## ISO/IEC JTC1/SC2/WG2 N3xxx

L2/06-xxx 2006-04-02

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 characters for Myanmar minority languages in the UCS

**Source:** Michael Everson

**Status: Individual Contribution** 

**Replaces: N2768, N1883R** 

Action: For consideration by JTC1/SC2/WG2 and UTC

Date: 2006-04-02

Since the Myanmar script was first encoded, it has been known that a number of additions used by minority languages would be needed. This proposal requests their addition. It contains the proposal summary form.

The languages supported by this proposal are Mon, Shan, Pali in Shan script, S'gaw Karen, Western Pwo Karen, Eastern Pwo Karen, and Kayah. Most of the characters proposed are spacing letters, but combining vowel signs, combining medial consonant signs, and combining tone marks are also proposed.

The history of the Myanmar script is not one of a single line of development. The Mon script, in fact, is probably the oldest of the traditions, though the letterforms of the former have been "Burmicized" due to later influence, particularly that of typography. A number of language-specific differences arose during the period of development, much as has happened with the Arabic and Cyrillic scripts. Most of the letters are used in common, but some letters have language-specific forms. These are not unifiable with "standard" Myanmar letters, and books in Burmese about Mon, for instance, use both of them concurrently. In the discussion of the additions below, the language-specific letters are listed, in the brief shorthand "x contrasts with Burmese y".

The new rendering model proposed in N3043 which disunifies visible  $\S$  ASAT from stacking  $\square$  VIRAMA applies to the characters here. The general implication of this is that when following VIRAMA, the LETTERS here should be rendered below and usually slightly smaller than the base letter. Where there is a departure from the mainstream stacking behaviour for Burmese, characters have been added here to facilitate minority-language usage. For example,  $\S$  LETTER NA and  $\S$  LETTER MA subjoin as expected in Burmese  $\S$   $\S$  so the special Mon subjoined forms  $\S$  CONSONANT SIGN MON MEDIAL NA and  $\S$  CONSONANT SIGN MON MA are encoded as dependent consonant signs.

#### **Additions for Mon**

A number of characters contrast with Burmese characters: @ LETTER MON E contrasts with Burmese e LETTER E; Ç LETTER MON NGA contrasts with Burmese c LETTER NGA; & LETTER MON JHA contrasts with Burmese of LETTER JHA; VOWEL SIGN MON II contrasts with Burmese vowel sign II. Other characters are unique to Mon: Vowel sign mon o represents open o; LETTER MON BBA and G LETTER MON BBE represent bilabial implosives; Consonant sign mon na and Consonant sign mon ma are mentioned above; Consonant sign mon medial la is used in S'gaw Karen for ya, where it contrasts with Burmese Consonant sign medial ya proposed in N3043. A clarification should be given regarding C LETTER

MON NGA which merits discussion. While Burmese c NGA stacks as expected  $\infty$ , Mon  $\infty$  NGA stacks by rendering only its "diacritic"  $\infty$ ; that is, \* $\infty$  does not occur. It could be possible to treat the diacritic as a medial (like  $\infty$  CONSONANT SIGN MON MEDIAL NA and  $\infty$  CONSONANT SIGN MON MA) but this is impractical due to attested glyph variants of the letter. Both  $\infty$  and  $\infty$  are attested glyph variants of the letter, and if \* $\infty$  were encoded as a medial, it could be confusing as to whether it should be applied to c LETTER NGA or to  $\infty$  VOWEL SIGN E (which isn't an independent letter, but a dependent vowel). The encoding proposed here is simpler. (Figures 1, 2, 3, 9, and 10.)

#### **Additions for Shan**

A number of characters contrast with Burmese characters:  $\infty$  Letter shan a contrasts with Burmese  $\infty$  Letter a;  $\infty$  Letter shan ka contrasts with Burmese  $\infty$  Letter ka;  $\infty$  Letter shan ka contrasts with Burmese  $\infty$  Letter ka;  $\infty$  Letter shan ka contrasts with Burmese  $\infty$  Letter ca;  $\infty$  Letter shan nya contrasts with Burmese  $\infty$  Letter nya;  $\infty$  Letter shan na contrasts with Burmese  $\infty$  Letter na;  $\infty$  Letter shan pha contrasts with Burmese  $\infty$  Letter pha;  $\infty$  Letter shan tha contrasts with Burmese  $\infty$  Letter shan ha contrasts with Burmese  $\infty$  Letter shan ha contrasts with Burmese  $\infty$  Letter shan ha contrasts with Burmese  $\infty$  Letter na;  $\infty$  Lette

