ISO/IEC JTC1/SC2/WG2 N3781 L2/10-080 2010-02-28

Title: Preliminary Proposal to Encode the Pau Cin Hau Script in ISO/IEC 10646

Author: Anshuman Pandey (pandey@umich.edu)

Status: Individual Contribution **Action:** For consideration by UTC

Date: 2010-02-28

1 Introduction

The purpose of this document is to bring the Pau Cin Hau writing system to the attention of the Unicode Technical Committee. Very little information about this script is available in English, but there appears to be a good number of materials written in Chin languages. This document provides a historical background of Pau Cin Hau, briefly describes its structure, presents tentative character names and glyph shapes, and enumerates several encoding issues. Research on the script is ongoing and the present author is in contact with specialists of Chin languages and culture. The information given here is subject to change. It is requested that any information regarding Pau Cin Hau be submitted to the author at the email address given above.

2 Background

The Pau Cin Hau script is the ecclesiastical script of the Laipian religious tradition, which developed in the Chin Hills region at the turn of the 20th century. The script is named after Pau Cin Hau (1859–1948), a 'Sukte' (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 Pau Cin Hau revised his script twice and developed the third and final form in 1931.

The script was designed to represent Tedim ([ctd]), a northern language of the Kuki-Chin group of the Tibeto-Burman family, which is spoken in Chin State, Myanmar. It 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 Pau Cin Hau himself is referred to as *laipianpa* "script creator". The name is also romanized as 'Pau Chin Hau' and is known in one source as 'Bow-chinhow'.

There appear to be logographic and alphabetic forms of Pau Cin Hau. The logographic form is believed to be the original script revealed to Pau Cin Hau (see Figure 6). It is reported that the logographic script consisted of 1,050 characters, which were reduced to a repertoire consisting of 57 characters.⁴ This 57-character repertoire belongs to the alphabetic script described in this document. The transformation of the logographic script into an alphabetic form was the final revision made by Pau Cin Hau in 1931. Some characters of both the logographic and alphabetic scripts resemble those found in Burmese, Latin, and other writing systems, but these occurances are coincidental. Neither form has a genetic relationship with any other script.

The alphabetic script formalized in 1931 was used for writing and printing Laipian and Christian literature. A primer for the script was printed for the purpose of teaching the script (see Figure 3). Books of Laipian ritual songs were written in the script and it is believed that some of these were also printed.⁵ In 1931, the Baptist and Foreign Bible Society printed the "Sermon on the Mount" from the book of St. Matthew in

¹ Bennison 1933: 217. ² Pau: 11. ³ American Bible Society 1938: 82. ⁴ Pau: 10. ⁵ Banks 1967: 46.

the 'Kamhow' (now known as Tedim) dialect in the Pau Cin Hau script (see Figure 4). At least one metal font was developed in order to print the materials of the Baptist and Foreign Bible Society. The script charts shown in Figure 2 and Figure 3 appear to be printed using different fonts, but there is insufficient information to ascertain the accuracy of this assumption.

The use of the Pau Cin Hau script diminished with the decline of the Laipian tradition and the rise of Christian missionary activity. Baptist missionaries introduced Latin-based orthographies for languages of the Chin Hills and in many cases developed the first orthographies for these languages. A Latin-based script for Tedim was introduced by J. H. Cope, an American Baptist missionary. This Latin alphabet is the regular script for Tedim and has replaced Pau Cin Hau. Although practice of Laipian and usage of the Pau Cin Hau script have declined, both still survive to an extent. It is reported that both the logographic and syllabic systems are still used in a very limited fashion by the Laipian religious community. Information on the size of the user community was unavailable to Bennison in 1931; the same is true at present. In addition to the religious uses of Pau Cin Hau, the script appears to enjoy scholarly attention, as is evidenced by articles published online by members of Chin-speaking communities and in websites dedicated to the script.

