

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 additional Miao characters in the SMP of the UCS****Source: Michael Everson and Erich Finkle****Status: Individual Contribution****Action: For consideration by JTC1/SC2/WG2 and UTC****Date: 2010-08-08**

1. Introduction. This proposal requests the addition of a number of letters used for Hei Yi in the Miao script. If accepted, the following characters will exist:

Ყ U+16F0C MIAO LETTER YI TTA
• used in Hei Yi

Შ U+16F12 MIAO LETTER YI NNA
• used in Hei Yi

Ჩ U+16F31 MIAO LETTER YI DZHA
• used in Hei Yi

Ც U+16F56 MIAO LETTER AHH
• used in Gan Yi

Ძ U+16F5B MIAO LETTER WO
• used in Hei Yi

2. The “wart”. Although in N3789 it is stated that “Some Yi users do the same thing with a vertical stroke mark resembling a serif on the letters. A given word in a particular dialect can be written with a wart, a dot, or (rather more rarely and with little standardization) the serif-like mark, or without being considered a spelling change.”, further research indicates that the “wart” proper should only be considered to be either a ring fused to the letter, or a free-standing dot (typically) to the right of the consonant. The “serif-like mark” referred to above should not be considered to be a “wart”. In Hei Yi, for example, both the balloting U+16F2E Ღ YI TSHA and the proposed U+16F31 Ჩ YI DZHA are used. Three consonants with this “serif-like mark” are proposed here.

3. Additional vowels. A glottalizing vowel Ც AHH and a sort of labialized Ძ WO are used in Hei Yi and were not proposed in N3789. We recommend that these characters be inserted into the appropriate places in the code chart, since the whole set is still under ballot. We also recommend a new column be added.

4. Unicode Character Properties. These should be checked carefully by the UTC against the list they are currently using for the other Miao characters.

```
16F0C;MIAO LETTER YI TTA;Lo;0;L;;;;N;;;;;
16F12;MIAO LETTER YI NNA;Lo;0;L;;;;N;;;;;
16F31;MIAO LETTER YI DZHA;Lo;0;L;;;;N;;;;;
16F56;MIAO LETTER AHH;Mc;0;L;;;;N;;;;;
16F5B;MIAO LETTER WO;Mc;0;L;;;;N;;;;;
```

Figures.

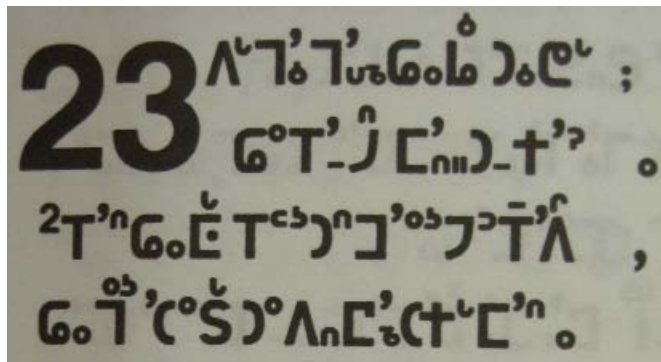


Figure 1. Sample showing LETTER WO, LETTER YI DZHA and LETTER YI NNA. “Hei Yi Book of Psalms”. Yunnan Sheng Jidujiao Xiehui 云南省基督教协会 “Yunnan Province Christian Association”, 2009.

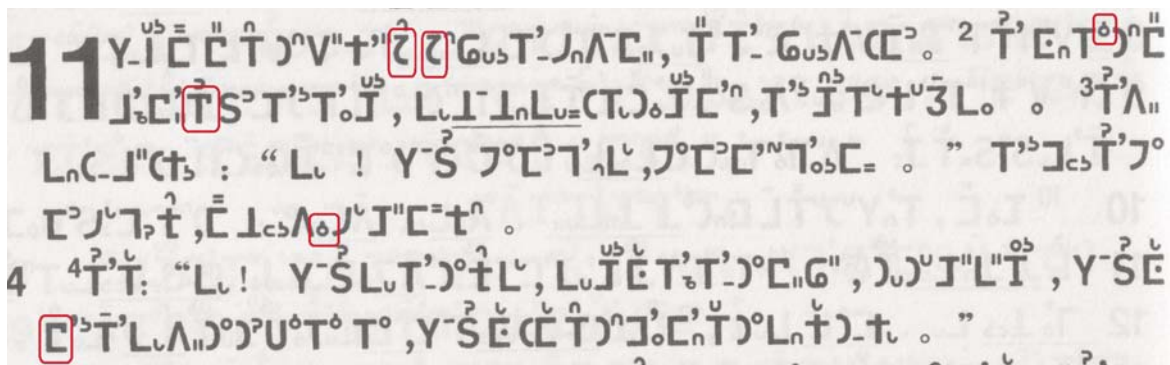


Figure 2. Sample showing LETTER YI TTA, YI NNA, LETTER YI DZHA, and LETTER WO. “Hei Yi Book of Genesis”. Yunnan Sheng Jidujiao Xiehui 云南省基督教协会 “Yunnan Province Christian Association”, 2009

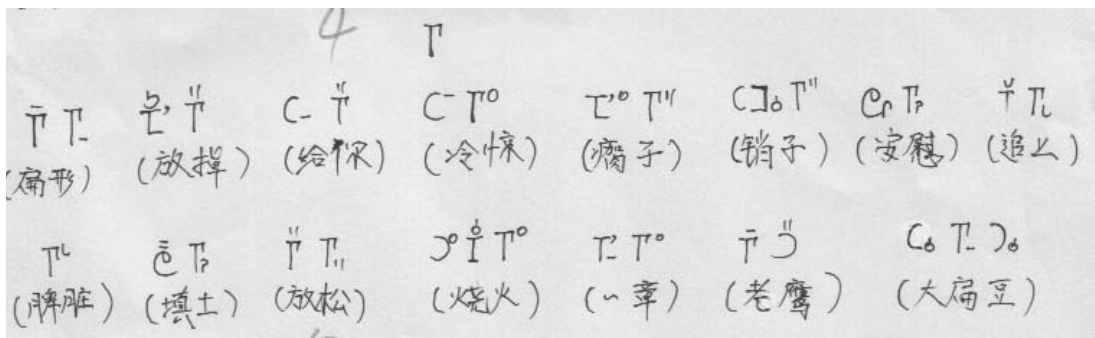


Figure 3. Sample showing MIAO LETTER YI TTA and MIAO LETTER WO.

Eastern Lisu (Lipo) word	Meaning in Mandarin Chinese	Meaning in English
𐌃	黄色	Color yellow
𐌃	可以	Can, able to
𐌃	花	flower
𐌃	哈	Ha of hallelujah, “the sound of loving”

Figure 4. Chart showing minimal pairs in Gan Yi, including MIAO LETTER AHH.