#### **Additions for Shan Pali**

A number of characters used for Pali in Shan context contrast with Burmese characters:  $\wp$  Letter shan Pali GA contrasts with Burmese  $\wp$  Letter GA;  $\wp$  Letter shan Pali GHA contrasts with Burmese  $\wp$  Letter shan Pali JA contrasts with Burmese  $\wp$  Letter shan Pali JHA contrasts with Burmese  $\wp$  Letter JHA;  $\wp$  Letter shan Pali TTA contrasts with Burmese  $\wp$  Letter TTA;  $\wp$  Letter shan Pali TTHA contrasts with Burmese  $\wp$  Letter TTHA;  $\wp$  Letter shan Pali DDA contrasts with Burmese  $\wp$  Letter DDA;  $\wp$  Letter shan Pali DDA contrasts with Burmese  $\wp$  Letter shan Pali DA contrasts with Burmese  $\wp$  Letter DDA;  $\wp$  Lette

#### Additions for S'gaw Karen

Three characters used in S'gaw Karen are not used in Burmese: § LETTER SGAW KAREN SHA contrasts with Burmese § *rha* (which is § LETTER RA + § CONSONANT SIGN MEDIAL HA); p SIGN SGAW KAREN HATHI and 1 SIGN SGAW KAREN KE PHO are tone marks. (Figures 7, 8, 10, 11, 12, and 14.)

#### **Additions for Western Pwo Karen**

Four characters used in Western Pwo Karen are not used in Burmese:  $\circ$  Letter western Pwo Karen that contrasts with Burmese  $\otimes$  sa;  $\circ$  Letter western Pwo Karen Pwa is the last letter of the Western Pwo Karen alphabet;  $\circ$  vowel sign western Pwo Karen eu and  $\circ$  vowel sign western Pwo Karen ue are vowel signs;  $\circ$  sign western Pwo Karen tone-1,  $\circ$  sign western Pwo Karen tone-2,  $\circ$ 1 sign western Pwo Karen tone-3,  $\circ$ 1 sign western Pwo Karen tone-4, and  $\circ$ 5 sign western Pwo Karen tone-5 are tone marks. (Figures 13, 14, 15, and 16.)

#### **Additions for Eastern Pwo Karen**

#### **Additions for Kayah**

One character used in Kayah contrasts with a Burmese character: § vowel sign kayah u is used alongside Burmese vowel sign u. Two characters are unique to Kayah: § vowel sign kayah ae and § vowel sign kayah ee. (Figure 20.)

#### **Ordering**

The unified order for the Myanmar script incorporating the characters here is given below. Ordering is syllable-based, so this is indicative of only one level of ordering.

```
ka < shan-ka < kha < shan-kha < ga < shan-pali-ga < gha < shan-pali-gha < nga < mon-nga <
ca < shan-ca < cha < ja < shan-pali-ja < jha < shan-pali-jha < mon-jha < sgaw-karen-sha < nya < shan-nya < nnya <
tta < shan-pali-tta < ttha < shan-pali-ttha < dda < shan-pali-dda <
ddha < shan-pali-ddha < nna < shan-pali-nna < eastern-pwo-karen-nna <
ta < tha < da < shan-pali-da < dha < shan-pali-dha < na < shan-na <
pa < pha < shan-pha < shan-fa < ba < bha < shan-pali-bha < ma <
ya < ra < la < wa < shan-tha < sha < ssa < western-pwo-karen-tha < sa < great-sa < ha < shan-ha <
lla < shan-pali-lla < mon-bba < eastern-pwo-karen-ywa < eastern-pwo-karen-gwa <
a < shan-a < shan-hsipaw-fa < shan-hsipaw-ra < i < ii < u < uu <
vocalic-r < vocalic-rr < vocalic-ll < vocalic-ll < e < mon-bbe < western-pwo-karen-pwa < mon-e < o < au
```

#### **Issues**

Several characters look as though they could be sequences of a base character plus the proposed \*U+103E  $\circ$  Consonant sign medial ha. These are:  $\wp$  letter shan fa,  $\wp$  letter shan ha,  $\wp$  letter sgaw Karen sha,  $\phi$  letter western pwo karen pwa,  $\phi$  letter eastern pwo karen ywa, and  $\phi$  letter EASTERN PWO KAREN GHWA. At the meeting in Yangon, this was discussed at length. All of the letters have their own place in the alphabet of the language which uses them (see Figures 4, 7, 12, and 18). In Pwo Karen, for instance, η LETTER SGAW KAREN SHA and φ LETTER WESTERN PWO KAREN PWA, are both treated as separate letters of the alphabet, sorted rather far from the base letters q ra and o pa, both of which do, as it happens, take genuine medials ( $\circ$  pya,  $\circ$  pla,  $\circ$  pra,  $\circ$  pwa, and  $\circ$  rwa) which are sorted as expected under o pa and q ra. In better typography and in handwriting, q LETTER SGAW KAREN SHA contrasts with qrha (though much printed matter is not exemplary of the best typography). Shan uses two letters, & LETTER SHAN FA and o LETTER SHAN HA, both of which can be seen with alternate shapes up and o. It's Kam Mong 2004) a shape like a, where the top part isn't analyzable to any other letter. Since the Myanmar script is to be a unified set to deal with all of these languages, we judge it best to let of be used in its traditional productive medial role in the Burmese, Mon, and S'gaw Karen languages, but to encode as unique letters the ones used non-productively in Pwo Karen. In Pwo Karen, the angled marks are typically fused to the letters, which also suggests a difference. Precedent for this can be found in the letterforms used for jha: In Burmese, alongside  $\varphi_i$  is found the form  $\varphi_i$ , which looks as though it is a ligature of o ca and g -ya (but it doesn't), and in Mon the forms g and g are often found, probably better rendered  $e_j$  in a generic font; it is not  $e_j$   $a_j$   $a_$ 

