L2/23-101

Text representation and shaping manual of the Mongolian script
蒙古文字的文本表记与成形手册

Kushim JIANG 姜兆勤, 2023–04–01

Contents 目录

1 Introduction 引言

Two decades ago, ISO/IEC 10646 and Unicode added support for Mongolian with a seemingly reasonable character set. At the time, little was known about how exactly these characters should be rendered. Since then, various vendors have struggled to make their own sense of the characters, while users have suffered both from the lack of interoperability between the vendors' implementations and from the resulting lack of native support on the major platforms.
ISO/IEC 10646 和 Unicode 标准于约二十年前起开始支持蒙古文字，提供了看似合理的字符集，而当时人们对这些字符应如何呈现知之甚少。此后，各厂商一直努力界定这些字符的功能，用户则一直为厂商实现之间欠妥的兼容性与主要平台上由此产生的本地化欠支持而痛苦。

Experts have uncovered problems with Mongolian encoding and proposed various patches, especially in recent years. However, until recently (MGC/01–01, 2018), there have been few complete specifications available for the community to discuss, develop, and ultimately agree upon.
字符集设计者们一直提出蒙古文字字符集相关的议题，并在近几年提出了许多修正方案。而直至最近（MGC/01–01 标准，2018）才出现完整规范可供讨论与发展，以便最终达成共识。

This *draft* specification demonstrates and proposes how to specify both comprehensive guidelines for text representation and coherent rules for text shaping. It deals mainly with the first three layers in the overall technical architecture of text rendering in the Mongolian script:
本规范草案展示并提供全面的文本表记与一致的文本成形规则。主要涉及使用蒙古文字呈现文本的整体技术架构的前三层：

- The Unicode Standard and ISO/IEC 10646, a synchronized pair of standards, specify the identities of encoded characters and standardized variation sequences.
  Unicode 标准与 ISO/IEC 10646，一对规定字符身份及其标准变体序列的同步标准。

- The Unicode Standard and its various supplementary standards further provide the characters with additional behavioral specifications, including character properties (general category, cursive joining type, etc.) and algorithms (normalization, collation, line breaking, text segmentation, bidirectional, vertical text layout, etc.).
  Unicode 标准及诸补充标准，进一步为字符提供其他行为规范，含字符属性（通用类别、连写连接类型等）与算法（规范化、排序、移行、分段、双文向、竖排等）。

- The required complex shaping is implemented in fonts and text shaping engines, according to OpenType Layout (OTL) or other complex shaping models, such as AAT (Apple Advanced Typography) and Graphite.
  根据 OTL 或 AAT、Graphite 等复杂成形模型，在字库与文本成形引擎中实现所需的复杂文本成形。

- Mongolian text is treated inline as horizontal and left–to–right, while text layout engines are responsible for setting lines vertically and arranging multiple lines in the preferred left–to–right

order.
蒙古文字文本在行内视为水平自左向右的文向。排版引擎负责自左向右地排列纵行。

Although the architectural flaws cannot be fixed without migrating to a radically different encoding model, careful specification can help eliminate unwanted differences between vendor implementations. Specifically:
尽管在不迁移至具备根本差异的字符集模型的条件下无法根本解决架构缺陷，细致的规范仍可帮助消除厂商实现间不必要的差异。具体而言：

- Retain the phonetic encoding principle of the existing model. That is, encode the underlying phonetic letters of a given text as characters, then contextually shape and manually control the encoded characters to reproduce the original written forms in the given texts. If this principle is not maintained, we would be better off migrating to a graphetic encoding model designed from scratch.
保持现有音位模型。即以音位字母定义字符，通过上下文成形过程与手动控制过程呈现文本的具体书写形式。若不保持该模型，迁移至与该模型无关联的形位模型更优。

- Minimize confusability by restricting the ways in which a given script can be graphically encoded. In particular, limit the overlap between the variant sets of phonetic letters. Also, moderately normalize orthographic inconsistencies in how the underlying phonetic letters of certain words are interpreted.
通过限制书写形式的形位分析以最小化可混淆性，尤其限制音位字母变体集之间的重叠情况。同时适度规范正写法在分析部分单词的音位字母上的分歧。

- Prioritize the modern orthography in day–to–day usage. Such a problematic encoding model needs to be standardized to be as restrictive as possible, and should not be so extensible as to include marginal and scholarly use cases. Compromising historical orthography when phonetically representing them harms the text exchange of modern orthography.
在日常使用中优先考虑现代正写法。字符集模型应尽可能受控，以减低对边缘情况与学术要求的扩展性。使用音位模型向历史正写法妥协无益于现代正写法下的文本交流。

- Formally define modularized and ordered text shaping rules that are unambiguous and have well–defined fallback mechanisms. Use the Twelve Syllabaries as a foundation. Avoid introducing contextual rules that seem productive enough in modern orthography (especially loanwords) but are absent in the Twelve Syllabaries. Minimize dictionary–based shaping to reduce unnecessary assumptions. Systematic guidelines for text representation will become a natural consequence of the strict shaping rules.
形式化定义模块化、层次化、无歧义、优先级明确的文本成形规则。以十二字头为基础，避免引入对现代正写法看似有效（主要为借词情形）但不见于十二字头的规则。最小化字典式成形规则，以减少不必要的假设。系统化的文本表记成为严格的成形规则的自然结果。

2 Architecture 架构

The encoding models developed for each writing system in the Mongolian script have homogeneous structures, so the generalized architecture is given in this chapter, and descriptions of specific writing systems are given in the following chapters.
对蒙古文字的各书写系统建立的编码模型具有等价架构，因此本章给出一般架构，后续章节对具体书写系统加以描述。

2.1 Data files and the standard toolchain 数据文件与标准工具链

The Mongolian shaping data files formally capture the additional identity information and shaping behavior of Mongolian characters in a machine–readable format. Appendix X, *A tutorial on font production*, demonstrates how to utilize the data files in font production.
蒙古文字成形数据文件以机器可读格式形式化描述蒙古文字字符的额外信息与成形行为。附录 X《字库生产导引》将演示如何在字库生产过程中利用数据文件。

2.2 Character set 字符集

A typical Mongolian writing system implementation requires the characters shown in Table 1. The table includes both Mongolian–specific characters as well as characters that are shared with other scripts.
典范蒙古文字书写系统实现需要表 1 所示的字符，含特有字符与共用字符。

Table 1 Required characters
表 1 所需字符

Script 文字	Type of characters 字符类别	Character example 字符示例	Note 备注
General	Space	space	
	Punctuation	middle dot, …	
	Format controls	ZWJ, ZWNJ, …	participate in shaping
	Digits	digit one, …	
Mongolian	Punctuation	birga, …	less used now
	Format controls	nirugu, FVS, …	participate in shaping
	Digits	digit one, …	less used now
	Phonetic letters	letter *a*, …	participate in shaping
CJK	Punctuation	question mark, …	
	Digits	digit one, …	

2.2.1 Phonetic letters and written units 音位字母与书写单位

Writing systems in the Mongolian script do not have a well–recognized system of typical letters (i.e., user–perceived primary units of writing) that is common to most writing systems today. Instead, users are accustomed to identifying letters at a much more phonetic level, where letters are not reliably related to writing and are therefore considered to be phonetic letters in this specification.
使用蒙古文字的书写系统通常没有大多数书写系统所常见的典型字母系统（即用户认为的主要书写单位），用户更习惯于在语音层面上识别字母。字母与书写之间没有固有联系，因此在本规范中被认为是音位字母。

Multi–to–one confusion. The system of phonetic letters is largely based on the historical phonemes of the Mongolian written language, as reflected in the conservative orthography, rather than on how the under–differentiated writing system actually works with its limited set of writing units. As a result, many phonetic letters do not have distinct written forms and can be confused in writing, but are still identified as distinct letters because they are intended to be the abstract representatives of distinct phonemes.
多对一的混淆性。音位字母系统主要基于传统正写法所反映的书面语言的历史音位，而非书写系统在其有限的书写单位中的实际运作。因此一些音位字母没有代表性的书面形式，在书写中可能会出现混淆，但仍被认定为独立的字母，因为它们是不同音位的抽象指代物。

One–to–multi unpredictability. Furthermore, the phonetic letters can be written with multiple different written units and/or written unit sequences. Exactly which written form is used to represent a phonetic letter is determined by a combination of complex predictive rules and arbitrary variations. Many phonetic letters have largely unpredictable correspondences to their written forms.
一对多的不可预测性。此外，音位字母可用多个不同的书写单位及其序列来书写。由于具体使用何种书写形式来表示一个音位字母是由复杂的预测规则和自由的变形过程共同决定的，因此一些音位字母与其书写形式的对应关系基本上是不可预测的。

2.2.2 Format controls 格式控制符

Zero Width Non–Joiner (ZWNJ), Zero Width Joiner (ZWJ), and nirugu. U+200C ZERO WIDTH NON–JOINER and U+200D ZERO WIDTH JOINER are Unicode's standard cursive joining controls. Note that ZWJ also breaks interaction (such as ligation) between two consecutive characters since it is treated as an invisible character. U+180A MONGOLIAN NIRUGU is a Mongolian–specific modifier letter that behaves exactly like ZWJ but is visible as a piece of stem stroke. ZWNJ and ZWJ should not be accessible to the average user on common keyboard layouts, as everyday text does not require these tricky characters.
零宽禁连符（ZWNJ）、零宽连接符（ZWJ）与尼茹股。ZWNJ 与 ZWJ 为 Unicode 的标准连写控制字符。注意 ZWJ 还会破坏相邻字母之间的联系（如连字形式），因其视作不可见字母。尼茹股是蒙古文字特有的修饰字母，其行为与 ZWJ 完全一样，但作为可见的字干而存在。ZWNJ 与 ZWJ 应避免通过键盘布局被普通用户访问，因为日常文本不包含这些字符。

The visible character nirugu should be used to cause joining in everyday text. A common use case is to end a patronymic abbreviation that is the initial syllable body (i.e., an optional onset plus the first vowel) or just the initial consonant letter of the father's name.
可见的尼茹股应在日常文本中使用以形成连接。常见的使用情况是用于父名缩写末，该缩写可能为词首音节（即可选的声母辅音加上词首元音）或词首辅音字母。

Vowel Separator (MVS) and Narrow No–Break Space (NNBSP). MVS is a Mongolian–specific format control for requesting the chachlag variation. It is transcribed as "·" (a middle dot). NNBSP is a Mongolian–specific whitespace and format control used to mark and shape particles. It is transcribed as "–" (an en–dash). Use of the NNBSP is discouraged in preference for the MVS.
元音分隔符（MVS）与窄宽不移行空格（NNBSP）。MVS 是蒙古文字特有的格式控制符，用于产生分写变体。MVS 在本文中转写为"·"（一个间隔号）。NNBSP 是蒙古文字特有空格与格式控制符，用于表记和呈现词缀。NNBSP 在本文中转写为"–"（一条半身线）。目前，NNBSP 不被推荐使用，相关功能由 MVS 完全接管。

Free Variation Selectors (FVS's). FVS's are Mongolian–specific format controls. As combining marks, they are applied to certain characters to request the forms not captured by the predictive shaping rules.
自由变体选择符（FVS）。FVS 是蒙古文字特有的格式控制符。作为结合记号，它们出现于某些字符之后，用于产生未被正写法成形规则捕获的形式。

2.3 Shaping process 变形流程

The Mongolian text shaping process is based on the well–implemented technology foundation for general scripts and cursive scripts, while an additional phase of Mongolian–specific shaping steps is inserted into the ordinary shaping process required by cursive scripts. The minimal shaping process consists of a number of steps as shown in Table 2.
蒙古文字的文本成形过程基于一般文字与连写文字的实现成熟的技术基础，并在连写文字所需的一般成形过程中插入了蒙古文字的特定文本成形步骤阶段。最低限度的变形流程由若干步骤组成，如表 2 所示。

2.3.1 General shaping phases 通用变形阶段

These are the basic mechanisms in fonts that apply to all scripts.
通用变形阶段是字库的基础机制，适用于所有文字。

The basic character–to–glyph mapping (phase IA) is typically controlled by the TrueType/OpenType table `cmap`. The Unicode representative glyphs can be used here as the default glyph mappings for phonetic letters, but these representative glyphs are essentially irrelevant to the final rendering.
基本的字符 – 字图映射（IA 阶段）通常由 TrueType/OpenType 的 `cmap` 表控制。可以使用

Unicode 名义字图作为音位字母映射后的默认字图，但这些名义字图与最终渲染完全无关。

Vertical forms of punctuation marks (phase IB) are critical to the proper setting of Mongolian text, but are not part of the complex shaping between letters and format controls.
标点符号的竖排形式（IB 阶段）对蒙古文字的正确排版至关重要，但不属于复杂成形过程。

Table 2 Overview of shaping process
表 2 变形过程总览

Shaping phase **变形阶段**		**Shaping step** **变形步骤**
IA. General		· Basic character–to–glyph mapping
IIA. Cursive script		· Initiation of cursive positions
III. Mongolian–specific *Reduction of phonetic letters to written units*	*Phonetic*	1. Chachlag
		2. Syllabic
		3. Particle
	Graphemic	4. Devsger
		5. Post–bowed
	Uncaptured	6. FVS–selected
IIB. Cursive script (continued) *Sub–written–unit variations*		1. Variation involving bowed written units
		2. Cleanup of format controls
		3. Optional treatments
IB. General (continued) *Typography*		1. Vertical forms of punctuation marks
		2. Optional treatments

2.3.2 Cursive script shaping phases 连写文字变形阶段

On top of the general shaping mechanisms, complex scripts require additional shaping phases to be inserted after the basic character–to–glyph mapping and before typographical treatments. In particular, cursive scripts all undergo the cursive joining mechanism.
在通用变形机制的基础上，复杂文字需要在基本的字符 – 字图映射之后和文字设计处理之前插入额外的成形阶段。特别地，连写文字都须经历连写连接机制。

Cursive joining. Written forms exhibit the cursive joining mechanism (phase IIA). Both sides of a written form can either be joined to an adjacent written form or not, with up to four different states. Or, more abstractly, each written form is in one of the four cursive positions:
连写连接。连写文字的书写形式将经历连写连接机制（IIA 阶段）。其中，一个书写形式的两侧既可与相邻书写形式连接，也可不连接，从而最多表现出四种状态。即每个书写形式都处于以下四个连写点位之一：

- *isolated*, abbreviated as “isol”: not joined forward, not joined backward;
  非连形，缩写为“isol”：前方不连，后方不连（即为蒙古文字中的上方不连，下方不连，后同）；
- *initial*, “init”: not joined forward, joined backward;
  后连形，缩写为“init”：前方不连，后方连；
- *medial*, “medi”: joined forward, joined backward;
  双连形，缩写为“medi”，前方连，后方连；
- *final*, “fina”: joined forward, not joined backward.
  前连形，缩写为“fina”，前方连，后方不连。

Cursive positions are irrelevant to word boundaries, although they are usually consistent with word–wise positions in Mongolian because cursive joining breaks within a word are limited in the writing system.
连写点位与词边界通常无关，尽管在蒙古文字中二者通常相关，因为蒙古文字的词内不连情形有限。

Implementation. The nominal glyph of each phonetic letter will be mapped to the default glyph of that letter at a given cursive position.
实现。每个音位字母的名义字图将被映射为该字母在该连写点位下的默认字图。

Graphemic variation after bowed written units. Before the sub–written–unit variation, bowed written units may first cause a vowel to change its form.
圆头书写单位后形位变化。在次书写单位变形前，圆头书写单位可能先引起元音形式变化。

2.3.3 Mongolian–specific shaping phases 蒙古文字特有的变形阶段

Phase III consists a series of steps for Mongolian–specific shaping requirements, and within each step there may be more than one set of non–overlapping rules, each for a different group of letters. Forms not captured by the predictive conditions are requested with FVS's.
III 阶段由一系列蒙古文字特定的变形要求的步骤组成，每个步骤中可能有不止一组互斥规则，每组规则针对不同的字母。欲产生未被预测条件捕捉到的形式，则要求使用 FVS。

