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

Doc Type: **Working Group Document**
Title: **Proposal for encoding the Kayah Li script in the BMP of the UCS**
Source: **Michael Everson**
Status: **Individual Contribution**
Replaces: **N3024**
Action: **For consideration by JTC1/SC2/WG2 and UTC**
Date: **2006-03-09**

The Kayah Li script is used to write Eastern and Western Kayah Li languages, which are members of Karenic branch of the Sino-Tibetan language family. They are also known as Red Karen and Karenni. Eastern Kayah Li is spoken by about 360,000 people, and Western Kayah Li by about 210,000 people, mostly in the Kayah and Karen states of Myanmar, but also by people living in Thailand.

Origin
Kayah Li script was devised by Htae Bu Phae in March 1962, in part in response to the appearance of Latin-based orthographies which had appeared after 1950. It is taught in schools in refugee camps in Thailand. Kayah Li's relation to Brahmic scripts can be seen in its ordering and the shapes of some of its letters, although the shapes of most of them were developed independently. At least nine of its characters bear a relation to characters in the Myanmar script: compare Kayah Li ᑕ NGA and Myanmar ᄃ NGA, ᑫ NYA and ᑳ NYA, ᔕ NA and ᔓ NA, ᗯ MA and ᖇ MA, ᑫ LA and ᗄ LA, ᑯ THA and ᑫ TTHA, ᑕ A and ᘁ A, ᐛ E and ᐝ AI, and ᱹ PLOPHU and ᱺ DOT BELOW. A Myanmar-script orthography also exists for Kayah Li; it uses a number of extensions also used for S'gaw Karen, along with three vowel diacritics unique for Kayah Li: ᱹ [ɤ], ᱺ [u], and ᱻ [e]. It may be the case that the last two of these are related to Kayah Li ᱻ [u] and ᱼ [e], but if so, it is not clear which script may have borrowed from which, as both orthographies are relatively young. In any case, Kayah Li diacritical marks are considered script-specific in this proposal. (A proposal to encode Myanmar-script extensions is being prepared as N3044 to encode the three other diacritics and many other characters besides.)

Structure
Unlike the Myanmar script, the Kayah Li script is a true alphabet; consonants have no inherent /a/ vowel sound. Four of the vowels are written with spacing letters (ᑕ *a*, ᗯ *o'*, ᘁ *i*, ᘓ *ô*), and five are written with diacritics applied above the letter ᑕ A, which serves as a vowel-carrier (ᒉ *u'* [ɯ], ᒇ *e* [ɛ], ᒈ *u* [u], ᒊ *ê* [e], ᒋ *o* [ɔ]). Above the letter ᗯ OE, the same diacritics are used to represent sounds found in loanwords (ᔉ *o'u* [ɤɯ], ᔊ *u'e* [ɯɛ], ᔋ *o'u* [ɤu], ᔌ *o'ê* [ɤe], ᔍ *o'o* [ɤɔ]). Users consider the base letters and the diacritics to be quite distinct. When spelling words, for example, the diacritical marks are named separately from the vowel carrier: ᖴᖰ *phe³* 'father' is spelled as "*pha, a, e, calya-plophu*", indicating ᖴ + ᑕ + ᐝ + ᱺ. The tone marks ᱹ PLOPHU, ᱻ CALYA, and ᱼ CALYA PLOPHU are applied below the vowel vowel letters ᑕ A, ᗯ OE, ᘁ I, and ᘓ OO. Where tone marks and vowel diacritics co-occur on the base letters ᑕ A and ᗯ OE, the vowels are conventionally entered first, though applications must recognize that for ᖰ *e*, ᑕ + ᱻ + ᐝ is equivalent to ᑕ + ᐝ + ᱻ. Although the vowel diacritic ᱻ *o* and the mid-tone mark ᱼ could be analyzed as combinations of ᱻ + ᐛ and ᱺ + ᱻ respectively (compare Myanmar ᱻ *o*, a combination of ᱹ *i* and ᱺ *u*), they are encoded as units for simplicity and to prevent ambiguity: *ᱻ and *ᱼ do not occur, but could be spelling errors if users expected to compose the diacritics. In addition, they appear as units in didactic materials produced by the user community.

