Title: Proposal to Encode the Pau Cin Hau Alphabet in ISO/IEC 10646
Source: Script Encoding Initiative (SEI)
Author: Anshuman Pandey (pandey@umich.edu)
Status: Liaison Contribution
Action: For consideration by UTC and WG2
Date: 2011-04-27

1 Introduction

This is a proposal to encode the Pau Cin Hau Alphabet in the Universal Character Set. The information presented here is based upon and supersedes that given in the following documents:

- N3781 L2/10-080 "Preliminary Proposal to Encode the Pau Cin Hau Script in ISO/IEC 10646"
- N3784R L2/10-092R "Defining Properties for Tone Marks of the Pau Cin Hau Script"
- N3865R L2/10-073R1 "Allocating the Pau Cin Hau Scripts in the Unicode Roadmap"
- N3960 L2/10-437 "Preliminary Proposal to Encode the Pau Cin Hau Alphabet in ISO/IEC 10646"

Several changes have been made since N3960, which include: revision of character names; clarification of tone mark values; addition of collation information and other character data; and the inclusion of new specimens.

The Pau Cin Hau Alphabet is related to another writing system, which is discussed in "Introducing the Logographic Script of Pau Cin Hau" (N3961 L2/10-438).

2 Background

The Pau Cin Hau Alphabet is a liturgical script of the Laipian religious tradition, which emerged in the Chin Hills region of present-day Chin State, Myanmar at the turn of the 20th century. The script is named after Pau Cin Hau (1859–1948), a Tedim Chin, who founded the Laipian tradition and developed the script in order to convey his teachings. In an account given by J. J. Bennison in the 1931 *Census of India* report for Burma, Pau Cin Hau stated that the characters of his script were revealed to him in a dream in 1902.[1]

The script was designed to represent Tedim (ISO 639: ctd), a language of the northern branch of the Kuki-Chin group of the Tibeto-Burman family, which is spoken in Chin State. Tedim is the modern name for the language previously known as Tiddim; it also refers to the Tedim dialects Kamhau (Kamhow) and Sokte. The script is known traditionally as *pau cin hau lai* "script of Pau Cin Hau" and *tual lai* "local script". The Tedim word *lai* means "writing" and *laipian* is the 'script-based religion'; Pau Cin Hau himself is referred to as *laipianpa* "creator of *laipian*".[2] The name of the script is also romanized as 'Pau Chin Hau' and is spelled in one source as 'Bow-chinhow'.[3] While the script was developed for writing Tedim, several letters and tone marks represent sounds that are not attested in Tedim, but which do exist in other Chin languages. This suggests that the alphabet may have been created as a universal script for the Chin languages.

---

[1] Bennison 1933: 217.   [2] Pau: 11.   [3] American Bible Society 1938: 82.

There are two distinct writing systems associated with Pau Cin Hau and the Laipian tradition. One is a logographic script and the other is the alphabetic system discussed here. Both are attested in manuscript and printed sources. The original script devised by Pau Cin Hau in 1902 was reformed at least twice. The logographic script is the product of the first reformation. It consists of 1,050 characters, a number that is based upon a traditional count of the characters used in a Laipian recitation text (see Figure 11). The final revision was made in 1931, when the logographic script was greatly simplified.[4] The result is the 57-character alphabet, which is proposed here for encoding. Aside from the derivation of the alphabetic script from the logographic script, these two scripts do not have any genetic relationship with any other writing system. The glyph shapes of certain letters bear resemblance to Myanmar and Latin characters; however, the correspondences are largely graphical. The similarities are likely the result of contact with Burmese communities, as well as influence from the Latin alphabet that was proliferated in the region by Western missionaries.

Pau Cin Hau's alphabet was used for writing and publishing Laipian and Christian literature. Pau Cin Hau authored a script primer with the help of Tham Cin Kham that was published in 1932 in Rangoon by Dorothy Sherrat and the Baptist Foreign Bible Society (BFBS) under the title *Kamhow-Sokte Spelling Book* (see figures 2 and 3). In 1931, the BFBS printed the "Sermon on the Mount" from the Christian text of St. Matthew in Tedim ('Kamhow-Sokte') using the alphabet (see figure 6). These documents are hand-written. It is reported that books of Laipian ritual songs were printed in the alphabet,[5] however, there is no documented evidence for this claim. The digitized font used in this proposal is the first typeface developed for the Pau Cin Hau Alphabet.

The use of the Pau Cin Hau Alphabet diminished with the decline of the Laipian tradition and the rise of Christian missionary activity. Indeed, very little information is available about the script. For instance, the record for BFBS's 'Kamhow-Sokte' version of the "Sermon on the Mount" in the catalogue of the Cambridge University Library offers a note that the text might be written in the "Pollard syllabic script". Beginning in the late 19th century, Baptist missionaries introduced Latin-based scripts for Chin languages and in many cases developed the first written forms for these languages. A Latin orthography for Tedim was introduced by Joseph H. Cope, an American Baptist missionary. It remains the regular script for Tedim. Although practice of Laipian and usage of its liturgical scripts have declined, both traditions continue to survive to some extent. It is reported that both the logographic and alphabetic systems are still in limited used by the Laipian community.[6] Information on the size of the user community was unavailable to Bennison in 1931;[7] the same is true at present. The script also enjoys some scholarly attention, as is evidenced by articles published on several websites by members of Chin-speaking communities.

3   Script Details

3.1   Script Name

The name of the script is 'Pau Cin Hau Alphabet.'

3.2   Character Repertoire

The Pau Cin Hau Alphabet has 57 characters: 21 consonant letters, 7 vowel letters, 9 final-consonant letters, and 20 tone marks. A code chart and names list is provided in Figure 1.

---

[4] Pau: 10.   [5] Banks 1967: 46.   [6] Button 2009: 23 fn. 20.   [7] Bennison 1933: 217.

3.3   Character Names

Transliterated Latin names for letters of the Pau Cin Hau Alphabet were first assigned by Dorothy Sherrat in 1932 (see Figure 2). It is likely that these names were adopted by J. J. Bennison with minor modifications (see Figure 8). Unfortunately, several of these names are idiosyncratic and do not effectively describe the characters. For instance, in Sherrat's system initial consonant letters are written with an 'A', as is the convention for Indic scripts. Such usage suggests an inherent-vowel feature that does not exist for these letters. Second, the names for some consonant letters are confusive, ie. '*tga*' for '*ca*'. Also, the Latin letters in the names for aspirated consonant letters are transposed, eg. '*hpa*' for '*pha*'; this may be influenced by old transliteration conventions for Burmese. In some names, aspiration is indicated where none exists, eg. '*hsa*' for '*sa*'. Moreover, the names for final consonants are misleading. Codas in Tedim can be only voiceless, liquid, or nasal consonants. Therefore, names such as *ab*, *ag*, *ad*, etc. contradict basic phonological properties of the letters; also, the use of Latin '*a*' in the names does not reflective any actual value and it is likely a prosthetic element intended to assist in the pronunciation of the names for final-consonant letters in isolation. Lastly, neither Sherrat or Bennison provide names for the 'tone marks', apart from Sherrat's description of the tone contours as '*sang*', '*kui*', '*phei*', '*niam*'.

New names have been adopted for the proposed characters. These names follow the conventions for alphabetic systems. Major changes include the removal of the 'A' in Sherrat's names for consonant and final-consonant letters. Names for aspirated consonant letters and final-consonant letters more accurately describe their properties. Changes to names of vowels and other consonants reflect their phonology. Incidentally, the new names are similar to modern Latin orthography for Tedim (see Zomi 2009). Each 'tone mark' has been given a descriptive name (see Section 4.5). Shown below are historical names for the letters used by Sherrat and Bennison, and the new names assigned to the letters:

	Old	New		Old	New		Old	New
Ɡ	pa	P	И	tga	C	ꔇ	ua	UA
ꔈ	ka	K	ᒪ	ta	T	Λ	ia	IA
C	la	L	Ɡ	hta	TH	ꖎ	ab	FINAL P
ꔈ	ma	M	ꔆ	na	N	Τ	ag	FINAL K
B	da	D	Z	hpa	PH	ꖜ	ad	FINAL T
U	ya	Z	?	ra	R	V	am	FINAL M
ꓵ	va, ba	V	ꖹ	fa	F	Ɛ	an	FINAL N
Ƨ	nga	NG	ꖜ	cha	CH	ꓵ	al	FINAL L
ə	ha	H	ƴ	á, à	A	P	au	FINAL W
ᴦ	ga	G	Ƨ	é, ài	E	ꖱ	ang	FINAL NG
Ҳ	hka, xa	KH	ꓵ	i	I	O	ai	FINAL Y
℮	hsa	S	Λ	aw	O			
Ɛ	ba	B	ꓷ	u	U			

3.4   Encoding Order

The encoding order is based upon the traditional arrangement of the characters in alphabet charts. The sort order differs from the encoding order (see Section 4.8).

4   Writing System

4.1   Structure

The Pau Cin Hau Alphabet is written from left to right. Vowels, consonants, and tone marks are written linearly as independent characters. The script was designed for Tedim, whose syllable canon may be described as $(C_1)V_1(V_2)(C_2)T$.[8] See Button (2011) for a detailed description of Tedim phonology.

4.2   Vowels

The nucleus ($V_1(V_2)$) is represented using the following vowels and diphthongs:[9]

୪	A	/ɐ/ (/a/)		A	O	/ɔ/ (/o/)		� UA	/ʊɑ/

୪   A   /ɐ/ (/a/)   A   O   /ɔ/ (/o/)   �   UA   /ʊɑ/

ᔐ   E   /ɛ/ (/e/)   d   U   /ʊ/ (/u/)   Λ   IA   /ɪɑ/

ꓶ   I   /ɪ/ (/i/)

Length distinctions are given in parentheses. Vowel length is indicated using 'tone marks' (see Section 4.5).

4.3   Consonants

The onset (C_1) is represented by one of the following consonants.[10]

Ꞡ   P   /p/   Ƨ   NG   /ŋ/   L   T   /t/

Ꝺ   K   /k/   ə   H   /h/   Ꞡ   TH   /tʰ/

C   L   /l/   ⲅ   G   /g/   Ꞓ   N   /n/

Ꝋ   M   /m/   Ж   KH   /x/ (/kʰ/)   Z   PH   /pʰ/

B   D   /d/   e   S   /s/   ꜟ   R   /r/

U   Z   /z/ (/j/)   Ɛ   B   /b/   h   F   /f/

Ꞑ   V   /v/ (/w/)   И   C   /tʃ/   Ԋ   CH   /tʃʰ/

Phonetic values are based upon modern Tedim. Historical values are given in parentheses. The letters R, F, CH represent phonemes not found in Tedim.

4.4   Final Consonants

The coda (C_2) is represented by one of the following final-consonant letters:[11]

Ꝉ   FINAL P   /p/   V   FINAL M   /m/   P   FINAL W   /w/

Т   FINAL K   /k/   Ꝝ   FINAL N   /n/   ꝛ   FINAL NG   /ŋ/

Ԋ   FINAL T   /t/   ∩   FINAL L   /l/   O   FINAL Y   /j/

Tedim has a word-final glottal stop, but it is represented using the glottal-stop mark described in Section 4.5.

---

[8] Thang 2001: 33.   [9] Button 2011: 15–16.   [10] Button 2011: 23.   [11] Button 2011: 19.

4.5 'Tone' Marks

The tone (T) is represented using one of the 20 'tone' marks. These marks are used for indicating vowel length, tone, and glottal stop and punctuation. Of these, 15 represent tones and 5 represent glottal stop.

