ISO/IEC JTC1/SC2/WG2 N4414
L2/13-069
2013-04-22

Title: Revised Proposal to Encode the Soyombo Script in ISO/IEC 10646
Source: Script Encoding Initiative (SEI)
Author: Anshuman Pandey (pandey@umich.edu)
Status: Liaison Contribution
Action: For consideration by UTC and WG2
Date: 2013-04-22

1 Introduction

A request to include Soyombo in the Universal Character Set (ISO/IEC 10646) was made by the Mongolia and Japan national bodies in September 1998 in the document WG2 N1855 L2/98-358. An update on the request was provided in January 2000 by Takayuki K. Sato (Japan), who stated that the project for encoding Soyombo had stalled due to funding issues (WG2 N2163 L2/00-055). The script was allocated to the Roadmap to the Supplementary Multilingual Plane (SMP) in WG2 Meeting 38 in March 2000 (WG2 N2203 L2/00-234). There was no further action. The present effort aims to fulfill the original request. This proposal supersedes the following documents:

- N3949 L2/10-399: Preliminary Proposal to Encode the Soyombo Script in ISO/IEC 10646
- N3986 L2/11-054: Determining the Encoding Model for Soyombo Vowels
- N4026 L2/11-125: Revised Preliminary Proposal to Encode Soyombo in the UCS
- N4142 L2/11-412: Proposal to Encode the Soyombo Script in ISO/IEC 10646

The major changes from N4142 L2/11-412 include the renaming of some characters, reordering of the characters, and a significant revision of the text of the proposal.

The Soyombo font used here is based upon the font developed by Oliver Corff in August 1996 for his "Soyombo for LaTeX" package. The proposal author has made modifications to Corff's original font, which consist of the addition of new characters and glyphs.

2 Background

Soyombo (Mongolian: Соёмбо бичиг ᠰᠣᠶᠣᠮᠪᠣ ᠪᠢᠴᠢᠭ *soyombo bicig*) is a script used for writing Mongolian, Sanskrit, and Tibetan. It was used mainly for producing ornamental Buddhist texts. The script was designed in 1686 by Zanabazar (1635–1723), the first spiritual leader of Tibetan Buddhism in Mongolia, who also developed the Horizonal Square (Xewtee Dörböljin) script. The name *soyombo* ( ) is derived from Sanskrit स्वयंभु *svayaṃbhu* 'self-existing'. There are several records in Soyombo, which consist of manuscripts and inscriptions, such as that shown in figure 23. It is also te subject of academic studies, such as Boldsaikhan, et al. (2005) and Shagdarsürüng (2001). The script is currently in use, as is attested by a poem that was recently typeset using a digitized font (see figure 25).

3 Proposal Details

There are 117 characters proposed for encoding in the Soyombo block. A code chart and names list are attached. Names for characters are based upon Latin transliterations given in sources, such as Shagdarsürüng

(2001), with descriptors added for distinguishing characters with names that are transliterated identically. An attempt has been made to align Soyombo characters names with those for Indic and Tibetan characters in the UCS, and in parallel to those proposed for the Mongolian Square script (see N4413 L2/13-068).

4 Script Details

4.1 Structure of the Script

Soyombo is an alphasyllabic script that is written from left to right. In some cases, it is written vertically, from top to bottom (see section 4.14). Independent vowels are written using a vowel-carrier letter to which vowel signs are attached. Vowel length is indicated by a length mark that is attached to a base letter or to a sequence consisting of a base letter and a dependent vowel sign. Consonant letters possess the inherent vowel *a*, but in some language-specific contexts are purely alphabetic. The phonetic value of a consonant letter is changed by attaching a vowel sign to it. Consonant clusters are rendered as stacks, which are written using prefixed and subjoined letters. Syllable-final Mongolian consonants are written as dependent signs.

4.2 Structure of Characters

Each Soyombo letter consists of a frame and a nucleus that represents a distinctive phoneme, eg. . In the proposed encoding this combination of a frame and nucleus is considered an atomic letter, eg. *ga*. Vowels, final consonants, and other phonetic features are written as signs that attach to various positions on a letter: above the frame: ; below the nucleus: and ; inside the frame: ; to the right of the frame: and ; to the terminal of the frame: . Multiple signs may occur in a single location: , , . Nuclei of other letters attach beneath the nucleus of the base letter for representing consonant clusters, eg. , . These signs and their behaviors are described throughout the proposal.

Although a frame-nucleus combination is considered an atomic letter in this proposal, there is an alternate model for representing letters in which the frame and nucleus may be considered separate atomic elements. This approach mimics the visual method of writing Soyombo, where a letter is produced by drawing the frame and then writing the nucleus. However, this model has some drawbacks. First, it requires two characters for representing a single Soyombo letter, eg. *frame* + *ga-nucleus* instead of *ga*. Secondly, the visual model complicates the identification and conceptualization of letters: the Soyombo *ga* inherently refers to , not to its constituent elements + .

4.3 Vowel Letter

The LETTER A represents both the vowel *a* and a zero vowel. It functions as a vowel carrier for writing independent and initial vowels, and in such contexts it assumes the phonetic value of the combining vowel sign.

4.4 Vowel Length Mark

The VOWEL LENGTH MARK indicates vowel length. When attached to LETTER A or a consonant, it represents the lengthening of the inherent vowel *a* to *ā*, eg ā and *kā*. When it is written in combination with a vowel sign, the mark always occurs after the latter, eg. *ī* is encoded as <LETTER A, VOWEL SIGN I, VOWEL LENGTH MARK>. It attaches to the terminal of letter frame.

4.5 Vowel Signs

There are 10 dependent vowel signs:

VOWEL SIGN I	VOWEL SIGN O	VOWEL SIGN VOCALIC R
VOWEL SIGN E	VOWEL SIGN OE	VOWEL SIGN VOCALIC L
VOWEL SIGN UE	VOWEL SIGN AI	
VOWEL SIGN U	VOWEL SIGN AU	

Initial and independent forms of vowels are represented by attaching vowel signs to LETTER A. Long vowels are represented by placing the VOWEL LENGTH MARK after a vowel sign. The 10 vowel signs and the VOWEL LENGTH MARK can be used for writing the 20 vowel sounds provided for by the script:

a ā i ī ü ǖ u ū e ē o ō ö ȫ ai au

r̥ r̥̄ l̥ l̥̄

The signs VOWEL SIGN AI and VOWEL SIGN AU represent the diphthongs *ai* and *au*, but when used for writing Mongolian they may also be used as secondary vowel signs for writing *i* and *u* in vowel sequences:

riul < RA, VOWEL SIGN I, VOWEL SIGN AU, CONSONANT SIGN L>

gei < GA, VOWEL SIGN E, VOWEL SIGN AI>

dīin < DA, VOWEL SIGN I, VOWEL LENGTH MARK, VOWEL SIGN AI, CONSONANT SIGN N>

As indicated in the above encoded representations, the sign for the secondary vowel is always placed after the first vowel sign and any accompanying length mark and before any final-consonant sign.

4.6 Consonant Letters

There are 40 consonant letters:

GA	PA	GALIG GA	GALIG DHA
KA	MA	GALIG GHA	GALIG BA
NGA	YA	GALIG JA	GALIG BHA
JA	RA	GALIG JHA	GALIG SSA
CA	VA	GALIG TTA	GALIG TSA
NYA	LA	GALIG TTHA	GALIG TSHA
DA	SHA	GALIG DDA	GALIG DZA
TA	SA	GALIG DDHA	GALIG ZHA
NA	HA	GALIG NNA	GALIG ZA
BA	GALIG KSSA	GALIG DA	GALIG SMALL A

The order of the consonant letters adheres to that given in traditional charts. The letters GA .. HA are used in common for writing Mongolian, Sanskrit, and Tibetan. The letters with names containing the descriptor 'GALIG' (from гали *galig*, a Mongolian term for the transcription of non-Mongolian sounds) are used for writing Sanskrit and Tibetan. The term distinguishes letters used for writing the same sound, but in different languages, eg. GA and GALIG GA both represent /g/, but the latter is reserved for Sanskrit and Tibetan.

4.6.1 Notes on consonants

GALIG KSSA The letter GALIG KSSA represents the Sanskrit cluster *kṣa* (/kʂa/). In Soyombo, this letter represents a phoneme that is phonetically a consonant cluster, but, it has the structure of an atomic letter. It is encoded as a consonant letter because in all cases consonant conjunct forms are written as stacks in Soyombo, not as ligatures. While in some scripts the written form for Sanskrit /kʂa/ has an encoded representation as a character sequence, such an approach would not be consistent with this script.

GALIG SMALL A The letter GALIG SMALL A corresponds to U+0F60 TIBETAN LETTER -A.

4.6.2 Representation of Sanskrit and Tibetan

Mongolian is written using the common letters:

ga	GA	*ta*	TA	*la*	LA
ka	KA	*na*	NA	*va*	VA
ṅa	NGA	*ba*	BA	*ša*	SHA
ǰa	JA	*pa*	PA	*sa*	SA
ča	CA	*ma*	MA	*ha*	HA
ña	NYA	*ya*	YA		
da	DA	*ra*	RA		

Sanskrit and Tibetan are represented using a mix of common and *galig* letters. The common letters for voiced sounds (eg. GA, JA, DA, BA) are used for Sanskrit voiceless unaspirated stops, while the letters for voiceless sounds (eg. KA, CA, TA, PA) are used for the voiceless aspirated counterparts. The *galig* letters are used for the voiced unaspirated and aspirated pairs. The Sanskrit repertoire is as follows:

ka	GA	*ḍa*	GALIG DDA	*ma*	MA
kha	KA	*ḍha*	GALIG DDHA	*ya*	YA
ga	GALIG GA	*ṇa*	GALIG NNA	*ra*	RA
gha	GALIG GHA	*ta*	DA	*la*	LA
ṅa	NGA	*tha*	TA	*va*	VA
ca	JA	*da*	GALIG DA	*śa*	SHA
cha	CA	*dha*	GALIG DHA	*ṣa*	GALIG SSA
ja	GALIG JA	*na*	NA	*sa*	SA
jha	GALIG JHA	*pa*	BA	*ha*	HA
ña	NYA	*pha*	PA	*kṣa*	GALIG KSSA
ṭa	GALIG TTA	*ba*	GALIG BA		
ṭha	GALIG TTHA	*bha*	GALIG BHA		

Tibetan is represented by adding the following to the Sanskrit repertoire:

tsa	GALIG TSA	*dza*	GALIG DZA	*za*	GALIG ZA
tsha	GALIG TSHA	*zha*	GALIG ZHA	*'a*	GALIG SMALL A

4.7 Final Consonant Signs

Syllable-final consonants in Mongolian are written using the following 12 combining signs:

CONSONANT SIGN G	CONSONANT SIGN N	CONSONANT SIGN L
CONSONANT SIGN K	CONSONANT SIGN B	CONSONANT SIGN SH
CONSONANT SIGN NG	CONSONANT SIGN M	CONSONANT SIGN S
CONSONANT SIGN D	CONSONANT SIGN R	CONSONANT SIGN SMALL A

The majority of signs attach to the frame below the nucleus of a letter. The exception is CONSONANT SIGN SMALL A, which attaches to the right of the frame. These signs are shown below combined with A:

ag ak aṅ ad an ab am ar al aś as āṅ

The final-consonant sign always occurs after a vowel sign or the VOWEL LENGTH MARK in encoded text, eg. *kīng* is represented as <KA, VOWEL SIGN I, VOWEL LENGTH MARK, CONSONANT SIGN NG>.

4.8 Gemination Mark

Geminated consonants are written using the GEMINATION MARK, which is stacked above the triangle of the letter frame: <GA + GEMINATION MARK> → *kka*. In encoded text, the sign occurs immediately after the base letter before any combining sign.

4.9 Consonant Conjuncts

Consonant clusters are written as conjuncts, which are rendered as vertical stacks. When some consonants are cluster-initial, they are written as prefixed forms that are joined to the regular form of the following letter. A non-initial consonant is written using a subjoined form, which is the distinctive body of a letter without the frame, hereafter the 'nucleus'. This nucleus is written beneath the nucelus of the initial letter, eg. GA + RA is written as *gra*. Clusters consisting of multiple consonants are represented by writing the nucleus of each non-initial consonant beneath the subjoined form of the previous consonant, eg. GA + RA + VA is written as *grva*.

The consonants that are written as prefixed forms in the initial position of a cluster are LA, SHA, SA, RA. In such cases, C_2 functions as the base letter, but C_1 is parsed first in the logical order.

- LA takes the shape when it is C_1, eg. LA + GA is rendered as *lka*.
- SHA takes the shape when it is C_1, eg. SHA + GA is rendered as *śka*.
- SA takes the shape when it is C_1, eg. SA + GA is rendered as *ska*.
- takes the shape when it is C_1, eg. RA + GA is rendered as *rka*.

4.9.1 Proposed Encoding Model

Given the stacking behavior of consonant letters, the context-specific forms of LA, SHA, SA, RA, and the underlying structure of Soyombo, the proposed model for encoding conjuncts is based upon the subjoined-letter model for Tibetan, with some script-specific modifications. The subjoined model requires that a special form for each consonant letter be encoded separately, as well as for letters that take special forms when they are cluster initial. This approach allows for the writing of any combination or any number of consonants in a cluster. Moreover, the subjoined-letter model complements the frame-nucleus structure of the script and adheres to the method of writing Soyombo by hand.

The proposed model requires the accommodation of prefixed forms of LA, SHA, SA, RA, which are proposed for encoding as independent characters:

PREFIXED LETTER LA

PREFIXED LETTER SA

PREFIXED LETTER SHA

PREFIXED LETTER RA

It also requires the independent encoding of subjoined forms of each consonant letter:


SUBJOINED LETTER GA


SUBJOINED LETTER KA


SUBJOINED LETTER NGA


SUBJOINED LETTER JA


SUBJOINED LETTER CA


SUBJOINED LETTER NYA


SUBJOINED LETTER DA


SUBJOINED LETTER TA


SUBJOINED LETTER NA


SUBJOINED LETTER BA


SUBJOINED LETTER PA


SUBJOINED LETTER MA


SUBJOINED LETTER YA


SUBJOINED LETTER RA


SUBJOINED LETTER VA


SUBJOINED LETTER LA


SUBJOINED LETTER SHA


SUBJOINED LETTER SA


SUBJOINED LETTER HA


SUBJOINED LETTER GALIG KSSA


SUBJOINED LETTER GALIG GA


SUBJOINED LETTER GALIG GHA


SUBJOINED LETTER GALIG JA


SUBJOINED LETTER GALIG JHA


SUBJOINED LETTER GALIG TTA


SUBJOINED LETTER GALIG TTHA


SUBJOINED LETTER GALIG DDA


SUBJOINED LETTER GALIG DDHA


SUBJOINED LETTER GALIG NNA


SUBJOINED LETTER GALIG DA


SUBJOINED LETTER GALIG DHA


SUBJOINED LETTER GALIG BA


SUBJOINED LETTER GALIG BHA


SUBJOINED LETTER GALIG SSA


SUBJOINED LETTER GALIG TSA


SUBJOINED LETTER GALIG TSHA


SUBJOINED LETTER GALIG DZA


SUBJOINED LETTER GALIG ZHA


SUBJOINED LETTER GALIG ZA


SUBJOINED LETTER GALIG SMALL A

4.9.2 Consonant conjuncts in script charts

The following conjuncts are shown in traditional charts of Soyombo:

They are not independent characters, but conjuncts. Moreover, they are not the only conjuncts used in Soyombo, as others are attested in manuscripts. They are likely shown in order to illustrate the manner of representing consonant clusters, particularly the prefixed forms of LA, SHA, SA, RA; the principle of using subjoined characters for writing non-initial consonants; and indicating geminate consonants. These conjuncts are to be represented in encoded text as:

1. *kya* = <GA, ◌ SUBJOINED YA>
2. *kra* = <GA, ◌ SUBJOINED RA>
3. *kla* = <GA, ◌ SUBJOINED LA>
4. *kva* = <GA, ◌ SUBJOINED VA>
5. *kka* = <GA, ◌ GEMINATION MARK>
6. *ṅka* = <NGA, ◌ SUBJOINED GA>.
7. *ñca* = <NYA, ◌ SUBJOINED JA>.
8. *ṇṭa* = <GALIG NNA, ◌ SUBJOINED GALIG TTA>.
9. *nta* = <NA, ◌ SUBJOINED DA>.
10. *mpa* = <MA, ◌ SUBJOINED BA>.
11. *lka* = < PREFIXED LA, GA>.
12. *śka* = < PREFIXED SHA, GA>.
13. *ska* = < PREFIXED SA, GA>.
14. *rka* = < PREFIXED RA, GA>.

4.10 Various Signs

- ◌ SIGN ANUSVARA is used for indicating nasalization in words of Sanskrit origin, eg. *oṃ* <A, ◌ SIGN O, ◌ SIGN ANUSVARA>.
- ◌ SIGN VISARGA is used for indicating post-vocalic aspiration in Sanskrit words.

4.11 Punctuation

The following characters are used for punctuation:

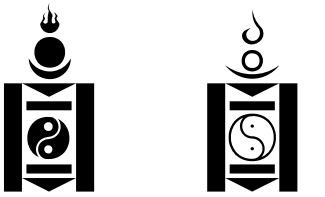
- TSHEG is used for marking the end of a syllable. It corresponds to U+0F0B TIBETAN MARK INTER-SYLLABIC TSHEG.
- SHAD indicates the end of a phrase or sentence. It corresponds to U+0F0D TIBETAN MARK SHAD and U+0964 DEVANAGARI DANDA.
- DOUBLE SHAD indicates the end of a text section. It corresponds to U+0F0E TIBETAN MARK NYIS SHAD and U+0965 DEVANAGARI DOUBLE DANDA.

4.12 Head Marks

There are two 'head marks': HEAD MARK and TERMINAL MARK. These characters are used for indicating the beginning and end of section of text.

4.13 Symbol

The SYMBOL SVAYAMBHU was design by Zanabazar, the creator of the script. It is an officially recognized symbol in Mongolia and appears on the flag and coat of arms of the country (see figures 30 and 31). The SYMBOL SVAYAMBHU appears in textual environments and must be represented in plain text (see figure 29. The SYMBOL SVAYAMBHU is similar to the HEAD MARK:



SVAYAMBHU HEAD MARK

The chief difference between the SYMBOL SVAYAMBHU and the HEAD MARK is the top.

4.14 Vertical Text

Soyombo may be written vertically, from top to bottom and left to right. There are two ways of rendering Soyombo in vertical environments. The first is by writing one syllable beneath the other. The example below shows the syllables *dzür ka* written vertical, as in the seal in figure 22.

LETTER GALIG DZA, VOWEL SIGN UE, CONSONANT SIGN R, KA

This sort of layout can be produced using vertical environments.

The second way of rendering Soyombo vertically is by writing multiple syllables using subjoined forms of consonants. Such stacking behavior departs from the general rule for writing syllables and mimicks

consonant conjuncts. The example below shows the word *temedeg* 'symbol' written vertically. Its three syllables have the same vowel, which is represented using a single vowel sign above the base letter; the consonants of non-initial syllables are written using subjoined letters, while the final consonant is written using a final-consonant sign:

TA, ◌̄ VOWEL SIGN E, MA, ◌̄ VOWEL SIGN E, DA, ◌̄ VOWEL SIGN E, ◌ FINAL CONSONANT SIGN G

A similar stacking practice is used for writing syllables that contain different vowels. The example below shows the word *yenu*, whose two syllables have different vowels that are written at different locations:

YA, ◌̄ VOWEL SIGN E, NA, ◌ VOWEL SIGN U

It is likely that the stacking of syllables occurs only when the syllables have the same vowel or when the vowels signs are written in different positions. It is unlikely that syllables would be stacked if the vowels signs were to clash.

The representation of vertical text by joining multiple syllables to form a word is currently not supposed by the proposed encoding. This sort of vertical representation could be achieved in the future through the use of a control character, such as a vertical syllable joiner. Although these are not consonant clusters, it may also be possible to use subjoined forms of letters. Additional research is required regarding this matter.

5 Glyph Interactions

5.1 Vowel Signs

When vowel-signs that appear at the base of the letter frame are written with final-consonants signs, then the size and position of the vowel signs are modified in order to accommodate the placement of both signs. This rule affects two vowel signs: ◌ VOWEL SIGN UE and ◌ VOWEL SIGN U.

- Shaping and positioning of VOWEL SIGN UE: The regular shape ◌ is compressed horizontally as ◌

KA + VOWEL SIGN UE + FINAL CONSONANT SIGN NG → *küng*

- Shaping and positioning of VOWEL SIGN U: The regular shape ◌ is compressed horizonally as ◌

KA + VOWEL SIGN U + FINAL CONSONANT SIGN NG → *kung*

5.2 Gemination Sign

The triangle of the base letter may be lowered to accommodate the GEMINATION MARK within the normal letter height: compare  with .

5.3 Glyph Shaping of Base Letters

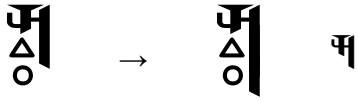
The glyphs of base letters in a conjunct may be modified in order to accommodate the placement of subjoined letters within the letter frame.

1. When 𑩐 GA, 𑩑 KA, 𑩒 NGA, 𑩓 GALIG GA, 𑩔 GALIG GHA are base letters, the nucleus is slightly truncated and vertically compressed, eg. the regular shape 𑩐 of GA is altered to 𑩐, as in writing 𑩐𑪙 *kya* <𑩐 GA, ◌𑪙 SUBJOINED LETTER YA>.

2. When 𑩐 GA, 𑩑 KA, 𑩒 NGA, 𑩓 GALIG GA, 𑩔 GALIG GHA occur as a base and subjoined letter pair, the nucleus of the base letter is compressed, the descender of its frame is broken, and the subjoined letter attaches to the bottom poriton of the descender, eg. 𑩒𑪜 *ṅka* <𑩒 NGA ◌𑪜 SUBJOINED GA>; 𑩒𑪝 *ṅga* <𑩒 NGA ◌𑪝 SUBJOINED GALIG GA>.

5.4 Glyph Shaping of Subjoined Letters

The glyphs for subjoined letters represent the base element of the regular letter. In some cases, it may be necessary to include part of the frame stem in the glyphs for subjoined letters in order to maintain the Soyombo graphical structure. For example, the syllable *grva* consists of a base letter and two subjoined letters. The position of the second subjoined letter would straddle the baseline or fall below it. In such cases, the stem of the base-letter frame would be extended.



𑩐 GA, ◌𑪚 SUBJOINED LETTER RA, ◌𑪰 SUBJOINED LETTER VA

This could be handled by creating glyphs for subjoined letters that contain a segment of the frame stem, eg. ◌𑪰 for ◌𑪰 SUBJOINED LETTER VA.

6 Character Data

6.1 Character Properties

```
11800;SOYOMBO LETTER A;Lo;0;L;;;;;N;;;;;
11801;SOYOMBO VOWEL SIGN I;Mn;0;NSM;;;;;N;;;;;
11802;SOYOMBO VOWEL SIGN E;Mn;0;NSM;;;;;N;;;;;
11803;SOYOMBO VOWEL SIGN UE;Mn;0;NSM;;;;;N;;;;;
11804;SOYOMBO VOWEL SIGN U;Mn;0;NSM;;;;;N;;;;;
11805;SOYOMBO VOWEL SIGN O;Mn;0;NSM;;;;;N;;;;;
11806;SOYOMBO VOWEL SIGN OE;Mn;0;NSM;;;;;N;;;;;
11807;SOYOMBO VOWEL SIGN AU;Mn;0;NSM;;;;;N;;;;;
11808;SOYOMBO VOWEL SIGN AI;Mn;0;NSM;;;;;N;;;;;
11809;SOYOMBO VOWEL SIGN VOCALIC R;Mn;0;NSM;;;;;N;;;;;
1180A;SOYOMBO VOWEL SIGN VOCALIC L;Mn;0;NSM;;;;;N;;;;;
1180B;SOYOMBO VOWEL LENGTH MARK;Mn;0;NSM;;;;;N;;;;;
1180C;SOYOMBO LETTER GA;Lo;0;L;;;;;N;;;;;
1180D;SOYOMBO LETTER KA;Lo;0;L;;;;;N;;;;;
1180E;SOYOMBO LETTER NGA;Lo;0;L;;;;;N;;;;;
1180F;SOYOMBO LETTER JA;Lo;0;L;;;;;N;;;;;
11810;SOYOMBO LETTER CA;Lo;0;L;;;;;N;;;;;
11811;SOYOMBO LETTER NYA;Lo;0;L;;;;;N;;;;;
11812;SOYOMBO LETTER DA;Lo;0;L;;;;;N;;;;;
```

```
11813;SOYOMBO LETTER TA;Lo;0;L;;;;;N;;;;;
11814;SOYOMBO LETTER NA;Lo;0;L;;;;;N;;;;;
11815;SOYOMBO LETTER BA;Lo;0;L;;;;;N;;;;;
11816;SOYOMBO LETTER PA;Lo;0;L;;;;;N;;;;;
11817;SOYOMBO LETTER MA;Lo;0;L;;;;;N;;;;;
11818;SOYOMBO LETTER YA;Lo;0;L;;;;;N;;;;;
11819;SOYOMBO LETTER RA;Lo;0;L;;;;;N;;;;;
1181A;SOYOMBO LETTER VA;Lo;0;L;;;;;N;;;;;
1181B;SOYOMBO LETTER LA;Lo;0;L;;;;;N;;;;;
1181C;SOYOMBO LETTER SHA;Lo;0;L;;;;;N;;;;;
1181D;SOYOMBO LETTER SA;Lo;0;L;;;;;N;;;;;
1181E;SOYOMBO LETTER HA;Lo;0;L;;;;;N;;;;;
1181F;SOYOMBO LETTER KSSA;Lo;0;L;;;;;N;;;;;
11820;SOYOMBO LETTER GALIG GA;Lo;0;L;;;;;N;;;;;
11821;SOYOMBO LETTER GALIG GHA;Lo;0;L;;;;;N;;;;;
11822;SOYOMBO LETTER GALIG JA;Lo;0;L;;;;;N;;;;;
11823;SOYOMBO LETTER GALIG JHA;Lo;0;L;;;;;N;;;;;
11824;SOYOMBO LETTER GALIG TTA;Lo;0;L;;;;;N;;;;;
11825;SOYOMBO LETTER GALIG TTHA;Lo;0;L;;;;;N;;;;;
11826;SOYOMBO LETTER GALIG DDA;Lo;0;L;;;;;N;;;;;
11827;SOYOMBO LETTER GALIG DDHA;Lo;0;L;;;;;N;;;;;
11828;SOYOMBO LETTER GALIG NNA;Lo;0;L;;;;;N;;;;;
11829;SOYOMBO LETTER GALIG DA;Lo;0;L;;;;;N;;;;;
1182A;SOYOMBO LETTER GALIG DHA;Lo;0;L;;;;;N;;;;;
1182B;SOYOMBO LETTER GALIG BA;Lo;0;L;;;;;N;;;;;
1182C;SOYOMBO LETTER GALIG BHA;Lo;0;L;;;;;N;;;;;
1182D;SOYOMBO LETTER GALIG SSA;Lo;0;L;;;;;N;;;;;
1182E;SOYOMBO LETTER GALIG TSA;Lo;0;L;;;;;N;;;;;
1182F;SOYOMBO LETTER GALIG TSHA;Lo;0;L;;;;;N;;;;;
11830;SOYOMBO LETTER GALIG DZA;Lo;0;L;;;;;N;;;;;
11831;SOYOMBO LETTER GALIG ZHA;Lo;0;L;;;;;N;;;;;
11832;SOYOMBO LETTER GALIG ZA;Lo;0;L;;;;;N;;;;;
11833;SOYOMBO LETTER GALIG SMALL A;Lo;0;L;;;;;N;;;;;
11834;SOYOMBO CONSONANT SIGN G;Mn;0;NSM;;;;;N;;;;;
11835;SOYOMBO CONSONANT SIGN K;Mn;0;NSM;;;;;N;;;;;
11836;SOYOMBO CONSONANT SIGN NG;Mn;0;NSM;;;;;N;;;;;
11837;SOYOMBO CONSONANT SIGN D;Mn;0;NSM;;;;;N;;;;;
11838;SOYOMBO CONSONANT SIGN N;Mn;0;NSM;;;;;N;;;;;
11839;SOYOMBO CONSONANT SIGN B;Mn;0;NSM;;;;;N;;;;;
1183A;SOYOMBO CONSONANT SIGN M;Mn;0;NSM;;;;;N;;;;;
1183B;SOYOMBO CONSONANT SIGN R;Mn;0;NSM;;;;;N;;;;;
1183C;SOYOMBO CONSONANT SIGN L;Mn;0;NSM;;;;;N;;;;;
1183D;SOYOMBO CONSONANT SIGN SH;Mn;0;NSM;;;;;N;;;;;
1183E;SOYOMBO CONSONANT SIGN S;Mn;0;NSM;;;;;N;;;;;
1183F;SOYOMBO CONSONANT SIGN SMALL A;Mn;0;NSM;;;;;N;;;;;
11840;SOYOMBO SIGN ANUSVARA;Mn;0;NSM;;;;;N;;;;;
11841;SOYOMBO SIGN VISARGA;Mc;0;L;;;;;N;;;;;
11842;SOYOMBO GEMINATION SIGN;Mn;0;NSM;;;;;N;;;;;
11843;SOYOMBO PREFIXED LETTER LA;Lo;0;L;;;;;N;;;;;
11844;SOYOMBO PREFIXED LETTER SHA;Lo;0;L;;;;;N;;;;;
11845;SOYOMBO PREFIXED LETTER SA;Lo;0;L;;;;;N;;;;;
11846;SOYOMBO PREFIXED LETTER RA;Lo;0;L;;;;;N;;;;;
11847;SOYOMBO SUBJOINED LETTER GA;Mn;0;NSM;;;;;N;;;;;
11848;SOYOMBO SUBJOINED LETTER KA;Mn;0;NSM;;;;;N;;;;;
11849;SOYOMBO SUBJOINED LETTER NGA;Mn;0;NSM;;;;;N;;;;;
1184A;SOYOMBO SUBJOINED LETTER JA;Mn;0;NSM;;;;;N;;;;;
1184B;SOYOMBO SUBJOINED LETTER CA;Mn;0;NSM;;;;;N;;;;;
1184C;SOYOMBO SUBJOINED LETTER NYA;Mn;0;NSM;;;;;N;;;;;
1184D;SOYOMBO SUBJOINED LETTER DA;Mn;0;NSM;;;;;N;;;;;
1184E;SOYOMBO SUBJOINED LETTER TA;Mn;0;NSM;;;;;N;;;;;
1184F;SOYOMBO SUBJOINED LETTER NA;Mn;0;NSM;;;;;N;;;;;
```

```
11850;SOYOMBO SUBJOINED LETTER BA;Mn;0;NSM;;;;;N;;;;;
11851;SOYOMBO SUBJOINED LETTER PA;Mn;0;NSM;;;;;N;;;;;
11852;SOYOMBO SUBJOINED LETTER MA;Mn;0;NSM;;;;;N;;;;;
11853;SOYOMBO SUBJOINED LETTER YA;Mn;0;NSM;;;;;N;;;;;
11854;SOYOMBO SUBJOINED LETTER RA;Mn;0;NSM;;;;;N;;;;;
11855;SOYOMBO SUBJOINED LETTER VA;Mn;0;NSM;;;;;N;;;;;
11856;SOYOMBO SUBJOINED LETTER LA;Mn;0;NSM;;;;;N;;;;;
11857;SOYOMBO SUBJOINED LETTER SHA;Mn;0;NSM;;;;;N;;;;;
11858;SOYOMBO SUBJOINED LETTER SA;Mn;0;NSM;;;;;N;;;;;
11859;SOYOMBO SUBJOINED LETTER HA;Mn;0;NSM;;;;;N;;;;;
1185A;SOYOMBO SUBJOINED LETTER GALIG KSSA;Mn;0;NSM;;;;;N;;;;;
1185A;SOYOMBO SUBJOINED LETTER GALIG GA;Mn;0;NSM;;;;;N;;;;;
1185B;SOYOMBO SUBJOINED LETTER GALIG GHA;Mn;0;NSM;;;;;N;;;;;
1185C;SOYOMBO SUBJOINED LETTER GALIG JA;Mn;0;NSM;;;;;N;;;;;
1185D;SOYOMBO SUBJOINED LETTER GALIG JHA;Mn;0;NSM;;;;;N;;;;;
1185E;SOYOMBO SUBJOINED LETTER GALIG TTA;Mn;0;NSM;;;;;N;;;;;
1185F;SOYOMBO SUBJOINED LETTER GALIG TTHA;Mn;0;NSM;;;;;N;;;;;
11860;SOYOMBO SUBJOINED LETTER GALIG DDA;Mn;0;NSM;;;;;N;;;;;
11861;SOYOMBO SUBJOINED LETTER GALIG DDHA;Mn;0;NSM;;;;;N;;;;;
11862;SOYOMBO SUBJOINED LETTER GALIG NNA;Mn;0;NSM;;;;;N;;;;;
11863;SOYOMBO SUBJOINED LETTER GALIG DA;Mn;0;NSM;;;;;N;;;;;
11864;SOYOMBO SUBJOINED LETTER GALIG DHA;Mn;0;NSM;;;;;N;;;;;
11865;SOYOMBO SUBJOINED LETTER GALIG BA;Mn;0;NSM;;;;;N;;;;;
11866;SOYOMBO SUBJOINED LETTER GALIG BHA;Mn;0;NSM;;;;;N;;;;;
11867;SOYOMBO SUBJOINED LETTER GALIG SSA;Mn;0;NSM;;;;;N;;;;;
11868;SOYOMBO SUBJOINED LETTER GALIG TSA;Mn;0;NSM;;;;;N;;;;;
11869;SOYOMBO SUBJOINED LETTER GALIG TSHA;Mn;0;NSM;;;;;N;;;;;
1186A;SOYOMBO SUBJOINED LETTER GALIG DZA;Mn;0;NSM;;;;;N;;;;;
1186B;SOYOMBO SUBJOINED LETTER GALIG ZHA;Mn;0;NSM;;;;;N;;;;;
1186C;SOYOMBO SUBJOINED LETTER GALIG ZA;Mn;0;NSM;;;;;N;;;;;
1186D;SOYOMBO SUBJOINED LETTER GALIG SMALL A;Mn;0;NSM;;;;;N;;;;;
1186F;SOYOMBO TSHEG;Po;0;L;;;;;N;;;;;
11870;SOYOMBO SHAD;Po;0;L;;;;;N;;;;;
11871;SOYOMBO DOUBLE SHAD;Po;0;L;;;;;N;;;;;
11872;SOYOMBO HEAD MARK;Po;0;ON;;;;;N;;;;;
11873;SOYOMBO TERMINAL MARK;Po;0;ON;;;;;N;;;;;
11874;SOYOMBO SYMBOL SVAYAMBHU;So;0;L;;;;;N;;;;;
```

6.2 Linebreaking Properties

Linebreaking properties given in the data format of `LineBreak.txt`:

```
11800; AL             # LETTER A
11801..11809; CM      # VOWEL SIGN I .. VOWEL LENGTH MARK
1180A..1181C; AL      # GA .. HA
1181D..11828; CM      # CONSONANT SIGN G .. CONSONANT SIGN SMALL A
11829..1183D; AL      # GALIG GA .. KSSA
11842; CM             # GEMINATION SIGN
11843..1186E; CM      # PREFIXED LETTER LA .. SUBJOINED LETTER GALIG SMALL A
1186F; BA             # TSHEG
11870..11871; BA      # SHAD .. DOUBLE SHAD
11872; BB             # HEAD MARK
11873; BA             # TERMINAL MARK
11874; AL             # SYMBOL SVAYAMBHU
```

7 References

Atwood, Christopher P. 2004. *Encyclopedia of Mongolia and the Mongol Empire*. New York: Facts on File.

Boldsaikhan, B., B. Batsana, and Ts. Oyuntsetseg. 2005. *Соёмбо Нууц ба Синергетик* [= *Secret of Soyombo and Synergetic*]. (Түвэд, монгол бичгийн эхийг орчуулан хавсаргав). [illegible]. Ed. by T. Bulgan. Ulaanbaatar: Shambala Association, System Science Research Institute, Mongolian University of Science and Technology.

Corff, Oliver. "Soyombo for LaTeX", v. 0.9. September 2, 1996. `http://userpage.fu-berlin.de/~corff/im/Soyombo/overview.Soyombo.html`

"Histoire du livre". 2010. "Histoire du livre et histoire de l'écriture: l'écriture, entre pratique, symbole et économie" (26 August 2010). `http://histoire-du-livre.blogspot.com/2010/08/lecriture-entre-le-symbole-et-la.html`

Kapaj, Luigi. 2002. "Mongol Scripts". `http://www.viahistoria.com/SilverHorde/main.html?research/MongolScripts.html`

Mongolwiki. 2008. "Сайн чанарын зураг байхгүй байна". `http://wiki.ecm-outsourcing.com/index.php?title=Дурс:Soyombo_usgiin_bar.gif`

Mongolia and Japan. 1988 "Addition of Soyombo Script". N1855 L2/98-358. September 21, 1998. `http://std.dkuug.dk/jtc1/sc2/wg2/docs/n1855.pdf`

Pandey, Anshuman. 2010a. "Preliminary Proposal to Encode the Xawtaa Dorboljin Script in ISO/IEC 10646". N3956 L2/10-411. October 23, 2010. `http://std.dkuug.dk/jtc1/sc2/wg2/docs/n3956.pdf`

———. 2010b. "Preliminary Proposal to Encode the Soyombo Script in ISO/IEC 10646". N3949 L2/10-399. October 30, 2010. `http://std.dkuug.dk/jtc1/sc2/wg2/docs/n3949.pdf`

———. 2011a. "Determining the Encoding Model for Soyombo Vowels". N3986 L2/11-054. February 5, 2011. `http://std.dkuug.dk/jtc1/sc2/wg2/docs/n3986.pdf`

———. 2011b. "Revised Preliminary Proposal to Encode Soyombo in the UCS". N4026 L2/11-125. April 25, 2011. `http://std.dkuug.dk/jtc1/sc2/wg2/docs/n4026.pdf`

———. 2011c. "Proposal to Encode the Soyombo Script in ISO/IEC 10646". N4142 L2/11-412. October 25, 2011. `http://std.dkuug.dk/jtc1/sc2/wg2/docs/n4142.pdf`

———. 2013. "Proposal to Encode the Mongolian Square Script in ISO/IEC 10646". N4413 L2/13-068. April 22, 2013. `http://std.dkuug.dk/jtc1/sc2/wg2/docs/n4413.pdf`

Rintschen. 1953. "A propos de la sigillographie mongole". *Acta Orientalia Academiae Scientiarum Hungaricae*, vol. III., pp. 25-31.

Sato, Takayuki K. "Soyombo and Pagba (old Mongol scripts)". N2163 L2/00-055. January 6, 2000. `http://std.dkuug.dk/jtc1/sc2/wg2/docs/n2163.doc`

Shagdarsürüng, Tseveliin. 2001. *Study of Mongolian Scripts* (Graphic Study or Grammatology). Enl. 2nd ed. Bibliotheca Mongolica: Monograph 1. Ed. by Sharaviin Choimaa. Ulaanbaatar: Center for Mongol Studies, National University of Mongolia.

Wikimedia. 2006. "1000 Tugriks - Recto". `http://en.wikipedia.org/wiki/File:1000_Tugriks_-_Recto.jpg`

———. 2008. "Flag of Mongolia". http://en.wikipedia.org/wiki/File:Flag_of_Mongolia.svg

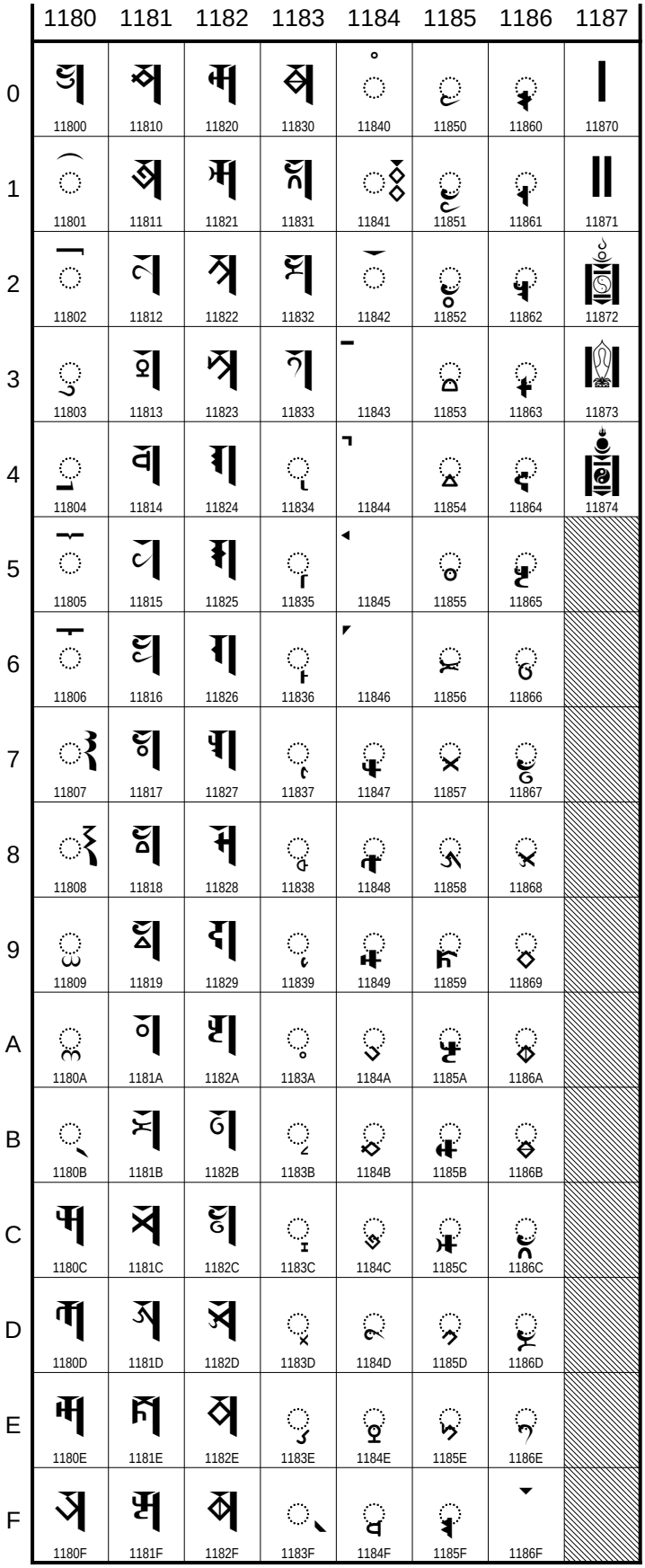
———. 2009a. "Coat of Arms of Mongolia". http://en.wikipedia.org/wiki/File:Coat_of_Arms_of_Mongolia.svg

———. 2009b. "Green Tara poem typed in Soyombo script". http://en.wikipedia.org/wiki/File:Green_Tara_Poem_Soyombo.jpg

8 Acknowledgments

I express my gratitude to Shriramana Sharma, Peter Constable (Microsoft), and György Kara (Indiana University, Bloomington) for reviewing a draft of this proposal and for providing detailed comments on the encoding model and various other aspects of the encoding.

This project was made possible in part by a grant from the United States National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at the University of California, Berkeley). Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment for the Humanities.


1180 1181 1182 1183 1184 1185 1186 1187
0 11800 11810 11820 11830 11840 11850 11860 11870
1 11801 11811 11821 11831 11841 11851 11861 11871
2 11802 11812 11822 11832 11842 11852 11862 11872
3 11803 11813 11823 11833 11843 11853 11863 11873
4 11804 11814 11824 11834 11844 11854 11864 11874
5 11805 11815 11825 11835 11845 11855 11865
6 11806 11816 11826 11836 11846 11856 11866
7 11807 11817 11827 11837 11847 11857 11867
8 11808 11818 11828 11838 11848 11858 11868
9 11809 11819 11829 11839 11849 11859 11869
A 1180A 1181A 1182A 1183A 1184A 1185A 1186A
B 1180B 1181B 1182B 1183B 1184B 1185B 1186B
C 1180C 1181C 1182C 1183C 1184C 1185C 1186C
D 1180D 1181D 1182D 1183D 1184D 1185D 1186D
E 1180E 1181E 1182E 1183E 1184E 1185E 1186E
F 1180F 1181F 1182F 1183F 1184F 1185F 1186F

Vowel carrier

11800 𑩐 SOYOMBO LETTER A

Vowel signs

11801 ◌𑩰 SOYOMBO VOWEL SIGN I
11802 ◌𑩱 SOYOMBO VOWEL SIGN E
11803 ◌𑩲 SOYOMBO VOWEL SIGN UE
11804 ◌𑩳 SOYOMBO VOWEL SIGN U
11805 ◌𑩴 SOYOMBO VOWEL SIGN O
11806 ◌𑩵 SOYOMBO VOWEL SIGN OE
11807 ◌𑩶 SOYOMBO VOWEL SIGN AU
11808 ◌𑩷 SOYOMBO VOWEL SIGN AI
11809 ◌𑩸 SOYOMBO VOWEL SIGN VOCALIC R
1180A ◌𑩹 SOYOMBO VOWEL SIGN VOCALIC L

Vowel length mark

1180B ◌𑩺 SOYOMBO VOWEL LENGTH MARK

Consonants

1180C 𑩻 SOYOMBO LETTER GA
• Used for Sanskrit ka
1180D 𑩼 SOYOMBO LETTER KA
• Used for Sanskrit kha
1180E 𑩽 SOYOMBO LETTER NGA
1180F 𑩾 SOYOMBO LETTER JA
• Used for Sanskrit ca
11810 𑩿 SOYOMBO LETTER CA
• Used for Sanskrit cha
11811 𑪀 SOYOMBO LETTER NYA
11812 𑪁 SOYOMBO LETTER DA
• Used for Sanskrit ta
11813 𑪂 SOYOMBO LETTER TA
• Used for Sanskrit tha
11814 𑪃 SOYOMBO LETTER NA
11815 𑪄 SOYOMBO LETTER BA
• Used for Sanskrit pa
11816 𑪅 SOYOMBO LETTER PA
• Used for Sanskrit pha
11817 𑪆 SOYOMBO LETTER MA
11818 𑪇 SOYOMBO LETTER YA
11819 𑪈 SOYOMBO LETTER RA
1181A 𑪉 SOYOMBO LETTER VA
1181B 𑪊 SOYOMBO LETTER LA
1181C 𑪋 SOYOMBO LETTER SHA
1181D 𑪌 SOYOMBO LETTER SA
1181E 𑪍 SOYOMBO LETTER HA
1181F 𑪎 SOYOMBO LETTER GALIG KSSA
11820 𑪏 SOYOMBO LETTER GALIG GA
11821 𑪐 SOYOMBO LETTER GALIG GHA
11822 𑪑 SOYOMBO LETTER GALIG JA
11823 𑪒 SOYOMBO LETTER GALIG JHA
11824 𑪓 SOYOMBO LETTER GALIG TTA
11825 𑪔 SOYOMBO LETTER GALIG TTHA
11826 𑪕 SOYOMBO LETTER GALIG DDA
11827 𑪖 SOYOMBO LETTER GALIG DDHA
11828 𑪗 SOYOMBO LETTER GALIG NNA
11829 𑪘 SOYOMBO LETTER GALIG DA
1182A 𑪙 SOYOMBO LETTER GALIG DHA
1182B 𑪚 SOYOMBO LETTER GALIG BA
1182C 𑪛 SOYOMBO LETTER GALIG BHA
1182D 𑪜 SOYOMBO LETTER GALIG SSA
1182E 𑪝 SOYOMBO LETTER GALIG TSA
1182F 𑪞 SOYOMBO LETTER GALIG TSHA
11830 𑪟 SOYOMBO LETTER GALIG DZA
11831 𑪠 SOYOMBO LETTER GALIG ZHA
11832 𑪡 SOYOMBO LETTER GALIG ZA
11833 𑪢 SOYOMBO LETTER GALIG SMALL A

Final consonant signs

11834 ◌𑪣 SOYOMBO CONSONANT SIGN G
11835 ◌𑪤 SOYOMBO CONSONANT SIGN K
11836 ◌𑪥 SOYOMBO CONSONANT SIGN NG
11837 ◌𑪦 SOYOMBO CONSONANT SIGN D
11838 ◌𑪧 SOYOMBO CONSONANT SIGN N
11839 ◌𑪨 SOYOMBO CONSONANT SIGN B
1183A ◌𑪩 SOYOMBO CONSONANT SIGN M
1183B ◌𑪪 SOYOMBO CONSONANT SIGN R
1183C ◌𑪫 SOYOMBO CONSONANT SIGN L
1183D ◌𑪬 SOYOMBO CONSONANT SIGN SH
1183E ◌𑪭 SOYOMBO CONSONANT SIGN S
1183F ◌𑪮 SOYOMBO CONSONANT SIGN SMALL A

Signs for Sanskrit

11840 ◌𑪯 SOYOMBO SIGN ANUSVARA
11841 ◌𑪰 SOYOMBO SIGN VISARGA

Gemination mark

11842 ◌𑪱 SOYOMBO GEMINATION MARK

Prefixed letters

11843 ◌𑪲 SOYOMBO PREFIXED LETTER LA
11844 ◌𑪳 SOYOMBO PREFIXED LETTER SHA
11845 ◌𑪴 SOYOMBO PREFIXED LETTER SA
11846 ◌𑪵 SOYOMBO PREFIXED LETTER RA

Subjoined letters

11847 ◌𑪶 SOYOMBO SUBJOINED LETTER GA
11848 ◌𑪷 SOYOMBO SUBJOINED LETTER KA
11849 ◌𑪸 SOYOMBO SUBJOINED LETTER NGA
1184A ◌𑪹 SOYOMBO SUBJOINED LETTER JA
1184B ◌𑪺 SOYOMBO SUBJOINED LETTER CA
1184C ◌𑪻 SOYOMBO SUBJOINED LETTER NYA
1184D ◌𑪼 SOYOMBO SUBJOINED LETTER DA
1184E ◌𑪽 SOYOMBO SUBJOINED LETTER TA
1184F ◌𑪾 SOYOMBO SUBJOINED LETTER NA
11850 ◌𑪿 SOYOMBO SUBJOINED LETTER BA
11851 ◌𑫀 SOYOMBO SUBJOINED LETTER PA
11852 ◌𑫁 SOYOMBO SUBJOINED LETTER MA
11853 ◌𑫂 SOYOMBO SUBJOINED LETTER YA
11854 ◌𑫃 SOYOMBO SUBJOINED LETTER RA
11855 ◌𑫄 SOYOMBO SUBJOINED LETTER VA
11856 ◌𑫅 SOYOMBO SUBJOINED LETTER LA
11857 ◌𑫆 SOYOMBO SUBJOINED LETTER SHA
11858 ◌𑫇 SOYOMBO SUBJOINED LETTER SA
11859 ◌𑫈 SOYOMBO SUBJOINED LETTER HA
1185A ◌𑫉 SOYOMBO SUBJOINED LETTER GALIG KSSA
1185B ◌𑫊 SOYOMBO SUBJOINED LETTER GALIG GA
1185C ◌𑫋 SOYOMBO SUBJOINED LETTER GALIG GHA
1185D ◌𑫌 SOYOMBO SUBJOINED LETTER GALIG JA
1185E ◌𑫍 SOYOMBO SUBJOINED LETTER GALIG JHA
1185F ◌𑫎 SOYOMBO SUBJOINED LETTER GALIG TTA
11860 ◌𑫏 SOYOMBO SUBJOINED LETTER GALIG TTHA
11861 ◌𑫐 SOYOMBO SUBJOINED LETTER GALIG DDA
11862 ◌𑫑 SOYOMBO SUBJOINED LETTER GALIG DDHA
11863 ◌𑫒 SOYOMBO SUBJOINED LETTER GALIG NNA
11864 ◌𑫓 SOYOMBO SUBJOINED LETTER GALIG DA
11865 ◌𑫔 SOYOMBO SUBJOINED LETTER GALIG DHA
11866 ◌𑫕 SOYOMBO SUBJOINED LETTER GALIG BA
11867 ◌𑫖 SOYOMBO SUBJOINED LETTER GALIG BHA
11868 ◌𑫗 SOYOMBO SUBJOINED LETTER GALIG SSA
11869 ◌𑫘 SOYOMBO SUBJOINED LETTER GALIG TSA
1186A ◌𑫙 SOYOMBO SUBJOINED LETTER GALIG TSHA
1186B ◌𑫚 SOYOMBO SUBJOINED LETTER GALIG DZA
1186C ◌𑫛 SOYOMBO SUBJOINED LETTER GALIG ZHA

1186D SOYOMBO SUBJOINED LETTER GALIG ZA
1186E SOYOMBO SUBJOINED LETTER GALIG SMALL A

Punctuation

1186F SOYOMBO TSHEG
→ 0F0B tibetan mark intersyllabic tsheg
11870 SOYOMBO SHAD
11871 SOYOMBO DOUBLE SHAD

Head marks

11872  SOYOMBO HEAD MARK
• written at the beginning of Soyombo text
11873  SOYOMBO TERMINAL MARK
• written at the end of Soyombo text

Symbol

11874  SOYOMBO SYMBOL SVAYAMBHU

Figure 1: Photograph of a chart of the Soyombo script (from "Histoire du livre" 2010).

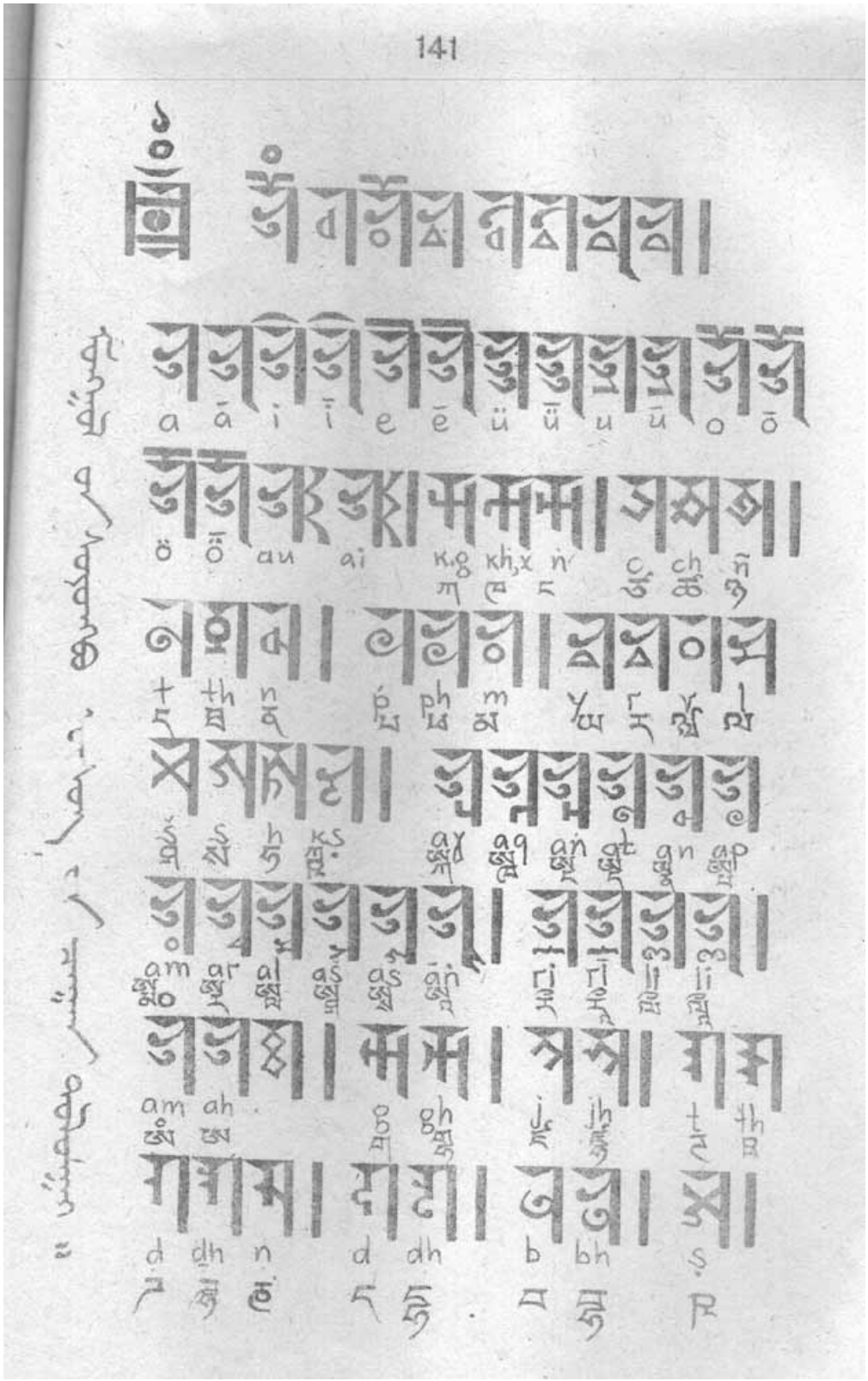


Figure 2: Characters of the Soyombo script (from Kapaj 2002).

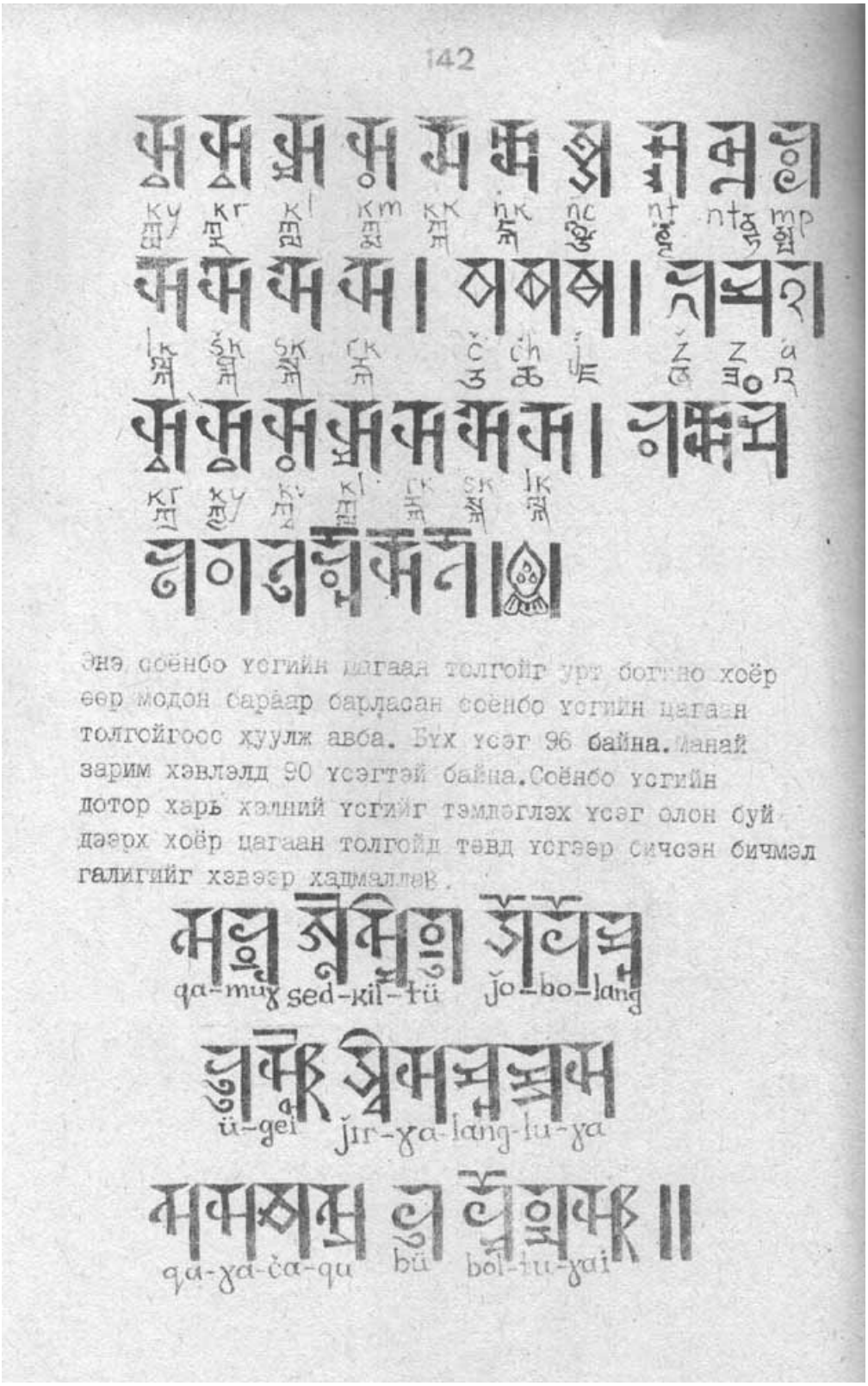


Figure 3: Characters of the Soyombo script (from Kapaj 2002).

Figure 4: Traditional chart of Soyombo (from Shagdarsürüng 2001: 152).

Хуудас эхлэсний тэмдэг. Номд бярга буюу эгчим ( ) -тэй адил үүрэгтэй. Соёмбо бичгийн энэ тэмдэг монголчуудын хувьд тусгаар тогтнолын бэлгэдэл болжээ.

Figure 5: Description of Soyombo vowels (from Shagdarsürüng 2001: 133).

1. 

tl. A; **tc.** mong., tib., sans. a
Энэ хэлбэр нь а эгшигийн бие даасан (IF) буюу үгийн (зарим тохиолдолд үеийн) эхинд тохиолдоно. Жишээлбэл:

tl. A-li$_1$;
tc. a-li.

(Ø) **tl.** a$_o$; **tc.** a. Энэ нь а эгшигийн гол хэлбэр (MF) буюу үгийн (зарим тохиолдолд үеийн) дунд болоод адагт тохиолдох нууц буюу тэг (Ø) хэлбэр. Жишээлбэл:

tl. Ga$_o$-Ja$_o$r;
tc. ga-jar

1.a. 

tl. ā; **tc**. ā. Урт а эгшигийн бие даасан буюу (IF) хэлбэр. Жишээ нь:

tl. A-Da$_o$r;
tc. ā-dar.

tl. ¯a$_o$; **tc**. ā. Энэ нь угтаа эгшигийн уртын тэмдэг. а эгшигтэй тохиолдвол, тэрхүү а эгшиг нь нууц буюу "тэг" (Ø) хэлбэртэй байдаг. Жишээлбэл:

tl. Kā$_o$n;
tc. kān.

2. 

tl. I; **tc**. mong, tib., sans: i. Энэ нь і эгшигийн (IF) хэлбэр. Жишээлбэл:

tl. I-Te-Gel;
tc. i-te-gel.

tl. i$_1$; **tc.** i. Энэ нь і эгшигийн гол хэлбэрийн нэг буюу (MF$_1$) хэлбэр. Зөвхөн гийгүүлэгч (C)-ийн дараа буюу (C + i$_1$) нөхцөлд л тохиолдоно. Жишээлбэл:

tl. A-Ci$_1$-tu$_1$;
tc. a-či-tu.

tl. i$_2$; **tc.** i. Энэ нь і эгшигийн гол хэлбэрийн нэг буюу (MF$_2$) хэлбэр. Зөвхөн эгшиг (V)-ийн дараа буюу (V + i$_2$) нөхцөлд тохиолдолдоно. Жишээлбэл:

tl. Bol-Tu$_1$-Ga$_o$i$_2$;
tc. bol-tu-gai.

2.a. 

tl. I; **tc.** mong., sans: ī. Урт і эгшигийн бие даасан (IF) хэлбэр. Жишээлбэл:

tl. Ī-Me;
tc. ī-me.

Figure 6: Description of Soyombo vowels (from Shagdarsürüng 2001: 134).

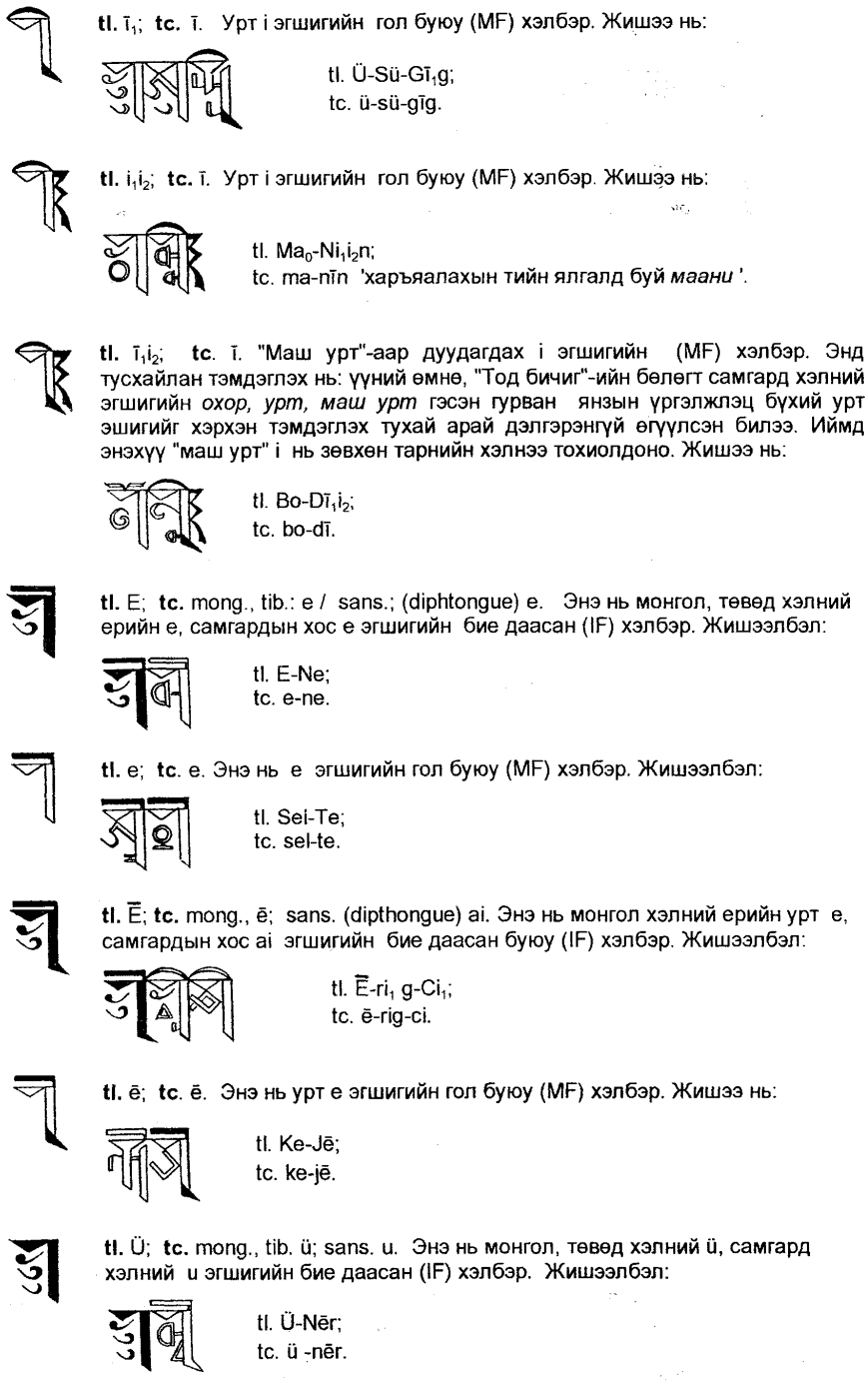


Figure 7: Description of Soyombo vowels (from Shagdarsürüng 2001: 135).

tl. ü; **tc.** ü. Энэ нь ü эгшигийн гол буюу (MF) хэлбэр. Жишээ нь:

tl. Ün-Dü-Sü;
tc. ün-dü-sü.

4.a. 

tl. Ū; **tc.** mong. ū, sans. ū. Монгол хэлний урт ū, самгардын урт ū эгшигийн бие даасан буюу (IF) хэлбэр. Жишээ нь:

tl. Ūr;
tc. ūr.

tl. ū; **tc.** ū. Монгол хэлний урт ū, самгардын урт ū эгшигийн гол буюу (MF) хэлбэр. Жишээ нь:

tl. Ü-Jū-Lüg-Sen;
tc. ü-jū-lüg-sen.

5. 

tl. U; **tc.** mong. u. Монгол хэлний u эгшигийн бие даасан буюу (IF) хэлбэр. Жишээ нь:

tl. U-sun;
tc. u-sun.

tl. u_1; **tc.** u. Энэ нь u эгшигийн гол буюу (MF) хэлбэрийн нэг. Зөвхөн гийгүүлэгчийн дараа буюу (C + u_1) нөхцөлд тохиолдоно. Жишээ нь:

tl. Ya_0-Bu_1-Da_0l;
tc. ya-bu-dal

tl. u_2; **tc.** u. Энэ нь u эгшигийн гол буюу (MF) хэлбэрийн нэг. Зөвхөн эгшигийн дараа буюу (V + u_2) нөхцөлд тохиолдоно. Жишээ нь:

tl. Jo-Ri_1u_2l-Ji_1u_2;
tc. jo-riul-ju.

tl. Ya_0-Ga_0u_2l;
tc. ya-gaul 'шалтгаан, үндэс '[15]

5.a. 

tl. ū; **tc.** mong., sans. ū. Энэ нь монгол болон самгардын урт ū эгшигийн бие даасан буюу (IF) хэлбэр. Жишээ нь:

[15] Соёмбо бичгийн *"Итгэл"*-д төвөд хэлний tib. rgyu гэдэг үгийг ya-gaul (<Mo. *yaγaγul) хэмээн орчуулсан нь буй. Энэ үгийг Ойродын Зая Бандидын орчуулсан тод *"Итгэл"*-д ündüsün гэж, буриад бичмэлд siltaγan хэмээн орчуулжээ. Энэ тухай G. Kara, *Un texte mongol en écriture soyombo*, - AOH, Tomus. IX, Fasc. 1, Budapest, 1959, pp. 1-38 болон Chagdarsureng, *Sur quelques traductions mongoles du "Natha"*, - Studia Mongolica, Tom. 2 (10), Fasc. 11, Ulan-Bator, 1975, p. 183 (N. 86) -д үзмүү.

Figure 8: Description of Soyombo vowels (from Shagdarsürüng 2001: 136).

tl. ū-Da$_0$m;
tc. ū-dam.

tl. ū; **tc.** ū. Урт ū эгшигийн (MF) хэлбэр. Жишээ нь:

tl. A-Rūl;
tc. ā-rūl.

6. 

tl. O; **tc.** mong., tib. o, sans. diphtongue o. Энэ нь монгол, төвөд хэлний ерийн о эгшиг, самгард хэлний хос о эгшигийн бие даасан (IF) хэлбэр. Жишээлбэл:

tl. O-Ron;
tc.o-ron.

tl. o; **tc.** o. Энэ нь о эгшигийн гол буюу (MF) хэлбэр. Жишээ нь:

tl. Oi-Ro;
tc.oi-ro.

6.a. 

tl. Ō; **tc.** mong. ō, sans. diphtongue au. Энэ нь монгол хэлний урт о болон самгард хэлний au хос эгшигийн бие даасан (IF) хэлбэр. Жишээлбэл:

tl. ō-sor;
tc.ō-sor.

tl. ō; **tc.** ō. Энэ нь урт о эгшигийн гол буюу (MF) хэлбэр. Жишээ нь:

tl. Kō-son;
tc. kō-son.

7. 

tl. Ö; **tc.** mong. ö. Энэ нь ö эгшигийн бие даасан (IF) хэлбэр бөгөөд зөвхөн монгол хэлнээ тохиолдоно. Жишээлбэл:

tl. Ö-mög;
tc. ö-mög.

tl. ö; **tc.** ö. Монгол хэлний ö эгшигийн гол буюу (MF) хэлбэр. Жишээ нь:

tl. Ö-mö-nö;
tc. ö-mö-nö.

7 a. 

tl. Ō; **tc.** mong. ō. Монгол хэлний урт ö эгшигийн бие даасан (IF) хэлбэр. Жишээлбэл:

tl. Ōr;
tc. ōr.

Figure 9: Description of Soyombo vowels (from Shagdarsürüng 2001: 137).

tl. o̲; **tc.** o̲. Урт ө эгшигийн гол буюу (MF) хэлбэр. Жишээ нь:

tl. Bo̲r;
tc. bo̲r.

8. 

tl. Au_2. **tc.** mong. au буюу diphtongue 'waw '. Энэхүү хос эгшигийн талаар цагаан толгойн жагсаалтын 5-ын u_2-аас лавшруулан үзнэ үү.

9. 

tl. Ai_2; **tc.** mong. ai буюу diphtongue 'yod'. Энэхүү хос эгшигийн талаар цагаан толгойн жагсаалтын 2-ын i_2-аас лавшруулан үзнэ үү.

10. 

tl. mong. G(a), tib., sans: K(a); **tc.** mong. ga, tib., sans; ka. Энэ нь монгол хэлний g(a), төвөд, самгардын к(a) гийгүүлэгчийн бие даасан (IF) хэлбэр.

tl. Ga_0-Ja_0r;
tc. ga-jar.

tl. g; **tc.** g. Энэ нь монгол хэлний гийгүүлэгчийн гол (MF) хэлбэр бөгөөд амьгүй буюу дэвсгэрлэж ордог хэлбэр. Жишээ нь:

tl. Ka_0-mu_1g;
tc. ka-mug.

tl. sans. k; **tc.** sans. k. Зөвхөн самгард хэлнээ тохиолдоно. Жишээ нь:

tl. kKa_0;
tc. kka.

11. 

tl. mong. K(a), tib., sans.: KH(a); **tc.** mong. ka, tib., sans.: kha.
Монгол хэлний ка; төвөд, самгард хэлний гийгүүлэгчийн бие даасан (IF) хэлбэр. Жишээлбэл:

tl. Ka_0-mu_1g;
tc. ka-mug.

tl. (k) / kh; **tc.** (k) / kh. Энэ нь (к) / kh гийгүүлэгчийн гол буюу (MF) хэлбэр. Монгол бичгийн хэлний зөв бичих зүйд энэ гийгүүлэгч дэвсгэрлэж ордоггүй. Харин Тод бичигт дэвсгэрлэж ордог ёсон буй.

12. 

tl. tib., sans: N(a); **tc.** tib., sans.: na. Энэ нь төвөд, самгард хэлний гүн угийн гийгүүлэгчийн бие даасан буюу (IF) хэлбэр. Монгол хэлнээ эл гийгүүлэгч нь үгийн болон үеийн эхинд тохиолддоггүй учир монгол бичвэрт ийм хэлбэр хэрэглэгдэхгүй.

tl. mong. n̲ / tib., sans.: ṅ; **tc.** mong. n̲ / tib., sans.: ṅ. Энэ нь n̲/ṅ гийгүүлэгчийн гол буюу (MF) хэлбэр. Монгол хэлнээ зөвхөн үеийн эцэст тохиолдоно.

tl. Ji_1 r-Ga_0-La_0 n̲;
tc. jir-ga-la̲n.

Figure 10: Description of Soyombo consonants (from Shagdarsürüng 2001: 138).

13. **tl.** mong. J(a), tib., sans.: C(a).
tc. mong. ja (i -гээс бусад эгшигийн өмнө) / ja(i эгшигийн өмнө), tib., sans.: ca. Энэ нь монгол хэлний ja/ ja, төвөд, самгардын са гийгүүлэгчийн бие даасан (IF) хэлбэр. Жишээлбэл:


tl. $Ja_0r\text{-}Li_1g$;
tc.jar-lig.


tl. $Ji_1r\text{-}Ga_0\text{-}La_0n$;
tc. jir-ga-lan.

14. **tl.** mong. C(a), tib., sans.: CH(a);
tc. mong. ca (i -гээс бусад эгшигийн өмнө) / ča (зөвхөн i эгшигийн өмнө); tib., sans.: cha. Энэ нь монгол хэлний са / ča, төвөд, самгард хэлний cha гийгүүлэгчийн бие даасан (IF) хэлбэр. Жишээлбэл:


tl. $Ca_0\gamma\text{-}ga_0\text{-}ca_0$;
tc. ca-ga-ca (<Mo. čaγ-ača).


tl. $Ci_1\text{-}Na_0r$;
tc. či-nar.


tl. $Ci_1u_2l\text{-}Ga_0\text{-}Nu_1$;
tc. čiul-ga-nu
(<Mo. ciγulγan-u).

15. **tl.** tib., sans.: $\bar{N}$(a);
tc. tib., sans.: ña. Төвөд, самгард хэлний гийгүүлэгчийн бие даасан (IF) хэлбэр.

16. **tl.** mong. D(a), tib., sans.: t(a);
tc. mong. da, tib., sans. ta. Монгол хэлний da, төвөд самгард хэлний ta гийгүүлэгчийн бие даасан буюу (IF) хэлбэр. Жишээлбэл:


tl. $Da_0\text{-}ga_0n$;
tc. da-gan (< Mo. daγa-).

tl. mong. d, tib., sans.: t;
tc.mong. d, tib., sans.: t. Монгол хэлний d, төвөд, самгард хэлний t гийгүүлэгчийн гол буюу (MF) хэлбэр. Жишээлбэл:


tl. Ki-Ged;
tc.ki-ged (< Mo. kiged).

17. **tl.** mong. T(a); tib., sans.: TH(a);
tc. mong. ta, tib., sans.: tha. Монгол хэлний ta, төвөд, самгард хэлний tha гийгүүлэгчийн бие даасан буюу (IF) хэлбэр. Жишээлбэл:


tl. $Ta_0\text{-}Ra_0g\text{-}Sa_0n$;
tc. ta-rag-san (< Mo. tara-).

18. **tl.** mong., tib., sans.: N(a);
tc. mong., tib., sans.: na. Монгол, төвөд, самгард хэлний na гийгүүлэгчийн бие даасан (IF) хэлбэр. Жишээ нь:


tl. $Na_0\text{-}Ra_0$;
tc. na-ra.

Figure 11: Description of Soyombo consonants (from Shagdarsürüng 2001: 139).

tl. n.
tc. n. Монгол, төвөд, самгард хэлний n гийгүүлэгчийн гол буюу (MF) хэлбэр. Жишээ нь:

tl. Na_0-Ya_0n;
tc. na-yan.

19. 

tl. mong. B(a), tib., sans.: P(a);
tc. mong. ba, tib., sans.: pa. Монгол хэлний ba, төвөд, самгард хэлний pa гийгүүлэгчийн бие даасан буюу (IF) хэлбэр. Жишээлбэл:

tl. Ba_0-Ri_1-
tc. ba-ri- (Mo. bari-).

tl. mong. b, tib., sans.: p;
tc. mong. b, tib., sans.: p.
Монгол хэлний b, төвөд, самгард хэлний p гийгүүлэгчийн гол буюу (MF) хэлбэр. Жишээлбэл:

tl. El-Deb;
tc. el-deb.

20. 

tl. mong. P(a), tib., sans.: PH(a);
tc. mong. pa, tib., sans.: pha. Монгол хэлний pa, төвөд, самгард хэлний pha гийгүүлэгчийн бие даасан буюу (IF) хэлбэр нь.

21. 

tl. mong., tib., sans.: M(a);
tc. mong., tib., sans.: ma. Монгол, төвөд, самгард хэлний ma гийгүүлэгчийн бие даасан (IF) хэлбэр. Жишээ нь:

tl. Ma_0-$Ši_1$;
tc. ma-si.

tl. m;
tc. m. Энэ нь m гийгүүлэгчийн гол буюу (MF) хэлбэр. Жишээлбэл:

tl. Er-Dem;
tc. er-dem.

tl. tib., sans.: (O)m;
tc. tib., sans. : (o)m. Энэ нь төвөд, самгард бичвэрт, тарнийн хэлнээ тохиолдоно. Уугуул монгол хэлнээ тохиолдохгүй.

22. 

tl. mong., tib.: Y(a), sans.: semi-vowel Y(a);
tc. mong., tib., ya, sans.: semi-vowel ya. Монгол, төвөд хэлний y гийгүүлэгч, самгард хэлний заримдаг (тал) y эгшигийн бие даасан (IF) хэлбэр.

tl. Ya_0-Ga_0u_2l;
tc. ya-gaul (Энэ бөлөгийн 15-р зүүлтээс тодруулна уу.)

Figure 12: Description of Soyombo consonants (from Shagdarsürüng 2001: 140).

tl. tib., sans.: y(a);
tc. tib., sans.: ya. Хэдийгээр энэ нь монгол, төвөд хэлний хувьд у гийгүүлэгч, самгард хэлний тухайд заримдаг (буюу тал) у эгшигийн гол (MF) хэлбэр мөн боловч хэрэглээ нь хязгаартай. Үүнд: соёмбо бичигийн цагаан толгойгоос үзэхэд нэгд, монгол хэлний дэвсгэр гийгүүлэгчид багтаагаагүй ажээ; хоёрт, төвөд болон самгард хэлнээ зүүлтэт үсэгт тохиолдохоор тэмдэглэжээ.

tl. Kya_0;
tc.kya.

23. 

tl. mong., tib., sans.: R(a);
tc. mong., tib.: ra, sans.: semi-vowel ra. Энэ нь монгол, төвөд хэлний ra гийгүүлэгч, самгард хэлний ra хэмээх заримдаг (буюу тал) эгшигийн бие даасан (IF) хэлбэр. Жишээлбэл:

tl. Ja_0u_2-Ra_0-Da_0;
tc.jau-ra-da (< Mo. jaγuradu ~ jaγurada)

tl. mong., tib.: r, sans.: (semi-vowel) r;
tc. mong., tib.: r, sans. (semi-vowel) r. Монгол, төвөд хэлний r гийгүүлэгч, самгард хэлний r хэмээх заримдаг (буюу тал) эгшигийн гол (MF) хэлбэр нь. Монгол хэлнээ дэвсгэрлэж ороходоо энэхүү гол хэлбэрийг дундуур нь таллаж, тал гурвалжин хэлбэрээр тэмдэглэнэ. Жишээ нь:

tl. Er-Dem;
tc. er-dem.

Харин төвөд, самгард хэлний давхар үсгийн тухайд, r гийгүүлэгч (буюу заримдаг эгшиг)-ийн гол хэлбэрийг хялбарчлахгүй, яг хэвээр нь толгой, зүүлт болгож бичдэг. Жишээлбэл:

tl. Kra_0;
tc. kra.

tl. rKa_0;
tc.rka.

24. 

tl. tib., V(a), sans.: (semi-vowel): V(a);
tc. tib. va, sans.: (semi-vowel): va. Төвөдийн va гийгүүлэгч, самгардын заримдаг (буюу тал) va эгшигийн бие даасан (IF) хэлбэр нь.

tl. tib., sans.: v(a);
tc. tib., sans.: va.
Энэ нь төвөд, самгардын давхар үсэгт v(a) зүүлт болж ордог гол (MF) хэлбэр. Үүнийг эл бөлөгийн No. 63-аас тодруулан үзмүү.

25. 

tl. mong., tib.: L(a), sans.: (semi-vowel): L(a);
tc. mong., tib.: la. sans.: (semi-vowel): la. Энэ нь Монгол, төвөд хэлний l гийгүүлэгч, самгард хэлний l хэмээх заримдаг (буюу тал) эгшигийн бие даасан (IF) хэлбэр нь. Жишээлбэл:

tl. Nom-La_0g-Sa_0n;
tc. nom-lag-san.

Figure 13: Description of Soyombo consonants (from Shagdarsürüng 2001: 141).

tl. mong., tib.: l, sans.: (semi-vowel): l;
tc. mong., tib.: l, sans.: (semi-vowel): l.
Энэ нь монгол, төвөд хэлний l гийгүүлэгч, самгардын заримдаг (буюу тал) l эгшигийн гол (MF) хэлбэр. Жишээлбэл:

tl. Bel-ge;
tc. bel-ge.

Харин энэхүү (MF) хэлбэр нь төвөд, самгардийн давхар үсэгт толгой болох ороходоо ганц хөндлөн зураас болж ордог. Тухайлбал:

tl. lKa$_0$;
tc. lka.

26. 

tl. mong. Š(a), tib., sans. Ç(a) / Ś(a).
tc. mong. ša, tib., sans.: ça / śa. Энэ нь монгол хэлний ša, төвөд, самгард хэлний ça буюу śa гийгүүлэгийн бие даасан (IF) хэлбэр. Жишээлбэл:

tl. Teg-Še;
tc. teg-še.

tl. mong. š;
tc. mong. š. Энэ нь š гийгүүлэгчийн монгол хэлэнд тохиолдох гол буюу (MF) хэлбэр. Соёмбо бичигийн цагаан толгойноос үзэхэд үүнийг монгол хэлний дэвсгэр үсэгт багтаажээ (Тод бичигт ч бас ийм буй). Тухайлбал:

tl. Aš;
tc. aš.

Самгард хэлнээ, давхар үсэгт энэ гийгүүлэгчийг толгой болгож залгахдаа (MF) хэлбэрийг таллаж арай хялбарчлан тэмдэглэдэг бөлгөө.

tl. çKa$_0$;
tc. çka.

27. 

tl. mong., tib., sans. S(a);
tc. mong., tib., sans. sa. Sa гийгүүлэгчийн бие даасан (IF) хэлбэр.

tl. Sa$_0$-Ra$_0$;
tc. sa-ra.

tl. s; **tc.** s. Энэ нь sa гийгүүлэгчийн гол буюу (MF) хэлбэр. Монгол хэлний тухайд дэвсгэрлэж орох хэлбэр гэсэн үг. Жишээлбэл:

tl. Te-güs;
tc. te-güs.

Төвөд, самгард хэлний давхар үсэгт sa толгойг бичихдээ бяцхан гурвалжин болгож доорхи байдлаар тэмдэглэнэ.

Figure 14: Description of Soyombo consonants (from Shagdarsürüng 2001: 142).

tl. sKa$_o$;
tc. ska.

28. 

tl. mong., tib., H(a), sans.: (semi-vowel): H(a);
tc. mong., tib. ha, sans.: (semi-vowel): ha. Энэ нь монгол, төвөд хэлний ha гийгүүлэгч, самгард хэлний заримдаг (буюу тал) ha эгшигийн бие даасан (IF) хэлбэр.

29. 

tl. sans. (compound consonant): KS(a);
tc. sans. (compound consonant): ksa.
Самгард хэлний нийлмэл гийгүүлэгч ksa -гийн бие даасан (IF) хэлбэр.

30. 

tl. mong., Ag, tib., sans.: Ak;
tc. mong. ag, tib., sans.: ak. Монгол хэлний дэвсгэр -g гийгүүлэгчийн хэлбэр (Уг үсгийн талаар жагсаалтын No. 10-аас тодруулан үзнэ үү).

31. 

tl. mong. Ak, tib., sans.: Akh;
tc. mong. ak, tib., sans. akh. Дэвсгэрлэж орох -k гийгүүлэгчийн хэлбэр (Уг үсгийн талаар жагсаалтын No. 11-аас тодруулан үзнэ үү).

32. 

tl. mong. An̲, tib., sans.: Aṅ;
tc. mong. an̲, tib., sans. aṅ. Монгол хэлний дэвсгэр -n гийгүүлэгчийн хэлбэр (Уг үсгийн талаар жагсаалтын No. 12-аас тодруулан үзнэ үү).

33. 

tl. mong. Ad, tib., sans.: At;
tc. mong. ad, tib., sans.: at; Монгол хэлний дэвсгэр -d гийгүүлэгчийн хэлбэр (Уг үсгийн талаар жагсаалтын No. 16-аас тодруулан үзнэ үү).

34. 

tl. mong. An, tib., sans.: An;
tc. mong., an, tib., sans.: an. Монгол хэлний дэвсгэр -n гийгүүлэгчийн хэлбэр (Уг үсгийн талаар жагсаалтын No. 18-аас тодруулан үзнэ үү).

35. 

tl. mong. Ab, tib., sans.: Ap;
tc. mong. ab; tib., sans.: ap. Монгол хэлний дэвсгэр -b гийгүүлэгчийн хэлбэр (Уг үсгийн талаар жагсаалтын No. 19-аас тодруулан үзнэ үү).

36. 

tl. mong., tib., sans.: Am;
tc. mong., tib., sans.: am. Монгол хэлний дэвсгэр -m гийгүүлэгчийн хэлбэр (Уг үсгийн талаар жагсаалтын No. 21-аас тодруулан үзнэ үү).

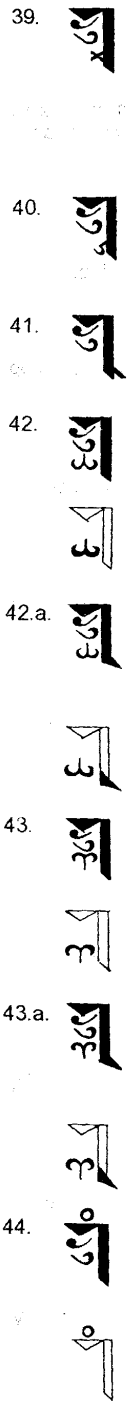
37. 

tl. mong., tib., sans.: Ar;
tc. mong., tib., sans.: ar. Монгол хэлний дэвсгэр -r гийгүүлэгчийн хэлбэр (Уг үсгийн талаар жагсаалтын No. 23-аас тодруулан үзнэ үү).

38. 

tl. mong., tib., sans.: Al;
tc. mong., tib., sans.: al. Монгол хэлний дэвсгэр -l гийгүүлэгчийн хэлбэр (Уг үсгийн талаар жагсаалтын No. 25-аас тодруулан үзнэ үү).

Figure 15: Description of Soyombo consonants (from Shagdarsürüng 2001: 143).



39. **tl.** mong., tib., sans.: Aš;
tc. mong., tib., sans. aš. Дэвсгэрлэж орох -š гийгүүлэгчийн хэлбэр (Уг үсгийн талаар жагсаалтын No. 26-аас тодруулан үзнэ үү). Монгол бичигийн зөв бичих зүйд энэхүү š гийгүүлэгч дэвсгэрлэж ордоггүй; харин тод бичгт дэвсгэрлэдэг бөгөөд "арааны сийгээн" гэдэг тусгай нэр устай. Энэ талаар "Тод бичиг" хэмээх бөлөгийн холбогдох хэсгээс тодруулан үзмүү.

40. **tl.** mong., tib., sans.: As;
tc. mong., tib., sans. as. Монгол хэлний дэвсгэр -s гийгүүлэгчийн хэлбэр (Уг үсгийн талаар жагсаалтын No. 27-аас тодруулан үзнэ үү).

41. **tl.** mong. An̲; tib. A'n̲ (ཨཾ).
tc. mong. ān̲. < (?) Mo. -iyan/-iyen; -ban/-ben.

42. **tl.** sans. Ṛ ~Ṛi;
tc. ṛ. ~ ṛi. Самгард хэлний ṛ ~ ṛi эгшигийн бие даасан буюу (IF) хэлбэр.

tl. sans. ṛ ~ ṛi;
tc. sans. ṛ ~ ṛi. Самгард хэлний ṛ ~ ṛi эгшигийн гол буюу (MF) хэлбэр.

42.a. **tl.** sans. R̄ ~ R̄i;
tc. sans. r̄ ~ r̄i. Самгард хэлний урт ṛ ~ṛi эгшигийн бие даасан буюу (IF) хэлбэр.

tl. sans. r̄ ~ r̄i;
tc. sans. r̄ ~ r̄i. Самгард хэлний урт ṛ ~ ṛi эгшигийн гол буюу (MF) хэлбэр.

43. **tl.** sans. Ḷ ~ Ḷi;
tc. sans. ḷ ~ ḷi. Самгард хэлний ḷ ~ḷi эгшигийн бие даасан буюу (IF) хэлбэр.

tl. sans. Ḷ ~Ḷi;
tc. sans. ḷ ~ ḷi. Самгард хэлний ḷ ~ ḷi эгшигийн гол буюу (MF) хэлбэр.

43.a. **tl.** sans. L̄ ~ L̄i;
tc. sans. l̄ ~ l̄i. Самгард хэлний урт ḷ ~ḷi эгшигийн бие даасан буюу (IF) хэлбэр.

tl. sans. l̄ ~ l̄i;
tc. sans. l̄ ~ l̄i. Самгард хэлний урт ḷ ~ ḷi эгшигийн гол буюу (MF) хэлбэр.

44. **tl.** sans. Aṃ ~ Aṁ;
tc. sans. aṃ ~ aṁ. Самгард хэлний anu-svara -гийн бие даасан буюу (IF) хэлбэр.

tl. sans. ṃ ~ ṁ;
tc. sans. ṃ ~ ṁ.

Figure 16: Description of Soyombo consonants (from Shagdarsürüng 2001: 144).

Самгард хэлний anu-svara -гийн гол буюу (MF) хэлбэр. Жишээлбэл:

tl. Sva_0-ya_0ṁ-bhü;
tc. sva-yaṁ-bhü.

45. 
tl. sans. AḤ;
tc. sans. aḥ. Самгард хэлний vi-sarga гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

46. 
tl. tib., sans.: G(a);
tc. tib., sans.: ga. Төвөд, самгард хэлний ga гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

47. 
tl. sans. GH(a);
tc. sans. gha. Самгард хэлний gha гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

48. 
tl. tib., sans.: J(a);
tc. tib., sans.: ja. Самгард хэлний ja гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

49. 
tl. sans. JH(a);
tc. sans. jha. Самгард хэлний jha гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

50. 
tl. sans. Ṭ(a);
tc. sans. ṭa. Самгард хэлний ṭa гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

51. 
tl. sans. ṬH(a);
tc. sans. ṭha. Самгард хэлний ṭha гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

52. 
tl. sans. Ḍ(a);
tc. sans. ḍa. Самгард хэлний ḍa гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

53. 
tl. sans. ḌH(a);
tc. sans. ḍha. Самгард хэлний ḍha гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

54. 
tl. sans. Ṇ(a);
tc. sans. ṇa. Самгард хэлний ṇa гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

55. 
tl. tib., sans.: D(a);
tc. tib., sans.: da.
Төвөд, самгард хэлний da гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

Figure 17: Description of Soyombo consonants (from Shagdarsürüng 2001: 145).

56. 
tl. sans. DH(a);
tc. sans. dha. Самгард хэлний dha гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

57. 
tl. tib., sans.: B(a);
tc. tib., sans.: ba. Төвөд, самгард хэлний ba гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

58. 
tl. sans. BH(a);
tc. sans. bha. Самгард хэлний bha гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

59. 
tl. sans. Ṣ(a);
tc. sans. ṣa. Самгард хэлний ṣa гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

60. 
tl. tib., sans. Ky(a);
tc. tib., sans. kya. Төвөд, самгард хэлний kya гийгүүлэгчийн бие даасан буюу (IF) хэлбэр. (Энэ жагсаалтын No. 22-оос лавлагтун!)

61. 
tl. tib., sans.: Kr(a);
tc. tib., sans.: kra. Төвөд, самгард хэлний kra гийгүүлэгчийн бие даасан буюу (IF) хэлбэр. (Энэ жагсаалтын No. 23-аас лавлагтун!)

62. 
tl. tib., sans.: Kl(a);
tc. tib., sans.: kla. Төвөд, самгард хэлний kla гийгүүлэгчийн бие даасан буюу (IF) хэлбэр. (Энэ жагсаалтын No. 25-аас лавлагтун!)

63. 
tl. tib., sans.: Kv(a);
tc. tib., sans.: kva. Төвөд, самгард хэлний kva гийгүүлэгчийн бие даасан буюу (IF) хэлбэр. (Энэ жагсаалтын No. 24-өөс лавлагтун!)

64. 
tl. sans. kK(a);
tc. sans. kka. Самгард хэлний kka гийгүүлэгчийн бие даасан буюу (IF) хэлбэр. (Энэ жагсаалтын No. 10-аас лавлагтун!)

65. 
tl. sans. Ṅ K(a);
tc. sans. ṅka. Самгард хэлний нийлмэл ṅka гийгүүлэгчийн бие даасан буюу (IF) хэлбэр. (Энэ жагсаалтын No. 10, 12-оос лавлагтун!)

66. 
tl. sans. N̄C(a);
tc. sans. n̄ca. Самгар хэлний нийлмэл nca гийгүүлэгчийн бие даасан буюу (IF) хэлбэр. (Энэ жагсаалтын No. 13, 15-аас лавлагтун!)

67. 
tl. sans. Ṇ Ṭ(a);
tc. sans. ṇṭa. Самгард хэлний нийлмэл ṇṭa гийгүүлэгчийн бие даасан буюу (IF) хэлбэр. (Энэ жагсаалтын No. 50, 54-өөс лавлагтун!)

Figure 18: Description of Soyombo consonants (from Shagdarsürüng 2001: 146).

68. 
tl. sans. NT(a);
tc. sans. nta. Самгард хэлний нийлмэл nta гийгүүлэгчийн бие даасан буюу (IF) хэлбэр. (Энэ жагсаалтын No. 16, 18-аас лавлагтун!)

69. 
tl. sans. MP(a);
tc. sans. mpa. Самгард хэлний нийлмэл mpa гийгүүлэгчийн бие даасан буюу (IF) хэлбэр. (Энэ жагсаалтын No. 19, 21-ээс лавлагтун!)

70. 
tl. tib., sans.: lK(a);
tc. tib., sans.: lka. Төвөд, самгард хэлний lka гийгүүлэгчийн бие даасан буюу (IF) хэлбэр. (Энэ жагсаалтын No. 25-аас лавлагтун!)

71. 
tl. sans. çK(a) ~ śK(a);
tc. sans. çka ~ śka. Самгард хэлний çka ~ ska гийгүүлэгчийн бие даасан буюу (IF) хэлбэр. (Энэ жагсаалтын No. 10, 26-аас лавлагтун!)

72. 
tl. tib., sans.: sK(a);
tc. tib., sans.: ska. Төвөд, самгард хэлний ska гийгүүлэгчийн бие даасан буюу (IF) хэлбэр. (Энэ жагсаалтын No. 10, 27-оос лавлагтун!)

73. 
tl. tib., sans.: rK(a);
tc. tib., sans.: rka. Төвөд, самгард хэлний rka гийгүүлэгчийн бие даасан буюу (IF) хэлбэр. (Энэ жагсаалтын No. 10, 23-аас лавлагтун!)

74. 
tl. tib. Č(a);
tc. tib. ča. Зөвхөн төвөд хэлний ča гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

75. 
tl. tib. ČH(a);
tc. tib. čha. Зөвхөн төвөд хэлний čha гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

76. 
tl. tib. J(a);
tc. tib. ja. Зөвхөн төвөд хэлний ja гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

77. 
tl. tib. Ž(a);
tc. tib. ža. Зөвхөн төвөд хэлний ža гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

78. 
tl. tib. Z(a);
tc. tib. za. Зөвхөн төвөд хэлний za гийгүүлэгчийн бие даасан буюу (IF) хэлбэр.

Figure 19: Description of Soyombo consonants (from Shagdarsürüng 2001: 147).

79. 
tl. tib. '(a);
tc. tib. 'a. Төвөд хэлний "шанаган a" хэмээх 'a-čhun буюу 'a гийгүүлэгчийн бие даасан (IF) хэлбэр.

80. 
Номыг төгсгөх тэмдэг.

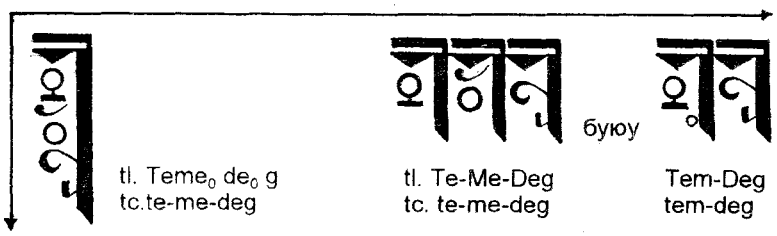
Figure 20: Description of Soyombo consonants (from Shagdarsürüng 2001: 148).

Figure 21: Character elements used in Soyombo characters (from Shagdarsürüng 2001: 153).

Энэхүү тамгын дардас дээрхи бичвэрийг латин галигаар сэргээн үзүүлбэл доорхи мэт болно.

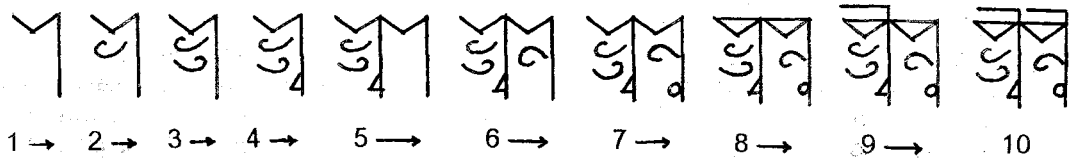
I. naya - selel
II. kör - yenu jür - ka
III. grva - chanun temedeg

Үүнээс харахул, "босоо" соёмбоор бичихэд монгол хэлний эгшиг зохицох ёсыг ашиглаж өгүүлэх эрхтэний ойр төстэй оролцоогоор нь нэгэн толгойд бөлөглөж бичдэг зарчим байсан нь харагдана. Тухайлбал: So. temedeg (<Mo. temdeg) гэдэг үгийг "босоо" болон "хэвтээ"-гээр хэрхэн бичих байдлыг харьцуулан үзүүлэе.



Соёмбо бичгээр хичээнгүйлэн дармаллаж бичихийн хажуугаар ер энгийн байдлаар татлан бичдэг бөгөөд аливаа бичиг үсэгт хэрхэн бичдэг *Үсэгийн дүрэм* байдагчилан соёмбо бичигт ч тусгайлан баримтлах тогтсон журам буй. Үүнийг "er-dem" гэдэг үгээр жишээлэн үзүүлсүгэй.

er-dem:



Соёмбоор буй монгол хэлний дурсгалын зүйлийг хэлний талаас авч үзэхүл соёмбо бичгээр монгол үгийг бичихдээ тухайн цагийн монгол хэлний аман дуудалгын байдал

Figure 22: Method of writing Soyombo vertically (from Shagdarsürüng 2001: 150). The specimen at top is a seal containing vertical text in which the syllables of words are written as conjuncts.

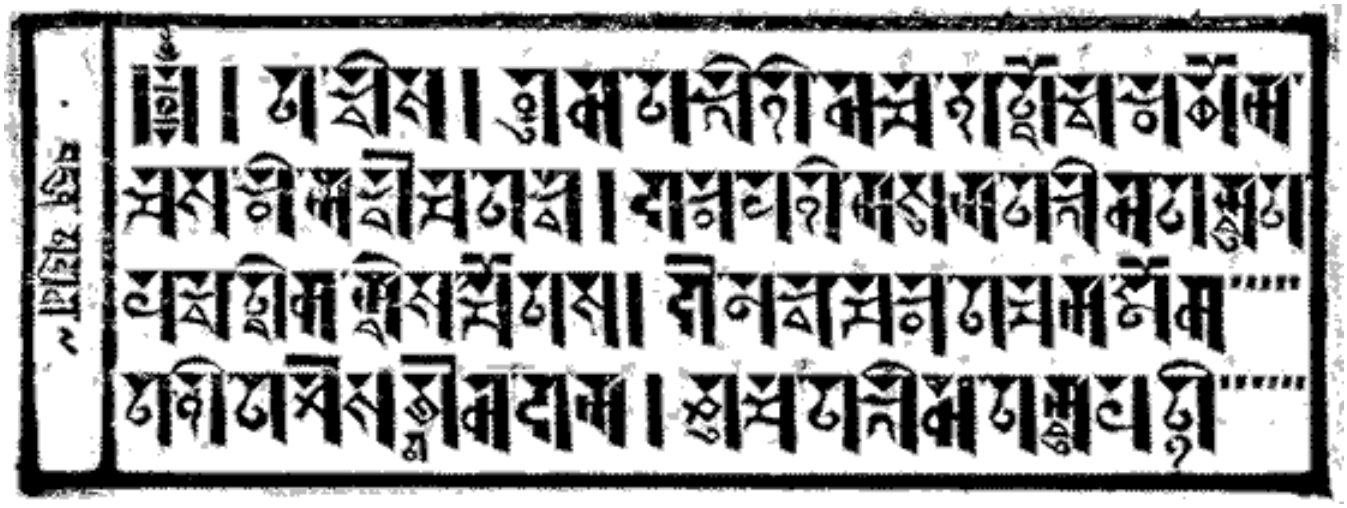


Figure 23: A folio from a Soyombo manuscript (from Mongolwiki 2008).

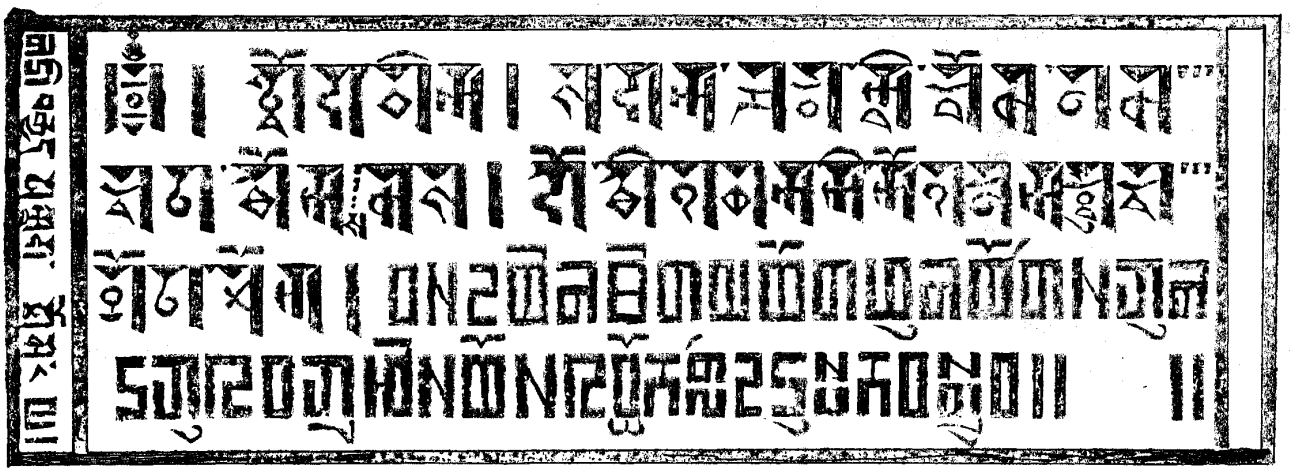


Figure 24: A manuscript containing text in Soyombo and Mongolian Square (from Shagdarsürüng 2001: 173).

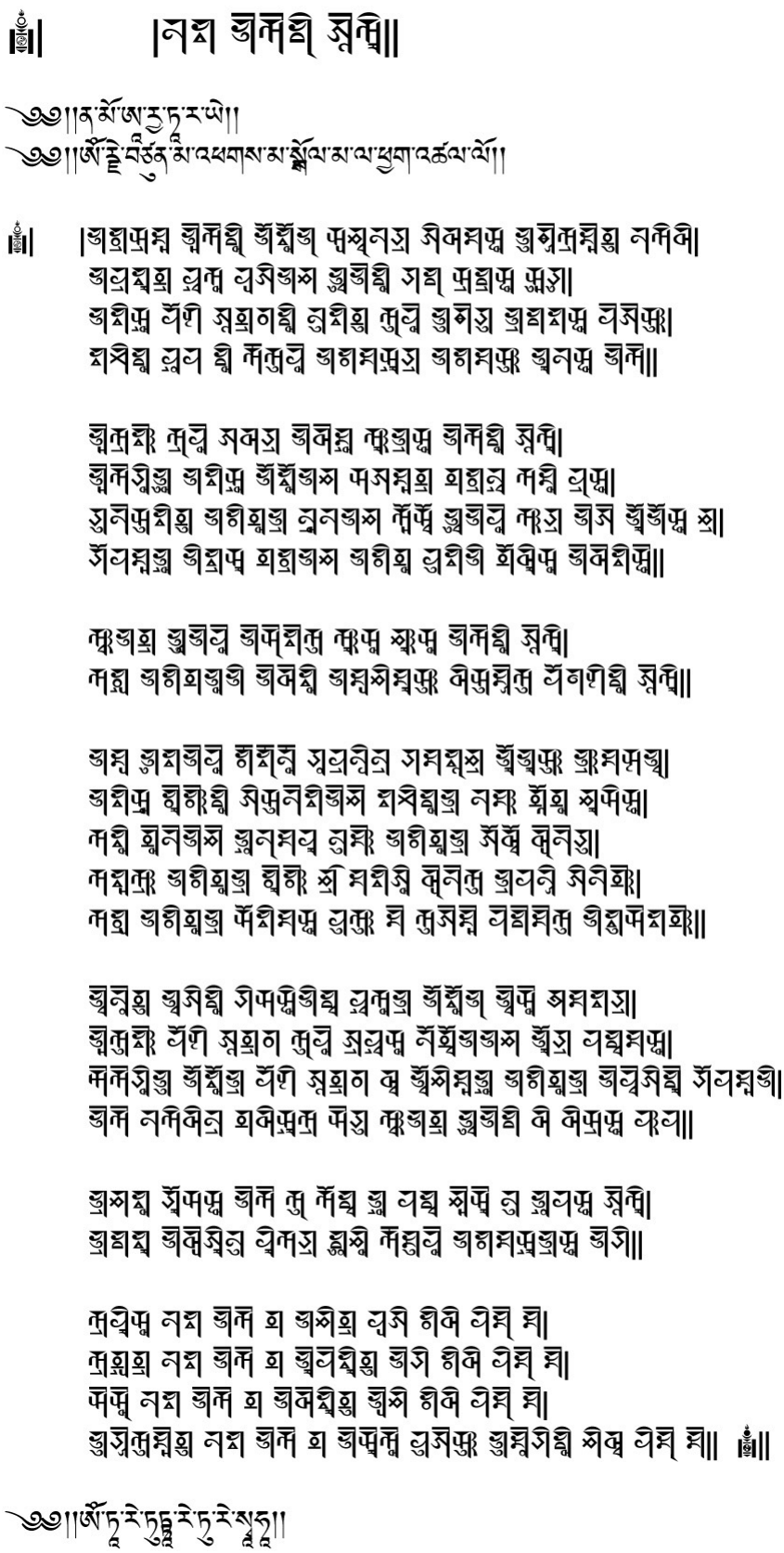


Figure 25: Poem for the Green Tara typeset in a digitized Soyombo font (from Wikimedia 2009b).

Figure 26: Xylograph (block print) of a book cover in Soyombo, Mongolian Square, Mongolian, and Cyrillic (from Boldsaikhan 2005: 330). The title is Sanskrit written in Soyombo: *Mongal-svayaṃbhu-jyoti-varṇa-lipiḥ*. The Mongolian Square represents Tibetan, the Mongolian represents Mongolian, and the Cyrillic represents Modern (Khalkha) Mongolian.

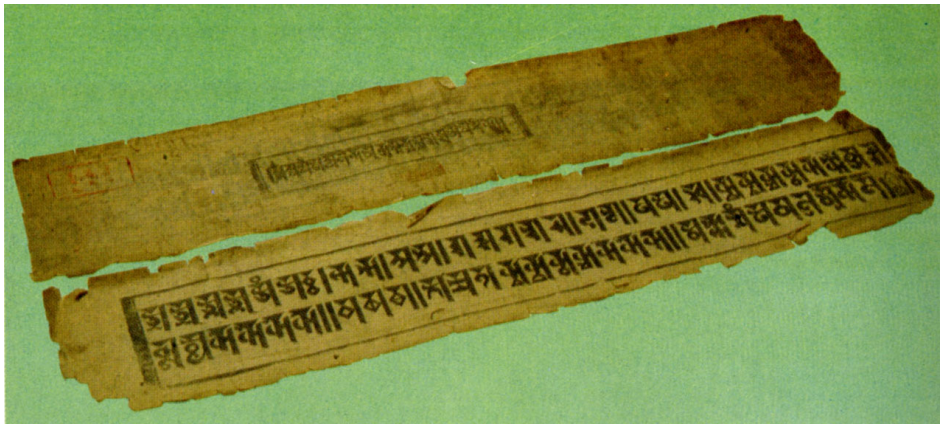


Figure 27: Folios of a Soyombo manuscript (from Boldsaikhan 2005: 378).

Figure 28: Folios of a Soyombo manuscript (from Boldsaikhan 2005: 378).

Figure 29: The SVAYAMBHU symbol with Mongolian text (from Rintschen 1953: 8).

Figure 30: The flag of Mongolia showing the SVAYAMBHU symbol (from Wikimedia 2008).

Figure 31: The coat of arms of Mongolia showing the SVAYAMBHU symbol (from Wikimedia 2009a).

Figure 32: An imprint of the SVAYAMBHU symbol in the center on the recto face of a 1,000 Mongolian tögrög (tugrik) note (from Wikimedia 2006).

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://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*.**

A. Administrative

1. **Title:** *Revised Proposal to Encode the Soyombo Script in ISO/IEC 10646*
2. Requester's name: *Script Encoding Initiative (SEI) / Anshuman Pandey (pandey@umich.edu)*
3. Requester type (Member body/Liaison/Individual contribution): *Liaison contribution*
4. Submission date: *2013-04-22*
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): *Yes*
      Proposed name of script: *Soyombo*
   b. The proposal is for addition of character(s) to an existing block:
      Name of the existing block:
2. Number of characters in proposal: *117*
3. Proposed category (select one from below - see section 2.2 of P&P document):
   A-Contemporary ; B.1-Specialized (small collection) ; B.2-Specialized (large collection) X
   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?
   Anshuman Pandey
   b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
   Anshuman Pandey (pandey@umich.edu)
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)? *Yes*

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: N3902-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)

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? *György Kara <gkara@indiana.edu>*
   If YES, available relevant documents: *See text of proposal*
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? *Yes*
   Reference: *Size of user community is unknown. Script is used in print and digital publications.*
4. The context of use for the proposed characters (type of use; common or rare) *Common*
   Reference: *The script is used for writing Mongolian, Sanskrit, and Tibetan.*
5. Are the proposed characters in current use by the user community? *Yes*
   If YES, where? Reference: *Users worldwide engaged in Mongolian studies*
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 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? *Yes*
    If YES, is a rationale for such use provided? *Yes*
    If YES, reference: *Combining vowel and final-consonant signs*
    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: