

Title: Proposal for encoding the Beria Erfe (Zaghawa) script in the SMP of the UCS
Authors: Siddick Adam Issa, Adam Ahmed, Alsadig Sadig Ahmed Osman, Izdin Bormay, Nasradin Mohammed Adam, Mahamat Ahmat Hamat Gaga, Andrew Wolfe, Lorna Evans, Andrij Rovenchak, Oreen Yousuf
Date: 26 January 2024

A preliminary proposal was described in [L2/08-265](#) by Lorna Priest.

Summary: This is a proposal for 50 characters for the Beria Erfe script. The roadmap already has 3 columns allocated for the Beria script (U+16EA0..U+16ECF). Here we request 4 columns to be allocated for the script. The adjacent column U+16ED0..U+16EDF appears to be available.

1. Introduction

The Zaghawa and their Language

The Zaghawa, a transnational ethnic group, are mainly located in Darfur, Sudan and Eastern and Northern Chad. They are a minority in both Sudan and in Chad, but they are highly mobilized for language development in Darfur (Sany & Desai, 2008). The Zaghawa language (ISO 639-3 code: *zag*), also known as Beria, is part of the Eastern Saharan branch of the larger Saharan language family. The Zaghawa language is spoken by almost 300,000 people in Chad and Sudan (Harir, 1987). Approximately 180,000 of them are located in Darfur (Osman, 2006). The Zaghawa language is divided into five main dialects. *Wegi* is the largest group and they are mostly located in Darfur in Sudan (Osman, 2006). *Kube* are located on the border between Sudan and Chad but the majority of them live in Chad. *Tuba* live in Chad in the area north of Kube and some of them mix with Kube. The *Dirong* are mainly settled in Chad (Osman, 2006). Blacksmiths (known as *Mai*) are spread out among all of the Zaghawa regions (Tubiana 2008). Their speech variety is distinct from that of adjacent non-*Mai* Zaghawa communities (Osman, 2006). In addition to the five main dialects, there are minor variations within these speech communities, particularly for the *Brogat* subgroup in the western part of the Tuba dialect area and the *Guruf* subgroup in the western villages of the Dirong dialect area (Anonby & Johnson 2001).

Background to Creating the Writing System for Zaghawa

The Zaghawa language does not have a standardized orthography (Issa, 2021). However, two versions of the writing system have been suggested. The first versions are the Latin-based alphabets from 1912 by McMichael, then Tubiana (Issa, 2021) with a recent version not widely disseminated in the speech community proposed by Sabri Abdelkerim Bichara (Wolfe and Bichara ms, 2017). See Figure 11a for a very informal comparison of the Beria Erfe script, Latin script, and Arabic script. In the 1950s, another form of writing, often called Beria Giray Erfe (to be called “Beria Erfe” in Unicode), was under development by a teacher named Adam Tajir that was inspired by Zaghawa herding traditions. He had the idea to use branding marks on livestock (especially camels) as the basis for characters in the language. To this end, he collected over 300 symbols which have been used by the Zaghawa over the course of history to differentiate among their animals and properties. An improved second version of the script was worked on by Siddick Adam Issa and introduced to the computer via the help of David Faris and SIL International (Issa, 2021). Since 2012 the community activists and leaders in the diaspora, especially in Israel and Egypt but also in Turkey

and Cameroon and elsewhere, started learning the script in groups and individuals via social media and in person. They published learning booklets such as Beria Erfe book by Alsadik Sidik and Izedin Bormay, flyers created by them and Nasradin Mohammed Adam for the Zaghawa community in Egypt and Sudan during the International Mother Language Day in 2022, and recorded videos on social media such as Facebook and YouTube. Alsadik Sidik and Bormay are among the most influential actors writing material, teaching, training teachers and raising awareness toward preserving the Zaghawa language through the use of the Beria Erfe script, but there are other stakeholders and content creators in the script community from the Chadian side of the border, notably Mahamat Ahmat Hamat Gaga (living in Turkey) and Siddick Adam Issa himself (living in N'Djamena, Chad).

The current status of the script is that its repertoire is stable and the glyph designs are stable across the different communities.

2. Structure

The Beria Erfe script is an alphabet written left to right. The total number of letters is 25, of which five represent vowels: **፪** (/i/), **፫** (/ɛ/), **፩** (/a/), **፳** (/ɔ/), and **፭** (/u/).

Consonants

The consonants are **፻** (/b/), **፼** (/d/), **፾** (/f/), **፷** (/g/), **፸** (/h/), **፹** (/j/), **፻** (/k/), **፻** (/l/), **፻** (/m/), **፻** (/n/), **፻** (/p/), **፻** (/t/), **፻** (/v/), **፻** (/w/), and **፻** (/ʃ/).

Some notes on specific characters and their phonemes:

- **፸**: There is some subphonemic articulatory variation in the production of /h/ in Beria, where the /h/ exhibits coarticulation with the advanced tongue root position of the following vowel sound. Some in the speech community, particularly those formally educated in Arabic, hear these two **፸** sounds as distinct sounds, parallel to the pharyngeal/laryngeal distinction in Arabic. The Beria Erfe community of practice has opted not to capture this variation in the script, because they do not feel it is needed.
- **፻**: The character for /t/ was added to the script more recently, with an eye toward the needs of the Chadian dialects. The **፻** sound, which has lateral, rhotic, and retroflex properties, is found primarily in the dialects prevalent on the Chadian side of the border (Kube, Tuba, and Dirong; cf. Wolfe 2001:19ff). The sound is not found in most neighboring languages, and Chad's Zaghawas often have metalinguistic awareness around the uniqueness of this sound and the difficulty that non-native speakers have in reproducing it. In place names, it is sometimes romanized as an R (as in "Iriba" **兕**ብራ [hiqiba], the capital of Wadi Fira) and sometimes as an L (as in "Dar Bilia" **دار**ብሊѧ [biqja]). Because Beria Erfe began with the Sudanese dialect, the script needed this slight adaptation for the Chadian milieus.

Vowels

The vowel system of the Zaghawa language exhibits ATR harmony and contains nine vowels (or ten, in varieties where schwa is used), namely:

+ATR ("heavy"): **፪** (/i/) **፫** (/ɛ/) **፩** (/a/) **፳** (/ɔ/) **፭** (/u/)

–ATR (“light”): ፻ (/ɪ/) ፻ (/ɛ/) ፻ (/a/) ፻ (/ə/) ፻ (/ʊ/)

To represent the +ATR vowels, the combining macron diacritic is used by some in the community of practice, notably the Chadian script users (located in Chad and elsewhere). The combining macron then appears only on the first vowel of the word, because the ATR value of the following vowels will be predictable from the first vowel. If the first vowel is tense, the rest of the vowels will be tense. If the first vowel is lax, the rest of the vowels almost always will be lax. If there is a tense suffix like -፻፻ (/‐nu/) for the third person plural negative marker or -፻፻፻ (/‐egi/) for the first person singular possessive, then the first vowel of this +ATR suffix gets the macron.

The SIL font includes diacritics for tone marking. As of this writing, these marks are not widely used, except for the combining acute accent used to mark (roughly) high tone. This is necessary in the Chadian dialects, where the difference between perfective and imperfective verbal aspect is marked only by tone, so ፻፻፻ ([ségí]) ‘I ate’ (perfective) vs. ፻፻፻ ([ségí]) ‘I eat’ (imperfective). These distinctions are less critical in Sudan’s Wegi dialect, where verbal aspect is also expressed in the vowel segment itself, i.e., ፻ ([ɪ]) vs ፻ ([ɛ]).

Certain minimal pairs in Beria are distinguished only by lexical tone, for instance most frequently ፻ ([ɔ̄]) ‘person’ vs ፻ ([ɔ̄]) ‘milk’. As in the orthographies of many tonal languages, in Beria Erfe currently these pairs are not consistently distinguished. In addition to the combining acute accent, the SIL font also allows for combining grave accent and combining dot above. We propose using the common tone marks from the 0300 block (in particular U+0300, U+0301, and U+0307) in case the community of practice decides to mark lexical tone consistently. If tone marks are added, they will likely be optional, and not affect the collating order of characters.

There are no known words where the combining macron for +ATR and the combining acute accent for high tone (or any other tonal marking) are needed on the same vowel. This is because ATR is nearly always shown on the first vowel of the word while contrastive tone is nearly always marked on the second vowel of the word. The majority of words in Beria are disyllabic.

While the use of diacritics varies slightly among communities of practice, nobody objects to their inclusion in the Unicode Standard, and the overall repertoire of characters including diacritics to be used is stable.

Casing

Much as in standard English and French orthographic norms, the intent of the community of practice is that capitalization is to be used for the first letter of proper names of people, places, and things, as well as to mark the beginning of sentences. Actual usage is somewhat inconsistent, as it sometimes is even in Latin script use in English.

3. Proposed Script Name: Beria Erfe

The use of “Beria Erfe” in the proposed Unicode Standard script and character names bears some explanation, however. Zagħawa (الزغافة) is an exonym given by adjacent Arab communities to the Beri community, and now widely used in other languages, whenever the Beri community is mentioned in the international press in conjunction with Chad (e.g., the family of the late Idriss Deby Itno) or with Sudan (e.g., ongoing Darfur genocide coverage).

