

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 to encode three Bopomofo letters for Hmu and Ge  
**Source:** Andrew West  
**Status:** Individual Contribution  
**Action:** For consideration by JTC1/SC2/WG2 and UTC  
**Date:** 2009-01-25

## 1. Introduction

This is a proposal to encode three Bopomofo letters that are used in an obsolete orthography for representing the Hmu and Ge languages that was developed by the Australian missionary Maurice H. Hutton in the 1920s and 1930s. The Hmu (Chinese Hemu 赫目), also known as the "Black Miao", and the Ge 革 (also known as Gedou 革兜) peoples, who are officially classified as members of the Miao 苗 nationality, mostly live in southeastern Guizhou. Hutton developed an orthography for representing Hmu based on the National Phonetic Script (i.e. Bopomofo) in about 1920, and during the late 1920s and early 1930s he published Hmu translations of the Four Gospels (1928), a Catechism (1928), a Hymn Book (1928), and eventually the complete New Testament (1934). He also applied this orthography to the Ge language (called *Keh Deo* by Hutton). He first included some hymns in Ge in his *Black Miao Hymnary* of 1928, and later published a *Keh Deo Catechism and Hymnary* (1937) and the Gospels of St. Mark and St. John (both in 1937). Although Hutton's orthography for Hmu and Ge is now long since obsolete, it is reported that there is a community of about fifty Hmu Christians in the vicinity of Kaili (in western Guizhou) that still make use of Hutton's 1928 *Black Miao Hymnary* printed in his extended Bopomofo script.

Hutton's orthography for Hmu and Ge uses most of the standard Bopomofo letters, together with three special letters that he devised in order to represent sounds not found in standard Chinese (see Fig. 1 for a list of all the letters used in Hutton's orthography for Hmu):

ㄗ representing [ɣ] (called "Z hard" by Hutton)

ㄌ representing [ʎ] (called "LI" by Hutton)

ㄗ representing [ʒ] (called "yee" by Hutton)

All three of these letters are used for both Hmu and Ge (see Figs. 2 and 3 for Hmu samples and Fig. 4 for a Ge sample).

In addition to devising additional Bopomofo letters, Hutton also used two special tone marks, ㄛ for the level mid high tone 44, and ㄛ for the level low tone 22. However, these can be represented by the combining sequences <302B 302A> ㄛ and <302A 302D> ㄛ respectively, so do not need encoding separately.

Proposed Code Point	Proposed Character Name	Glyph
31B8	BOPOMOFO LETTER XX	Z
31B9	BOPOMOFO LETTER LL	L
31BA	BOPOMOFO LETTER ZZ	+

Unicode Character Properties	Script	Line Break Property
31B8;BOPOMOFO LETTER XX;Lo;0;L;;;;;N;;;;;	Bopomofo	ID
31B9;BOPOMOFO LETTER LL;Lo;0;L;;;;;N;;;;;	Bopomofo	ID
31BA;BOPOMOFO LETTER ZZ;Lo;0;L;;;;;N;;;;;	Bopomofo	ID

