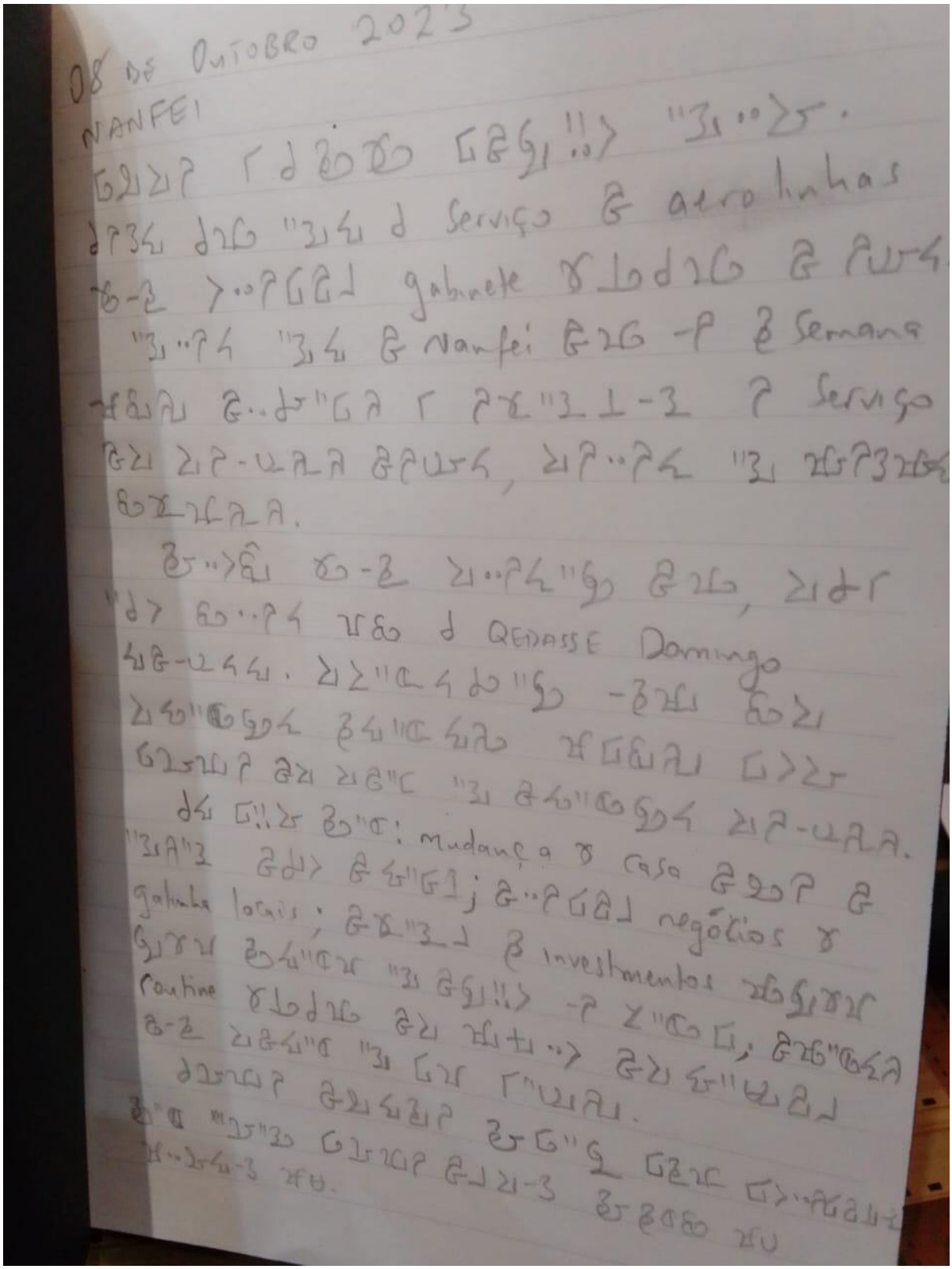


Figure 11. 2023 Chichewa-language handwriting by a Malawian user. This image showcases some intralinear Portuguese words in the Latin script (important to typographers) as the user is a businessman that conducts international business in Portuguese. The topic is business.

Figure 12. A journal entry from a user in Blantyre.