#### **Unicode Character Properties**

```
1022; MYANMAR LETTER SHAN A;Lo;0;L;;;;N;;;;
1028; MYANMAR LETTER MON E;Lo;0;L;;;;N;;;;
1033; MYANMAR VOWEL SIGN MON II;Mn;0;NSM;;;;N;;;;
1034; MYANMAR VOWEL SIGN MON O;Mn;0;NSM;;;;N;;;;
1035; MYANMAR VOWEL SIGN SHAN E;Mc;0;L;;;;N;;;;
105A; MYANMAR LETTER MON NGA;Lo;0;L;;;;N;;;;
```

```
105B; MYANMAR LETTER MON JHA; Lo; 0; L;;;; N;;;;
105C; MYANMAR LETTER MON BBA; Lo; 0; L;;;;; N;;;;
105D; MYANMAR LETTER MON BBE; Lo; 0; L;;;;; N;;;;;
105E; MYANMAR CONSONANT SIGN MON MEDIAL NA; Mn; 0; NSM;;;;; N;;;;;
105F; MYANMAR CONSONANT SIGN MON MEDIAL MA; Mn; 0; NSM; ;; ;; N; ;; ;;
1060; MYANMAR CONSONANT SIGN MON MEDIAL LA; Mn; 0; NSM;;;;; N;;;;
1061; MYANMAR LETTER SHAN KA; Lo; 0; L;;;;; N;;;;;
1062; MYANMAR LETTER SHAN KHA; Lo; 0; L;;;;; N;;;;;
1063; MYANMAR LETTER SHAN CA; Lo; 0; L;;;;; N;;;;;
1064; MYANMAR LETTER SHAN NYA; Lo; 0; L;;;;; N;;;;
1065; MYANMAR LETTER SHAN NA; Lo; 0; L;;;;; N;;;;;
1066; MYANMAR LETTER SHAN PHA; Lo; 0; L;;;;; N;;;;;
1067; MYANMAR LETTER SHAN FA; Lo; 0; L;;;;; N;;;;
1068; MYANMAR LETTER SHAN THA; Lo; 0; L;;;;; N;;;;;
1069; MYANMAR LETTER SHAN HA; Lo; 0; L;;;; N;;;;
106A; MYANMAR LETTER SHAN HSIPAW FA; Lo; 0; L;;;;; N;;;;
106B; MYANMAR LETTER SHAN HSIPAW RA; Lo; 0; L;;;;; N;;;;
106C; MYANMAR VOWEL SIGN SHAN A; Mc; 0; L;;;;; N;;;;;
106D; MYANMAR VOWEL SIGN SHAN AA; Mc; 0; L;;;;; N;;;;;
106E; MYANMAR VOWEL SIGN SHAN EE ABOVE; Mn; 0; NSM;;;;; N;;;;
106F; MYANMAR VOWEL SIGN SHAN E ABOVE; Mn; 0; NSM;;;;; N;;;;
1070; MYANMAR VOWEL SIGN SHAN FINAL Y; Mn; 0; NSM;;;;; N;;;;
1071; MYANMAR SIGN SHAN TONE-2; Mc; 0; L;;;;; N;;;;
1072; MYANMAR SIGN SHAN TONE-3; Mc; 0; L;;;; N;;;;
1073; MYANMAR SIGN SHAN COUNCIL TONE-4; Mc; 0; L;;;;; N;;;;
1074; MYANMAR SIGN SHAN TONE-5; Mc; 0; L;;;;; N;;;;;
1075; MYANMAR SIGN SHAN TONE-6; Mc; 0; L;;;;; N;;;;
1076; MYANMAR SIGN SHAN COUNCIL EMPHATIC TONE; Mn; 0; NSM;;;;; N;;;;
1077; MYANMAR CONSONANT SIGN SHAN MEDIAL WA; Mn; 0; NSM;;;;; N;;;;
1078; MYANMAR LETTER SHAN PALI GA; Lo; 0; L;;;;; N;;;;
1079; MYANMAR LETTER SHAN PALI GHA; Lo; 0; L;;;;; N;;;;
107A; MYANMAR LETTER SHAN PALI JA; Lo; 0; L;;;;; N;;;;
107B; MYANMAR LETTER SHAN PALI JHA; Lo; 0; L;;;;; N;;;;
107C; MYANMAR LETTER SHAN PALI TTA; Lo; 0; L;;;;; N;;;;
107D; MYANMAR LETTER SHAN PALI TTHA; Lo; 0; L;;;;; N;;;;
107E; MYANMAR LETTER SHAN PALI DDA; Lo; 0; L;;;;; N;;;;;
107F; MYANMAR LETTER SHAN PALI DDHA; Lo; 0; L;;;;; N;;;;
1080; MYANMAR LETTER SHAN PALI NNA; Lo; 0; L;;;;; N;;;;;
1081; MYANMAR LETTER SHAN PALI DA; Lo; 0; L;;;;; N;;;;;
1082; MYANMAR LETTER SHAN PALI DHA; Lo; 0; L;;;;; N;;;;
1083; MYANMAR LETTER SHAN PALI BA; Lo; 0; L;;;;; N;;;;;
1084; MYANMAR LETTER SHAN PALI LLA; Lo; 0; L;;;;; N;;;;
1085; MYANMAR LETTER SGAW KAREN SHA; Lo; 0; L;;;;; N;;;;
1086; MYANMAR SIGN SGAW KAREN HATHI; Mc; 0; L;;;;; N;;;;
1087; MYANMAR SIGN SGAW KAREN KE PHO; Mc; 0; L;;;;; N;;;;
1088; MYANMAR LETTER WESTERN PWO KAREN THA; Lo; 0; L;;;;; N;;;;
1089; MYANMAR LETTER WESTERN PWO KAREN PWA; Lo; 0; L;;;;; N;;;;
108A; MYANMAR VOWEL SIGN WESTERN PWO KAREN EU; Mc; 0; L;;;;; N;;;;
108B; MYANMAR VOWEL SIGN WESTERN PWO KAREN UE; MC; 0; L;;;;; N;;;;
108C; MYANMAR SIGN WESTERN PWO KAREN TONE-1; Mc; 0; L;;;;; N;;;;
108D; MYANMAR SIGN WESTERN PWO KAREN TONE-2; Mc; 0; L;;;;; N;;;;
108E; MYANMAR SIGN WESTERN PWO KAREN TONE-3; Mc; 0; L;;;;; N;;;;
108F; MYANMAR SIGN WESTERN PWO KAREN TONE-4; Mc; 0; L;;;; N;;;;
1090; MYANMAR SIGN WESTERN PWO KAREN TONE-5; Mc; 0; L;;;;; N;;;;
1091; MYANMAR LETTER EASTERN PWO KAREN NNA; Lo; 0; L;;;;; N;;;;
1092; MYANMAR LETTER EASTERN PWO KAREN YWA; Lo; 0; L;;;;; N;;;;
1093; MYANMAR LETTER EASTERN PWO KAREN GHWA; Lo; 0; L;;;;; N;;;;
1094; MYANMAR VOWEL SIGN KAYAH OE; Mn; 0; NSM;;;;; N;;;;
1095; MYANMAR VOWEL SIGN KAYAH U; Mn; 0; NSM;;;;; N;;;;
1096; MYANMAR VOWEL SIGN KAYAH EE; Mn; 0; NSM; ; ; ; ; N; ; ; ;
```

