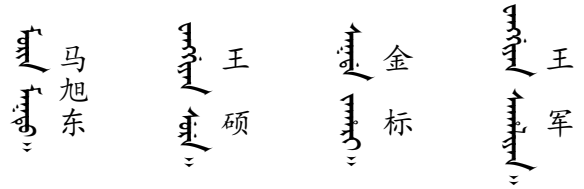
Proposal to Encode One Sibe Letter and One Mongolian Format Control, and Add Ten Variants to Three Mongolian Letters

MA Xudong (xudong.ma@gmail.com) – *MORIN Maktu*
WANG Shuo (150374704@qq.com) – *WANGGIYAN Sure*
JIN Biao (63316836@qq.com) – *SAKDA Jahari*
WANG Jun (sahalin@qq.com) – *WANGGIYAN Sahaliyan*



1. Proposed Characters

Two characters are proposed in order to present Sibe correctly.

Code Point	Proposed Character Name	Representative Glyph
1878	MONGOLIAN LETTER SIBE UUE	
181A	MONGOLIAN FREE VARIATION SELECTOR FOUR	FV S4

2. Character Properties

General Properties

```
1878;MONGOLIAN LETTER SIBE UUE;Lo;0;L;;;;;N;;;;;
181A;MONGOLIAN FREE VARIATION SELECTOR FOUR;MN;0;NSM;;;;;N;;;;;
```

Line Break Properties

```
1878;AL           # Lo           MONGOLIAN LETTER SIBE UUE
181A;CM           # Mn           MONGOLIAN FREE VARIATION SELECTOR FOUR
```

Shaping Behavior – Part 1

	Isolate Form	Initial Form	Medial Form	Medial Form 1	Medial Form 2	Final Form	Final Form 1	Final Form 2	Final Form 3	Final Form 4
U+1878										

1878		first form (isolate)
1878 ZWJ		first form (initial)
ZWJ 1878 ZWJ		first form (medial)
ZWJ 1878 FVS1 ZWJ		second form (medial)
ZWJ 1878 FVS2 ZWJ		third form (medial)

ZWJ 1878	first form (final)
ZWJ 1878 FVS1	second form (final)
ZWJ 1878 FVS2	third form (final)
ZWJ 1878 FVS3	fourth form (final)
ZWJ 1878 FVS4	fifth form (final)

This is the fifth and also the last vowel letter of the Sibe alphabet[1]. It is in fact the combination of the fifth vowel letter ᠥ *u* (U+1860 MONGOLIAN LETTER SIBE UE) and the sixth vowel letter ᠤ *v* (U+1861 MONGOLIAN LETTER SIBE U) of the Manchu alphabet without the isolate and the initial forms of the latter. It makes the usage of the Sibe system of five vowel letters and the consonant letters *k*, *g*, *h* and *k‘*, *g‘*, *h‘* in the Sibe way possible. See “**3. Details**” below.

It has five distinct final forms, so Mongolian Free Variation Selector Four (abbreviated FVS4) is necessary and should be encoded. (FVS4 is also necessary for Manchu. See our future proposals.)

Shaping Behavior – Part 2

	Isolate Form	Isolate Form 1	Initial Form	Initial Form 1	Medial Form	Medial Form 1	Final Form
U+183A	ᠠ	ᠠ	ᠠ	ᠠ	ᠠ	ᠠ	ᡆ
U+186C	ᠡ᠂	ᠡ᠂	ᠡ᠂	ᠡ᠂	ᠡ᠂	ᠡ᠂	ᡇ᠂
U+186D	ᠣ᠊	ᠣ᠊	ᠣ᠊	ᠣ᠊	ᠣ᠊	ᠣ᠊	