	16F0	16F1	16F2	16F3	16F4	16F5	16F6	16F7	16F8	16F9
0	𪛀 16F00	𪛁 16F10	𪛂 16F20	𪛃 16F30	𪛄 16F40	𪛅 16F50	𪛆 16F60	𪛇 16F70		𪛈 16F90
1	𪛉 16F01	𪛊 16F11	𪛋 16F21	𪛌 16F31	𪛍 16F41	𪛎 16F51	𪛏 16F61	𪛐 16F71		𪛑 16F91
2	𪛒 16F02	𪛓 16F12	𪛔 16F22	𪛕 16F32	𪛖 16F42	𪛗 16F52	𪛘 16F62	𪛙 16F72		𪛚 16F92
3	𪛛 16F03	𪛜 16F13	𪛝 16F23	𪛞 16F33	𪛟 16F43	𪛠 16F53	𪛡 16F63	𪛢 16F73		𪛣 16F93
4	𪛤 16F04	𪛥 16F14	𪛦 16F24	𪛧 16F34	𪛨 16F44	𪛩 16F54	𪛪 16F64	𪛫 16F74		𪛬 16F94
5	𪛭 16F05	𪛮 16F15	𪛯 16F25	𪛰 16F35		𪛱 16F55	𪛲 16F65	𪛳 16F75		𪛴 16F95
6	𪛵 16F06	𪛶 16F16	𪛷 16F26	𪛸 16F36		𪛹 16F56	𪛺 16F66	𪛻 16F76		𪛼 16F96
7	𪛽 16F07	𪛾 16F17	𪛿 16F27	𪜀 16F37		𪜁 16F57	𪜂 16F67	𪜃 16F77		𪜄 16F97
8	𪜅 16F08	𪜆 16F18	𪜇 16F28	𪜈 16F38		𪜉 16F58	𪜊 16F68	𪜋 16F78		𪜌 16F98
9	𪜍 16F09	𪜎 16F19	𪜏 16F29	𪜐 16F39		𪜑 16F59	𪜒 16F69	𪜓 16F79		𪜔 16F99
A	𪜕 16F0A	𪜖 16F1A	𪜗 16F2A	𪜘 16F3A		𪜙 16F5A	𪜚 16F6A	𪜛 16F7A		𪜜 16F9A
B	𪜝 16F0B	𪜞 16F1B	𪜟 16F2B	𪜠 16F3B		𪜡 16F5B	𪜢 16F6B	𪜣 16F7B		𪜤 16F9B
C	𪜥 16F0C	𪜦 16F1C	𪜧 16F2C	𪜨 16F3C		𪜩 16F5C	𪜪 16F6C	𪜫 16F7C		𪜬 16F9C
D	𪜭 16F0D	𪜮 16F1D	𪜯 16F2D	𪜰 16F3D		𪜱 16F5D	𪜲 16F6D	𪜳 16F7D		𪜴 16F9D
E	𪜵 16F0E	𪜶 16F1E	𪜷 16F2E	𪜸 16F3E		𪜹 16F5E	𪜺 16F6E	𪜻 16F7E		𪜼 16F9E
F	𪜽 16F0F	𪜾 16F1F	𪜿 16F2F	𪝀 16F3F		𪝁 16F5F	𪝂 16F6F		𪝃 16F8F	𪝄 16F9F

Consonant onsets

16F00	ɹ	MIAO LETTER PA	• used for ba in Dry Yi
16F01	ɹ̥	MIAO LETTER BA	
16F02	ɹ̥	MIAO LETTER YI PA	• used for pa in Dry Yi
16F03	ɹ̥	MIAO LETTER PLA	• used in Sichuan Hmong
16F04	ɹ̥	MIAO LETTER MA	
16F05	ɹ̥	MIAO LETTER MHA	
16F06	ɹ̥	MIAO LETTER ARCHAIC MA	• used in Pollard's early orthography
16F07	ɹ̥	MIAO LETTER FA	
16F08	ɹ̥	MIAO LETTER VA	
16F09	ɹ̥	MIAO LETTER VFA	• used in Black Yi
16F0A	ɹ̥	MIAO LETTER TA	• used for da in Dry Yi
16F0B	ɹ̥	MIAO LETTER DA	
16F0C	ɹ̥	MIAO LETTER YI TTA	• used in Hei Yi
16F0D	ɹ̥	MIAO LETTER YI TA	• used for ta in Dry Yi
16F0E	ɹ̥	MIAO LETTER TTA	
16F0F	ɹ̥	MIAO LETTER DDA	
16F10	ɹ̥	MIAO LETTER NA	
16F11	ɹ̥	MIAO LETTER NHA	
16F12	ɹ̥	MIAO LETTER YI NNA	• used in Hei Yi
16F13	ɹ̥	MIAO LETTER ARCHAIC NA	• used in Pollard's early orthography
16F14	ɹ̥	MIAO LETTER NNA	
16F15	ɹ̥	MIAO LETTER NNHA	
16F16	ɹ̥	MIAO LETTER LA	
16F17	ɹ̥	MIAO LETTER LYA	• used in Black Yi
16F18	ɹ̥	MIAO LETTER LHA	
16F19	ɹ̥	MIAO LETTER LHYA	• used in Black Yi
16F1A	ɹ̥	MIAO LETTER TLHA	
16F1B	ɹ̥	MIAO LETTER DLHA	
16F1C	ɹ̥	MIAO LETTER TLHYA	
16F1D	ɹ̥	MIAO LETTER DLHYA	
16F1E	ɹ̥	MIAO LETTER KA	• used for ga in Dry Yi
16F1F	ɹ̥	MIAO LETTER GA	
16F20	ɹ̥	MIAO LETTER YI KA	• used for ka in Dry Yi
16F21	ɹ̥	MIAO LETTER QA	
16F22	ɹ̥	MIAO LETTER QGA	
16F23	ɹ̥	MIAO LETTER NGA	
16F24	ɹ̥	MIAO LETTER NGHA	
16F25	ɹ̥	MIAO LETTER ARCHAIC NGA	• used in Pollard's early orthography
16F26	ɹ̥	MIAO LETTER HA	
16F27	ɹ̥	MIAO LETTER XA	• archaic character used in a 1949 S reformed orthography
16F28	ɹ̥	MIAO LETTER GHA	
16F29	ɹ̥	MIAO LETTER GHHA	
16F2A	ɹ̥	MIAO LETTER TSSA	
16F2B	ɹ̥	MIAO LETTER DZZA	
16F2C	ɹ̥	MIAO LETTER NYA	
16F2D	ɹ̥	MIAO LETTER NYHA	