Naming

Character names use the usual UCS conventions, with some adjustments which relate to Latin orthographies used for Kayah Li. The glyphs, names, and phonetic values (for Western Kayah Li) are given here, along with the values of the set of vowels used in loanwords: ꤊ KA [k], ꤋ KHA [kʰ], ꤌ GA [g], ꤍ NGA [ŋ], ꤎ SA [s], ꤏ SHA [sʰ], ꤐ ZA [ʑ], ꤑ NYA [ɲ], ꤒ TA [t], ꤓ HTA [tʰ], ꤔ NA [n], ꤕ PA [p], ꤖ PHA [pʰ], ꤗ MA [m], ꤘ DA [d], ꤙ BA [b], ꤚ RA [r], ꤛ YA [j], ꤜ LA [l], ꤝ WA [w], ꤞ THA [ɕ]/[θ]/[s], ꤟ HA [h], ꤠ VA [v], ꤡ CA [tɕ], ꤢ A [a], ꤣ OE [ɣ], ꤤ I [i], and ꤥ OO [o].

Digits and punctuation

Digits have distinctive forms. Kayah Li uses punctuation like COMMA, QUESTION MARK, EXCLAMATION MARK, HYPEN, PARENTHESES, and QUOTATION MARKs. It also has two unique punctuation marks, ⌢ CWI which indicates extended intonation, and ꤮ SHYA which functions as a phrase or sentence delimiter.

Ordering

Kayah Li sorts at the first level as shown below. It is likely that a set of contractions will be necessary to account for the vowel and tone behaviour in ISO/IEC 14651 and the UCA.

<div align="center">

ꤊ *ka* < ꤋ *kha* < ꤌ *ga* < ꤍ *nga* < ꤎ *sa* < ꤏ *sha* < ꤐ *za* < ꤑ *nya* <
ꤒ *ta* < ꤓ *hta* < ꤔ *na* < ꤕ *pa* < ꤖ *pha* < ꤗ *ma* < ꤘ *da* < ꤙ *ba* <
ꤚ *ra* < ꤛ *ya* < ꤜ *la* < ꤝ *wa* < ꤞ *tha* < ꤟ *ha* < ꤠ *va* < ꤡ *ca* <
ꤢ a^{55} < ꤢ́ $a^{55ʔ}$ < ꤢ̀ a^{11} < ꤢ̂ a^{31} < ꤣ o^{55} < ꤣ́ $o^{55ʔ}$ < ꤣ̀ o^{11} < ꤣ̂ o^{31} <
ꤤ i^{55} < ꤤ́ $i^{55ʔ}$ < ꤤ̀ i^{11} < ꤤ̂ i^{31} < ꤥ $ô^{55}$ < ꤥ́ $ô^{55ʔ}$ < ꤥ̀ $ô^{11}$ < ꤥ̂ $ô^{31}$ <
ꤢ̇ u^{55} < ꤢ̇́ $u^{55ʔ}$ < ꤢ̇̀ u^{11} < ꤢ̇̂ u^{31} < ꤢ̆̀ e^{55} < ꤢ̆̀ $e^{55ʔ}$ < ꤢ̆̀ e^{11} < ꤢ̆̀ e^{31} <
ꤢ̂̇ u^{55} < ꤢ̂̇́ $u^{55ʔ}$ < ꤢ̂̇̀ u^{11} < ꤢ̂̇̂ u^{31} < ꤢ̂̆ $ê^{55}$ < ꤢ̂̆́ $ê^{55ʔ}$ < ꤢ̂̆̀ $ê^{11}$ < ꤢ̂̆̂ $ê^{31}$ <
ꤣ̆ o^{55} < ꤣ̆́ $o^{55ʔ}$ < ꤣ̆̀ o^{11} < ꤣ̆̂ o^{31} < ꤝ̇ $ơu^{55}$ < ꤝ̇́ $ơu^{55ʔ}$ < ꤝ̇̀ $ơu^{11}$ < ꤝ̇̂ $ơu^{31}$ <
ꤝ̀ $uʾe^{55}$ < ꤝ̀́ $uʾe^{55ʔ}$ < ꤝ̀̀ $uʾe^{11}$ < ꤝ̀̂ $uʾe^{31}$ < ꤝ̂ $ơu^{55}$ < ꤝ̂́ $ơu^{55ʔ}$ < ꤝ̂̀ $ơu^{11}$ < ꤝ̂̂ $ơu^{31}$ <
ꤝ̆ $ơê^{55}$ < ꤝ̆́ $ơê^{55ʔ}$ < ꤝ̆̀ $ơê^{11}$ < ꤝ̆̂ $ơê^{31}$ < ꤝ̆ $ơo^{55}$ < ꤝ̆́ $ơo^{55ʔ}$ < ꤝ̆̀ $ơo^{11}$ < ꤝ̆̂ $ơo^{31}$