ǀ	RISING TONE LONG	·	MID-LEVEL TONE
ǀ·	RISING TONE	:	GLOTTAL STOP VARIANT
ǀː	SANDHI GLOTTAL STOP	⇉	MID-LEVEL TONE LONG FINAL
ǁ	RISING TONE LONG FINAL	⇉	MID-LEVEL TONE FINAL
ǁ·	RISING TONE FINAL	↓	LOW-FALLING TONE LONG
ǁː	SANDHI GLOTTAL STOP FINAL	↓·	LOW-FALLING TONE
↑	SANDHI TONE LONG	↓ː	GLOTTAL STOP
↑·	SANDHI TONE	↓↓	LOW-FALLING TONE LONG FINAL
↑↑	SANDHI TONE LONG FINAL	↓↓·	LOW-FALLING TONE FINAL
↑↑·	SANDHI TONE FINAL	↓↓ː	GLOTTAL STOP FINAL

The traditional arrangement of tone marks (see Figure 2) may be analyzed as follows:

	BASIC			SENTENCE FINAL		
	long	regular	glottal	long	regular	glottal
rising tone (II)	ǀ	ǀ·	ǀː	ǁ	ǁ·	ǁː
sandhi tone	↑	↑·		↑↑	↑↑·	
mid-level tone (I)		·	:	⇉	⇉	
low-falling (III)	↓	↓·	↓ː	↓↓	↓↓·	↓↓ː

True tone marks are shown in black, while glottal-stop marks are in gray. True tone marks have two features: tone contour (basic or sandhi) and vowel length (regular and long). The 'sentence-final' marks are contextual variants of the corresponding 'basic' marks that are used at the end of sentences or as punctuation.

The remaining 5 marks are used for writing sentence-final glottal stop. Although orthographically paired with tone marks, they have no actual correlation with tones. As shown below, there is one basic glottal stop in Tedim, one sandhi variant, and another variant that is possibly used for representing other languages:

	BASIC	SENTENCE FINAL
	glottal	glottal
sandhi variant	ǀː	ǁː
non-Tedim variant	:	
regular	↓ː	↓↓ː

The structure of 'sentence-final' marks shows influence from Indic scripts, particularly Burmese, in which basic punctuation marks are doubled to indicate major sections, eg. ၊ U+104A MYANMAR SIGN LITTLE SECTION is doubled to produce ။ U+104B MYANMAR SIGN SECTION.

4.6 Digits

Script-specific digits are unattested. Latin digits are used.

4.7 Punctuation

Word boundaries are indicated using spaces. End of sentence is marked with final forms of tone marks. Latin punctuation is used. In some cases, sentence-final tone marks may be followed by a Latin period, etc. Other script-specific punctuation is unattested.

4.8 Collation

Consonant and vowel letters are sorted as primary:

ƴ A < Ƨ E < ꓵ I < A O < d U < � UA < ꓥ IA < Ɠ P < Ø K < C L < Ө M <
B D < U Z < ꓵ V < �printNG < ə H < ꓩ G < X KH < e S < Ɛ B < ꓲ C < L T <
Ɠ TH < ⅭⅭ N < Z PH < ʔ R < ꓺ F < ꓧ CH < ꓴ FINAL P < ꓔ FINAL K <
V FINAL M < Ꮹ FINAL N < ꓵ FINAL L < P FINAL W < �२ FINAL NG < ꓧ FINAL T <
O FINAL Y

Tone marks are sorted as secondary. The collation is based upon the ordering of tone contours for Tedim:

· MID-LEVEL TONE < ꓛ MID-LEVEL TONE FINAL < ꓛ MID-LEVEL TONE LONG FINAL <
ǂ RISING TONE < ǂ RISING TONE FINAL < ǀ RISING TONE LONG <
ǁ RISING TONE LONG FINAL < ǂ LOW-FALLING TONE < ǁ LOW-FALLING TONE FINAL <
ǂ LOW-FALLING TONE LONG < ǁ LOW-FALLING TONE LONG FINAL < ǂ SANDHI TONE <
ꓮ SANDHI TONE FINAL < ꓒ SANDHI TONE LONG < ꓮ SANDHI TONE LONG FINAL <
ǂ GLOTTAL STOP < ꓮ GLOTTAL STOP FINAL < ǂ SANDHI GLOTTAL STOP <
ꓮ SANDHI GLOTTAL STOP FINAL < ː GLOTTAL STOP VARIANT

4.9 Line Breaking

Line breaking should occur at spaces. Words are not broken at end-of-line and no hyphen is used or attested. All tone marks have the linebreaking property 'AL'. No breaking may occur between a tone mark and the character that precedes it.