## 2. Examples

ㄅ	<b>B</b> [p]	ㄋ	<b>ng</b> [ŋ]	ㄗ	<b>Ts'</b> [tʂʰ]	ㄛ	<b>eh</b> [ɔ]
ㄆ	<b>p'</b> [pʰ]	ㄏ	<b>H' or Hor</b> [χ]	ㄘ	<b>S</b> [s]	ㄝ	<b>eh</b> [e]
ㄇ	<b>M</b> [m]	ㄐ	<b>g or J</b> [tɕ]	ㄗ	<b>Z</b> (hard)	ㄞ	<b>ai</b> [ai]
ㄉ	<b>F</b> [f]	ㄑ	<b>ch'</b> [tʂʰ]	ㄛ	<b>Ll</b>	ㄟ	<b>ei</b> [ei]
ㄌ	<b>V</b> [w]	ㄒ	<b>N</b> [nʰ]	ㄜ	<b>yee</b>	ㄠ	<b>ao</b> [au]
ㄎ	<b>D</b> [t]	ㄒ	<b>sh</b> [ç]			ㄡ	<b>eo</b> [əu]
ㄏ	<b>T'</b> [tʰ]	ㄣ	<b>Ts</b> [tʂ]	ㄝ	<b>E or ih</b> [i]	ㄢ	<b>an</b> [an]
ㄌ	<b>N</b> [n]	ㄤ	<b>Ts'</b> [tʂʰ]	ㄨ	<b>oo</b> [u]	ㄣ	<b>en</b> [ən]
ㄎ	<b>L</b> [l]	ㄨ	<b>Hs or hsih</b> [ʃ]	ㄩ	<b>ü</b> [y]	ㄤ	<b>ang</b> [aŋ]
ㄎ	<b>K, or gore</b> [k]	ㄨ	<b>R or rī</b> [ʀ]	ㄩ	<b>a</b> [a]	ㄤ	<b>eng</b> [əŋ]
ㄎ	<b>K' or Kore</b> [kʰ]	ㄨ	<b>Ts</b> [tʂ]	ㄩ	<b>ore</b> [o]	ㄤ	<b>er</b> [ə]

Fig. 1 : Joakim Enwall, *A Myth Become Reality* Vol.1 (Stockholm, 1994) page 198

(Table of Letters in Hutton's Orthography for Hmu)

*Sample Text in Hutton's NPS-Writing for Hmu*

St Mark 1:7, New Testament, 1934.

ni <sup>55</sup> he	qan <sup>44</sup> tell	li <sup>22</sup> truth	tcu <sup>21</sup> speak,	me <sup>55</sup> there-is	i <sup>33</sup> one	le <sup>55</sup> class.	n'an <sup>33</sup> be
ve <sup>22</sup> I	taŋ <sup>22</sup> after	qe <sup>33</sup>	lo <sup>22</sup> come,	te <sup>22</sup> ability	yu <sup>21</sup>	pi <sup>55</sup> than	ve <sup>22</sup> I
me <sup>55</sup> have	tho <sup>33</sup> big	no <sup>22</sup> much,	ve <sup>22</sup> I	qa <sup>55</sup> just	to <sup>21</sup> be	qon <sup>22</sup> bend	tu <sup>33</sup> waist
to <sup>55</sup> with	ni <sup>55</sup> he	tha <sup>44</sup> solve	tcu <sup>55</sup> class.	tha <sup>22</sup> string	ha <sup>33</sup> shoe,	sei <sup>55</sup> also	to <sup>21</sup> be
a <sup>55</sup> not	p'ei <sup>22</sup> worthy	pie <sup>44</sup> part.					

'And he would preach, saying: "After me someone stronger than I is coming; I am not fit to stoop and untie the laces of his shoes.'"

Fig. 2 : Joakim Enwall, *A Myth Become Reality* Vol.1 (Stockholm, 1994) page 212

(Passage from St. Mark in Hmu)