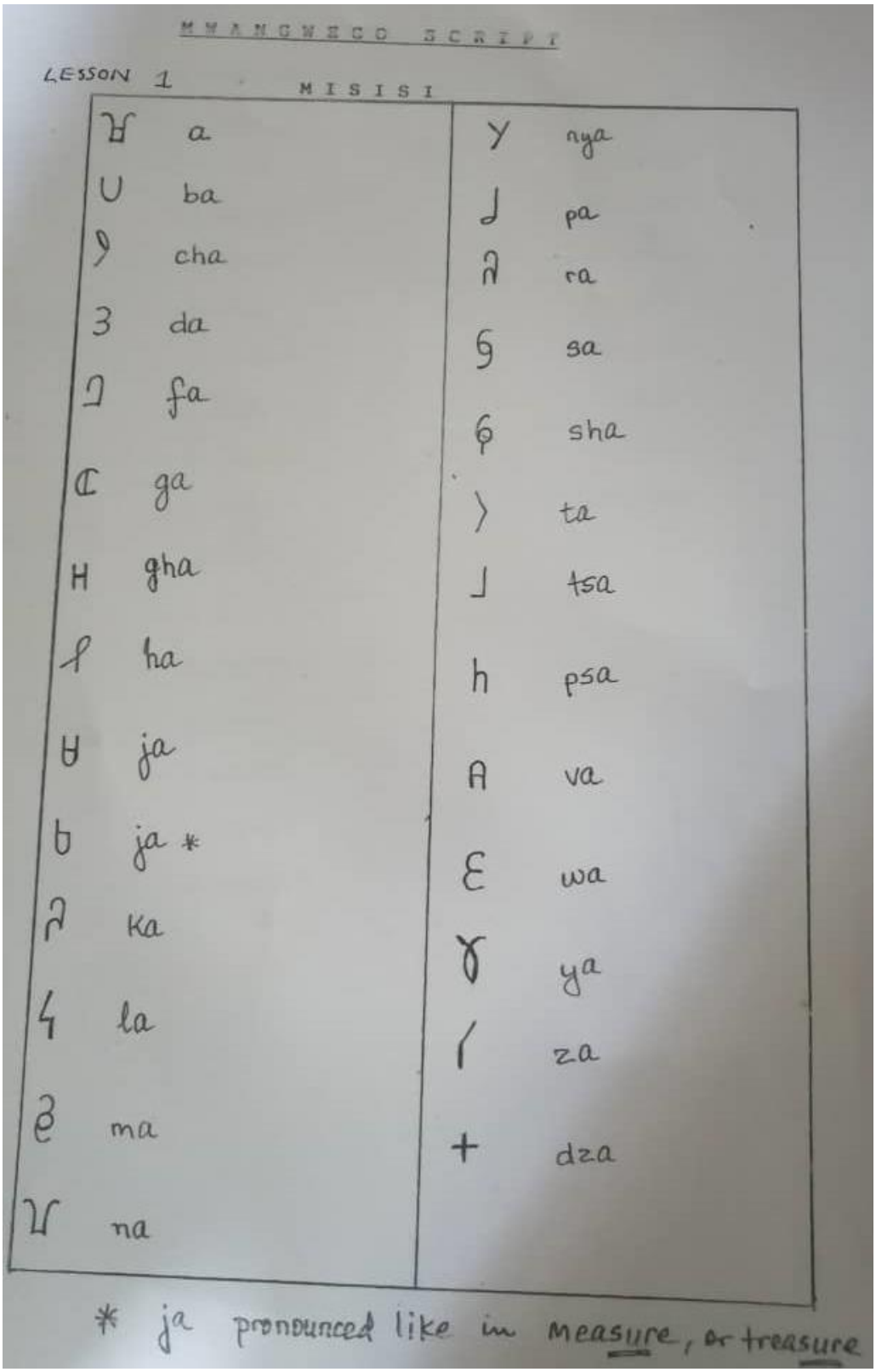


Figure 13. Educational material for Misisi (lesson 1).