2.4 Notation 凡例

Phonemic letters are transcribed as Italicized lowercase letters, and written units are transcribed as Roman small capital letters. For transcriptions without diacritics, phonemic letters are transcribed as all lowercase letter sequences, and written units are transcribed as beginning uppercase letter sequences.
音位字母将转写为意大利体小写字母，书写单位将转写为罗马体小型大写字母。就无变音符号的转写而言，音位字母转写为全小写字母序列，书写单位转写为开头大写字母序列。

In the phonetic letter table, written forms that are not used in the orthographic shaping process are highlighted in gray, and written forms that are used only historically are highlighted in yellow. Gray numbers after the transcription of written forms indicate the corresponding FVS serial number when called manually, and exclamation marks indicate that the written form can be called only by FVS.
音位字母表中，正字法成形过程中不被使用的书写形式标为灰色，仅历史使用的书写形式标为黄色。书写形式转写后的灰色数字表示在手动调用书写形式时对应的 FVS 序号，感叹号表示该书写形式仅可通过 FVS 调用。

The capital letters enclosed in square brackets in the comments indicate that:
备注中用方括号括住的大写字母分别表示：

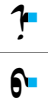
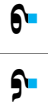
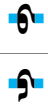
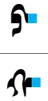
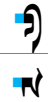
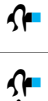
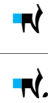
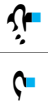
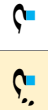
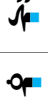
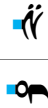
- *Chachlag*, [C]: Found in cases where the written form is separated by an inserted space within a non–affixed word segment.
  分写，[C]：见于非词缀的词段内书写形式分开的情形。
- *Particle*, [P]: Found as an auxiliary or grammatical appositive.
  助词，[P]：见于助词或格附加成分。
- *Lexical*, [L]: Found in special words, such as loanwords.
  词形，[L]：见于借词等特殊词形。
- *Initial*, [I]: Found in initial syllables.
  词首，[I]：见于首音节。
- *Devsger*, [D]: Found in the consonant letter at the end of a Mongolian syllable, or in the vowel letter following a vowel letter.
  韵尾，[D]：见于音节末辅音或元音后元音。
- *Root*, [R]: Found in the second root of a double root word.
  词根，[R]：见于双词根词的第二词根。
- *Feminine*, [F]: Found in feminine words.
  阴性，[F]：见于阴性词。
- *Ligature*, [G]: Found in ligature forms.
  连字，[G]：见于连字形式。
- *Historical*, [H]: Found in early modern orthographies.
  历史，[H]：见于前现代正字法。

3 Hudum writing system 传统蒙古文

3.1 Character set 字符集

3.1.1 Written units 书写单位

Table 3 Hudum written units
表 3 传统蒙古文书写单位

Written unit		Positional forms: .isol, .init, .medi, .fina				Sub–written–unit variants				Comments 备注
A										
Á	Aa									
I										
Į	Ix									Not recommanded.
O										
U										
Ų	Ux									Not recommanded.
Ü	Ue									
N										
B										Unified:  ‹B2 (B2)›.
P										
H										
Ḧ	Hx									
G										
Ğ	Gx									Early modern orthography.
M										
L										
S										
Ŝ	Sz									
Ś	Sh									
T										
D										

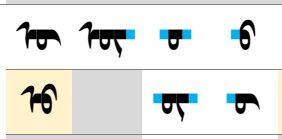
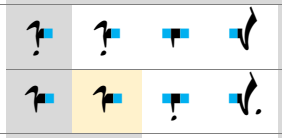
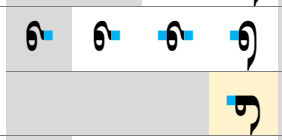
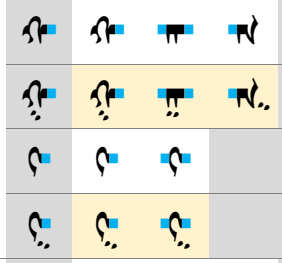
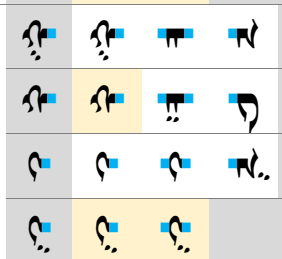
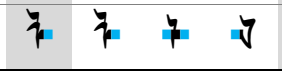
Đ	Dd													
Ć	Ch													
J														
Y														
R													?	
W														
F														
K														Unified: ‹K2 (K2)›.
C														
Z														
Ħ	Hr													
Ř	Rh													
Ž	Zr													
Č	Cr													

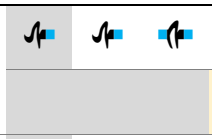
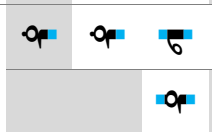
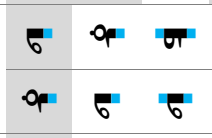
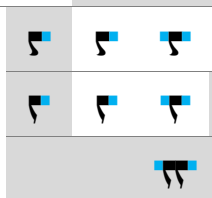
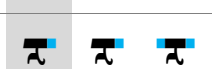
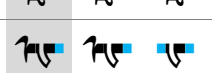
3.1.2 Phonetic letters 音位字母

Table 4 Hudum phonetic letters

表 4 传统蒙古文音位字母

Ph. le. 音位字母		Cp. 码位	Written forms 书写形式								Comments 备注
a		1820					AA 3	AA 2	A	A 2	Á.isol [C]: *tan·a*, TAN Á. A.isol, A.init [P]: *a*, A; *abu–aća*, AABO AĆA. AA.medi [R]: *buyanarbin*, BOYAAAARBIA. Á.fina [G]: *ba*, BÁ.
							A 1!	A 1	AA 1!	Á 1	
							Á 2				
e		1821					A 2	A	A	A 2	Á.isol [C]: *egün·e*, AGON Á. AA.init [H, UM]: *erte*, AARDA. Á.fina [G]: *be*, BÁ.
							Á 1	AA 1!		Á 1	
é	ee	1827					AW	AW	W	W	
i		1822					AI 3	AI 2	I 3	I	I.isol, I.init [P]: *gen–i*, GAA I; *ger–iyer*, GAR IIAR. II.medi [D], AI.medi [R]: *sainirögel*, SAIIAAIROGAL. I.medi [L]: *naima*, NAIMA. I.init [H]: *man·i*, MAN I.
							I 1	I 1	AI 1!		
							Į 2!		II 2		

o		1823		AO	AO	O	U 2	AO.medi [R]: *uranodo*, AORAAAODU. U.fina [C]: *ćino·a*, ĆINU Á. O.fina [L, I, G]: *ḳino*, K2INO; *no*, NO.
						AO 1!	O 1	
u		1824		AO 3	AO 2	O	U 2	U.isol, O.init [P]: *man–u*, MAA U; *mal–un*, MAL OA. U.fina [C]: *liŋhu·a*, LIAGHU Á. AO.medi [R]: *nasunurtu*, NASOAAORDU. O.fina [L, I, G]: *taḳanu*, TAK2ANO; *nu*, NO. Ụ.isol [H]: *dagun·u*, TAḤON U.
				U 1	O 1	AO 1!	O 1	
				Ụ 2!				
ö	oe	1825		AÜ	AOI	O 3	U 3	OI.medi, Ü.fina [I]: *törö*, TOIRU; *lö*, LÜ. OI.medi, AOI.medi [R]: *sainböhe*, SAIIABOIGÁ; *sanagaöljei*, SANAḤAAOILJAI. AU.isol [H]: *ö*, AU. O.fina [G]: *söhö*, SOIGO.
				AU 1!		OI 1	Ü 1	
						AOI 2!	O 2	
ü	ue	1826		AÜ	AOI 2	O 3	U 3	AOI.init, O.init, U.isol [P]: *yabul–ügei*, YABOL AOIHAI; *ger–ün*, GAR OA; *tegün–ü*, TAGOA U. OI.medi, Ü.fina [I]: *nür*, NOIR; *lü*, LÜ. AOI.medi [R]: *ćiŋünen*, ĆAIAGAOINAA. AU.isol, OI.medi [L]: *ü*, AU; *séḳünd*, SWK2OIND. Ụ.isol [H]: *tegün·ü*, TAGON U. O.fina [G]: *hihü*, GIGO.
				AU 1!	O 1	OI 1	Ü 1	
				U 2		AOI 2!	O 2	
				Ụ 3!				
n		1828		N–	N 2	A 2	A 2	A.medi [D]: *tanda*, TAADA. N.medi [L]: *séḳünd*, SWK2OIND. N.fina [C]: *tan·a*, TAN Á. A.init [H]: *nara*, AARA.
				A– 1!	A 1!	N 1	N 1	
ŋ	ng	1829		–AG–	–AG–	AG	AG	
b		182A		B–	B	B	B	B2.fina [H]: *ab*, AAB2.
							B2 1!	
p		182B		P–	P	P	P	
h		182C		H–	H 3	H 3	H	H.fina [C]: *aŋh·a*, AAGH Á. G.init, G.medi [F]: *hei*, GAI; *sehe*, SAGÁ. G.medi [L]: *téhnig*, TWGNIG. Ḧ.init, Ḧ.medi, Ḧ.fina [H]: *hada*, ḤADA; *sahal*, SAḤAL; *bah·a*, BAḤ Á. G̈.init, G̈.medi [H]: *hereg*, G̈ARAG; *sahidag*, SAG̈IDAH.
				Ḧ– 1!	Ḧ 1!	Ḧ 1!	Ḧ 1!	
				G– 2!	G 2	G 2		
				G̈– 4!	G̈ 4!	G̈ 4!		
g		182D		Ḧ–	Ḧ 3	H 3	H 1	H.medi [D]: *agta*, AAHDA. G.init, G.medi, G.fina [F]: *ghir*, GGIR; *üge*, AOIGÁ; *beg*, BAG. G.init [L]: *gram*, GRAM. Ḧ.fina [C]: *bag·a*, BAḤ Á. Ḧ.init, G̈.init, G̈.medi [H]: *goyo*, ḤOYU; *genen*, G̈ANAA; *egće*, AG̈ĆA.
				H– 1!	H 1!	Ḧ 1	G 2	
				G– 2!	G 2	G 2	Ḧ 3	
				G̈– 4!	G̈ 4!	G̈ 4!		
m		182E		M–	M	M	M	
l		182F		L–	L	L	L	

s		1830		S–	S	S	S	Ŝ.fina [H]: *eres*, ARAŜ.
ś	sh	1831		Ś–	Ś 2	Ś 2	Ś	
				S– 1!	S 1	S 1		
t		1832		T–	T	D 2	T	T.medi [L]: *métr*, MWTR.
						T 1		
d		1833		D–	T 2	Đ 2	Đ	Đ.medi [D]: *adgau*, AAĐĦAU. D.init, D.medi, D.fina [L]: *daliy·a*, DALII Á; *adrés*, ADRWS; *séḳünd*, SWK2OIND.
				T– 1!	D 1	D 1	D 1!	
ć	ch	1834		Ć–	Ć	Ć	Ć	
j		1835		I–	I	J	J	I.isol [C]: *bui–j·a*, BOI I Á.
				I 1			I 1!	
y		1836		Y–	Y 2	Y 3	I	I.init, I.medi [P]: *huda–yin*, HODA IIA; *ger–iyar*, GAR IIAR.
				I– 1!	I 1	I 1		
						II 2!		
r		1837		R–	R	R	R	
w		1838		W–	W	W	W	U.fina [C]: *ow·a*, AOU Á.
						O 1!	U 1	
f		1839		F–	F	F	F	
ḳ	k2	183A		K2–	K2	K2	K2	
k		183B		K–	K	K	K	
c		183C		C–	C	C	C	
z		183D		Z–	Z	Z	Z	
ħ	hh	183E		AH–	AH	H	H	
ř	rh	183F		Ř–	Ř	Ř	Ř	
ɬ	lh	1840		LH–	LH	LH	–LH–	
ž	zr	1841		Ž–	Ž	Ž–	Ž–	
č	cr	1842		Č–	Č	Č–	Č–	

Isolated ɪ̣ and isolated ᴜ̣. Isolated ɪ̣ and isolated ᴜ̣ are used historically, with examples shown in [Dorji *et al.*, 2017]. Considering the examples *man·i*, *balgasun·u*, *man·u*, *erten·ü* and *tegün·ü*, as well as the modern *gadan·a*, *hulugan·a*, *bühün·e* and *emün·e*, the dot in ɪ̣ and ᴜ̣ should be consistently analyzed as belonging to *n* rather than *i*, *u* or *ü*. Therefore, these written forms are marked in gray.
非连形 ɪ̣ 与非连形 ᴜ̣。非连形 ɪ̣ 与非连形 ᴜ̣ 用于历史文献，示例参见文献 [Dorji *et al.*, 2017]。考察文献中的示例 *man·i*、*balgasun·u*、*man·u*、*erten·ü* 与 *tegün·ü*，同时考虑现代文献中的 *gadan·a*、*hulugan·a*、*bühün·e* 与 *emün·e*，ɪ̣ 与 ᴜ̣ 中的点应被一致地分析为属于字母 *n*，而非属于字母 *i*、*u* 或 *ü*。因此这些书写形式被标灰。

Final ɪ in letter *j*. Final ɪ in letter *j* is introduced in [Dorji *et al.*, 2017], where the written form of *j* in the example *bui–j·a* is not final, but isolated. Since there are currently no examples exist to prove its existence, the written form is marked in gray, but the shaping process still supports the theoretical existence of a word with final ɪ in letter *j* and chachlag.
字母 *j* 的前连形 ɪ。字母 *j* 的前连形 ɪ 在 [Dorji *et al.*, 2017] 中引入，其中示例 *bui–j·a* 内的 *j* 并非前连形，而是非连形。由于暂无示例证明字母 *j* 的前连形 ɪ 的存在性，该书写形式被标灰。但变形过程仍然支持理论上存在的含有字母 *j* 的前连形 ɪ 与分写左撇的词汇。

Medial ɪɪ in letter *y* and medial ᴏ in letter *w*. Medial ɪɪ in letter *y* and medial ᴏ in letter *w* are introduced in [Dorji *et al.*, 2017], where the written form ɪɪ of letter *y* in the examples *ayl* and *aymag*, and the written form ᴏ of letter *w* in the example *tawlay* are all differences derived from letter analysis. Since *ail*, *aimag* and *taulai* are preferred, these written forms are marked in gray.
字母 *y* 的双连形 ɪɪ 与字母 *w* 的双连形 ᴏ。字母 *y* 的双连形 ɪɪ 与字母 *w* 的双连形 ᴏ 在 [Dorji *et al.*, 2017] 中引入，其中示例 *ayl* 与 *aymag* 中 *y* 的 ɪɪ 以及示例 *tawlay* 中 *w* 的 ᴏ 均源于字母分析差异。由于本手册倾向于分析成 *ail*、*aimag* 与 *taulai*，这些书写形式被标灰。

Letter *ḳ* (ᴋ2) and letter *k* (ᴋ). Letter *ḳ* in form ᴋ2 is used to represent /kʰ/ (Latin: *k*) in Hudum literature in China, while letter *k* in form ᴋ is used to represent /kʰ/ (Cyrillic: *к*) in Hudum literature in Mongolian. The reason for disunifying the two letters is that Hudum Aligali uses ᴋ2 to record /g/ and uses ᴋ to record /kʰ/.
字母 *ḳ* (ᴋ2) 与 *k* (ᴋ)。字母 *ḳ* (ᴋ2) 在中国的传统蒙古文文献中用于记录音位 /kʰ/ (*k*)，字母 *k* (ᴋ) 在蒙古国的传统蒙古文文献中用于记录音位 /kʰ/ (*к*)。区分二者的原因在于传统蒙古文阿礼嘎礼用 ᴋ2 记录音位 /g/，用 ᴋ 记录音位 /kʰ/。

Medial ʜ in letter *ħ* and initial ž in letter *ž*. Medial ʜ in letter *ħ* and initial ž in letter *ž* are similar, so they are unified in proposals such as [SAH, 2017]. The reason for disunifying the two written units is that Hudum Aligali uses ʜ to record /ɦ/ and so on, and uses ž to record /ʈ/.
字母 *ħ* 的双连形 ʜ 与字母 *ž* 的后连形 ž。字母 *ħ* 的双连形 ʜ 与字母 *ž* 的后连形 ž 因形似从而在诸如 [SAH, 2017] 的提案中被认同。区分二者的原因在于传统蒙古文阿礼嘎礼用 ʜ 记录 /ɦ/ 等音位，用 ž 记录音位 /ʈ/。