Beri (በርሃዊት) is the name used by the community itself to refer to its culture and ethnicity. The name of the language is derived from the addition of *-a* (አ), meaning ‘mouth,’ to the word *Beri*: the ‘mouth of the *Beri*.’ In some varieties of Beria (particularly in Sudan), the word is pronounced *Bera* ([be:ra]). There was some initial discussion (and an earlier draft of this proposal) in which we contemplated using “Zaghawa” to abstract away from these differences. However, all major stakeholders, even those who speak the Sudanese dialect, now agree that “Beria” is either preferable or acceptable, with the proviso that the word “Zaghawa” should appear in the main Core Specification and/or at the top of the names list, since this name is how the community is known to most of the world.

The term *erfe* ‘branding mark’ also varies among dialects, with *erbe* appearing in both the Wogi and Tuba dialects. Even community leaders who speak dialects that use *erbe* have agreed that *erfe* is acceptable for the script, since that version has become commonplace in script teaching.

The word *giray*, meaning ‘writing,’ distinguishes livestock-branding *erfe* from written *erfe*. An earlier version of this proposal included the entire name “Beria Giray Erfe” in the Unicode character names, but in the interest of brevity, and in response to comments from the Script Ad Hoc Committee, community stakeholders have agreed that it is acceptable to omit this word from the script name in the Unicode Standard. The name “Beria Giray Erfe” will remain in use within the community, but in the Unicode context, the word “writing” is redundant.

4. Proposed Character Repertoire

The proposed character repertoire consists of 50 upper and lowercase letters.

As mentioned above, the character for /l/ (BERIA ERFE LETTER ERIGO TAMURA) was added to the script more recently, with an eye toward the needs of the Chadian dialects. [This character is only represented in Figure 12 to demonstrate why it is needed because until recently no font has supported that character.](#) However, it is essential for the needs in the Chadian dialects.

Character Data

The characters of Beria Erfe have traditional names that are being used throughout the script use community. They are listed below as official character names in the Unicode Character Properties. There is a set of simplified character names that are sometimes used for ease of pedagogy, but the major stakeholders agree that the Unicode Standard should include the longer names, and the simplified names will remain in informal use.

[Two codepoints are left between the end of the capital letters and the beginning of the lower case letters to allow for the possibility of encoding new case pairs.](#)

The character repertoire, and the design of the characters are stable.

Unicode Character Properties

Character Properties: `UnicodeData.txt`

```
16EA0;BERIA ERFE CAPITAL LETTER ARKAB;Lu;0;L;;;;;N;;;;;16EBB;  
16EA1;BERIA ERFE CAPITAL LETTER BASIGNA;Lu;0;L;;;;;N;;;;;16EBC;  
16EA2;BERIA ERFE CAPITAL LETTER DARBAI;Lu;0;L;;;;;N;;;;;16EBD;  
16EA3;BERIA ERFE CAPITAL LETTER EH;Lu;0;L;;;;;N;;;;;16EBE;
```

16EA4;BERIA ERFE CAPITAL LETTER FITKO;**Lu**;0;L;;;;;**N**;;;;16EBF;
16EA5;BERIA ERFE CAPITAL LETTER GOWAY;**Lu**;0;L;;;;;**N**;;;;16EC0;
16EA6;BERIA ERFE CAPITAL LETTER HIRDEABO;**Lu**;0;L;;;;;**N**;;;;16EC1;
16EA7;BERIA ERFE CAPITAL LETTER I;**Lu**;0;L;;;;;**N**;;;;16EC2;
16EA8;BERIA ERFE CAPITAL LETTER DJAI;**Lu**;0;L;;;;;**N**;;;;16EC3;
16EA9;BERIA ERFE CAPITAL LETTER KOTO;**Lu**;0;L;;;;;**N**;;;;16EC4;
16EAA;BERIA ERFE CAPITAL LETTER LAKKO;**Lu**;0;L;;;;;**N**;;;;16EC5;
16EAB;BERIA ERFE CAPITAL LETTER MERI;**Lu**;0;L;;;;;**N**;;;;16EC6;
16EAC;BERIA ERFE CAPITAL LETTER NINI;**Lu**;0;L;;;;;**N**;;;;16EC7;
16EAD;BERIA ERFE CAPITAL LETTER GNA;**Lu**;0;L;;;;;**N**;;;;16EC8;
16EAE;BERIA ERFE CAPITAL LETTER NGAY;**Lu**;0;L;;;;;**N**;;;;16EC9;
16EAF;BERIA ERFE CAPITAL LETTER OI;**Lu**;0;L;;;;;**N**;;;;16ECA;
16EB0;BERIA ERFE CAPITAL LETTER PI;**Lu**;0;L;;;;;**N**;;;;16ECB;
16EB1;BERIA ERFE CAPITAL LETTER ERIGO;**Lu**;0;L;;;;;**N**;;;;16ECC;
16EB2;BERIA ERFE CAPITAL LETTER ERIGO TAMURA;**Lu**;0;L;;;;;**N**;;;;16ECD;
16EB3;BERIA ERFE CAPITAL LETTER SERI;**Lu**;0;L;;;;;**N**;;;;16ECE;
16EB4;BERIA ERFE CAPITAL LETTER SHEP;**Lu**;0;L;;;;;**N**;;;;16ECF;
16EB5;BERIA ERFE CAPITAL LETTER TATASOU;**Lu**;0;L;;;;;**N**;;;;16ED0;
16EB6;BERIA ERFE CAPITAL LETTER UI;**Lu**;0;L;;;;;**N**;;;;16ED1;
16EB7;BERIA ERFE CAPITAL LETTER WASSE;**Lu**;0;L;;;;;**N**;;;;16ED2;
16EB8;BERIA ERFE CAPITAL LETTER AY;**Lu**;0;L;;;;;**N**;;;;16ED3;
16EBB;BERIA ERFE SMALL LETTER ARKAB;**Ll**;0;L;;;;;**N**;;;;16EA0;;16EA0
16EBC;BERIA ERFE SMALL LETTER BASIGNA;**Ll**;0;L;;;;;**N**;;;;16EA1;;16EA1
16EBD;BERIA ERFE SMALL LETTER DARBAI;**Ll**;0;L;;;;;**N**;;;;16EA2;;16EA2
16EBE;BERIA ERFE SMALL LETTER EH;**Ll**;0;L;;;;;**N**;;;;16EA3;;16EA3
16EBF;BERIA ERFE SMALL LETTER FITKO;**Ll**;0;L;;;;;**N**;;;;16EA4;;16EA4
16EC0;BERIA ERFE SMALL LETTER GOWAY;**Ll**;0;L;;;;;**N**;;;;16EA5;;16EA5
16EC1;BERIA ERFE SMALL LETTER HIRDEABO;**Ll**;0;L;;;;;**N**;;;;16EA6;;16EA6
16EC2;BERIA ERFE SMALL LETTER I;**Ll**;0;L;;;;;**N**;;;;16EA7;;16EA7
16EC3;BERIA ERFE SMALL LETTER DJAI;**Ll**;0;L;;;;;**N**;;;;16EA8;;16EA8
16EC4;BERIA ERFE SMALL LETTER KOTO;**Ll**;0;L;;;;;**N**;;;;16EA9;;16EA9
16EC5;BERIA ERFE SMALL LETTER LAKKO;**Ll**;0;L;;;;;**N**;;;;16EAA;;16EAA
16EC6;BERIA ERFE SMALL LETTER MERI;**Ll**;0;L;;;;;**N**;;;;16EAB;;16EAB
16EC7;BERIA ERFE SMALL LETTER NINI;**Ll**;0;L;;;;;**N**;;;;16EAC;;16EAC
16EC8;BERIA ERFE SMALL LETTER GNA;**Ll**;0;L;;;;;**N**;;;;16EAD;;16EAD
16EC9;BERIA ERFE SMALL LETTER NGAY;**Ll**;0;L;;;;;**N**;;;;16EAE;;16EAE
16ECA;BERIA ERFE SMALL LETTER OI;**Ll**;0;L;;;;;**N**;;;;16EAF;;16EAF
16ECB;BERIA ERFE SMALL LETTER PI;**Ll**;0;L;;;;;**N**;;;;16EB0;;16EB0
16ECC;BERIA ERFE SMALL LETTER ERIGO;**Ll**;0;L;;;;;**N**;;;;16EB1;;16EB1
16ECD;BERIA ERFE SMALL LETTER ERIGO TAMURA;**Ll**;0;L;;;;;**N**;;;;16EB2;;16EB2
16ECE;BERIA ERFE SMALL LETTER SERI;**Ll**;0;L;;;;;**N**;;;;16EB3;;16EB3
16ECF;BERIA ERFE SMALL LETTER SHEP;**Ll**;0;L;;;;;**N**;;;;16EB4;;16EB4
16ED0;BERIA ERFE SMALL LETTER TATASOU;**Ll**;0;L;;;;;**N**;;;;16EB5;;16EB5
16ED1;BERIA ERFE SMALL LETTER UI;**Ll**;0;L;;;;;**N**;;;;16EB6;;16EB6
16ED2;BERIA ERFE SMALL LETTER WASSE;**Ll**;0;L;;;;;**N**;;;;16EB7;;16EB7
16ED3;BERIA ERFE SMALL LETTER AY;**Ll**;0;L;;;;;**N**;;;;16EB8;;16EB8

Line breaking and spaces

Line breaking generally occurs at spaces (U+0020) or hyphens (U+002D). Long words may optionally be broken at syllable boundaries with Latin hyphen-minus to facilitate better typesetting

in paragraphs of texts. Line breaks cannot occur within numbers. Line breaks always come after punctuation, not before.