</div>

Linebreaking

Kayah Li uses spaces between graphemic 'word' units and so inter-word line-breaking can be performed based on character properties. Opportunities for hyphenation are lexical, but a hyphen may not be inserted between a base vowel and a diacritic.

Implementations

William G. Kauffman implemented Kayah Li in an eight-bit DOS font in 1990. J. Fraser Bennett implemented Kayah Li in a number of eight-bit fonts between 1991 and 1993, together with a keyboard layout design (see Figure 2).

Issues

Vowel diacritics have been assigned to canonical combining class 230, which is the standard class for above marks. It is proposed that the tone diacritics be assigned to class 240, which has been used for below marks, though up to now only in the exceptional case of U+0345 COMBINING GREEK YPOGEGRAMMENI (iota subscript). The rationale for using class 240 for the iota subscript is that this mark is derived from a vowel iota that orthographically would follow the initial vowel—that is, the vowel base to which the iota subscript is applied; thus, the iota subscript logically follows the above breathing or accent marks that apply to the vowel base. By using class 240 for this below mark, canonically-ordered sequences follow the logical order that would be expected by users in editing and needed for various linguistic processes.

This same rationale is applicable to the Kayah Li tone marks: the tone marks logically apply to the syllable as a whole and so logically follow the vowel marks. This is the order users expect when editing text, and the order that would be most amenable to various linguistic processes such as sorting. By

assigning the tone marks to class 240, normalization will result in the logical ordering of Consonant-Vowel-Tone that users expect and that linguistic processes require.

If the standard below class of 220 were used, then users would face a counter-intuitive order of marks when editing normalized text. For example, a sequence ε + ◌̀ + ◌ U+A922 U+A927 U+A92D would be normalized to ε + ◌ + ◌̀ U+A922 U+A92D U+A927. If a user receives such data and goes to edit it, say, backspacing to delete the last mark, then he or she would experience the counter-intuitive result that the vowel diacritic is deleted and the tone left behind, rather than the opposite. While it is true that users may encounter data having marks in either order, regardless of which classes are assigned, by assigning the tone diacritics to class 240 so as to give logically-correct order when normalized will reduce the likelihood of users encountering problems.

Similarly, if the standard below class of 220 were used, then normalization would not produce the ordering that is most useful for linguistic processes. Linguistic processes such as collation would begin by applying normalization, but would still require some further mechanism utilized after normalization to allow the vowel and tone marks to have their effect in the logically-correct order. For example, if a collation weight is to be assigned to a sequence ε + ◌̀ + ◌ U+A922 U+A927 U+A92D but normalization re-orders this as ε + ◌ + ◌̀ U+A922 U+A92D U+A927, then the collation process either will need to reverse the order of the marks so that weights apply in the correct order, or the collation process will need to assign single weights to every possible syllable. This problem does not arise if the tones diacritics are assigned to class 240, however. By assigning the tones to class 240, normalization will immediately provide the order that is needed for collation and other such linguistic processes.

If a class of 240 is deemed unacceptable, then we propose that the vowel marks be assigned to class 0 and the tone marks to class 220. The assignment of vowel marks to class 0 is consistent with the use of class 0 for vowel marks in other Brahmic and Brahmic-influenced scripts to which these vowel marks are historically related. By assigning above and below marks to separate classes with the tone marks in a higher-numbered class than the vowel marks, then normalization will still result in marks that are in correct logical order, and so the concerns mentioned above will still be accommodated.