16F2E	ɹ̥	MIAO LETTER TSHA	• used for dzha in Dry Yi
16F2F	ɹ̥	MIAO LETTER DZHA	
16F30	ɹ̥	MIAO LETTER YI TSHA	• used for tsha in Dry Yi
16F31	ɹ̥	MIAO LETTER YI DZHA	• used in Hei Yi
16F32	ɹ̥	MIAO LETTER REFORMED TSHA	• archaic character used in a 1949 S reformed orthography
16F33	ɹ̥	MIAO LETTER SHA	
16F34	ɹ̥	MIAO LETTER SSA	
16F35	ɹ̥	MIAO LETTER ZHA	• used in Black Yi
16F36	ɹ̥	MIAO LETTER ZSHA	• used in Black Yi
16F37	ɹ̥	MIAO LETTER TSA	• used for dza in Dry Yi
16F38	ɹ̥	MIAO LETTER DZA	
16F39	ɹ̥	MIAO LETTER YI TSA	• used for tsa in Dry Yi
16F3A	ɹ̥	MIAO LETTER SA	
16F3B	ɹ̥	MIAO LETTER ZA	
16F3C	ɹ̥	MIAO LETTER ZSA	• used in Black Yi
16F3D	ɹ̥	MIAO LETTER ZZA	
16F3E	ɹ̥	MIAO LETTER ZZSA	• used in Black Yi
16F3F	ɹ̥	MIAO LETTER ARCHAIC ZZA	• used in Pollard's early orthography
16F40	ɹ̥	MIAO LETTER ZZYA	• used in Black Yi
16F41	ɹ̥	MIAO LETTER ZZSYA	• used in Black Yi
16F42	ɹ̥	MIAO LETTER WA	
16F43	ɹ̥	MIAO LETTER AH	• glottal stop
16F44	ɹ̥	MIAO LETTER HHA	• used in Black Yi

Modifiers

16F50	ɹ̥	MIAO LETTER NASALIZATION	
16F51	ɹ̥	MIAO LETTER ASPIRATION	
16F52	ɹ̥	MIAO LETTER REFORMED VOICING	• archaic character used in a post-1949 reformed orthography
16F53	ɹ̥	MIAO LETTER REFORMED ASPIRATION	• archaic character used in a post-1949 reformed orthography





Vowels and finals

16F54	ɹ̥	MIAO LETTER A	
16F55	ɹ̥	MIAO LETTER AA	• used in Eastern Lisu
16F56	ɹ̥	MIAO LETTER AHH	• used in Hey Yi
16F57	ɹ̥	MIAO LETTER AN	
16F58	ɹ̥	MIAO LETTER ANG	• also used for aw
16F59	ɹ̥	MIAO LETTER O	
16F5A	ɹ̥	MIAO LETTER OO	
16F5B	ɹ̥	MIAO LETTER WO	• used in Hei Yi