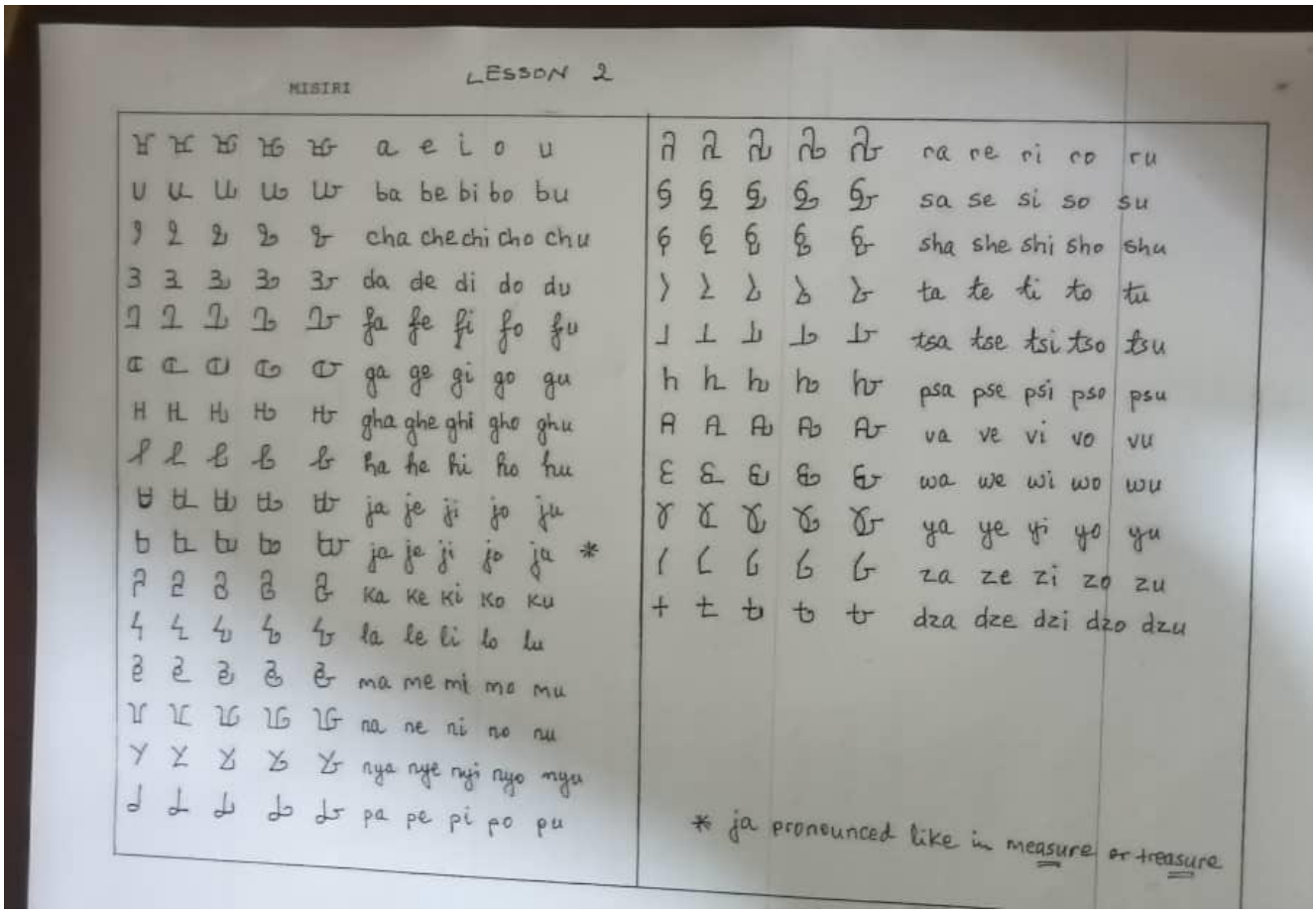
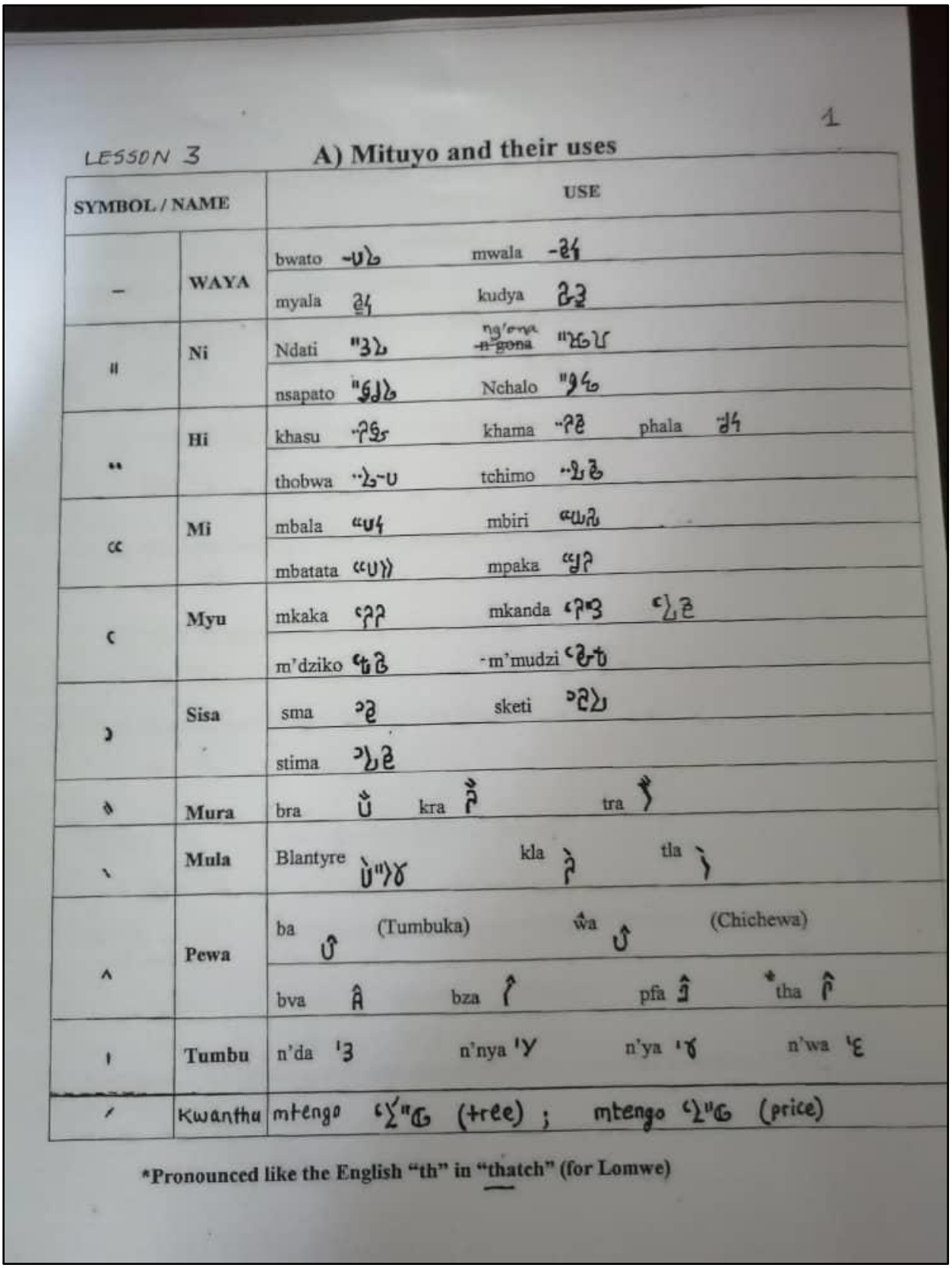


Figure 14. Educational material for Misiri (lesson 2).



1

LESSON 3

A) Mituyo and their uses

SYMBOL / NAME		USE
–	WAYA	bwato mwala myala kudya
‖	Ni	Ndati ng'ona / n'gona nsapato Nchalo
"	Hi	khasu khama phala thobwa tchimo
‹‹	Mi	mbala mbiri mbatata mpaka
‹	Myu	mkaka mkanda m'dziko m'mudzi
›	Sisa	sma sketi stima
◊	Mura	bra kra tra
\	Mula	Blantyre kla tla
^	Pewa	ba (Tumbuka) ŵa (Chichewa) bva bza pfa *tha
ı	Tumbu	n'da n'nya n'ya n'wa
′	Kwantha	mtengo (tree) ; mtengo (price)

*Pronounced like the English "th" in "thatch" (for Lomwe)

Figure 15. Educational material for Mituyo (lesson 3).