Linebreak Properties: LineBreak.txt

```
16EA0..16EB8 ; AL # Lu [25] BERIA ERFE CAPITAL LETTER ARKAB..BERIA ERFE CAPITAL LETTER AY  
16EBB..16ED3 ; AL # Ll [25] BERIA ERFE SMALL LETTER ARKAB..BERIA ERFE SMALL LETTER AY
```

Combining Marks

Combining marks from the Combining Diacritical Marks block are used, by some communities, on vowels (*arkab*, *eh*, *i*, *oi*, and *ui*). These include: U+0300 COMBINING GRAVE ACCENT, U+0301 COMBINING ACUTE ACCENT, U+0304 COMBINING MACRON, and U+0307 COMBINING DOT ABOVE.

These should be added to `ScriptExtensions.txt` for the Beria Erfe script.

As mentioned earlier, there are no known words where the combining macron for +ATR and the combining acute accent for high tone (or any other tonal marking) are needed on the same vowel. See Figure 11b for a table of the various vowels and their combining marks.

Collation

Collating order is as shown in the code chart below. It is based on the traditional arrangement of letters. “Heavy” (+ATR) vowels written with the combining macron are sorted after their “light” (-ATR) counterparts. Tone marks do not affect the collating order of characters.

```
ꝝ <<< ꝛ << Ꝝ <<< Ꝝ < ꝗ <<< ꝗ < Ꝛ <<< Ꝛ < Ꝕ <<< Ꝕ < Ꝗ <<<  
ꝙ < ꝑ <<< ꝑ < Ꝕ <<< Ꝕ < ꝕ <<< ꝕ < Ꝗ <<< Ꝗ < ꝗ <<< ꝗ < Ꝙ <<< Ꝙ < ꝙ <<<  
ꝝ <<< ꝝ < ꝝ <<<
```

Punctuation and digits

There are no script-specific punctuation or digits. Word boundaries are indicated using spaces. Latin script punctuation marks are also used: full stop (U+002E), comma (U+002C), exclamation mark (U+0021), question mark (U+003F), single (U+2018, U+2019) and double (U+201C, U+201D) curly quotation marks, colon (U+003A), semicolon (U+003B), left (U+0028) and right (U+0029) parentheses, em-dash (U+2014), and forward slash (U+002F).

The hyphen-minus is used by some in the script community to separate enclitics from their phonological hosts, in particular (but not exclusively) in the case of the light verb -ꝑ- (g̊in̊i) ‘he did’ with its lexical verb root.

European digits (U+0030..U+0039) are used. Originally there was some brief consideration for adding script-specific digits. At this time, no script-specific digits have been used or proposed.

	16EA	16EB	16EC	16ED
0	፩ 16EA0	፪ 16EB0	፫ 16EC0	፬ 16ED0
1	፪ 16EA1	፪ 16EB1	፮ 16EC1	፯ 16ED1
2	፭ 16EA2	፪ 16EB2	፯ 16EC2	፪ 16ED2
3	፯ 16EA3	፪ 16EB3	፪ 16EC3	፯ 16ED3
4	፭ 16EA4	፪ 16EB4	፯ 16EC4	
5	፫ 16EA5	፪ 16EB5	፯ 16EC5	
6	፮ 16EA6	፯ 16EB6	፪ 16EC6	
7	፭ 16EA7	፪ 16EB7	፪ 16EC7	
8	፪ 16EA8	፪ 16EB8	፪ 16EC8	
9	፯ 16EA9		፪ 16EC9	
A			፪ 16ECA	
B	፪ 16EAB	፩ 16EBB	፪ 16ECB	
C	፪ 16EAC	፪ 16EBC	፪ 16ECC	
D	፪ 16EAD	፪ 16EBD	፪ 16ECD	
E	፪ 16EAE	፪ 16EBE	፪ 16ECE	
F	፪ 16EAF	፪ 16EBF	፪ 16ECF	

The Beria Erfe script is used for the language of the Zagawa people of Sudan and Chad. The language is known as Beria, Bera, or Zagawa.

Various letters

16EA0	፩	BERIA ERFE CAPITAL LETTER ARKAB
16EA1	፪	BERIA ERFE CAPITAL LETTER BASIGNA
16EA2	፭	BERIA ERFE CAPITAL LETTER DARBAI
16EA3	፯	BERIA ERFE CAPITAL LETTER EH
16EA4	፯	BERIA ERFE CAPITAL LETTER FITKO
16EA5	፩	BERIA ERFE CAPITAL LETTER GOWAY
16EA6	፮	BERIA ERFE CAPITAL LETTER HIRDEABO
16EA7	፪	BERIA ERFE CAPITAL LETTER I
16EA8	፪	BERIA ERFE CAPITAL LETTER DJAI
16EA9	፯	BERIA ERFE CAPITAL LETTER KODO
16EAA	፯	BERIA ERFE CAPITAL LETTER LAKKO
16EAB	፪	BERIA ERFE CAPITAL LETTER MERI
16EAC	፪	BERIA ERFE CAPITAL LETTER NINI
16EAD	፩	BERIA ERFE CAPITAL LETTER GNA
16EAE	፳	BERIA ERFE CAPITAL LETTER NGAY
16EAF	፳	BERIA ERFE CAPITAL LETTER OI
16EB0	፪	BERIA ERFE CAPITAL LETTER PI
16EB1	፪	BERIA ERFE CAPITAL LETTER ERIGO
16EB2	፪	BERIA ERFE CAPITAL LETTER ERIGO TAMURA
16EB3	፪	BERIA ERFE CAPITAL LETTER SERI
16EB4	፪	BERIA ERFE CAPITAL LETTER SHEP
16EB5	፯	BERIA ERFE CAPITAL LETTER TATASOUYE
16EB6	፩	BERIA ERFE CAPITAL LETTER UI
16EB7	፪	BERIA ERFE CAPITAL LETTER WASSE
16EB8	፯	BERIA ERFE CAPITAL LETTER AY
16EB9	፳	<reserved>
16EBA	፳	<reserved>
16EBB	፩	BERIA ERFE SMALL LETTER ARKAB
16EBC	፳	BERIA ERFE SMALL LETTER BASIGNA
16EBD	፭	BERIA ERFE SMALL LETTER DARBAI
16EBE	፯	BERIA ERFE SMALL LETTER EH
16EBF	፯	BERIA ERFE SMALL LETTER FITKO
16EC0	፩	BERIA ERFE SMALL LETTER GOWAY
16EC1	፮	BERIA ERFE SMALL LETTER HIRDEABO
16EC2	፳	BERIA ERFE SMALL LETTER I
16EC3	፪	BERIA ERFE SMALL LETTER DJAI
16EC4	፯	BERIA ERFE SMALL LETTER KODO
16EC5	፯	BERIA ERFE SMALL LETTER LAKKO
16EC6	፪	BERIA ERFE SMALL LETTER MERI
16EC7	፪	BERIA ERFE SMALL LETTER NINI
16EC8	፩	BERIA ERFE SMALL LETTER GNA
16EC9	፩	BERIA ERFE SMALL LETTER NGAY
16ECA	፳	BERIA ERFE SMALL LETTER OI
16ECB	፳	BERIA ERFE SMALL LETTER PI
16ECC	፳	BERIA ERFE SMALL LETTER ERIGO
16ECD	፳	BERIA ERFE SMALL LETTER ERIGO TAMURA
16ECE	፪	BERIA ERFE SMALL LETTER SERI
16ECF	፪	BERIA ERFE SMALL LETTER SHEP
16ED0	፯	BERIA ERFE SMALL LETTER TATASOUYE
16ED1	፩	BERIA ERFE SMALL LETTER UI
16ED2	፩	BERIA ERFE SMALL LETTER WASSE
16ED3	፯	BERIA ERFE SMALL LETTER AY

5. Design

Lowercase letters sit on a baseline (with punctuation such as full stop, comma, and question mark). The uppercase letters descend below the baseline (blue characters are uppercase). Example:

₩>X SHALL WE GO ON? WE WILL,
WE WILL GO ON. WE WILL GO ON?

When Latin script is mixed with Beria Erfe characters, the baseline should be at the bottom of the lowercase Beria Erfe characters. The uppercase characters descend below the baseline. Punctuation should follow standard positioning for alignment with the baseline of the lowercase letters.

I X O X E T X A L E S H E ? Laegae lakourei?
(Good morning?)