16F5C	⦶	MIAO LETTER W
16F5D	⦶c	MIAO LETTER E
16F5E	⦶e	MIAO LETTER EN
16F5F	⦶e	MIAO LETTER ENG
16F60	⦶c	MIAO LETTER OEY
16F61	⦶n	MIAO LETTER I
16F62	⦶n	MIAO LETTER IA
16F63	⦶n _u	MIAO LETTER IAN
16F64	⦶n _u	MIAO LETTER IANG
		• also used for iaw
16F65	⦶o	MIAO LETTER IO
16F66	⦶z	MIAO LETTER IE
16F67	⦶m	MIAO LETTER II
		• used in Eastern Lisu
16F68	⦶n	MIAO LETTER IU
16F69	⦶n _u	MIAO LETTER ING
		• also used for in
16F6A	⦶u	MIAO LETTER U
16F6B	⦶r	MIAO LETTER UA
16F6C	⦶u _r	MIAO LETTER UAN
16F6D	⦶u _r	MIAO LETTER UANG
		• also used for uaw
16F6E	⦶w	MIAO LETTER UU
		• used in Eastern Lisu
16F6F	⦶u	MIAO LETTER UEI
16F70	⦶u _u	MIAO LETTER UNG
16F71	⦶s	MIAO LETTER Y
16F72	⦶s	MIAO LETTER YI
16F73	⦶p	MIAO LETTER AE
16F74	⦶c	MIAO LETTER AEE
		• used in Eastern Lisu
16F75	⦶r	MIAO LETTER ERR
16F76	⦶r	MIAO LETTER ROUNDED ERR
		• used in Eastern Lisu
16F77	⦶r	MIAO LETTER ER
16F78	⦶r	MIAO LETTER ROUNDED ER
		• used in Eastern Lisu
16F79	⦶i	MIAO LETTER AI
16F7A	⦶u	MIAO LETTER EI
16F7B	⦶ii	MIAO LETTER AU
16F7C	⦶q	MIAO LETTER OU
16F7D	⦶c	MIAO LETTER N
16F7E	⦶e	MIAO LETTER NG

Positioning tone marks

These are used to position the vowel off of the baseline position to indicate a changed tone.

16F8F		MIAO LETTER TONE RIGHT
16F90		MIAO LETTER TONE TOP RIGHT
16F91		MIAO LETTER TONE ABOVE
16F92		MIAO LETTER TONE BELOW

Baseline tone marks

These are used in Chuxiong Ahmao instead of the four above

16F93	τ	MIAO LETTER TONE-2
16F94	ς	MIAO LETTER TONE-3
16F95	η	MIAO LETTER TONE-4
16F96	Ϛ	MIAO LETTER TONE-5
16F97	ς	MIAO LETTER TONE-6
16F98	ι	MIAO LETTER TONE-7
16F99	τ	MIAO LETTER TONE-8

Archaic baseline tone marks

These are archaic characters used in a post-1949 reformed orthography

16F9A	Ϛ	MIAO LETTER REFORMED TONE-1
16F9B	Ϛ	MIAO LETTER REFORMED TONE-2
16F9C	Ϛ	MIAO LETTER REFORMED TONE-4
16F9D	Ϛ	MIAO LETTER REFORMED TONE-5
16F9E	Ϛ	MIAO LETTER REFORMED TONE-6
16F9F	Ϛ	MIAO LETTER REFORMED TONE-8

A. Administrative

1. Title

Proposal for encoding additional Miao characters in the SMP of the UCS

2. Requester's name

Michael Everson and Erich Fickle

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

Individual contribution.

4. Submission date

2010-08-08

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.

1b. Proposed name of script

1c. The proposal is for addition of character(s) to an existing block

Yes.

1d. Name of the existing block

Miao.

2. Number of characters in proposal

5.

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?

Yes.

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

Yes.

7. 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.

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 <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database <http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

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?

Yunnan Minority Language Commission.

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?

Yes.

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

Common.

4b. Reference

Scriptures, dictionaries, song books, textbooks, signs

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

Yes.

5b. If YES, where?

In Northeastern Yunnan, China.

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

No.

6b. If YES, is a rationale provided?

6c. If YES, reference

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

Yes.

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?

No.

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?

Yes.

11b. If YES, is a rationale for such use provided?

No.

11c. If YES, reference

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?