#### **Bibliography**

Khu Myar Reh. 2005. The greedy dog and his bone. [Chiang Mai]: [Payap University].

Nai Pan Hla. 1989. *An introduction to Mon language*. Kyoto: Center for Southeast Asian Studies, Kyoto University.

```
ပပို့ဆုရြားလွိဳ,ထီး ပၿရာမို့ (Shan-Thai reader).

ဖျံုလိခင်းထား (Eastern Pwo Karen reader).

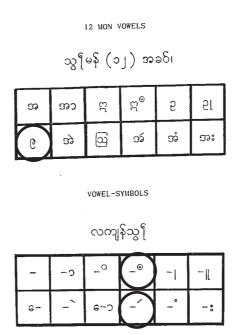
မွန်ဘာသာ: သင်ကြားနည်း (Mon grammar book).

သွရာ်မေ,လွိဳ,ထီး 200 ခုဝိ,မူင်း (Shan-Thai reader).

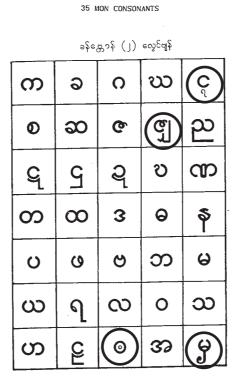
ပှဖြို့၏အလံုးခဲ့၊ထီး ဖိုး၁၊ဆအဂး (Western Pwo Karen reader).

ပှဖြို့၏ဆမျို့ ဒဲ ပယဒ္ဒအခ်ီးပယီး = ပိုးကရင်-မြန်မာအဘိဓါန် = Pwo Kayin-Myanmar Dictionary. 1989. (Western Pwo Karen).
```

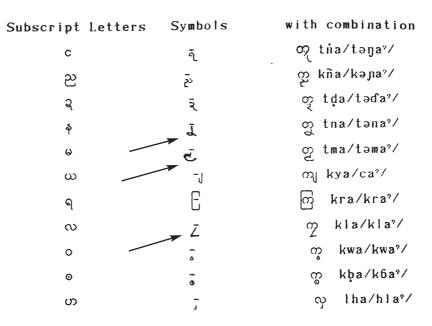
## Figures



**Figure 1.** Sample from Nai Pan Hla's 1989 Mon dictionary, showing LETTER MON E, VOWEL SIGN MON II, and VOWEL SIGN MON O.



