UC Berkeley
Proposals from the Script Encoding Initiative

Title
Proposal for encoding Myanmar characters for Karen and Kayah in the UCS

Permalink
https://escholarship.org/uc/item/8213t58x

Authors
Everson, Michael
Hosken, Martin

Publication Date
2006-09-08

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

ISO/IEC JTC1/SC2/WG2 N3142
L2/06-303
2006-09-08

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 Myanmar characters for Karen and Kayah in the UCS**
Source: **UC Berkeley Script Encoding Initiative (Universal Scripts Project)**
Authors: **Michael Everson and Martin Hosken**
Status: **Liaison Contribution**
Replaces: **N3080**
Action: **For consideration by JTC1/SC2/WG2 and UTC**
Date: **2006-09-08**

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 the addition of characters for a number of them. It contains the proposal summary form.

The languages supported by this proposal are the Tibeto-Burman languages Western Pwo Karen, Eastern Pwo Karen, Geba Karen, and Kayah. Five of the characters proposed are spacing letters, but six combining vowel signs, and five combining tone marks are also proposed.

The history of the Myanmar script is not one of a single line of development. 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 Karen, 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*".

Additions for Western Pwo Karen

Four characters used in Western Pwo Karen are not used in Burmese: ၥ LETTER WESTERN PWO KAREN THA contrasts with Burmese ၐ *sa*; ၡ LETTER WESTERN PWO KAREN PWA is the last letter of the Western Pwo Karen alphabet; ◌ၤ VOWEL SIGN WESTERN PWO KAREN EU and ◌ၢ VOWEL SIGN WESTERN PWO KAREN UE are vowel signs; ◌ၣ SIGN WESTERN PWO KAREN TONE-1, ◌ၥ SIGN WESTERN PWO KAREN TONE-2, ◌ၦ SIGN WESTERN PWO KAREN TONE-3, ◌ၧ SIGN WESTERN PWO KAREN TONE-4, and ◌ၨ SIGN WESTERN PWO KAREN TONE-5 are tone marks. (Figures 1, 2, 3, 4, and 5.) Note that the dot on ◌ၢ VOWEL SIGN WESTERN PWO KAREN UE is an integral part of the character; it is not U+1036 ◌း MYANMAR SIGN ANUSVARA. The character U+1037 ◌့ MYANMAR SIGN DOT BELOW is used with with ◌ၣ SIGN WESTERN PWO KAREN TONE-1, ◌ၥ SIGN WESTERN PWO KAREN TONE-2, and ◌ၦ SIGN WESTERN PWO KAREN TONE-3; it must follow them in the text stream—otherwise the dot would rest on the base consonant (see Figure 5).

Additions for Eastern Pwo Karen

One character used in Eastern Pwo Karen contrasts with a Burmese character: ၓ LETTER EASTERN PWO KAREN NNA contrasts with Burmese ဏ NNA. Two other characters are unique to Eastern Pwo Karen: ၺ LETTER EASTERN PWO KAREN YWA and ၻ LETTER EASTERN PWO KAREN GHWA. (Figures 6 and 7.)

Addition for Geba Karen

One character used in Geba Karen contrasts with a Burmese character: ◌ꩻ VOWEL SIGN GEBA KAREN I, used for a tense "î", contrasts with Burmese ◌ိ VOWEL SIGN I. (Figure 9.)

Additions for Kayah

One character used in Kayah contrasts with a Burmese character: ◌ꤰ VOWEL SIGN KAYAH U is used alongside Burmese ◌ု VOWEL SIGN U. Two characters are unique to Kayah: ◌꤮ VOWEL SIGN KAYAH OE and ◌꤯ VOWEL SIGN KAYAH EE. (Figure 8.)

Ordering

The unified order for the Myanmar script incorporating the characters here (and those previously accepted for encoding) is given below. Ordering is syllable-based, so this is indicative of only one level of ordering. The Shan and Palaung characters proposed in N3xxx are shown here underscored.

ka < shan-ka < kha < shan-kha < ga < gha < nga < mon-nga <
ca < shan-ca < cha < ja < jha < mon-jha < sgaw-karen-sha < nya < shan-nya < nnya <
tta < ttha < dda < ddha < nna < *eastern-pwo-karen-nna* <
ta < tha < da < dha < na < shan-na <
pa < pha < shan-pha < shan-fa < rumai-palaung-fa < ba < bha < ma <
ya < ra < la < wa < shan-tha < sha < ssa < *western-pwo-karen-tha* < sa < great-sa < ha < shan-ha <
lla < mon-bba < *eastern-pwo-karen-ywa* < *eastern-pwo-karen-gwa* <
a < shan-a < *kayah-oe* < i < ii < u < *kayah-u* < *kayah-ee* < uu <
vocalic-r < vocalic-rr < vocalic-l < 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 ◌ྐ CONSONANT SIGN MEDIAL HA. These are like ꤫ LETTER SGAW KAREN SHA which has already been accepted for encoding: the letters are ꤊ LETTER WESTERN PWO KAREN PWA, ꤋ LETTER EASTERN PWO KAREN YWA, and ꤌ LETTER EASTERN PWO KAREN GHWA. At a "Workshop on Myanmar Language Processing", held in Yangon 13-15 February 2006 (cf N3043R), this was discussed at length. All of the letters have their own place in the alphabet of the language which uses them (see Figures 1, 2, and 6). In Western Pwo Karen, for instance, ꤫ LETTER SGAW KAREN SHA and ꤊ LETTER WESTERN PWO KAREN PWA, are both treated as separate letters of the alphabet, sorted rather far from the base letters ရ *ra* and ပ *pa*, both of which do, as it happens, take genuine medials (ျ *pya*, ြ *pla*, ြ *pra*, ွ *pwa*, and ွ *rwa*) which are sorted as expected under ပ *pa* and ရ *ra*. Since the Myanmar script is to be a unified set to deal with all of these languages, we judge it best to let ◌ྐ 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, often to the right as opposed to centred, which also suggests a difference. See Figure 10 showing the distinctive LETTER PWO KAREN PWA with the tail fused toward the right of the character, quite differently to the Burmese MEDIAL HA in ဉ *nha* and မ *mha* which occurs in the same document.