I A E X X E T X E G C E T X E ? Louwai lae deilae?
(What are you doing now?)

Underlining is sometimes used. However, at this point there is no standard for where the position of the underline should be. The position of the blue line in the sample below seems to be acceptable by the community as the proper position for an underline. Placing it through the descender was found to be undesirable.

123 Q A L E 8 T X E . G ouylae.

6. Acknowledgement

This project was partly supported by a grant from the United States National Endowment for the Humanities (PR-268710-20), which funds the Universal Scripts Project (part of the Script Encoding Initiative at the University of California, Berkeley). Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment for the Humanities.

7. Bibliography

- Anonby, Erik John & Eric Johnson. 2001. A sociolinguistic survey of the Zaghawa (Beria) of Chad and Sudan. N'Djamena, Chad: Association SIL Tchad.
- Harir, S. A. 1987. Liquid phonemes (rs and ls) in Beri (Zaghawa). Bergen, Norway: Universitetet i Bergen Institutt for fonetikk og lingvistikk.
- Issa, Siddick Adam. 2007. Vers un orthographe beria qui est fidèle à ses origines nubiens, meroïtiques et africains. Article submitted for the Nilo-Saharan Colloquium. Paris, France.

- Issa, S. A. 2021. New vision for Zaghawa language orthography. <https://blogs.ed.ac.uk/nilotic/wp-content/uploads/sites/3585/2021/05/Issa.pdf>
- Mathes, T. K. 2015. Acoustic phonetic aspects of a 7/9 vowel inventory: A study of the [ATR] harmony language Zaghawa. In *Selected Proceedings of the 44th Annual Conference on African Linguistics*, ed. Ruth Kramer et al., 180-187. Somerville, MA: Cascadilla Proceedings Project. <http://www.lingref.com/cpp/acal/44/paper3136.pdf>
- Osman, S. N. 2006 (April). Proverbs and Idiomatic Phrases in Zaghawa Language. In *7th International Sudan Studies conference*.
- Rovenchak, A. and J. Glavy. 2011. “Beria”. In *African Writing Systems of the Modern Age: The Sub-Saharan Region*. New Haven, Buena Park, New Rochelle, London, Lviv, Abidjan: Athinkra, 79–82. ISBN 978-0-9818294-1-8
- Sany, J., and S. Desai. 2008. Transnational ethnic groups and conflict: The Zaghawa in Chad and Sudan. *Conflict Trends* 2008(2), 25–29.
- Tubiana, Marie-José. 2008. Hommes sans voix: Forgerons du nord-est du Tchad et de l'est du Niger. Paris: Éditions L'Harmattan.
- Wolfe, Andrew M. 2001. Toward a generative phonology and morphology of the dialects of Beria. Cambridge, MA: Harvard College B.A. honors thesis.
- Wolfe, Andrew M. & Sabri Abdelkerim Bichara. 2017. Beria orthography proposal (Proposition d'orthographe beria). Unpublished draft. N'Djamena, Chad: SIL Tchad, ms.

Online resources:

- <https://www.facebook.com/watch/?v=395571291909586&ref=sharing>
- <https://www.facebook.com/watch/?v=1393880144385625&ref=sharing>
- <https://www.facebook.com/watch/?v=553221529079575&ref=sharing>
- <https://www.youtube.com/watch?v=NEIcjw5pukg&t=62s>
- <https://www.youtube.com/watch?v=-G7bRxBlfxc&t=1s>
- <https://www.youtube.com/watch?v=TrC94sjpAm8>
- <https://www.youtube.com/watch?v=UWUZv3VSVYk>
- <https://youtu.be/DE72CISRwEs?si=2AhYDzsSUg7oswuS>
- <https://youtu.be/m5tNPK7c7vk?si=qFDtK6pF8-uIWQ0L>
- <https://youtu.be/5Qq6CHp543I?si=naGqP0crH5ucGEzD>

8. Additional Images/Figures



Figure 1. Bilingual Beria/Arabic description of a Beria Erfe keyboard at International Mother Language Day in N'Djamena, Chad - February 2022

Note the use of standard left and right double quotation marks in the second line of Beria Erfe script, and the use of the standard colon at the bottom left of the Beria Erfe script section. Title case as in a headline is used on this flyer; upper-case letters have been marked here with red arrows, and a red baseline has been added to bring the upper-case descenders into sharper relief.

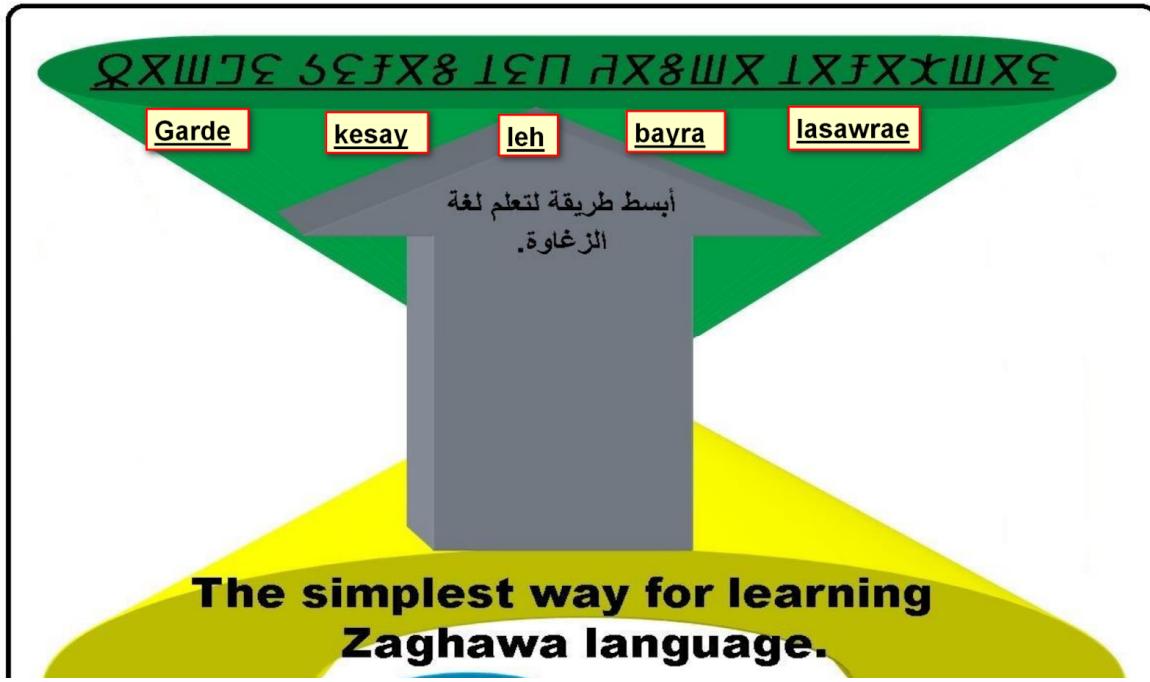


Figure 2a. Booklet, "The simplest way for learning Zaghawa language" developed by Nasradin Mohammed Adam to teach the Zaghawa community in Egypt.

Note the use of full stop after each bullet point. Note use of sentence case. Only the initial letter is capitalized. (See Latin transliteration added below.) Note use of underlining that touches the descender of the upper-case letter *goway*.