4.10 Character Properties

```
11AC0;PAU CIN HAU LETTER P;Lo;0;L;;;;;N;;;;;
11AC1;PAU CIN HAU LETTER K;Lo;0;L;;;;;N;;;;;
11AC2;PAU CIN HAU LETTER L;Lo;0;L;;;;;N;;;;;
11AC3;PAU CIN HAU LETTER M;Lo;0;L;;;;;N;;;;;
11AC4;PAU CIN HAU LETTER D;Lo;0;L;;;;;N;;;;;
11AC5;PAU CIN HAU LETTER Z;Lo;0;L;;;;;N;;;;;
11AC6;PAU CIN HAU LETTER V;Lo;0;L;;;;;N;;;;;
11AC7;PAU CIN HAU LETTER NG;Lo;0;L;;;;;N;;;;;
```

```
11AC8;PAU CIN HAU LETTER H;Lo;0;L;;;;;N;;;;;
11AC9;PAU CIN HAU LETTER G;Lo;0;L;;;;;N;;;;;
11ACA;PAU CIN HAU LETTER KH;Lo;0;L;;;;;N;;;;;
11ACB;PAU CIN HAU LETTER S;Lo;0;L;;;;;N;;;;;
11ACC;PAU CIN HAU LETTER B;Lo;0;L;;;;;N;;;;;
11ACD;PAU CIN HAU LETTER C;Lo;0;L;;;;;N;;;;;
11ACE;PAU CIN HAU LETTER T;Lo;0;L;;;;;N;;;;;
11ACF;PAU CIN HAU LETTER TH;Lo;0;L;;;;;N;;;;;
11AD0;PAU CIN HAU LETTER N;Lo;0;L;;;;;N;;;;;
11AD1;PAU CIN HAU LETTER PH;Lo;0;L;;;;;N;;;;;
11AD2;PAU CIN HAU LETTER R;Lo;0;L;;;;;N;;;;;
11AD3;PAU CIN HAU LETTER F;Lo;0;L;;;;;N;;;;;
11AD4;PAU CIN HAU LETTER CH;Lo;0;L;;;;;N;;;;;
11AD5;PAU CIN HAU LETTER A;Lo;0;L;;;;;N;;;;;
11AD6;PAU CIN HAU LETTER E;Lo;0;L;;;;;N;;;;;
11AD7;PAU CIN HAU LETTER I;Lo;0;L;;;;;N;;;;;
11AD8;PAU CIN HAU LETTER O;Lo;0;L;;;;;N;;;;;
11AD9;PAU CIN HAU LETTER U;Lo;0;L;;;;;N;;;;;
11ADA;PAU CIN HAU LETTER UA;Lo;0;L;;;;;N;;;;;
11ADB;PAU CIN HAU LETTER IA;Lo;0;L;;;;;N;;;;;
11ADC;PAU CIN HAU LETTER FINAL P;Lo;0;L;;;;;N;;;;;
11ADD;PAU CIN HAU LETTER FINAL K;Lo;0;L;;;;;N;;;;;
11ADE;PAU CIN HAU LETTER FINAL T;Lo;0;L;;;;;N;;;;;
11ADF;PAU CIN HAU LETTER FINAL M;Lo;0;L;;;;;N;;;;;
11AE0;PAU CIN HAU LETTER FINAL N;Lo;0;L;;;;;N;;;;;
11AE1;PAU CIN HAU LETTER FINAL L;Lo;0;L;;;;;N;;;;;
11AE2;PAU CIN HAU LETTER FINAL W;Lo;0;L;;;;;N;;;;;
11AE3;PAU CIN HAU LETTER FINAL NG;Lo;0;L;;;;;N;;;;;
11AE4;PAU CIN HAU LETTER FINAL Y;Lo;0;L;;;;;N;;;;;
11AE5;PAU CIN HAU RISING TONE LONG;Lo;0;L;;;;;N;;;;;
11AE6;PAU CIN HAU RISING TONE;Lo;0;L;;;;;N;;;;;
11AE7;PAU CIN HAU SANDHI GLOTTAL STOP;Lo;0;L;;;;;N;;;;;
11AE8;PAU CIN HAU RISING TONE LONG FINAL;Lo;0;L;;;;;N;;;;;
11AE9;PAU CIN HAU RISING TONE FINAL;Lo;0;L;;;;;N;;;;;
11AEA;PAU CIN HAU SANDHI GLOTTAL STOP FINAL;Lo;0;L;;;;;N;;;;;
11AEB;PAU CIN HAU SANDHI TONE LONG;Lo;0;L;;;;;N;;;;;
11AEC;PAU CIN HAU SANDHI TONE;Lo;0;L;;;;;N;;;;;
11AED;PAU CIN HAU SANDHI TONE LONG FINAL;Lo;0;L;;;;;N;;;;;
11AEE;PAU CIN HAU SANDHI TONE FINAL;Lo;0;L;;;;;N;;;;;
11AEF;PAU CIN HAU MID-LEVEL TONE;Lo;0;L;;;;;N;;;;;
11AF0;PAU CIN HAU GLOTTAL STOP VARIANT;Lo;0;L;;;;;N;;;;;
11AF1;PAU CIN HAU MID-LEVEL TONE LONG FINAL;Lo;0;L;;;;;N;;;;;
11AF2;PAU CIN HAU MID-LEVEL TONE FINAL;Lo;0;L;;;;;N;;;;;
11AF3;PAU CIN HAU LOW-FALLING TONE LONG;Lo;0;L;;;;;N;;;;;
11AF4;PAU CIN HAU LOW-FALLING TONE;Lo;0;L;;;;;N;;;;;
11AF5;PAU CIN HAU GLOTTAL STOP;Lo;0;L;;;;;N;;;;;
11AF6;PAU CIN HAU LOW-FALLING TONE LONG FINAL;Lo;0;L;;;;;N;;;;;
11AF7;PAU CIN HAU LOW-FALLING TONE FINAL;Lo;0;L;;;;;N;;;;;
11AF8;PAU CIN HAU GLOTTAL STOP FINAL;Lo;0;L;;;;;N;;;;;
```

4.11   'Confusable' Characters

Some characters of the Pau Cin Hau Alphabet resemble those already encoded in the UCS:

```
PAU CIN HAU LETTER K                    ; 1005 MYANMAR LETTER CA
PAU CIN HAU LETTER L                    ; 0043 LATIN CAPITAL LETTER C
PAU CIN HAU LETTER L                    ; 1004 MYANMAR LETTER NGA
PAU CIN HAU LETTER M                    ; 1013 MYANMAR LETTER DHA
PAU CIN HAU LETTER D                    ; 0042 LATIN CAPITAL LETTER B
PAU CIN HAU LETTER Z                    ; 0055 LATIN CAPITAL LETTER U
```

```
PAU CIN HAU LETTER H                   ; 1001 MYANMAR LETTER KHA
PAU CIN HAU LETTER S                   ; 1054 MYANMAR LETTER VOCALIC L
PAU CIN HAU LETTER B                   ; 0190 LATIN LETTER OPEN E
PAU CIN HAU LETTER T                   ; 004C LATIN CAPITAL LETTER L
PAU CIN HAU LETTER TH                  ; 1027 MYANMAR LETTER E
PAU CIN HAU LETTER PH                  ; 005A LATIN CAPITAL LETTER Z
PAU CIN HAU LETTER R                   ; 0294 LATIN LETTER GLOTTAL STOP
PAU CIN HAU LETTER F                   ; 0068 LATIN SMALL LETTER H
PAU CIN HAU LETTER O                   ; 0041 LATIN CAPITAL LETTER A
PAU CIN HAU LETTER FINAL P             ; 1016 MYANMAR LETTER PHA
PAU CIN HAU LETTER FINAL K             ; 0054 LATIN CAPITAL LETTER T
PAU CIN HAU LETTER FINAL M             ; 0056 LATIN CAPITAL LETTER V
PAU CIN HAU LETTER FINAL Y             ; 004F LATIN CAPITAL LETTER O
PAU CIN HAU LETTER FINAL Y             ; 101D MYANMAR LETTER WA
PAU CIN HAU RISING TONE LONG           ; 104A MYANMAR SIGN LITTLE SECTION
PAU CIN HAU RISING TONE LONG FINAL ; 104B MYANMAR SIGN SECTION
PAU CIN HAU MID-LEVEL TONE             ; 00B7 MIDDLE DOT
PAU CIN HAU GLOTTAL STOP VARIANT    ; 003A COLON
```

5   References

American Bible Society. 1938. *The Book of a Thousand Tongues*. Edited by Eric M. North. New York and London: Harper & Brothers.