Unicode Character Properties

```
1065;MYANMAR LETTER WESTERN PWO KAREN THA;Lo;0;L;;;;;N;;;;;
1066;MYANMAR LETTER WESTERN PWO KAREN PWA;Lo;0;L;;;;;N;;;;;
1067;MYANMAR VOWEL SIGN WESTERN PWO KAREN EU;Mc;0;L;;;;;N;;;;;
1068;MYANMAR VOWEL SIGN WESTERN PWO KAREN UE;Mc;0;L;;;;;N;;;;;
1069;MYANMAR SIGN WESTERN PWO KAREN TONE-1;Mc;0;L;;;;;N;;;;;
106A;MYANMAR SIGN WESTERN PWO KAREN TONE-2;Mc;0;L;;;;;N;;;;;
106B;MYANMAR SIGN WESTERN PWO KAREN TONE-3;Mc;0;L;;;;;N;;;;;
106C;MYANMAR SIGN WESTERN PWO KAREN TONE-4;Mc;0;L;;;;;N;;;;;
106D;MYANMAR SIGN WESTERN PWO KAREN TONE-5;Mc;0;L;;;;;N;;;;;
106E;MYANMAR LETTER EASTERN PWO KAREN NNA;Lo;0;L;;;;;N;;;;;
106F;MYANMAR LETTER EASTERN PWO KAREN YWA;Lo;0;L;;;;;N;;;;;
```

```
1070;MYANMAR LETTER EASTERN PWO KAREN GHWA;Lo;0;L;;;;;N;;;;;
1071;MYANMAR VOWEL SIGN GEBA KAREN I;Mn;0;NSM;;;;;N;;;;;
1072;MYANMAR VOWEL SIGN KAYAH OE;Mn;0;NSM;;;;;N;;;;;
1073;MYANMAR VOWEL SIGN KAYAH U;Mn;0;NSM;;;;;N;;;;;
1074;MYANMAR VOWEL SIGN KAYAH EE;Mn;0;NSM;;;;;N;;;;;
```

Bibliography

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

ဖုံ လိခင်းထ (Eastern Pwo Karen reader).

ပှဖျိုးရှိအလံၤဝှ့ထံး ဖိၤဃဆအဂး (Western Pwo Karen reader).

ပှဖျိုးရှိဆးချွ့ ဒဲ ပယးအခီၤပယီၤ = ပိုးကရင်-မြန်မာအဘိဓါန် = *Pwo Kayin-Myanmar Dictionary*. 1989. (Western Pwo Karen).

Acknowledgements

This project was made possible in part by a grant from the U.S. National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at UC Berkeley), and also by support from Payap University, Chiang Mai.