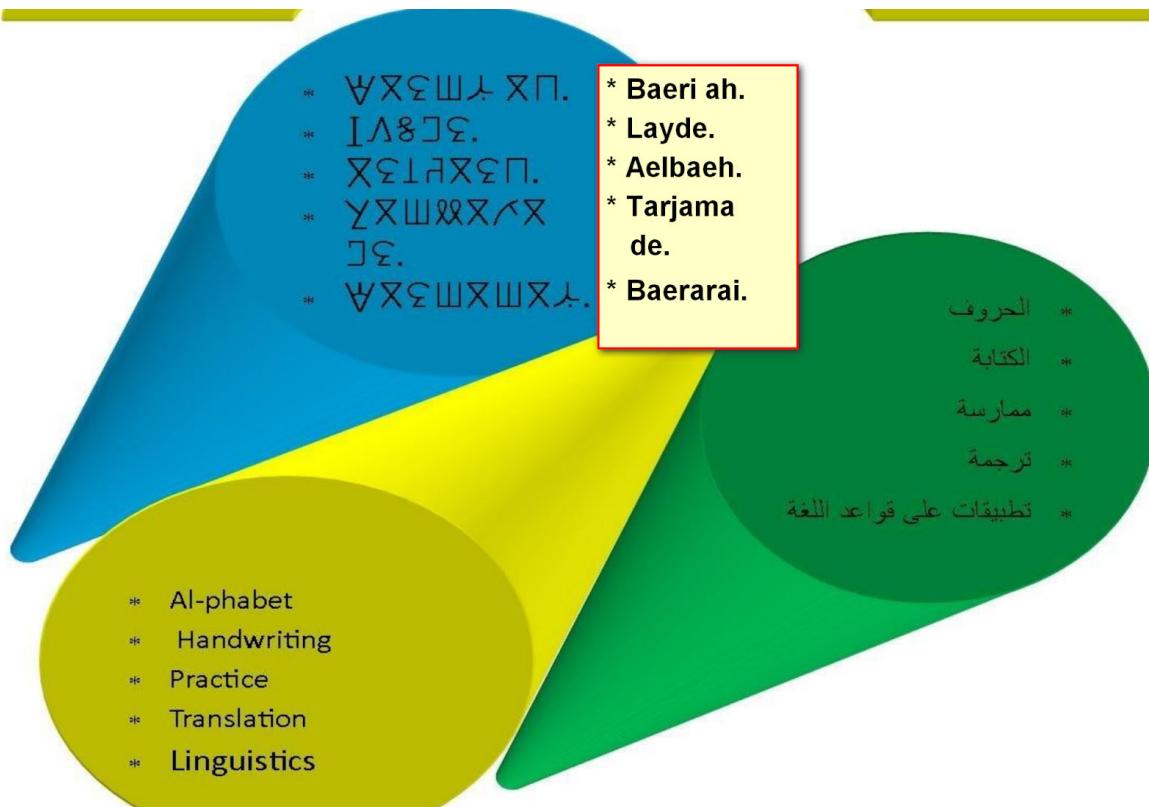


Figure 2b. Bottom of same page. Bullet points have sentence case, as shown in added Latin transliteration. Added thin red line shows baseline of lower-case letters.

<u>كدر لاکوري؟</u> Kadr lakourei?	<u>كدر لاکورو؟</u> Kadr lakourou?
<u>اودو باء لقى لو</u> Oudo baar laegae. Good night تصبح على خير	<u>اودو باء لقى لو</u> Oudo baar laegao ! Good night ! تصبحون على خير

Nasradin Mohammed's Latin transcription matches his Beria Giray Erfe exactly throughout, except that he writes "baar" in Latin where his Beria Giray Erfe version has only "bar." He consistently uses sentence casing.

The double capital in the word Oudo appears to be a typo, as he capitalizes only the first letter of this sequence elsewhere.

Figure 2c. Same booklet, p.10. Note sentence case throughout. The word “Oudo” appears once as “OUdo” and another time as “Oudo” in this same image. This appears to be a typo.

6) 盍ڙڻش تلڙڙخ؟	Habr louay?	How are things?
كيف الامور؟	حبار لواي	
Oufouy 7). 盍ڙڙچڙ.	oufouy.	Not bad.
ليس سينا	أوفوي	
اهڙ	aey	ابي
نعم	yes	نعم
بر	no	نوه
No		لا
• تيرالو لادي؟		teirralo ladi ?
what's your name ?		تبيرولو لادي؟

Nasradin Mohammed Adam

Page

نصرالدين محمد ادم

13

Figure 2d. Same booklet, p.13. Illustrates that the initial double capital from Fig. 2c is not standard.

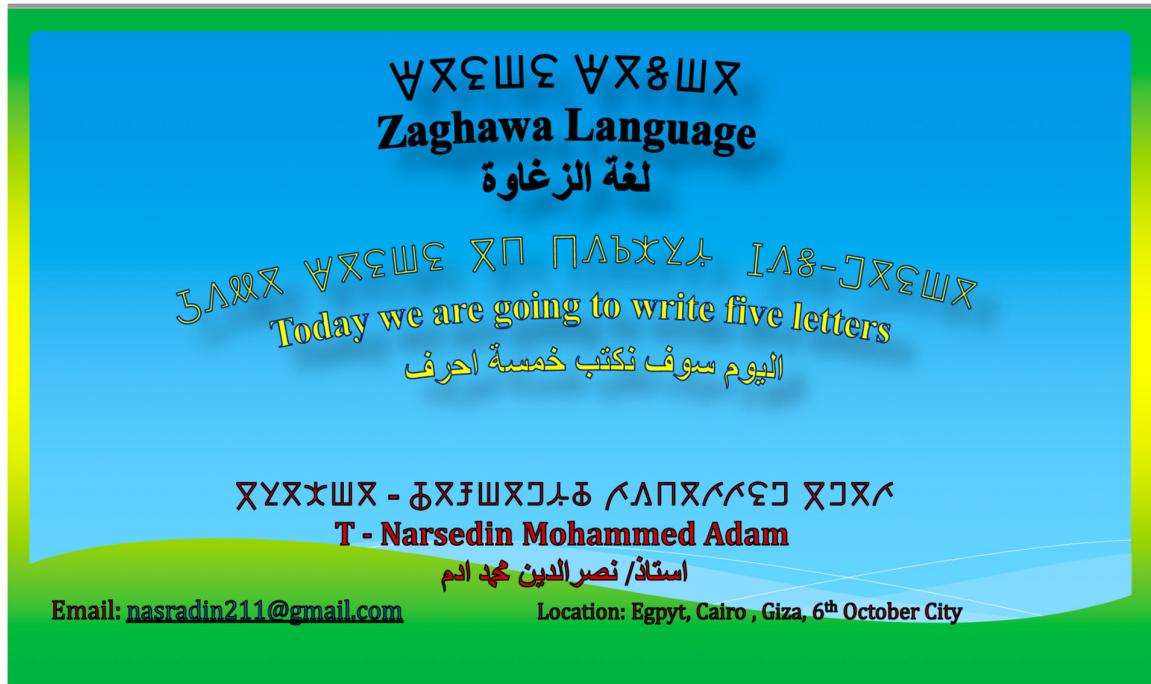


Figure 3a. Booklet developed by Nasradin Mohammed Adam to teach the Zaghawa community in Egypt.

Note use of hyphens.



Figure 3b. p.5 from the same booklet. Note the sentence case as shown in added Latin transliteration. (Red line added here to accentuate baseline visually.)

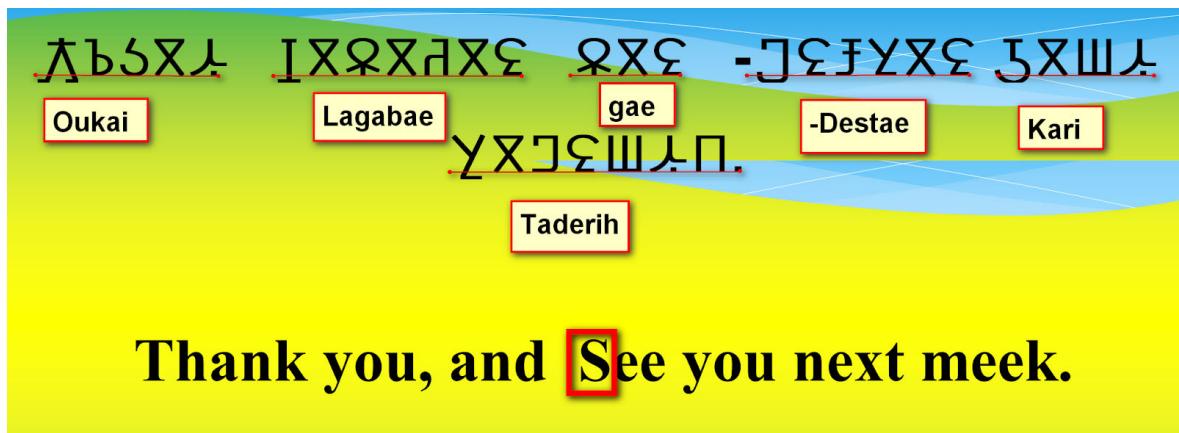


Figure 3c, p.4 from the same booklet. Use of sentence case is not entirely consistent within existing materials; this page from the same document uses something closer to title case, as illustrated in the Latin transliteration with underlining added.



Figure 4a, material developed by Alsadik Sidik and Izedin Bormay to teach the Zaghawa community in Israel and other places mainly in diaspora, p.5. Latin transliteration added in

text boxes, and red underlining added to accentuate lower-case baseline. Note consistent use of sentence case, comma and full stop. Capitals are used to start new sentences after periods, but not after commas.