183A	first form (isolate)
183A FVS1	second form (isolate)
183A ZWJ	first form (initial)
183A FVS1 ZWJ	second form (initial)
ZWJ 183A ZWJ	first form (medial)
ZWJ 183A FVS1 ZWJ	second form (medial)
ZWJ 183A ZWJ	first form (final)
186C	first form (isolate)
186C FVS1	second form (isolate)
186C ZWJ	first form (initial)
186C FVS1 ZWJ	second form (initial)
ZWJ 186C ZWJ	first form (medial)

[1] Abkai Transliteration is used for Manchu and Sibe to make the descriptions easier. The following is the concise comparison of Abkai Transliteration, the Möllendorff transliteration system, and the transliteration of *A Comprehensive Manchu-Chinese Dictionary* (CMCD). Note that the transliterations for Manchu and for Sibe are a bit different.

Abkai:	v	k‘	g‘	h‘	q	x	z	c	r“	jy‘	qy‘	zi	cy‘
Möllendorff:	ū	kʿ	gʿ	hʿ	c	š	dz	tsʿ	ž	jy	cʿy	dzi	ts
CMCD:	uu	kk	gg	hh	ch	sh	z	c	rr	zhy	chy	zy	cy

ZWJ 186C FVS1 ZWJ	ᢎ	second form (medial)
ZWJ 186C ZWJ	ᢏ	first form (final)
186D	ᢐ	first form (isolate)
186D FVS1	ᢑ	second form (isolate)
186D ZWJ	ᢐ	first form (initial)
186D FVS1 ZWJ	ᢑ	second form (initial)
ZWJ 186D ZWJ	ᢐ	first form (medial)
ZWJ 186D FVS1 ZWJ	ᢐ	second form (medial)

The cells and lines highlighted in yellow are proposed to add to the current Unicode standard. These forms are only used in Sibe.

The cell and line highlighted in gray are also proposed to add to the current Unicode standard. The final form of U+186C is only used in Manchu Ali Gali.

U+183A, U+186C and U+186D have no final form in Sibe. The final form of U+183A is used only in Mongolian so *baiti* is used to display the form. U+186D has no final form at all.

3. Details

The Sibe alphabet was converted from the Manchu alphabet with a few modifications and additions. In Sibe, the former sixth vowel letter ᡶ *v* (U+1861 MONGOLIAN LETTER SIBE U) of the Manchu alphabet is rarely used except after ᡴ *k*, ᡤ *g* or ᡥ *h*, and when used after ᡴ *k*, ᡤ *g* or ᡥ *h*, it shares the same pronunciation with the fifth vowel letter ᡠ *u* (U+1860 MONGOLIAN LETTER SIBE UE). The letter ᡶ *v* itself is removed from the Sibe alphabet but the medial and final forms persist and are considered as variants of the fifth vowel letter ᡠ *u*. Thus, there are five vowel letters in the Sibe alphabet, and there are six vowel letters in the Manchu alphabet.

In Manchu, the letters *k*, *g* and *h* in ᡴᡝ *ke*/ᡴᡳ *ki*/ᡴᡠ *ku*, ᡤᡝ *ge*/ᡤᡳ *gi*/ᡤᡠ *gu* and ᡥᡝ *he*/ᡥᡳ *hi*/ᡥᡠ *hu* are written and pronounced differently as in ᡴᠠ *ka*/ᡴᠣ *ko*/ᡴᡶ *kv*, ᡤᠠ *ga*/ᡤᠣ *go*/ᡤᡶ *gv* and ᡥᠠ *ha*/ᡥᠣ *ho*/ᡥᡶ *hv*, while they are pronounced the same as the letters ᡴ *k‘*, ᡤ *g‘* and ᡥ *h‘* in ᡴ *k‘a*/ᡠ *k‘o*, ᡤ *g‘a*/ᡠ *g‘o* and ᡥ *h‘a*/ᡠ *h‘o*.