Figures

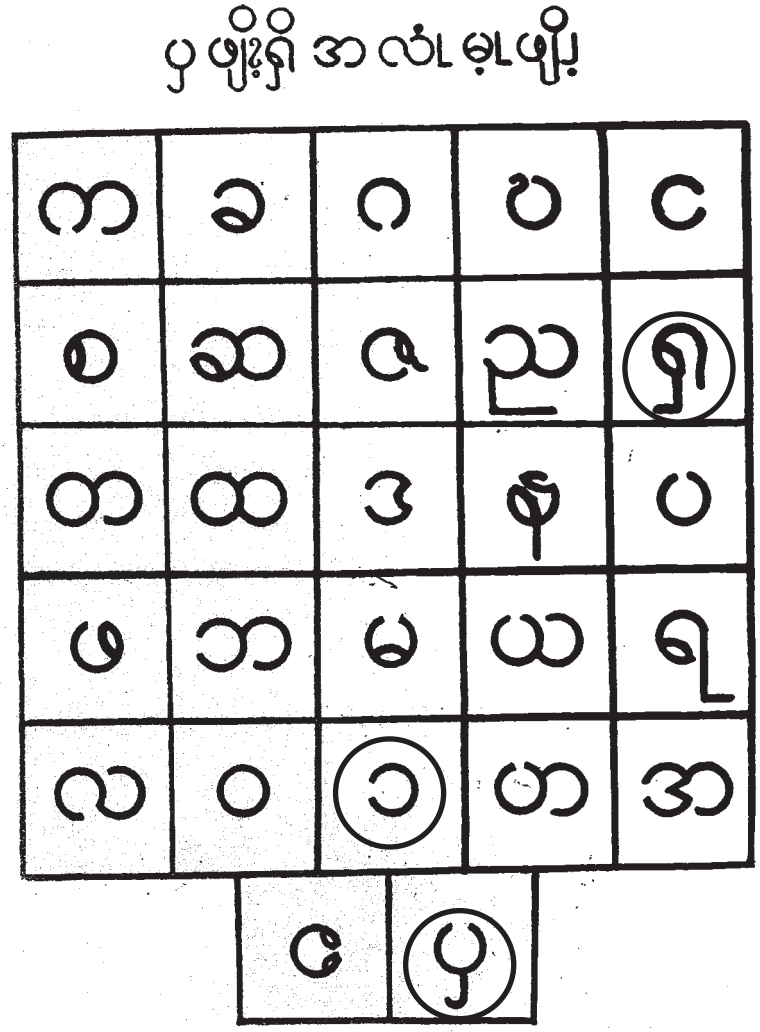


Figure 1. Sample from a reader in Western Pwo Karen, showing LETTER SGAW KAREN SHA, LETTER WESTERN PWO KAREN THA, and LETTER WESTERN PWO KAREN PWA. Note the "umbrella handle" glyph for PWA—not typical for ◌ྭ MYANMAR MEDIAL HA.

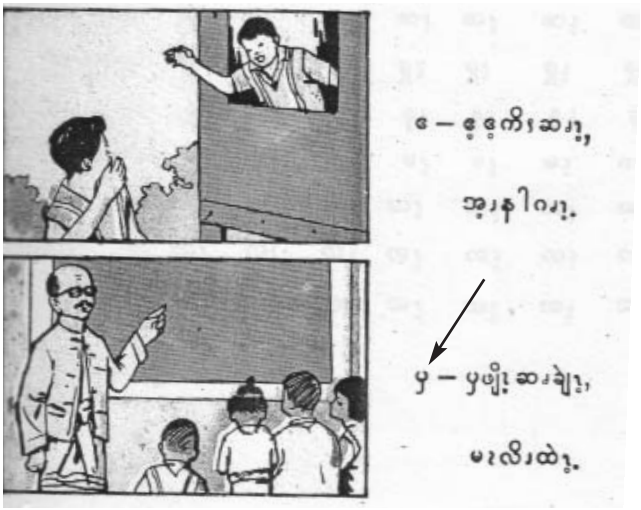


Figure 2. Sample from a reader in Western Pwo Karen, showing LETTER WESTERN PWO KAREN PWA listed as a separate letter of the alphabet: ၎ *pwa* is for ၎ဖျိုး *pwaphli* 'Pwo Karen'.

Figure 3. Sample from a reader in Western Pwo Karen, showing LETTER WESTERN PWO KAREN THA, LETTER SGAW KAREN SHA, and LETTER WESTERN PWO KAREN PWA. In PWA the tail is shifted right.

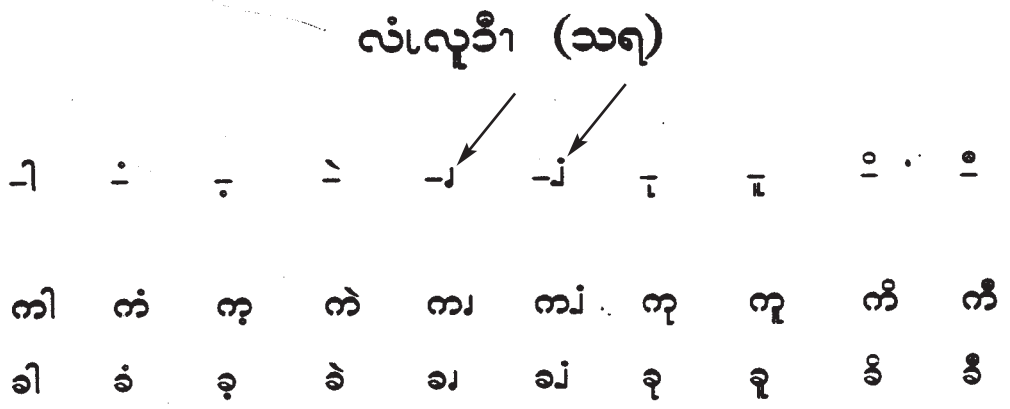


Figure 4. Sample from a reader in Western Pwo Karen, listing the vowel signs and showing VOWEL SIGN WESTERN PWO KAREN EU, and VOWEL SIGN WESTERN PWO KAREN UE.