British and Foreign Bible Society. 1931. ꤢꤟꤤ ꤜꤘꤢꤚ ꤛꤢꤛ ꤚꤣꤟꤤ ꤜꤢꤞ ["The Sermon on the Mount" (Kamhow-Sokte)]. Rangoon.

Banks, E. Pendleton. 1967. "Pau Cin Hau: A Case of Religious Innovation among the Northern Chin." In *American Historical Anthropology: Essays in Honor of Leslie Spier*. London and Amsterdam: Southern Illinois University Press.

Bennison, J. J. 1933. "The Pau Cin Hau Movement in the Chin Hills." *Census of India, 1931*, vol. XI, Burma, pt. 1, ch. XI. pp. 217–218.

Button, Christopher T. J. 2009. "A Reconstruction of Proto Northern Chin in Old Burmese and Old Chinese Perspective". Ph.D. dissertation. School of Oriental and African Studies, University of London. `http://victoria.linguistlist.org/~lapolla/downloads/Button-Proto-Northern-Chin.pdf`

———. 2011. *Proto Northern Chin*, Volume 1: An Old Burmese and Old Chinese Perspective. STEDT Monograph Series #10. Berkeley: Sino-Tibetan Etymological Dictionary and Thesaurus Project.

Cambridge University Library. "Bible. N.T. Matthew. Chin, Tiddim. Selections. 1931." `http://ul-newton.lib.cam.ac.uk/vwebv/holdingsInfo?searchId=3581&recCount=25&recPointer=4&bibId=4376880`

Go, Khup Za. 2008. *Zo Chronicles: A Documentary Study of History and Culture of the Kuki-Chin-Lushai Tribe*. New Delhi: Mittal Publications.

Pandey, Anshuman. 2010a. "Preliminary Proposal to Encode the Pau Cin Hau Script in ISO/IEC 10646" (N3781 L2/10-080), February 28, 2010. `http://std.dkuug.dk/JTC1/SC2/WG2/docs/n3781.pdf`

———. 2010b. "Defining Properties for Tone Marks of the Pau Cin Hau Script" (N3784R L2/10-092R), May 22, 2010. `http://std.dkuug.dk/JTC1/SC2/WG2/docs/n3784.pdf`

———. 2010c. "Allocating the Pau Cin Hau Scripts in the Unicode Roadmap" (N3865R L2/10-073R1), July 26, 2010. `http://std.dkuug.dk/JTC1/SC2/WG2/docs/n3865.pdf`

———. 2010d. "Preliminary Proposal to Encode the Pau Cin Hau Alphabet in ISO/IEC 10646" (N3960 L2/10-437), October 27, 2010. `http://std.dkuug.dk/JTC1/SC2/WG2/docs/n3960.pdf`

———. 2010e. "Introducing the Logographic Script of Pau Cin Hau" (N3961 L2/10-438), October 27, 2010. `http://std.dkuug.dk/JTC1/SC2/WG2/docs/n3961.pdf`

Pau Chin Hau and Tham Chin Kham. 1932. ꗀꗥꖺ ꗕꘐꗪꕴ ꘖꗖ‖ [Spelling Book in Kamhow-Sokte]. Toungoo: P. R. G. Press.

Pau, Pum Khan. "Religious conversions reconsidered: Christian missions and indigenous movement among the Zo people in colonial Chin Hills." `http://www.scribd.com/doc/25568052/Religious-Conversion-Reconsidered`

"Pu Pau Cin Hau Lai." Zomi Online Library, Item #200. `http://zomilibrary.com/main/items/show/200`

Thang, Khoi Lam. 2001. "A Phonological Reconstruction of Proto-Chin". M.A. thesis. Payap University, Chiang Mai, Thailand. `http://ic.payap.ac.th/graduate/linguistics/theses/Khoi_Lam_Thang_Thesis.pdf`

"Thu Tatuam." 2008. `http://thutatuam.multiply.com/`

Zomi, Taang. 2009. "Zokam and Kawlkam Consonant Phonemes". Zomi Online Library, Item #27., `http://zomilibrary.com/main/items/show/273`

6   Acknowledgments

	11AC	11AD	11AE	11AF
0	ⴳ 11AC0	ⵉ 11AD0	ⴳ 11AE0	: 11AF0
1	ⴰ 11AC1	Z 11AD1	∩ 11AE1	⇉ 11AF1
2	C 11AC2	? 11AD2	P 11AE2	⇉ 11AF2
3	ⴱ 11AC3	h 11AD3	ⴲ 11AE3	↓ 11AF3
4	B 11AC4	Ⱨ 11AD4	O 11AE4	↓ 11AF4
5	U 11AC5	ⵖ 11AD5	I 11AE5	↓ 11AF5
6	∩ 11AC6	ⴹ 11AD6	I· 11AE6	ⱖ 11AF6
7	ⴹ 11AC7	ⴖ 11AD7	I: 11AE7	ⱖ 11AF7
8	ə 11AC8	A 11AD8	‖ 11AE8	ⱖ: 11AF8
9	ⵕ 11AC9	d 11AD9	‖· 11AE9	
A	Ⅹ 11ACA	ⴺ 11ADA	‖: 11AEA	
B	e 11ACB	Λ 11ADB	↑ 11AEB	
C	Ɛ 11ACC	ⴚ 11ADC	↑: 11AEC	
D	И 11ACD	T 11ADD	⋀ 11AED	
E	L 11ACE	Ⱨ 11ADE	⋀: 11AEE	
F	ⴳ 11ACF	V 11ADF	· 11AEF	