Unicode Character Properties

```
A900;KAYAH LI DIGIT ZERO;Nd;0;L;;0;0;0;N;;;;;
A901;KAYAH LI DIGIT ONE;Nd;0;L;;1;1;1;N;;;;;
A902;KAYAH LI DIGIT TWO;Nd;0;L;;2;2;2;N;;;;;
A903;KAYAH LI DIGIT THREE;Nd;0;L;;3;3;3;N;;;;;
A904;KAYAH LI DIGIT FOUR;Nd;0;L;;4;4;4;N;;;;;
A905;KAYAH LI DIGIT FIVE;Nd;0;L;;5;5;5;N;;;;;
A906;KAYAH LI DIGIT SIX;Nd;0;L;;6;6;6;N;;;;;
A907;KAYAH LI DIGIT SEVEN;Nd;0;L;;7;7;7;N;;;;;
A908;KAYAH LI DIGIT EIGHT;Nd;0;L;;8;8;8;N;;;;;
A909;KAYAH LI DIGIT NINE;Nd;0;L;;9;9;9;N;;;;;
A90A;KAYAH LI LETTER KA;Lo;0;L;;;;;N;;;;;
A90B;KAYAH LI LETTER KHA;Lo;0;L;;;;;N;;;;;
A90C;KAYAH LI LETTER GA;Lo;0;L;;;;;N;;;;;
A90D;KAYAH LI LETTER NGA;Lo;0;L;;;;;N;;;;;
A90E;KAYAH LI LETTER SA;Lo;0;L;;;;;N;;;;;
A90F;KAYAH LI LETTER SHA;Lo;0;L;;;;;N;;;;;
A910;KAYAH LI LETTER ZA;Lo;0;L;;;;;N;;;;;
A911;KAYAH LI LETTER NYA;Lo;0;L;;;;;N;;;;;
A912;KAYAH LI LETTER TA;Lo;0;L;;;;;N;;;;;
A913;KAYAH LI LETTER HTA;Lo;0;L;;;;;N;;;;;
A914;KAYAH LI LETTER NA;Lo;0;L;;;;;N;;;;;
A915;KAYAH LI LETTER PA;Lo;0;L;;;;;N;;;;;
A916;KAYAH LI LETTER PHA;Lo;0;L;;;;;N;;;;;
A917;KAYAH LI LETTER MA;Lo;0;L;;;;;N;;;;;
A918;KAYAH LI LETTER DA;Lo;0;L;;;;;N;;;;;
A919;KAYAH LI LETTER BA;Lo;0;L;;;;;N;;;;;
A91A;KAYAH LI LETTER RA;Lo;0;L;;;;;N;;;;;
A91B;KAYAH LI LETTER YA;Lo;0;L;;;;;N;;;;;
A91C;KAYAH LI LETTER LA;Lo;0;L;;;;;N;;;;;
A91D;KAYAH LI LETTER WA;Lo;0;L;;;;;N;;;;;
```

```
A91E;KAYAH LI LETTER THA;Lo;0;L;;;;;N;;;;;
A91F;KAYAH LI LETTER HA;Lo;0;L;;;;;N;;;;;
A920;KAYAH LI LETTER VA;Lo;0;L;;;;;N;;;;;
A921;KAYAH LI LETTER CA;Lo;0;L;;;;;N;;;;;
A922;KAYAH LI LETTER A;Lo;0;L;;;;;N;;;;;
A923;KAYAH LI LETTER OE;Lo;0;L;;;;;N;;;;;
A924;KAYAH LI LETTER I;Lo;0;L;;;;;N;;;;;
A925;KAYAH LI LETTER OO;Lo;0;L;;;;;N;;;;;
A926;KAYAH LI VOWEL UE;Mn;230;L;;;;;N;;;;;
A927;KAYAH LI VOWEL E;Mn;230;L;;;;;N;;;;;
A928;KAYAH LI VOWEL U;Mn;230;L;;;;;N;;;;;
A929;KAYAH LI VOWEL EE;Mn;230;L;;;;;N;;;;;
A92A;KAYAH LI VOWEL O;Mn;230;L;;;;;N;;;;;
A92B;KAYAH LI TONE PLOPHU;Mn;240;L;;;;;N;;;;;
A92C;KAYAH LI TONE CALYA;Mn;240;L;;;;;N;;;;;
A92D;KAYAH LI TONE CALYA PLOPHU;Mn;240;L;;;;;N;;;;;
A92E;KAYAH LI SIGN CWI;Po;0;L;;;;;N;;;;;
A92F;KAYAH LI SIGN SHYA;Po;0;L;;;;;N;;;;;
```