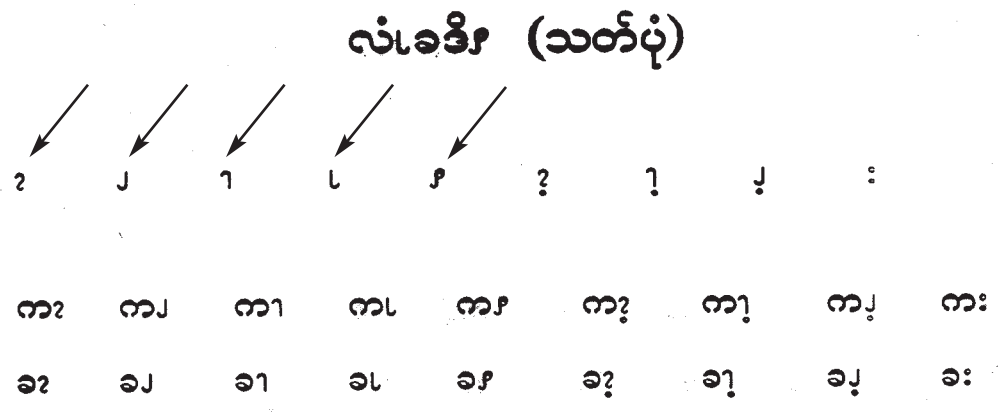


Figure 5. Sample from a reader in Western Pwo Karen, listing the tone marks and showing SIGN WESTERN PWO KAREN TONE-1, SIGN WESTERN PWO KAREN TONE-2, SIGN WESTERN PWO KAREN TONE-3, SIGN WESTERN PWO KAREN TONE-4, SIGN WESTERN PWO KAREN TONE-5.

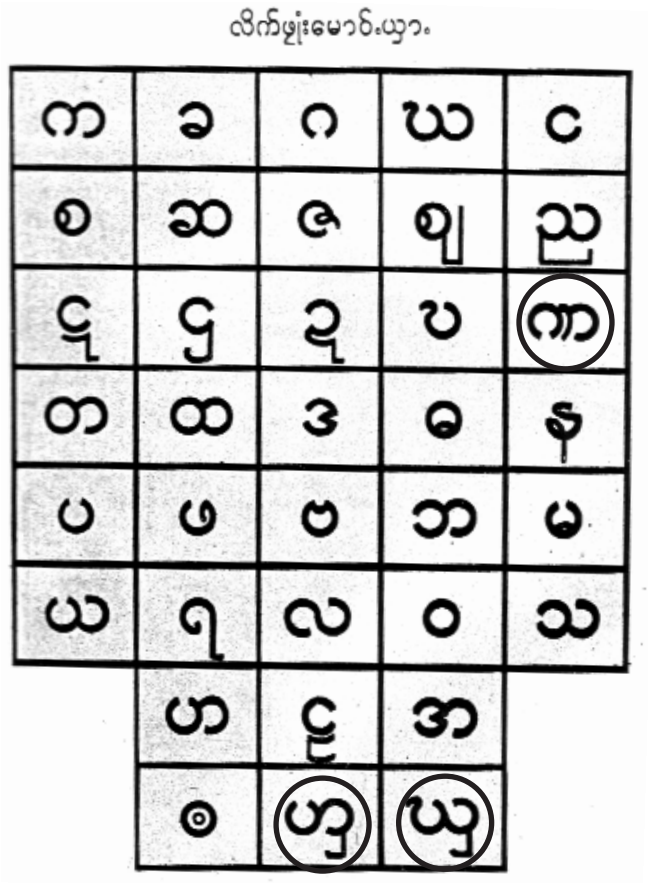


Figure 6. Sample from a reader in Eastern Pwo Karen, showing LETTER EASTERN PWO KAREN NNA, LETTER EASTERN PWO KAREN YWA, and LETTER EASTERN PWO KAREN GHWA.