Consonants

11AC0 ⴳ PAU CIN HAU LETTER P
11AC1 ⴰ PAU CIN HAU LETTER K
11AC2 C PAU CIN HAU LETTER L
11AC3 ⴱ PAU CIN HAU LETTER M
11AC4 B PAU CIN HAU LETTER D
11AC5 U PAU CIN HAU LETTER Z
11AC6 ∩ PAU CIN HAU LETTER V
11AC7 ⴹ PAU CIN HAU LETTER NG
11AC8 ə PAU CIN HAU LETTER H
11AC9 ⵕ PAU CIN HAU LETTER G
11ACA Ⅹ PAU CIN HAU LETTER KH
11ACB e PAU CIN HAU LETTER S
11ACC Ɛ PAU CIN HAU LETTER B
11ACD И PAU CIN HAU LETTER C
11ACE L PAU CIN HAU LETTER T
11ACF ⴳ PAU CIN HAU LETTER TH
11AD0 ⵉ PAU CIN HAU LETTER N
11AD1 Z PAU CIN HAU LETTER PH
11AD2 ? PAU CIN HAU LETTER R
11AD3 h PAU CIN HAU LETTER F
11AD4 Ⱨ PAU CIN HAU LETTER CH

Vowels

11AD5 ⵖ PAU CIN HAU LETTER A
11AD6 ⴹ PAU CIN HAU LETTER E
11AD7 ⴖ PAU CIN HAU LETTER I
11AD8 A PAU CIN HAU LETTER O
11AD9 d PAU CIN HAU LETTER U
11ADA ⴺ PAU CIN HAU LETTER UA
11ADB Λ PAU CIN HAU LETTER IA

Final consonants

11ADC ⴚ PAU CIN HAU LETTER FINAL P
11ADD T PAU CIN HAU LETTER FINAL K
11ADE Ⱨ PAU CIN HAU LETTER FINAL T
11ADF V PAU CIN HAU LETTER FINAL M
11AE0 ⴳ PAU CIN HAU LETTER FINAL N
11AE1 ∩ PAU CIN HAU LETTER FINAL L
11AE2 P PAU CIN HAU LETTER FINAL W
11AE3 ⴲ PAU CIN HAU LETTER FINAL NG
11AE4 O PAU CIN HAU LETTER FINAL Y

Tone marks

11AE5 I PAU CIN HAU RISING TONE LONG
11AE6 I· PAU CIN HAU RISING TONE
11AE7 I: PAU CIN HAU SANDHI GLOTTAL STOP
11AE8 ‖ PAU CIN HAU RISING TONE LONG FINAL
11AE9 ‖· PAU CIN HAU RISING TONE FINAL
11AEA ‖: PAU CIN HAU SANDHI GLOTTAL STOP FINAL
11AEB ↑ PAU CIN HAU SANDHI TONE LONG
11AEC ↑: PAU CIN HAU SANDHI TONE
11AED ⋀ PAU CIN HAU SANDHI TONE LONG FINAL
11AEE ⋀: PAU CIN HAU SANDHI TONE FINAL
11AEF · PAU CIN HAU MID-LEVEL TONE
11AF0 : PAU CIN HAU GLOTTAL STOP VARIANT
11AF1 ⇉ PAU CIN HAU MID-LEVEL TONE LONG FINAL
11AF2 ⇉ PAU CIN HAU MID-LEVEL TONE FINAL
11AF3 ↓ PAU CIN HAU LOW-FALLING TONE LONG
11AF4 ↓ PAU CIN HAU LOW-FALLING TONE
11AF5 ↓ PAU CIN HAU GLOTTAL STOP
11AF6 ⱖ PAU CIN HAU LOW-FALLING TONE LONG FINAL
11AF7 ⱖ PAU CIN HAU LOW-FALLING TONE FINAL
11AF8 ⱖ: PAU CIN HAU GLOTTAL STOP FINAL

Figure 1: Proposed code chart and names list for the Pau Cin Hau Alphabet.
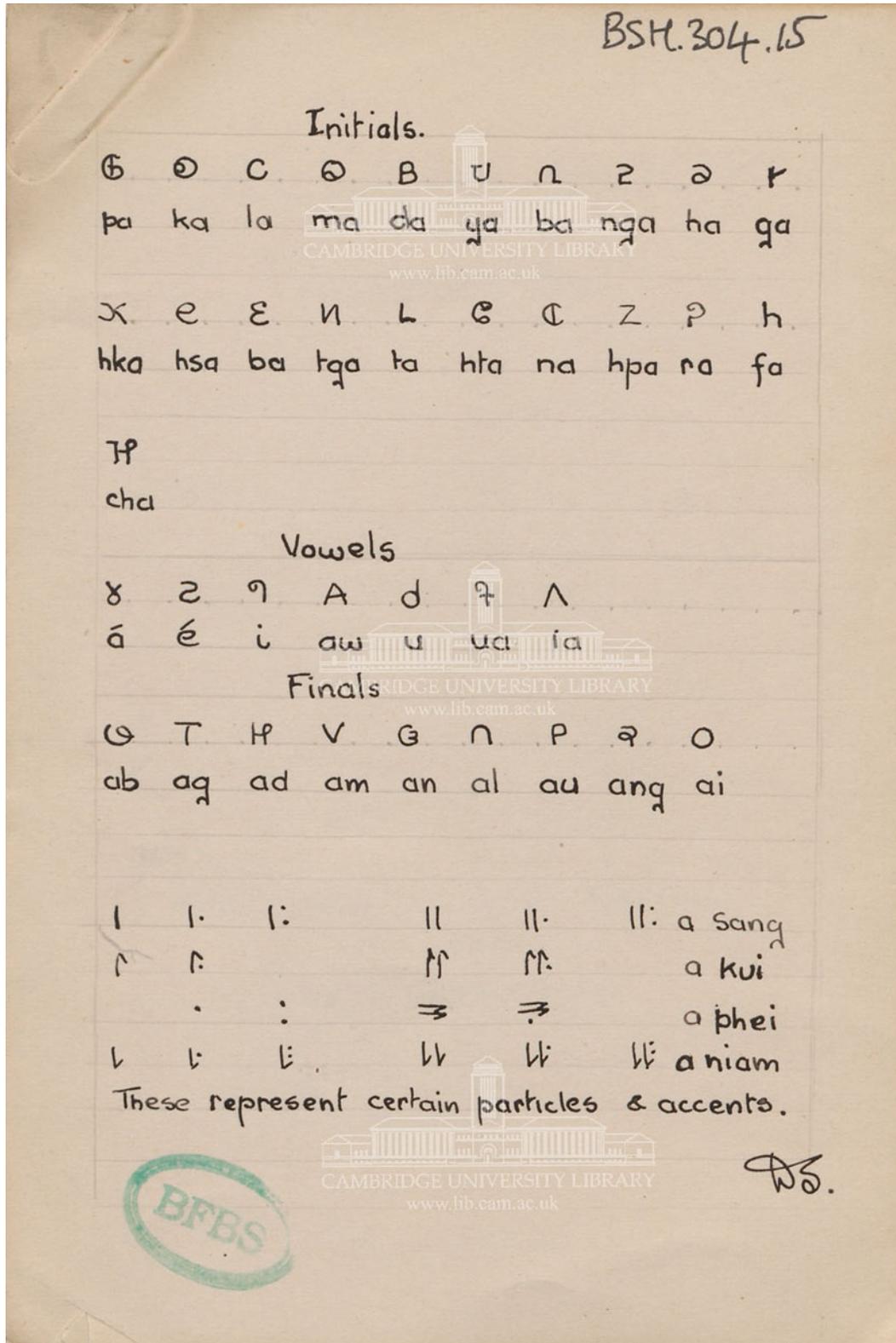
Figure 2: A plate showing the letters of the alphabet with Latin names given by Dorothy Sherrat (from Pau Cin Hau 1932). Tedim names for the tone mark groups are also given.

