

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 one additional Myanmar character for Pa’o Karen in the UCS**

**Source: Michael Everson and Martin Hosken**

**Status: Individual Contribution**

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

**Date: 2009-03-24**

The Pa’o Karen language is spoken by about 560,000 people, mostly in Myanmar. One tone mark used by them needs to be encoded in the UCS. It is the mirror image of U+1087 MYANMAR SIGN SHAN TONE-1, and the Pa’o use it alongside U+108F MYANMAR SIGN RUMAI PALAU TONE-5. The character is in current use as can be seen in the example below, from a recent Bible translation.

**Unicode Character Properties**

AA7B;MYANMAR SIGN PAO KAREN TONE;Mc;0;L;;;;;N;;;;;

# ဟွိုန်ဟောင်သတင်၊ မတ်ကျမ်၊

**ငေါင်းတယ်းအွဉ်**

မတ်ကတံမ်း ခါ၊ဒါး ဟွိုန်ဟောင်သတင်၊ နောင်၊ နိုင်း ကျမ်၊ လီ၊ကို ဗွေး ထားဝရဘုရာ၊ ကအွဉ်း သေ၊ခါ၊ ဂတိတေား အတိုင်း၊ ကအီး ကျ၊ လွဉ်ဒါး ထာ၊ ကယ်းတင်း အဗွေး နောင်၊ ထာ ဗွေးယေသျှ ချားခြောင်၊ ကျော်းညား သတင်၊ ဟောင်နောင်၊ သျှ။ ကကျော်းညား သတင်၊ နိုင်းနောင်၊ နောင်၊ ဗွေးယေသျှ ယိုကလျှးဖွာ၊ သက်မွန်သော့၊ ယုဒလို၊မျိုး ခမ်းကိုလျှမ်း၊ တကားတောင်း၊ ဒေါ၊နယ် ကျော်းညား ပါဟ်းတန်လောကသား ဖုံးနောင်၊ သျှ။

မတ်နောင်၊ စီးစဉ်း တမ်းမတ်ခါ၊ တပ်ဟောင် မွေးနေး ကျမ်၊ ကိုယိုသျှ။ ကျမ်၊ ယိုနောင်၊ တယ်း စအီစွဉ်း ဗွေး ယေသျှ လျှးဖွာ၊ လွဉ်ခရား တင်ခံး ဗတ္တိဇံး ထီမင်္ဂလား ခရားတွမ်း ဝေ၊ ဗွေး နောင်၊ ခံးစားတန်း သေ့၊ ကိုင်း၊ဖုံးသျှ။ ထျှနောင်၊ ဝေ၊ ဗွေး ကမာ၊ အမူ၊ထာ၊မာ၊ ကထွာ၊ အဒေါ၊ဟော်၊ သွဉ်သင်စွဉ်း ကလျှမ်းလာ၊မာ၊ ဟောင်မာ၊ရေ၊ဖေး ဒါးလို၊ ဆာဖုံး ကာလိလဲခမ်းကို ခရား သျှ။ ယိုထျှနောင်၊ ဟွိုန်ဟောင် သတင်၊ တမ်း မတ်ခါ၊ ဗွေးယေသျှအောင်၊ ကာလိလဲခမ်း ကျ၊လေ့၊ ယေရုသလမ်၊ ဝေ၊ ကို ခရား တုံနောင်၊ အဆိုင်၊သွတ်၊ အပတ်ကိုယ်တွင်၊ ဗွေးယေသျှခံး သီကင်း ထွန်းလောင်း ခရား တွမ်းအောင်၊ ထာ၊ သီကိုထိုင်၊ မွန်လွဉ်ချာ ခရား ဖုံး နောင်၊သျှ။

ဟွိုန်ဟောင်သတင်၊ ကျမ်၊ ယိုနောင်၊ ဒင်း ပျနယ် ဗွေး ယေသျှယိုကထာ ဆရာတန်မျှတ် အတိုင်၊ ဝေ၊ ဗွေး နောင်၊ လ၊အခွင်၊ အားဏား ကဖော်းထန်း နယ်နောင်၊ ဗွေး ထားဝရဘုရာ၊ အပညတ်ကျမ်၊ ယိုအမိပျယ်း ဖုံးတင်၊ ဟော်နယ်နောင်၊ ဗွေး ထားဝရဘုရာ၊ အခမ်းထီ တေား ခရား နောင်၊ သျှ။ ကျမ်၊ ယိုဝေ၊ ဗွေး အထာ၊ သွဉ်သင်ဖုံးနောင်၊ အခန်း၊ကဏ္ဍအောင်၊ (၅)ချက် စုဗူးခါ၊နိုင်း ဆလယိုသျှ။ နောင်၊ ဖုံးနောင်၊ (၁) တရာ၊ ကဟော်၊ ကောင်လောင်း ပတ်သံစီဒါး ခမ်းထီ ကင်ဟောင် သားဖုံး အကျင်၊ စရိတ္တ၊ တားဝွန်း ဝတ္တရာ၊ အခွင်၊အရေ၊ တွမ်းဝေ၊ ကအီး ထာလွဉ်ခရား နောင်၊ ဖော်း နယ်ဖေးသျှ။ အခန်း၊တန် (၅-၇) (၂) ငေါင်း ထွန်းဖေး တပည့်တေား တဆီနီပါ၊ ဖေး ဝေ၊ သီးကမာ၊ သားသနား ထာ၊မာ၊ (အခန်း၊ တန် ၁၀) (၃) ငေါင်းတက်တိုင်၊ ခမ်း ကင်ဟောင်ခရား (အခန်း၊တန် ၃) (၄) သွဉ်သင်ကထာ တပည့်တေား အနက်အမိပျယ်း (အခန်း၊တန် - ၁၄) (၅) သွဉ်သင်ဖေး ခေတ်ကားလ ယိုခါကအီး သွတ်၊လိုထို၊ ခရား တွမ်း ဗွေး ထားဝရဘုရာ၊ ခမ်းထီ ကအီး တယ်း ထောင်း လွဉ်ခရား (အခန်း၊တန် ၂၄-၂၅)

**Figure 1.** Sample from the book of Matthew in Pa’o Karen.

## A. Administrative

1. Title

**Proposal for encoding one additional Myanmar character for Pa'o Karen in the UCS.**

2. Requester's name

**Michael Everson and Martin Hosken**

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

**Individual contribution.**

4. Submission date

**2009-03-12**

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 Extended-A.**

2. Number of characters in proposal

**1**

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. Is a repertoire including character names provided?

**Yes.**

4b. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?

**Yes.**

4c. Are the character shapes attached in a legible form suitable for review?

**Yes.**

5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

**Michael Everson.**

5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

**Michael Everson, Fontographer.**

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

**No.**

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

**Yes.**

7. Special encoding issues: Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

**Yes.**

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.

**No.**

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?

**The publisher of the bible.**

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?

**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?

**Yes.**

10c. If YES, reference

**Like some of the Shan tone marks, this look superficially like a punctuation mark, but it is a combining character and typically has a hollow dot.**

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?