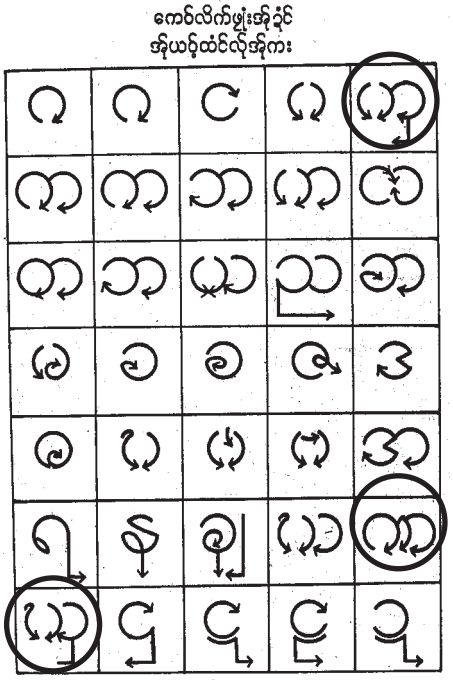
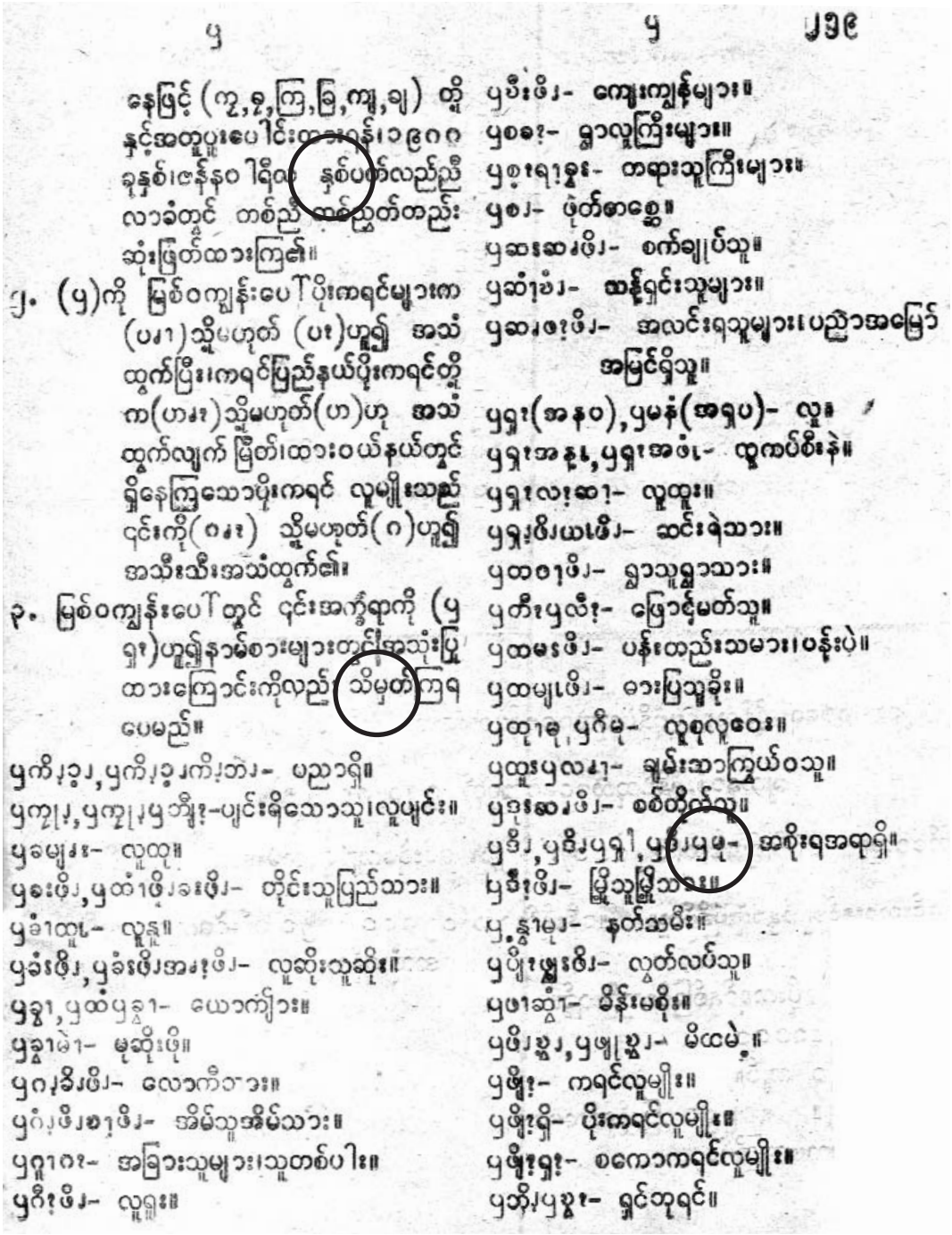


Figure 7. Sample from a reader in Eastern Pwo Karen, showing the ductus for writing Eastern Pwo Karen letters. Highlighted are LETTER EASTERN PWO KAREN YWA, LETTER EASTERN PWO KAREN NNA, and LETTER EASTERN PWO KAREN GHWA.

Figure 8. 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.

Figure 9. Sample from a Geba Karen primer published in 2005, showing VOWEL SIGN GEBA KAREN I alongside VOWEL SIGN I.



ပ္ပ ပ္ပ ၂၅၉

နေဖြင့် (က္ခ,ဂ္ဂ,ကြ,ခြ,ကျ,ချ) တို့
နှင့်အတူပူးပေါင်းထားရန်၊ ၁၉၈၈
ခုနှစ်၊ ဇန်နဝါရီလ နှစ်ပတ်လည်ညီ
လာခံတွင် တစ်ညီ ထဆ်ညီတ်တည်း
ဆုံးဖြတ်ထားကြ၏။

၅. (ပ္ပ)ကို မြစ်ဝကျွန်းပေါ်ပိုးကရင်မျိုးက
(ပဖ၊)သို့မဟုတ် (ပး)ဟူ၍ အသံ
ထွက်ပြီး၊ကရင်ပြည်နယ်ပိုးကရင်တို့
က(ဟး)သို့မဟုတ်(ဟ)ဟု အသံ
ထွက်လျက် မြေတ်၊ထားဝယ်နယ်တွင်
ရှိနေကြသောပိုးကရင် လူမျိုးသည်
ဂင်းကို(ဂး) သို့မဟုတ်(ဂ)ဟူ၍
အသီးသီးအသံထွက်၏။