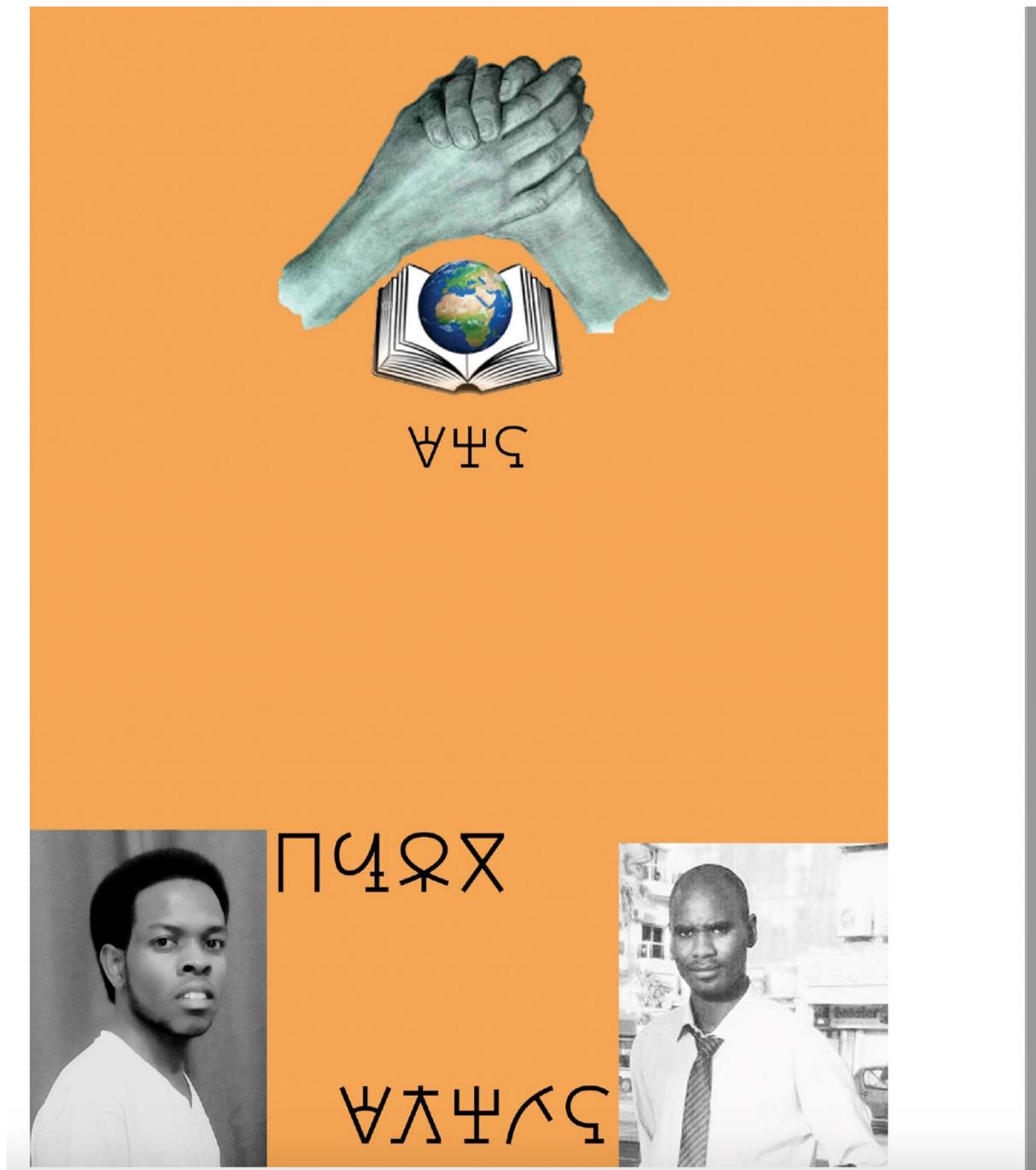


Figure 4b, cover of same booklet. Uses all capitals for both logo and authors' names, though this is not an intentional and agreed convention.

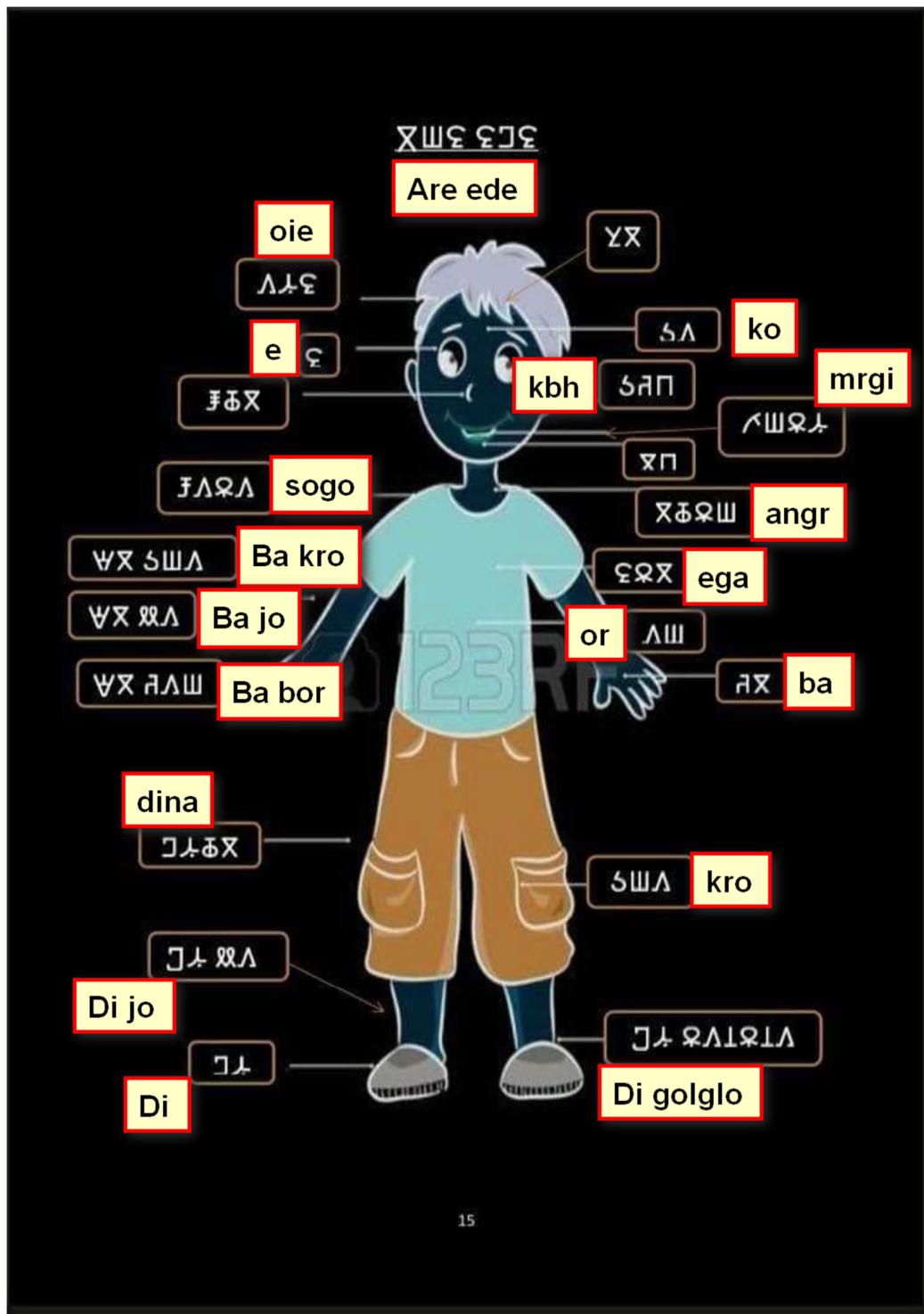


Figure 4c, same booklet, p. 15. Case use is somewhat inconsistent here; see added Latin transliterations.



Figure 5, Flyer for learning the Zaghawa language. While font size varies here, it is upper-case letters that are used throughout.



Figure 6, Flyer for learning the Zaghawa language. Here, lower case is used throughout.



Figure 7, Flyer for learning the Zaghawa language. Here, lower case is used throughout.

In addition, there are three separate active WhatsApp groups dedicated to learning the Beria script, with 112, 170, and 43 members.

LES NOMBRES EN BERIA					
nombre	En lettre	nombre	En lettre	nombre	En lettre
0	፩፭፻፲፲	11	፩፭፻፱፲፲፳፭፻፱፲	30	፩፭፻፲
1	፩፭፻፱፲	12	፩፭፻፱፲፲፭፻፱፲፲፲	40	፩፭፻፱፲፭
2	፩፭፻፲	13	፩፭፻፱፲፲፭፻፱፲፲፲	50	፩፭፻፱፲፭
3	፩፭፻	14	፩፭፻፱፲፲፭፻፱፲፲፭	60	፩፭፻፱፲፭
4	፩፭፻፱	15	፩፭፻፱፲፲፭፻፱፲፲፭	70	፩፭፻፱፲፭
5	፩፭፻፲	16	፩፭፻፱፲፲፭፻፱፲፲፭	80	፩፭፻፱፲፭
6	፩፭፻፲፭	17	፩፭፻፱፲፲፭፻፱፲፲፭	90	፩፭፻፱፲፭
7	፩፭፻፲፭	18	፩፭፻፱፲፲፭፻፱፲፲፭	100	፩፭፻፱፲፭
8	፩፭፻፲፭	19	፩፭፻፱፲፲፭፻፱፲፲፭	200	፩፭፻፱፲፭፻፱፲፭
9	፩፭፻፱፲፭	20	፩፭፻፱፲፭	300	፩፭፻፱፲፭
10	፩፭፻፱፲	21	፩፭፻፱፲፲፭፻፱፲፲	1000	፩፭፻፱፲፭፻፱፲፭

Figure 8: Numbers in Beria chart produced by Mahamat Hamat Ahmat Gaga.

Note the use of European digits. Note also the use of the +ATR macron with the number 9, የ፩፭፻፱፲ ([disti]). It appears on the first vowel but not the second, as described above. This word appears in the Beria translations of 9, 19, and 90. The +ATR [i] vowel also appears in the word for 1,000, የ፩፭፻፱፲፭ ([tim]). All diacritics are indicated with red arrow.