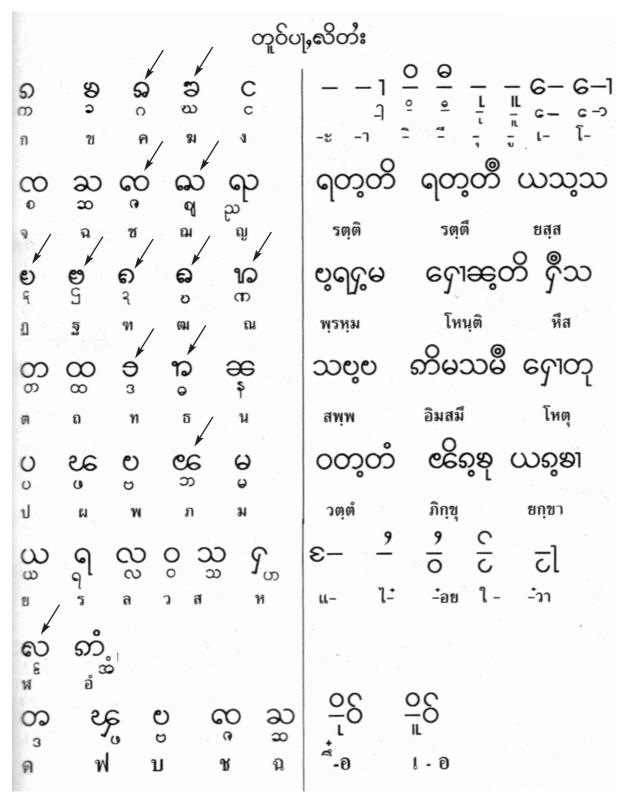
**Figure 2.** Sample from Nai Pan Hla's 1989 Mon dictionary, showing letter Mon NGA, letter Mon Jha, Letter Mon BBA, and letter Mon BBE.



**Figure 3.** Sample from Nai Pan Hla's 1989 Mon dictionary, showing CONSONANT SIGN MON MEDIAL NA, CONSONANT SIGN MON MEDIAL MA, and CONSONANT SIGN MON MEDIAL LA.



**Figure 4.** Sample from a Shan-Thai reader, showing letter shan ka, letter shan kha, letter shan ca, letter shan nya, letter shan na, letter shan pha, letter shan fa, letter shan ha, letter shan a, letter shan pali ga, letter shan hsipaw fa, letter shan hsipaw ra, letter shan tha, vowel sign shan e, vowel sign shan a, vowel sign shan a, vowel sign shan final y, vowel sign shan ee above, vowel sign shan e above, consonant sign shan medial wa, sign shan tone-2, sign shan tone-3, sign shan tone-5, and sign shan tone-6.



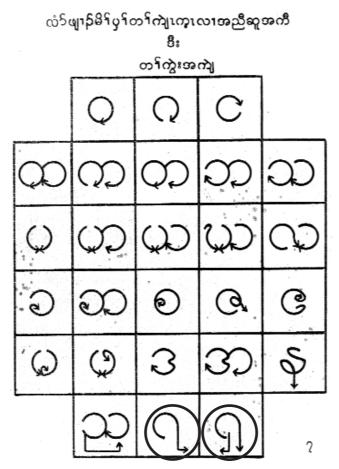
**Figure 5.** Sample from a reader in Shan and Thai, showing letter shan pali gha, letter shan pali ja, letter shan pali jha, letter shan pali tta, letter shan pali ttha, letter shan pali dda, letter shan pali ddha, letter shan pali dha, letter shan pali dha, letter shan pali dha, letter shan pali bha, and letter shan pali lla.



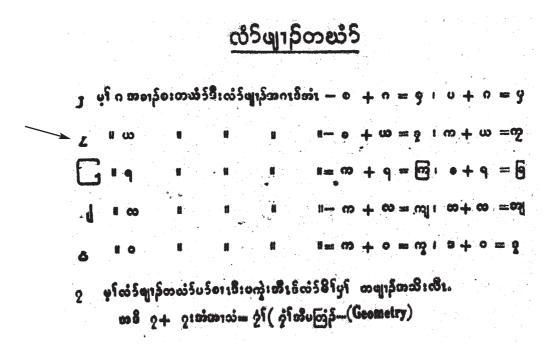
**Figure 6.** Sample from a Shan Council reader, showing SIGN SHAN COUNCIL EMPHATIC TONE and SIGN SHAN TONE-4.



Figure 7. Sample from a reader in S'gaw Karen, showing LETTER SGAW KAREN SHA.