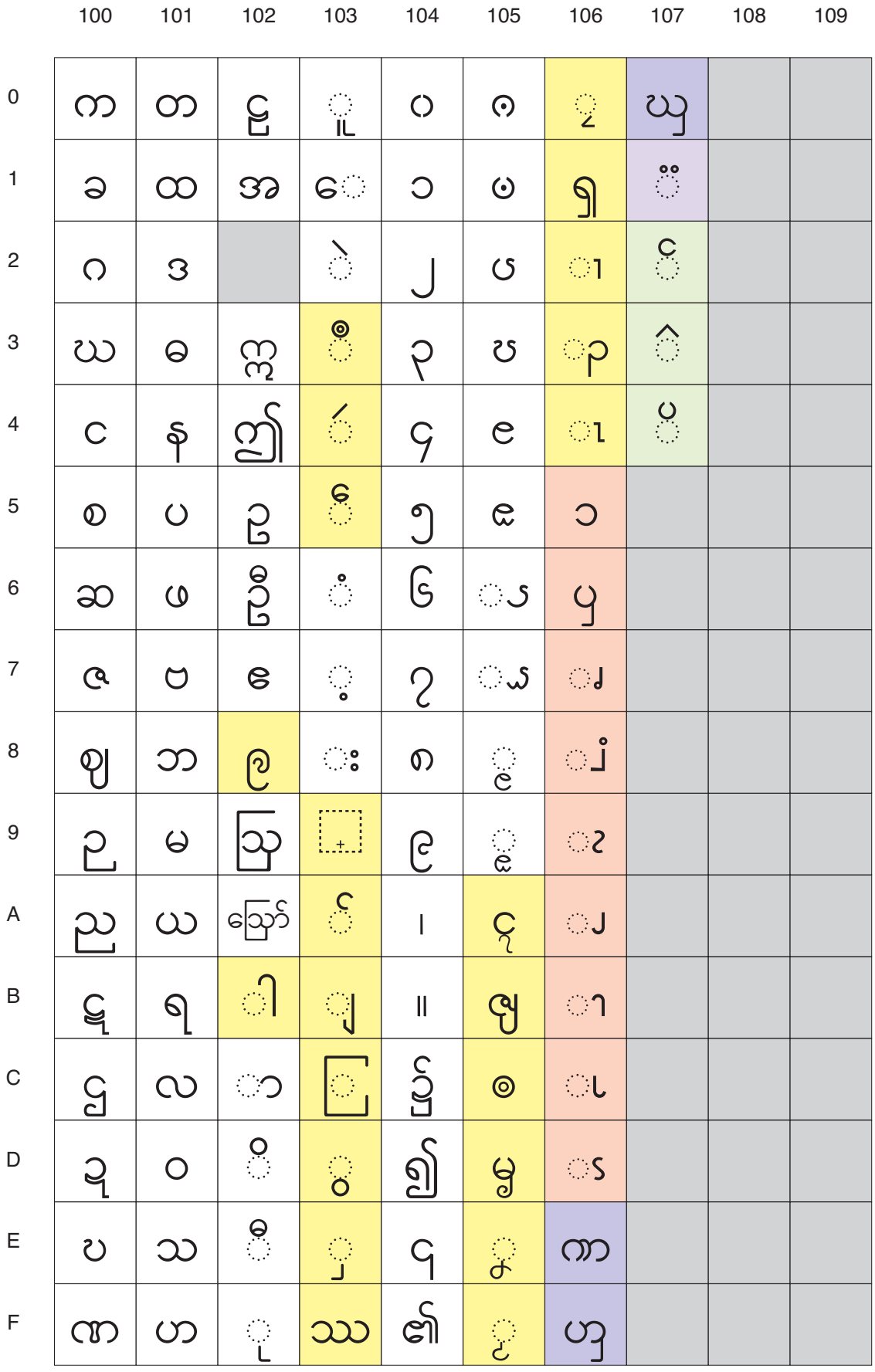
၃. မြစ်ဝကျွန်းပေါ်တွင် ဂင်းအက္ခရာကို (ပ္ပ
ရှ)ဟူ၍နာမ်စားများတွင်ပုံသုံးပြု
ထားကြောင်းကိုလည်း သိမှတ်ကြရ
ပေမည်။

ပ္ပကိ၊ခွ၊,ပ္ပကို၊ခွ၊ကို၊ဘဲ၊- မညာရှိ။
ပ္ပကူ၊,ပ္ပကူ၊ပ္ပဘျီး-ပျင်းဓိသောသူ၊လူပျင်း။
ပ္ပခမျှး- လူထု။
ပ္ပခးဖို၊,ပ္ပထံဖို၊ခးဖို၊- တိုင်းသူပြည်သား။
ပ္ပခံထူး- လူခု။
ပ္ပခံးဖို၊,ပ္ပခံးဖိုအး၊ဖို၊- လူဆိုးသူဆို။
ပ္ပခွာ,ပ္ပထံပ္ပခွာ- ယောကျ်ား။
ပ္ပခွာမဲာ- မုဆိုးဖို။
ပ္ပဂ၊ခီ၊ဖို၊- လောကီဘဝ။
ပ္ပဂံ၊ဖို၊စ၊ဖို၊- အိမ်သူအိမ်သား။
ပ္ပဂ္ဂါဂး- အခြားသူမျိုး၊သူတစ်ပါး။
ပ္ပဂီး၊ဖို၊- လူရွှေး။