፳፻፻፻፻፻፻	፳፻፻፻፻፻፻	français
፳፻/፳፻	ay/ey	je
፳፻	na	tu
፳፻፻	ber	il/elle
፳፻፻፻/፳፻፻	tougouy/tow	nous
፳፻፻	now	vous
፳፻፻፻፻	berr	ils/elles

Figure 9, Pronoun chart produced by Mahamat Hamat Ahmat Gaga

Note the use of forward slash to suggest alternatives (in this case, variability for Kube versus Tuba dialect isoglosses). This also illustrates the use of the doubled final letter to indicate the tonal modifications associated with pluralization (compare ፳፻ “il/elle” versus ፳፻፻፻ “ils/elles”).

The following examples come from the “Milk Maid” (፳፻ ፳፻፻፻) book produced by Siddick Adam Issa and David Faris with help from SIL Tchad.

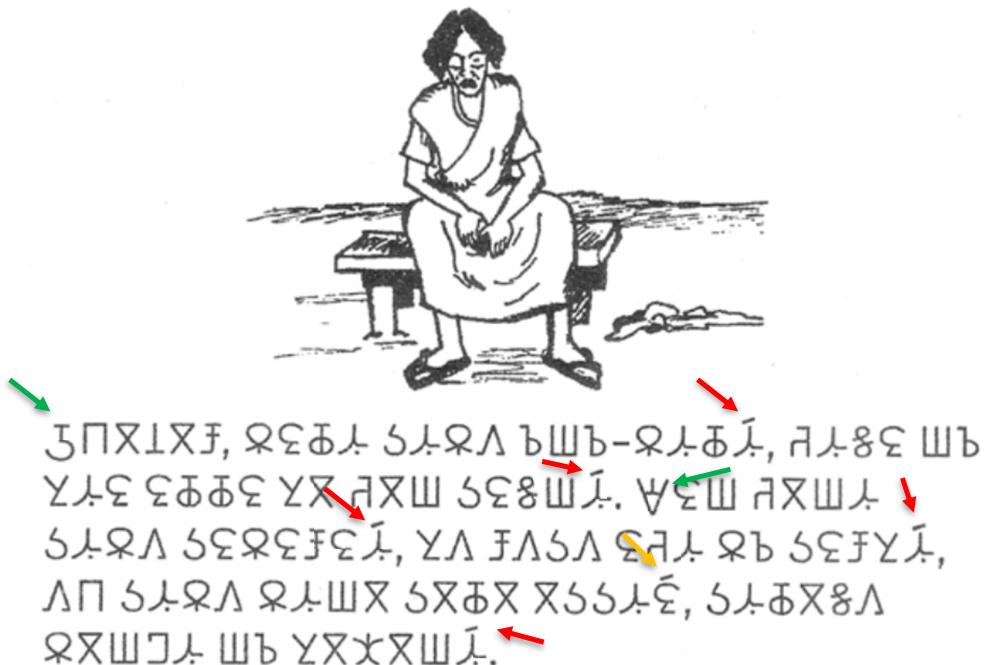


Figure 10a, *Oh Barta*, by Siddick Adam Issa, p.11

Note the use of the hyphen, the comma, and the acute accent signaling perfective aspect in the penultimate word of the top line, նշե-օշուկի (/nsh-e-ošu-k'i/). The acute diacritic appears five more times on this page, usually but not always on the high vowel ִ ([i]). It also appears on the lax mid front vowel ֵ ([ɛ]). These acute accent diacritic examples are indicated with red and orange arrows. Upper case letters are indicated in green arrows.

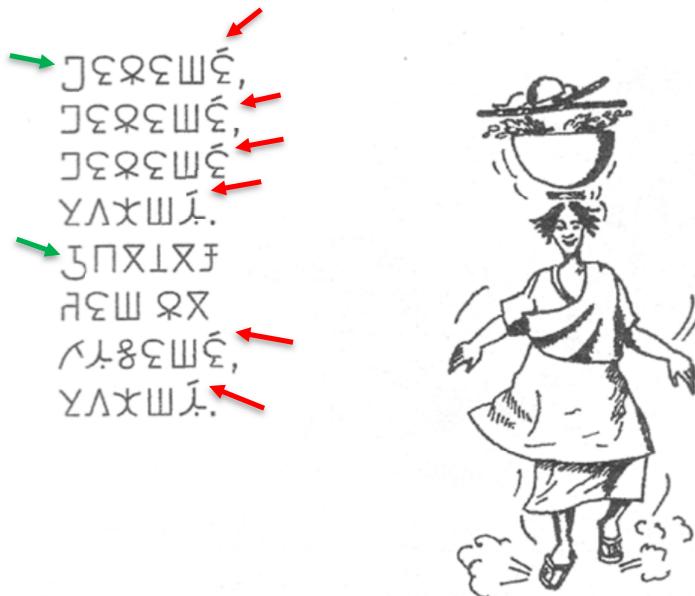
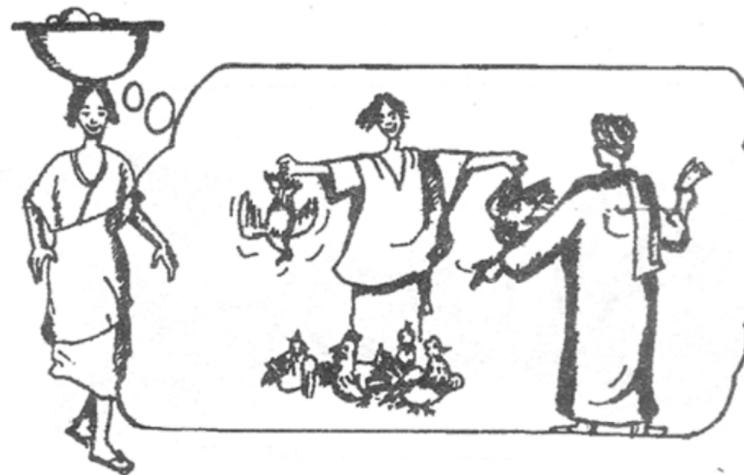


Figure 10b, *Oh Barta*, by Siddick Adam Issa, p.9

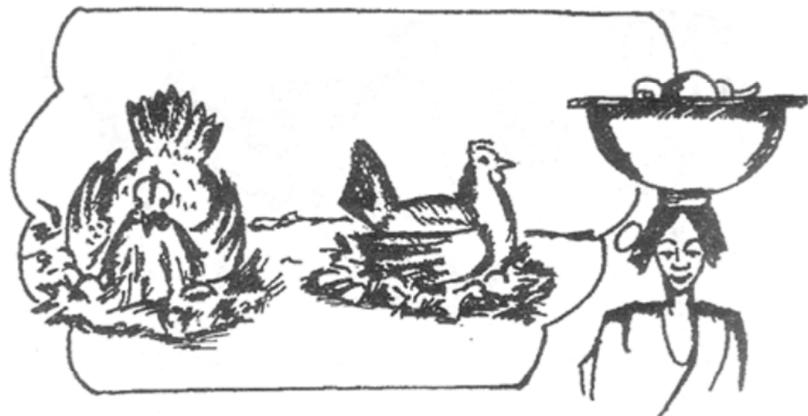
On this page, the acute diacritic (red arrows) appears on nearly every final. Note the sentence capitalization (green arrows) and the full stop at the end of the page.



၁၆၂၁-၁၂၁၃၁, ၂၁၂ ၁၂၁၂၁၃၁၃၁. ၁၁၁၁၁၃-
၁၂၁၃၁၃၁၃၁, ၂၁၁၁၁၃၁၃၁.

Figure 10c, *Oh Barta*, by Siddick Adam Issa, p.6

This page twice exhibits the use of hyphen to separate the light verb enclitic as described above. (See blue arrows.) In one example, it is further suffixed with the conditional marker ံ (-ŋ/).)



የ አለበት ስያዊን ሽበ አድራሻ-ቁሳቁ, ይኝቀብ
ስናዬት, “የ ይኝቀብ አድራሻቁ, እሱተቁት.”

Figure 10d, *Oh Barta*, by Siddick Adam Issa, p.11

On this page, the protagonist brackets off her thoughts with left and right double quotation marks.



እ ዘመን በቃልም ስለ ንጉሴ መሸጻች እና ዘመን ተረጋግጧል ነው,
“እና ከ ዘመን ስለ ንጉሴ ይፈጸማል,
ለሁ የዚህ ቀን በዚህ ዘመን ተረጋግጧል”

Figure 10e, *Oh Barta*, by Siddick Adam Issa, p.11

This page contains examples of the combining grave accents, on, for instance, the final word of the page. See dark red arrow.

The following chart gives a phonetic representation of the Beria Erfe script as well as listing an unofficial Latin script orthography and informal usage of Arabic script. However, these should not be considered authoritative as the script has not been standardized in either Latin script or Arabic script.

