Title: Defining Properties for Tone Marks of the Pau Cin Hau Script
Author: Anshuman Pandey (pandey@umich.edu)
Status: Individual Contribution
Action: For consideration by UTC and WG2
Date: 2010-05-22

1 Introduction

This document provides additional details about the tone marks of the Pau Cin Hau Syllabary, as introduced in N3781 L2/10-080. The information presented here is based upon personal correspondence with Christopher Button, a specialist of Chin linguistics, from March 2010. Questions regarding the encoding of these characters are identified in Section 3, for which advice is sought from the Unicode Technical Committee.

2 Background

The Pau Cin Hau Syllabary has 20 tone marks, which are said to represent the complete tonal system of the Tedim language that is spoken in Chin State, Myanmar. There is a paucity of information about these characters. Charts of the script simply show the tone marks arranged in a matrix of four rows and six columns, as shown below, but do not explain the logic behind the arrangement or reveal the values of the marks:

The term 'tone mark' is not a completely accurate description of the characters. Of the 20 marks, 15 represent tone (columns 1, 2, 4, 5), while 5 represent glottal stop (columns 3 and 6). Based upon glyph shapes, the 15 tone marks may be divided into two groups of two columns (1–2 and 4–5), which indicate positional and vocalic features. The rows indicate tone categories. The 15 tone marks may be classified as follows:

	WORD FINAL		SENTENCE FINAL	
	LONG VOWEL	SHORT VOWEL	LONG VOWEL	SHORT VOWEL
TONE II	Ⅰ	Ⅰ·	Ⅱ	Ⅱ·
SANDHI TONE	Ⲅ	Ⲅ:	ⲅⲅ	ⲅⲅ:
TONE I		·	⇉	⇉
TONE III	Ⳑ	Ⳑ:	ⳑⳑ	ⳑⳑ:

The tone marks in the 'word-final' class represent (a) basic tone + vowel length and (b) sandhi tone + vowel length. The 'sentence-final' marks are contextual variants of the 'word-final' marks and correspond to them by column. When a 'word-final' tone mark appears at the end of a sentence, it is replaced with the corresponding 'sentence-final' mark.

The five marks used for indicating glottal stop may be classified as shown below. The glottal stop has regular and sandhi-variant forms.

	WORD FINAL	SENTENCE FINAL
	GLOTTAL STOP	GLOTTAL STOP
SANDHI VARIANT	╟:	╟:
SANDHI VARIANT?	:	
REGULAR	╟	╟

The placement of glottal stop marks along with tone marks in the traditional arrangement is peculiar. The glottal stop has no actual correlation with tones in the phonology of the Tedim language. Connected speech allows fluctuations in pronunciation that are functionally matched with the tones orthographically, but with which there is no real phonological link. The regular glottal stop mark ╟ is paired with tone III orthographically (row 4) and the sandhi-variant form ╟: is paired with tone II orthographically (row 1). The sandhi-variant form :, which is orthographically paired with tone I (row 3), does not have real representation in Tedim, as there is only one possible sandhi variant of words with glottal stops.

The glyph shapes of the 'sentence-final' marks imply their function as punctuation marks. The 'sentence-final' marks are produced by doubling the basic graphical element of the 'word-final' mark, eg. ╵ → ‖, ╟ → ╟, etc. This practice shows the influence of Burmese orthography, namely the shape of ‖ U+104B MYANMAR SIGN SECTION, which may be decomposed as a doubling of ╵ U+104A MYANMAR SIGN LITTLE SECTION. The exceptions to this composition rule are the shapes of marks in the tone I category, for which the 'sentence-final' marks are produced by introducing doubles of graphical elements not found in the 'word-final' mark, eg. ∅ → ⇉ and · → ⇉. Alternately, the horizontal doubled lines may represent the absence of an orthographic symbol in non-sentence final position.

While the Pau Cin Hau Syllabary may have been designed for writing Tedim, it appears that the tone-mark repertoire of the script was intended for comprehensive representation of the northern Chin languages. For example, as indicated above, not all tone marks have actual representation in Tedim. There is only one possible sandhi variant for word-final glottal stop in the language, but there are three marks for syllable-final glottal stop in the script. It is likely that the mark : was included for writing other Chin languages that possess a sandhi variant of glottal stop that is orthographically paired with tone I (row 1).

The contours of the basic tone categories of the northern Chin languages vary by language, as shown below:[1]

---
[1] Adapted from Button (2009: 37).

	MIZO	ZAHAU	THADO	ZO	TEDIM	SIZANG
TONE I	˥	˧	˧	˦	˦	˦
TONE IIA	˧	˥	˨	˧	˧	˧
TONE IIB	˨	˨				
TONE III	˦	˦	˦	˨	˦	˨

Based upon the differences in the correlation between tone contour and tone category, the semantic value of Pau Cin Hau tone marks is dependent upon the language being represented.

3   Questions

Naming Convention   Is there a preferred convention for naming characters that represent tones? Tone characters that are already encoded in the UCS for other scripts have names that include 'mark' and 'sign', while some names do not contain such descriptors; some characters are named serially, while others are named according to the tone contours they represent, eg.

- U+07EB NKO COMBINING SHORT HIGH TONE
- U+0EC8 LAO TONE MAI EK
- U+1063 MYANMAR TONE MARK SGAW KAREN HATHI
- U+1087 MYANMAR SIGN SHAN TONE-2
- U+1970 TAI LE LETTER TONE-2
- U+19C8 NEW TAI LUE TONE MARK-I

Both U+1063 MYANMAR TONE MARK SGAW KAREN HATHI and U+1087 MYANMAR SIGN SHAN TONE-2 are used for representing tone, have identical properties in the Unicode Character Database, and are part of the same script block, but they are named differently. Are there criteria for distinguishing a tone 'mark' from a tone 'sign'?

Character Names   In the preliminary proposal the tone marks were given tentative names based upon their serial order in script charts, eg. PAU CIN HAU TONE MARK ONE ... PAU CIN HAU TONE MARK TWENTY. Now that values for the tone marks have been established, should the generic names be replaced with more descriptive names? Are the names given below suitable?

ǀ	PAU CIN HAU TONE-2 LONG VOWEL		·	PAU CIN HAU TONE-I SHORT VOWEL
ǀ·	PAU CIN HAU TONE-2 SHORT VOWEL		:	PAU CIN HAU TONE-I GLOTTAL STOP
ǀ:	PAU CIN HAU TONE-2 GLOTTAL STOP		⇉	PAU CIN HAU TONE-I LONG VOWEL FINAL
ǁ	PAU CIN HAU TONE-2 LONG VOWEL FINAL		⇉	PAU CIN HAU TONE-I GLOTTAL STOP FINAL
ǁ·	PAU CIN HAU TONE-2 SHORT VOWEL FINAL		↳	PAU CIN HAU TONE-3 LONG VOWEL
ǁ:	PAU CIN HAU TONE-2 GLOTTAL STOP FINAL		↳	PAU CIN HAU TONE-3 SHORT VOWEL
⌐	PAU CIN HAU SANDHI TONE LONG VOWEL		↳	PAU CIN HAU TONE-3 GLOTTAL STOP
⌐·	PAU CIN HAU SANDHI TONE SHORT VOWEL		⌐⌐	PAU CIN HAU TONE-3 LONG VOWEL FINAL
⌐⌐	PAU CIN HAU SANDHI TONE LONG VOWEL FINAL		⌐⌐·	PAU CIN HAU TONE-3 SHORT VOWEL FINAL
⌐⌐:	PAU CIN HAU SANDHI TONE SHORT VOWEL FINAL		⌐⌐:	PAU CIN HAU TONE-3 GLOTTAL STOP FINAL

Character Properties   What is the appropriate way to describe the properties of tone marks that are also used as punctuation?

Joining Behavior   Should the tone marks be treated as combining characters?

Ordering   Should the tone marks be ordered according to the traditional arrangement given in script charts or by tone contour?

4   References

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`

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

5   Acknowledgments