Bibliography

Bennett, J. Fraser. 1993. *Kayah Li script: a brief description.* Urbana-Champaign: University of Illinois.

Karenni Literature Department. 1994. *ꤊꤢ꤬ꤛꤢꤩꤚꤢ꤭-ꤊꤢꤩꤛꤢꤩ ꤞꤛꤣꤜꤣꤨꤛꤢꤩ: The modern Western Kayah Li-English lexicon.* [Chiang Mai]: Payap University.

Solnit, David B. 1997. *Eastern Kayah Li: grammar, texts, glossary.* Honolulu: University of Hawai'i Press. ISBN 0-8248-1743-5

Acknowledgements

Figures

Figure 1. Kayah Li keyboard layout, from Bennett 1993, highlighting combining diacritical marks.

Figure 2. Sample text from Bennett 1993 of the folk tale "The Tortoise and the Monkey". The punctuation mark CWI can be seen here, as can European question mark and quotation marks.

Notes On The Use Of The Lexicon

A lexicon is a very useful tool for those who know how to glean from its resources. People can look up words to find their meanings, appropriate spellings, and how words are used. It takes time and patience for the person who is not accustomed to using a lexicon, but once mastered, this skill will reap a lifetime of benefit.

This lexicon is in the alphabetical order of the Western Kayah Li language. The order of consonants begins with /n/ and ends with /N/. The vowels follow the consonants.

Order of Consonants:

Order of Vowels:

An orthography and pronunciation guide begins on page 7 while a useful list of words in various semantic domains occurs in the appendix section of the lexicon. This latter section is useful for those who are learning Western Red Karen.

Figure 3. Example in Kayah Li and Latin scripts from the front matter of the 1994 *Western Kayah Li-English Lexicon*. Alphabetical order is shown, as is the punctuation character SHYA.

TABLE XXX - Row A9: KAYAH LI

	A90	A91	A92
0	O	ß	Ϻ
1	ā	⊇	Π
2	V	⅋	Ɛ
3	ℤ	⊔	⊔
4	�𝖓	⅊	Ɔ
5	Γ	℮	θ
6	⅁	⋒	◌̇
7	⊔	⩊	◌̀
8	∟	ℤ	◌̂
9	∪	⅁	◌̆
A	⅏	ⴄ	◌̈
B	Ə	⋂	◌̣
C	Ꞁ	⅊	◌̑
D	⊏	⅂	◌̤
E	℘	⅁	⌣
F	⅁	℘	I

