

# ISO/IEC JTC1/SC2/WG2 N2719

L2/04-112
-----------

<b>Title</b>	<b>Response to Comments on Phags-pa Proposal in N2706</b>
<b>Source</b>	Andrew C. West
<b>Document Type</b>	Individual Contribution
<b>Date</b>	10th March 2004

## Background

This document presents clarifications on the proposal to encode the Phags-pa script given in document N2622, in response to the eight points raised by the Chinese National Body in document N2706 :

1. Alphabet not integrated, 2 vowels ö and ü used for recording Mongolian missing.
2. Some presentation forms are regarded as cononical forms by error, e.g., A868 and A869 are presentation forms A85C, they should not be placed at current locations.
3. Letters sorting wrong. For instance: AB68 AB69 AB6A AB6B A862 A863 A864 A865.
4. Typeface in this PDAM had never been used in history of Phags-pa script.
5. Some characters' shapes are not the best choice for cononical characters, shapes selected should be those appear frequently. For instance: A851 and A86C.
6. Character names should conform to that of Mongolian (1800-18AF). For instance: A86F ANUSVARA, A870 BIRGA, etc..
7. At least one "Free Variation Form Selector" is needed. A "Syllable Connector" different to "general space" is also needed.
8. Phags-pa script should not be regarded as variant of Tibetan script.

Before addressing each of these points individually, I would like to take this opportunity to acknowledge my indebtedness to Chinese scholarship on the Phags-pa script. My understanding of the Phags-pa script has been largely based on the work of Chinese scholars, in particular Professor Junast 照册斯图, China's foremost expert on the Phags-pa script. I warmly welcome comments on my proposal from my Chinese colleagues, and hope that together we can agree on a sensible and workable encoding of the Phags-pa script in order to facilitate the advancement of Phags-pa studies both within China and throughout the world.












## Point 1

*Alphabet not integrated, 2 vowels ö and ü used for recording Mongolian missing.*












The Mongolian script has seven basic vowel letters : **a** [U+1820], **e** [U+1821], **i** [U+1822], **o** [U+1823], **u** [U+1824], **ö** [U+1825] and **ü** [U+1826]. On the other hand the Phags-pa script only has five vowel letters : **i** [A86A PHAGS-PA LETTER I], **u** [A86B PHAGS-PA LETTER U], **e** [A86C PHAGS-PA LETTER E], **o** [A86D PHAGS-PA LETTER O] and **e** [A86E PHAGS-PA LETTER EE]. In Mongolian texts written using the Phags-pa script the vowel sounds ö and ü are represented by the digraphs **eo** <A86E A86D> and **eu** <A86E A86B> respectively. For example :

- Classical Mongolian **ögbei** ᠣᠭᠪᠡᠢ <1825 182D 182A 1821 1822> "gave" is written as '**eo** **be**' <A861 A86E A86D A842 0020 A852 A86E A86C> in the Phags-pa script.
- Classical Mongolian **ülü** ᠤᠯᠤ <1826 182F 1826> "not" is written as '**eu** **lu**' <A861 A86E A86B 0020 A85D A86B> in the Phags-pa script.

My position with regards to these two digraphs follows established Chinese scholarship, as can be seen from the following remarks by Professor Junast :

The Phags-pa script has five vowels :  [i] <A86A>,  [u] <A86B>,  [e] <A86C>,  [o] <A86D> and  [e] <A86E>. ... In Phags-pa orthography certain sounds are represented by means of digraphs, including both vowels and consonants. For example, the Mongolian "feminine" vowels ö and ü are represented by the combination sequences  <A86E> and  <A86D> (i.e.  eo) and  <A86E> and  <A86B> (i.e.  eu) respectively.

Junast, "Basibawen yuanyin zimu zixing wenti shang de liangzhong tixi"; in Minzu Yuwen 1987.4.

八思巴文有五个元音字母： [i] <A86A>， [u] <A86B>， [e] <A86C>， [o] <A86D>， [e] <A86E>。... 在八思巴文的书写法中也用双字母表示某些语音，其中包括元音和辅音，例如蒙古语的阴性元音 ö 和 ü，分别用  <A86E> 与  <A86D> 的组合即  eo 和  <A86E> 与  <A86B> 的组合即  eu 表示。




照那斯图，《八思巴文元音字母字形问题上的两种体系》。载《民族语文》1987.4。

The Phags-pa digraphs **eo** and **eu** are also used to represent the vowels **ö** and **ü** in Chinese words, for example the Chinese ideographs 荀 *xún* and 卷 *juǎn* are represented in the Phags-pa script as **seun** <A85F A86E A86B A84F> and **geon** <A842 A86E A86D A84F> respectively. In the edition of the 14th-century Phags-pa rhyming dictionary of Chinese, Menggu Ziyun 蒙古字韵, edited by Professors Junast and Yang Naisi 杨耐思 (Beijing, 1987) the Phags-pa spellings for these two ideographs are transliterated as **sèun** and **gèon** (in this work A86C is transliterated as **e**, and A86E is transliterated as **è**), again indicating that the **ö** and **ü** sounds are regarded as being represented by digraphs in the Phags-pa script.



In conclusion, I believe that it is unnecessary and inappropriate to encode pre-composed "eo" and "eu" digraphs to represent the Mongolian vowel sounds "ö" and "ü".

## Point 2

*Some presentation forms are regarded as cononical forms by error, e.g., A868 and A869 are presentation forms A85C, they should not be placed at current locations.*

A868  [PHAGS-PA SUBJOINED LETTER RA] and A869  [PHAGS-PA SUPERFIXED LETTER RA] are proposed for encoding in addition to A85C  [PHAGS-PA LETTER RA]. The rationale for encoding these forms of the Phags-pa letter RA was provided in Section 6 of document N2622, but I will attempt to further clarify why it is necessary to encode the superfixed and subjoined forms of the letter RA separately.

To put it simply, as the Phags-pa script does not have an explicit letter to represent the vowel **a** it would be impossible to distinguish, and hence correctly render, normal and subjoined/superfixed forms of the letters RA in a syllable with an inherent "a" vowel. Take for example the following two Tibetan words :

- Tibetan **rang**  <0F62 0F44> "self" = Phags-pa **rang** <A85C A843> (normal RA)
- Tibetan **rnga**  <0F62 0F94> "drum" = Phags-pa **rnga** <A869 A843> (superfixed RA)

If only a single form of the Phags-pa letter RA were to be encoded, both **rang** and **rnga** would have to be represented as <A85C A843> in the Phags-pa script, even though the two words should be rendered distinctly, the former with the normal form of the letter RA, the latter with the superfixed T-shaped form of the letter RA.

Likewise, the subjoined forms of the letters WA, YA and RA need to be encoded separately in order for the rendering engine to be able to correctly render the form of these letters in words with an inherent "a" vowel. Compare the following pairs of Phags-pa syllables :

- Chinese **hǎi** 海 "sea" = Phags-pa **hay** <A860 A85B> (normal YA)
- Chinese **xià** 夏 "summer" = Phags-pa **hya** <A860 A867> (subjoined YA)
- Chinese **kǎo** 考 "to examine" = Phags-pa **khaw** <A841 A857> (normal WA)
- Chinese **kuā** 誇 "to boast" = Phags-pa **khwa** <A841 A866> (subjoined WA)
- Sanskrit **sar** = Phags-pa **sar** <A85F A85C> (normal RA)
- Sanskrit **sra** = Phags-pa **sra** <A85F A868> (subjoined RA)

In each case it would be impossible to distinguish or correctly render the Phags-pa words with a subjoined WA/YA/RA unless the subjoined forms are encoded separately. This is illustrated by the following detail from the mid 14th-century Sanskrit Phags-pa inscription on the East Wall of Juyong Guan 居庸关 in which the Sanskrit syllables **sra** (in the word **sa ha sra**) and **sar** (in the word **sar wa**) occur in close proximity :



As there is no explicit vowel **a** in the Phags-pa script, the two syllables **sra** and **sar** are written as the letter SA followed by a subjoined letter RA <A85F A868> and the letter SA followed by a normal letter RA <A85F A85C>

respectively. If the subjoined letter RA was considered to be simply a presentation form of the letter RA, both syllables would have to be encoded as <A85F A85C>, and it would be impossible for a rendering engine to render both syllables correctly.

## Point 3

*Letters sorting wrong. For instance: AB68 AB69 AB6A AB6B A862 A863 A864 A865.*

There are two 14th-century Chinese sources that list the forty-one original Phags-pa letters (the History of the Yuan Dynasty states that the Phags-pa Lama devised a script comprising 41 letters, but does not list them) :

- Fashu Kao 法书考: a work on calligraphy composed by the Yuan dynasty Uighur official Sheng Ximing 盛熙明, first published in 1334.
- Shushi Huiyao 书史会要: a work on the history of calligraphy by the late Yuan / early Ming author Tao Zongyi 陶宗仪, first published in 1376, eight years after the fall of the Yuan dynasty.

These two works both list the forty-one Phags-pa letters according to the traditional order of the Tibetan letters :

1. KA ༀ <A840>	15. BA ༂ <A852>	29. HA ༌ <A860>
2. KHA ༆ <A841>	16. MA ༄ <A853>	30. A ༎ <A861>
3. GA ༈ <A842>	17. TSA ༊ <A854>	31. I ༐ <A86A>
4. NGA ༌ <A843>	18. TSHA ༎ <A855>	32. U ༒ <A86B>
5. CA ༐ <A844>	19. DZA ༔ <A856>	33. E ༖ <A86C>
6. CHA ༒ <A845>	20. WA ༘ <A857>	34. O ༘ <A86D>
7. JA ༔ <A846>	21. ZHA ༚ <A858>	35. QA ༜ <A862>
8. NYA ༖ <A847>	22. ZA ༞ <A859>	36. XA ༞ <A863>
9. TA ༘ <A84C>	23. -A ༠ <A85A>	37. FA ༠ <A864>
10. THA ༢ <A84D>	24. YA ༣ <A85B>	38. GGA ༣ <A865>
11. DA ༤ <A84E>	25. RA ༥ <A85C>	39. EE ༦ <A86E>
12. NA ༨ <A84F>	26. LA ༨ <A85D>	40. Subjoined WA ༨ <A866>
13. PA ༰ <A850>	27. SHA ༰ <A85E>	41. Subjoined YA ༰ <A867>
14. PHA ༲ <A851>	28. SA ༲ <A85F>	

Letters 1-30 are consonants derived from the thirty basic Tibetan consonants (and ordered according to traditional Tibetan consonant order); letters 31-34 are vowels derived from the four primary Tibetan vowels (and ordered according to traditional Tibetan vowel order); letters 35-38 are additional consonants that are used to represent sounds found in Mongolian, Chinese and other languages; letter 39 is an additional vowel used for writing Mongolian, Chinese and Uigur; and letters 40-41 are subjoined letters used for writing Tibetan, Sanskrit and Chinese.

The order of letters in my proposal is based on the order given in these two early sources, with the following modifications :

- The series of reversed letters used to represent the Sanskrit retroflex letters **ṭa**, **ṭha**, **ḍa** and **ṇa** (letters TTA, TTTHA, DDA and NNA [A848..A84B]) has been interpolated into the sequence between the letters NYA and TA. This corresponds to the ordering of letters in the Unicode Tibetan block, where the Tibetan letters TTA, TTTHA, DDA, DDHA and NNA [0F4A..0F4E] are placed between the Tibetan letters NYA and TA.
- The four additional non-Tibetan consonants QA, XA, FA and GGA have been moved up to the end of the thirty-four Tibetan-derived consonants, so that all the consonants are grouped together.
- The subjoined letters WA and YA have been moved between the base consonants and the vowels, and the additional subjoined letter RA and superfixed letter RA have been added after subjoined letter YA.
- The additional non-Tibetan vowel EE has been moved up to the end of the four Tibetan-derived vowels, so that all the vowels are grouped together.

I believe that grouping all the base consonants together, followed by all the subjoined/superfixed consonants, followed by all the vowels is sensible.

---

## Point 4

*Typeface in this PDAM had never been used in history of Phags-pa script.*

The typeface used is of necessity an artificial typeface, but it does try to follow the style of letters used in printed Yuan dynasty Chinese Phags-pa texts, specifically the 1340 edition of Baijiaxing Mengguwen 百家姓蒙古文 "The Hundred Chinese Family Names in Mongolian Script".

---







## Point 5



*Some characters' shapes are not the best choice for cononical characters, shapes selected should be those appear frequently. For instance: A851 and A86C.*

Unfortunately it is difficult to determine the canonical shapes of many Phags-pa characters as they often vary in shape from one text or inscription to another, and in the earliest surviving descriptions of the Phags-pa script (Menggu Ziyun, Fashu Kao and Shushi Huiyao) many of the glyph shapes are clearly corrupt. As discussed below, the letters PHA [A851] and E [A86C] are particularly problematic.

## Letter PHA









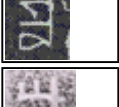
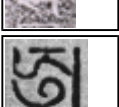
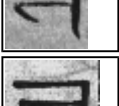
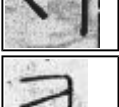
The letter PHA is not used for writing Mongolian, and although it is used for writing Chinese, it occurs in the spelling of few common ideographs. Moreover, even in texts that do use the letter PHA, it is often virtually indistinguishable in shape from the letter BA [A852]. Therefore there are not many examples of the letter PHA suitable to use as a model. Some sample images of the letter PHA from various sources are given below :

No.	Image	Reading	Source
1.		sph-a	Sanskrit Phags-pa inscription on the East Wall of Juyong Guan (1345).
2.		phu	Baijiaxing Mengguwen 百家姓蒙古文 "The Hundred Chinese Family Names in Mongolian Script" in the 1340 edition of <u>Shilin Guangji</u> 事林广记
3.		pha	Shushi Huiyao 书史会要 (1376).
4.		phué	<u>Menggu Ziyun</u> 蒙古字韵 (manuscript copy of edition published in 1308).
5.		phué	Inscription at Shandong of an edict conferring posthumous titles on Confucius dated 1307.
6.		phué	Inscription at Hebei of an edict conferring posthumous titles on Confucius dated 1307.




Having reviewed these and other sources, I accept that the shape of the glyph for A851 in my font () does not closely match the letter shapes shown above, and I therefore suggest that the representative glyph for PHAGS-PA LETTER PHA be changed to  (cf. No.6 above).

## Letter E





The letter E occurs in a wide range of shapes, as shown below :

No.	Image	Reading	Source
1.		èng	Mongolian Phags-pa inscription on the East Wall of Juyong Guan (1345).
2.		tè	Sanskrit Phags-pa inscription on the West Wall of Juyong Guan (1345).
3.		è	Stone inscription of the edict of Buyantu Khan dated 1314.
4.		jué	Inscription at Hebei of an edict conferring posthumous titles on Confucius dated 1307.
5.		è	A manuscript edict dated the Year of the Dragon (1328 ?) in the Kashag Historical Archives of Tibet.
6.		'uè	Chinese Phags-pa inscription dated 1298.
7.		jué	Inscription at Shandong of an edict conferring posthumous titles on Confucius dated 1307.
8.		nèn	Chinese Phags-pa inscription dated 1363.
9.		yè	A 'Nestorian' Christian tombstone in Quanzhou (Marco Polo's Zayton).
10.		'uè	Baijiaxing Mengguwen 百家姓蒙古文 "The Hundred Chinese Family Names in Mongolian Script" in the 1340 edition of Shilin Guangji 事林广记
11.		è	Shushi Huiyao 书史会要 (1376).
12.		è	Menggu Ziyun 蒙古字韵 (manuscript copy of edition published in 1308).





Based on an analysis of twenty-three Mongolian Phags-pa inscriptions and texts Professor Junast identifies three main forms of the letter E :

-  double-toothed form (Nos.1-4 above)
-  Short-hooked form (Nos.5-7 above)
-  Long-hooked form (Nos.8-11 above)


In his analysis of the origins of the form of the letter E, Professor Junast states :


If we wanted to select the most correct letter shape [out of the three main forms of the letter E] then without a doubt it would be  or . As to  and , they are probably simplifications of the former letter shape.

Junast, "Youguan Basiba zimu è de jige wenti". In Minzu Yuwen 1988.1

如果就中再找一个最规范的形式，那无疑是  或  了。至于  和 ，大概是它的简化形式。

照册斯图，《有关八思巴字母è的几个问题》。载《民族语文》1988.1。

Whilst the double-toothed form of the letter E is not the most common form of the letter, it is according to Professor Junast the most correct form of the letter. In selecting the shape of the letter E to use in the font I was heavily influenced by Professor Junast's important monograph on the Phags-pa letter E, which is the main reason why the glyph for letter E in my font () is based on the double-toothed form of the letter occurring in the Juyong Guan inscriptions.


Nevertheless, I accept that it may be more appropriate to use one of the more common letter forms; and as a long-hooked form would be more in keeping with the style of the font face, I suggest that the representative glyph for PHAGS-PA LETTER E be changed to  (cf. Nos.8-11 above).

## Point 6


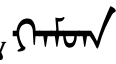
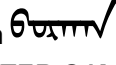




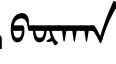




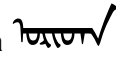



*Character names should conform to that of Mongolian (1800-18AF). For instance: A86F ANUSVARA, A870 BIRGA, etc..*

As the Phags-pa script was derived from the Tibetan script, there is an obvious one-to-one correspondence between most Phags-pa letters and the letters of the Tibetan script, which is why the proposed Phags-pa character names largely follow the corresponding Tibetan names.

On the other hand, there is no similar one-to-one correspondence between the letters of the Phags-pa script and the Mongolian script. Indeed, the Phags-pa script bears very little relationship with the Uighur-derived Mongolian script other than its writing orientation (vertical, left-to-right). Not only do Phags-pa letters such as A847 NYA, A851 PHA, A856 DZA and A858 ZHA not correspond to any letter in the Mongolian script, but due to the fact that the Mongolian and Phags-pa scripts employ different orthographic rules for writing Mongolian there is a many-to-one or many-to-many relationship between some letters of the two scripts. For example :

- 182C [MONGOLIAN LETTER QA] (transliterated as **k** in "feminine" words and **q** in "masculine" words) corresponds to either A840 [PHAGS-PA LETTER KA] or A841 [PHAGS-PA LETTER KHA] or A862 [PHAGS-PA LETTER QA] or A863 [PHAGS-PA LETTER XA], depending on the individual word :
  - Mongolian **yeke**  "large, great" = Phags-pa **yè ke** <A85B A86C 0020 A840 A86E>



- Mongolian **ken**  "who" = Phags-pa **khen** <A841 A86E A84F>
- Mongolian **qamuy**  "all" = Phags-pa **qa muq** <A862 0020 A853 A86B A862>
- Mongolian **burqan**  "Buddha" = Phags-pa **pur xan** <A850 A86B A85C 0020 A863 A84F>
- 182D [MONGOLIAN LETTER GA] (transliterated as **g** in "feminine" words and **y** in "masculine" words) corresponds to either A840 [PHAGS-PA LETTER KA] or A842 [PHAGS-PA LETTER GA] or A85A [PHAGS-PA LETTER -A] or A862 [PHAGS-PA LETTER QA], depending on the individual word :
  - Mongolian **galab**  "kalpa, aeon" = Phags-pa **kal bu dun** (gen. pl.) <A840 A85D 0020 A852 A86B 0020 A84E A86B A84F>
  - Mongolian **bičig**  "written document, book" = Phags-pa **bi chig** <A852 A86A 0020 A845 A86A A842>
  - Mongolian **qayan**  "khan, emperor" = Phags-pa **qa -an** <A862 0020 A85A A84F> or **q-an** <A862 A85A A84F>
  - Mongolian **čayan**  "white" = Phags-pa **cha q-an** <A845 0020 A862 A85A A84F>
- 182A [MONGOLIAN LETTER BA] corresponds to either A850 [PHAGS-PA LETTER PA] or A852 [PHAGS-PA LETTER BA], depending on the individual word :
  - Mongolian **burqan**  "Buddha" = Phags-pa **pur xan** <A850 A86B A85C 0020 A863 A84F>
  - Mongolian **basa**  "then, still, also" = Phags-pa **ba sa** <A852 0020 A85F>
- 1830 [MONGOLIAN LETTER SA] corresponds to either A85F [PHAGS-PA LETTER SA] or A859 [PHAGS-PA LETTER ZA], depending on the individual word :
  - Mongolian **ečüs**  "end, goal" = Phags-pa **he chus** <A860 A86E 0020 A845 A86B A85F>
  - Mongolian **sara**  "month" = Phags-pa **za ra** <A859 0020 A85C>
- 1833 [MONGOLIAN LETTER DA] corresponds to either A84C [PHAGS-PA LETTER TA] or A84E [PHAGS-PA LETTER DA], depending on the individual word :
  - Mongolian **dalai**  "sea, ocean" = Phags-pa **ta layi** <A84C 0020 A85D A85B A86A>
  - Mongolian **urida**  "former, previous" = Phags-pa **u ri da** <A86B 0020 A85C A86A A84E>
- A84E [PHAGS-PA LETTER DA] corresponds to either 1832 [MONGOLIAN LETTER TA] or 1833 [MONGOLIAN LETTER DA], depending on the individual word :
  - Phags-pa **deng ri** <A84E A86C A843 0020 A85C A86A> "heaven" = Mongolian **tengri** 
  - Phags-pa **u ri da** <A86B 0020 A85C A86A A84E> "former, previous" = Mongolian **urida** 
- A862 [PHAGS-PA LETTER QA] corresponds to either 182C [MONGOLIAN LETTER QA] or 182D [MONGOLIAN LETTER GA], depending on the individual word :
  - Phags-pa **qa muq** <A862 0020 A853 A86B A862> "all" = Mongolian **qamuy** 

From these examples it can be seen that attempting to name the Phags-pa letters after Mongolian letters would be highly problematic.

## A86F

A86F corresponds in usage to both the candrabindu and anusvara signs in the Tibetan and Devanagari scripts, but its proposed name is CANDRABINDU because it corresponds in shape to the candrabindu sign found in Sanskrit, Tibetan and other Brahmic scripts (cf. 0901, 0981, 0A81, 0B01, 0C01, 0F83). Whilst U+1880 is named MONGOLIAN LETTER ALI GALI ANUSVARA ONE, I note that not only is the letter shaped like a candrabindu sign rather than an anusvara sign, but that in the Unicode code charts it is actually cross-referenced to U+0F83, which is the Tibetan form of the candrabindu.

Note also that where A86F is used in the Phags-pa version of the Buddhist texts inscribed at Juyong Guan, the corresponding Tibetan version of the text uses a candrabindu sign rather than the more normal anusvara :



The invocation **om bhrūm bhrūm bhrūm** in the Phags-pa version (left) and Tibetan version (right) of the Buddhist text on the East Wall at Juyong Guan. Note how A86F in the Phags-pa text corresponds to U+0F83 [TIBETAN SIGN SNA LDAN = CANDRABINDU], not to U+0F7E [TIBETAN SIGN RJES SU NGA RO = ANUSVARA], in the Tibetan version of the text.

Instead of changing the name of A86F, I suggest that an additional cross-reference to U+1880 [MONGOLIAN LETTER ALI GALI ANUSVARA ONE] be added to this character, as indicated in the Appendix at the bottom of this document.

## A870

As far as I am aware, A870 does not occur in any early (13th and 14th century) Phags-pa texts or inscriptions, but is only found in more recent Tibetan Phags-pa texts (see Examples 4 and 5 in N2622). Thus I do not believe that it would be appropriate to give the character a Mongolian name (BIRGA).

Instead of changing the names of A870 and A871, I suggest that an additional cross-reference to U+1800 [MONGOLIAN BIRGA] be added to these characters, as indicated in the Appendix at the bottom of this document.

## Point 7

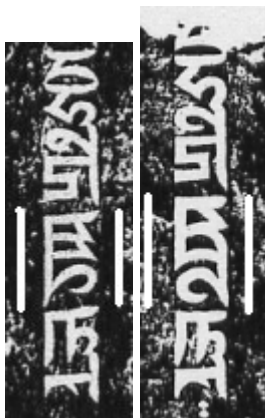
*At least one “Free Variation Form Selector” is needed. A “Syllable Connector” different to “general space” is also needed.*



### Free Variation Selector

I agree that a Phags-pa Free Variation Selector (FVS) may be needed.

As discussed in Section 8 of N2622, A860 [PHAGS-PA LETTER HA], A867 [PHAGS-PA SUBJOINED LETTER YA], A86A [PHAGS-PA LETTER I], A86B [PHAGS-PA LETTER U] and A86C [PHAGS-PA LETTER E] occur in reversed forms after A848..A84B [PHAGS-PA LETTER TTA, TTHA, DDA and NNA]. As this glyph reversal is contextual there is no need to encode separate characters for these reversed letter forms. However, a Phags-pa FVS would be very useful, in order to show a reversed letter HA, Subjoined YA, I, U or E in isolation (e.g. for metalanguage descriptions of the Phags-pa script).

Furthermore, the letter I occurs in both reversed and unreversed forms after the letter TTHA (see examples below), and so it would be useful to have a mechanism for inhibiting contextual letter reversal where normally expected.














These two images from the Juyong Guan inscriptions show the word **dhish tthi te** (Sanskrit **dhīṣṭhite**) written with a reversed letter I  after the letter TTHA [left] and an unreversed letter I  after the letter TTHA [right].

A Phags-pa Free Variation Selector could operate such that when applied to one of the letters HA, Subjoined YA, I, U or E it :

- selects the reversed glyph form in situations where the unreversed glyph form would normally be selected (e.g. for the letter in isolation);
- selects the unreversed glyph form in situations where the reversed glyph form would normally be selected (i.e. after one of the letters TTA, TTHA, DDA or NNA).

Thus the presence or absence of a Phags-pa Free Variation Selector could be used to produce reversed or unreversed glyph forms as shown below :

Character Sequence	Glyph	Description
HA, FVS		PHAGS-PA LETTER HA reversed form
SUBJOINED-YA, FVS		PHAGS-PA SUBJOINED LETTER YA reversed form
ZWJ, I, FVS		PHAGS-PA LETTER I reversed final form
ZWJ, U, FVS		PHAGS-PA LETTER U reversed final form
ZWJ, E, FVS		PHAGS-PA LETTER E reversed final form
DDA, HA		Syllable DDHA with reversed Letter HA (automatic contextual reversal of Letter HA after Letter DDA)
NNA, SUBJOINED-YA		Syllable NNYA with reversed Letter Subjoined YA (automatic contextual reversal of Subjoined Letter YA after Letter NNA)
DDA, E		Syllable DDE with reversed Letter E (automatic contextual reversal of Letter E after Letter DDA)
NNA, U		Syllable NNU with reversed Letter U (automatic contextual reversal of Letter U after Letter NNA)
TTHA, I		Syllable TTHI with reversed Letter I (automatic contextual reversal of Letter I after Letter TTHA)
TTHA, I, FVS		Syllable TTHI with unreversed Letter I (FVS selects unreversed form of Letter I after Letter TTHA)

If a Phags-pa Free Variation Selector were to be considered appropriate, the following additional character could be added to the PHAGS-PA block :

Code Point	Glyph	Character Name	Unicode Properties	Notes
A874		PHAGS-PA FREE VARIATION SELECTOR	Mn;0;NSM;;;;;N;;;;;	

## Syllable Connector

I do not believe that a separate "Syllable Connector" character is needed. The whitespace between the component syllables of a Phags-pa word can be represented by NO-BREAK SPACE [U+00A0] or NARROW NO-BREAK SPACE [U+202F], or indeed simply by SPACE [U+0020] as lines do break on intra-word syllable divisions in traditional texts and inscriptions.

## Point 8

*Phags-pa script should not be regarded as variant of Tibetan script.*

My proposal in no way considers the Phags-pa script to be a variant of the Tibetan script, but it is impossible to ignore its Tibetan heritage :

- The script was created by a Tibetan monk, who based all the Phags-pa letter shapes on Tibetan letters.
- In the earliest Chinese descriptions of the script ([Fashu Kao](#) and [Shushi Huiyao](#)) the letters are ordered according to the traditional Tibetan letter order.

Nevertheless, I appreciate that the cross-references to the corresponding Tibetan character for each Phags-pa character in Table 1 of N2622 might give the mistaken impression that the Phags-pa script is considered to be a variant of the Tibetan script. These cross-references are only intended to indicate what Tibetan character a given Phags-pa character derives from, and are not intended to imply that the particular Phags-pa letter is only or mainly used for writing Tibetan.<sup>1</sup> As there is a many-to-one or many-to-many correspondence between many of the letters of the Phags-pa and Mongolian scripts (see Point 6 above), it is not practical to annotate the relationship between Phags-pa letters and Mongolian letters in a similar way. However, it may be useful to annotate the linguistic usage of each letter, to show which out of the Mongolian, Chinese, Uighur, Tibetan and Sanskrit languages it is used for writing.<sup>2</sup> This would more clearly indicate that the Phags-pa script is not simply a variant of the Tibetan script, but was an international script designed to be used for writing Mongolian and other languages. My proposed annotations are given in the Appendix at the bottom of this document.

1. "Cross references (preceded by -->) are used to indicate a related character of interest, but without indicating the nature of the relation" ([The Unicode Standard](#) 4.0 Section 16.1).

2. "An informative note may include a list of the language(s) using that character where this information is considered useful" ([The Unicode Standard](#) 4.0 Section 16.1).

---

## Appendix

The following table is a revision of Table 1 in Document N2622, in response to comments by the Chinese National Body. The following changes, shown in red text, have been incorporated :

- Change to the reference glyphs for A851 and A86C.
- Addition of general notes indicating language usage of Phags-pa characters.
- Addition of cross-references to Mongolian characters where appropriate.

Note that there are no changes to character repertoire, character names or character properties.









Revised Table of Proposed PHAGS-PA Characters

Code Point	Glyph	Character Name	Unicode Properties	Notes
A840	ཀ	PHAGS-PA LETTER KA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Tibetan, Sanskrit --> 0F40 ཀ tibetan letter ka
A841	ཁ	PHAGS-PA LETTER KHA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan, Sanskrit --> 0F41 ཁ tibetan letter kha
A842	ག	PHAGS-PA LETTER GA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan, Sanskrit --> 0F42 ག tibetan letter ga
A843	ང	PHAGS-PA LETTER NGA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Tibetan, Sanskrit --> 0F44 ང tibetan letter nga
A844	ཅ	PHAGS-PA LETTER CA	Lo;0;L;;;;;N;;;;;	* Chinese, Tibetan --> 0F45 ཅ tibetan letter ca
A845	ཆ	PHAGS-PA LETTER CHA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan --> 0F46 ཆ tibetan letter cha
A846	ཇ	PHAGS-PA LETTER JA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan --> 0F47 ཇ tibetan letter ja
A847	ཉ	PHAGS-PA LETTER NYA	Lo;0;L;;;;;N;;;;;	* Chinese, Tibetan, Sanskrit --> 0F49 ཉ tibetan letter nya
A848	ཏ	PHAGS-PA LETTER TTA	Lo;0;L;;;;;N;;;;;	* Sanskrit --> 0F4A ཏ tibetan letter tta
A849	ཐ	PHAGS-PA LETTER TTHA	Lo;0;L;;;;;N;;;;;	* Sanskrit --> 0F4B ཐ tibetan letter ttha
A84A	ཌ	PHAGS-PA LETTER DDA	Lo;0;L;;;;;N;;;;;	* Sanskrit --> 0F4C ཌ tibetan letter dda
A84B	ཎ	PHAGS-PA LETTER NNA	Lo;0;L;;;;;N;;;;;	* Sanskrit --> 0F4E ཎ tibetan letter nna

A84C	ཏ	PHAGS-PA LETTER TA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Tibetan, Sanskrit --> 0F4F ཏ tibetan letter ta
A84D	ཐ	PHAGS-PA LETTER THA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan, Sanskrit --> 0F50 ཐ tibetan letter tha
A84E	ཌ	PHAGS-PA LETTER DA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan, Sanskrit --> 0F51 ཌ tibetan letter da
A84F	ཎ	PHAGS-PA LETTER NA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan, Sanskrit --> 0F53 ཎ tibetan letter na
A850	ཡ	PHAGS-PA LETTER PA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Tibetan, Sanskrit --> 0F54 ཡ tibetan letter pa
A851	མ	PHAGS-PA LETTER PHA	Lo;0;L;;;;;N;;;;;	* Chinese, Tibetan, Sanskrit --> 0F55 མ tibetan letter pha
A852	འ	PHAGS-PA LETTER BA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan, Sanskrit --> 0F56 འ tibetan letter ba
A853	མ	PHAGS-PA LETTER MA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan, Sanskrit --> 0F58 མ tibetan letter ma
A854	ཨ	PHAGS-PA LETTER TSA	Lo;0;L;;;;;N;;;;;	* Chinese, Tibetan, Sanskrit --> 0F59 ཨ tibetan letter tsa
A855	ཨ	PHAGS-PA LETTER TSHA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Tibetan, Sanskrit --> 0F5A ཨ tibetan letter tsha
A856	ཨ	PHAGS-PA LETTER DZA	Lo;0;L;;;;;N;;;;;	* Chinese, Tibetan, Sanskrit --> 0F5B ཨ tibetan letter dza
A857	ཨ	PHAGS-PA LETTER WA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan, Sanskrit --> 0F5D ཨ tibetan letter wa
A858	ཨ	PHAGS-PA LETTER ZHA	Lo;0;L;;;;;N;;;;;	* Chinese, Tibetan --> 0F5E ཨ tibetan letter zha
A859	ཨ	PHAGS-PA LETTER ZA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Tibetan --> 0F5F ཨ tibetan letter za
A85A	ཨ	PHAGS-PA LETTER -A	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan, Sanskrit --> 0F60 ཨ tibetan letter -a
A85B	ཨ	PHAGS-PA LETTER YA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan, Sanskrit

				--> 0F61 ཡ tibetan letter ya
A85C	𑀢	PHAGS-PA LETTER RA	Lo;0;L;;;;;N;;;;;	* Mongolian, Uighur, Tibetan, Sanskrit --> 0F62 ར tibetan letter ra
A85D	𑀣	PHAGS-PA LETTER LA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan, Sanskrit --> 0F63 ལ tibetan letter la
A85E	𑀤	PHAGS-PA LETTER SHA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan, Sanskrit --> 0F64 ཤ tibetan letter sha
A85F	𑀥	PHAGS-PA LETTER SA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan, Sanskrit --> 0F66 ས tibetan letter sa
A860	𑀦	PHAGS-PA LETTER HA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan, Sanskrit --> 0F67 ཏ tibetan letter ha
A861	𑀧	PHAGS-PA LETTER A	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan, Sanskrit --> 0F68 ཏ tibetan letter a
A862	𑀨	PHAGS-PA LETTER QA	Lo;0;L;;;;;N;;;;;	* Mongolian, Uighur
A863	𑀩	PHAGS-PA LETTER XA	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese
A864	𑀪	PHAGS-PA LETTER FA	Lo;0;L;;;;;N;;;;;	* Chinese, Uighur
A865	𑀫	PHAGS-PA LETTER GGA	Lo;0;L;;;;;N;;;;;	
A866	𑀬	PHAGS-PA SUBJOINED LETTER WA	Lo;0;L;;;;;N;;;;;	* Chinese, Tibetan, Sanskrit --> 0FAD སྐ tibetan subjoined letter wa
A867	𑀭	PHAGS-PA SUBJOINED LETTER YA	Lo;0;L;;;;;N;;;;;	* Chinese, Tibetan, Sanskrit --> 0FB1 སྑ tibetan subjoined letter ya
A868	𑀮	PHAGS-PA SUBJOINED LETTER RA	Lo;0;L;;;;;N;;;;;	* Tibetan, Sanskrit --> 0FB2 སྒ tibetan subjoined letter ra
A869	𑀯	PHAGS-PA SUPERFIXED LETTER RA	Lo;0;L;;;;;N;;;;;	* Tibetan --> 0F62 ར tibetan letter ra
A86A	𑀰	PHAGS-PA LETTER I	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan, Sanskrit --> 0F72 ཨ tibetan vowel sign i
A86B	𑀱	PHAGS-PA LETTER U	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan, Sanskrit --> 0F74 ཨ tibetan vowel sign u
				* Mongolian, Chinese, Uighur, Tibetan,



A86C		PHAGS-PA LETTER E	Lo;0;L;;;;;N;;;;;	Sanskrit --> 0F7A ཀྲི tibetan vowel sign e
A86D		PHAGS-PA LETTER O	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur, Tibetan, Sanskrit --> 0F7C ཀྲོ tibetan vowel sign o
A86E		PHAGS-PA LETTER EE	Lo;0;L;;;;;N;;;;;	* Mongolian, Chinese, Uighur
A86F		PHAGS-PA LETTER CANDRABINDU	Lo;0;L;;;;;N;;;;;	* Sanskrit --> 0F83 ཀྲིྱ tibetan sign sna ldan --> 0F7E ཀྲིྱ tibetan sign rjes su nga ro --> 1880 ཀྲིྱ mongolian letter ali gali anusvara one
A870		PHAGS-PA SINGLE HEAD MARK	Po;0;ON;;;;;N;;;;;	* Tibetan * marks beginning of text --> 0F04 ཀྲིྱ tibetan mark initial yig mgo mdun ma --> 1800 ཀྲིྱ mongolian birga
A871		PHAGS-PA DOUBLE HEAD MARK	Po;0;ON;;;;;N;;;;;	* Tibetan * marks beginning of text --> 0F04 ཀྲིྱ tibetan mark initial yig mgo mdun ma --> 0F05 ཀྲིྱ tibetan mark closing yig mgo sgab ma --> 1800 ཀྲིྱ mongolian birga
A872		PHAGS-PA MARK SHAD	Po;0;ON;;;;;N;;;;;	* Tibetan --> 0F0D ཀྲིྱ tibetan mark shad
A873		PHAGS-PA MARK DOUBLE SHAD	Po;0;ON;;;;;N;;;;;	* Tibetan --> 0F0E ཀྲིྱ tibetan mark nyis shad