3 Character Repertoire

The alphabetic Pau Cin Hau script has a repertoire of 57 characters, which consists of 21 consonant letters, 7 vowel letters, 9 coda letters, and 20 tone marks. A preliminary code chart and nameslist is provided in Figure 1.

Character names are tentatively assigned according to those given in the script charts. The only exceptions are the names of Y AA and Z AAI, which are shown in the charts as 'á' and 'ái', respectively. Names for the tone marks have been assigned serially based upon the order in which they appear in the charts.

The character names are problematic; please see section 7 for further details.

4 Structure

Pau Cin Hau is written from left to right. It is said that the script was designed to represent the Tedim language, whose syllable canon may be described as $(C_1)V_1(V_2)(C_2)T$. The presence of consonant letters, such as FA and CHA, that represent phonemes not used in Tedim suggest that the script was intended for writing other Chin languages, such as Mizo, as well.

The onset (C_1) is represented by one of the following consonant letters. Only these letters may appear in the initial consonant position.¹⁰

Œ	PA	/p/	2	NGA	/ŋ/	L	TA	/t/
0	KA	/k/	9	HA	/h/	G	HTA	$/t^h/$
С	LA	/1/	r	GA	/g/	C	NA	/n/
9	MA	/m/	\propto	XA	/x/	Z	HPA	$/p^h/$
В	DA	/d/	е	HSA	/s/	7	RA	/r/
U	YA	/j/	3	BA	/b/	h	FA	/ f /
Ω	VA	$/_{ m V}/$	И	TGA	$/_{ m I}/$	ዠ	CHA	/ch/

The nucleus (V_1) is represented using the following vowel letters. The letters UA and IA represent dipthongs, which are coded as $V_1(V_2)$.

⁶ Banks 1967: 59. ⁷ Button 2009: 23, fn. 20. ⁸ Bennison 1933: 217. ⁹ Thang 2001: 33. ¹⁰ Button 2009: 34. ¹¹ Button 2009: 26.

8	AA	/a/	Α	\mathbf{AW}	/o/	٨	IA	/I a /
S	AAI	/e/	d	U	/u/			
ገ	I	/i/	Ŧ	UA	/va/			

The coda (C_2) is represented using one of the following consonant letters. These letters may appear only as the final consonant.¹² The letters AU and AI are glides.¹³

૭	AB	/p/	V	AM	/m/	ρ	AU	$/_{\rm W}/$
Т	AG	/k/	Ġ	AN	/n/	ન	ANG	/ŋ/
Ъ	AD	/t/	Λ	AL	/1/	0	ΑI	/j/

The tone (T) is represented using one of the 20 so-called 'tonal signs', shown below. Script charts show the tone marks arranged into four rows. The significance of this arrangement is unknown. The glyphs of a particular row exhibit a common structure, except for those of the third row:

ROW A	ı	ŀ	 :	II	·	:
ROW B	1	ŀ		11	11	
ROW C		•	:	3	3	
ROW D	l	ŀ	ij	ll	lli	LÜ

Christopher Button, a specialist of northern Chin linguistics, has shed much light upon the properties and arrangement of the tonal signs. ¹⁴ He explains that the signs are used to mark vowel length, tone, and syllable-final glottal stop (written in the Latin orthography using 'h'). His description alludes to a sub-grouping of the signs. The ten signs in the left-most three columns are used for marking "3 basic + 1 sandhi tones with 2 vowel distinctions"; the "remaining 2 are then left for codas ending in a glottal stop" of which one is a "sandhi alternate". The signs in the right-most three columns have the same features as those in the first three, but are used at the end of sentences. Button suggests that the doubling of lines is borrowed from Burmese, where two vertical lines are used to mark end of sentence. Further research is required in order to determine the actual semantic values of each tonal sign, but based on the preceding information, the signs may be tentatively described as follows:

	SHORT			SEN	INAL		
	SHORT	LONG	SANDHI	SHORT	LONG	SANDHI	
RISING	L	ŀ	ŀ	ll	lli	ناا	
MID	ı	ŀ	 :	II	II٠	:	
FALLING	1	1:		11	11		
CODA W/ GLOTTAL STOP		•	:	3	큿		

Digits and punctuation are unattested. However, punctuation may be inherently reflected in some tonal signs,

Button 2009: 30. 13 Button 2009: 27. 14 Button, February 2010, personal correspondence.

as described above. Figure 4 shows the use of Latin digits and punctuation (period, etc.) in printed Pau Cin Hau texts.

5 Source for Glyph Shapes

The proposed glyphs for Pau Cin Hau characters shown in the code chart in Figure 1 are based upon the forms shown in script charts. The font was designed by the present author and is a work in progress.

6 Character Properties

The tentative properties for Pau Cin Hau characters in the Unicode Character Database format are given below. The properties of tone marks are not accurate.

```
11B00; PAU CIN HAU LETTER PA; Lo; 0; L;;;;; N;;;;;
11B01; PAU CIN HAU LETTER KA; Lo; 0; L;;;;; N;;;;;
11B02; PAU CIN HAU LETTER LA; Lo; 0; L;;;;; N;;;;;
11B03; PAU CIN HAU LETTER MA; Lo; 0; L;;;;; N;;;;
11B04; PAU CIN HAU LETTER DA; Lo; 0; L;;;;; N;;;;
11B05; PAU CIN HAU LETTER YA; Lo; 0; L;;;;; N;;;;
11B06; PAU CIN HAU LETTER VA; Lo; 0; L;;;;; N;;;;;
11B07; PAU CIN HAU LETTER NGA; Lo; 0; L;;;;; N;;;;;
11B08; PAU CIN HAU LETTER HA; Lo; 0; L;;;;; N;;;;
11B09; PAU CIN HAU LETTER GA; Lo; 0; L;;;;; N;;;;
11B0A; PAU CIN HAU LETTER XA; Lo; 0; L; ;; ;; N; ;; ;;
11B0B; PAU CIN HAU LETTER HSA; Lo; 0; L;;;;; N;;;;
11B0C; PAU CIN HAU LETTER BA; Lo; 0; L;;;;; N;;;;
11BOD; PAU CIN HAU LETTER TGA; Lo; 0; L;;;;; N;;;;;
11B0E; PAU CIN HAU LETTER TA; Lo; 0; L; ;; ;; N; ;; ;;
11B0F; PAU CIN HAU LETTER HTA; Lo; 0; L;;;;; N;;;;;
11B10; PAU CIN HAU LETTER NA; Lo; 0; L;;;;; N;;;;
11B11; PAU CIN HAU LETTER HPA; Lo; 0; L;;;;; N;;;;;
11B12; PAU CIN HAU LETTER RA; Lo; 0; L;;;;; N;;;;
11B13; PAU CIN HAU LETTER FA; Lo; 0; L;;;;; N;;;;;
11B14; PAU CIN HAU LETTER CHA; Lo; 0; L;;;;; N;;;;;
11B15; PAU CIN HAU LETTER AA; Lo; 0; L;;;;; N;;;;
11B16; PAU CIN HAU LETTER AAI; Lo; 0; L;;;;; N;;;;;