ပ္ပဒီးဖို၊- ကျေးကျွန်မျိုး။
ပ္ပစခး- ရွာလူကြီးမျိုး။
ပ္ပစ့ားရး၊ခွး- တရားသူကြီးမျိုး။
ပ္ပစ၊- ဖုံတ်စာစွဲ။
ပ္ပဆးဆး၊ဖို၊- စက်ချုပ်သူ။
ပ္ပဆံ၊ခံ၊- ဆန်ရှင်းသူမျိုး။
ပ္ပဆး၊စး၊ဖို၊- အလင်းရသူမျိုး၊ပညာအမြော်
အမြင်ရှိသူ။
ပ္ပရှး(အနဝ),ပ္ပမနံ(အရှပ)- လူး
ပ္ပရှးအနှး,ပ္ပရှးအဖံး- လူကပ်စီးနဲ့။
ပ္ပရှးလး၊ဆ၊- လူထူး။
ပ္ပရှ၊ဖို၊ယးဖို၊- ဆင်းရဲသား။
ပ္ပထစ၊ဖို၊- ရွာသူရွာသား။
ပ္ပကီးပ္ပလီး၊- ဖြောင့်မတ်သူ။
ပ္ပထမးဖို၊- ပန်းတည်းသမား၊ဝန်းပဲ့။
ပ္ပထမျ၊ဖို၊- ဓားပြသူခိုး။
ပ္ပထူးမု,ပ္ပဂီမု- လူစုလူဝေး။
ပ္ပထူးပ္ပလး၊- ချမ်းသာကြွယ်ဝသူ။
ပ္ပဒုးဆး၊ဖို၊- စစ်တိုက်သူ။
ပ္ပဒီ၊,ပ္ပဒီ၊ပ္ပရှါ,ပ္ပဒီ၊ပ္ပမု- အစိုးရအထူရှိ။
ပ္ပဒီး၊ဖို၊- မြို့သူမြို့သား။
ပ္ပန္ဒါမှ၊- နတ်သမီး။
ပ္ပပိုး၊ဖျွးဖို၊- လှတ်လပ်သူ။
ပ္ပဖၢဆုံၢ- မိန်းမစိုး။
ပ္ပဖိ၊ဘွ၊,ပ္ပဖျ့ဘွ၊- မိဆမဲ့။
ပ္ပဖျိး- ကရင်လူမျိုး။
ပ္ပဖျိး၌- ပိုးကရင်လူမျိုး။
ပ္ပဖျိး၌း- စကောကရင်လူမျိုး။
ပ္ပဘိ၊ပ္ပဘွး- ရှင်ဘုရင်။

Figure 10. Sample from a Pwo Karen-Myanmar published in 1989, showing the distinctive LETTER PWO KAREN PWA with the tail fused toward the right of the character. Compare this shape with the Burmese MEDIAL HA in ⟨ နှ *nha* ⟩ and ⟨ မှ *mha* ⟩.

TABLE XX - Row 10: MYANMAR



G = 00
P = 00

TABLE XX - Row 10: MYANMAR

hex	Name
00	MYANMAR LETTER KA
01	MYANMAR LETTER KHA
02	MYANMAR LETTER GA
03	MYANMAR LETTER GHA
04	MYANMAR LETTER NGA
05	MYANMAR LETTER CA
06	MYANMAR LETTER CHA
07	MYANMAR LETTER JA
08	MYANMAR LETTER JHA
09	MYANMAR LETTER NYA
0A	MYANMAR LETTER NNYA
0B	MYANMAR LETTER TTA
0C	MYANMAR LETTER TTHA
0D	MYANMAR LETTER DDA
0E	MYANMAR LETTER DDHA
0F	MYANMAR LETTER NNA
10	MYANMAR LETTER TA
11	MYANMAR LETTER THA
12	MYANMAR LETTER DA
13	MYANMAR LETTER DHA
14	MYANMAR LETTER NA
15	MYANMAR LETTER PA
16	MYANMAR LETTER PHA
17	MYANMAR LETTER BA
18	MYANMAR LETTER BHA
19	MYANMAR LETTER MA
1A	MYANMAR LETTER YA
1B	MYANMAR LETTER RA
1C	MYANMAR LETTER LA
1D	MYANMAR LETTER WA
1E	MYANMAR LETTER SA
1F	MYANMAR LETTER HA
20	MYANMAR LETTER LLA
21	MYANMAR LETTER A
22	(This position shall not be used)
23	MYANMAR LETTER I
24	MYANMAR LETTER II
25	MYANMAR LETTER U
26	MYANMAR LETTER UU
27	MYANMAR LETTER E
28	MYANMAR LETTER MON E
29	MYANMAR LETTER O
2A	MYANMAR LETTER AU
2B	MYANMAR VOWEL SIGN TALL AA
2C	MYANMAR VOWEL SIGN AA
2D	MYANMAR VOWEL SIGN I
2E	MYANMAR VOWEL SIGN II
2F	MYANMAR VOWEL SIGN U
30	MYANMAR VOWEL SIGN UU
31	MYANMAR VOWEL SIGN E
32	MYANMAR VOWEL SIGN AI
33	MYANMAR VOWEL SIGN MON II
34	MYANMAR VOWEL SIGN MON O
35	MYANMAR VOWEL SIGN E ABOVE
36	MYANMAR SIGN ANUSVARA
37	MYANMAR SIGN DOT BELOW
38	MYANMAR SIGN VISARGA
39	MYANMAR SIGN VIRAMA
3A	MYANMAR SIGN ASAT
3B	MYANMAR CONSONANT SIGN MEDIAL YA
3C	MYANMAR CONSONANT SIGN MEDIAL RA
3D	MYANMAR CONSONANT SIGN MEDIAL WA
3E	MYANMAR CONSONANT SIGN MEDIAL HA
3F	MYANMAR LETTER GREAT SA
40	MYANMAR DIGIT ZERO
41	MYANMAR DIGIT ONE
42	MYANMAR DIGIT TWO
43	MYANMAR DIGIT THREE
44	MYANMAR DIGIT FOUR
45	MYANMAR DIGIT FIVE
46	MYANMAR DIGIT SIX
47	MYANMAR DIGIT SEVEN
48	MYANMAR DIGIT EIGHT
49	MYANMAR DIGIT NINE
4A	MYANMAR SIGN LITTLE SECTION
4B	MYANMAR SIGN SECTION
4C	MYANMAR SYMBOL LOCATIVE
4D	MYANMAR SYMBOL COMPLETED
4E	MYANMAR SYMBOL AFOREMENTIONED
4F	MYANMAR SYMBOL GENITIVE
50	MYANMAR LETTER SHA
51	MYANMAR LETTER SSA
52	MYANMAR LETTER VOCALIC R
53	MYANMAR LETTER VOCALIC RR
54	MYANMAR LETTER VOCALIC L
55	MYANMAR LETTER VOCALIC LL
56	MYANMAR VOWEL SIGN VOCALIC R
57	MYANMAR VOWEL SIGN VOCALIC RR
58	MYANMAR VOWEL SIGN VOCALIC L