In Sibe, the former Manchu letters ᡴ *k*, ᡤ *g* and ᡥ *h* in ᡴᡝ *ke*/ᡴᡳ *ki*/ᡴᡠ *ku*, ᡤᡝ *ge*/ᡤᡳ *gi*/ᡤᡠ *gu* and ᡥᡝ *he*/ᡥᡳ *hi*/ᡥᡠ *hu* are considered the same as the Manchu letters ᡴ *k‘*, ᡤ *g‘* and ᡥ *h‘* in ᡴ *k‘a*/ᡠ *k‘o*, ᡤ *g‘a*/ᡠ *g‘o* and ᡥ *h‘a*/ᡠ *h‘o* according to the pronunciations, so the Manchu letters ᡴ *k*, ᡤ *g* and ᡥ *h* are the variants of the Sibe letters *k‘*, *g‘* and *h‘*. At the same time, the Manchu syllables ᡶ *kv*, ᡶ *gv* and ᡶ *hv* are the Sibe syllables ᡶ *ku*, ᡶ *gu* and ᡶ *hu*.

Vow. / Con.	U+1820		U+185D	U+1873 / U+185E	U+1823		U+1860 / U+1878	U+1861 / U+1878	Manchu / Sibe
	ᠠ	ᡡ	ᡢ	ᡣ	ᡤ	ᡥ	ᡠ	ᡶ	
U+1863	*ka*		*ke*	*ki*	*ko*		*ku*	*kv*	Manchu
U+183A		*k‘a*				*k‘o*			
U+1863	*ka*		~~*ke*~~	~~*ki*~~	*ko*		~~*ku*~~	~~*kv*~~ *ku*	Sibe
U+183A		*k‘a*	*k‘e*	*k‘i*		*k‘o*	*k‘u*		

	ᡤᠠ	ᡤᡝ	ᡤᡳ	ᡤᠣ	ᡤᡠ	ᡤᡡ	
U+1864	*ga*		*ge*	*gi*	*go*	*gu*	*gv*	Manchu
U+186C		*g‘a*				*g‘o*		
U+1864	*ga*		~~*ge*~~	~~*gi*~~	*go*	~~*gu*~~	~~*gv*~~ *gu*	Sibe
U+186C		*g‘a*	*g‘e*	*g‘i*		*g‘o*	*g‘u*	
	ᡥᠠ	ᡥᡝ	ᡥᡳ	ᡥᠣ	ᡥᡠ	ᡥᡡ	
U+1865	*ha*		*he*	*hi*	*ho*	*hu*	*hv*	Manchu
U+186D		*h‘a*				*h‘o*		
U+1865	*ha*		~~*he*~~	~~*hi*~~	*ho*	~~*hu*~~	~~*hv*~~ *hu*	Sibe
U+186D		*h‘a*	*h‘e*	*h‘i*		*h‘o*	*h‘u*	

The strikethrough syllables should be replaced by the ones highlighted in yellow.