3.2 Shaping process 变形流程

Table 5 Hudum phonetic letter classes
表 5 传统蒙古文音位字母类

Class	Members
masculine vowel	*a o u*
feminine vowel	*e ö ü é*
neuter vowel	*i*
vowel	{*masculine vowel*} {*feminine vowel*} {*neuter vowel*}
consonant	*n ŋ b p h g m l s ś t d ć j y r w f ķ k c z ħ ř ŧ ž č*

Table 6 Hudum shaping: Mongolian–specific phase
表 6 传统蒙古文成形过程：蒙古文字特有的变形阶段

Shaping step	Letters	Conditions	Lookups
1. Chachlag	*a e*	**if** follows an MVS:	`Chachlag`
2. Syllabic	*o u ö ü*	**if** follows an initial *consonant* or an initial *consonant* cluster:	`Marked`
	d	**if** precedes a final *vowel*:	`Marked`
	n j w	**if** precedes an MVS that precedes an isolated *a* or isolated *e*:	`Chachlag_Onset`
	h g	**if** precedes an MVS that precedes an isolated *a*:	`Chachlag_Onset`
	g	**if** precedes an MVS that precedes an isolated *e*:	`Chachlag_Devsger`
	n t d	**if** precedes a *vowel*:	`Onset`
		else if follows a *vowel*:	`Devsger`
	h g	**if** precedes a *masculine vowel*:	`Masculine_Onset`
		else if precedes a *feminine vowel* or *neuter vowel*:	`Feminine`
		else if follows a *masculine vowel*:	`Masculine_Devsger`
		else if follows a *feminine vowel*:	`Feminine`
		else if remotely follows a *masculine vowel* without a blocking *feminine vowel*:	`Masculine_Devsger`
		else if remotely follows a *feminine vowel* without a blocking *masculine vowel*:	`Feminine`
		else if remotely precedes a *masculine vowel* without a blocking *feminine vowel*:	`Masculine_Devsger`
		else if initial form precedes a *consonant*:	`Feminine`
		else:	`Feminine`
	t	**if** precedes *é* or a *consonant*:	`Devsger`
	ś	**if** precedes *i*:	`Dotless`
	g	**if** follows *s* or *d*:	`Dotless`
3. Particle	*a e i u ü d y*	**if** is in the specific particle dictionary that follows an MVS:	`Particle`
	u ü	**if** is in the specific particle dictionary:	`Particle`

4.	**Devsger**	*i*	**if** follows a *vowel* that does not or will not end with a written unit I:	Devsger
5.	**Post–bowed**	*o u ö ü*	**if** is in the written form of U and follows a (bowed) written unit G/K/B/P/F:	Post_Bowed
		a e	**if** is in the written form of A and follows a (bowed) written unit G/K/B/P/F:	Post_Bowed

Table 7 Hudum shaping: lookups

表 7 传统蒙古文成形过程：条件变形

Ph. le.	Cp.	Lookup	Written forms	Comments
a	1820	Default	AA AA A A	
		Chac	Á	
		DiPa	Á A	
		PoBo	Á	
e	1821	Default	A A A A	
		Chac	Á	
		DiPa	Á	
		PoBo	Á	
é ee	1827	Default	AW AW W W	
i	1822	Default	AI AI I I	
		DiPa	I I	
		Devs	II	
o	1823	Default	AO AO O U	
		Mark	O	
		PoBo	O	
u	1824	Default	AO AO O U	
		Mark	O	
		DiPa	U O U	
		PoBo	O	
ö oe	1825	Default	AÜ AOI O U	
		Mark	OI Ü	
		PoBo	O	

ü	ue	1826	Default					AÜ	AOI	O	U
			Mark							OI	Ü
			DiPa					U	O	O	U
			PoBo								O
n		1828	Default					N–	N	A	A
			ChOn								N
			Onse							N	
			Devs							A	A
ŋ	ng	1829	Default					–AG–	–AG–	AG	AG
b		182A	Default					B–	B	B	B
p		182B	Default					P–	P	P	P
h		182C	Default					H–	H	G	H
			ChOn								H
			MaOn						H	H	
			MaDe							H	H
			Femi						G	G	
g		182D	Default					Ḧ–	Ḧ	G	G
			ChOn								Ḧ
			ChDe								H
			MaOn						Ḧ	Ḧ	
			MaDe							H	H
			Dotl							H	H
			Femi						G	G	G
m		182E	Default					M–	M	M	M
l		182F	Default					L–	L	L	L
s		1830	Default					S–	S	S	S
ś	sh	1831	Default					Ś–	Ś	Ś	Ś
			Dotl						S	S	

`Masculine_Devsger` affects all of the medial forms and final forms (including masculine forms and feminine forms), and `Dotless` only affects masculine forms with dots.

t		1832	Default					T–	T	D	T
			Devs							T	
d		1833	Default					D–	T	Đ	Đ
			Mark						D		
			Onse						T	D	
			Devs							Đ	Đ
			DiPa						D		
ć	ch	1834	Default					Ć–	Ć	Ć	Ć
j		1835	Default					I–	I	J	J
			ChOn					I			I
y		1836	Default					Y–	Y	Y	I
			DiPa						I	I	
r		1837	Default					R–	R	R	R
w		1838	Default					W–	W	W	W
			ChOn								U
f		1839	Default					F–	F	F	F
ķ	k2	183A	Default					K2–	K2	K2	K2
k		183B	Default					K–	K	K	K
c		183C	Default					C–	C	C	C
z		183D	Default					Z–	Z	Z	Z
ħ	hh	183E	Default					AH–	AH	H	H
ř	rh	183F	Default					Ř–	Ř	Ř	Ř
ɬ	lh	1840	Default					LH–	LH	LH	–LH–
ž	zr	1841	Default					Ž–	Ž	Ž–	Ž–
č	cr	1842	Default					Č–	Č	Č–	Č–

Function of the FVS's. In the shaping steps required by MGC/01–01, the appearance of FVS not only determines the presentation form of a Mongolian letter, but also changes the presentation forms of other Mongolian letters. For this reason, it is also necessary to adjust back the unexpected shaping steps caused

by FVS's after each shaping step is completed. These extra steps are not included in Table 6, but are completed in the specific implementation.

FVS 的功能。在 MGC/01–01 指定的变形步骤中，FVS 的出现不仅定下某一字母的书写形式，还会同时改变其他蒙古文字母的书写形式。为此，在每一个变形步骤完成以后，还需要将 FVS 导致的意料之外的变形步骤调整回去。这些额外步骤不列入表 6，但已在具体实现中完成。

3.3 Reference 参考文献

- 全国信息技术标准化技术委员会 / National Information Technology Standardization Technical Committee (NITS) (2022). 信息技术 传统蒙古文名义字符、变形显现字符和控制字符使用规则 (征求意见稿) / *Information technology — Traditional Mongolian nominal characters, presentation characters and use rules of controlling characters*. GB/T 25914—xxxx.
- 内蒙古自治区民族事务委员会 / Ethnic Affairs Commision of Inner Mongolia Autonomous Region (EAC), 内蒙古大学 / Inner Mongolia University (2021). 信息技术 传统蒙古文名义字符到变形显现字符的转换补充规则 / *Information technology — Transferring rules of Traditional Mongolian nominal form to variant form*. MGC/01–01 (version 1.0.5).
- 梁金宝 / LIANG Jinbao (2021). 蒙古文名义字符到变形显现字符转换规则管理系统 (传统蒙古文) (version 1.0.5), rule.mongoltoli.cn.
- 全国信息技术标准化技术委员会 / NITS (2010). 信息技术 传统蒙古文名义字符、变形显现字符和控制字符使用规则 / *Information technology — Traditional Mongolian nominal characters, presentation characters and use rules of controlling characters*. GB/T 25914—2010.
- 确精扎布 / [Hudum] *Ćoijuŋjab* (2013). "传统蒙古文编码". In 确精扎布蒙古文信息处理专辑, 呼和浩特: 内蒙古教育出版社, pp. 276–294.
- 内蒙古大学蒙古学研究院蒙古语文研究所 (1999). 蒙汉词典 (增订本). 呼和浩特: 内蒙古大学出版社.
- Ядамжавын Цэвэл (1966). *Монгол Хэлний Товч Тайлбар Толь*. Улаанбаатар: Улсын Хэвлэлийн Хэрэг Эрхлэх Хороо.
- 梁海 / LIANG Hai (2019). Towards a well–formed Mongolian specification that allows interoperable implementations. WG2 N5104 = L2/19–130 = MWG/3–N8R.

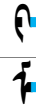
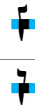
4 Todo writing system 托忒文

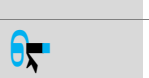
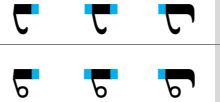
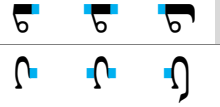
4.1 Character set 字符集

4.1.1 Written units 书写单位

Table 8 Todo written units

表 8 托忒文书写单位

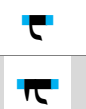
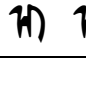
Written unit		Positional forms: .isol, .init, .medi, .fina				Sub–written–unit variants				Comments 备注
A										
Á	Aa									
E										
I										
Î	Ip									
O										
Ó	Op									
Ọ	Ob									
Ọ́	Ot									
U										
Ú	Up									
Ł	Lv									
N										Unified: ‹N2 (N2)›.
B										
P̂	Pp									
Ḧ	Hx									
Ḥ	Hb									
Ĥ	Hp									
G										
K										
M										
L										

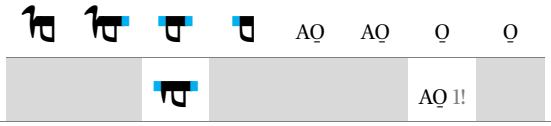
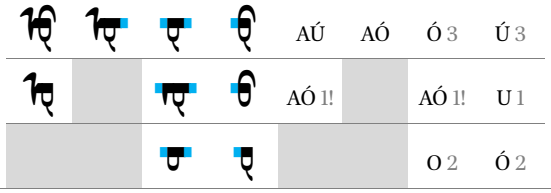
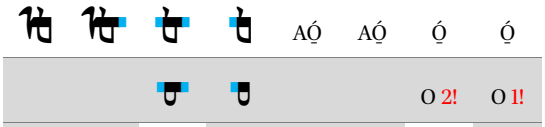
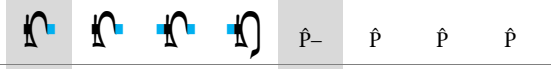
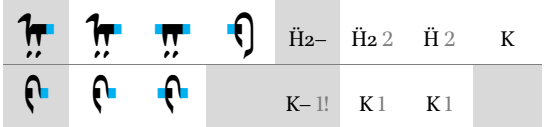
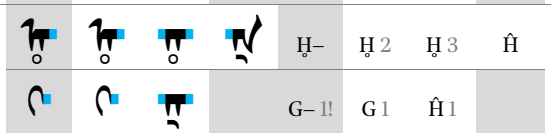
S											
Ś	Sh										
Ṭ	Tp										
Ḍ	Dp										
Ć	Ch										
Ĉ	Cp										
J											
J̌	Jb										
Y											Not recommanded.
R											
W											
Ẉ	Wb										
K̂	Kp										
Ĝ	Gp										
H	Hr										
Ɣ	Yp										
Ñ	Ny										
Ż	Zz										

4.1.2 Phonetic letters 音位字母

Table 9 Todo phonetic letters

表 9 托忒文音位字母

Ph. le. 音位字母	Cp. 码位	Written forms 书写形式								Comments 备注
a	1820					AA 3	AA 2	A	A 2	A.isol [P]: *a*, A. AA.medi [R]: *buyanarbin*, BÓIAAAARBÎA. Á.fina [G]: *ba*, BÁ.
						A !!		AA !!	Á 1	
e	1844					AE	AE	E	E	AE.medi [R].
								AE !!		
i	1845					AI 3	AÎ	Î 2	I 3 2	AÎ.medi [D]: *bain*, BAAÎA. AÎ.medi [R].

						AI 1!	AÎ 1	Î 1			Î.fina [G]: *śabi*, ŚABÎ.
							I 3!				
o		1846				AǪ	AǪ	Ǫ	Ǫ		AǪ.medi [R]: *altanorgil*, AALṬAAAǪRGÎL.
								AǪ 1!			
u		1847				AÚ	AÓ	Ó 3	Ú 3		AÓ.isol [I]: *u*, AÓ. AÓ.medi [R]: *harausun*, ḤARAAÓSÓA. U.fina, O.medi [D]: *tatuu*, ṬAṬÓU; *gaśuun*, ĤAŚÓOA.
						AÓ 1!		AÓ 1!	U 1		
								O 2	Ó 2		
ö	oe	1848				AǪ	AǪ	Ǫ	Ǫ		AǪ.medi [R]: *höhüöndör*, KǪKOAǪAḌǪR.
								O 2!	O 1!		
								AǪ 1!			
ü	ue	1849				AU	AO	O	U 2		AO.isol [I]: *u*, AO. AO.medi [R]: *ćiŋünen*, ĵÎAGAONEA.
						AO 1!		AO 1!	O 1		
ł	lvs	1843				Ł	Ł	Ł	Ł		
n		1828				N–	N 2	A 2	A 2		A.medi [D]: *front*, WRǪAṬ. N2.init [P]: *beljin–ni*, BEŁĈÎA N2I3.
							N2 3	N 1			
ŋ	ng	184A				–AG–	–AG–	AG	AG2		
b		184B				B–	B	B	B2		
p		184C				P̂–	P̂	P̂	P̂		
h		184D				Ḧ2–	Ḧ2 2	Ḧ 2	K		K.init, K.medi [F]: *hicełł*, KÎJEŁL; *yehe*, IEKE.
						K– 1!	K 1	K 1			
g		184E				Ḫ–	Ḫ 2	Ḫ 3	Ĥ		G.init, G.medi [F]: *gerel*, GEREL; *egećі*, AEGEJI3. Ĥ.medi [D]: *cagdał*, JAĤḌAŁ.
						G– 1!	G 1	Ĥ 1			
								G 2			
m		184F				M–	M	M	M2		
l		182F				L–	L	L	L		
s		1830				S–	S	S	S		
ś	sh	1831				Ś–	Ś 2	Ś 2	Ś		
t		1850				Ṭ–	Ṭ	Ṭ	Ṭ		

d		1851					Ḍ–	Ḍ	Ḍ	Ḍ
ć	ch	1852					J̣–	J̣	J̣	J̣
j		1853					Ĉ–	Ĉ	Ĉ	Ĉ
z		1834					Ć–	Ć	Ć	Ć
c		1854					J–	J	J	J
y		1855					I–	I	I	I3
							Y– 1!	Y 1!	Y 1!	
r		1837					R–	R	R	R
w		1856					Ẉ–	Ẉ	Ẉ	Ẉ
									O 1!	U 1!
f		1838					W–	W	W	W
ḳ	kh	1857					K̂–	K̂	K̂	K̂
ġ	gh	1858					Ĝ–	Ĝ	Ĝ	Ĝ
ḥ	hh	1859					AH̱–	AH̱	AH̱	AH̱
j̃	jy	185A					BY̱–	BY̱	BY̱	–BY̱–
ñ	ny	185B					Ñ–	Ñ	Ñ	–Ñ–
ż	zr	185C					Ż–	Ż	Ż	Ż
ɬ	lh	1840					LH̱–	LH̱	LH̱	–LH̱–

Final I3 in letter *y*, medial O and final U in letter *w*. Final I3 in letter *y*, medial O and final U in letter *w* are introduced in GB/T 36649—2018, where the written form I3 of letter *y* and the written form O of letter *w* in the examples *tuwlay*, and the written form U of letter *w* in the example *daruw* are all differences derived from letter analysis. Since *tuulai* and *daruu* are preferred, the written forms are marked in gray.

字母 *y* 的前连形 I3、字母 *w* 的双连形 O 与前连形 U。字母 *y* 的前连形 I3、字母 *w* 的双连形 O 与前连形 U 在 GB/T 36649—2018 中引入，其中示例 *tuwlay* 中 *y* 的 I3 与 *w* 的 O 以及示例 *daruw* 中 *w* 的 U 均源于字母分析差异。由于本手册倾向于分析成 *tuulai* 与 *daruu*，这些书写形式被标灰。

Written form Ẉ in letter *w*. Written form Ẉ in letter *w* is similar to the written form D in Hudum letter *d*, but since the form Ẉ comes from the bent W and thus from beth, and the form D from lamedh, the two are represented differently.

字母 *w* 的书写形式 Ẉ。字母 *w* 的书写形式 Ẉ 与传统蒙古文字母 *d* 的书写形式 D 形似，但因为书写形式 Ẉ 源自弯曲的书写形式 W 进而源自字母 beth，而书写形式 D 源自字母 lamedh，因而

这两个书写形式被区别表示。

Initial ᠠᡳ and medial ᡳ in letter *i*, medial ᠣ and final ᠣ in letter *ö*, written form ᡭ in letter *y*. These written forms are introduced in GB/T 36649—2018, showing the cognition between Todo *e* and Hudum *é*, Todo *i* and Hudum *i*, Todo *ö* and Hudum *ö*, Todo *y* and Hudum *y*. Since cognition is not considered a basis for establishing a written form system, the letters in the running text are not analyzed for these forms with cognition associations, and therefore these written forms are marked in gray.

字母 *i* 的后连形 ᠠᡳ 与双连形 ᡳ、字母 *ö* 的双连形 ᠣ 与前连形 ᠣ、字母 *y* 的书写形式 ᡭ。这些书写形式在 GB/T 36649—2018 中引入，指示托忒文字母 *i* 与传统蒙古文字母 *i*、托忒文字母 *ö* 与传统蒙古文字母 *ö*、托忒文字母 *y* 与传统蒙古文字母 *y* 之间的理据关联。由于理据关联不被视作建立书写形式系统的基础，实际行文中的字母不应被分析为这些具有理据关联的形式，进而这些形式被标灰。

4.2 Shaping process 变形流程

Table 10 Todo phonetic letter classes

表 10 托忒文音位字母类

Class	Members
masculine vowel	*a o u*
feminine vowel	*e ö ü*
neuter vowel	*i*
vowel	{*masculine vowel*} {*feminine vowel*} {*neuter vowel*}
consonant	*n ŋ b p h g m l s ś t d ć j z c y r w f ķ ġ ħ ǰ ñ ż ł*

Table 11 Todo shaping: Mongolian–specific phase

表 11 托忒文成形过程：蒙古文字特有的变形阶段

Shaping step	Letters	Conditions	Lookups
2. Syllabic	*n*	**if** precedes a *vowel*:	`Onset`
		else if follows a *vowel*:	`Devsger`
	h g	**if** precedes a *masculine vowel*:	`Masculine_Onset`
		else if precedes a *feminine vowel* or *neuter vowel*:	`Feminine`
	g	**else if** follows a *vowel*:	`Masculine_Devsger`
3. Particle	*n*	**if** follows an MVS:	`Particle`
4. Devsger	*i*	**if** follows a *vowel*:	`Devsger`
	u	**if** follows *u*:	`Devsger`
5. Post–bowed	*a i u ü*	**if** follows a (bowed) written unit B/P̂/G/K/Ĝ/K̂:	`Post_Bowed`

Table 12 Todo shaping: lookups

表 12 托忒文成形过程：条件变形

Ph. le.		Cp.	Lookup	Written forms								Comments
a		1820	Default					AA	AA	A	A	
			PoBo								Á	
e		1844	Default					AE	AE	E	E	
i		1845	Default					AI3	AÎ	Î	I3	
			Devs							AÎ		
			PoBo								Î	
o		1846	Default					AǪ	AǪ	Ǫ	Ǫ	
u		1847	Default					AÚ	AÓ	Ó	Ú	
			Devs							O	U	
			PoBo								Ó	
ö	oe	1848	Default					AǪ́	AǪ́	Ǫ́	Ǫ́	
ü	ue	1849	Default					AU	AO	O	U	
			PoBo								O	
ɬ	lvs	1843	Default					Ł	Ł	Ł	Ł	
n		1828	Default					N–	N	A	A	
			Onse							N		
			Devs							A		
			Part						N2			
ŋ	ng	184A	Default					–AG–	–AG–	AG	AG2	
b		184B	Default					B–	B	B	B2	
p		184C	Default					P̂–	P̂	P̂	P̂	
h		184D	Default					Ḧ2–	Ḧ2	Ḧ	K	
			MaOn						Ḧ2	Ḧ		
			Femi						K	K		
g		184E	Default					Ḥ–	Ḥ	Ḥ	Ĥ	
			MaOn						Ḥ	Ḥ		

			Femi						G	G	
			MaDe							Ĥ	
m		184F	Default					M–	M	M	M2
l		182F	Default					L–	L	L	L
s		1830	Default					S–	S	S	S
ś	sh	1831	Default					Ś–	Ś	Ś	Ś
t		1850	Default					Ṭ–	Ṭ	Ṭ	Ṭ
d		1851	Default					Ḍ–	Ḍ	Ḍ	Ḍ
ć	ch	1852	Default					J̣–	J̣	J̣	J̣
j		1853	Default					Ĉ–	Ĉ	Ĉ	Ĉ
z		1834	Default					Ć–	Ć	Ć	Ć
c		1854	Default					J–	J	J	J
y		1855	Default					I–	I	I	I3
r		1837	Default					R–	R	R	R
w		1856	Default					Ẉ–	Ẉ	Ẉ	Ẉ
f		1838	Default					W–	W	W	W
ḳ	kh	1857	Default					K̂–	K̂	K̂	K̂
ġ	gh	1858	Default					Ĝ–	Ĝ	Ĝ	Ĝ
ħ	hh	1859	Default					AH––	AH	AH	AH
j̃	jy	185A	Default					BY––	BY	BY	–BY–
ñ	ny	185B	Default					Ñ–	Ñ	Ñ	–Ñ–
ż	zr	185C	Default					Ż–	Ż	Ż	Ż
ɬ	lh	1840	Default					LH––	LH	LH	–LH–

Function of the sign *l*. The sign *l* behaves as a base character in some cases, but its logic in the shaping process behaves as a combining mark. For this purpose, the necessary preprocessing and postprocessing should be introduced: before the Mongolian–specific steps begin, *l* is shaped in mark class, modifying the joining type of the characters before and after *l*; after the Mongolian–specific steps end, *l* that should be represented as mark is shaped to ligature with the neighboring letters. These extra steps are not included in Table 11, but are completed in the specific implementation.

长音号 *t* 的功能。长音号 *t* 在一些情况下表现为基字，但其在成形过程中的逻辑表现为结合记号。为此须要引入必要的预处理与后处理过程：在蒙古文字特有的成形步骤开始之前，将 *t* 变形为记号类，修改 *t* 前后的字符的连写点位；在蒙古文字特有的成形步骤结束之后，将应当表现为记号的 *t* 与邻近字母变换为连字形式。这些额外步骤不列入表 11，但已在具体实现中完成。

4.3 Reference　参考文献

- 全国信息技术标准化技术委员会 / NITS (2018). 信息技术 托忒文名义字符、变形显现字符和控制字符使用规则 / *Information technology — Todo nominal characters, presentation characters and use rules of controlling characters*. GB/T 36649—2018.
- 梁金宝 / LIANG Jinbao (2020). 蒙古文名义字符到变形显现字符转换规则管理系统 (托忒文) (version 0.1.0), rule.mongoltoli.cn.
- 确精扎布 / [Hudum] *Ćoijuŋjab* (2013). “托忒文编码”. In 确精扎布蒙古文信息处理专辑, 呼和浩特: 内蒙古教育出版社, pp. 295–309.
- 确精扎布 / [Hudum] *Ćoijuŋjab*, 纳 · 格日勒图 (1998). 卫拉特方言词汇. 呼和浩特: 内蒙古大学出版社.
- 额尔德尼, 格日勒玛 (1988). 蒙古语常用词汇. 乌鲁木齐: 新疆人民出版社.
- 乌恩奇, 齐 · 艾仁才 (2005). 四体卫拉特方言鉴. 乌鲁木齐: 新疆人民出版社.
- 塞尔格 / [Hudum] *Sergei* (2016). 托忒文及其历史演变研究 / *Research on the Todo script and its development of history*. Doctoral dissertation, 内蒙古大学 / Inner Mongolia University.
- 明再 / [Todo] *Miŋżei* (2019). 托忒编码与字体构想 / *Тод бичгийн баримжаа кодын тухай санууламж*. L2/19–136.

5 Sibe writing system 锡伯文

5.1 Character set 字符集

5.1.1 Written units 书写单位

Table 13 Sibe written units

表 13 锡伯文书写单位

Written unit		Positional forms: .isol, .init, .medi, .fina				Sub–written–unit variants				Comments 备注
A										
Á	Aa									
Ȧ	Ah									
Å	At									
Ă	Ai									
I										Unified:  ‹I2 (I2)›.
Ĭ	Ic									
O										
Ȯ	Oh									
U										
Ů	Uh									
Ü	Ue									
N										
N̈	Nx									
H										
Ḣ	Hh									
H̊	Hc									
G										Unified:  ‹G3 (G3)›.
Ġ	Gh									
G̊	Gc									
G̈	Gx									Never used.
B										

P	Pb									
S										
Ŝ	Sp									
Ç	Cs									
Ẓ	Zs									
T										
Ṫ	Th									
Ṯ	Tb									
Ṯ̇	Tt									
D										
Ḋ	Dh									
Ḍ	Db									
Ḍ̇	Dt									
Đ	Dd									
L										
M										
Ć	Ch									
Ć̊	Cc									
J										
Y										
R										
V										
W										
K2										
Ḱ	Kh									
Ḱ̊	Kc									
R̂	Rr									

5.1.2 Phonetic letters 音位字母

Table 14 Sibe phonetic letters
表 14 锡伯文音位字母

Ph. le. 音位字母	Cp. 码位	Written forms 书写形式								Comments 备注
’	sbm	1807					–A–	–A–	A	–A–	A.medi: *ju’i*, IȮAI.
a		1820					AA 3	AA 2	A	A 2	Á.fina [G]: *amba*, AAMBÁ.
										Á 1	
e		185D					A	A	Å 2	Å 4	A.medi, A.fina [F]: *dergi*, ṬARĠI; *erde*, ARḊA; *gege*, ĠAĠÁ. Å.fina, Á.fina [G]: *dube*, ṬOBÅ; *gege*, ĠAĠÁ.
									A 1	A 1	
										Å 2	
										Á 3	
i		185E					AI 2	AI	I 3	I 3	I.isol [P]: *boo–i*, BOU I. AI.medi [D]: *weile*, WEAILE. I2.isol [P]: *piyazi*, ṖIYAẐI2.
							I 1		AI 2	I2 1	
										AI 2!	
ï	ii	185F					–Ă–	–Ă–	Ă	Ă	
o		1823					AO	AO	O	U 2	O.fina [I, G]: *so*, SO; *kurbo*, GORBO.
										O 1	
u		1860					AȮ	AȮ	Ȯ 2	Ů 4	O.medi, U.fina [F]: *tugi*, ṬOĠI; *albatu*, AALBAḌU. Ȯ.fina [I, G]: *su*, SȮ; *kurbu*, GORBȮ. O.fina [F, I, G]: *tu*, ṬO; *niŋgu*, NIAGĠO.
									O 1	U 1	
										Ȯ 2	
										O 3	
ü	ue	1861					AÜ	AOI	OI	Ü	
n		1828					N–	N 2	A 2	A 2	A.medi [D]: *anda*, AAAḊA. N2.fina [P]: *en*, AN2.
									N 1	N2 1!	
ŋ	ng	1862					–AG–	–AG–	AG	AG3	
k		1863					H–	H 3	H 3	Ñ	G.init, G.medi [F]: *keb*, GAB; *ikiri*, AIGIRI. Ñ.medi [D]: *akdan*, AAÑḊAA.
							G– 1!	G 1	Ñ 1	G4 1!	
							Ğ– 2!	Ğ 2!	G 2	Ğ 2!	

									Ġ 4!		
g		1864					Ḣ–	Ḣ 2	Ḣ 2	–Ḣ–	Ġ.init, Ġ.medi [F]: *gege*, ĠAĠÁ.
							Ġ– 1!	Ġ 1	Ġ 1		
h		1865					Ḣ–	Ḣ 2	Ḣ 2	–Ḣ–	Ġ.init, Ġ.medi [F]: *hehe*, ĠAĠÁ.
							Ġ– 1!	Ġ 1	Ġ 1		
b		182A					B–	B	B	B	
p		1866					Ṗ–	Ṗ	Ṗ	–Ṗ–	
s		1830					S–	S	S	S	
ś	sh	1867					Ŝ–	Ŝ	Ŝ	Ŝ	
t		1868					Ṯ–	Ṯ 2	Ḏ 3	Đ	Ṯ.init, Ḏ.medi [F]: *tetele*, ṮAḎALÀ. Đ.medi [D]: *tuttu*, ṮOĐḎU.
							T– 1!	T 1	D 1		
									Đ 2		
d		1869					Ṫ–	Ṫ 2	Ḋ 2	–Ḋ–	Ṫ.init, Ḋ.medi [F]: *dedun*, ṪAḎOA.
							Ṯ– 1!	Ṫ 1	Ḋ 1		
l		182F					L–	L	L	L	
m		182E					M–	M	M	M	
ć	ch	1834					Ć–	Ć	Ć	–Ć–	
j		186A					I–	I	J2	–J2–	
y		1836					Y–	Y 2	Y 3	–Y–	
r		1837					R–	R	R	R	
f		186B					V–	V	V	–V–	
w		1838					W–	W	W	W	
ķ	kh	183A					K2–	K2	K2	–K2–	
ġ	gh	186C					Ḱ–	Ḱ	Ḱ	–Ḱ–	
ħ	hh	186D					Ḵ–	Ḵ	Ḵ	–Ḵ–	
c		186E					Ç–	Ç	Ç	–Ç–	
z		186F					Ẓ–	Ẓ	Ẓ	Ẓ	
ř	rh	1870					R̂–	R̂	R̂	–R̂–	

č	cr	1871					Č–	Č	Č	–Č–
ž	zr	1872					Ǐ–	Ǐ	Ǐ	–Ǐ–

Final AI in letter *i*. Final AI in letter *i* is introduced in GB/T 36641—2018, where the written form AI of letter *i* in the examples *jui* is the difference derived from letter analysis. Since *ju'i* is preferred, the written form is marked in gray.
字母 *i* 的前连形 AI。字母 *i* 的前连形 AI 在 GB/T 36641—2018 中引入，其中示例 *jui* 中 *i* 的 AI 源于字母分析差异。由于本手册倾向于分析成 *ju'i*，该书写形式被标灰。

Letter *j*. Since the establishment of the Sibe–Solon Cultural Association in 1947, the Association has made modifications to the Sibe orthography. In the scheme given by the Association, both the initial and medial forms of the letter *j* use the written form I. However, in the subsequent practice of using the Sibe writing system, it was found that this modification would cause confusion of words. Therefore, in all kinds of recently published books, newspapers and dictionaries, the written form J2 is used instead of I.
字母 *j*。自 1947 年锡伯索伦文化协会成立以来，该协会对锡伯文正字法进行修改，其中字母 *j* 的后连形与双连形都使用书写形式 I。但在此后的锡伯文应用实践中，人们发现这一修改在词汇中引入了混淆。因此在近期出版的书籍、报纸和字典中，书写形式 J2 代替了书写形式 I。

Letter *z*. Unlike Manchu, in the Sibe *zi*, the stroke in *z* evolve into a long tooth, forming a *zi* ligature. In the literature, the *i* part of the Sibe *zi* ligature is considered equivalent to the *i* part of the ligature with bowed letter, which led GB/T 36641—2018 to introduce the upper part of the *zi* ligature that is considered to belong to *z*. However, this part never appears separately in non–ligature places and is therefore not considered as separate written forms of the letter *z*.
字母 *z*。与满文不同，锡伯文音节 *zi* 中 *z* 的笔画逐渐发展成长牙，进而形成 *zi* 连字。一些文献将 *zi* 连字中 *i* 的部分分析为与带圆头辅音的连字中 *i* 的部分等同，这使得 GB/T 36641—2018 引入了锡伯文 *zi* 连字的上半部分。但由于这一部分从未在非连字的地方单独出现，这些形式不被认定为字母 *z* 的额外的书写形式。

5.2 Shaping process 变形流程

Table 15 Sibe phonetic letter classes
表 15 锡伯文音位字母类

Class	Members
masculine vowel	*a o ü*
feminine vowel	*e u*
neuter vowel	*i*
vowel	{*masculine vowel*} {*feminine vowel*} {*neuter vowel*} *ï*
consonant	*' n ŋ k g h b p s ś t d l m ć j y r f w ķ ġ ħ c z ř č ž*

Table 16 Sibe shaping: Mongolian–specific phase

表 16 锡伯文成形过程：蒙古文字特有的变形阶段

Shaping step	Letters	Conditions	Lookups
2. Syllabic	*o u*	**if** follows an initial *consonant*:	Marked
	i	**if** follows *z*:	Marked
	e u	**if** follows *t*/*d*/*k*/*g*/*h*:	Feminine
	n	**if** precedes a *vowel*:	Onset
		else if follows a *vowel*:	Devsger
	k g h	**if** precedes a *masculine vowel*:	Masculine_Onset
		else if precedes a *feminine vowel* or *neuter vowel*:	Feminine
	t d	**if** precedes *a*/*i*/*o*:	Masculine_Onset
		else if precedes *e*/*u*/*ü*:	Feminine
	k t	**else if** follows a *vowel*:	Devsger
3. Particle	*i*	**if** follows an MVS:	Particle
4. Devsger	*i*	**if** follows a *vowel*:	Devsger
5. Post–bowed	*a o*	**if** follows a (bowed) written unit B/P/K2/K̇/K̊:	Post_Bowed
	e i u	**if** follows a (bowed) written unit B/P/G/Ġ/G̊/G̈:	Post_Bowed

Table 17 Sibe shaping: lookups

表 17 锡伯文成形过程：条件变形

Ph. le.	Cp.	Lookup	Written forms								Comments
’	sbm 1807	Default					–A–	–A–	A	–A–	
a	1820	Default					AA	AA	A	A	
		PoBo								Á	
e	185D	Default					A	A	Ȧ	Ȧ	Post_Bowed step keeps the dotted or dotless status.
		Femi							A	A	
		PoBo								Å	
										Á	
i	185E	Default					AI	AI	I	I	
		Mark								I2	
		Part					I				
		Devs							AI		
		PoBo								I2	

ï	ii	185F	Default					–Ă–	–Ă–	Ă	Ă	
o		1823	Default					AO	AO	O	U	
			Mark								O	
			PoBo								O	
u		1860	Default					AȮ	AȮ	Ȯ	Ụ	Feminine step removes the dots in written units. Post_Bowed step keeps the dotted or dotless statuss.
			Mark								Ȯ	
			Femi							O	U	
											O	
			PoBo								Ȯ	
											O	
ü	ue	1861	Default					AÜ	AOI	OI	Ü	
n		1828	Default					N–	N	A	A	
			Onse							N		
			Devs							A		
ŋ	ng	1862	Default					–AG–	–AG–	AG	AG3	
k		1863	Default					H–	H	H	Ñ	
			MaOn						H	H		
			Femi						G	G		
			Devs							Ñ		
g		1864	Default					Ḣ–	Ḣ	Ḣ	–Ḣ–	
			MaOn						Ḣ	Ḣ		
			Femi						Ġ	Ġ		
h		1865	Default					Ḧ–	Ḧ	Ḧ	–Ḧ–	
			MaOn						Ḧ	Ḧ		
			Femi						Ġ̊	Ġ̊		
b		182A	Default					B–	B	B	B	
p		1866	Default					Ṗ–	Ṗ	Ṗ	–Ṗ–	
s		1830	Default					S–	S	S	S	

ś	sh	1867	Default					Ŝ–	Ŝ	Ŝ	Ŝ
t		1868	Default					Ṯ–	Ṯ	Ḍ	Đ
			MaOn						T	D	
			Femi						Ṯ	Ḍ	
			Devs							Đ	
d		1869	Default					Ṫ–	Ṫ	Ḋ	–Ḋ–
			MaOn						Ṫ	Ḋ	
			Femi						Ṯ	Ḍ	
l		182F	Default					L–	L	L	L
m		182E	Default					M–	M	M	M
ć	ch	1834	Default					Ć–	Ć	Ć	–Ć–
j		186A	Default					I–	I	J2	–J2–
y		1836	Default					Y–	Y	Y	–Y–
r		1837	Default					R–	R	R	R
f		186B	Default					V–	V	V	–V–
w		1838	Default					W–	W	W	W
ḳ	kh	183A	Default					K2–	K2	K2	–K2–
ġ	gh	186C	Default					Ġ–	Ġ	Ġ	–Ġ–
ħ	hh	186D	Default					Ḵ–	Ḵ	Ḵ	–Ḵ–
c		186E	Default					Ç–	Ç	Ç	–Ç–
z		186F	Default					Ẓ–	Ẓ	Ẓ	Ẓ
ř	rh	1870	Default					R̂–	R̂	R̂	–R̂–
č	cr	1871	Default					Č–	Č	Č	–Č–
ž	zr	1872	Default					Ǐ–	Ǐ	Ǐ	–Ǐ–

5.3 Reference 参考文献

- 全国信息技术标准化技术委员会 (2018). 信息技术 锡伯文名义字符、变形显现字符和控制字符使用规则 / *Information technology — Sibe nominal characters, presentation characters and use rules of controlling characters*. GB/T 36641—2018.
- 梁金宝 / LIANG Jinbao (2020). 蒙古文名义字符到变形显现字符转换规则管理系统 (托忒文) (version 0.1.0), rule.mongoltoli.cn.
- 确精扎布 / [Hudum] *Ćoijuŋjab* (2013). "锡伯文编码". In 确精扎布蒙古文信息处理专辑, 呼和浩特: 内蒙古教育出版社, pp. 310–320.
- 新疆维吾尔自治区民族语言文字工作委员会 (1992). 现代锡伯文学语言正字法. 乌鲁木齐: 新疆人民出版社.
- 杨震远 (2013). 锡汉教学词典. 乌鲁木齐: 新疆人民出版社.
- 佘吐肯 (2009). 锡伯语语法通论. 乌鲁木齐: 新疆人民出版社.
- 张泰镐 / JANG Taeho (2008). 锡伯语语法研究 / *Sibe Grammar*. 昆明: 云南民族出版社.

6 Manchu writing system 圈点满文

6.1 Character set 字符集

6.1.1 Written units 书写单位

Table 18 Manchu written units

表 18 圈点满文书写单位

Written unit		Positional forms: .isol, .init, .medi, .fina				Sub–written–unit variants				Comments 备注
A										
Á	Aa									
Ȧ	Ah									
Å	At									
Ă	Ai									
I										
Ï	Iy									
Ỉ	Ic									
O										
Ȯ	Oh									
U										
Ů	Uh									
Ü	Ue									
N										
H										
Ḣ	Hh									
Ḥ	Hc									
Ḧ	Hx									
G										Unified: ‹G›.
Ġ	Gh									
Ġ	Gc									
G̈	Gx									Historical literature.

B										
Ṗ	Pb									
S										
Ŝ	Sp									
Ç	Cs									
Ẓ	Zs									
T										
Ṫ	Th									
Ṭ	Tb									
Ṭ̇	Tt									
D										
Ḋ	Dh									
Ḍ	Db									
Ḍ̇	Dt									
Đ	Dd									
L										
M										
Ć	Ch									
Č	Cc									
J										
J̊	Jc									
Y										
R										
V										
W										
K2	K2									
K̇	Kh									
K̊	Kc									

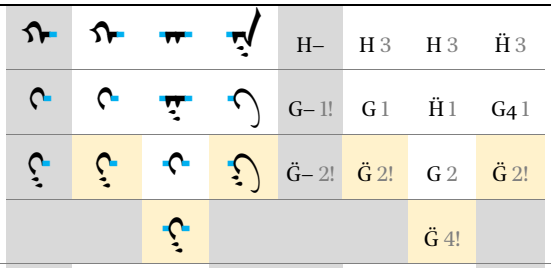
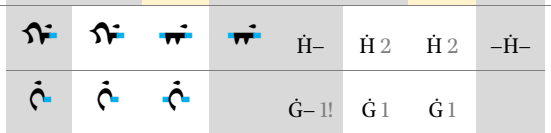
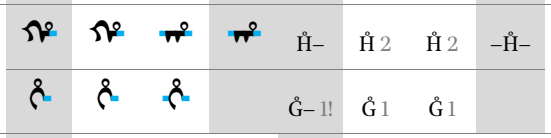
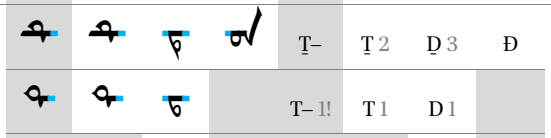
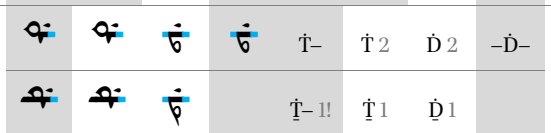
R̂	Rr									

6.1.2 Phonetic letters 音位字母

Table 19 Manchu phonetic letters

表 19 圈点满文音位字母

Ph. le. 音位字母	Cp. 码位	Written forms 书写形式								Comments 备注
’	sbm	1807					–A–	–A–	A	–A–	A.medi: *kui’i*, GOIIAI.
a	1820						AA 3	AA 2	A	A 2	Á.fina [G]: *amba*, AAMBÁ.
										Á 1	
e	185D						A	A	Ȧ 2	Ȧ 4	A.medi, A.fina [F]: *dergi*, ŢARĠI; *erde*, ARḊA; *gege*, ĠAĠÁ. Å.fina, Á.fina [G]: *be*, BÅ; *gege*, ĠAĠÁ.
									A 1	A 1	
										Å 2	
										Á 3	
i	1873						AI 2	AI	I 3	I 3	I.isol [P]: *boo–i*, BOU I. II.medi [D]: *weile*, WEIILE. Ï.fina [P]: *bin zi*, BIA ẔÏ.
							I 1		II 2	I₂ 1!	
										Ï 2	
										AI 4!	
ï	ii	185F					–Ă–	–Ă–	Ă	Ă	
o	1823						AO	AO	O	U 2	O.fina [I, G]: *so*, SO; *olbo*, AOLBO.
										O 1	
u	1860						AȮ	AȮ	Ȯ 2	Ü 4	O.medi, U.fina [F]: *tugi*, ŢOĠI; *albatu*, AALBAḊU. Ȯ.fina [I, G]: *su*, SȮ; *kurbu*, GORBȮ. O.fina [F, I, G]: *tu*, ŢO; *niŋgu*, NIAGĠO.
									O 1	U 1	
										Ȯ 2	
										O 3	
ü	ue	1861					AÜ	AOI	OI	Ü	
n	1828						N–	N 2	A 2	A 2	A.medi [D]: *anda*, AAAḊA. N₂.fina [P, L]: *en*, AN₂; *han*, ĤAN₂.
									N 1	N₂ 1!	
ŋ	ng	1829					–AG–	–AG–	AG	AG	

k		1874		H–	H 3	H 3	Ḧ 3	G.init, G.medi, G4.fina [F]: *keb*, GAB; *enetkek*, ANÀĐGAG4. Ḧ.medi [D]: *akdan*, AAḦḊAA.
				G– 1!	G 1	Ḧ 1	G4 1	
				Ġ– 2!	Ġ 2!	G 2	Ġ 2!	
						Ġ 4!		
g		1864		Ḣ–	Ḣ 2	Ḣ 2	–Ḣ–	Ġ.init, Ġ.medi [F]: *gege*, ĠAĠÁ.
				Ġ– 1!	Ġ 1	Ġ 1		
h		1865		Ḧ–	Ḧ 2	Ḧ 2	–Ḧ–	Ġ.init, Ġ.medi [F]: *hehe*, ĠAĠÁ.
				Ġ– 1!	Ġ 1	Ġ 1		
b		182A		B–	B	B	B	
p		1866		Ṗ–	Ṗ	Ṗ	–Ṗ–	
s		1830		S–	S	S	S	
ś	sh	1867		Ŝ–	Ŝ	Ŝ	Ŝ	
t		1868		Ṯ–	Ṯ 2	Ḍ 3	Đ	Ṯ.init, Ḍ.medi [F]: *tetele*, ṮAḌALÀ. Đ.medi [D]: *tuttu*, ṮOĐḌU.
				T– 1!	T 1	D 1		
						Đ 2		
d		1869		Ṫ–	Ṫ 2	Ḋ 2	–Ḋ–	Ṯ.init, Ḍ.medi [F]: *deo*, ṮAU; *adu*, AAḌU.
				Ṯ– 1!	Ṯ 1	Ḍ 1		
l		182F		L–	L	L	L	
m		182E		M–	M	M	M	
ć	ch	1834		Ć–	Ć	Ć	–Ć–	
j		1835		I–	I	J	–J–	
y		1836		Y–	Y 2	Y 3	–Y–	
r		1875		R–	R	R	R2	
f		1876		V–	V 2	V 2	–V–	W.init, W.medi [D]: *fina*, WINA; *anafu*, AANAWU̇.
				W– 1!	W 1	W 1		
w		1838		W–	W	W	W	
ḳ	kh	183A		K2–	K2	K2	–K2–	
ġ	gh	186C		K̇–	K̇	K̇	–K̇–	

ħ	hh	186D					Ḵ–	Ḵ	Ḵ	–Ḵ–	
c		186E					Ç–	Ç	Ç	–Ç–	
z		186F					Ẓ–	Ẓ	Ẓ	Ẓ	
ř	rh	1870					R̂–	R̂	R̂	–R̂–	
č	cr	1871					Č–	Č	Č	–Č–	
ž	zr	1877					J̌–	J̌	J̌	–J̌–	

Final AI in letter *i*. Final AI in letter *i* is introduced in GB/T 36645—2018, where the written form AI of letter *i* in the examples *jui* is the difference derived from letter analysis. Since *ju'i* is preferred, the written form is marked in gray.

字母 *i* 的前连形 AI。字母 *i* 的前连形 AI 在 GB/T 36645—2018 中引入，其中示例 *jui* 中 *i* 的 AI 源于字母分析差异。由于本手册倾向于分析成 *ju'i*，该书写形式被标灰。

6.2 Shaping process 变形流程

Table 20 Manchu phonetic letter classes

表 20 圈点满文音位字母类

Class	Members
masculine vowel	*a o ü*
feminine vowel	*e u*
neuter vowel	*i*
vowel	{*masculine vowel*} {*feminine vowel*} {*neuter vowel*} *ï*
consonant	*' n ŋ k g h b p s ś t d l m ć j y r f w ķ ġ ħ c z ř č ž*

Table 21 Manchu shaping: Mongolian–specific phase

表 21 圈点满文成形过程：蒙古文字特有的变形阶段

Shaping step	Letters	Conditions	Lookups
2. Syllabic	*o u*	**if** follows an initial *consonant*:	`Marked`
	f	**if** precedes *i*/*o*/*u*/*ü*:	`Marked`
	i	**if** follows *z*:	`Marked`
	e u	**if** follows *t*/*d*/*k*/*g*/*h*:	`Feminine`
	n	**if** precedes a *vowel*:	`Onset`
		else if follows a *vowel*:	`Devsger`
	k g h	**if** precedes a *masculine vowel*:	`Masculine_Onset`
		else if precedes a *feminine vowel* or *neuter vowel*:	`Feminine`
	k	**else if** follows *e* that follows *t*:	`Masculine_Devsger`

			else if follows *u* that follows *k*/*g*/*h*:	Feminine
			else if follows *e*/*ü*:	Feminine
			else:	Masculine_Devsger
		t d	**if** precedes *a*/*i*/*o*:	Masculine_Onset
			else if precedes *e*/*u*/*ü*:	Feminine
		t	**else if** follows a *vowel*:	Devsger
3.	**Particle**	*i*	**if** follows an MVS:	Particle
4.	**Devsger**	*i*	**if** follows a *vowel*:	Devsger
5.	**Post–bowed**	*a o*	**if** follows a (bowed) written unit B/P/K2/K̇/K̈:	Post_Bowed
		e u	**if** follows a (bowed) written unit B/P/G/Ġ/G̈/G̈:	Post_Bowed

Table 22 Manchu shaping: lookups

表 22 圈点满文成形过程：条件变形

Ph. le.		Cp.	Lookup	Written forms								Comments
’	sbm	1807	Default					–A–	–A–	A	–A–	
a		1820	Default					AA	AA	A	A	
			PoBo								Á	
e		185D	Default					A	A	Ȧ	Ȧ	Post_Bowed step keeps the dotted or dotless status.
			Femi							A	A	
			PoBo								Ȧ́	
											Á	
i		1873	Default					AI	AI	I	I	
			Mark								Ï	
			Part					I				
			Devs							II		
ï	ii	185F	Default					–Ă–	–Ă–	Ă	Ă	
o		1823	Default					AO	AO	O	U	
			Mark								O	
			PoBo								O	
u		1860	Default					AȮ	AȮ	Ȯ	Ù	Feminine step removes the dots in written units. Post_Bowed step keeps
			Mark								Ȯ	

			Femi							O	U	the dotted or dotless statuss.
											O	
			PoBo								Ȯ	
											O	
ü	ue	1861	Default					AÜ	AOI	OI	Ü	
n		1828	Default					N–	N	A	A	
			Onse							N		
			Devs							A		
ŋ	ng	1829	Default					–AG–	–AG–	AG	AG	
k		1874	Default					H–	H	H	Ḧ	
			MaOn						H	H		
			Femi						G	G	G4	
			MaDe							Ḧ	Ḧ	
g		1864	Default					Ḣ–	Ḣ	Ḣ	–Ḣ–	
			MaOn						Ḣ	Ḣ		
			Femi						Ġ	Ġ		
h		1865	Default					Ḧ–	Ḧ	Ḧ	–Ḧ–	
			MaOn						Ḧ	Ḧ		
			Femi						Ġ	Ġ		
b		182A	Default					B–	B	B	B	
p		1866	Default					Ṗ–	Ṗ	Ṗ	–Ṗ–	
s		1830	Default					S–	S	S	S	
ś	sh	1867	Default					Ŝ–	Ŝ	Ŝ	Ŝ	
t		1868	Default					T̲–	T̲	D̲	Đ	
			MaOn						T	D		
			Femi						T̲	D̲		
			Devs							Đ		
d		1869	Default					Ṫ–	Ṫ	Ḋ	–Ḋ–	

			MaOn						Ṫ	Ḋ		
			Femi						Ṯ	Ḏ		
l		182F	Default					L–	L	L	L	
m		182E	Default					M–	M	M	M	
ć	ch	1834	Default					Ć–	Ć	Ć	–Ć–	
j		1835	Default					I–	I	J	–J–	
y		1836	Default					Y–	Y	Y	–Y–	
r		1875	Default					R–	R	R	R2	
f		186B	Default					V–	V	V	–V–	
			Mark						W	W		
w		1838	Default					W–	W	W	W	
ķ	kh	183A	Default					K2–	K2	K2	–K2–	
ġ	gh	186C	Default					Ķ–	Ķ	Ķ	–Ķ–	
ħ	hh	186D	Default					Ḱ–	Ḱ	Ḱ	–Ḱ–	
c		186E	Default					Ç–	Ç	Ç	–Ç–	
z		186F	Default					Ẓ–	Ẓ	Ẓ	Ẓ	
ř	rh	1870	Default					R̂–	R̂	R̂	–R̂–	
č	cr	1871	Default					Č–	Č	Č	–Č–	
ž	zr	1877	Default					Ĭ–	Ĭ	J̆	–J̆–	

6.3 Reference 参考文献

- 全国信息技术标准化技术委员会 (2018). 信息技术 满文名义字符、变形显现字符和控制字符使用规则 / *Information technology — Manchu nominal characters, presentation characters and use rules of controlling characters*. GB/T 36645—2018.
- 梁金宝 / LIANG Jinbao (2020). 蒙古文名义字符到变形显现字符转换规则管理系统 (托忒文) (version 0.1.0), rule.mongoltoli.cn.
- 确精扎布 / [Hudum] *Ćoijuŋjab* (2013). "满文编码". In 确精扎布蒙古文信息处理专辑, 呼和浩特: 内蒙古教育出版社, pp. 321–334.
- 安双成 (1993). 满汉大辞典. 沈阳: 辽宁民族出版社.
- 刘景宪, 赵阿平, 赵金纯 (1997). 满语研究通论. 牡丹江: 黑龙江朝鲜民族出版社.
- 黄锡惠 (2012). "满文外音字 *zi* 的词头、词中形式研究" / Research on the initial form and medial form of *zi* in Manchu writing system. 满语研究 / *Manchu Studies*.

7 Hudum Aligali writing system 传统蒙古文阿礼嘎礼

7.1 Character set 字符集

7.1.1 Written units 书写单位

Table 23 Hudum Aligali written units

表 23 传统蒙古文阿礼嘎礼书写单位

Written unit		Positional forms: .isol, .init, .medi, .fina				Sub–written–unit variants				Comments 备注
A										Unified: ‹A2 (A2)›.
Á	Aa									
I										
O										
U										
G										
K										Unified: ‹K2 (K2)›.
N										
Ź	Zc									
C										
Z										
Ñ	Ny									
Ž	Zr									
Č	Cr									
Ḍ	Ds									
Ṿ	Wn									
Đ	Dv									
Ṭ	Tp									
Ď	Dq									
Ǧ	Bg									
Ǩ	Pg									
B										

M									
R									
L									
W									
W̄	Wp								
Ś	Sh								
S̈	Sx								
S									
Ħ	Hr								
J									
Ć	Ch								
Ř	Rh								
Ż	Zz								
Q									
V̀	Vi								

7.1.2 Phonetic letters 音位字母

Table 24 Hudum Aligali phonetic letters
表 24 传统蒙古文阿礼嘎礼音位字母

Ph. le. 音位字母		Cp. 码位	Written forms 书写形式								Comments 备注
a		1820					A 1!	AA 2	A	A 2	Á.isol [C]: *tʰa·a*, ṬÁ Á. A.fina [P]: *cla*, ŹLA.
							Á 2			Á 1	
ą	a2	1887								A2 1!	
i		1888					AI4	AI	I	I4	
u		1826					AU 1!	AOI 2	O 3	U 3	
									OI 1	O 2	
e		1827					AW	AW	W	W	
o		1823					AO	AO	O	U 2	

										O 1
k		1889					G–	G	G	GV̀
kʰ	kh	183B					K–	K	K	K
g		183A					K2–	K2	K2	K2V̀ 1!
ṅ	ng	1829					–AG–	–AG–	AG	AG
ṇ̇	ng2	188A					NG–	NG	NG	NG4
c		188B					Ź–	Ź	Ź	Ź
cʰ	ch	183C					C–	C	C	C
j		183D					Z–	Z	Z	Z
ñ	ny	185B					Ñ–	Ñ	Ñ	Ñ
ṭ	tr	188C					Ž–	Ž	Ž	Ž
ṭʰ	trh	188D					Č2–	Č2	Č2	–Č2–
ḍ	dr	188E					D̰–	D̰	D̰	D̰
ṇ	nr	188F					Ẉ–	Ẉ	Ẉ	Ẉ
t		1890					Ď̃–	Ď̃	Ď̃	Ď̃
tʰ	th	1850					Ṭ–	Ṭ	Ṭ	Ṭ
d		1891					Ď–	Ď	Ď	Ď
n		1828					N–	N 2	N 1	A 2
p		1892					B̌–	B̌	B̌	B̌
pʰ	ph	1893					P̌–	P̌	P̌	P̌
b		182A					B–	B	B	B2 1!
m		184F					M–	M	M	M2
y		1855					I–	I	I	I
r		1837					R–	R	R	R
l		182F					L–	L	L	L
v		1838					W–	W	W	W
ү	wa	18A6					W̄–	W̄	W̄	–W̄–
ś	sy	1831					Ś–	Ś 2	Ś 2	Ś

ṣ	sh	1894					Ṡ–	Ṡ	Ṡ	Ṡ
s		1830					S–	S	S	S
h		1859					AH–	AH	AH 2	H
									H 1	
č	cr	1854					J–	J	J	J
čʰ	crh	1834					Ć–	Ć	Ć	Ć
ž	zr	1895					Ř2–	Ř2	Ř2	Ř2
z		1896					Ż–	Ż	Ż	Ż
ʔ	q	1897					Q–	Q	Q	–Q–
ɬ	lh	1840					LH–	LH	LH	–LH–

7.2 Shaping process 变形流程

Table 25 Hudum Aligali phonetic letter classes

表 25 传统蒙古文阿礼嘎礼音位字母类

Class	Members
vowel	*a ạ i u e o*
consonant	*k kʰ g ṅ ṇ c cʰ j ñ ṭ ṭʰ ḍ ṇ t tʰ d n p pʰ b m y r l v ɣ ś ṣ s h č čʰ ž z ʔ ɬ*

Table 26 Hudum Aligali shaping: Mongolian–specific phase

表 26 传统蒙古文阿礼嘎礼成形过程：蒙古文字特有的变形阶段

Shaping step	Letters	Conditions	Lookups
1. **Chachlag**	*a*	**if** follows an MVS:	`Chachlag`
2. **Syllabic**	*u*	**if** follows an initial *consonant* or an initial *consonant* cluster:	`Marked`
	h	**if** follows *g*/*j*/*ḍ*/*d*/*b*:	`Initial`
	a	**if** precedes an MVS that precedes an isolated *a*:	`Chachlag_Onset`
5. **Post–bowed**	*ạ*	**if** follows *w*:	`Post_W`
	a o u	**if** follows a (bowed) written unit G/K/K2/Ǩ/Ǩ/B:	`Post_Bowed`

Table 27 Hudum Aligali shaping: lookups

表 27 传统蒙古文阿礼嘎礼成形过程：条件变形

Ph. le.	Cp.	Lookup	Written forms								Comments
a	1820	Default					A	AA	A	A	
		Chac					Á				
		PoBo								Á	
ạ	1887	Default								A2	
		PosW								A2	
i	1888	Default					AI4	AI	I	I4	
u	1826	Default					AU	AOI	O	U	
		Mark							OI		
		PoBo								O	
e	1827	Default					AW	AW	W	W	
o	1823	Default					AO	AO	O	U	
		PoBo								O	
k	1889	Default					G–	G	G	GV̀	
kʰ	183B	Default					K–	K	K	K	
g	183A	Default					K2–	K2	K2	K2V̀	
ṅ	1829	Default					–AG–	–AG–	AG	AG	
ṇ	188A	Default					NG–	NG	NG	NG4	
c	188B	Default					Ź–	Ź	Ź	Ź	
cʰ	183C	Default					C–	C	C	C	
j	183D	Default					Z–	Z	Z	Z	
ñ	185B	Default					Ñ–	Ñ	Ñ	Ñ	
ṭ	188C	Default					Ž–	Ž	Ž	Ž	
ṭʰ	188D	Default					Č2–	Č2	Č2	–Č2–	
ḍ	188E	Default					Ḍ–	Ḍ	Ḍ	Ḍ	
ṇ	188F	Default					Ṇ–	Ṇ	Ṇ	Ṇ	
t	1890	Default					Ď–	Ď	Ď	Ď	

tʰ	th	1850	Default					Ṭ–	Ṭ	Ṭ	Ṭ
d		1891	Default					Ď–	Ď	Ď	Ď
n		1828	Default					N–	N	N	A
p		1892	Default					Ǎ–	Ǎ	Ǎ	Ǎ
pʰ	ph	1893	Default					Ǐ–	Ǐ	Ǐ	Ǐ
b		182A	Default					B–	B	B	B
m		184F	Default					M–	M	M	M2
y		1855	Default					I–	I	I	I
r		1837	Default					R–	R	R	R
l		182F	Default					L–	L	L	L
v		1838	Default					W–	W	W	W
ṿ	w	18A6	Default					W̄–	W̄	W̄	–W̄–
ś	sy	1831	Default					Ś–	Ś	Ś	Ś
ṣ	sh	1894	Default					S̈–	S̈	S̈	S̈
s		1830	Default					S–	S	S	S
h		1859	Default					Ħ–	Ħ	AH	Ħ
			Init							Ħ	
č	cr	1854	Default					J–	J	J	J
čʰ	crh	1834	Default					Ć–	Ć	Ć	Ć
ž	zr	1895	Default					Ř2–	Ř2	Ř2	Ř2
z		1896	Default					Ż–	Ż	Ż	Ż
ʔ	q	1897	Default					Q–	Q	Q	–Q–
ɬ	lh	1840	Default					LH–	LH	LH	–LH–

7.3 Reference 参考文献

- (1929). 同文韻統. 上海涵芬楼影印清乾隆内府刻本.
- (1929). 漢滿蒙藏四體合璧阿禮嘎禮讀咒法. 上海涵芬楼影印清乾隆内府刻本.
- (1929). 漢滿蒙藏四體合璧大藏全咒. 上海涵芬楼影印清乾隆内府刻本.
- 乌日晗 (2019). 康熙版蒙古文《甘珠尔》字体复刻开发研究 / *Study on the revival font development of the Kangxi version of Mongolian Kanjur*. Master thesis, 内蒙古农业大学 / Inner Mongolia Agricultural University.
- Weizhe ZHENG, Yihua WANG (2017). *Glyph corrections for Ali Gali final forms in the Mongolian block*. L2/17–115R.
- Дьердь КАРА (1972). *Книги Монгольских Кочевников*. Москва: Издательства «Наука».

8 Manchu Aligali writing system 满文阿礼嘎礼

8.1 Character set 字符集

8.1.1 Written units 书写单位

Table 28 Manchu Aligali written units

表 28 满文阿礼嘎礼书写单位

Written unit		Positional forms: .isol, .init, .medi, .fina				Sub–written–unit variants				Comments 备注
A										Unified: ‹A2 (A2)›.
Á	Aa									
Ā	Al									
Ȧ	Ah									
Å	At									
Ă	Ai									
I										
Ï	Iy									
Ỉ	Ic									
Ï	Ih									
O										
Ȯ	Oh									
U										
Ů	Uh									
Ü	Ue									
K2	K2									
Ḱ	Kh									
Ḣ	Hh									Unified: ‹Ḣ2 (Hh2)›.
H̊	Hc									
H̃	Hy									Unified: ‹H̃2 (Hy2)›.
Ḧ	Hx									
G										Unified: ‹G›.

Ġ	Gh										
Ġ̊	Gc										
Ṇ̊	Nb										
N											
Ç	Cs										
Ç̊	Ct										
Ẓ	Zs										
Ẓ̊	Zt										
ɪ	Iq										
ɪ̊	It										
J											
J̈	Jq										
J̊	Jt										
J̊̈	Jc										
J̈̈	Jh										
Ć	Ch										
Ć̊	Cc										
Ẉ	Wn										
W											
T											
Ṫ	Th										
T̈	Tx										
T̰	Tb										
Ṫ̰	Tt										
T̰̈	Ts										
Ḍ	Dr										
Ḍ̈	Dy										
D											

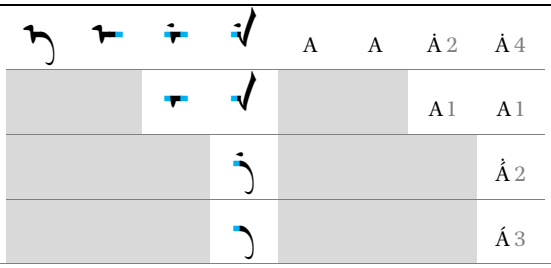
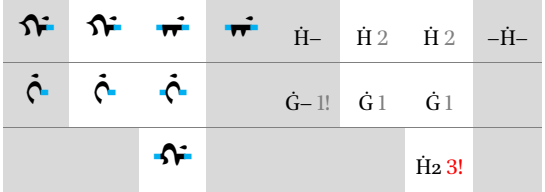
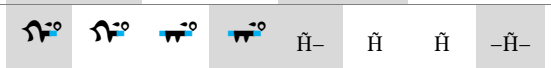
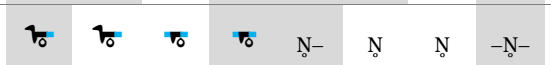
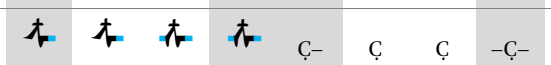
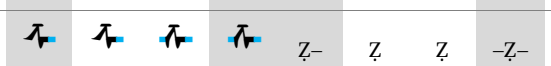
Ḋ	Dh									
D̈	Dx									
Ḍ	Db									
Ḍ̇	Dt									
Ḍ̈	Ds									
Đ	Dd									
B										
B̌	Bg									
B̊	Bc									
Ṗ	Pb									Unified:  ‹Ṗ2 (Pb2)›.
M										
Y										
R										Unified: ‹R3 (R3)›.
L										Unified:  ‹L2 (L2)›.
Ĺ	Lc									
S										Unified: ‹S3 (S3)›.
Ś	Sc									
Ŝ	Sp									
Ś̊	St									
S̈	Sx									

8.1.2 Phonetic letters 音位字母

Table 29 Manchu Aligali phonetic letters

表 29 满文阿礼嘎礼音位字母

Ph. le. 音位字母		Cp. 码位	Written forms 书写形式								Comments 备注
a		1820					AA 3	AA 2	A	A 2	Á.fina [G]: *ba*, BÁ.
										Á 1	
ā	aa	1887					AĀ	AA	A	Ā	

e		185D					A	A	Ȧ 2	Ȧ 4	A.medi, A.fina [F]: *tʰei*, TAI; *tʰe*, TA. Å.fina, Á.fina [G]: *be*, BÅ; *ḳʰe*, GÁ.
									A 1	A 1	
										Å 2	
										Á 3	
i		1873					AI 2	AI	I 3	I 3	Ï.fina [P]: *ji*, Ẓ Ï.
									II 2	Ï 2	
ï	ii	185F					–Ă–	–Ă–	Ă	Ă	
o		1823					AO	AO	O	U 2	O.fina [I, G]: *lo*, LO; *bo*, BO.
										O 1	
u		1860					AȮ	AȮ	Ȯ 2	Ü 4	O.medi, U.fina [F]: *tʰuu*, TOU; *niyu*, NIYU. Ȯ.fina [I, G]: *ḍu*, ḎȮ; *bu*, BȮ. O.fina [F, I, G]: *tu*, TO; *kʰu*, GO.
									O 1	U 1	
										Ȯ 2	
										O 3	
ü	ue	1861					AÜ	AOI	OI	Ü	
k		186C					Ḳ–	Ḳ	Ḳ	Ḳ	
kʰ	kh	183A					K2–	K2	K2	–K2–	
g		1864					Ḣ–	Ḣ 2	Ḣ 2	–Ḣ–	Ġ.init, Ġ.medi [F]: *gi*, ĠI; *sgi*, SĠI. Ḣ2.medi [L]: *ʔgeġs*, A2Ḣ2AGS.
							Ġ– 1!	Ġ 1	Ġ 1		
									Ḣ2 3!		
ġ	g2	1874					G– 1!	G 1	Ḧ 1	Ḧ 3	G.init, G.medi, G4.fina [F]: *ġi*, GI; *ʔgeġs*, A2Ḣ2AGS; *seġ*, SAG4.
									G 2	G4 1	
gʰ	gh	189A					H̃–	H̃	H̃	–H̃–	H̃2.medi [L]: *ʔgʰeġs*, A2H̃2AGS.
									H̃2 1!		
ṅ	ng	189B					Ṇ–	Ṇ	Ṇ	–Ṇ–	
ŋ	ng2	1829					–AG–	–AG–	AG	AG	
c		189C					Ç̊–	Ç̊	Ç̊	–Ç̊–	
cʰ	ch	186E					Ç–	Ç	Ç	–Ç–	
j		186F					Ẓ–	Ẓ	Ẓ	–Ẓ–	

jʰ	jh	189D		Ẓ–	Ẓ	Ẓ	–Ẓ–	
ṭ	tr	189E		Ǐ–	Ǐ	J̌	J̌	
ṭʰ	trh	1871		Č–	Č	Č	Č	
ḍ	dr	1877		Ǐ–	Ǐ	J̌	J̌	
ḍʰ	drh	189F		Ï–	Ï	J̈	J̈	
ṇ	nr	188F		�working–	Ṇ	Ṇ	Ṇ	
t		18A0		Ḍ–	Ḍ 2	Ḍ 2	Ḍ	Ḍ.init, Ḍ.medi [F]: *te*, ḌA; *steŋ*, SḌAAG.
				Ḍ– !!	Ḍ 1	Ḍ 1		
tʰ	th	1868		Ṯ–	Ṯ 2	Ḏ 3	Ð	Ṯ.init, Ḏ.medi [F]: *tʰe*, ṮA; *ʔtʰen*, A2ḎAA. Ð.medi, Ð.fina [D]: *otʰburub*, AOÐBÒRÒB; *ʔotʰ*, A2OÐ.
				T– !!	T 1	D 1		
						Ð 2		
d		1869		Ṫ–	Ṫ 2	Ḋ 2	–Ḋ–	Ṯ.init, Ḏ.medi [F]: *de*, ṮA; *rdul*, RḎOL.
				Ṯ– !!	Ṯ 1	Ḏ 1		
dʰ	dh	18A1		Ṯ–	Ṯ 2	Ḏ 2	–Ḏ–	Ṯ.init, Ḏ.medi [F]: *dʰe*, ṮA; *sdʰe*, SḎA.
				Ṯ– !!	Ṯ 1	Ḏ 1		
n		1828		N–	N 2	A 2	A 2	A.medi [D, Uyghur]: *kʰanᵗʰ*, K2AAÐ. N2.fina [D]: *byan*, BYAN2.
						N 1	N2 !!	
p		1892		Ǎ–	Ǎ	Ǎ	Ǎ	
pʰ	ph	1866		Ṗ–	Ṗ	Ṗ	–Ṗ–	
				Ṗ2– !!	Ṗ2 !!	Ṗ2 !!		
b		182A		B–	B	B	B	
bʰ	bh	18A8		Å–	Å	Å	Å	
m		182E		M–	M	M	M	
y		1836		Y–	Y 2	Y 3	–Y–	
r		1875		R–	R	R	R2	R3.init, R3.medi [Vocalic]: *ri*, R3I; *tʰirii*, TIR3II.
				R3– !!	R3 !!	R3 !!		
l		182F		L–	L	L	L	L2.init [Vocalic]: *li*, L2I.
				L2– !!	L2 !!			

v		1838					W–	W	W	W	
ś	sy	1867					Ŝ–	Ŝ	Ŝ	Ŝ	
ṣ	sh	18A2					S̈–	S̈	S̈	S̈	
s		1830					S–	S	S	S	S3.fina [P]: *cʰeṣ*, ÇÀS3.
										S3 2!	
h		1865					H̊–	H̊ 2	H̊ 2	–H̊–	Ġ.init, Ġ.medi [F]: *hi*, ĠI; *shi*, SĠI.
							Ġ– 1!	Ġ 1	Ġ 1		
č	cr	18A3					I–	I	J	J	
čʰ	crh	1834					Ć–	Ć	Ć	Ć	
ž	zr	1835					I–	I	J	J	
ẑ	zh	18A4					Ŝ–	Ŝ	Ŝ	Ŝ	
z		18A5					S̊–	S̊	S̊	S̊	
ʔ	q	1807					A2–	A2	A	–A–	
ɬ	lh	18AA					L̊–	L̊	L̊	–L̊–	

8.2 Shaping process 变形流程

Table 30 Manchu Aligali phonetic letter classes

表 30 满文阿礼嘎礼音位字母类

Class	Members
masculine vowel	*a o ü*
feminine vowel	*e u*
neuter vowel	*i*
vowel	{*masculine vowel*} {*feminine vowel*} {*neuter vowel*} *ā ï*
consonant	*k kʰ g ġ gʰ ṅ ŋ c cʰ j jʰ ṭ ṭʰ ḍ ḍʰ ṇ t tʰ d dʰ n p pʰ b bʰ m y r l v ś ṣ s h č čʰ ž ẑ z ʔ ɬ*

Table 31 Manchu Aligali shaping: Mongolian–specific phase

表 31 满文阿礼嘎礼成形过程：蒙古文字特有的变形阶段

Shaping step	Letters	Conditions	Lookups
2. Syllabic	*o u*	**if** follows an initial *consonant*:	`Marked`
	i	**if** follows *c*/*j*/*jʰ*:	`Marked`
	e u	**if** follows *t*/*tʰ*/*d*/*dʰ*/*g*/*ġ*/*gʰ*/*h*:	`Feminine`
	e	**if** follows *ṅ*:	`Feminine`

		n	**if** precedes a *vowel*:	Onset
			else if follows a *vowel*:	Devsger
		s	**if** follows a *vowel*:	Devsger
		g ġ h	**if** precedes a *masculine vowel* or *ï*:	Masculine_Onset
			else if precedes a *feminine vowel* or *neuter vowel*:	Feminine
		ġ	**else if** follows *e* that follows *t*:	Masculine_Devsger
			else if follows *e*/*ü*:	Feminine
			else:	Masculine_Devsger
		t tʰ d dʰ	**if** precedes *a*/*i*/*o*:	Masculine_Onset
			else if precedes *e*/*u*/*ü*:	Feminine
		tʰ	**else if** follows a *vowel*:	Devsger
4.	**Devsger**	*i*	**if** follows a *vowel*:	Devsger
		u	**If** follows a *vowel*:	Feminine
5.	**Post–bowed**	*a o*	**if** follows a (bowed) written unit Ḱ/K₂/Ǩ/Ṗ/B/B̊:	Post_Bowed
		e i u	**if** follows a (bowed) written unit Ġ/G/G̊/Ḣ/Ḣ₂/Ȟ/H̃₂/Ṇ/Ǩ/Ṗ/B/B̊:	Post_Bowed

Table 32 Manchu Aligali shaping: lookups

表 32 满文阿礼嘎礼成形过程：条件变形

Ph. le.		Cp.	Lookup	Written forms								Comments
a		1820	Default					AA	AA	A	A	
			PoBo								Á	
ā	aa	1887	Default					AĀ	AA	A	Ā	
e		185D	Default					A	A	Ȧ	Ȧ	Post_Bowed step keeps the dotted or dotless status.
			Femi							A	A	
			PoBo								Å̇	
											Á	
i		1873	Default					AI	AI	I	I	
			Mark								Ï	
			Devs							II		
ï	ii	185F	Default					–Ă–	–Ă–	Ă	Ă	
o		1823	Default					AO	AO	O	U	
			Mark								O	
			PoBo								O	

u		1860	Default					AȮ	AȮ	Ȯ	Ụ	Feminine step removes the dots in written units. Post_Bowed step keeps the dotted or dotless statuss.
			Mark								Ȯ	
			Femi							O	U	
											O	
			PoBo								Ȯ	
											O	
ü	ue	1861	Default					AÜ	AOI	OI	Ü	
k		186C	Default					Ḳ–	Ḳ	Ḳ	Ḳ	
kʰ	kh	183A	Default					K2–	K2	K2	–K2–	
g		1864	Default					Ḣ–	Ḣ	Ḣ	–Ḣ–	
			MaOn						Ḣ	Ḣ		
			Femi						Ġ	Ġ		
ġ	g2	1874	Default					H–	G	Ḧ	Ḧ	
			Femi						G	G	G4	
			MaDe							Ḧ	Ḧ	
gʰ	gh	189A	Default					Ḫ–	Ḫ	Ḫ	–Ḫ–	
ṅ	ng	189B	Default					Ṇ–	Ṇ	Ṇ	–Ṇ–	
ŋ	ng2	1829	Default					–AG–	–AG–	AG	AG	
c		189C	Default					Ç–	Ç	Ç	–Ç–	
cʰ	ch	186E	Default					Ç–	Ç	Ç	–Ç–	
j		186F	Default					Ẓ–	Ẓ	Ẓ	–Ẓ–	
jʰ	jh	189D	Default					Ẓ–	Ẓ	Ẓ	–Ẓ–	
ṭ	tr	189E	Default					Ǐ–	Ǐ	J̌	J̌	
ṭʰ	trh	1871	Default					Č–	Č	Č	Č	
ḍ	dr	1877	Default					Ǐ–	Ǐ	J̌	J̌	
ḍʰ	drh	189F	Default					Ï–	Ï	J̈	J̈	
ṇ	nr	188F	Default					Ṿ–	Ṿ	Ṿ	Ṿ	
t		18A0	Default					Ḍ–	Ḍ	Ḍ	Ḍ	

			MaOn							Ḍ	Ḍ	
			Femi							Ḍ	Ḍ	
t^{h}	th	1868	Default					T̤–	T̤	D̤	Đ	
			MaOn						T	D		
			Femi						T̤	D̤		
			Devs							Đ		
d		1869	Default					Ṫ–	Ṫ	Ḍ	–Ḍ–	
			MaOn						Ṫ	Ḍ		
			Femi						Ṯ	Ḏ		
d^{h}	dh	18A1	Default					T̈–	T̈	Ḍ	–Ḍ–	
			MaOn						T̈	Ḍ		
			Femi						Ṯ	Ḏ		
n		1828	Default					N–	N	A	A	
			Onse							N		
			Devs							A		
p		1892	Default					Ǎ–	Ǎ	Ǎ	Ǎ	
p^{h}	ph	1866	Default					Ṗ–	Ṗ	Ṗ	–Ṗ–	
b		182A	Default					B–	B	B	B	
b^{h}	bh	18A8	Default					B̊–	B̊	B̊	B̊	
m		182E	Default					M–	M	M	M	
y		1836	Default					Y–	Y	Y	–Y–	
r		1875	Default					R–	R	R	R2	
l		182F	Default					L–	L	L	L	
v		1838	Default					W–	W	W	W	
ś	sy	1867	Default					Ŝ–	Ŝ	Ŝ	Ŝ	
ṣ	sh	18A2	Default					S̈–	S̈	S̈	S̈	
s		1830	Default					S–	S	S	S	
			Devs								S3	

h		1865	Default					H–	H	H	–H–
			MaOn						H	H	
			Femi						G	G	
č	cr	18A3	Default					I–	I	J	J
čʰ	crh	1834	Default					Ć–	Ć	Ć	Ć
ž	zr	1835	Default					I–	I	J	J
ẑ	zh	18A4	Default					Š–	Š	Š	Š
z		18A5	Default					Š–	Š	Š	Š
ʔ	q	1807	Default					A2–	A2	A	–A–
ɬ	lh	18AA	Default					L–	L	L	–L–

8.3 Reference 参考文献

- (1929). 同文韻統. 上海涵芬楼影印清乾隆内府刻本.
- (1929). 漢滿蒙藏四體合璧阿禮嘎禮讀咒法. 上海涵芬楼影印清乾隆内府刻本.
- (1929). 漢滿蒙藏四體合璧大藏全咒. 上海涵芬楼影印清乾隆内府刻本.
- (1957). 御制五体清文鉴. 北京: 民族出版社.
- 王敌非 / Difei WANG (2011). “满文阿礼嘎礼字母发微” / “Origin of Manchu ālikāli Letters”. In 满语研究 / *Manchu Studies*, No. 2 (General No. 53).
- 郑维喆 / Weizhe ZHENG, 王奕桦 / Yihua WANG (2017). *Glyph corrections for Ali Gali final forms in the Mongolian block*. L2/17–115R.
- 郑维喆 / Weizhe ZHENG (2016). *Preliminary proposal to encode four letters in Todo and Manchu Ali Gali*. L2/16–331.
- 魏安 / Andrew Christopher WEST (2006). *Proposal to encode one Manchu Ali Gali letter*. WG2 N3041.
- China National Body (2006). *Supporting references for WG2 N3041, Manchu Ali Gali*. WG2 N3092.

9 Todo Aligali writing system 托忒文阿礼嘎礼

9.1 Character set 字符集

9.1.1 Written units 书写单位

Table 33 Todo Aligali written units
表 33 托忒文阿礼嘎礼书写单位

Written unit		Positional forms: .isol, .init, .medi, .fina				Sub–written–unit variants				Comments 备注
A										
Á	Aa									
E										
I										
Î	Ip									
O										
Ọ	Ob									
U										
Ł	Lv									
Ḥ	Hb									
Ĥ	Hp									
K										Unified: ‹K2 (K2)›
G										
N										
Ĉ	Cp									
C										
Ź	Zc									
Ñ	Ny									
Ž	Zr									
Č	Cr									
Ḍ	Ds									
Ṿ	Wn									

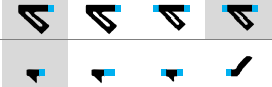
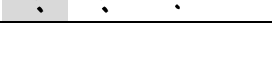
D̃	Dv											
Ṭ	Tp											
Ḍ	Dp											
B́	Bh											
Ṕ	Ph											
B												
M												
Y	Yp											
Ī	Ir											
R												
L												
Ẉ	Wb											
W̄	Wp											
Ś	Sh											
S̈	Sx											
S												
Ħ	Hr											
J												
Ć	Ch											
Ŕ	Rz											
Ż	Zz											
Q												

9.1.2 Phonetic letters 音位字母

Table 34 Todo Aligali phonetic letters

表 34 托忒文阿礼嘎礼音位字母

Ph. le. 音位字母	Cp. 码位	Written forms 书写形式				Comments 备注
a	1820					A 1! AA 2 A A 2

										Á 1
e		1844					AE	AE	E	E
i		1845					AI3	AÎ	Î 2	I3 2
										Î 1
o		1846					AO	AO	O	O
u		1849					AU	AO	O	U 2
							AO 1!			O 1
ɬ	lvs	1843					Ł	Ł	Ł	Ł
k		183A					K2–	K2	K2	–K2–
kʰ	kh	183B					K–	K	K	–K–
g		1889					G–	G	G	–G–
ġ	g2	184E					Ḥ–	Ḥ 2	Ḥ 3	Ĥ
							G– 1!	G 1	Ĥ 1	
								Ĥ 3!	G 2	
ṅ	ng	184A					–AG–	–AG–	AG	AG2
ṇ	ng2	188A					NG–	NG	NG	–NG–
c		1853					Ĉ–	Ĉ	Ĉ	–Ĉ–
cʰ	ch	183C					C–	C	C	–C–
j		188B					Ź– 1!	Ź 1!	Ź 1!	–Ź–
ñ	ny	185B					Ñ–	Ñ	Ñ	–Ñ–
ṭ	tr	188C					Ž–	Ž	Ž	–Ž–
ṭʰ	trh	188D					Č2–	Č2	Č2	–Č2–
ḍ	dr	188E					Ḏ–	Ḏ	Ḏ	–Ḏ–
ṇ	nr	188F					Ẉ–	Ẉ	Ẉ	–Ẉ–
t		1890					Ď–	Ď	Ď	Ď
tʰ	th	1850					Ṭ–	Ṭ	Ṭ	–Ṭ–
d		1851					Ḍ–	Ḍ	Ḍ	–Ḍ–
n		1828					N–	N 2	N 1	A 2

p		1892					Ƀ– 1!	Ƀ 1!	Ƀ 1!	–Ƀ–
pʰ	ph	184C					Ƥ– 1!	Ƥ 1!	Ƥ 1!	–Ƥ–
b		184B					B–	B	B	B2
m		184F					M–	M	M	M2
y		1855					I–	I	I	–I–
ẏ	ya	18A7					–Ɏ–	–Ɏ–	Ɏ	Ɏ
ï	ii	18A7								Ī 1!
r		1837					R–	R	R	R
l		182F					L–	L	L	L
v		1856					Ẉ–	Ẉ	Ẉ	–Ẉ–
ẏ	wa	18A6					–Ẅ	–Ẅ	–Ẅ	Ẅ
ś	sy	1831					Ś–	Ś 2	Ś 2	Ś
ṣ	sh	1894					Š–	Š	Š	–Š–
s		1830					S–	S	S	S
h		1859					AH–	AH	AH 2	–AH–
									Ħ 1	
č	cr	1854					J–	J	J	–J–
ǰ	jr	1834					Ć–	Ć	Ć	–Ć–
ž	zh	1899					Ŕ–	Ŕ	Ŕ	–Ŕ–
z		1896					Ż–	Ż	Ż	–Ż–
ʔ	q	1897					Q–	Q	Q	–Q–
ɬ	lh	1840					LH–	LH	LH	–LH–

9.2 Shaping process 变形流程

Table 35 Todo Aligali phonetic letter classes
表 35 托忒文阿礼嘎礼音位字母类

Class	Members
masculine vowel	*a o*
feminine vowel	*e u*
neuter vowel	*i*
vowel	{*masculine vowel*} {*feminine vowel*} {*neuter vowel*}
consonant	*k kʰ g ṅ c cʰ j ñ ṭ ṭʰ ḍ ṇ ṇ t tʰ d n p pʰ b m y ẏ r l v ṿ ś ṣ s h č ǰ ž z ʔ ɬ*

Table 36 Todo Aligali shaping: Mongolian–specific phase
表 36 托忒文阿礼嘎礼成形过程：蒙古文字特有的变形阶段

Shaping step	Letters	Conditions	Lookups
2. Syllabic	*ġ*	**if** precedes a *masculine vowel*:	Masculine_Onset
		else if precedes a *feminine vowel* or *neuter vowel*:	Feminine
		else if follows a *vowel*:	Masculine_Devsger
	h	**if** follows *g*/*j*/*ḍ*/*d*/*b*:	Initial
5. Post–bowed	*a*	**if** follows an initial *consonant*:	Post_Bowed
	i o u	**if** follows a (bowed) written unit K2/K/G/Ḱ/Ṕ/B:	Post_Bowed

Table 37 Todo Aligali shaping: lookups
表 37 托忒文阿礼嘎礼成形过程：条件变形

Ph. le.	Cp.	Lookup	Written forms								Comments
a	1820	Default					A	AA	A	A	
		PoBo								Á	
e	1844	Default					AE	AE	E	E	
i	1845	Default					AI3	AÎ	Î	I3	
		PoBo								Î	
o	1846	Default					AǪ	AǪ	Ǫ	Ǫ	
u	1849	Default					AU	AO	O	U	
		PoBo								O	
ɬ	lvs 1843	Default					Ł	Ł	Ł	Ł	

k		183A	Default					K2–	K2	K2	–K2–	
kʰ	kh	183B	Default					K–	K	K	–K–	
g		1889	Default					G–	G	G	–G–	
ġ	g2	184E	Default					Ḥ–	Ḥ	Ḥ	Ĥ	
			MaOn						Ḥ	Ḥ		
			Femi						G	G		
			MaDe							Ĥ		
ġ	g2	184E	Default					Ḥ–	Ḥ	Ḥ	Ĥ	
ṅ	ng	188A	Default					NG–	NG	NG	AG2	
c		1853	Default					Ĉ–	Ĉ	Ĉ	–Ĉ–	
cʰ	ch	183C	Default					C–	C	C	–C–	
j		188B	Default					Ź–	Ź	Ź	–Ź–	
ñ	ny	185B	Default					Ñ–	Ñ	Ñ	–Ñ–	
ṭ	tr	188C	Default					Ž–	Ž	Ž	–Ž–	
ṭʰ	trh	188D	Default					Č–	Č	Č	–Č–	
ḍ	dr	188E	Default					Ḍ–	Ḍ	Ḍ	–Ḍ–	
ṇ	nr	188F	Default					Ẉ–	Ẉ	Ẉ	–Ẉ–	
t		1890	Default					D–	D	D	D	
tʰ	th	1850	Default					Ṭ–	Ṭ	Ṭ	–Ṭ–	
d		1851	Default					Ḍ–	Ḍ	Ḍ	–Ḍ–	
n		1828	Default					N–	N	N	A	
p		1892	Default					Ɓ–	Ɓ	Ɓ	–Ɓ–	
pʰ	ph	184C	Default					Ṕ–	Ṕ	Ṕ	–Ṕ–	
b		184B	Default					B–	B	B	B2	
m		184F	Default					M–	M	M	M2	
y		1855	Default					I–	I	I	–I–	
ẏ	ya	18A7	Default					–Y–	–Y–	Y	Y	
r		1837	Default					R–	R	R	R	

l		182F	Default					L–	L	L	L
v		1856	Default					W̥–	W̥	W̥	–W̥–
y	wa	18A6	Default					–W2	–W2	–W2	W2
ś	sy	1831	Default					Ś–	Ś	Ś	Ś
ṣ	sh	1894	Default					S̈–	S̈	S̈	–S̈–
s		1830	Default					S–	S	S	S
h		1859	Default					AH–	AH	AH	–H–
			Init							H	
č	cr	1854	Default					J–	J	J	–J–
ǰ	jr	1834	Default					Ć–	Ć	Ć	–Ć–
ž	zh	1899	Default					Ŕ–	Ŕ	Ŕ	–Ŕ–
z		1896	Default					Ż–	Ż	Ż	–Ż–
ʔ	q	1897	Default					Q–	Q	Q	–Q–
ɬ	lh	1840	Default					LH–	LH	LH	–LH–

9.3 Reference 参考文献

- 叶尔达 (主编) (2017). *伊犁河流域额鲁特人托忒文文献荟萃*, 北京: 中国社会科学出版社.
- 丹碧 / [Hudum] *Dambipéljid*, 格・李杰 / [Hudum] *Ge. Lijei* (2008). 蒙汉对照托忒文字卫拉特蒙古历史文献译编. 乌鲁木齐: 新疆人民出版社.
- Хөхийн Бямбажав (2019). *Тод үсгийн сурах бичиг*. Улаанбаатар: Соёмбо Принтинг.
- 姜兆勤 / Kushim Jiang (2021). *Proposal to encode four Mongolian characters in UCS*. L2/21–244.
- 郑维喆 / Weizhe Zheng (2016). *Preliminary proposal to encode four letters in Todo and Manchu Ali Gali*. L2/16–331.
- 萨如拉 (2014). 托忒文阿礼嘎礼研究. Master thesis, 内蒙古大学 / Inner Mongolia University.

10 Comparison with related documents

10.1 Hudum writing system

Comparison with Ćoijunjjab (2013). Ćoijunjjab (2013) introduces MONGOLIAN LETTER NA MEDIAL SEARATE FORM for the letter *n*. This presentation form is used as the intial N2 in Todo *n*, which is never used in Hudum. Therefore, this written form is not included in the Hudum part of this document.

More importantly, the naming principle for written forms used by Ćoijunjjab (2013) differs significantly from this document, for example, naming the initial A of the letter *a* as the "third medial" form of the letter *a*. This is because it names the written form based on its position in the Mongolian word, where the front of chachlag and particles are not considered word boundaries. So the reason it says a written form is "medial" is because it is not preceded or followed by a word boundary. However, this document defines "medial" as joined both forward and backward.

Comparison with GB/T 25914—2010. This document contains all the written forms of all the letters in GB/T 25914—2010 and can be regarded as an implementation of it.

Comparison with GB/T 25914—2022. This document contains all the written forms of all the letters in GB/T 25914—2022 and can be regarded as an implementation of it.

Comparison with LIANG (2019). This document uses the shaping progress framework of LIANG (2019) and presents the written units, phonetic letters, and logic of the shaping process in the same format. The main difference is that this document introduces positional default forms, fully uses dictionary–based particle shaping logic, and adds shaping rules as required by GB/T 25914—2022; whereas LIANG (2019) uses fallback forms instead of default forms in the shaping progress, minimizes the use of dictionary–based particle shaping logic, and does not include these additional shaping rules.

10.2 Todo writing system

Comparison with Ćoijunjjab (2013). Ćoijunjjab (2013) introduces MONGOLIAN LETTER NA THIRD MEDIAL FORM for the letter *n*. This presentation form is used as the final N in Hudum *n*, which is never used in Todo. Therefore, this written form is not included in the Todo part of this document.

Ćoijunjjab (2013) introduces MONGOLIAN LETTER TODO LONG VOWEL SIGN FINAL FORM for the letter *ᠢ*. This presentation form consists of a chachlag and a long vowel sign, and is part of the ligatures like *baᠢ*, *paᠢ* and so on. Since the behavior of the long vowel sign when added to a ligature is considered similar to a combining mark, the chachlag is not considered to be part of the long vowel sign, and thus this written form is not included.

Comparison with GB/T 36649—2018. This document contains all the written forms of all the letters in

GB/T 36649—2018 and can be regarded as an implementation of it.

10.3 Sibe writing system

Comparison with Ćoijuŋjab (2013). Ćoijuŋjab (2013) introduces MONGOLIAN LETTER YA SECOND INITIAL FORM for the letter *y*. This presentation form is used as the medial ᡳ in Hudum *y*, which is never used in Sibe. Therefore, this written form is not included in the Sibe part of this document.

Ćoijuŋjab (2013) introduces MONGOLIAN LETTER SIBE I THIRD FINAL FORM for the letter *i*. This presentation form is used as the final ᡭ in Manchu *i*, which is never used in Sibe. Therefore, this written form is not included in the Sibe part of this document.

Ćoijuŋjab (2013) introduces MONGOLIAN LETTER SIBE ZA SECOND INITIAL FORM and MONGOLIAN LETTER SIBE ZA SECOND MEDIAL FORM for the letter *z*. These presentation forms are used in the upper part of Sibe ligature *zi*, which is never used alone in Sibe. Therefore, this written form is not included in the Sibe part of this document.

The presentation form MONGOLIAN LETTER SIBE ZHA MEDIAL FORM for letter *ž* in Ćoijuŋjab (2013) is a J2 form with a ring on the right. However, examining today's Sibe newspapers, books and dictionaries, this form has changed, so it is not included in the Sibe part of this document.

Comparison with GB/T 36641—2018. GB/T 36641—2018 introduces MONGOLIAN LETTER SIBE ZA SECOND INITIAL FORM and MONGOLIAN LETTER SIBE ZA SECOND MEDIAL FORM for the letter *z*. These presentation forms are used in the upper part of Sibe ligature *zi*, which is never used alone in Sibe. Therefore, this written form is not included in the Sibe part of this document.

10.4 Manchu writing system

Comparison with Ćoijuŋjab (2013). Ćoijuŋjab (2013) introduces MONGOLIAN LETTER SA THIRD FINAL FORM for the letter *s*. This presentation form is used as the final S3 in Manchu Aligali *s*, which is never used in Manchu. Therefore, this written form is not included in the Manchu part of this document.

Ćoijuŋjab (2013) introduces MONGOLIAN LETTER SIBE ANG FINAL FORM for the letter *ŋ*. This presentation form is used as the final AG3 in Sibe *ŋ*, which is never used in Manchu. Therefore, this written form is not included in the Manchu part of this document.

Ćoijuŋjab (2013) introduces MONGOLIAN LETTER SIBE ZA SECOND INITIAL FORM and MONGOLIAN LETTER SIBE ZA SECOND MEDIAL FORM for the letter *z*. For the same reasons as in section 10.3, these written forms are not included in the Manchu part of this document.

Ćoijuŋjab (2013) introduces MONGOLIAN LETTER SIBE ZHA MEDIAL FORM for the letter *ž*. This presentation form is different from the final ᡮ in Manchu *ž*, so this written form is not included in the Manchu part of this

document.

Ćoijuŋjab (2013) introduces MONGOLIAN LETTER MANCHU I SECOND MEDIAL FORM for the letter *i*. This presentation form is used as the medial AI in Sibe *i*, which does occur in Manchu literature, but is regarded as a non–normative writing variant, so this written form is not included in the Manchu part of this document.

Comparison with GB/T 36645—2018. […]

10.5 Hudum Aligali writing system

[…]

10.6 Manchu Aligali writing system

[…]

10.7 Todo Aligali writing system

[…]