hex	Name
59	MYANMAR VOWEL SIGN VOCALIC LL
5A	MYANMAR LETTER MON NGA
5B	MYANMAR LETTER MON JHA
5C	MYANMAR LETTER MON BBA
5D	MYANMAR LETTER MON BBE
5E	MYANMAR CONSONANT SIGN MON MEDIAL NA
5F	MYANMAR CONSONANT SIGN MON MEDIAL MA
60	MYANMAR CONSONANT SIGN MON MEDIAL LA
61	MYANMAR LETTER SGAW KAREN SHA
62	MYANMAR LETTER SGAW KAREN EU
63	MYANMAR SIGN SGAW KAREN HATHI
64	MYANMAR SIGN SGAW KAREN KE PHO
65	MYANMAR LETTER WESTERN PWO KAREN THA
66	MYANMAR LETTER WESTERN PWO KAREN PWA
67	MYANMAR VOWEL SIGN WESTERN PWO KAREN EU
68	MYANMAR VOWEL SIGN WESTERN PWO KAREN UE
69	MYANMAR SIGN WESTERN PWO KAREN TONE-1
6A	MYANMAR SIGN WESTERN PWO KAREN TONE-2
6B	MYANMAR SIGN WESTERN PWO KAREN TONE-3
6C	MYANMAR SIGN WESTERN PWO KAREN TONE-4
6D	MYANMAR SIGN WESTERN PWO KAREN TONE-5
6E	MYANMAR LETTER EASTERN PWO KAREN NNA
6F	MYANMAR LETTER EASTERN PWO KAREN YWA
70	MYANMAR LETTER EASTERN PWO KAREN GHWA
71	MYANMAR VOWEL SIGN GEBA KAREN I
72	MYANMAR VOWEL SIGN KAYAH OE
73	MYANMAR VOWEL SIGN KAYAH U
74	MYANMAR VOWEL SIGN KAYAH EE
75	(This position shall not be used)
76	(This position shall not be used)
77	(This position shall not be used)
78	(This position shall not be used)
79	(This position shall not be used)
7A	(This position shall not be used)
7B	(This position shall not be used)
7C	(This position shall not be used)
7D	(This position shall not be used)
7E	(This position shall not be used)
7F	(This position shall not be used)
80	(This position shall not be used)
81	(This position shall not be used)
82	(This position shall not be used)
83	(This position shall not be used)
84	(This position shall not be used)
85	(This position shall not be used)
86	(This position shall not be used)
87	(This position shall not be used)
88	(This position shall not be used)
89	(This position shall not be used)
8A	(This position shall not be used)
8B	(This position shall not be used)
8C	(This position shall not be used)
8D	(This position shall not be used)
8E	(This position shall not be used)
8F	(This position shall not be used)
90	(This position shall not be used)
91	(This position shall not be used)
92	(This position shall not be used)
93	(This position shall not be used)
94	(This position shall not be used)
95	(This position shall not be used)
96	(This position shall not be used)
97	(This position shall not be used)
98	(This position shall not be used)
99	(This position shall not be used)
9A	(This position shall not be used)
9B	(This position shall not be used)
9C	(This position shall not be used)
9D	(This position shall not be used)
9E	(This position shall not be used)
9F	(This position shall not be used)

A. Administrative

1. Title

Proposal for encoding Myanmar characters for Karen and Kayah in the UCS.

2. Requester's name

UC Berkeley Script Encoding Initiative (Universal Scripts Project); authors: Michael Everson and Martin Hosken

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

Liaison contribution.

4. Submission date

2006-09-08

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)

No.

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

16

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?

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, similar characters have been submitted before. See N2768, N3043, N3044, N3080

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), **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?

Yes.

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?

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?