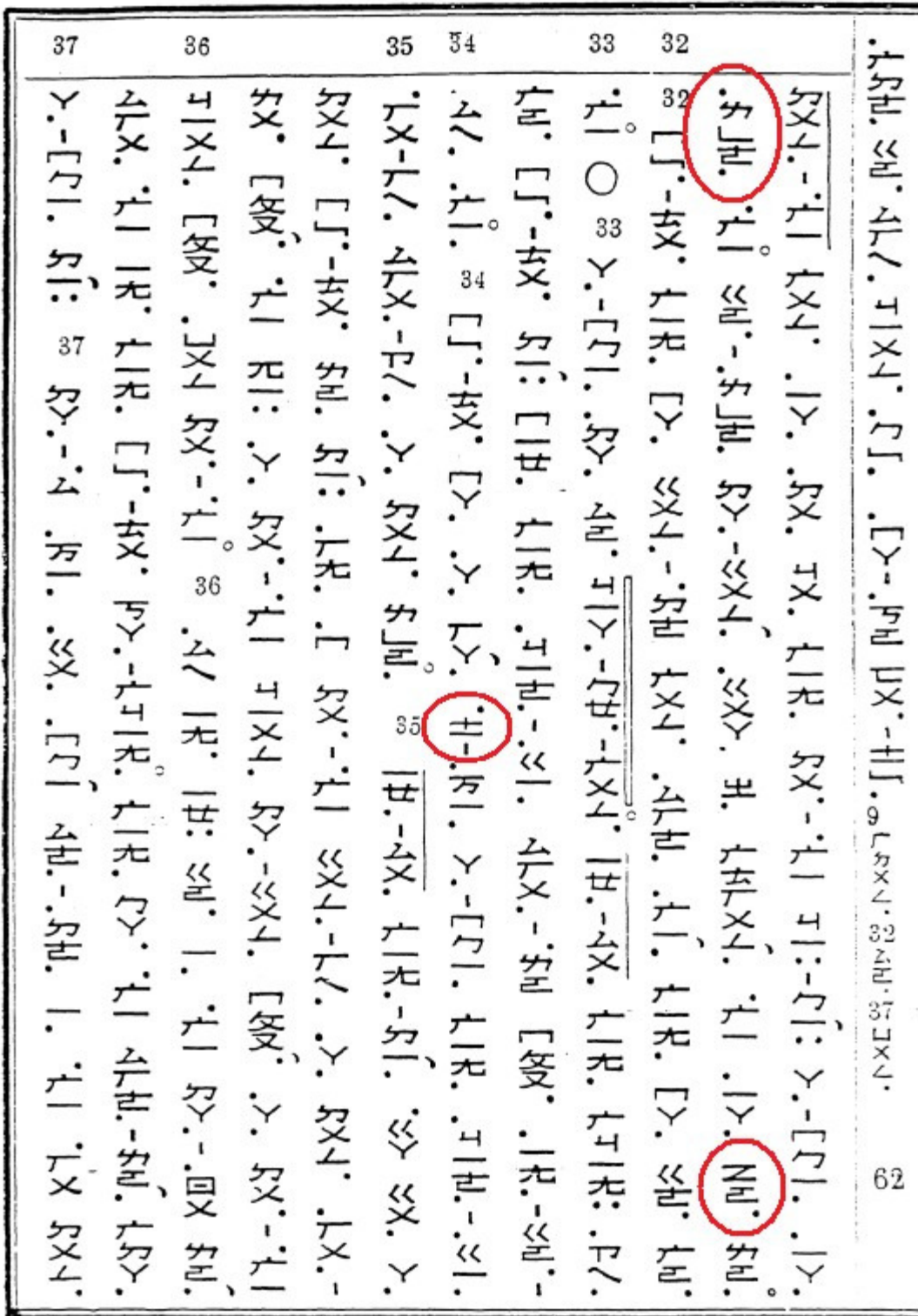


Fig. 4 : Joakim Enwall, "Introductory remarks on M.H. Hutton's Keh Deo writing" (1996) page 158

(Passage from St. Mark in Ge)

### 3. Bibliography

#### **Publications by Hutton in Black Miao (Hmu)**

*Black-Miao Hymnary*. Panghai: China Inland Mission, 1928.

*Black Miao St. Matthew*. London: British and Foreign Bible Society, 1928.

*Black Miao St. Mark*. London: British and Foreign Bible Society, 1928.

*Black Miao St. Luke*. London: British and Foreign Bible Society, 1928.

*Black Miao St. John*. London: British and Foreign Bible Society, 1928.

*Black Miao Hymnary*. Panghai: China Inland Mission, 1928.

*Black Miao Catechism*. Panghai: China Inland Mission, 1928.

*Black Miao New Testament*. London: British and Foreign Bible Society, 1934.

*Black Miao – English Dictionary*. Unpublished MS.

#### **Publications by Hutton in Keh Deo (Ge)**

*Catechism and Hymnary*. Shanghai: National Bible Society of Scotland, 1937.

*The Gospel of St. Mark*. Shanghai: National Bible Society of Scotland, 1937.

*The Gospel of St. John*. Shanghai: National Bible Society of Scotland, 1937.