G = 00
P = 00

TABLE XX - Row A9: KAYAH LI

hex	Name	hex	Name
00	KAYAH LI DIGIT ZERO		
01	KAYAH LI DIGIT ONE		
02	KAYAH LI DIGIT TWO		
03	KAYAH LI DIGIT THREE		
04	KAYAH LI DIGIT FOUR		
05	KAYAH LI DIGIT FIVE		
06	KAYAH LI DIGIT SIX		
07	KAYAH LI DIGIT SEVEN		
08	KAYAH LI DIGIT EIGHT		
09	KAYAH LI DIGIT NINE		
0A	KAYAH LI LETTER KA		
0B	KAYAH LI LETTER KHA		
0C	KAYAH LI LETTER GA		
0D	KAYAH LI LETTER NGA		
0E	KAYAH LI LETTER SA		
0F	KAYAH LI LETTER SHA		
10	KAYAH LI LETTER ZA		
11	KAYAH LI LETTER NYA		
12	KAYAH LI LETTER TA		
13	KAYAH LI LETTER HTA		
14	KAYAH LI LETTER NA		
15	KAYAH LI LETTER PA		
16	KAYAH LI LETTER PHA		
17	KAYAH LI LETTER MA		
18	KAYAH LI LETTER DA		
19	KAYAH LI LETTER BA		
1A	KAYAH LI LETTER RA		
1B	KAYAH LI LETTER YA		
1C	KAYAH LI LETTER LA		
1D	KAYAH LI LETTER WA		
1E	KAYAH LI LETTER THA		
1F	KAYAH LI LETTER HA		
20	KAYAH LI LETTER VA		
21	KAYAH LI LETTER CA		
22	KAYAH LI LETTER A		
23	KAYAH LI LETTER OE		
24	KAYAH LI LETTER I		
25	KAYAH LI LETTER OO		
26	KAYAH LI VOWEL UE		
27	KAYAH LI VOWEL E		
28	KAYAH LI VOWEL U		
29	KAYAH LI VOWEL EE		
2A	KAYAH LI VOWEL O		
2B	KAYAH LI TONE PLOPHU		
2C	KAYAH LI TONE CALYA		
2D	KAYAH LI TONE CALYA PLOPHU		
2E	KAYAH LI SIGN CWI		
2F	KAYAH LI SIGN SHYA		

A. Administrative

1. Title
Proposal for encoding the Kayah Li script in the BMP of the UCS.
2. Requester's name
Michael Everson
3. Requester type (Member body/Liaison/Individual contribution)
Individual contribution.
4. Submission date
2006-03-09
5. Requester's reference (if applicable)
6. Choose one of the following:
6a. This is a complete proposal
Yes.
6b. More information will be provided later
No.

B. Technical – General

1. Choose one of the following:
1a. This proposal is for a new script (set of characters)
Yes.
Proposed name of script
Kayah Li.
1b. The proposal is for addition of character(s) to an existing block
No.
1c. Name of the existing block
2. Number of characters in proposal
48
3. Proposed category (A-Contemporary; B.1-Specialized (small collection); B.2-Specialized (large collection); C-Major extinct; D-Attested extinct; E-Minor extinct; F-Archaic Hieroglyphic or Ideographic; G-Obscure or questionable usage symbols)
Category A.
4a. Proposed Level of Implementation (1, 2 or 3)
Level 3
4b. Is a rationale provided for the choice?
Yes.
4c. If YES, reference
Kayah Li uses combining diacritical marks.
5a. Is a repertoire including character names provided?
Yes.
5b. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?
Yes.
5c. Are the character shapes attached in a legible form suitable for review?
Yes.
6a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?
Michael Everson.
6b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:
Michael Everson, Fontographer.
7a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?
Yes.
7b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?
Yes.
8. 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.
9. 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.
See above.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.
Yes, a preliminary proposal was submitted in N3024.
2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?
Yes.

2b. If YES, with whom?

David Solnit and J. Fraser Bennett.

2c. If YES, available relevant documents

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

Yes.

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

Used to write the Kayah Li language.

4b. Reference

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

Yes.

5b. If YES, where?

In Thailand and Myanmar.

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

Yes. Positions A900-A92F are proposed.

6b. If YES, is a rationale provided?

Yes.

6c. If YES, reference

Contemporary use and accordance with the Roadmap.

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

Yes.

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

No.

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

8c. If YES, reference

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

No.

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

9c. If YES, reference

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

Yes.

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

Five vowel signs bear a superficial resemblence to some Latin diacritics. They have a significantly different range of acceptable glyphs and are encoded separately as a matter of principle.

10c. If YES, reference

See the Principles and Procedures document WG2 N3002, section F.6.

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

Yes.

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

Yes.

11c. If YES, reference

Vowels diacritics and tone marks.

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

Yes.

11e. If YES, reference

On page 2 the combinations of vowel diacritics, tone marks, and the letters A and OE are given.

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

No.

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

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

No.

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