Beria Erfe	Phonetic	Latin script	Arabic script	Beria Erfe	Phonetic	Latin script	Arabic script
خخ	/a/	Aa	أ / ا	ڦڻ	/ŋ/	Gg	قا
ڦڻ	/b/	Bb	ٻ	ڦڻ	/ŋ/	Jj	جا
ڏڻ	/d/	Dd	ڏ	ڦڻ	/ɔ/	Oo	عو
ڙڻ	/ɛ/	Ee	اى / ٻ	ڦڻ	/p/	Pp	ې
ڦڻ	/f/	Ff	ڦ	ڦڻ	/r/	Rr	ار
ڦڻ	/g/	Gg	ڦ	ڦڻ	/l/	ڦڻ	ر
ڦڻ	/h/	Hh	ڦ	ڦڻ	/s/	Ss	س
ڦڻ	/i/	Ii	اي	ڦڻ	/ʃ/	ش	ش
ڦڻ	/j/	Jj	جا	ڦڻ	/t/	Tt	ت
ڦڻ	/k/	Kk	ڪ	ڦڻ	/u/	Uu	او
ڦڻ	/l/	Ll	ڦ	ڦڻ	/w/	Ww	و
ڦڻ	/m/	Mm	ڦ	ڦڻ	/j/	Yy	ي
ڦڻ	/n/	Nn	ڦ				

Figure 11a, *Chart of Beria Erfe characters for encoding*

Beria Erfe	Phonetic	Latin script	Beria Erfe	Phonetic	Latin script
ꝩꝩ	/a/	Aa	ꝩꝩ	/ɔ/	Oo
ꝩꝩ	/à/	Àà	ꝩꝩ	/ɔ̄/	Òò
ꝩꝩ	/á/	Áá	ꝩꝩ	/ó/	Óó
ꝩꝩ	/ə/	Ââ	ꝩꝩ	/o/	Ôô
ꝩꝩ	/ε/	Ee	ꝩꝩ	/ʊ/	Uu
ꝩꝩ	/è/	Èè	ꝩꝩ	/ð/	Ùù
ꝩꝩ	/é/	Éé	ꝩꝩ	/ú/	Úú
ꝩꝩ	/e/	Êê	ꝩꝩ	/u/	Ûû
ꝩꝩ	/ɪ/	Ii			
ꝩꝩ	/ì/	Ìì			
ꝩꝩ	/í/	Íí			
ꝩꝩ	/i/	Îî			

Figure 11b, *Chart of Beria Erfe vowel characters*

The following page of text written by Mahamat Ahmat Hamat Gaga includes the distinction in the Chadian dialects between the erigo tamura (ሣ) underlined in red and the erigo (ሣ) underlined in green. It also shows the macron diacritic on all five vowels, in the words በቃዣ ‘cow’, በ ‘when’, ቃ ‘leg’, ለሱዣ ‘having entered’, and ቁርቃ ‘correct, excellent’, among others.

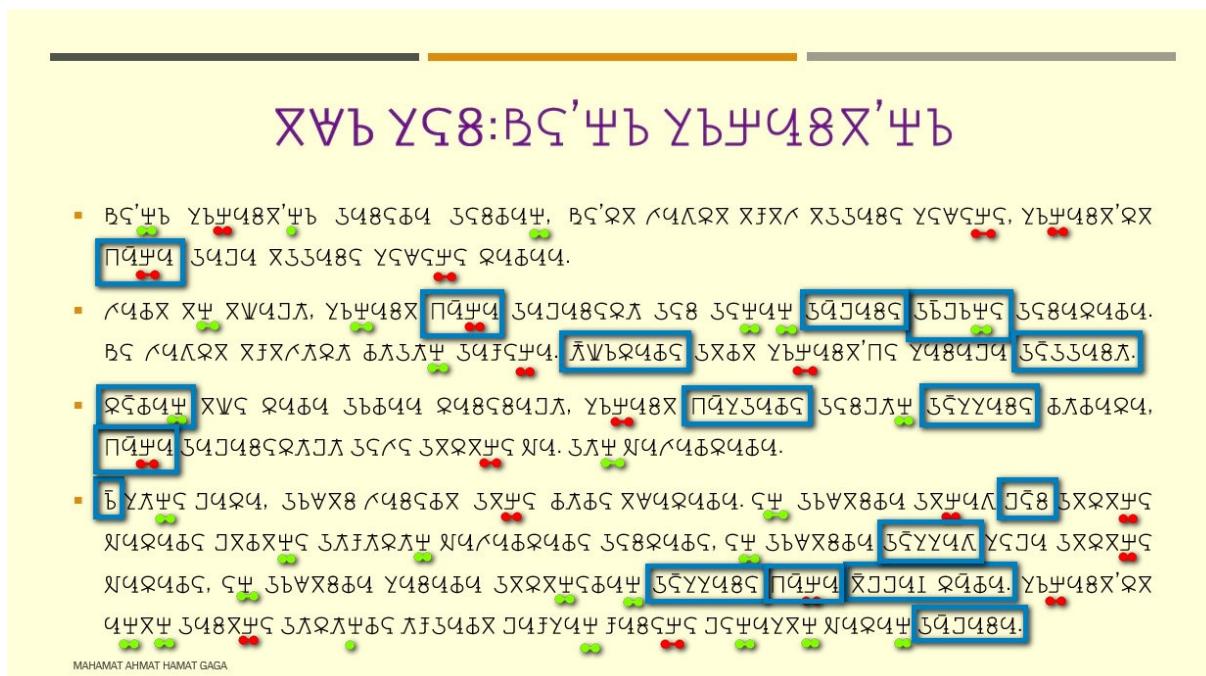


Figure 12, by Mahamat Ahmat Hamat Gaga

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹

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:	<i>Proposal for encoding the Beria Erfe (Zaghawa) script in the SMP of the UCS</i>	
2. Requester's name:	<i>Siddick Adam Issa, Adam Ahmed, Alsadig Sadig Ahmed Osman, Izdin Bormay, Nasradin Mohammed Adam, Mahamat Ahmat Hamat Gaga, Andrew Wolfe, Lorna Evans, Andrij Rovenchak, Oreen Yousuf</i>	
3. Requester type (Member body/Liaison/Individual contribution):	<i>Individual contribution</i>	
4. Submission date:	<i>25 January 2024 (revised)</i>	
5. Requester's reference (if applicable):	<i>N/A</i>	
6. Choose one of the following:		
This is a complete proposal:	<i>Yes</i>	
(or) More information will be provided later:	<i>No</i>	

B. Technical – General

1. Choose one of the following:		
a. This proposal is for a new script (set of characters):	<i>Beria Erfe</i>	
Proposed name of script:	<i>Beria Erfe</i>	
b. The proposal is for addition of character(s) to an existing block:	<i>No</i>	
Name of the existing block:	<i>50</i>	
2. Number of characters in proposal:	<i>50</i>	
3. Proposed category (select one from below - see section 2.2 of P&P document):		
A-Contemporary <input checked="" type="checkbox"/>	B.1-Specialized (small collection) <input type="checkbox"/>	B.2-Specialized (large collection) <input type="checkbox"/>
C-Major extinct <input type="checkbox"/>	D-Attested extinct <input type="checkbox"/>	E-Minor extinct <input type="checkbox"/>
F-Archaic Hieroglyphic or Ideographic <input type="checkbox"/>	G-Obscure or questionable usage symbols <input type="checkbox"/>	
4. Is a repertoire including character names provided?	<i>Yes</i>	
a. If YES, are the names in accordance with the “character naming guidelines” in Annex L of P&P document?	<i>Yes</i>	
b. Are the character shapes attached in a legible form suitable for review?	<i>Yes</i>	
5. Fonts related:		
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	<i>Lorna Evans, SIL International</i>	
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	<i>Open Font License</i>	
6. References:		
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<i>Yes</i>	
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<i>Yes</i>	
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)?	<i>Yes</i>	
	<i>See proposal</i>	
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.		

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)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?	<i>No</i>
If YES explain	<i>A preliminary introduction to the script was submitted L2/08-265</i>
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.)?	<i>Yes</i>
If YES, with whom?	<i>Many of the authors are from the user community.</i>
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?	<i>Yes</i>
Reference:	<i>See proposal.</i>
4. The context of use for the proposed characters (type of use; common or rare)	<i>Yes</i>
Reference:	<i>See proposal.</i>
5. Are the proposed characters in current use by the user community?	<i>Yes</i>
If YES, where? Reference:	<i>See proposal.</i>
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?	<i>No</i>
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)?	<i>Yes</i>
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?	<i>No</i>
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?	<i>No</i>
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?	<i>No</i>
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
11. Does the proposal include use of combining characters and/or use of composite sequences?	<i>Yes</i>
If YES, is a rationale for such use provided?	<i>Yes</i>
If YES, reference:	<i>See proposal</i>
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?	<i>No</i>
If YES, reference:	
12. Does the proposal contain characters with any special properties such as control function or similar semantics?	<i>No</i>
If YES, describe in detail (include attachment if necessary)	
13. Does the proposal contain any Ideographic compatibility characters?	<i>No</i>
If YES, are the equivalent corresponding unified ideographic characters identified?	
If YES, reference:	