4. Attestations

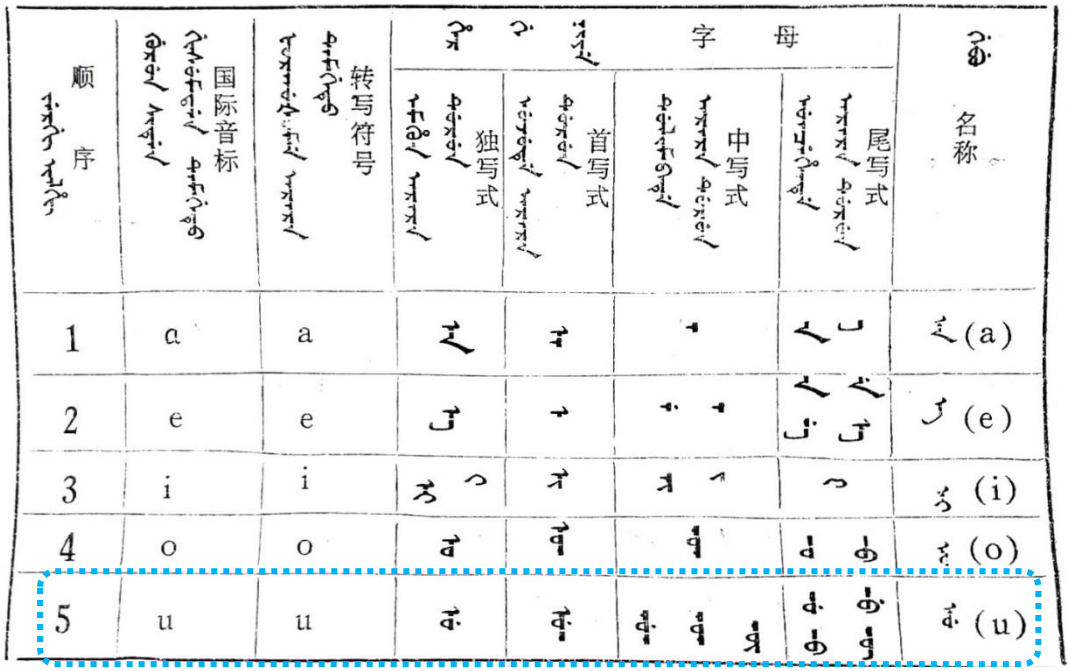
4.1 ᠨᡝᡳ ᡦᡠᠨ ᠰᡳᠪᡝ ᡧᡠ ᡨᠠᡴᡳᠨ ᡤᡳᠰᡠᠨ-ᡳ ᠠᡵᠠᡵᠠ ᡴᡠᡵᡳ

(***Nei Fon Sibe Xu Taqin G‘isun-i Arara Kouli***)

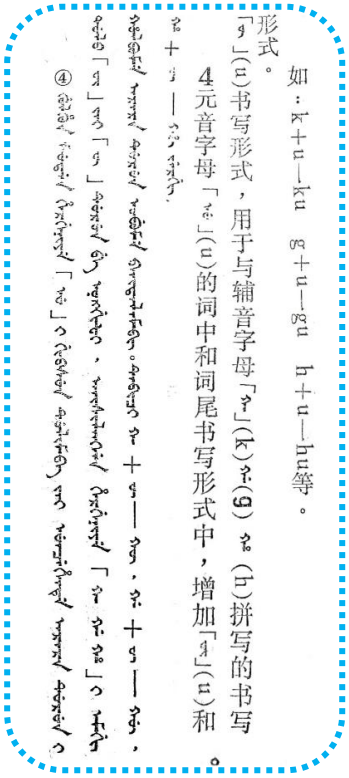
《现代锡伯文学语言正字法》

(***Xiàndài Xībó Wénxúe Yǔyán Zhèngzìfǎ***)

Modern Orthography of the Literary Language of Sibe



p. 2



如：k+u—ku　g+u—gu　h+u—hu等。

形式。

「ᡤᡠ」(u)书写形式，用于与辅音字母「ᡴ」(k)ᡤ(g) ᡥ (h)拼写的书写

4.元音字母「ᡠᡝ」(u)的词中和词尾书写形式中，增加「ᡤ」(u)和

ᡥᡠ + ᡠ — ᡥᡠ ᡳᡵᡝᡵᡝᠮᡝᡵᡝ。

ᡴᡠᡵᡳᡵᡝᡵᡝᠮᡝᡵᡝ ᡳᡵᡝᡵᡝᠮᡝᡵᡝ ᡴᡠᡵᡳᡵᡝᡵᡝᠮᡝᡵᡝ ᡴᡠᡵᡳᡵᡝᡵᡝᠮᡝᡵᡝ。ᡴᡠᡵᡳᡵᡝᡵᡝᠮᡝᡵᡝ ᡥᡠ + ᡠ — ᡥᡠ，ᡴᡠ + ᡠ — ᡴᡠ，

ᡴᡠᡵᡳᡵᡝᡵᡝᠮᡝᡵᡝ「ᡠᡝ」ᡳᡵᡝᡵᡝᠮᡝᡵᡝ「ᡠ」ᡴᡠᡵᡳᡵᡝᡵᡝᠮᡝᡵᡝ ᡴᡠᡵᡳᡵᡝᡵᡝᠮᡝᡵᡝ「ᡥᡠ ᡴᡠ ᡥᡠ」ᡳᡵᡝᡵᡝᠮᡝᡵᡝ

④ ᡴᡠᡵᡳᡵᡝᡵᡝᠮᡝᡵᡝ ᡴᡠᡵᡳᡵᡝᡵᡝᠮᡝᡵᡝ ᡴᡠᡵᡳᡵᡝᡵᡝᠮᡝᡵᡝ ᡴᡠᡵᡳᡵᡝᡵᡝᠮᡝᡵᡝ「ᡠᡝ」ᡳᡵᡝᡵᡝᠮᡝᡵᡝ

pp. 7-8

4.2 ᡥᡳᡵᡝᠮᡝ ᡥᡝᡵᡝᠨ ᡤᡠᡳᠰᡠᠨ ᡴᡝᠮᡴᡳᡥᡝ ᡥᡝᡵᡝᡴᡝᠨ ᡝᡳᠪᡳᡥᡝ ᠪᡳᡵᡝᡴᡠ ᠪᡳᡨᡥᡝ

(***Nikan Sibe G'isun Kamqiha Xolokon Bulek'u Bith'e***)

《汉锡简明对照词典》

(***Hàn-Xī Jiǎnmíng Duìzhào Cídiǎn***)

A Concise Chinese-Sibe Bilingual Dictionary

p. 11

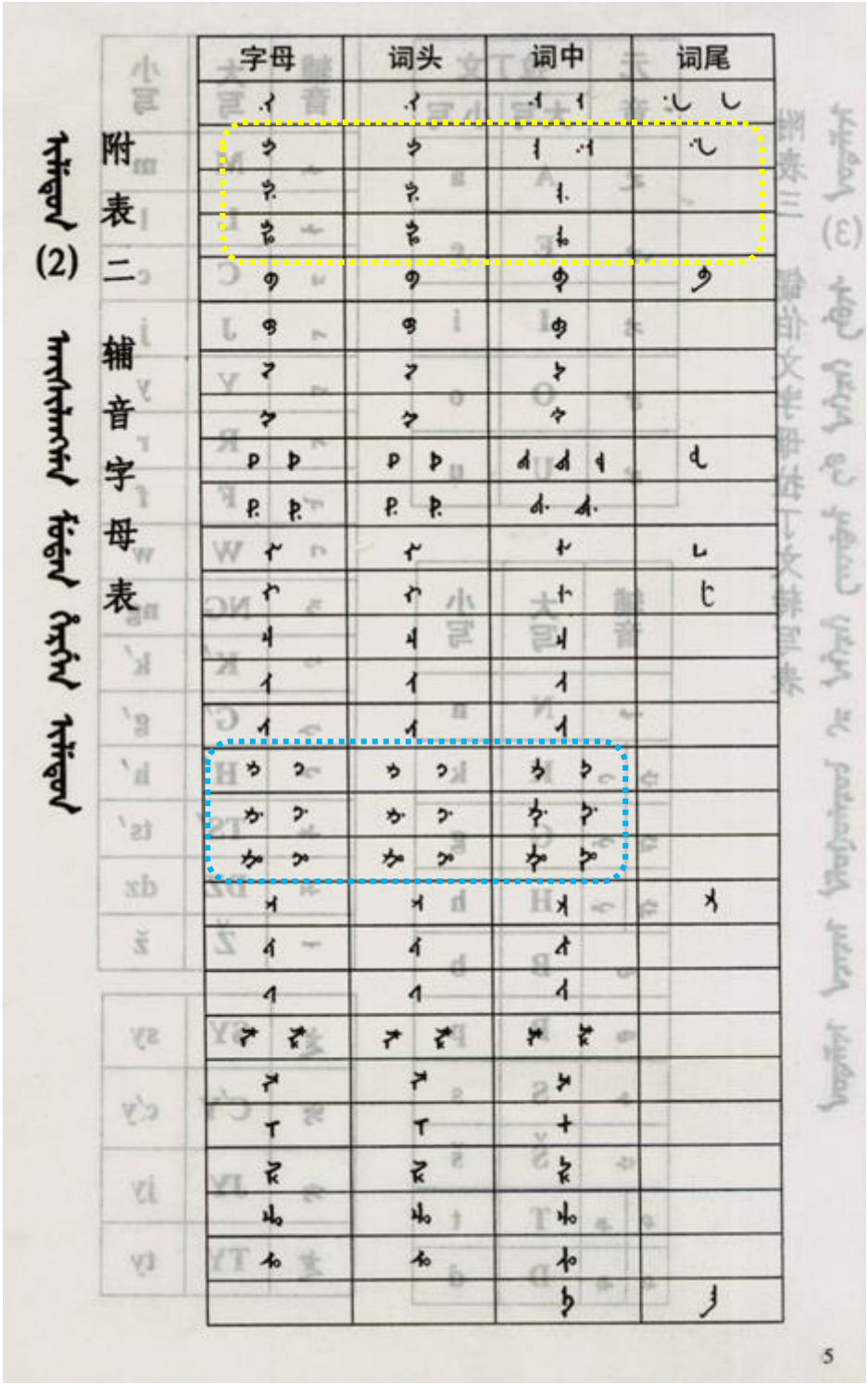
4.3 ᠵᠢᠭᠢᠨ᠎ᠡ ᠲᠠᠢᠬᠤᠸᠠ ᠲᠠᠢᠪᠢᠭᠤᠯ᠂ ᠨᠢᠶᠠᠮᠠᠩᠭᠠ ᠭᠢᠰᠤᠨ᠂ ᠲᠠᠢᠪᠢᠭᠤᠯᠤᠢ ᠲᠠᠢᠪᠢᠷᠢᠳᠡ ᠪᠠᠢᠲᠠᠯᠠᠷᠠ ᠪᠢᠲᠬᠡ᠂ ᠳᠡᠷᠭᠢ ᠳᠡᠪᠲᠡᠯᠢᠨ

(*Ajig'e Taqikui Taqibqen,* ***Niyamangga G'isun, Taqibukui Taqiburede Baitalara Bith'e****, Derg'i Debtelin*)

《义务教育课程标准试验教科书（试用）**语文（锡伯文）教师教学用书**·上册》

(***Yǔwén [Xībóweń] Jiàoshī-yòng Shū****, Shàngcè*)

Textbook of the Sibe language for primary schools, Teachers' book, Book 1, 2006



p. 5

p. 7

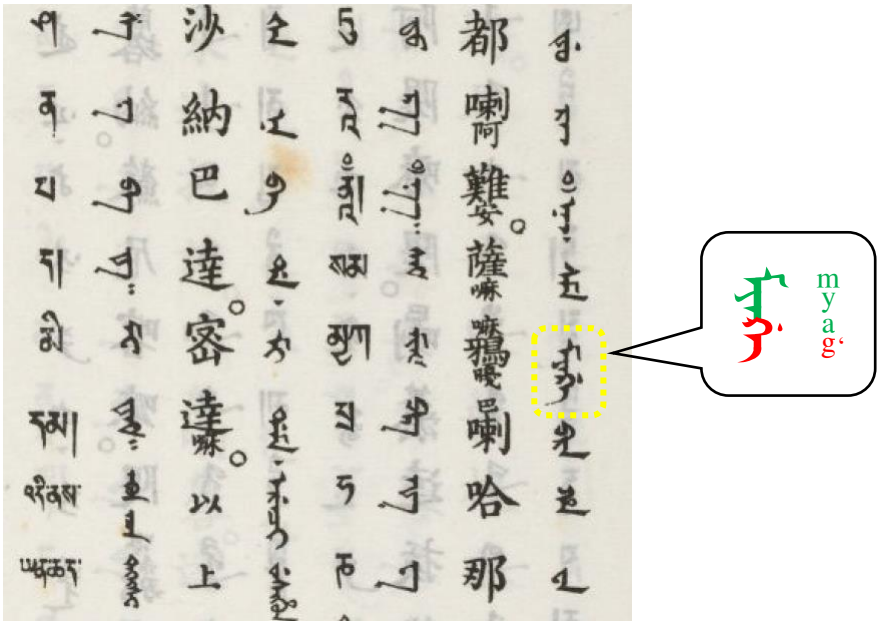
4.4 ᠨᠢᠷᠤ ᠣ ᠠᠷᠠᠪᠠ ᠮᠠᠨᠵᠤ ᠨᠢᠭᠠᠨ ᠮᠣᠩᠭᠣ ᠲᠠᠩᠭᠤᠲ ᠡᠷᠬᠡᠨ ᠤ ᠬᠠᠮᠭᠢᠬᠠ ᠶᠡᠬᠡ ᠭᠠᠩᠵᠤᠷ ᠤ ᠨᠣᠮᠤᠨ ᠤ ᠦᠭᠡᠷᠢ ᠲᠠᠷᠨᠢ᠃ ᠦᠵᠦᠢ ᠳᠣᠪᠲᠣᠨ ᠤ ᠳᠤᠢᠴᠢ ᠳᠡᠪᠲᠡᠯᠢᠨ᠃ ᠰᠤᠨᠵᠠᠭᠴᠢ ᠲᠠᠷᠨᠢ᠃

(***Han-i Araha Manju Nikan Monggo Tanggvt Hergen-i Kamqiha Amba G'anjur Nomun-i Uheri Tarni,*** *Ujui Dobton-i Duiqi Debtelin, Sunjaqi Tarni.*)

《**御制满汉蒙古西番合璧大藏全咒**・第一套第四卷・第五咒》

(***Yùzhì Mǎn-Hàn-Měnggǔ-Xīfān Hébì Dàzàng Quánzhòu,*** *Dì-Yī Tào Dì-Sì Juàn, Dì-Wǔ Zhòu*)

The Authorized Manchu-Chinese-Mongolian-Tibetan All Mantras in Mahāpiṭaka, Set 1, Volume 4, Mantra 5



p. 90

More examples can be found in this book.

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:** ***Proposal to encode one Sibe letter and one Mongolian format control, and add ten variants to three Mongolian letters***
2. Requester's name: *MA Xudong (xudong.ma@gmail.com)*
WANG Shuo (150374704@qq.com)
JIN Biao (63316836@qq.com)
WANG Jun (sahalin@qq.com)
3. Requester type (Member body/Liaison/Individual contribution): *Individual contribution*
4. Submission date: *2017-03-03*
5. Requester's reference (if applicable): *N/A*
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: *MONGOLIAN*
2. Number of characters in proposal: *1*
3. Proposed category (select one from below - see section 2.2 of P&P document):

A-Contemporary		B.1-Specialized (small collection)	X	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?
MA Xudong
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
6. References:
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? *NO*
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)? *YES*
See Further Information

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 (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.

[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)

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?
   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? *NO*
   Reference:
4. The context of use for the proposed characters (type of use; common or rare) *COMMON*
   Reference:
5. Are the proposed characters in current use by the user community? *NO*
   If YES, where? Reference:
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? *YES*
   If YES, is a rationale provided? *NO*
   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? *YES*
    If YES, is a rationale for its inclusion provided? *YES*
    If YES, reference: *See Further Information*
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?
    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:

(EOF)