Some Problems on the Encoding of Phags-pa Script

The 45th WG2 meeting held in June, 2004, adopted N2829 in which is mentioned, "The ad hoc discussion on encoding Phags-pa was unable to resolve all of the ballot comments on Phags-pa. There were many complex and difficult questions for which much further communication and discussion will be required. "

"With reference to changes requested to Phags-Pa encoding in PDAM 1 ballot comments, and the details given in document N2745 and associated documents N2719 and N2771, WG2 accepts the ad hoc group recommendations in document N2829. WG2 accepts the revised encoding of the 52 characters as shown on pages 79 to 81 in document N2832 for inclusion in Amendment 2 to ISO/IEC 10646: 2003 (removing it from Amendment 1). National bodies and liaison organizations are invited to provide feedback on the various open issues identified in document N2829."(Resolution M45.16). We think that such a resolution is extremely timely and extremely correct.

With a view of the problems summarized in N2829 and other relevant problems, experts from China and Mongolia have their second meeting on Phags-pa script encoding in Changsha Oct 24-25,2004, at which we carefully examine various opinions and now put forward, in the name of our two nations, our official and revised Proposal for Phags-pa Script Encoding and A Users' Agreement Related to Phags-pa Script Encoding.

The following are our opinions on some problems on Phags-pa script encoding that require further exchange of views and deepened discussion.

Part One   A Brief Account of the Phags-pa Script and Its Encoding

(1) Peculiarities and complexities of Phags-pa script: The Phags-pa script is a phonetic writing system created by the imperial teacher Phags-pa by a special edict of Emperor Khubilaii, Shizhu of the Yuan Dynasty and promulgated in 1269 for the purpose of "translating & writing" multi-lingual texts across the Yuan Empire. Materials discovered by now show that texts in Mongolian, Han.(Chinese), Tibetan, Uighur and some other foreign languages including Sanskrit have been "translated/written" in Phags-pa alphabet, of which Mongolian and Chinese texts are naturally the most frequently "translated/written" inasmuch as Mongolian was the "state language" of Yuan Dynasty and Chinese is used by the majority of ethnic groups in China.

As a special writing system for presenting multi-lingual texts, Phags-pa script has its outstanding features. Its alphabet has rich content, boasting more letters than enough for "translating/writing" any particular language, and that it has "compound letters", too. The Phags-pa script is written from left to right and from above downward. It differs from Mongolian in that it takes the syllable as a unit of ligature. While Tibetan uses the same punctuation mark to indicate the limits of a syllable and a word, Phags-pa alphabet needs no special punctuation marks for showing the limits of syllables or words.

In writing texts in different languages, Phags-pa script shows different characteristics not only in the number of letters used, but also in the way each language is pronounced or spelt. Thus, Mongolian and Han (Chinese) texts are transcribed according to their sounds, whereas Tibetan and Sanskrit are in principle transliterated word for word, though certain words are presented in

accordance with their sounds. Consequently, such inconsistency ought to be contained and reflected in the present encoding. For example, such letters as ꡝ, ꡖ and ꡢ never appear in Mongolian documents, but do appear in Chinese texts. In Mongolian and Sanskrit texts, ꡚ and ꡡ, ꡗ and ꡁ, ꡜ and ꡣ do not distinguish from each other. But in Chinese texts, these six letters are all letters having their distinguishing functions, viz., ꡚ = 禅, ꡡ = 审, ꡗ = 喻, ꡁ = 影, ꡜ = 晓, ꡣ = 匣. Letters ꡘ, ꡙ, ꡠ and ꡄ are all separate letters with distinctive functions in Sanskrit, but do not appear in Mongolian texts. So, in order to perfectly reflect such complicated details, we should treat these letters as "nominal glyphs" in our Phags-pa script encoding.

(2) Alphabet of the Phags-pa script: The encoding of Phags-pa script should reflect its letter system fully and correctly. We must know that the Phags-pa script "translates-writes" texts in different languages in not quite the same way. Their relationship is very complicated, far from being as simple as with a single language. We maintain that the letter system of the "nominal glyph" of Phags-pa script is a system characterized by pronunciation as its content and graphic figures as distinguishing features, whereas its letter system of "variant presentation glyphs" is a system characterized by mere sounds regardless of its external forms. We take ꡚ and ꡡ as two separate nominal glyphs mainly because they represent different sounds (though we notice the insignificant difference in their shapes). But we treat ꡄ and ꡅ as two "free variants" of one and the same letter mainly in accordance with their "sounds" (their difference in letter forms being a little greater than that between ꡚ and ꡡ). The same is true with such different letters as ꡊ, ꡋ, ꡌ, ꡍ, ꡎ, ꡏ and ꡐ which we regard as variants of the single letter ꡊ, especially according to their sounds (It should be admitted that their difference in letter forms is still greater than that between the above-mentioned two sets of glyphs).

(3) Uses of the Phags-pa script: Since texts in Phags-pa script are all historical documents, their spellings are not quite consistent or standardized. Nor is Phags-pa script a current writing system, so we do not have any realistic requirement or possibility to standardize it. We believe that the main purpose of the Phags-pa script encoding is to intactly preserve and represent those historical documents and materials for the convenience of research. We should by no means delete certain spellings or graphic symbols that do exist in historical documents, but instead, include as much as possible those diversified spellings and graphic symbols in our encoding. There are scholars who refuse such variants as ꡊ, ꡋ and ꡏ, all of which however are found in Mongolian, Tibetan and Sanskrit documents. Thus, there may appear several ways to spell a word, of which one spelling is correct, the others are not. But the letter spellings must be presented in the

encoding in order to represent what historically is true: ⟨glyph⟩ alongside of ⟨glyph⟩ and ⟨glyph⟩ alongside of ⟨glyph⟩. ⟨glyph⟩ appears 123 times and ⟨glyph⟩ 33 times in Mongolian documents. We mustn't change all their written forms into one single ⟨glyph⟩, must we ? if so, it would violate historical truth!

One opinion goes like this, "It is possible to easily translate a Tibetan document in Phags-pa alphabet back into Tibetan", which is actually an unthinkable "objective" excluded from the UCS. First of all, one should not "demand" that any encoding fulfill such a function, for it is an extravagant and unpractical hope. We clearly know that between Mongolian encoding and Mongolian texts in Phags-pa alphabet there does not exist any possibility of automatically transferring to each other. Moreover, what we have to emphasize is that there does not exist such "necessity"! So far as we know, in Tibetan documents written in Phags-pa letters there are a lot of words that do not meet the standards of Tibetan orthography. Thus, for the Tibetan person name "Rin chen", there can be five Phags-pa spellings: ⟨glyph⟩, ⟨glyph⟩, ⟨glyph⟩, ⟨glyph⟩ and ⟨glyph⟩ ; and for the Tibetan word "skor gsum", the Phags-pa script has two spellings: ⟨glyph⟩ and ⟨glyph⟩. How can we translate them into Tibetan letter for letter?

Instead, one should rather take into consideration the problem of how to link it up with the Latin transliteration in the Phags-pa script. We have to point out, then, that names in UCS are not the same thing as Latin transliteration of the text.

(4) Influence of Mongolian writing system on Phags-pa script: Some people think that the only possible relation between Phags-pa script and Mongolian writing is that they are both written in the same direction. However, the research of Chinese and Mongolian scholars shows that there exist many influences of Mongolian writing system on Phags-pa script. For example, the structure of the Phags-pa letter ⟨glyph⟩ OE and ⟨glyph⟩ UE are based on the Mongolian ⟨glyph⟩ OE and ⟨glyph⟩ UE; the separate, word-initial, word-medial and word-final forms of vowels in Phags-pa script are also designed in accordance with the vowel harmony in Mongolian; and the Phags-pa variant forms ⟨glyph⟩, ⟨glyph⟩ of the Mongolian vowels ⟨glyph⟩ OE and ⟨glyph⟩ UE in other than the first syllable are based on Mongolian vowel harmony, too. Besides, the Phags-pa letter ⟨glyph⟩ A856 is also modeled on Mongolian writing. Therefore, it is only natural that Phags-pa script presentation of Mongolian glyphs became the same as to Mongolian glyphs. Such similar features of Phags-pa script and Mongolian writing are not depend on our attitude to Mongolian writing system, nor our "fabrication", but it was the reflection of the Mongolian writing system to the Phags-pa script, which was created by the Great lama Phags-pa.

(5) Research on Phags-pa script: Before the 1980's, research on Phags-pa script had been conducted mainly outside China and Mongolia, and few materials in this script were discovered. Since the 1980's, however, we witnessed an unprecedented upsurge in the research on Phags-pa script in the Phags-pa's homeland China and Mongolia. According to incomplete statistics, over 110 treatises and monographs on Phags-pa script were published in China and Mongolia during this period, there are occurred less than 20 works in other countries. It was discovered more than 50 monuments in Phags-pa script during the past 20 years successively. Scholars in China and

Mongolia deepened their research on the Phags-pa script by relying on such rich materials and published many valuable monographs and treatises in which they put forward a series of new viewpoints. Chinese and Mongolian scholars have such a rich materials in their hands and their up-dated research achievements have provided us with an important scientific basis for the preparation of the present encoding of Phags-pa script.

However the studies on Phags-pa script are could not covered all the monuments in different languages, especially Chinese texts in Phags-pa script are requiring serious further investigations. Although scholars have reached common understanding on many problems concerning Phags-pa script, there nevertheless exist great divergences in quite a number of problems, thus, opinions do vary in the understanding of the letters �, ꧁and ꧂, and such divergences are very difficult to eliminate within a short time. The spellings of documents and sources in Phags-pa alphabet are not quite consistent or standardized, and what is more, since Phags-pa script is not a current official writing system, there is no need or possibility to "standardize" it. In view of this situation, we must not rigidly advocate one point of view and restrict another in preparing the Phags-pa script encoding. Instead, we have to adopt a tolerant attitude, treating all views equally without discrimination in problems in which there exist serious differences. In other words, it is undesirable to create a situation in which a certain point of view is restricted merely because of our way of encoding. Thus, there exist more than three explanations of the quality of the letter ꧃, in encoding we should tolerate all of them without attempting to clarify its quality. Different schools may explain and use ꧄in different ways. And again, different views on the letters ꧅ and ꧆ should also be tolerated without imposing any restriction on them. In a word, our encoding will permit to register the two letters ꧅ and ꧆ in three ways.

As soon as certain academic problems are solved, we may, then and not until then, revise, replenish and improve the Phags-pa encoding.

(6) Experience to be summed up and lessons to be drawn in the Phags-pa script encoding: During the late 1980's and early 1990's, Mongolian IT engineers in Mongolia, China and Germany prepared their respective Mongolian editing devices[1]. In China, during the practice of more than 10 years, rich experience has been accumulated with many lessons to be drawn in the preparation of Phags-pa script encoding as well as Phags-pa script information treatment. This is also very precious for us in developing the encoding. It is based on the experience and lessons in information treatment in the past 15 years that we put forward the proposal to include the treatment of ꧅ and ꧆ as "whole characters", and also to provide the "syllable delimiter" in the present encoding. Inasmuch as we prepare Phags-pa script encoding for the purpose of serving Phags-pa information treatment, it is only too natural that we consider and handle certain problems in terms of information treatment. To devise the "syllable delimiter" seems to go against the "status quo" of Phags-pa script, but actually to do so is absolutely advantageous with no harm to information treatment in the future.

[1] See "v Mong"(1978) in Germany, "Multi-lingual editing device"(1989) by Inner Mongolia University, "Founder BookMaker 9.1 for Mongolian"(1990) in China and "Light printing"(1990) in Mongolia.

(7) Two different ways to register letters: Quite a number of Phags-pa letters have their isolate, syllable-initial, syllable-medial and syllable-final forms. The isolate form of a vowel and the syllable-initial form of a consonant may be regarded as "nominal glyphs", yet the majority of syllable-initial forms and all syllable-medial and syllable-final forms cannot be "nominal glyphs". Under such conditions, there naturally appear two different ways to register letters: (a) Registration within a syllable or a word. Since it relies on its preceding and following glyphs so in most cases there is no need to use any control symbol, thus, the two variant presentation glyphs ᠬ and ᠪ in the words ᠬᠣᠯ ᠬᠣᠬ and ᠌ ᠬᠣ ᠬᠣ are registered without using the control symbol and (b) Registration of a single variant presentation glyph, which, without reliance on preceding and following glyphs, should use the control symbol, thus, ᠬ and ᠪ not within a word should be registered as ᠌ ᠬ ᠌ and ᠌ ᠣ ᠌. The UCS does not regulate any rules for the use of the control symbol. For unmistakable exchange between the users, however, it is necessary to specify certain rules for the users to follow. Such rules should include the two different ways to register letters, for its details please see the Reference Table in the Users' Agreement Related to Phags-pa Script Encoding.
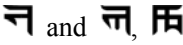
(8) The Users' Agreement Related to Encoding of Phags-pa Script: One opinion is that people should be able to register every letter with the only help of the Table of Nominal Glyphs of the UCS, there being no need to read the regulations in the Users' Agreement for the Phags-pa script encoding. In our eyes, this is a mere illusory wish, for according to our experience all encodings prepared in terms of nominal glyphs have variant presentation glyphs and a certain number of various control symbols not included in the UCS. In order that no different understanding might happen in the users' information exchange, there must be an agreement or a few regulations to unify the number and forms of the variant presentation glyphs and the usage of various kinds of control symbols. According to the UCS, if such an agreement or regulations are not included in it, then the users of a given encoding should reach such agreement or regulations through consultations. No doubt, the same is true with Phags-pa script encoding, otherwise it won't do to have only the Table of Glyphs for Phags-pa in the UCS. The author of 2719 seems to be persisting in such a viewpoint as "there being no need for the Users' Agreement", disdaining to take a glance at the Users' Agreement Related to the Encoding of Phags-pa Script. That's why out of the 17 example words he listed in Article 8 of 2719, 12 are mistaken.[2] If the author had read the Users' Agreement on Mongolian Encoding System, he would not have made such a hopeless mess in spelling the few most easy and simple Mongolian words. Perhaps from a negative side, his example confirms the very need for having a readers' agreement.

(9) Definition of nominal glyphs and variant presentation glyphs of the Phags-pa script: Both N2771 and we agree that Phags-pa script has nominal glyphs, but differ in the standards with which to define variant presentation glyphs. Since Phags-pa documents were written in historical periods with technical restrictions and lack of ascertained standards for its writing, it is a troublesome matter to try to define the variant presentation glyphs of certain letters of the script.

---

[2] He misspells ᠬ for ᠬ , ᠬᠮᠮ for ᠬᠮᠮ, ᠬᠣᠳ for ᠬᠣᠳ , ᠬᠮᠮ for ᠬᠮᠮ , ᠬᠣᠯ for ᠬᠣᠯ, ᠬᠣᠳ for ᠬᠣᠳ and ᠬᠣᠯ for ᠬᠣᠯ ! All this has resulted from his ignoring and not observing the regulations in the Users' Agreement on Mongolia Encoding System.

And it is evident that glyphs with slight differences between them cannot all be defined as variant presentation glyphs. In the view of the absence of ready standards for definition, we proposed in N2745 a set of standards for defining variant presentation glyphs (See [3] of N2745-1) to the effect that:

⒜ Strictly speaking, each letter in the Phags-pa script has several variant presentation glyphs. Out

of most consonant letters can be separated their respective four variants, viz., the isolate, syllable-(or word-)initial, syllable-(or word-)medial and syllable-(or word-)final forms. A few letters have less than four variants. Owing to different styles of writing, certain variants may take the same form.

⒝ According to the conditions under which each variant presentation glyph appears, the variant

presentation glyphs of Phags-pa script are divided into "conditional variants" and "free variants" (the latter include "positional variants" and "postpositive variants"). 卩 and 卩, 乕 and 乕 which differ evidently in strokes yet have the same pronunciation, are "free variants".

⒞ Certain glyphs are marked as having two sound forms in one language, but having only one

sound in another language. The former are actually two different letters while the latter two free variants of one and the same letter.

⒟ Variants, which differ slightly in size, thickness, length or turning angle, may be called "stroke

variants". Stroke variants which are not marked as having different pronunciations, or as having any variant forms of letter under certain conditions, are not indicated in the encoding. Of course, the above standards are open to discussion to see if they are feasible. However, we have to point out that for the definition of the variant presentation glyphs of letters in various languages, there should be a unified standard. A double standard is not to be used.

Part Two   Views on Some Concrete Problems

(1) **The vowel letters** 朙 **'OE' and** 朙 **'UE'** : In the Mongolian language, these two

vowels are indispensable basic vowels. Though a few variant presentation glyphs of these two letters are "compound letters" consisting of two or three lexemes, the majority of scholars list them in their Table of Letters, regarding them as compound yet independent letters, like N.Poppe

(1941,1957), B.Rinchen(1956), L.Ligeti(1964,1972), D.Čoijilsüreng(1974), Č.Šagdarsüreng

（1981,2001），Bulag（1983），Bao Xiang（1984），A.Damdinsüreng（1985），Tulgaguri（1998），Y.Jančib（2002）. Scholars who do not include these two letters in their Table of Letters, also point out that "in a few cases, a double letter shows one sound."[3]

The formation of these two letters 朙 and 朙 in Phags-pa script are modeled on the

---

[3] Junast, The Phags-pa Script and Mongolian Documents, Vol.I:A Collection of Research Essays, p.52.

Mongolian vowel letters ᠣ and ᠥ, too. Their pronunciations are not merely putting together the sounds of their original lexemes, like A+E+O and A+E+U, but are the single sounds ö and ü which have nothing to do with them. This point is in perfect conformity with the letter 豽

2864 in N2719. Because though the letter 豽 consists of two lexemes ᠬ and ᠦ, it is pronounced F, a single sound having nothing to do with H+U put together. It is quite right that N2719 treats as two nominal glyphs ᠬ+ᠦ, ᠬ+ᠦ glyphs ꡎꡜ ꡎꡜ which is very much like 豽 2864 in form, but puts together the original sounds of the lexemes (GA+U, CHA+U).

Therefore, as in the case of the letter 豽, we should treat them as two independent vowel letters instead of "compound letters". The Mongolian ᠣ ᠥ in UCS are treated in the same way.

The reason why we insist that there should be such compound characters as ꣓ and ꣖ in the encoding of Phags-pa script is that we were taught to do so by our own experience and lessons in the 1970's when we were casting type matrices for the Phags-pa script and later in the 1990's when we were developing computerized databank for it. For lack of experience in handling Phags-pa characters and words on the computer, we continued for too long a time to rely on the type mould technology for separating the letters ꣓ and ꣖ into ꡝ+ ꡊ+ꡜ and ꡝ+ ꡊ+ꡆ. Such practice of ours which was out of question in years of letterpress, caused us a lot of unsoluable difficulties in the handling of words and sentences in Phags-pa script. If in looking up words and characters and arranging the order of words we did not distinguish these two vowel letters that account for 14.8% of the total number of letters, we were unable to automatically tell apart ꡝ, ꡊ, ꡜ, ꡆ and ꣕, ꡅ, ꣓, ꣖, which brought us a lot of trouble in the follow-up information treatment in Phags-pa script. Thus, even if viewed from the angle of "automatic Latin transliteration for Phags-pa alphabet", these two vowel letters ought to be treated as separate letters. The Latin transliterations of ꣓ and ꣖ must never be "aeo", "'eo" or "aeu", "'eu". Such transliteration does not meet the requirement that "the transliterated text in Latin alphabet should be read in the pronunciation of the original text". The Latin transliteration of ꣓ and ꣖ (as well as their variants) can only be "ö" and "ü".

(2) **The letters ꡤ YA, ꡮ SHA, ꡜ HA and 豽 FA:** According to "Mongolian

Rhymes" in Phags-pa script, the eight glyphs ꡗ A857, ꡯ A86F, ꡚ A85A, ꡰ A870, ꡜ A85C, ꡱ A871, ꡲ A872 and ꡤ A864 mark respectively specific consonants of syllables of 36 characters in 1224 characters/words (ꡚ = 禅，ꡚ = 审 ，ꡗ = 喻 ，ꡯ = 影 ，ꡜ = 晓 ，ꡱ = 匣，ꡲ = 非，ꡤ = 敷), that's why we ought to prepare nominal glyphs for them. There are some scholars, however, who admit that these letters refer to specific initial consonants of syllables of the 36 characters on certain occasions, but insist they play roles in only one document, so instead of nominal glyphs, standardized variants should be provided for them. Here please note that the author of N2771 treats without hesitation the "reversed letters" ꡫ A86B, ꡬ A86C, ꡭ A86D and ꡮ A86E (marking at most 20-30 words) appearing only in the Phags-pa monument discovered in the Juyong Pass as nominal glyphs (to which we do agree). Then, why should he treat in quite different ways these two sets of glyphs that also have certain distinguishing function in materials of different languages? We remain perplexed in spite of much thought. Doesn't it result from practicing a double standard based on Tibetan? It's worth mentioning, too, that as early as the 14th century the "opposition" of ꡚ,ꡚ ;ꡗ ,ꡯ ;ꡜ,ꡱ; ꡲ, ꡤ were clearly stipulated in the Table of Letters in "Mongolian Rhymes" in Phags-pa script. We believe that it's right to treat ꡯ, ꡚ, ꡱ, ꡲ as independent nominal glyphs.

(3) **Variants of the three letters ꡂ, ꡛ and ꡧ:** The N2771 asserts that there is no need to provide the nine variant presentation glyphs ꡂ 0032, ꡂ0034, ꡂ 0037, ꡂ 0039, ꡂ 003B; ꡛ 0013, ꡛ 0014; ꡧ 0049, ꡧ 004B. We insist that there be such need:

(a) The five glyphs from ꡂ 0034 to ꡂ 003B are found in Mongolian, Tibetan and Sanskrit documents and that most frequently, thus, they appear 30 odd times in Mongolian documents. Though ꡂ and ꡂ do not play any distinguishing role, they differ evidently in their figures, so we must regard ꡂ as variant of ꡂ.

(b) The two glyphs ꡛ 0013 and ꡛ 0015 appear 33 times in Chinese words, and ꡛ and ꡛ are found in one and the same word in a document. It's true that ꡛ and ꡛ do not have any distinguishing function, but they differ evidently in figures. So we have to regard ꡛ as variant of ꡛ

(c) The two glyphs ꡧ 0049 and ꡧ 004B appear in Chinese and Tibetan words, but not so many times, so there are scholars who read them in one sound ꡦ A866. It is thus clear that they differ greatly from ꡨ A868 in figures, a fact not to be ignored.

(4) **Variant presentation glyphs of vowel letters:** N2622 and N2745 agree largely as to the number of variant presentation glyphs of the vowel figures. In Table 4 and Table 5 of N2622 are listed 19 glyphs of 4 positions of the four vowel letters ꡝ, ꡠ, ꡡ and ꡢ, while in N2745 are listed 22 glyphs of four positions of the above four vowel letters. Compared with N2622, N2745 have three additional glyphs, viz., ꡥ, ꡥ and ꡡ. We have just discussed above the problem concerning ꡧ of ꡥ and ꡥ, and though ꡡ appears rarely, it is very peculiar in its connection so it should be preserved in the form of a variant presentation glyph.

(5) **The syllable delimiter:** The Phags-pa script itself does not have any device to tell apart

the limits of a syllable and of a word, to do so requires artificial treatment. For the convenience in handling follow-up information of Phags-pa script, we maintain that it is necessary to provide as syllable delimiter a space in UCS different from a common space like NNB/SP. It shows the limits of syllables in a word and requires a gap equal to one third of a space. For the sake of registration, there must be a gap after each syllable and each word in Phags-pa texts. It is evident that the fact that we design another gap different from common space as syllable delimiter and make it a rule to use it will greatly facilitate the treatment of words, sentences and discourses in Phags-pa script. For example, 𖡣𖡣 𖡣 𖡣𖡣 𖡣𖡣 𖡣𖡣𖡣 (𖡣𖡣 𖡣is a disyllabic word and 𖡣𖡣𖡣 𖡣𖡣 𖡣𖡣𖡣 a trisyllabic word), they should be registered as follows: 𖡣𖡣𖡣NNB/SP 𖡣𖡣SP 𖡣𖡣𖡣NNB/SP 𖡣𖡣NNB/SP 𖡣𖡣𖡣SP. We regulate in the Users' Agreement that users in need of a syllable delimiter may use a certain space, which will bring great advantage to the treatment of follow-up information, without any negative effect. If one does not want to use the syllable delimiter, he may of course use a common space at the end of a syllable. But such an intention must not serve as a pretext on which to object to using the syllable delimiter, very convenient for users in handling Phags-pa information. Both will do: to use either of these ready spaces.

(6) **Free variant selectors and their numbers:** That Phags-pa script encoding needs free variant selectors is a common understanding that various schools have reached. The number of variant selectors depends ultimately upon that of the variant presentation glyphs which need variant selectors. Taking into consideration the actual need to register the four syllable-final forms 𖡣, 𖡣, 𖡣 and 𖡣 of the letter 𖡣A860, we suggest to use three variant selectors. Concretely, users may use FVS$_1$ 180B-180D, etc., or VS$_1$ FE00-FE02, etc.

(7) **The joiner:** We propose, for the nominal glyphs, to put a joiner (on the right side) with a variant on the left side of the glyph. The joiner after a common glyph connected with the latter on the left side and the joiner after a reversed letter connected on the left side do not require special selection; suffice it to register letters in their order (without pressing the key "joiner") so as to properly register the needed joiner. However, joiners after non-reversed letters connected on the left side and joiners after reversed letters connected on their left side require special selection. Such selection may be made in several ways. One is to press the key "joiner"; another is to press a certain "distinguishing symbol". Both will do if viewed only from the "joiner" angle. At the same time, we are considering still one more problem, i.e., seeing that, with the exception of nominal glyphs, almost every Phags-pa glyph has one or more than one (syllable-initial and syllable-final) variant connected below, we may have two ways to handle them: one is, apart from nominal glyphs, to design in the databank separate variant glyphs that connect below; the other is to prepare only nominal glyphs in the databank without putting any separate variant glyphs that connect below, and to connect the joiner, through a pre-fixed program, after the proper nominal glyphs so as to automatically offer a variant glyph with a joiner. We believe that the second way can economize a lot of variant presentation glyphs (at least 38 glyphs according to tentative calculation). In consideration of the above, we propose to put such a joiner.

(8) **The letter 𖡣 A:** To our knowledge, there exist about three views on this letter: (a) 𖡣 is a consonant letter; (b) 𖡣 is a vowel letter, and (c) apart from its function as a vowel or a consonant, 𖡣 also plays a role as *titim*(top of a character). Different schools insist on one of the three views, upon which an agreement does not seem to be reached within a short time. Therefore, we think that Phags-pa script encoding which is meant merely for the studies of ancient writings,

should adapt itself to all these views without being partial to one while restricting another. So it is inevitable to register in different ways, though if properly handled, confusion can be avoided. Considerate of schools with different views, we offer them such possibility to register in three different ways so that they may make their respective choices, but if we were not prepared to do so, it would have been a matter of principle.

Moreover, we cannot agree with the following assertion, "Any one who knows Tibetan and Sanskrit will easily find it wrong to say that ꦄ (in Phags-pa script) is a vowel". The following are our reasons: (a) Tibetan scholars themselves explain the letter ꦄ in Tibetan in different ways. Some regard it as a consonant, others regard it as a vowel like Sa-skya Kun-dgav rgyal-mtshan in the 13th century and Ngag-dbang bstan-dar in the 19th century, still others think it is a free sound,

i.e., both a consonant and a vowel, like Gser-tog in the 19th century. (b) The Sanskrit vowel अA is transliterated into ꦄ in Tibetan and the ꦄ so used is regarded as a vowel. The ꦄ in Phags-pa script does not play the same role in presenting texts in different languages, so we must not explain it according to the Tibetan encoding.

(9) **Punctuation marks:** If the Phags-pa encoding can transfer for use those four punctuation marks in N2745, we will of course not bother to include them in the Phags-pa encoding. And it is also possible to borrow for use the vertical ideographic full stop '｡'(U+02DA). There is no need to design two kinds of BIRGA again in encoding of Phags-pa script because it can be borrow the BIRGA '〽' and its variant '〽〽' in U+1800. As for the two different SHAD, they ought to be treated the same as in the Tibetan-Phags-pa script.

(10) **Style of writing:** We propose to use the Khubilai style because not only its strokes look smart and smooth, but also Emperor Khubilai played an unreplaceable role in creating and popularizing Phags-pa script. Besides, the historical fact that Khubilai style was installed and used in the Mongolian editing device in Mongolia during the second half of the 20th century should also be taken into our consideration. If N2771 recognizes that the Khubilai style is the most good-looking and a very common style of writing, isn't there any reason not to adopt it?

(11) **Variants of ⎺, ⌐ and ⌐:** We think that these three glyphs can be well handled if we regard all three as variants of ⌐, or regard the other two as nominal glyphs. According to the standards we proposed in N2745 for defining variant presentation glyphs, there will be no problem if all three are treated as variants without resulting in any confusion. It will also do if, as is proposed in N2662, two of them are handled as nominal glyphs and the third a variant. We do not insist on our own viewpoint, but let scholars decide it through consultations.

(12) **Order of encoding:** Our proposal for the encoding is based on most academic monographs on Phags-pa script published in recent years by scholars in China and Mongolia, since luckily it won't influence too much on the whole process of encoding. It's evident that no matter what order is agreed upon, it is inevitable for each language to formulate its order of procedure in line with its own habits and needs. Such being the case, we abandon the order of encoding proposed in N2745, and largely agree to that of N3696, only that those six nominal glyphs we put forward should be added therein, but the classification of glyphs be cancelled.

(13) **Names of glyphs:** On the names of the following several letters, scholars vary in their views. Our opinions are as follows:

(a) Names of ꧀ and ꧀: We agree with N2829 to name ꧀ A862 **QA** and ꧀ A863 **XA**.

(b) Names of ꧀ and ꧀: It is accepted that ꧀ A866 is to be called **EE** and ꧀ A860 called **E**.

(c) Name of ᡅ: It is accepted that ᡅ A85D is to be called "A". The ᡱ A856 can be called "**MINUSCULE A**".

(d) Names of ᢏ A86F, ᢐ A870, ᢑ A871, ᢒ A872: Names of these four letters should be different from names of the following four letters ᢏ **A857 PHAGS-PA LETTER YA，** ᢐ **A85A PHAGS-PA LETTER SHA，** ᢑ **A85C PHAGS-PA LETTER HA，** and ᢒ **A864 PHAGS-PA LETTER FA**. We are inclined to call ᢏ **A86F PHAGS-PA LETTER 0A**, ᢐ **A870 PHAGS-PA LETTER VOICELESS SHA**, ᢑ **A871 PHAGS-PA LETTER VOICED HA** and ᢒ **A872 PHAGS-PA LETTER ASPIRATED FA**. Our reason is that it is only when these four letters are used to register Chinese that they distinguish themselves from ᢏ **A857 YA，** ᢐ **A85A SHA，** ᢑ **A85C HA** and ᢒ **A864 FA**, so they should be named in line with their sound forms in Chinese. According to Wang Li, master in history of Chinese language, they are pronounced largely as ᢏ **A86F [∅], ** ᢐ **A870[ʃ], ** ᢑ **A871[ɣ] and ** ᢒ **A872[fʻ]** .

(e) Names of ᢇ and ᢈ: It is accepted that ᢇ A867 is to be called "**PHAGS-PA SUBJOINED LETTER WA**"and ᢈ A868 to be called **"PHAGS-PA SUBJOINED LETTER YA".**

(f) Name of ᢍ: Though appearing in various Tables of Letters, this glyph has actually never been found in any Phags-pa monuments. Not knowing its sound form, we tentatively call it "GGA".
(g) Name of ☋: We agree with N2829 to name "CANDRABINDU".