Figure 16. Educational material of paragraphs used for reading practice. The topic is on slavery and colonization.

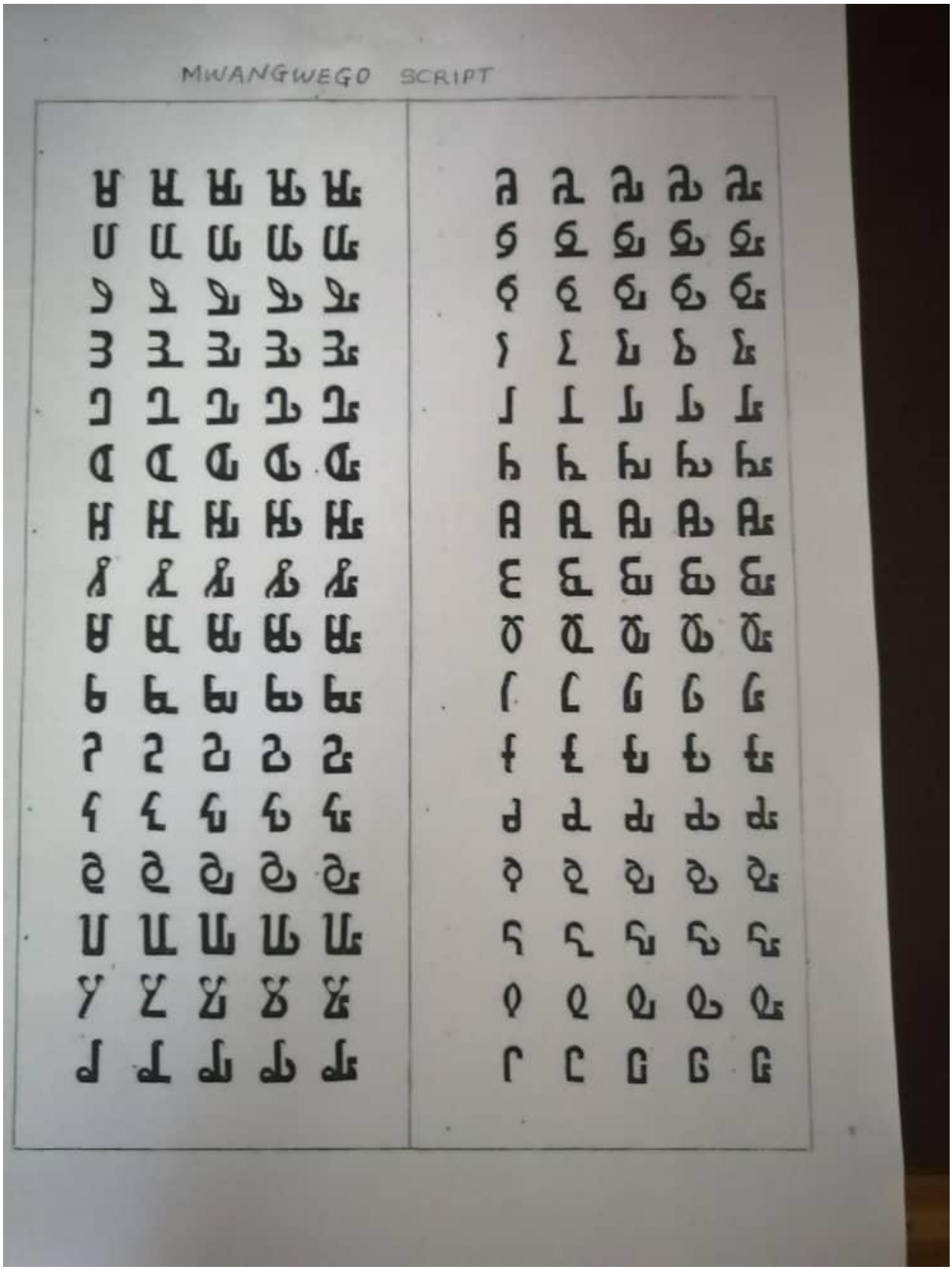


Figure 17. Educational material used to teach the basic syllabic graphemes of the Mwangwego script.

Figure 18. Most of the possible Mutuyo & Mituyo attachments onto Misisi (not according to collation/ordering).

