Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type:	**Working Group Document**
Title:	**Preliminary proposal for additions for Hei Yi to Miao block**
Source:	**Eiso Chan (**陈永聪**)**
Status:	**Individual Contribution**
Action:	**For consideration by WG2 & UTC**
Date:	**2022-05-04**

0. Introduction

In WG2 N3877 (L2/10-302) and N4845 (L2/17-345), Michael Everson, Erich Fickle and Adrian Cheuk proposed several letters used for Hei Yi, but when I checked the current Hei Yi textbook, some characters are not used currently and some characters have not been encoded yet. In WG2 N3669R (L2/09-253R), China NB proposed a voiced mark as U+16F28 as Fig. 1 shows, but this one was cancelled in WG2 N3789 (L2/10-093). I try to clarify these issues in this document.

1. Proposal

1.1. MIAO LETTER TSHA with inside dot

In Fig. 2, Miao Letter TSHA with inside dot which is after U+16F0C **T** is not encoded separately in the current code chart. In fact, this one has been showed in the evidence in WG2 N3877. Please see Figs. 3 and 4 in this document. In Section 2 in WG2 N3877, it looks this one was treated the same as Miao Letter TSHA with upper right voiced dot in the context. Please see Fig. 5. In Illustration 2 in WG2 N3669R, the Miao letters with upper right voiced dot are used in Miao not Hei Yi in Chuxiong (楚雄). Hei Yi is used in Luquan (禄劝) which is a county under Chuxiong before 1983. Now, Luquan is a county under Kunming (昆明). Miao is also used in Luquan, which the full name is Luquan Yi and Miao Autonomous County (禄劝彝族苗族自治县).

In WG2 N3789, the submitters treated the letters with the upper right voiced dot as the unifiable variants of the letters with the warts. Please see Fig. 7. Miao Letter Tsha with inside dot is an official letter in current Hei Yi orthography, so it's better to use a sequence to process it.

Table 1.1 Suggestion on MIAO LETTER TSHA WITH INSIDE DOT

Glyph	SVS	Sequence Name	Remark
	<U+16F2F U+FE00>	MIAO LETTER TSHA WITH INSIDE DOT	used in Hei Yi

Please also see Figs. 8, 10, 11, 12, 14.

1.2. MIAO LETTER YI ZHA

Table 1.2 Suggestion on MIAO LETTER YI ZHA

Glyph	Code Point	Character Name	Remark
	U+16F4B	MIAO LETTER YI ZHA	used in Hei Yi

In the 48 Hei Yi initials list as Fig. 2 shows. The last one is not found in the current code chart. As the local people told to me, the phonetic value is similar to U+16F35, but this one is only used in Hei Yi, so the character name should be MIAO LETTER YI ZHA. U+16F35 which the shape is the same as "R" is used in Miao in Chuxiong according to WG2 N3669R.

Please also see Figs. 9, 10, 13.

1.3. Other issue

When I check so many materials on the Miao script, I found that the current U+16F33 glyph doesn't match the sources in Miao, Hei Yi and so on.

Table 1.3 U+16F33 glyph

Current Glyph	Hei Yi Glyph

I don't request to do anything at this time, and I will do more research on this issue for the future document. On the other hand, the punctuations used with the Miao script should be clarified in future.

2. Unicode Character Properties

2.1. MIAO LETTER TSHA with inside dot

The SVS information should be added to StandardizedVariants.txt as below.

```
# Miao


16F2F FE00; Hei Yi style; #  MIAO LETTER DZHA
```

2.2. MIAO LETTER YI ZHA

The Unicode character properties of MIAO LETTER YI ZHA follow other Miao consonants.