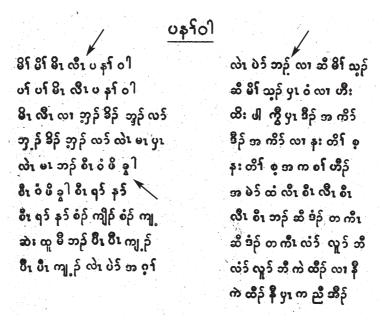
**Figure 8.** Sample from a reader in S'gaw Karen, showing the ductus for writing S'gaw Karen letters. Note the difference between LETTER RA and LETTER SGAW KAREN SHA.



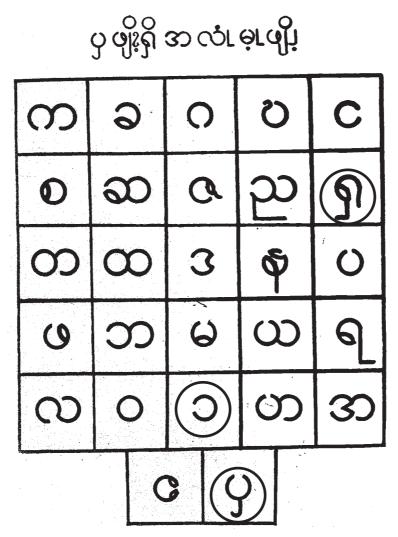
**Figure 9.** Sample from a reader in S'gaw Karen, showing CONSONANT SIGN MON MEDIAL LA (used in S'gaw Karen for *ya*) alongside characters proposed in N3043, CONSONANT SIGN MEDIAL HA (used in S'gaw Karen for *ga*), CONSONANT SIGN MEDIAL RA, CONSONANT SIGN MEDIAL YA (used in S'gaw Karen for *la*), and CONSONANT SIGN MEDIAL WA.

```
စား စာါ စားသံ 🛥 စာ်
               တားတါတားလံ 🕿 တာဉ်
= အးလံ
့ ဖျာဉ်ဆုံး
               စား စာါဖျာန်ဆံ၊ 🕿 စား
🗯 ဟၢသံ
               စားစားါဟ≀လံ ≡ အာβ்⊸
≖ ကူဉ်ဖိ
               တ၊ တါ ကုန် 8 m mr ←
               တား တာါ ေသံ 🛲 တင်
                တား တေါ်စူးဆံ 📟 တစ်
= #10i
                တားတာါနှူးလံ 📟 တာနိ
                စား စာါမႈလံ 🚾 အမြိ
                အး အေါ် လး သံ 🗪 အလိ
                အးေကါန္းသံ 🕿 အနီ
```

**Figure 10.** Sample from a reader in S'gaw Karen, showing vowels and tone marks, including vowel sign tall AA (proposed in N3043), vowel sign shan A, sign sgaw karen hathi, sign sgaw karen ke pho, and consonant sign mon medial la.



**Figure 11.** Sample from a reader in S'gaw Karen, showing SIGN SGAW KAREN HATHI and SIGN SGAW KAREN KE PHO. The CONSONANT SIGN MEDIAL WA (proposed in N3043) is shown with a triangular glyph.



**Figure 12.** Sample from a reader in Western Pwo Karen, showing LETTER SGAW KAREN SHA, LETTER WESTERN PWO KAREN THA, and LETTER WESTERN PWO KAREN PWA.

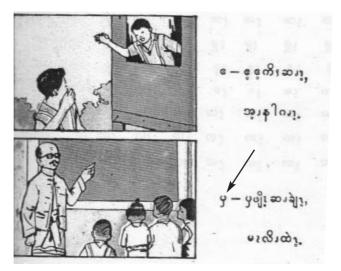


Figure 13. Sample from a reader in Western Pwo Karen, showing Letter western Pwo Karen Pwa listed as a separate letter of the alphabet:  $\varphi$  pwa is for  $\varphi \delta \varphi$  pwaphli 'Pwo Karen'.

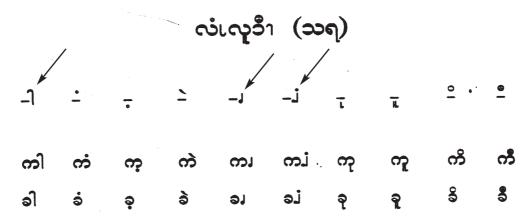
# ဆာမာလီ၊ - ၄၃



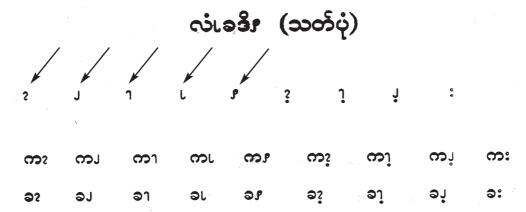
## **"**ധ്യീം രൂങ്ങറു:"

နးပံးတ်၊ – ဆာဟာသာရာန္ခု ဝင္စားဝံးဘဲးဘဲးတဲ့၊ နးပံးသီး – သာချန်ဆာကုံးပူး သာဘဲကဲ့အားပှာကိုး နးပံးတီး – သာချန်ဆာကုံးပူး ဝသာဘုံးတဲ့အားပှာကိုး နးပံးတီး – သာချန်ဆာကုံးပူး ဝသာဘုံးတဲ့အားပှာကိုး နးပံးတီး – သာချန်ဆာကုံးပူး ဝသာဘုံးတဲ့အားပွဲ၊ သာချာပော့၊ နးပံးတီး – သာချန်ဆာကုံးပူး ဝင္စားဝဲးဘုံးဘုံးတူး နေးပံးတီး – သာချန်ဆာကုံးပူး ဝင္စားဝဲးဘုံးဘုံးတူး နေးပံးသီး – မရွတ်းဂွားကချိန်းကုံးစားပွဲ၊ ဝင္စားဝဲးဘုံးဘုံးတိုး

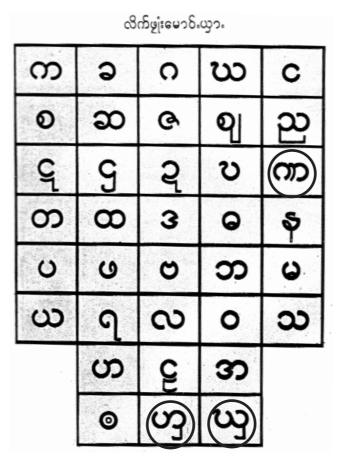
**Figure 14.** Sample from a reader in Western Pwo Karen, showing Letter western Pwo Karen tha, Letter sgaw karen sha, and letter western Pwo Karen Pwa.



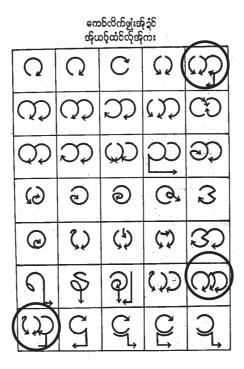
**Figure 15.** Sample from a reader in Western Pwo Karen, listing the vowel signs and showing vowel sign tall AA (proposed in N3043), as well as vowel sign western Pwo Karen Eu, and vowel sign western Pwo Karen UE.



**Figure 16.** Sample from a reader in Western Pwo Karen, listing the tone marks and showing Sign Western Pwo Karen tone-1, Sign Western Pwo Karen tone-2, Sign Western Pwo Karen tone-3, Sign Western Pwo Karen tone-4, Sign Western Pwo Karen tone-5.



**Figure 18.** Sample from a reader in Eastern Pwo Karen, showing Letter Eastern Pwo Karen NNA, Letter Eastern Pwo Karen Ywa, and Letter Eastern Pwo Karen Ghwa.



**Figure 19.** Sample from a reader in Eastern Pwo Karen, showing the ductus for writing Eastern Pwo Karen letters. Highlighted are LETTER EASTERN PWO KAREN YWA, LETTER EASTERN PWO KAREN NNA, and LETTER EASTERN PWO KAREN GHWA.



**Figure 20.** Sample from a story in Kayah, "The Greedy Dog and his Bone", showing VOWEL SIGN KAYAH OE, VOWEL SIGN KAYAH U, and VOWEL SIGN KAYAH EE.

**TABLE XX - Row 10: MYANMAR** 

	100	101	102	103	104	105	106	107	108	109
0	က	တ	<u> </u>	្	0	0	ূ	<b>9</b>	าด	ၭ
1	Э	$\infty$	အ	ေ	0	<b>(</b>	Ω	್ಕ	9	നാ
2	0	3	ဢ	े	J	C	ව	း	ಣ	ၯ
3	ဃ	0	33	<b>©</b>	9	ၓ	က	<b>:</b>	S	ၰ
4	С	<b>S</b>	නු	ं	9	е	ရာ	<b>.</b>	ಉ	ိ
5	D	C	5	ေဒ	၅	e	æ	ಿ	9	े
6	<b>x</b>	O	വാ	(_):	G	ు	ಎ	<u></u>	ၣ	္
7	Q	Ŋ	8	( ) ) · ( )	5	ು	35	ွ	ា	
8	ဈ	ဘ	ଡ		0	့	ႀ	ဂ	၁	
9	2	<b>Q</b>	သြ	+	၉	္	ß	ව	9	
Α	ည	ယ	ဪ	်	1	ç	တ	00	ា	
В	Ę	ବ	ി	্য	II	ၛ	ၿ	സ	ां	
С	G	8	ာ	୍ର		0	ា	W	ः	
D	ဍ	0	ಿ	္ဝ	ଶ	3	ા	W	ി	
E	ಬ	သ	ိ	្ន	9	ၞ	ိ	೧	ា	
F	ന	ဟ	ុ	ဿ	କ୍ଷ	ૃ	ိ	60	ು	

