

Universal Coded Character Set  
International Organization for Standardization  
Organisation Internationale de Normalisation  
Международная Организация по Стандартизации

**Doc Type:** Working Group Document  
**Title:** Proposal to deprecate the Mongolian standardized variants in UCD  
**Author:** Kushim Jiang  
**Status:** Individual Contribution  
**Action:** For consideration by UTC  
**Date:** 2026-04-02

## 1 Proposal

This proposal requests that all entries related to variants of Mongolian characters in UCD's [StandardizedVariants.txt](#) be deprecate, and update the relevant sections of this text file and the *Core Specification* in subsequent revisions.

## 2 Introduction

### 2.1 Description in the current *Core Specification*

As stated in the *Core Specification*:

The list of standardized variants in `StandardizedVariants.txt` for Mongolian has not yet been updated to synchronize with the requirements of current practice as stated in [UTN #57](#). This defect will be addressed in a future version of the Unicode Standard.

The background to this statement is that the standardized variants defined in the UCD reflect earlier analysis, while the variants described in the latest China national standards reflects the current analysis. Since there are discrepancies between the two, the data in the UCD will eventually need to be updated to reflect the current analysis in the latest China national standards, and use [UTN #57](#) to provide informative documentation.

The current analysis referred here pertains to the following four China national standards (For links to these documents, see also the [Topical Document List](#)):

- [GB/T 25914-2023](#): *Information technology — Traditional Mongolian nominal characters, presentation characters and use rules of controlling characters*
- [GB/T 36649-2018](#): *Information technology — Todo nominal characters, presentation characters and use rules of controlling characters*
- [GB/T 36641-2018](#): *Information technology — Sibe nominal characters, presentation characters and use rules of controlling characters*
- [GB/T 36645-2018](#): *Information technology — Manchu nominal characters, presentation characters and use rules of controlling characters*

The four China national standards listed above specify the characters, variant assignments, shaping rules, and ligature definitions for Hudum (with Ali Gali), Todo (with Ali Gali), Sibe, and Manchu (with Ali Gali), respectively. According to the editorial group, the final three national standards currently under revision have completed the public consultation phase and passed the preliminary review meeting within CESI, and have now been submitted to NITS for expert review.

## 2.2 Original motivation for establishing FVS

As can be seen from early proposals and the *User Agreement*, the original motivation for establishing FVS was to manually handle exceptions to the orthographic shaping logic, noted that the shaping of many loanwords and proper names does not conform to the rules of traditional orthography.

Furthermore, the functionality of FVS as defined in this context can be viewed as **switching** from one variant to another under a certain joining position, rather than **determining** a specific variant. In some cases, FVS1 is used when switching from variant A1 to variant A2, and FVS1 is also used when switching back from variant A2 to variant A1, because these two switching scenarios are mutually exclusive. In other examples, switching from variant A1 to variant A2 uses FVS1, switching from variant A3 to variant A4 uses FVS2, and in other cases it is undefined simply because no such switching scenario exists.

## 3 Reasons

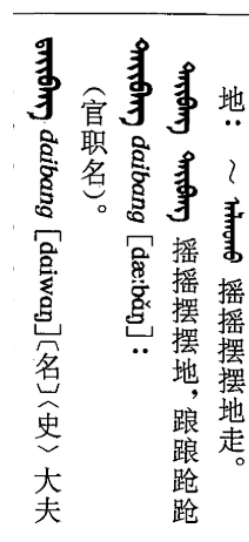
The Mongolian standardized variants in UCD need to be deprecated, because the initial motivation for FVS and the actual characteristics of Mongolian character variants do not meet the criteria for establishing standardized variants, and such variants should not be classified as free variants, for the following five key reasons.

### 3.1 Variants are orthographically required and semantically significant

Mongolian variants at specific joining positions are not ignorable but required by orthographic rules, and even distinguish different semantic meanings (for example, in “*daibang*,” as shown in the figure on the right, the variant of the first letter distinguishes the meaning. The former refers to the official rank of “*Dafu*,” while the latter means “staggeringly.”), making them essential rather than truly free variants.

### 3.2 FVS as a switcher rather than a decider

The original purpose of introducing FVS was to switch between character variants, not to determine them, which is inconsistent with the core logic of standardized variants. National standards have only recently begun to shift toward genuine variant determination, further highlighting the misalignment.



### 3.3 Lack of systematic effect on all variants

This reason applies to cursive script. Manichaean variants illustrate how standardized variants affect every variant of a character simultaneously, as shown in [WG2 N3644R](#). On the one hand, the variant does not alter the recognition of character unity. On the other hand, its occurrence is unpredictable—though such variants tend to appear at the end of lines in the manuscripts—and can therefore be regarded as free variants. From these two perspectives, the design of Manichaean variants aligns with the original intent of standardized variants.

Character	X <sub>n</sub>	X <sub>r</sub>	X <sub>m</sub>	X <sub>l</sub>
MEM	𐎌	𐎌	𐎌	𐎌
MEM + VS-1	𐎌	𐎌	𐎌	𐎌
DALETH	𐎎	𐎎		
DALETH + VS-1	𐎎	𐎎		
RESH	𐎐	𐎐		
RESH + VS-1	𐎐	𐎐		
NUN	𐎑			𐎑
NUN + VS-1	𐎑			𐎑
HE	𐎓			
HE + VS-1	𐎓			

However, such a systematic change is impossible for the Mongolian script, because its phonetic model does not organize variants based on graphemes; the graphemes corresponding to its FVS vary depending on the joining position. For example, the figure on the next page shows the arrangement of FVSs for u1823 according to the current national standard. Clearly, AO, O, and U do not correspond to specific FVSs; in other words, the function of FVSs is not to convert AO into O or U and then have O and U participate in cursive joining.

1823 ᠠᠣ	-	ᠠᠣ	ᠠᠣ	ᠣ	
	1			ᠠᠣ	ᠣ
	2				ᠤ

### 3.4 Mismatch between SVS function and font implementation expectation

The way standardized variants function in a font is completely different from how they are expected to function. The standardized variant operates during the CMAP phase, which precedes GSUB and GPOS; however, in practical implementations, it is generally expected to take effect during the GSUB phase.

This is because FVS are typically used to handle exceptional cases; therefore, they should be ignored during the orthographic shaping step and only take effect after that step is complete to determine a character's variant. In the original design of the encoding model, the FVS settings for one character do not affect the shaping process that should have been completed for other characters.

### **3.5 UCD cannot currently address writing system-specific requirements**

In the latest version of the four China nation standards, the expectation of character variants at specific positions is tied to specific writing systems. Because the Mongolian encoding model combines letters from different writing systems into the same character, different writing systems may expect different glyphs at certain joining positions for some characters. For example, if a letter in a particular writing system does not exist in the word-final position, that writing system would expect the glyph in the final position to be a virtual glyph, rather than a glyph from another writing system.

The complexity of the UCD data is currently insufficient to meet these demanding requirements. The definition of character variants at various joining positions in Mongolian depends on the writing system, and the definitions of their functions and conditions for occurrence are far more complex than those in SVS. There is no practical basis for the definitions in SVS, and these definitions should not be applied in SVS.

## **4 Feasibility**

As can be seen from the contents of [StandardizedVariants.txt](#), some variant definitions in this file have been defined and then commented out. Along with commenting out these variants, the line above explains the rationale for this decision and the relevant UTC resolution it complies with. It therefore appears that it is feasible to comment out the Mongolian-related standardized variants in the UCD via a UTC resolution.

**(End of Document)**