The property for UnicodeData.txt is shown as below.

```
16F4B;MIAO LETTER YI ZHA;Lo;0;L;;;;;N;;;;;


Script=Miao
```

3. Pictures

Combining marks

16F26	(◌	MIAO MARK NASALIZATION
16F27	◌'	MIAO MARK ASPIRATED
16F28	◌˙	MIAO MARK VOICED

Fig. 1 Combining marks in WG2 N3669
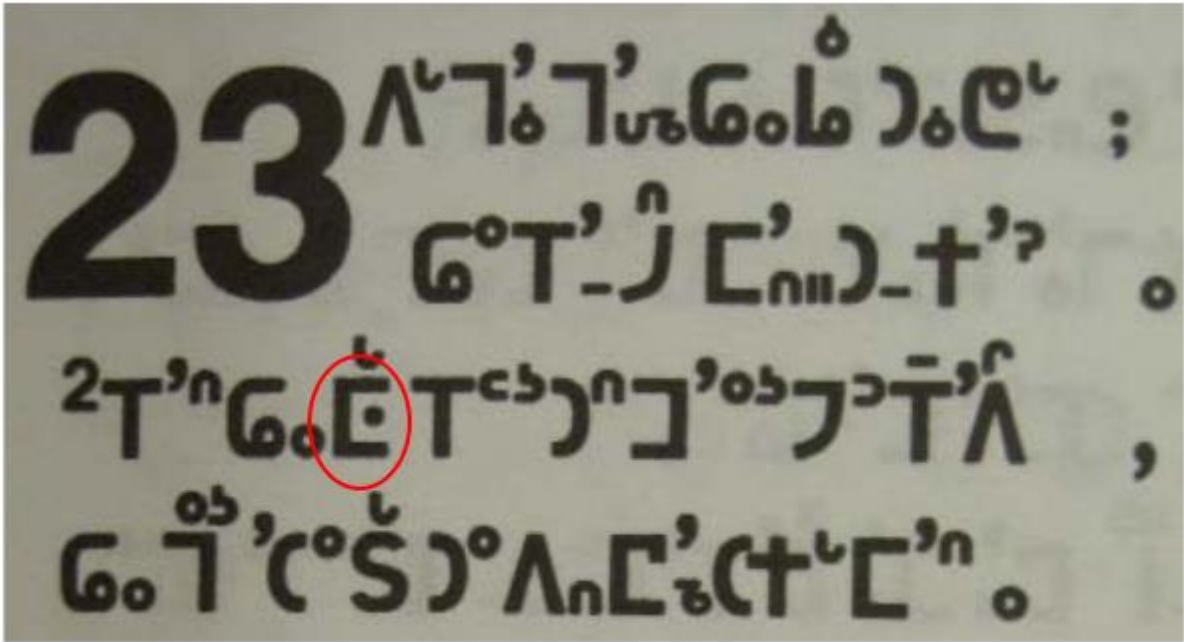

Fig. 2 «Ɛ˙ S° Sᵘ ʔ Sᵘ» (《柏格理黑彝文识字教材》), p. 75

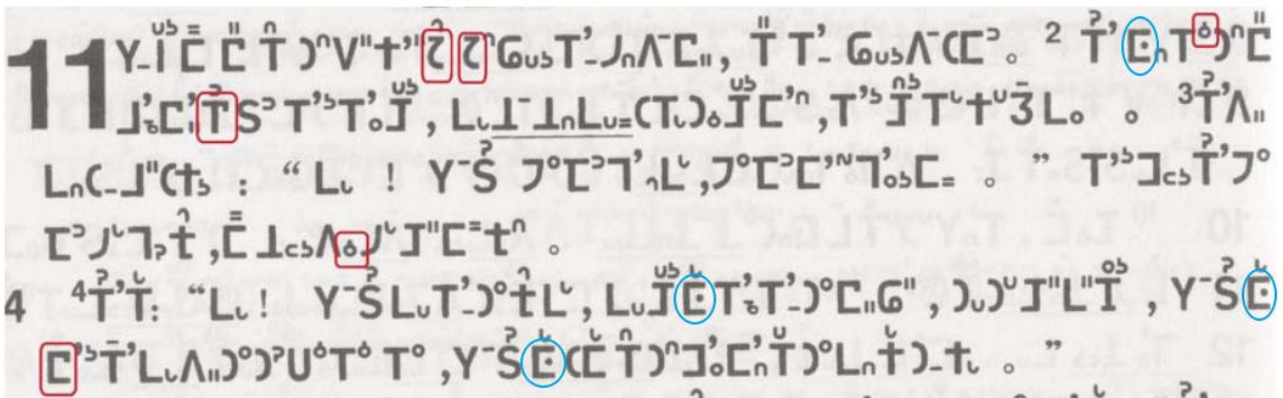*Fig. 3 Figure 1 in WG2 N3877*

Fig. 4 Figure 2 in WG2 N3877

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 ballotting U+16F2E Ꞓ YI TSHA and the proposed U+16F31 Ꞓ YI DZHA are used. Three consonants with this "serif-like mark" are proposed here.

Fig. 5 Section 2 in WG2 N3877

Fig. 6 Illustration 2 in WG2 N3669R

Fig. 7 Section 5 in WG2 N3789

Fig. 8 «ꊭ Sˏ Sꭒ ꌦ Sꭒ» (《柏格理黑彝文识字教材》), p. 63

Fig. 9 «ꊭ Sˏ Sꭒ ꌦ Sꭒ» (《柏格理黑彝文识字教材》), p. 72

5
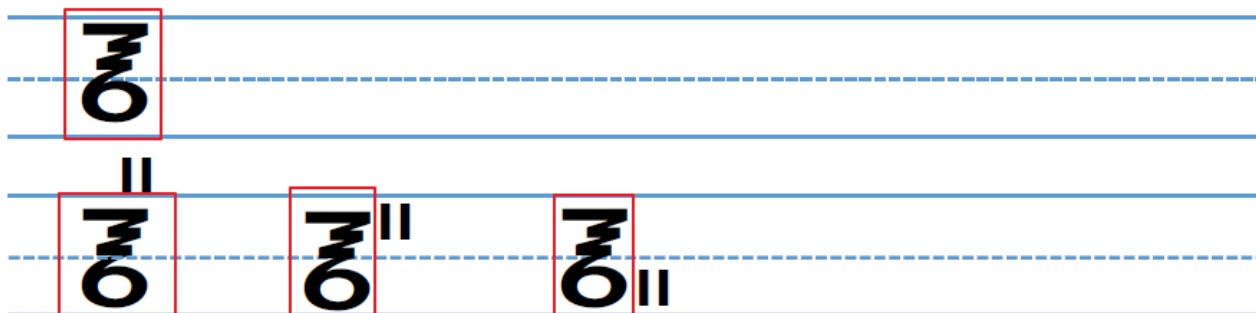
第十三课　　　　要寻找重要的事

Fig. 10 «...» (《柏格理黑彝文识字教材》), *p. 71*

11-12

Fig. 11 «ＳʔⅠ˅ ̈ＴＳ˅» (《頌主聖歌》)
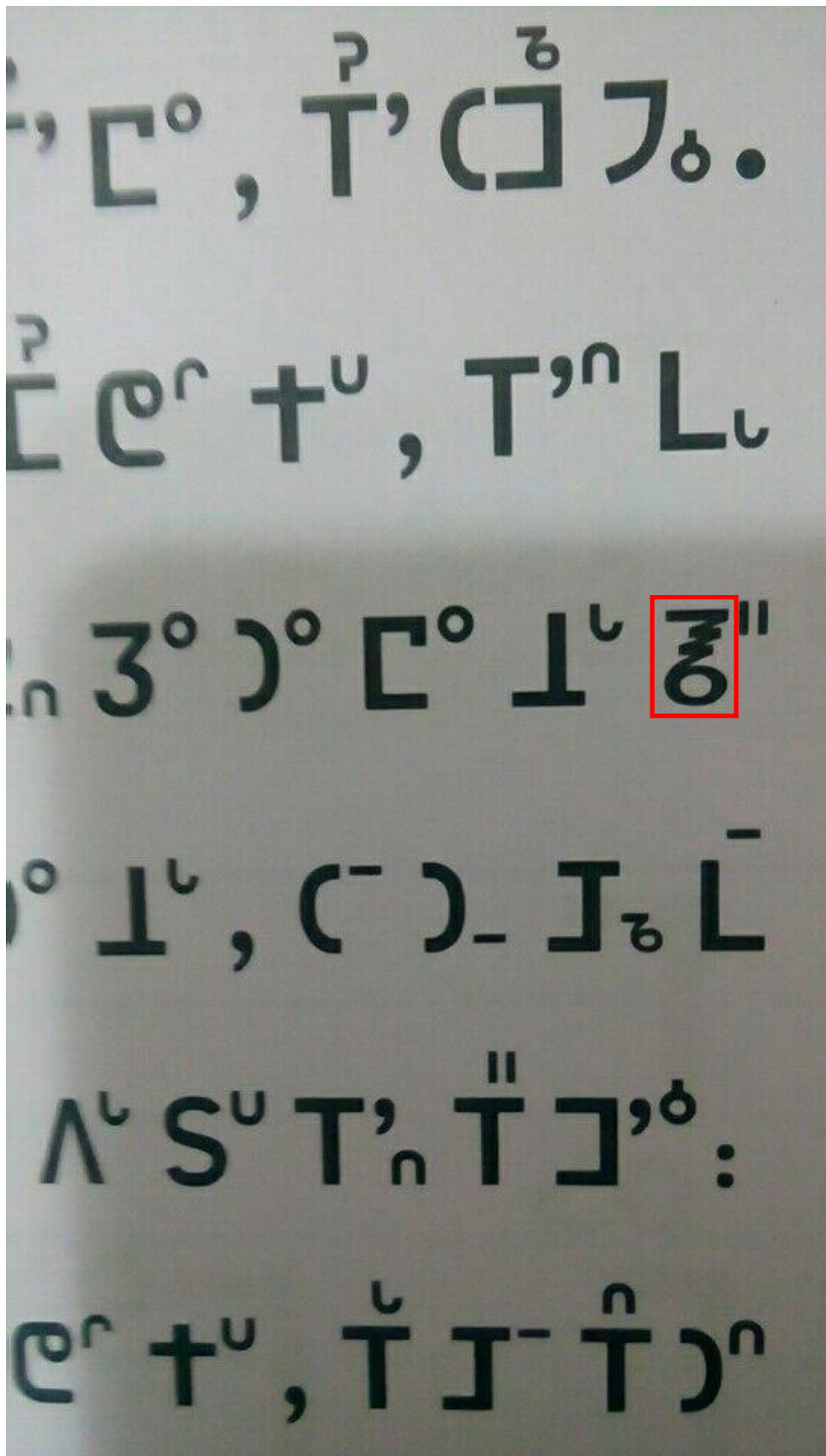
Fig. 12 «S꞉⊥ᵛ ẗ Sᵛ» (《頌主聖歌》)
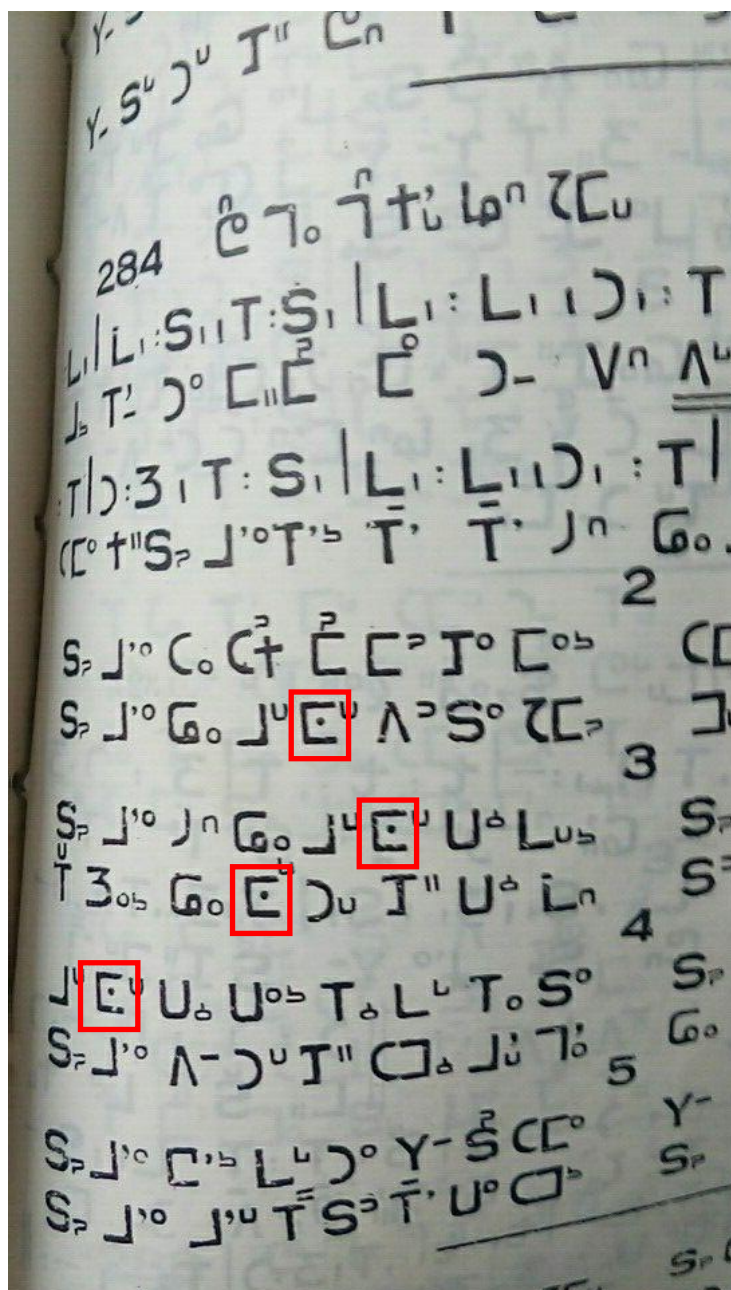
Fig. 13 «Sₔ⅃ᵛ ťSᵛ» (《頌主聖歌》)
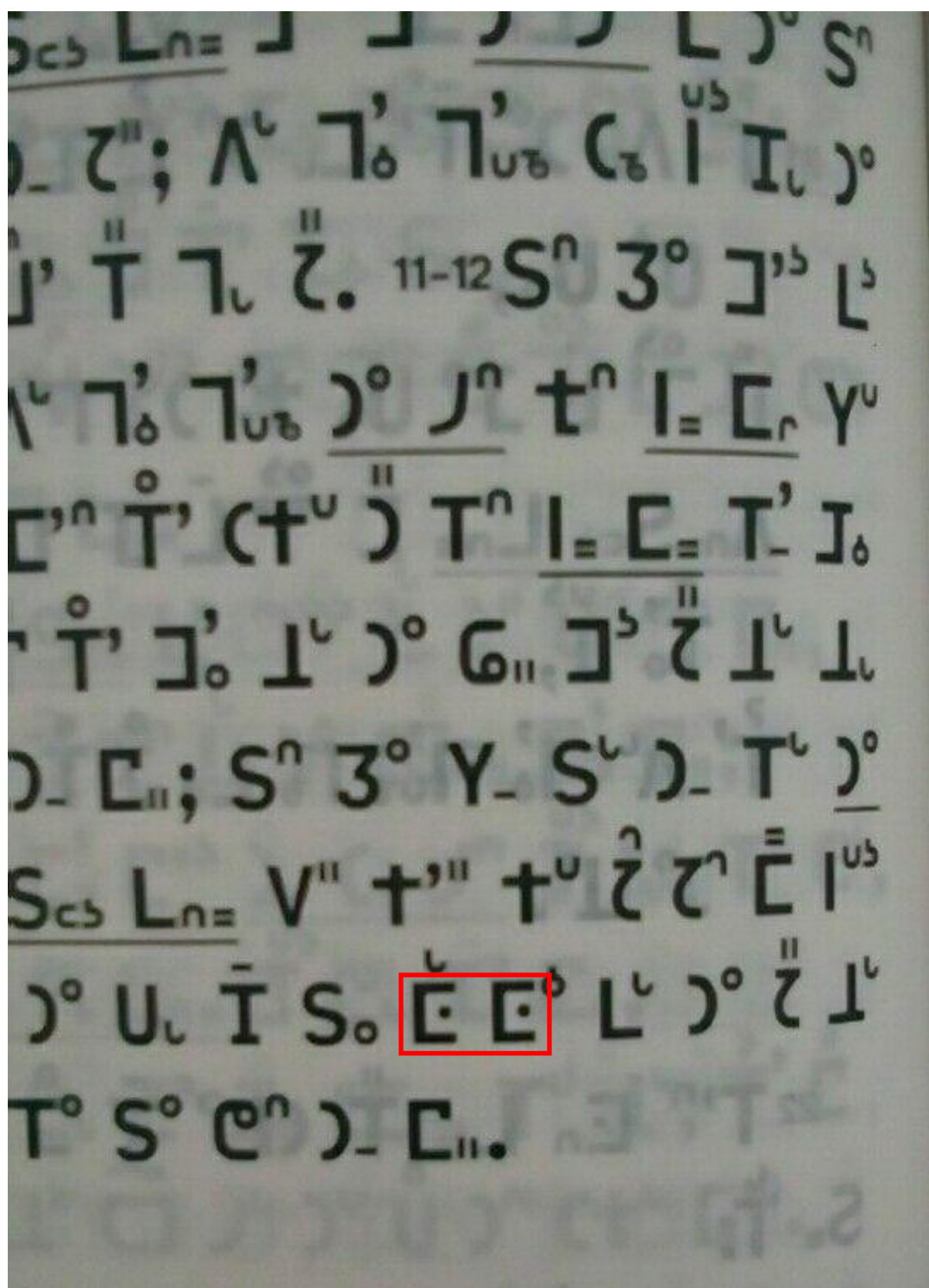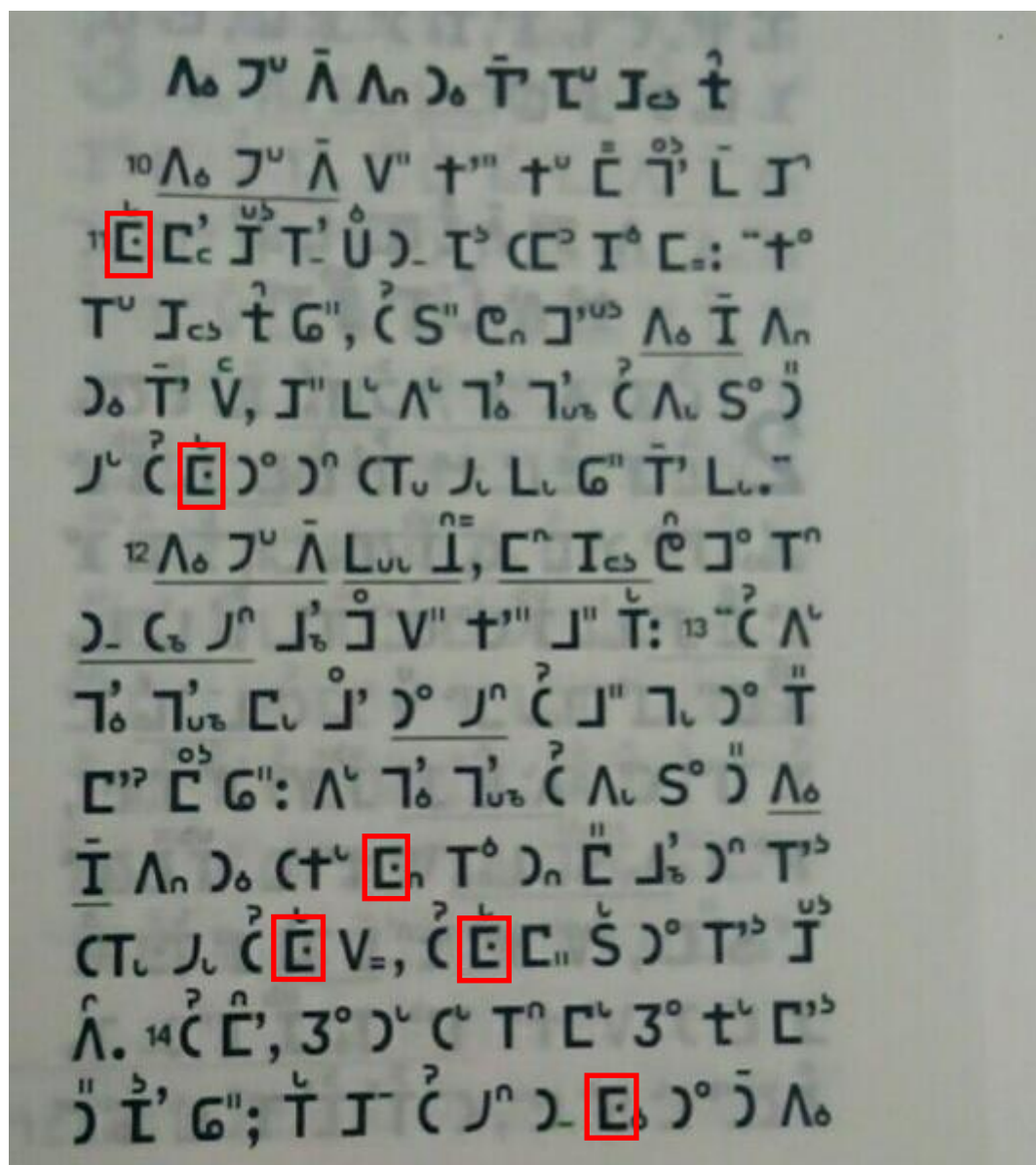
Fig. 14 «**S꞉Ꞁᵛ ꞇ̈ Sᵛ**» (《頌主聖歌》)

(End of Document)

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://std.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://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html **.**
See also http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps*.**

A. Administrative

1. **Title:**	***Preliminary additions for Hei Yi to Miao block***
2. Requester's name:	*Eiso CHAN*
3. Requester type (Member body/Liaison/Individual contribution):	*Individual contribution*
4. Submission date:	*2022-05-04*
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):   *NO*
       Proposed name of script:
   b. The proposal is for addition of character(s) to an existing block: *YES*
       Name of the existing block: *Miao*
2. Number of characters in proposal: *1*
3. Proposed category (select one from below - see section 2.2 of P&P document):

A-Contemporary	X	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		

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?
        Eiso CHAN
   b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
        Eiso CHAN, eisoch@126.com
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)?   *NO*

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 Unicode Character Database

---

[1]. Form number: N4502-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, 2009-11, 2011-03, 2012-01)

(http://www.unicode.org/reports/tr44/) 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?	*Language experts from Hei Yi*
If YES, available relevant documents:	*All the evidences in this document are provided by the language experts from Hei Yi.*
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?	*YES*
Reference:	*Ethnologue*
4. The context of use for the proposed characters (type of use; common or rare)	*Common*
Reference:	*Textbooks and scriptures*
5. Are the proposed characters in current use by the user community?	YES
If YES, where? Reference:	*Luquan Yi and Miao Autonomous County, Kunming City, Yunnan Province, PRC*
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, or could be confused with, 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?	*NO*
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 characters?	*NO*
If YES, are the equivalent corresponding unified ideographic characters identified?	
If YES, reference:	