11B17; PAU CIN HAU LETTER I; Lo; 0; L; ; ; ; ; N; ; ; ;
11B18; PAU CIN HAU LETTER AW; Lo; 0; L;;;;; N;;;;
11B19; PAU CIN HAU LETTER U; Lo; 0; L; ;; ;; N; ;; ;;
11B1A; PAU CIN HAU LETTER UA; Lo; 0; L;;;;; N;;;;;
11B1B; PAU CIN HAU LETTER IA; Lo; 0; L;;;;; N;;;;;
11B1C; PAU CIN HAU LETTER AB; Lo; 0; L;;;;; N;;;;;
11B1D; PAU CIN HAU LETTER AG; Lo; 0; L;;;;; N;;;;;
11B1E; PAU CIN HAU LETTER AD; Lo; 0; L;;;;; N;;;;;
11B1F; PAU CIN HAU LETTER AM; Lo; 0; L;;;;; N;;;;
11B20; PAU CIN HAU LETTER AN; Lo; 0; L;;;;; N;;;;;
11B21; PAU CIN HAU LETTER AL; Lo; 0; L;;;;; N;;;;
11B22; PAU CIN HAU LETTER AU; Lo; 0; L;;;;; N;;;;
11B23; PAU CIN HAU LETTER ANG; Lo; 0; L;;;;; N;;;;;
11B24; PAU CIN HAU LETTER AI; Lo; 0; L;;;;; N;;;;;
11B25; PAU CIN HAU TONE MARK ONE; Mc; 0; L;;;;; N;;;;
11B26; PAU CIN HAU TONE MARK TWO; Mc; 0; L;;;;; N;;;;;
11B27; PAU CIN HAU TONE MARK THREE; Mc; 0; L;;;;; N;;;;
11B28; PAU CIN HAU TONE MARK FOUR; Mc; 0; L;;;;; N;;;;
11B29; PAU CIN HAU TONE MARK FIVE; Mc; 0; L;;;;; N;;;;;
11B2A; PAU CIN HAU TONE MARK SIX; Mc; 0; L;;;;; N;;;;
11B2B; PAU CIN HAU TONE MARK SEVEN; Mc; 0; L;;;;; N;;;;;
```

```
11B2C; PAU CIN HAU TONE MARK EIGHT; Mc; 0; L; ;; ;; N; ;; ;; 11B2D; PAU CIN HAU TONE MARK NINE; Mc; 0; L; ;; ;; N; ;; ;; 11B2E; PAU CIN HAU TONE MARK TEN; Mc; 0; L; ;; ;; N; ;; ;; 11B2F; PAU CIN HAU TONE MARK ELEVEN; Mc; 0; L; ;; ;; N; ;; ;; 11B30; PAU CIN HAU TONE MARK TWELVE; Mc; 0; L; ;; ;; N; ;; ;; 11B31; PAU CIN HAU TONE MARK THIRTEEN; Mc; 0; L; ;; ;; N; ;; ;; 11B32; PAU CIN HAU TONE MARK FOURTEEN; Mc; 0; L; ;; ;; N; ;; ;; 11B33; PAU CIN HAU TONE MARK FIFTEEN; Mc; 0; L; ;; ;; N; ;; ;; 11B34; PAU CIN HAU TONE MARK SIXTEEN; Mc; 0; L; ;; ;; N; ;; ;; 11B35; PAU CIN HAU TONE MARK SEVENTEEN; Mc; 0; L; ;; ;; N; ;; ;; 11B36; PAU CIN HAU TONE MARK EIGHTEEN; Mc; 0; L; ;; ;; N; ;; ;; 11B37; PAU CIN HAU TONE MARK NINETEEN; Mc; 0; L; ;; ;; N; ;; ;; 11B38; PAU CIN HAU TONE MARK TWENTY; Mc; 0; L; ;; ;; N; ;; ;;
```

7 Outstanding Issues

More information about the Pau Cin Hau script must be collected. The outstanding issues are discussed below.

- 1. Origins What is the exact relationship between the logographic and alphabetic forms of the script?
- 2. Name Are the logographic and alphabetic forms of the script known by different names?
- 3. *Allocation* The script may be minor, but there is a user community associated with it. Should it be encoded in the Basic Multilingual Plane (BMP) or in the Supplementary Multilingual Plane (SMP)?
- 4. *Character Repertoire* Although the character repertoire mirrors that given in script charts, it is necessary to determine if any characters have been added since Pau Cin Hau finalized it in 1931.
- 5. Character Names Character names as given in the charts are problematic:
 - (a) The names for letters that represent the onset consonants possess an 'A', which is suggestive of Brahmic character names, eg. PA, KA, LA, etc. However, Pau Cin Hau characters do not possess an inherent vowel and the script is not based upon a Brahmic model.
 - (b) The 'presence' of Latin transliterations of voiced consonants in the character names for coda consonants contradict the basic phonological properties. For example, the letters AB, AG, AD are used only for representing codas. However, codas can be only voiceless, liquid, or nasal consonants. Moreover, the 'presence' of Latin 'A' in the names of coda letters suggests a phonological property that does not inhere in the letter.

It is possible that the character names are influenced by names of Burmese characters, and that the 'A' in names of consonant letters was added to assist in the pronunciation of the letter names.

Should character names reflect the original names as given in the charts or be changed to reflect actual phonetic values and character semantics?

- 6. *Character Properties* Does the possible use of some tone marks to represent punctuation at the end of sentences present any technical concerns?
- 7. *Phonetic Values* Need to verify the phonetic values of characters.

5

¹⁵ Button 2009: 30.

- 8. *Transliteration* Is there a standard transliteration for Tedim and other Chin languages that may be assigned to Pau Cin Hau characters?
- 9. *Collating Order* What is the collating order for Pau Cin Hau? Does it follow the order given in the script charts, which is mirrored in the encoding order? Does it follow the pattern used for Tedim?
- 10. Glyph Shapes Need to verify that glyph shapes are representative. Also, determine the nature of the nearly 'homoglyphic' letters: HP CHA and HP AD; P RA and P AU. The letters in the scripts charts upon which the digitized glyphs are based are distorted by the low resolution of the images.
- 11. *Tonal Signs* The value of each individual tonal sign remains to be identified.
- 12. *Specimens* Better reproductions hand-written and printed specimens are required. Is a complete printed version of the "Sermon on the Mount", an excerpt of which is shown in Figure 4, available anywhere? Also, is the book from which the specimen in Figure 3 was extracted still available?

8 References

- The American Bible Society. 1938. The Book of a Thousand Tongues: Being Some Account of the Translation and Publication of All or Part of The Holy Scriptures Into More Than a Thousand Languages and Dialects With Over 1100 Examples from the Text. Edited by Eric M. North. New York and London: Harper & Brothers.
- 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
- Go, Khup Za. 2008. Zo Chronicles: A Documentary Study of History and Culture of the Kuki-Chin-Lushai Tribe. New Delhi: Mittal Publications.
- 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/
- 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

9 Acknowledgments

I owe a great deal of appreciation to Christopher Button. He has patiently fulfilled my requests for information and overlooked my ignorance of Chin linguistics. He continues to explain the logic underlying the Pau Cin Hau system and his comments on the tonal signs are especially valuable. Last, but not least, he has provided me with specimens of the logographic form of the script.

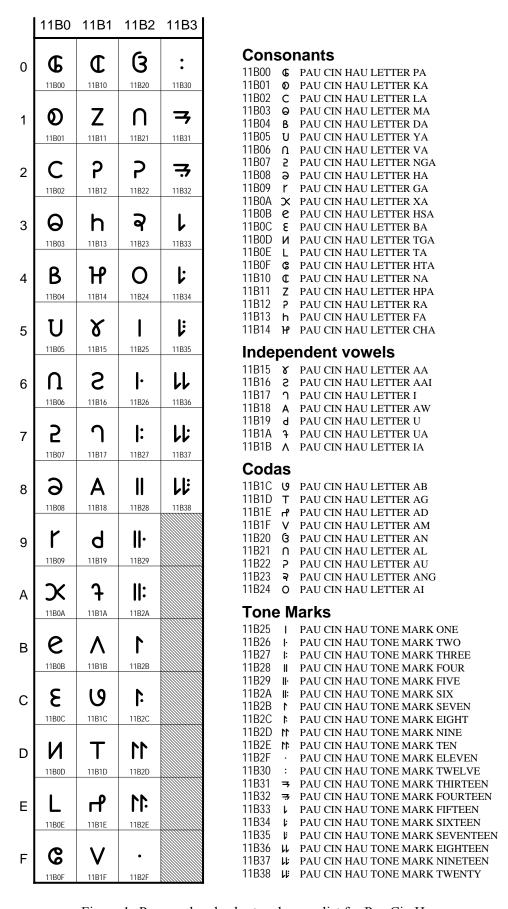


Figure 1: Proposed code chart and nameslist for Pau Cin Hau

						11111	AU CHIN HAL	, 50	••					
							(See faragraph 120).							
3 O C 6	១៩ ប	n e s n	×eε	Иι	s c z ·	የክዝኛ		pa ka la m	-	va nga h	a ga xa bs	be tga	ta hta na	hpa ra s
S &		A d	4 1					à ải i ave						
ד ט	Ħ	V (s	U P	, 3	0			ab ag ad	am an ai	an ang	aı			
1	ŀ	ij	ŧı		4	H:		1						
٢	t,		11		11			These a	re tonal	signs				
	• -	:	7		₹)		•				
. F .:	Ł	Ę	u		W.	W		-						
21.								på	pai	pi	baw.	pu	pua	pia
€x.	ડ્ય	61	© A	Gd	6 9	Œ ∧		kà	kai	ki	kaw	ku	kna	kia
ΦÃ	\$0	లా	ØA.	٥d	09	OΛ		諡	lai	li	law	lu	lua	lia
C)	CS	C٩	GA	Cd	Ca	GA.		mà	mai	m:	mav	B)ii	mua	mia
98	63	91	94	99	63	ΘΛ		dà	dai	di	daw	đu	dua	dia
BX	SB	81	BA	Bd	B9	BΛ		yà	yaı	yi	yaw	yu	you	yia
U	Sa	บา ดา	AD AA	rd ba	79 79	UA		và	vai	vi	vaw.	vu	yoa	via
ΩX	58	24	AS	59	54	5V UV		ngá	ngai	ngi	ngaw	ngu	ngua	nga
91 51	ક્ટ ક્ટ	&7 ∂ 9	94	99	24 24	9 V		hà	hai	hi	haw	hu	hua	hia
rs	42	r9	KA	rd	r3	41		gà	gai	gi	gaw	gu	gua	gui
×8	S×C	3 <1	3cA	×d	><1	× ^		ĸà	xai	хi	xaw	xu	xua	xia
68	89	64	eA.	ed	ê١	eΛ		lisā	hsai	hsi	hsaw	hsu	hsua	hsia
83	53	٤٩	EA	εd	έş	81		bà	hai	bi	pan.	bu	bua	bia
NX	SN	И9	AN	NA	ия	NV		tga	tga	tgi	tgaw	tgu	tgua	tgia
FR	1.2	Ĺ'n	LA	Ld	L٦	L۸		tà	tai	ti	taw	tu	tua	tia
e z	GS	œ9	A.D	Gq	G9	GA.		htá	htai	hti	htaw	htu	htua	htia
C.X	SD.	C٩	CA	æq	C?	ŒΛ		nå	nai	ni	naw	nu	17021	nio
ZY	ZZ	29	ZA	Zd	Ž٩	Z٨		hpd	hpai	hpir	hpaw	hpu	hpua	bpia
28	Pe	Pa	PA	29	2.4	PN		rà	rai	ř.	TAW	TU	rua	ria
ny	કત	กา	hA	ho	h9	h'A		£	fai		faw	fu	fua	lia
76 g	772	767	YA	rd	K9	K V		chá	chai	.chi	chaw 1	chu	chua	chia

Figure 2: The Pau Cin Hau script (from Bennison (1931), reprinted in Go (2008: 94)).

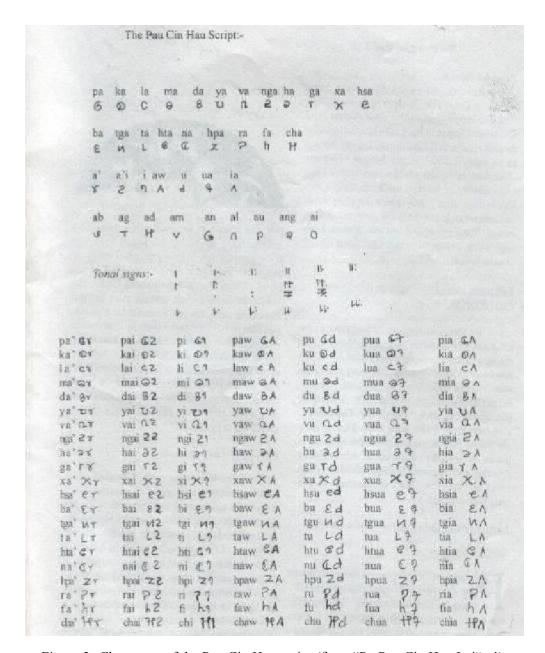


Figure 3: Characters of the Pau Cin Hau script (from "Pu Pau Cin Hau Lai": 4).

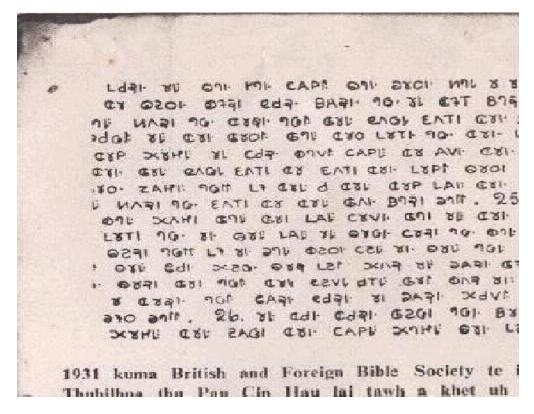


Figure 4: Excerpt of the "Sermon of the Mount" printed in Pau Cin Hau (from "Pu Pau Cin Hau Lai": 3).

LAAH AR OUH LUH CASE OUH OACH NU A A[...] (CX 0201 0171 647 BAR 10. XI (CT B17[...] [...] - ITA3 30A9 38D 30P - 1FXD - OP 1FAN 3P [?]dUI: YII CYO IOYO IOYO LYTI: 10. CAS XX41 AF CGS DJA: CVS1 CX VAI IOXO 1681 ILV3 RD ILV3 16V0 18D [?]YO. ZAHI: YUN LA CYI d CYI CYP LAI: CYI [?]; NAPI 10. EATI CY CY; GAI B1PI 211. OU! XVHI OU! RAI LA! CAN! OU! A! CA! LYTI 10. YI OY! LA! Y! OYU! CYTI 10. D1! 197 USG 18 13 1050 105 IS ET 1191 IESG [?] OY; GdI X2U· OY[?] L21 XA7 Y[?] OAPI G[...] [?] OVAI (VI 10[?] (V[?] C2VI dTI GYI OAA VI 1VPX IEVE IR 1EPO 1EVO 16L 1ERD & 310 3111. 26 VI CdI Cd71 G2UI 1UI BY[...] XXH; CX; 2AGI CXI CAP; X1H; OXI L[...]

Figure 5: The Pau Cin Hau text from Figure 4 rendered in a digitized font. There may be errors in the transcription on account of the low resolution of the source image.



Figure 6: A signboard in the logographic Pau Cin Hau script and in the Myanmar script. Photograph taken by Vungh Suan Mang in Tedim Township, Chin State, Myanmar on December 31, 2008.