G = 00 P = 00

### **TABLE XX - Row 10: MYANMAR**

00	hex	Name	hex	Name
	01 02 03 04 05 06 07 08 09 00 00 00 00 00 00 00 00 00 00 00 00	MYANMAR LETTER GA MYANMAR LETTER GA MYANMAR LETTER GA MYANMAR LETTER GA MYANMAR LETTER NOA MYANMAR LETTER OA MYANMAR LETTER OA MYANMAR LETTER JA MYANMAR LETTER JA MYANMAR LETTER JA MYANMAR LETTER JA MYANMAR LETTER ITA MYANMAR LETTER THA MYANMAR LETTER TA MYANMAR LETTER TA MYANMAR LETTER TA MYANMAR LETTER TA MYANMAR LETTER DOA MYANMAR LETTER PA MYANMAR LETTER PA MYANMAR LETTER BA MYANMAR BA	5A 5BC 5DE 5F 60 61 2 63 44 5 66 67 8 9 AB 6 6 6 6 7 7 1 2 3 4 7 5 6 7 7 8 7 7 AB 7 7 DE 7 7 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 9	MYANMAR LETTER MON JHA MYANMAR LETTER MON BBA MYANMAR LETTER MON BBA MYANMAR LETTER MON BBA MYANMAR CONSONANT SIGN MON MEDIAL NA MYANMAR CONSONANT SIGN MON MEDIAL LA MYANMAR CONSONANT SIGN MON MEDIAL LA MYANMAR CONSONANT SIGN MON MEDIAL LA MYANMAR LETTER SHAN KA MYANMAR LETTER SHAN KA MYANMAR LETTER SHAN KA MYANMAR LETTER SHAN NA MYANMAR LETTER SHAN NA MYANMAR LETTER SHAN NA MYANMAR LETTER SHAN NA MYANMAR LETTER SHAN PHA MYANMAR LETTER SHAN HA MYANMAR VOWEL SIGN SHAN A MYANMAR VOWEL SIGN SHAN EA MYANMAR SIGN SHAN TONE-5 MYANMAR SIGN SHAN TONE-6 MYANMAR LETTER SHAN PALL GHA MYANMAR LETTER SHAN PALL DHA MYANMAR LET

#### A. Administrative

1. Title

Proposal for encoding characters for Myanmar minority languages vin the UCS.

2. Requester's name

Michael Everson

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

Individual contribution.

4. Submission date

2006-03-30

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

Nο

#### B. Technical – General

1. Choose one of the following:

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

Nο

Proposed name of script

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

Yes.

1c. Name of the existing block

Myanmar.

2. Number of characters in proposal

66

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 2

4b. Is a rationale provided for the choice?

Yes.

4c. If YES, reference

Brahmic Level 2 implementation.

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?

No.

7b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

No. The characters to be represented are not new to SC2.

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)?

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. See N2827, N1883R, and Unicode 1.0, which are all relevant documents.

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?

San Lwin (Director General, Myanmar Language Commission), Tun Tint (Myanmar Language Commission), Thein Oo (President, Myanmar Computer Federation), Kyaw Thein (Vice-President, Myanmar Computer Federation), Myint Myint Than (Director, Myanmar Computer Federation), Zaw Htut (Myanmar Computer Professional Association, Myanmar's NET), Htoo Myint Naung (MyMyanmar Project, Technomation Studios, Universities of Computer Studies Yangon), Myint Thu (MyMyanmar Project, Myanmar Heritage Publications), Ngwe Tun (Mon—Myanmar Computer Professional Association, Solveware Solution, Myanmar Info-Tech), Maung Maung Thant (Myanmar Computer Professional Association), Jai Pah Bung Mein (Shan—SSi Technologies), Saw Hare Sei (S'gaw Karen—Ayeyawady Data Centre), Saw Baldwin Khaing Oo (S'gaw Karen—Ayeyawady Data Centre), Nant Silver Tun (Western Pwo Karen—Pwo Kayin Conference), William Wai Lin Kyaw (Myanmar Computer Professional Association, Myanmar Linux Users Group), Ye Myat Thu (Alpha Mandalay, Alpha Info-Tech), Martin Hosken (Payap University), Keith Stribley (Thanlwinsoft).

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?

People in Myanmar.

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

Common

4b. Reference

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

Yes

5b. If YES, where?

In Myanmar.

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

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)?

N/A.

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

Yes.

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

This proposal requests disunifications and a change in the sequences currently specified for Myanmar because those sequences do not work. Apart from that, the proposed U+1088 MYANMAR LETTER WESTERN PWO KAREN THA looks like U+1041 MYANMAR DIGIT ONE but UCS encoding principles always differentiate such pairs. Compare U+101D MYANMAR LETTER WA and U+1040 MYANMAR DIGIT ZERO.

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?

Yes.

10c. If YES, reference

See "Issues" above.

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

Brahmic vowel and consonant signs.

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

No

11e. If YES, reference

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

No.

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

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

No

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