Figure 3: A page from the 'Kamhow-Sokte' primer of Pau Cin Hau (1932: 1). Shown are the letters and tone marks of the alphabet in traditional order. Also shown are consonant-vowel combinations.

Figure 4: A page from Pau Cin Hau's primer showing tone patterns (1932: 8).

Figure 5: A page from Pau Cin Hau's primer showing the alphabet used in running text (1932: 9).

Figure 6: Page of the "Sermon on the Mount" in the Pau Cin Hau alphabet (from British and Foreign Bible Society 1931: 1).

5.

1. Mi hĭng ă tăm pi ă' mu ciăng ĭn mual tung a ă' ka to a ă' tu ciăng ĭn ă' nung jui tĕ ăm ma kiăng a hăwng pai a 2 ă' kăm ka ă' thu hĭl ĭn a ;

2 Mi hĭng ă lung ki niĕm sak thu pha ă' nga dĭng hi. Băng dĭng hiăm i ci lĕ van tung găm ăm ma tĕ găm ă' hi dĭng hi. 4 Mi hĭng ă' da tĕ ĭn thu pha ă' nga dĭng hi. Băng dĭng hiăm i ci lĕ ăm ma tĕ ĭn lung dăm na ă' nga dĭng hi. 5 Mi hĭng ă' lung nĕm tĕ ĭn thu pha ă' nga dĭng hi. Băng dĭng hiăm i ci lĕ ăm ma tĕ ĭn le tung găm ă' lua dĭng hi. 6 Mi hĭng thu man thu tăk du a ă' kiĕl tĕ ĭn thu pha ă' nga dĭng hi. Băng dĭng hiăm i ci lĕ ăm ma tĕ ĭn ă' kham dĭng hi. 7 Mi dăng ă' hĕ pi the tĕ ĭn thu pha ă' nga dĭng hi. Băng dĭng hiăm i ci lĕ ăm ma tĕ in hĕ pi na ă' nga dĭng hi. 8 Mi hĭng ă' lung shiăng tho tĕ thu pha ă' nga dĭng hi. Băng dĭng hiăm i ci lĕ ăm ma tĕ in Pashiăn ă' mu dĭng hi. 9 Mi hĭng ki lĕm sak the tĕ ĭn thu pha ă' nga dĭng hi. Băng dĭng hiăm i ci lĕ Pashiăn ta pa ki ci na ă' nga dĭng hi. 10 Thu man thu tăk shung a mi ă' bawl shiă na ă' thuawk tĕ ĭn thu pha ă' nga dĭng hi. Băng dĭng hiăm i ci le van tung găm ăm ma tĕ găm ă' hi dĭng hi. 11 Ke hang a mi ĭn hawng gĕn shiă, hăwng bawl shiă, ă' hi lo lăm hăwng maw sak ciăng ĭn thu pha n' nga dĭng hi. 12 Lung dăm ŭn, nuăm ŭn. Băng dĭng hiăm n' ci lĕ van tung nawp tuăm liĕn n' nga dĭng hi. N' mai piăng sa Pashiăn kăm sang tĕ ă' tu a băng ma ĭn mi ă' bawl shiă hi.