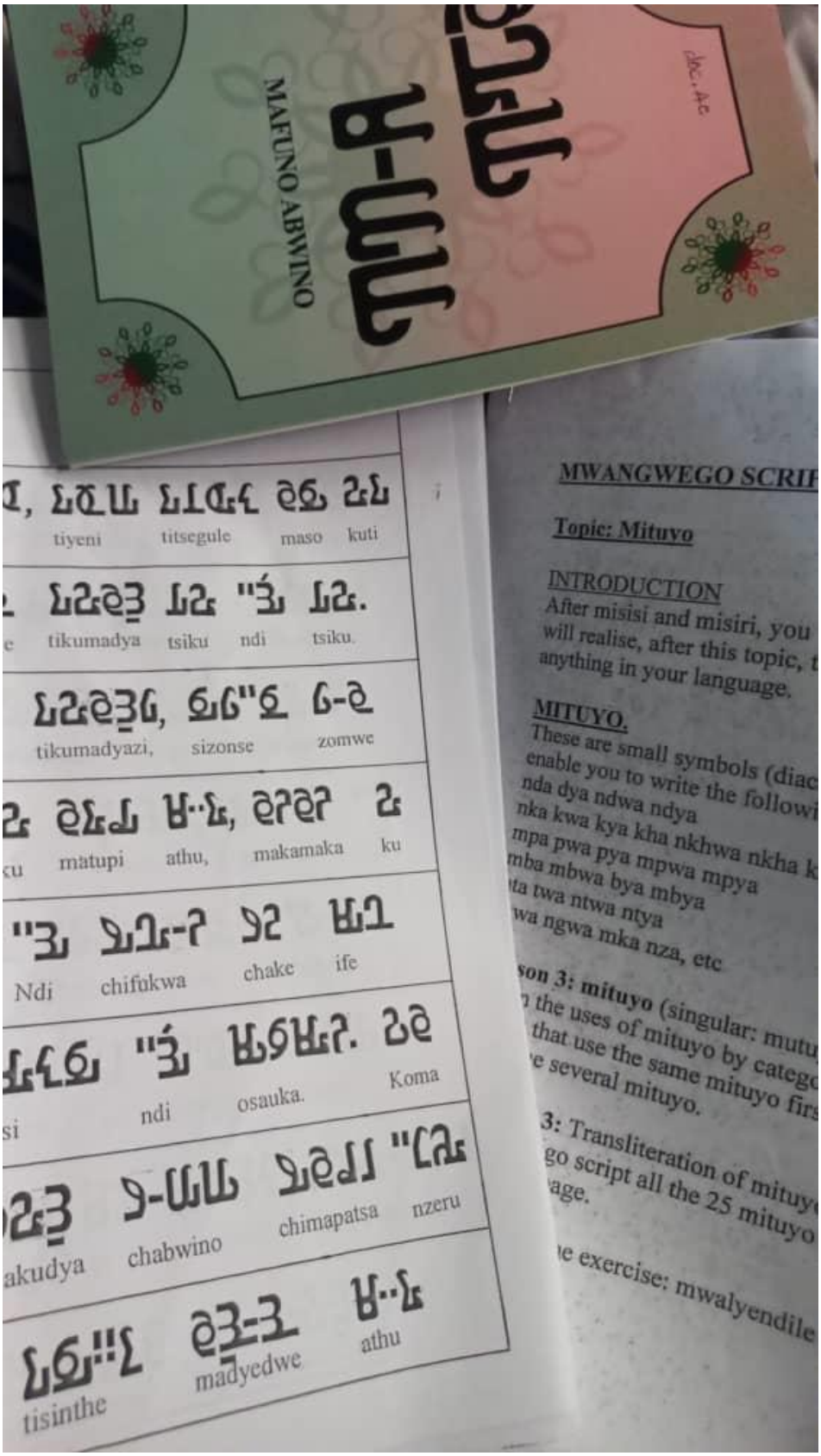


Figure 19. Educational materials used by teachers to teach the Mwangwego script.

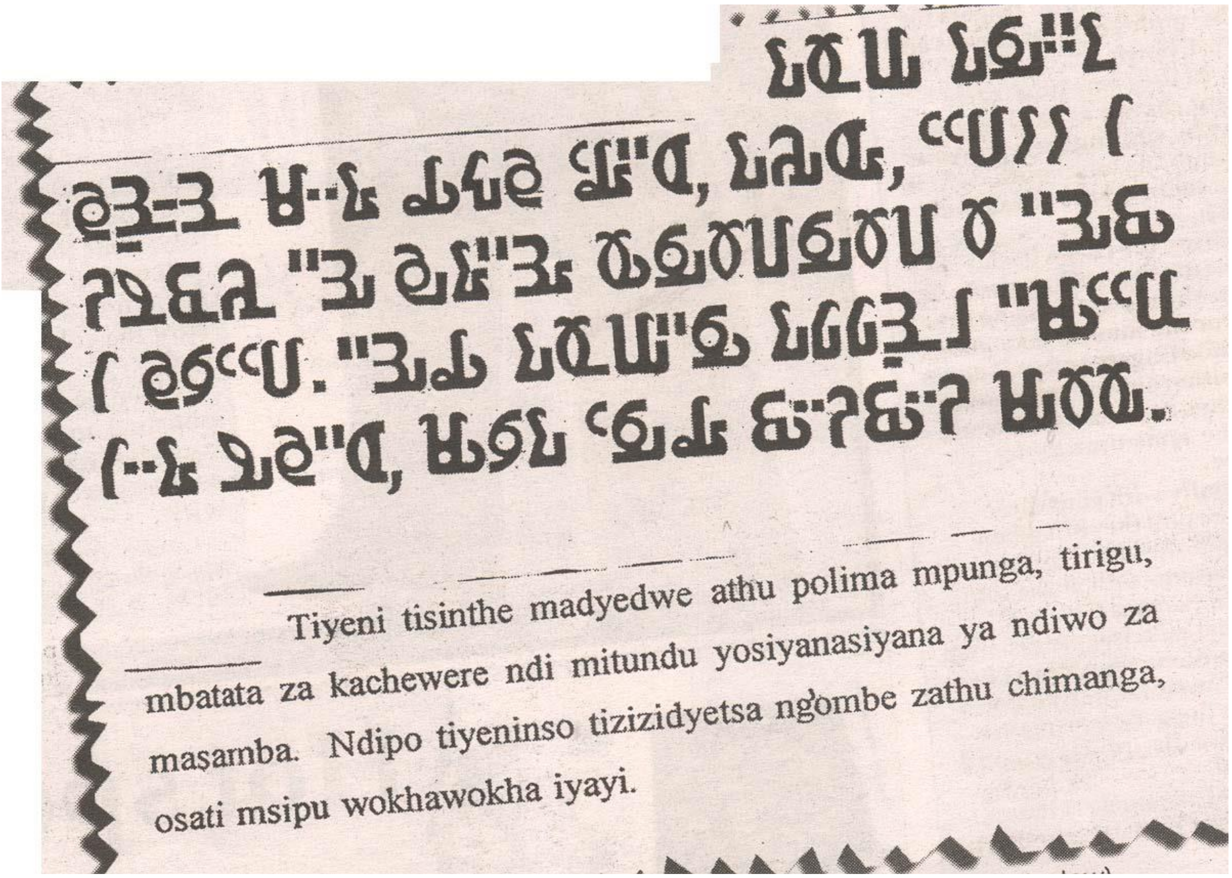


Figure 20. Excerpt from “[illegible]?” with Latin transliteration below hosted on a news site discussing the script. English translation: “*Let’s diversify our diet by growing rice, wheat, Irish potatoes and different types of vegetables. Let us feed our cattle with grains as well, not just fodder.*”

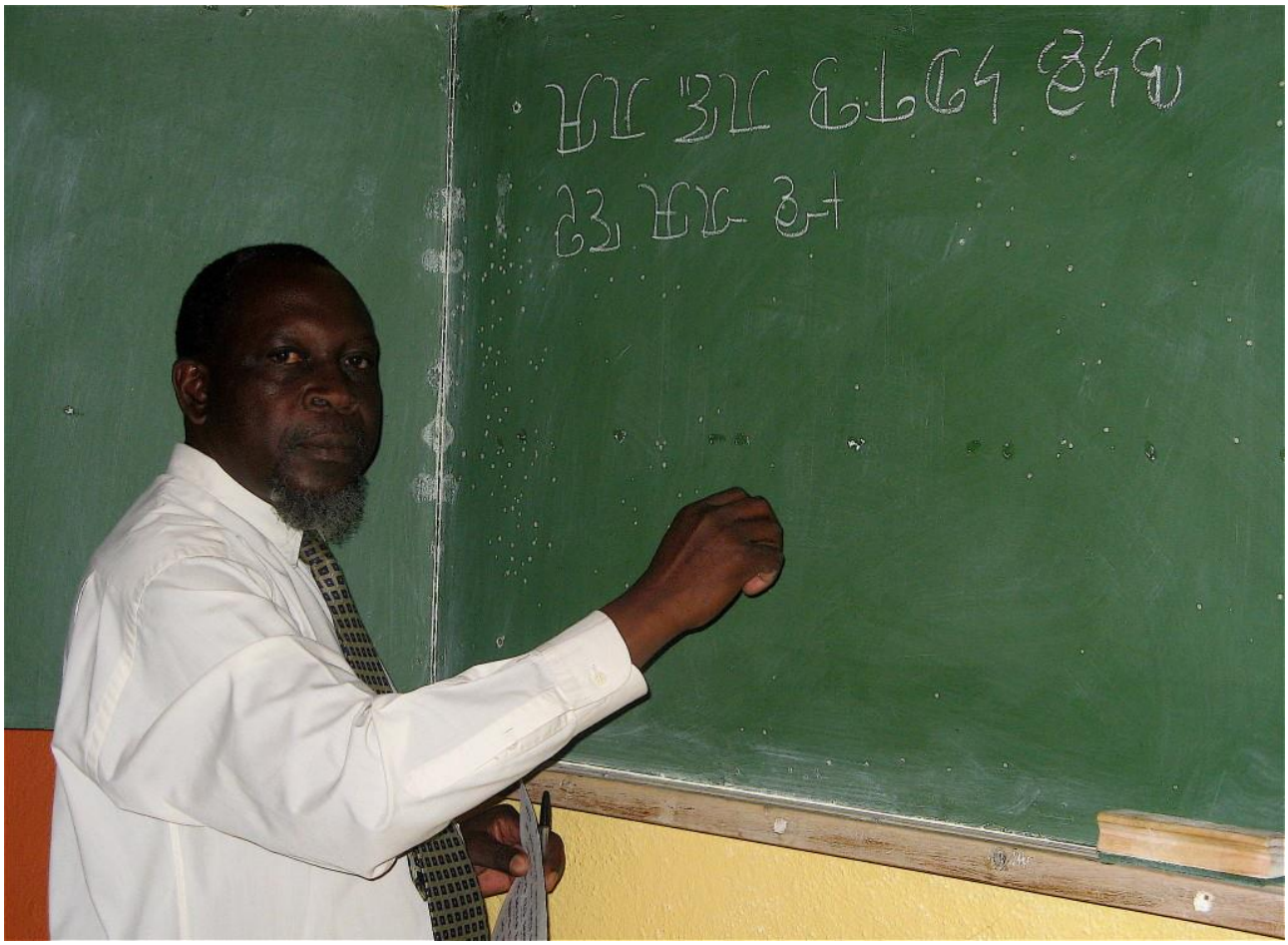
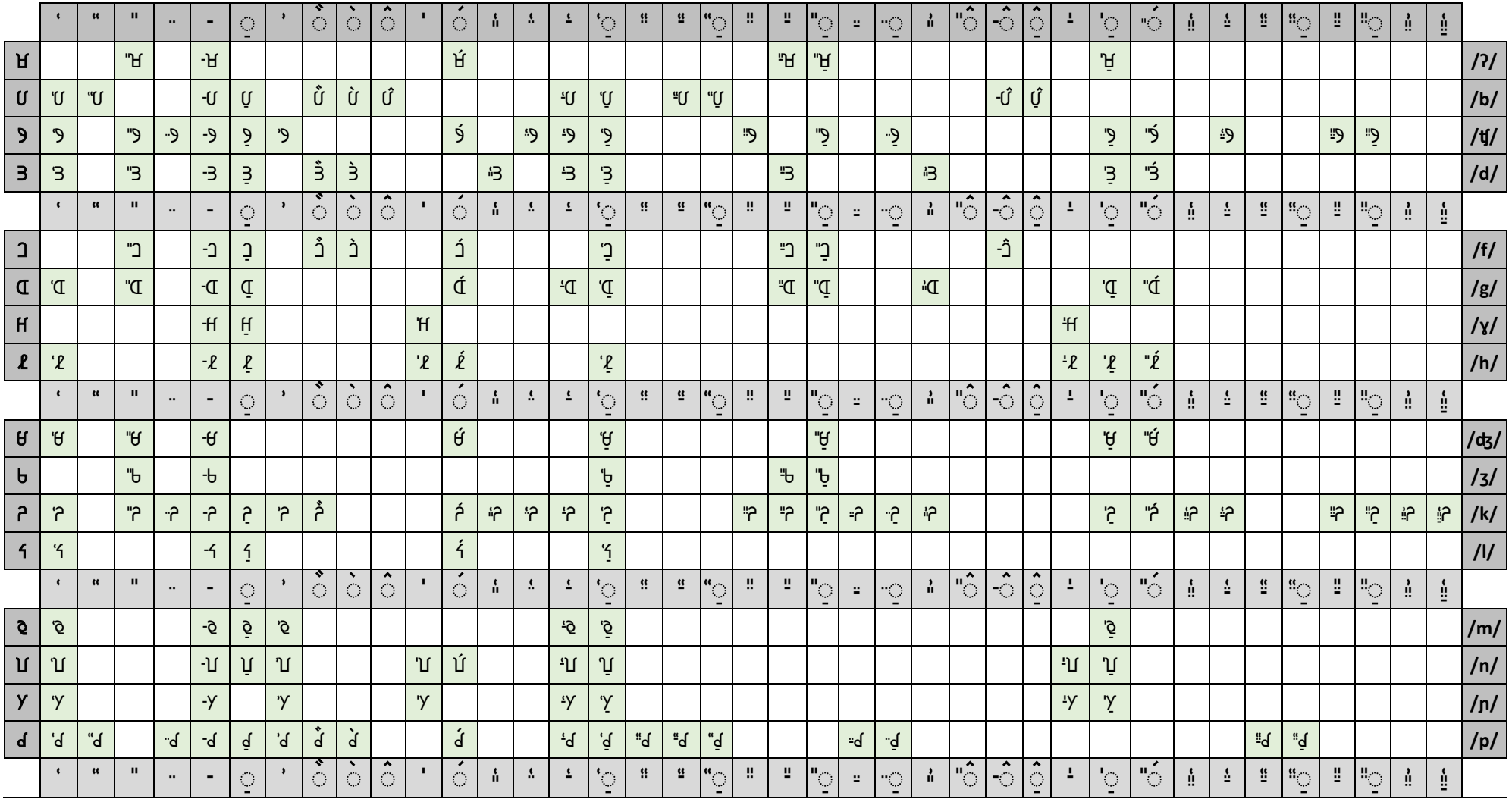


Figure 21. Nolence Mwangwego teaching a classroom the Mwangwego script sometime before late-2010.

X Linguistically Realized Mutuyo and Mituyo Stack Pairings with Misisi

The following table presents the Mutuyo and Mituyo stack pairings with the Misisi letters as found in the primary languages of Malawi. The table is equally valid when any of the Masiri are present. The table is offered as a reference for developers working on lexical projects such as spelling validation and input methods.



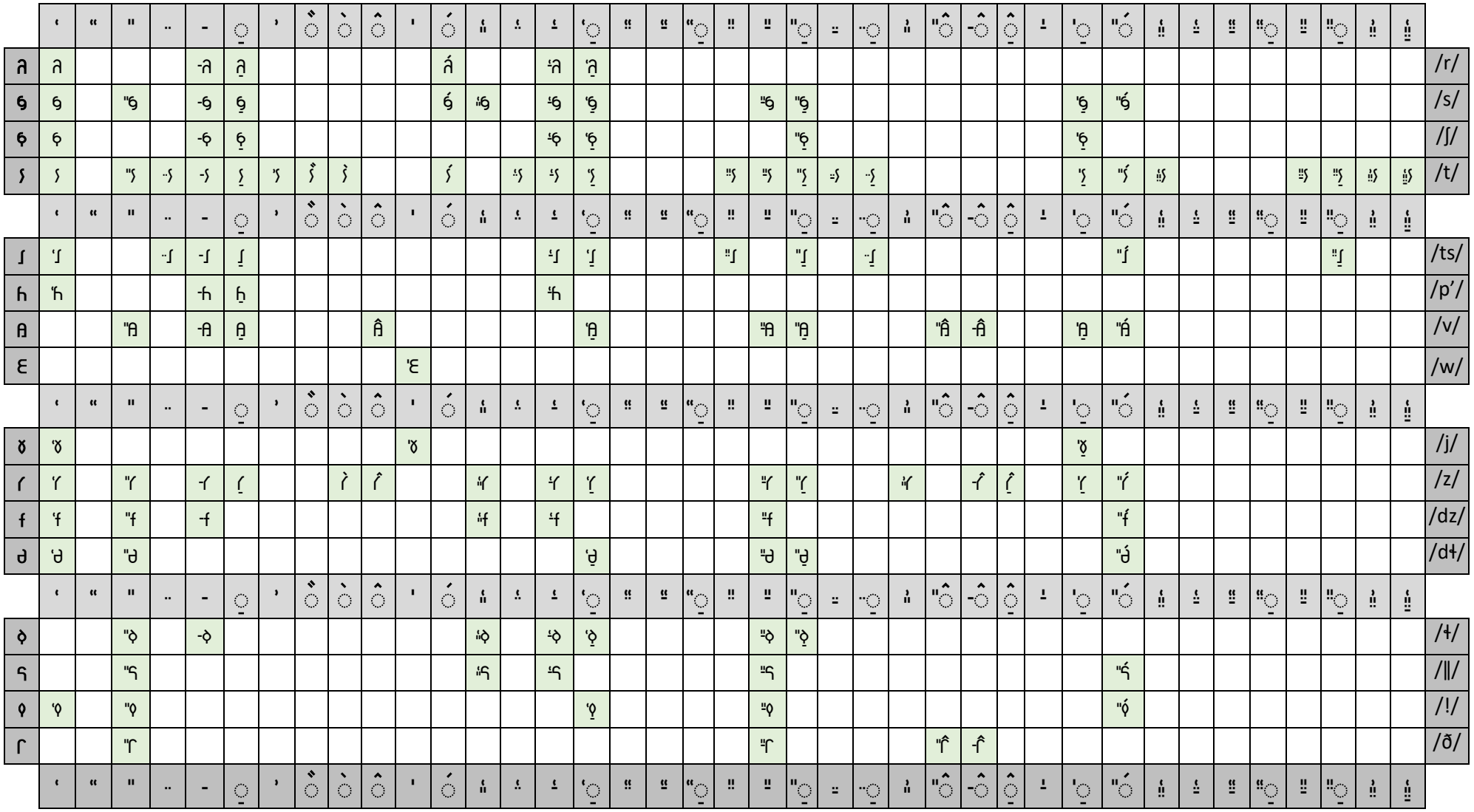


Table 10. Mwangwego letter and modifier occurences found in the phonology of Malawi languages.

XI ISO Proposal Summary Forms

ISO/IEC JTC 1/SC 2/WG 2

PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS

FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from http://std.dkuug.dk/JTC1/SC2/WG2/docs/principles.html **for guidelines and details before filling this form.**

Please ensure you are using the latest Form from http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html **.**

See also http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps*.**

A. Administrative

1. **Title:**	***Proposal for Encoding the Mwangwego Script in the UCS***
2. Requester's name:	*Oreen Yousuf, Daniel Yacob*
3. Requester type (Member body/Liaison/Individual contribution):	*Individual Contribution*
4. Submission date:	*2025-12-05*
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	*Yes*
(or) More information will be provided later:	

B. Technical – General

1. Choose one of the following:

a. This proposal is for a new script (set of characters):	*Yes*
Proposed name of script:	*Mwangwego*
b. The proposal is for addition of character(s) to an existing block:	
Name of the existing block:	

2. Number of characters in proposal: *64*

3. Proposed category (select one from below - see section 2.2 of P&P document):

A-Contemporary	X	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			

4. Is a repertoire including character names provided? *Yes*

[1] Form number: N4502-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

a. If YES, are the names in accordance with the “character naming guidelines”

in Annex L of P&P document?

b. Are the character shapes attached in a legible form suitable for review? *Yes*

5. Fonts related:

a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?

Athinkra

b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):

Athinkra, LLC, yacob@geez.org, https://github.com/athinkra/mwangwego-book

6. References:

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

b. Are published examples of use (such as samples from newspapers, magazines, or other sources)

of proposed characters attached? *Yes*

7. Special encoding issues:

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)? *Yes*

A sorting description is enclosed.

8. Additional Information:

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/reports/tr44/) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?	*Yes*
If YES explain	*Latest proposal L2/12-311*
2. 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*
If YES, with whom?	*Script creator, script users*
If YES, available relevant documents:	*Enclosed in the proposal*
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?	*Yes*
Reference:	*Enclosed in the proposal*
4. The context of use for the proposed characters (type of use; common or rare)	*Rare*
Reference:	*Enclosed in the proposal*
5. Are the proposed characters in current use by the user community?	Yes
If YES, where? Reference:	*Malawi*
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?	*No*
If YES, is a rationale provided?	
If YES, reference:	
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	*Yes*
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?	*No*
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?	*No*
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character?	*Yes*
If YES, is a rationale for its inclusion provided?	*Yes*

If YES, reference:	*Enclosed in the proposal*
11. Does the proposal include use of combining characters and/or use of composite sequences?	*No*
If YES, is a rationale for such use provided?	
If YES, reference:	
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?	
If YES, reference:	
12. Does the proposal contain characters with any special properties such as control function or similar semantics?	*No*
If YES, describe in detail (include attachment if necessary)	
13. Does the proposal contain any Ideographic compatibility characters?	*No*
If YES, are the equivalent corresponding unified ideographic characters identified?	
If YES, reference:	