#### **Secondary Sources**

Eugene Albert Nida and Eric McCoy North, *The Book of a Thousand Tongues* (revised ed.). London: United Bible Societies, 1972.

Joakim Enwall, *A Myth Become Reality : History and Development of the Miao Written Language*. Stockholm: Institute of Oriental Languages, 1994-1995.

Joakim Enwall, "Introductory remarks on M.H. Hutton's Keh Deo writing"; in *Mon-Khmer Studies* 28 (1996) pages 152-159.

---

**ISO/IEC JTC 1/SC 2/WG 2**  
**PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS**  
**FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646**

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.

**Form number: N3452-F** (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05)

### A. Administrative

1. Title:	<i>Proposal to encode three Bopomofo letters for Hmu and Ge</i>	
2. Requester's name:	<i>Andrew West</i>	
3. Requester type (Member body/Liaison/Individual contribution):		<i>Individual contribution</i>
4. Submission date:		<i>2009-01-25</i>
5. Requester's reference (if applicable):	<i>N/A</i>	
6. Choose one of the following:		
This is a complete proposal:		<i>YES</i>
(or) More information will be provided later:		<i>NO</i>

### B. Technical - General

1. Choose one of the following:		
a. This proposal is for a new script (set of characters):		<i>NO</i>
Proposed name of script:	<i>N/A</i>	
b. The proposal is for addition of character(s) to an existing block:		<i>YES</i>
Name of the existing block:	<i>Bopomofo Extended</i>	
2. Number of characters in proposal:		<i>3</i>
3. Proposed category (select one from below - see section 2.2 of P&P document):		
A-Contemporary	B.1-Specialized (small collection) <input checked="" type="checkbox"/>	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?		<i>YES</i>
a. If YES, are the names in accordance with the "character naming guidelines"?		<i>YES</i>
b. Are the character shapes attached in a legible form suitable for review?		<i>YES</i>
5. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?	<i>Andrew West</i>	
If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:		
6. References:		
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?		<i>YES</i>
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<i>YES</i>	
7. Special encoding issue		
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)?		<i>N/A</i>
8. 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 <a href="http://www.unicode.org">http://www.unicode.org</a> for such information on other scripts. Also see <a href="http://www.unicode.org/Public/UNIDATA/UCD.html">http://www.unicode.org/Public/UNIDATA/UCD.html</a> 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? If YES explain _____	<i>N/A</i>	_____	<i>NO</i>
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.)? If YES, available relevant documents: _____	<i>N/A</i>	_____	<i>NO</i>
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference: _____	<i>N/A</i>	_____	<i>N/A.</i>
4. The context of use for the proposed characters type of use; common or rare? Reference: _____	<i>N/A</i>	_____	<i>Rare</i>
5. Are the proposed characters in current use by the user community? If YES, where? Reference: _____	<i>N/A</i>	_____	<i>Possibly</i>
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If Yes, reference: _____	<i>N/A</i>	_____	<i>YES</i> <i>NO</i>
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?		_____	<i>N/A</i>
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If Yes, reference: _____	<i>N/A</i>	_____	<i>NO</i> <i>N/A</i>
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If Yes, reference: _____	<i>N/A</i>	_____	<i>NO</i> <i>N/A</i>
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? If YES, is a rationale for its inclusion provided? If Yes, reference: _____	<i>N/A</i>	_____	<i>NO</i> <i>N/A</i>
11. Does the proposal include use of combining characters and/or use of composite sequences? If YES, is a rationale for such use provided? If Yes, reference: _____	<i>N/A</i>	_____	<i>NO</i> <i>N/A</i>
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? If Yes, reference: _____	<i>N/A</i>	_____	<i>N/A</i>
12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary) _____		_____	<i>NO</i> <i>N/A</i>
13. Does the proposal contain any Ideographic compatibility character(s)? If YES, is the equivalent corresponding unified ideographic character(s) identified? If Yes, reference: _____	<i>N/A</i>	_____	<i>NO</i> <i>N/A</i>