13 No ma tĕ ĭn le tung ci n' hi. Ci ĭn ă' ăl măng lĕ băng ci ăl kik dĭng hiăm. Dai puă ă' pai khiă a mi khĕ taw shik gawp shĭm lo ă' dăng măn dĭng ĭn ă' ŭm ke. 14 No tĕ ĭn le tung khuaw vak n' hi. Mual dăwn a ă' ŭm khuaw ĭn ă' ki liă the ke hi. 15 Mi hĭng me vak ki dĕ ciăng ĭn a sĕng ki khu ke hi. Me vak koi na ă' koi a ĭn shung mi hĭng khĕm piu a dĭng ă' tang dĭng hi. 16 Tu a băng mă n' găm ta ă' pha mi hĭng ă' mu ciăng ĭn vang tung ă' ŭm
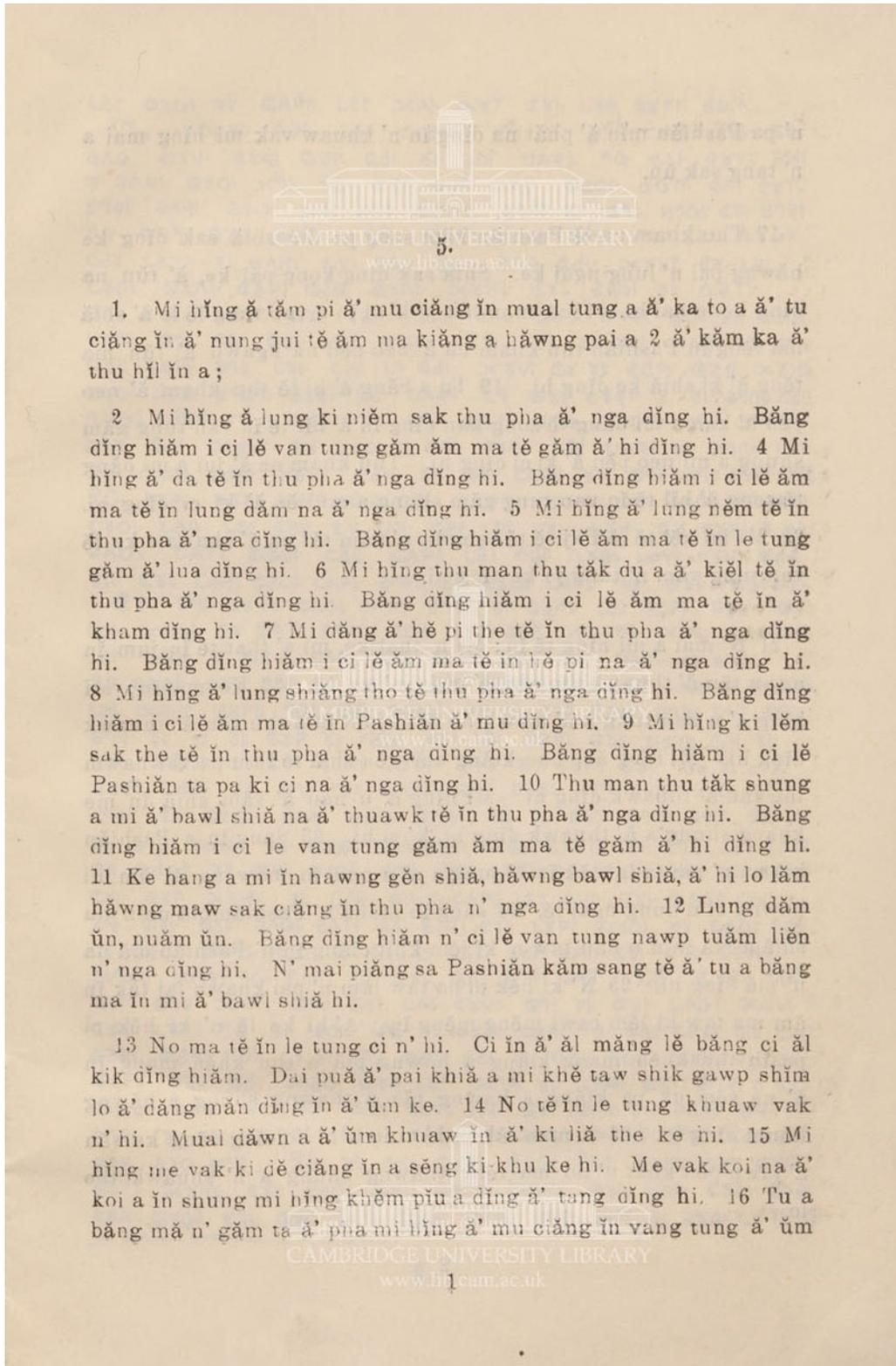
Figure 7: Page of the "Sermon on the Mount" showing Latin transliteration of the Tedim text (from British and Foreign Bible Society 1931: 2).

16

Figure 8: The Pau Cin Hau Alphabet (from Bennison (1931), reprinted in Go (2008: 94)).

Figure 9: Characters of the Pau Cin Hau Alphabet (from "Pu Pau Cin Hau Lai": 4).

Figure 10: A painting of Pau Cin Hau (from "Thu Tatuam" 2008). The inscription at the bottom of the image is the name 'Pau Cin Hau' in the logographic script.

Figure 11: A page from a Laipian recitation text written in the logographic script of Pau Cin Hau. Courtesy of Christopher Button.

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]
Please fill all the sections A, B and C below.
Please read Principles and Procedures Document (P & P) from http://www.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://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html**.**
See also http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps*.**

A. Administrative

1. **Title:** *Proposal to Encode the Pau Cin Hau Alphabet in ISO/IEC 10646*
2. Requester's name: *Script Encoding Initiative (SEI) / Anshuman Pandey <pandey@umich.edu>*
3. Requester type (Member body/Liaison/Individual contribution): *Liaison contribution*
4. Submission date: *2011-04-27*
5. Requester's reference (if applicable):
6. Choose one of the following:
   This is a complete proposal: *Yes*
   (or) More information will be provided later:

B. Technical – General

1. Choose one of the following:
   a. This proposal is for a new script (set of characters): *Yes*
      Proposed name of script: *Pau Cin Hau Alphabet*
   b. The proposal is for addition of character(s) to an existing block:
      Name of the existing block:
2. Number of characters in proposal: *57*
3. Proposed category (select one from below - see section 2.2 of P&P document):
   A-Contemporary ____ B.1-Specialized (small collection) __X__ 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 ____
4. Is a repertoire including character names provided? *Yes*
   a. If YES, are the names in accordance with the "character naming guidelines"
      in Annex L of P&P document? *Yes*
   b. Are the character shapes attached in a legible form suitable for review? *Yes*
5. Fonts related:
   a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?
      Anshuman Pandey, author of the proposal
   b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
      Anshuman Pandey <pandey@umich.edu>
6. References:
   a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? *Yes*
   b. 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,*
   see proposal for additional details

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 http://www.unicode.org/Public/UNIDATA/UCD.html and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? *No*

   If YES explain

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.)? *Yes*

   If YES, with whom? *Dr. Christopher Button <chris.button@hotmail.com>*

   If YES, available relevant documents: *See citations in text of proposal*

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

   Reference: *See text of proposal*

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

   Reference: *See text of proposal*

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

   If YES, where? Reference: *Limited use by Laipian community in Chin State, Myanmar*

6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely
   in the BMP? *No*

   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)? *Yes*

8. Can any of the proposed characters be considered a presentation form of an existing
   character or character sequence? *No*

   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? *No*

   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 an existing character? *No*

    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? *No*

    If YES, is a rationale for such use provided?

    If YES, reference:

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

    If YES, reference:

12. Does the proposal contain characters with any special properties such as
    control function or similar semantics? *No*

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

13. Does the proposal contain any Ideographic compatibility character(s)? *No*

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

    If YES, reference: