提交 UTC 文档：

UTC Document Submission:

关于在蒙古文编码中添加字符的提议
A Proposal to Add Characters to Mongolian Coding

Ma Fuquan

March 30, 2020

蒙古文字是从上到下书写，从左到右换行，并且字母之间用词脊相连的拼音文字。蒙古文字虽然有复杂的字母变形，但现代蒙古文字已经形成了稳定的蒙古文正字法。蒙古文正字法直接影响蒙古文编码的定义以及字符变形规则。本人认为蒙古文编码存在以下四个方面的问题：

Mongolian characters are alphabetic characters that are written from top to bottom, with lines breaking from left to right and letters of them are connected by word ridges. Although there is complicated letter variation in Mongolian characters, the modern Mongolian characters have formed a stable Mongolian orthography. Mongolian orthography directly

affects the definition of Mongolian coding and the variation rules of characters. There are four problems existing in Mongolian coding.

一、 蒙古文借词字母变形问题

I. Variation of alphabetic letters in Mongolian loanwords

蒙古文编码中控制符 FVS1、FVS2、FVS3 主要用于借词字母强制变形。蒙古文借词分为两种，一种是纯蒙古文字母书写的借词，这种借词遵循蒙古文正字法，输入时不需要强制变形控制符；另一种是蒙古文中借词专用字母及蒙古文字母混合使用的借词，这种借词需要使用控制符强行变形，如：MONGOLIAN LETTER U、MONGOLIAN LETTER TA，MONGOLIAN LETTER DA、MONGOLIAN LETTER NA 等字母需要强行变形。

本文主要说明借词字母及蒙古文字母混合使用的借词，简称借词。

The format controls FVS1, FVS2 and FVS3 in Mongolian coding are mainly used for the forced variation of alphabetic letters in loanwords. There are two kinds of loanwords in Mongolian language. One is the loanwords written in pure Mongolian letters. This kind of loanwords follows the Mongolian orthography and does not need the format controls. The other is the letters for loanwords in Mongolian combined with pure Mongolian letters. They use the format controls to forcibly vary the letters such as MONGOLIAN LETTER U，MONGOLIAN LETTER TA，MONGOLIAN LETTER DA and MONGOLIAN LETTER NA. In the following, we study the second one, namely the loanwords forming by combing the letters from the source language with Mongolian letters,
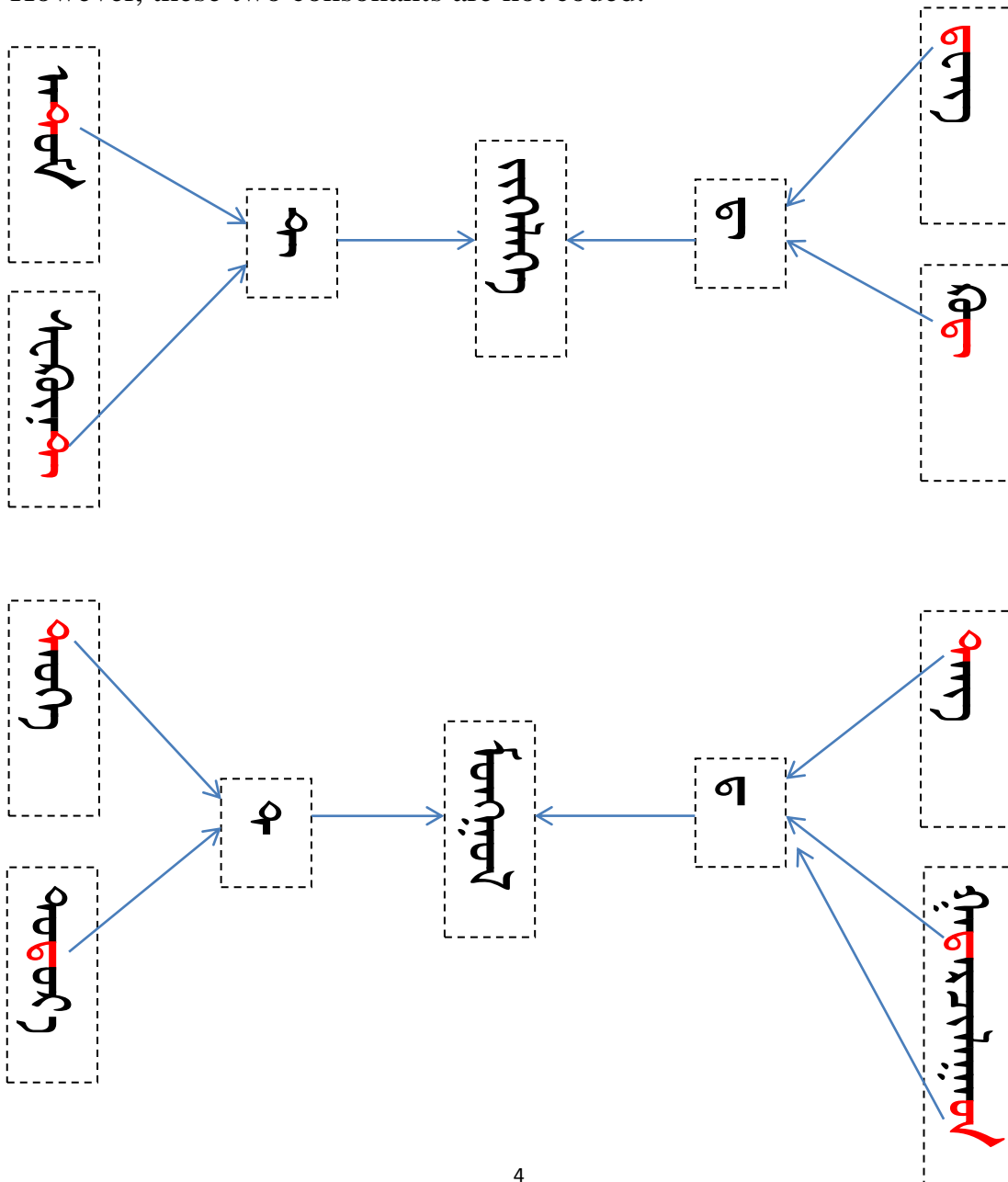
which we call loanwords for short.

（一）借词中元音字母及编码

1. Vowel letters in loanwords and their codes

借词一般使用《 ᠠ a、ᠡ E、ᠢ i、ᠣ o、ᠦ Ü》五个元音。其中两个阳性元音、两个阴性元音和一个中性元音，编码依次为《1820、1827、1822、1823、X 》，其中 X 代表没有编码。借词中阴性元音在词首、词中和词尾的字形与蒙古文阴性元音在词首、词中和词尾的字形有着明显的差别。

The five vowels " ᠠ a、ᠡ E、ᠢ i、ᠣ o、ᠦ Ü " are generally used in loanwords. They are two masculine vowels, two feminine vowels and one neutral consonant. The codes of the five vowels are 1820, 1827, 1822, 1823 and X in sequence, in which X represents no code. There is a clear distinction between the glyphs of the initial, medial and final feminine vowels in loanwords and those in Mongolian words.

（二）借词中辅音字母及编码

2. Consonant letters in loanwords and their codes

借词中辅音字母共有 26 个，包括 14 个蒙古文通用辅音字母和 10 个专用辅音字母，另外 MONGOLIAN LETTER TA 和 MONGOLIAN LETTER DA 是借词专用辨别字母，这两个辅音没有编码。

There are 26 consonant letters in loanwords in total, of which 14 are the same as the Mongolian consonant letters and 10 are for special use. The remaining two, MONGOLIAN LETTERTA and MONGOLIAN LETTER DA, in loanwords are only used to distinguish between letters. However, these two consonants are not coded.

（三）借词中复辅音的使用规则

3. Rules for the consonant clusters in loanwords

1.借词词首或词中复辅音第一个字母《ᠺ》使用 MONGGOLIAN LETTER ALI GALI KA。例如，ᠺᠥᠪ · ᠺᠯᠠᠰ · ᠺᠷᠠ᠊ · ᠺᠣᠯᠠᠰᠠᠷ · ᠺᠣᠨᠠᠷᠠ · ᠺᠯᠢᠨ 等。部分借词缀加蒙古文词缀。例如，ᠺᠯᠠᠰᠠᠨ · ᠺᠣᠯᠠᠰᠠᠨ 等。

***Here consonant clusters only refers to the clusters of two consonants.

3.1. MONGGOLIAN LETTER ALI GALI KA is used for "ᠺ" if it is the first letter of the initial and medial consonant clusters in loanwords. For example, the words ᠺᠥᠪ · ᠺᠯᠠᠰ · ᠺᠷᠠ᠊ · ᠺᠣᠯᠠᠰᠠᠷ · ᠺᠣᠨᠠᠷᠠ · ᠺᠯᠢᠨ follow this rule. Some Mongolian affixes are added to loanwords. For example, ᠺᠯᠠᠰᠠᠨ · ᠺᠣᠯᠠᠰᠠᠨ, etc.

2. 借词复辅音第一个辅音字母《ᡭ》在阳性元音或中性元音后使用不带点变形字形，词中或词尾音节中阴性元音后使用带点字形。例如，ᡭᠣᠯᠢᠰᠣ · ᠭᡭᠡᠷᠲ · ᠺᡭᠥᠷᠲ · ᠺᡭᠣᠷᠠᠳᠣᠷ 等。词首阴性音节《ᡭᠥᠷ》后面使用带点变形字形。例如，ᡭᠥᠷᠺ 等。

3.2 The first consonant letter "ᡭ" of consonant clusters in loanwords loses its dot after the masculine vowels or neutral vowels, but it restores its dot when following the feminine vowels in the middle and at the end of the syllables. For instance, the words ᡭᠣᠯᠢᠰᠣ · ᠭᡭᠡᠷᠲ · ᠺᡭᠥᠷᠲ · ᠺᡭᠣᠷᠠᠳᠣᠷ follow the above rule. The feminine syllable "ᡭᠥᠷ " is followed by the

dotted varied characters. The word ᡥᠣᠨᡳᠨ is such an example.

解决方法：保留 FVS1、FVS2、FVS3 等三个格式控制符用于强制变形蒙古文古文字形，蒙古文编码里添加如下三个特殊变形字母，属性设置为与蒙古文对应几个字母的大小写字母关系

Solution: The format controls FVS1, FVS2, FVS3 are still used in forced variation in ancient Mongolian literature. The following three special varied glyphs are added in Mongolian coding and their orientations are set to the correspondence of several Mongolian upper-case and lower-case letters.

ᡤ MONGOLIAN CAPITAL LETTER GE

ᡐ MONGOLIAN CAPITAL LETTER TA

ᡑ MONGOLIAN CAPITAL LETTER DA

二、双根词问题

II. The existing problems of two root words

《蒙古文诠释》（1979 年嘎拉桑著）、《蒙文原理》（1987 年策•诺尔金著）、《蒙古语书面语法》（1990 年博•仁亲著）、《明慧宝椟》（2009 格•拉西色楞著）、《现代蒙古语》（涛高、援朝、图力古尔、额尔敦初古拉、嘎拉桑等著）等语法书中均讲到字冠（ᠬᠤᠪᡳ）的隐形原理、变形原理、恢复原理，证明字冠是蒙古文正字法双根词连接字符。

The principles of invisibility, variation and restoration of the alpha (ᠬᠤᠪᡳ) have also been discussed in the grammar books listed in the following: *Interpretation of Mongolian Language* (Galseng, 1979), *The*

Principles of Mongolian Language (Tsur Norking, 1987), *The Written Grammar of Mongolian Language* (Bo Renqin, 1990), Minghui Baobi (Ge Racisereng, 2009), *The Modern Mongolian* (Togao, Yuanchao, Tulgurr, Erdun Chogla, Galseng, etc.). All these books prove this alpha ᠣᠣᠷ is a character for connecting in Mongolian orthography.

解决办法：1800 区添加《字冠 ᠣᠣᠷ》字符（MONGOLIAN Aleph）。

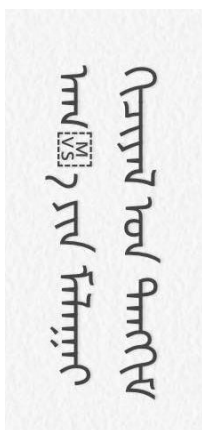**Solution:** add the prefix ᠣᠣᠷ (Mongolian Aleph) in the zone 1800.

三、NNBSP 不间断空格符使用不稳定

III. The unstable use of the uninterrupted space character NNBSP

苹果系统中 NNBSP 下文附加成分起首字母不变形。

The initial letters of the suffixes after NNBSP in the Apple system do not vary.

解决方法：把 NNBSP 移至 1800 区，并命名为 MONGOLIAN SPACE (MSP) , MSP 不仅不间断换行，还影响上下文变形。MSP 和 MVS 的字形改为空格字形。

Solution: To move NNBSP to 1800 zone and name it MONGOLIAN SPACE (MSP), which can not only change lines without breaks, but also affect the context variation. The glyphs of MSP and MVS are changed to space glyphs.
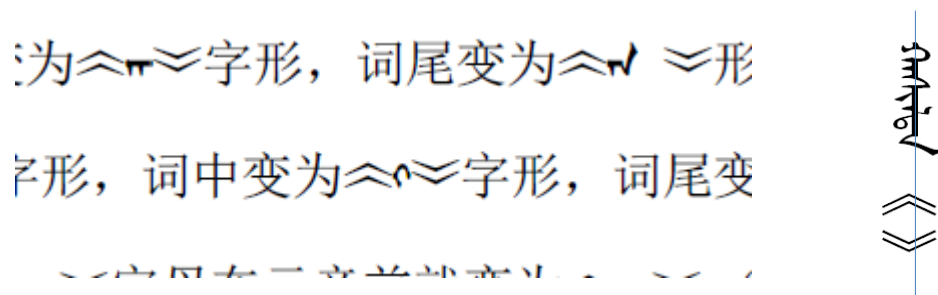
四、通用标点符号问题

IV. The problems of CJK Symbols and Punctuations

蒙古文借用通用标点符号存在两种问题，一是旋转属性与蒙古

文不一致；二是符号中心线与蒙古文词脊不一致。

There are two problems existing when using CJK Symbols and Punctuations in Mongolian. One is that the rotation orientation is not consistent with Mongolian language. Second, the center line of the punctuation marks is not consistent with the ridge of Mongolian words.

解决方法：在表情符号区添加蒙古文标点符号《• ?! !? ; ！ ？|（ ）《 》〈 〉［ ］ 》等标点符号。

Solution: add Mongolian punctuation marks "•", "?!", "!?", ";" ,"!", "?", "|"," （ ）"," 《 》" ," 〈 〉", " ［ ］ "   in the emoticon zone.

总之，FVS1、 FVS2、 FVS3 等控制符的使用不符合现代蒙古文正字法，所以导致字体厂家使用控制符规则不统一，标点符号带来的问题影响用户正常使用蒙古文编码。

In short, the use of the format controls such as FVS1, FVS2, FVS3 does not meet the modern Mongolian orthography, which causes the fonts manufactures to employ different rules for format controls. And the punctuation problems affect the normal use of the Mongolian coding.