Representing Mongolian in Unicode
蒙古文字的文本表记与成形规则

Kushim JIANG 姜兆勤, 2024–07–12

Contents 目录

1 Introduction 引言

The ISO/IEC 10646 and the Unicode Standard have supported Mongolian script for about twenty years, with the code charts defining the basic information about characters and their variants. However, due to the lack of written, complete, and standardized shaping rules at the time, there was no consensus on how these characters should be presented. Since then, various vendors have been striving to define the functions of these characters, while users have been troubled by the incompatibilities between vendors and the insufficient support for Mongolian script on major platforms.
ISO/IEC 10646 和 Unicode 标准自约二十年前开始支持蒙古文字，其码表中定义了字符及其变体的基本信息。然而，由于当时缺乏成文的、完整的、规范化的变形规则，人们对于这些字符的呈现方式无法达成一致共识。此后，各厂商一直在努力界定这些字符的功能，而用户则因厂商之间的互不兼容以及主要平台对蒙古文字的支持不足而感到困扰。

Experts have been dedicated to improving the character set of the Mongolian script, clarifying the shaping rules and the function of each character, and proposing numerous modifications and supplements. The step-by-step shaping system for Hudum can be found in L2/19–368, while the flat shaping requirements for Hudum, Todo, Sibe, and Manchu scripts is detailed in "Research on Mongolian Encodings". Recently, with the release of GB/T 25914—2023, the flat shaping requirements for Hudum and Hudum Ali Gali have been established as the Chinese national standard.
有关专家一直致力于完善蒙古文字的字符集，澄清变形规则与字符功能，并提出了许多修改与增补方案。其中，针对传统蒙古文所构建的分步变形规则系统可参见 L2/19–368，针对传统蒙古文、托忒文、锡伯文与满文所构建的变形要求可参见"蒙古文编码研究"。最近，随着 GB/T 25914—2023 的发布，传统蒙古文及其阿礼嘎礼的变形要求被确立为中国国家标准。

This document demonstrates and proposes a system of text representation and text shaping rules for Mongolian script. For text representation, this document will list the characters and their variants for letters and marks used in each writing system; for text shaping, this document will provide step-by-step shaping rules, detailing the behavior of each step and the presentation conditions of each variant.
本文将提供蒙古文字的文本表记与文本成形规则系统。从文本表记方面，本文将列举用于各书写系统的字母与记号的字符及其变体；从文本成形方面，本文将提供分步的变形规则，详细说明各步骤的行为与各变体的呈现条件。

The specifications provided in this document are oriented towards the first three layers of a five-layer system of text processing techniques, i.e.:
本文提供的规范面向五层的文本处理技术体系的前三层，即：

- The code charts of the Unicode Standard and ISO/IEC 10646 specify the code points, names, with annotation about characters and character sequences.
  Unicode 标准与 ISO/IEC 10646 的码表，规定字符与字符序列的码位、名称等信息。

The characters specified in this document correspond to the characters in the code charts.
本文所规定的字符与码表中的字符相对应。

- 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.), thus specifying the association of characters with glyphs and showing how this association is affected by factors such as character properties, contextual environment and localization.
  Unicode 标准及其他补充文件，提供字符属性（如通用类别、连写连接类型等）和算法（如规范化、排序、换行、分段、双向文本、竖排等）等信息，从而规定字符与字形的关联，并说明这种关联如何受字符属性、上下文环境及本地化等因素的影响。

  This document describes the association of Mongolian characters with glyphs, thus its content is closely related to the elements of this layer.
  本文描述蒙古文字字符与字形的关联，从而其内容与本层的要素密切相关。

- The text shaping engine and the font project file implement the rules for associating characters and glyphs as glyph images in the design space, substitution and positioning relations between glyph images, and text shaping operations performed by the text shaping engine.
  文本成形引擎与字库工程文件，将字符与字形的关联规则实现为设计空间中的字形图像、字形图像之间的转换关系，以及由文本成形引擎执行的文本成形操作。

  Based on specific design decisions, there is freedom in constructing glyph image sets and writing shaping rules, so this document cannot directly specify the framework of glyph image sets and shaping rules that each font project file should contain. However, the step-by-step shaping rules provided in this document can directly correspond to the shaping rules in the font project file, and therefore can provide a reference for the construction and conversion of the font project files.
  基于特定的设计决策，构建字形图像集和编写字形图像转换关系具有一定的自由度，因此本文无法直接规定各字库工程文件应包含的字形图像集与转换关系框架。然而，本文提供的分步变形规则可以直接对应于字形图像的转换关系，因此可以为字库工程文件的构建与转换提供参考。

- The typesetting application invokes the text shaping engine with the font files to render the glyph images in the typeset pages.
  排版应用调用文本成形引擎与字库文件，将字形图像呈现在排版的页面中。

- Devices such as monitors and printers convert typeset pages into physical objects.
  显示器、打印机等设备将排版的页面转换为物理对象。

The rule system provided in this document has the following features:
本文提供的规则系统具有以下特点：

- Retain the phonetic encoding principle of the existing model. Characters are defined in terms of the phonemes of the written language as recorded by each writing system. Each character contains variants that represent the written forms corresponding to the phoneme of that written language under specific phonemic analysis. Therefore, the model does not inherit the written unit analysis of Old Uyghur script, Sogdian script, and so on.
  沿用音位模型。以各书写系统所记录的书面语音位来定义字符，每个字符包含的变体代表该书面语音位在音位分析下对应的书写形式。因此，该模型不继承回鹘文字、粟特文字等的书写形式分析方法。

- Supplement the requirements of national standards on the basis of orthographic shaping. The basic shaping rules are constructed for each writing system based on the orthographic analysis. And the shaping rules are extended according to the requirements of national standards while ensuring no coupling. This makes it possible to explicitly indicate which steps derive from the additional requirements of the national standards in the step-by-step shaping rules.
  在正字法变形的基础上补充国家标准的要求。基于正字法分析，为各书写系统构建基础的变形规则，并在确保无耦合的前提下，按照国家标准的要求扩充变形规则。这使得在分步变形规则中，可以明确指示哪些步骤源自国家标准的额外要求。

- Provides font test tools and font generation tools. This document maintains and provides information on Mongolian characters and glyphs, text shaping rules, test sets and font test tools, and font generation tools based on a given set of glyph images and correspondence data. In addition, this document introduces how to use the above information and tools.
  提供字库测试工具与字库生成工具。本文维护并提供蒙古文字的各字符与字形信息、文本成形规则信息、测试集与字库测试工具，以及基于给定字形图像集与对应关系数据的字库生成工具等。此外，本文还将介绍上述信息与工具的使用方法。

- Formally define modular, hierarchical, unambiguous and prioritized text shaping rules. Flat shaping rules partition the contextual environment in which each character is placed into multiple parts and specify the glyphs corresponding to the characters in each part of the environment. Such a specification is not conducive to understanding and writing shaping rules, and thus to font production. Effectively defined step-by-step text shaping rules facilitate the writing of shaping rules with consistent and strict behavior during font production.
  形式化定义模块化、层次化、无歧义且优先级明确的文本成形规则。扁平的变形规则将各字符所处的上下文环境分割为多个部分，并规定各部分环境中字符所对应的字形。这种规定不利于理解和编写变形规则，进而不利于字库生产。有效定义的分步文本成形规则，有助于在字库生产过程中编写行为一致且严格的变形规则。

2 Architecture 架构

Each writing system using Mongolian script and the unified writing system formed by integrating them have the same architecture of text representation and text shaping. The text representation and text shaping rules contained in the individual writing systems and the integrated writing system are the concrete realization of this common architecture. This chapter describes this common architecture, and subsequent chapters describe the specific rules of each writing system in detail.
使用蒙古文字的各个书写系统，以及综合这些书写系统所形成的统一的书写系统，均具有相同的文本表记与文本成形架构。每种书写系统及其综合系统所包含的文本表记与文本成形规则，都是该共同架构的具体实现。本章将介绍这一共同架构，后续章节将详细说明各个书写系统的具体规则。

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

The text shaping data for Mongolian script are in the form of machine-readable files that provide the character attribute information, text shaping rules, font generation program, and font testing program for Mongolian script.
蒙古文字的文本成形数据以机器可读的文件形式，提供蒙古文字的字符属性信息、文本成形规则信息、字库生成程序和字库测试程序等内容。

The character attribute information and text shaping rules of Mongolian script provide the correspondence between Mongolian characters and variants in each writing system, and the correspondence between Mongolian character strings and glyph sequences. The font generation program is capable of completing and generating a usable font by providing the basic design file, the correspondence information between glyph images and variants, and the setting information. The font testing program can test how the font supports each writing system by providing the font file and the correspondence between glyph images and variants.
蒙古文字的字符属性信息和文本成形规则信息，详细描述了各书写系统中蒙古文字字符与字形的对应关系，以及蒙古文字字符串与字形序列的对应关系。字库生成程序在提供基础字库文件、字形图像与字形的对应关系信息和设定信息的情况下，能够增补并生成可用的字库。字库测试程序则在提供字库文件和字形图像与字形对应关系信息的情况下，能够测试字库对各书写系统的文本成形规则的支持情况。

2.2 Character set 字符集

The characters included in each writing system of the Mongolian script can be categorized into Mongolian-specific characters and characters shared with other scripts according to Table 1.
蒙古文字的各书写系统所包含的字符可按表 1 归类，分为特有字符和共用字符。

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, …	

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

Due to the introduction of *matres lectionis*, for the writing systems using Mongolian script, both vowels and consonants of the written language correspond to the actual texts, and both are similar in behavior and equal in status. Therefore, although the original writing system was abjad, each writing system using Mongolian script has the characteristics of an alphabet.

由于 *matres lectionis* 的引入，对于使用蒙古文字的书写系统来说，其记录的书面语中的元音音位和辅音音位均对应于文本元素，并且这两种文本元素的行为相似、地位均等。因此，即使原有的书写系统属于 abjad，使用蒙古文字的各个书写系统也因此带有了 alphabet 的特征。

These writing systems, with the characteristics of both abjad and alphabet, result in two ways of analyzing and encoding the Mongolian script. One is to determine the graphemes used by each writing system by chronologically comparing Mongolian script with Old Uyghur script and Old Turkic script, and to identify each grapheme as a character; the model formed by this method is called the **graphetic model**, and the basic unit in this model is called the **writing unit**. The other is to analyze the phonemes of the written language recorded by the writing system, to group together the glyphs that record the same phoneme, and to identify each phoneme as a character; the model formed by this method is called a **phonetic model**, and the basic unit in the model is called the **phonetic letter**.

这些兼具 abjad 和 alphabet 特性的书写系统，产生了两种分析与编码蒙古文字的方式。一种通过将蒙古文字与回鹘文字、突厥文字等进行历时比较，确定各书写系统使用的形位，并将各个形位确定为字符，这种方法形成的模型称为**形位模型**，该模型中的基本单位称为**书写单位**。另一种通过分析书写系统所记录的书面语音位，归纳记录同一音位的字形，并将各个音位确立为字符，这种方法形成的模型称为**音位模型**，该模型中的基本单位称为**音位字母**。

The specification used in this manual is based on the phonemic model, in which the characters correspond to the phonemes of the written language as recorded in the writing systems, with some compromises to maintain relative stability with the existing system. For example, the final ᡳ in Manchu records the phoneme /ɨ/, is the same as the phonetic letter ***ï***, but because it corresponds to ⟨i⟩ rather than ⟨y⟩ in the Möllendorff transcription, it is treated as corresponding to the phonetic letter ***i*** rather than ***ï***. This analysis is inherited in this specification.
本规范基于音位模型，其中各个蒙古文字的字符对应于其书写系统所记录的书面语音位，但也出于保持现有系统的相对稳定有所取舍。例如，圈点满文中前连形 ᡳ 所记录的书面语音位为 /ɨ/，虽与音位字母 ***ï*** 相同，但由于穆麟德转写中该形位对应于 ⟨i⟩ 而非 ⟨y⟩，该形位被视作对应于音位字母 ***i*** 而非 ***ï***。这一分析在本规范中得到继承。

Multi-to-one and one-to-multi confusion. Since the writing units and the phonemes do not have a one-to-one correspondence, the phenomenon of the same phoneme corresponding to more than one written unit and the same written unit corresponding to more than one phoneme may occur. For example, in Hudum, the phonemic letter ***n*** is usually written as N when it appears as an onset and as A when it appears as a coda (except for loanwords such as SWK2OIND); and the grapheme sequence AORDU may correspond to ***ordo*** "palace", ***urtu*** "long", and ***urdu*** "south".
多对一与一对多的混淆性。由于形位模型中的书写单位与音位模型中的音位字母并非一一对应，采用音位模型时，会出现同一音位字母对应多个书写形式，以及同一书写形式对应多个音位字母的现象。以传统蒙古文为例，音位字母 ***n*** 用于音节首时通常显示为 N，而在用于音节尾时则通常显示为 A（除外来词如 SWK2OIND）；形位序列 AORDU 则可对应于 ***ordo***“宫殿”、***urtu***“长的”和 ***urdu***“南面”。

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 character 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 characters.
零宽禁连符（ZWNJ）、零宽连接符（ZWJ）与尼茹股。ZWNJ 与 ZWJ 是 Unicode 的标准连写控制字符，属于不可见字符。需要注意的是，ZWJ 会破坏相邻字符之间的联系（如连字形式）。尼茹股是蒙古文字特有的字符，其控制连写的功能与 ZWJ 完全相同，但它作为可见的字干存在。应避免普通用户通过键盘输入 ZWNJ 和 ZWJ，因为日常文本不包含这些字符。

Word Root Delimiter. In Hudum and Todo, words with two roots are presented as the cursive joining of the glyphs of each root. For example, *sainirögel* in Hudum consists of the glyph SAIIA for *sain* and the glyph AIROGAL for *irögel*, i.e., SAIIAAIROGAL. If the glyphs are not triggered manually, a format control

character should be introduced to indicate the root delimiter. MGC/01–01 suggests the use of ZWJ or NIRUGU, etc. As an interim solution, this document supports the use of ZWJ as a root delimiter.
词根分界符。在传统蒙古文和托忒文中，双词根词显示为各词根的字形的连写。例如，传统蒙古文中的 *sainirögel* 由 *sain* 的字形 SAIIA 和 *irögel* 的字形 AIROGAL 连写组成，即 SAIIAAIROGAL。如果不手动改变字形，则应引入格式控制符来指示词根的分界。MGC/01–01 建议使用 ZWJ 或尼茹股等临时方案，本文支持使用 ZWJ 作为词根分界符。

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写。常见的使用情况是用于父名缩写末尾，该缩写可能是词首音节（即可能存在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 whitespace and format control used to represent and present particles. It is transcribed as "–" (an en-dash). Use of the NNBSP is discouraged in preference for the MVS, as it sometimes produces anomalous shaping in various contexts.
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。在本文中，MVS 转写为"·"（一个间隔号）。NNBSP 是蒙古文字使用的空格与格式控制符，用于表记和呈现词缀。在本文中，NNBSP 转写为"–"（一条半身线）。目前，由于 NNBSP 的文种属性推断容易出错，容易在各种情况下出现成形异常，NNBSP 不被推荐在蒙古文字中使用，其相关功能完全由 MVS 实现。

Free Variation Selectors (FVS's). FVS's are Mongolian-specific format controls. They are applied to follow 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，它们出现在某些字符之后，主要用于显示与正字法成形规则不同的字形。

2.2.3 Numbers and Punctuation 数字与标点符号

The document does not contain specifications for numbers, punctuation, damaru, ubadama, and so on; these will be added in subsequent versions.
本文暂未包含关于数字、标点符号、扎嘛噜、乌巴达嘛等的字形规范，这些内容将在后续版本中补充。

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成熟技术基础，并在连写文字的一般成形过程中插入了与蒙古文字相关的文本成形步骤。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 – 字形图像映射（IA 阶段）通常由 TrueType/OpenType 的 `cmap` 表控制。可以使用 Unicode 名义字形图像作为音位字母映射后的默认字形图像，但这些名义字形图像与最终渲染完全无关。

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

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。

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 – 字形图像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 (above, in Mongolian), not joined backward (below, in Mongolian);
  非连形，缩写为“isol”：前方不连，后方不连（即为纵排文本中的上方不连，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写的情形有限。

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形图像将被映射为该字母在该连写点位下的默认字形图像。

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 an FVS.
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。特别地，红色数字表示未见于

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。

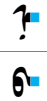
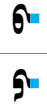
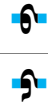
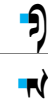
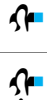
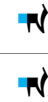
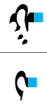
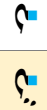
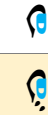
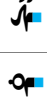
- *Initial*, [I]: Found in initial syllables.
  词首，[I]：见于首音节。
- *Devsger*, [D]: Found in the consonant letter at the end of a Mongolian syllable, or in the vowel letter following a vowel letter.
  韵尾，[D]：见于音节末辅音或元音后元音。
- *Root*, [R]: Found in the second root of a double root word.
  词根，[R]：见于双词根词的第二词根。
- *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									Early modern orthography.
O										
U										
Ų	Ux									Early modern orthography.
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										

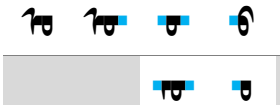
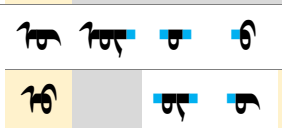
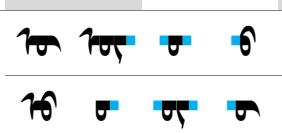
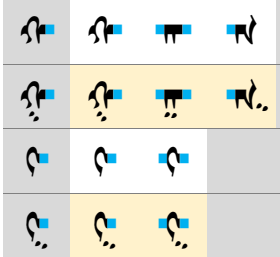
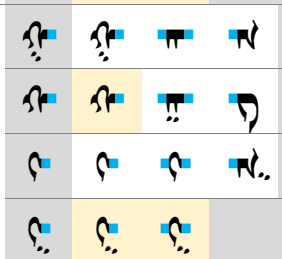
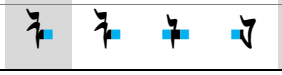
Đ	Dd												
Ć	Ch												
J													
Y													
R													
W													
F													
K													Unified: ‹K2 (K2)›.
C													
Z													
Ħ	Hr												
Ř	Rh												
Ž	Zr												
Č	Cr												

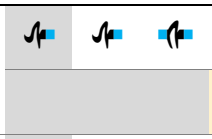
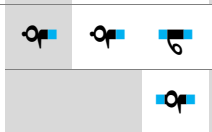
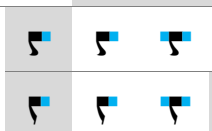
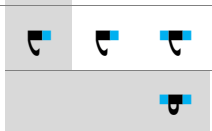
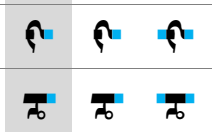
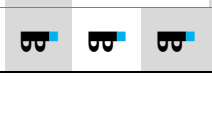
3.1.2 Phonetic letters 音位字母

Table 4 Hudum phonetic letters

表 4 传统蒙古文音位字母

Phon. let. 音位字母		Unic. 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. Į.init [H]: *man–i*, MAA Į.
							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*, TAẗOA Ụ.
				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*, SANAẗAAOILJAI. 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*, TAGOA Ụ. 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	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*, ḦADA; *sahal*, SAḦAL; *bah·a*, BAḦ 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*, BAḦ Á. Ḧ.init, G̈.init, G̈.medi [H]: *goyo*, ḦOYU; *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–	Ž.medi: *možiḳan*, MOŽIK2AA.
č	cr	1842		Č–	Č	Č–	Č–	

Isolated ᵢ and isolated ᵤ. Isolated ᵢ and isolated ᵤ are used historically, with examples shown in [Dorji *et al.*, 2017]. In some literature like [Haishan, 1917], even if the letter n before isolated ᵢ or ᵤ is omitted, the left point is retained and therefore has to be consistently analyzed as belonging to *i*, *u* or *ü*.
非连形 ᵢ 与非连形 ᵤ。非连形 ᵢ 与非连形 ᵤ 用于历史文献，示例参见文献 [Dorji *et al.*, 2017]。在一些文献如 [海山, 1917] 中，即使非连形 ᵢ 与非连形 ᵤ 之前的字母 *n* 被省略，居左的点仍被保留，因此该点不得不被一致性地分析为属于字母 *i*、*u* 或 *ü*。

Final ᵢ in letter *j*. Final ᵢ 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 ᵢ in letter *j* and chachlag.
字母 *j* 的前连形 ᵢ。字母 *j* 的前连形 ᵢ 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 ᵢ 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 ᵢ 与分写左撇的词汇。

Medial ᵢᵢ in letter *y* and medial ₒ in letter *w*. Medial ᵢᵢ in letter *y* and medial ₒ in letter *w* are introduced in [Dorji *et al.*, 2017], where the written form ᵢᵢ of letter *y* in the examples *ayl* and *aymag*, and the written form ₒ 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 ᵢᵢ 与字母 *w* 的双连形 ₒ。字母 *y* 的双连形 ᵢᵢ 与字母 *w* 的双连形 ₒ 在 [Dorji *et al.*, 2017] 中引入，其中示例 *ayl* 与 *aymag* 中 *y* 的 ᵢᵢ 以及示例 *tawlay* 中 *w* 的 ₒ 均源于字母分析差异。由于本文倾向于分析成 *ail*、*aimag* 与 *taulai*，这些书写形式被标灰。

Letter *ķ* (ₖ₂) and letter *k* (ₖ). Letter *ķ* in form ₖ₂ is used to represent /kʰ/ (Latin: *k*) in Hudum literature in China, while letter *k* in form ₖ is used to represent /kʰ/ (Cyrillic: *к*) in Hudum literature in Mongolia. The reason for disunifying the two letters is that Hudum Alikali uses ₖ₂ to record /g/ and uses ₖ to record /kʰ/.
字母 *ķ* (ₖ₂) 与 *k* (ₖ)。字母 *ķ* (ₖ₂) 在中国的传统蒙古文文献中用于记录音位 /kʰ/ (*k*)，字母 *k* (ₖ) 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 ₖ₂ 记录音位 /g/，用 ₖ 记录音位 /kʰ/。

Medial ₕ in letter *ħ* and initial ž in letter *ǯ*. Medial ₕ in letter *ħ* and initial ž in letter *ǯ* are similar, so they are unified in proposals such as [SAH, 2017]. The reason for disunifying the two written units is that Hudum Alikali uses ₕ to record /ɦ/ and so on, and uses ž to record /ʈ/.
字母 *ħ* 的双连形 ₕ 与字母 *ǯ* 的后连形 ž。字母 *ħ* 的双连形 ₕ 与字母 *ǯ* 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 ₕ 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 ł 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`
	a e	**if** follows an MVS and precedes an FVS:	`Default`
2. Syllabic	*o u ö ü*	**if** follows an initial *consonant*:	`Marked`
	o u ö ü	**if** precedes an FVS or follows an FVS:	`Default`
	ö ü	**if** final form precedes an FVS2 or FVS4 that precedes *h*/*g*:	`Marked`
	ö ü	**if** medial form follows a *consonant* cluster that starts with an initial *consonant*:	`Marked`
	ö ü	**if** follows a ZWJ or a medial consonant that follows a ZWJ:	`Marked`
	vowel	**if** follows a ZWJ:	`Toothed`
	d	**if** precedes a final *vowel* that does not precedes an FVS:	`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:	`Feminine`

		t	**if** precedes *é* or a *consonant*:	Devsger
		ś	**if** initial form precedes medial *i*, or medial form 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 end with a written unit I:	Devsger
		i	**if** follows a *vowel* that will not (consider FVS) 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

Phon. let.	Unic.	Lookup	Written forms				Comments
a	1820	Default	AA	AA	A	A	
		Chachlag	Á				
		Toothed			AO		
		Particle	Á	A			
		Pos_Bow				Á	
e	1821	Default	A	A	A	A	
		Chachlag	Á				
		Particle	Á				
		Pos_Bow				Á	
é	1827	Default	AW	AW	W	W	
i	1822	Default	AI	AI	I	I	
		Toothed			AI		
		Particle	I	I			
		Devsger			II		
o	1823	Default	AO	AO	O	U	
		Marked				O	
		Toothed			AO		

			Pos_Bow							O	
u		1824	Default					AO	AO	O	U
			Marked								O
			Toothed							AO	
			Particle					U	O		U
			Pos_Bow								O
ö	oe	1825	Default					AÜ	AOI	O	U
			Marked							OI	Ü
			Toothed							AOI	
			Pos_Bow								O
ü	ue	1826	Default					AÜ	AOI	O	U
			Marked							OI	Ü
			Toothed							AOI	
			Particle					U	O	O	U
			Pos_Bow								O
n		1828	Default					N–	N	A	A
			Cha_Ons								N
			Onset							N	
			Devsger							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
			Cha_Ons								H
			Mas_Ons						H	H	
			Mas_Dev							H	H
			Feminine						G	G	
g		182D	Default					Ë–	Ë	G	G	Masculine_Devsger affects all of the

			Cha_Ons								Ḧ	medial forms and final forms (including masculine forms and feminine forms), and Dotless only affects masculine forms with dots.
			Cha_Dev								H	
			Mas_Ons							Ḧ	Ḧ	
			Mas_Dev							H	H	
			Dotless							H	H	
			Feminine						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	
			Dotless						S	S		
t		1832	Default					T–	T	D	T	
			Devsger							T		
d		1833	Default					D–	T	Đ	Đ	
			Marked						D			
			Onset						T	D		
			Devsger							Đ	Đ	
			Particle						D			
ć	ch	1834	Default					Ć–	Ć	Ć	Ć	
j		1835	Default					I–	I	J	J	
			Cha_Ons					I			I	
y		1836	Default					Y–	Y	Y	I	
			Particle						I	I		
r		1837	Default					R–	R	R	R	
w		1838	Default					W–	W	W	W	
			Cha_Ons								U	
f		1839	Default					F–	F	F	F	
ḳ	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					ᠠᠨ–	ᠠᠨ	ᠨ	ᠨ
ř	rh	183F	Default					Ř–	Ř	Ř	Ř
ɬ	lh	1840	Default					ᠯᠨ–	ᠯᠨ	ᠯᠨ	–ᠯᠨ–
ž	zr	1841	Default					Ž–	Ž	Ž	Ž–
č	cr	1842	Default					Č–	Č	Č–	Č–

Shaping rules required by EAC. The grey shaping steps are specified by the EAC, see GB/T 25914—2023 and MGC/01–01. These shaping steps either introduce rules other than orthographic analysis, or cause the FVS to change the glyphs of characters other than those that precede it.
由 EAC 规定的变形规则。灰色的变形步骤由 EAC 规定，参见 GB/T 25914—2023 与 MGC/01–01。这些变形步骤或是引入了正字法分析以外的规则，或是使得 FVS 改变除了它之前的字符以外的其他字符的字形。

Shaping rules specified by UTN. The green shaping steps are specified by this UTN to temporarily support the shaping requirements for words with two roots.
由 UTN 规定的变形规则。绿色的变形步骤由本 UTN 规定，用于暂时支持与双词根词相关的变形要求。

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—2023.
- 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/allchars-mon.php.
- 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.
- 海山 / [Hudum] *Bayantemür–ün Haisaŋ* (1917). 蒙漢合璧五方元音. 民国六年仲春北京初版.

4 Todo writing system 托忒文

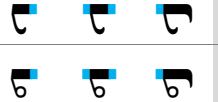
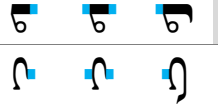
4.1 Character set 字符集

4.1.1 Written units 书写单位

Table 8 Todo written units

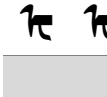
表 8 托忒文书写单位

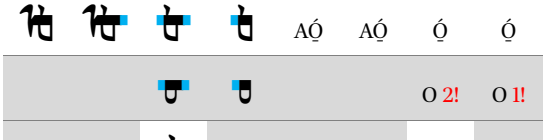
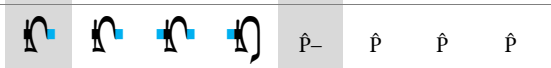
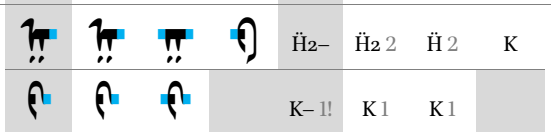
Written unit		Positional forms: .isol, .init, .medi, .fina				Sub-written-unit variants				Comments 备注
A										
Á	Aa									
E										
I										
Î	Ip									
O										
Ó	Op									
Ọ	Ob									
Ọ́	Ot									
U										
Ú	Up									
Ł	Lv									May behave as a mark.
N										Unified: ‹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											
Ĵ	Jb										
Y											Not recommended.
R											
W											
Ẉ	Wb										
K̂	Kp										
Ĝ	Gp										
Ħ	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

Phon. let. 音位字母	Unic. 码位	Written forms 书写形式								Comments 备注
a	1820					AA 3 A 1!	AA 2	A AA 1!	A 2 Á 1	A.isol [P]: *a*, A. AA.medi [R]: *buyanarbin*, BÓIAAAARBÎA. Á.fina [G]: *ba*, BÁ.
e	1844					AE	AE	E AE 1!	E	AE.medi [R].
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*, ʒÎAGAONEA.
						AO 1!		AO 1!	O 1	
ł	lvs	1843				Ł	Ł	Ł	Ł	
n		1828				N–	N	A 2	A 2	A.medi [D]: *front*, WRǪAṬ. N2.init [P]: *beljin–ni*, BEŁĈÎA N2I3.
							N2	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				Ḥ–	Ḥ 2	Ḥ 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。由于本文倾向于分析成 *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 AI and medial I in letter *i*, medial O and final O in letter *ö*, written form Y 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 AI 与双连形 I、字母 *ö* 的双连形 O 与前连形 O、字母 *y* 的书写形式 Y。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。

Initial N2 in letter *n*. This written unit is used for particle ***–ni*** in Todo, and is regarded as a stylistic variant of the written unit N, thus is not triggered by FVS.
字母 *n* 的前连形 N2。该书写形式用于托忒文的格附加形式 ***–ni***，视作书写形式 N 的风格变体，从而不可通过 FVS 调用。

Letter *ñ*. The letter *ñ* is introduced in [IMU, 1999], but is not found in Todo literature. The appearance in this [IMU, 1999] is subsequently succeeded by GB/T 36649—2018. The character is used by Hudum Ali Gali and Todo Ali Gali.
字母 *ñ*。字母 *ñ* 在《蒙汉词典（增订本）》中引入，但未见于托忒文献。《蒙汉词典（增订本）》中的字母表随后继承至 GB/T 36649—2018 内。传统蒙古文阿礼嘎礼与托忒文阿礼嘎礼使用该字符。

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 ɬ*

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	*vowel*	**if** follows a ZWJ:	`Toothed`
	n	**if** precedes a *vowel*:	`Onset`
		else if precedes a *consonant*:	`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

Phon. let.	Unic.	Lookup	Written forms								Comments
a	1820	`Default`					AA	AA	A	A	
		`Toothed`							AA		
		`Pos_Bow`								Á	
e	1844	`Default`					AE	AE	E	E	
		`Toothed`							AE		
i	1845	`Default`					AI3	AÎ	Î	I3	
		`Toothed`							AÎ		
		`Devsger`							AÎ		
		`Pos_Bow`								Î	
o	1846	`Default`					AO	AO	O	O	
		`Toothed`							AO		
u	1847	`Default`					AÚ	AÓ	Ó	Ú	
		`Toothed`							AÓ		
		`Devsger`							O	U	
		`Pos_Bow`								Ó	

ö	oe	1848	Default					AO̧	AO̧	O̧	O̧
			Toothed							AO̧	
ü	ue	1849	Default					AU	AO	O	U
			Toothed							AO	
			Pos_Bow								O
t	lvs	1843	Default					Ł	Ł	Ł	Ł
n		1828	Default					N–	N	A	A
			Onset							N	
			Devsger							A	
			Particle						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
			Mas_Ons						Ḧ2	Ḧ	
			Feminine						K	K	
g		184E	Default					Ḥ–	Ḥ	Ḥ	Ĥ
			Mas_Ons						Ḥ	Ḥ	
			Feminine						G	G	
			Mas_Dev							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					J̨–	J̨	J̨	J̨
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
ẖ	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 *t*. The sign *t* behaves as a base character in some cases, but its logic in the shaping process behaves as a combining mark. For this reason, the necessary preprocessing and postprocessing should be introduced: before the Mongolian-specific steps begin, *t* is shaped in mark class, modifying the joining type of the characters before and after *t*; after the Mongolian-specific steps finish, those signs *t* that need to be represented as marks are 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。

Shaping rules specified by UTN. The green shaping steps are specified by this UTN to temporarily support the shaping requirements for words with two roots.

由 UTN 规定的变形规则。绿色的变形步骤由本 UTN 规定，用于暂时支持与双词根词相关的变形要求。

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/allchars-todo.php.
- 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.
- 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.

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									
Ả	At									
Ă	Ai									
I										Unified:  ‹I2 (I2)›.
Ǐ	Ic									
O										
Ȯ	Oh									
U										
Ů	Uh									
Ü	Ue									
N										
N̈	Nx									
H										
Ḣ	Hh									
H̊	Hc									
G										Unified:  ‹G3 (G3)›.
Ġ	Gh									
Ġ̊	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										
K̇	Kh									
K̊	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

Phon. let. 音位字母	Unic. 码位	Written forms 书写形式								Comments 备注
ʔ	sbm	1807					–A–	–A–	A	–A–	A.medi: *juʔi*, 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.fina [P]: *piyazi*, ṖIYAẔI2.
							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	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	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–	Ṫ.init, Ḍ.medi [F]: *dedun*, ṪAḌOA.
							Ṯ– 1!	Ṫ 1	Ḍ 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 differing form 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。由于本文倾向于分析成 *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 Mas_Dev 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* evolved 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	*ʔ 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
		t d	**if** precedes *a*/*i*/*o*:	Masculine_Onset
			else if precedes *e*/*u*/*ü*:	Feminine
			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
			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

Phon. let.	Unic.	Lookup	Written forms								Comments
ʔ	sbm	1807	Default					–A–	–A–	A	–A–	
a		1820	Default					AA	AA	A	A	
			Pos_Bow								Á	
e		185D	Default					A	A	Ȧ	Ȧ	Post_Bowed step keeps the dotted or dotless status.
			Feminine							A	A	
			Pos_Bow								Ȧ́	
											Á	
i		185E	Default					AI	AI	I	I	
			Marked								I2	
			Particle					I				
			Devsger							AI		

			Pos_Bow								I2	
ï	ii	185F	Default					–Ă–	–Ă–	Ă	Ă	
o		1823	Default					AO	AO	O	U	
			Marked								O	
			Pos_Bow								O	
u		1860	Default					AȮ	AȮ	Ȯ	Ů	Feminine step removes the dots in written units. Post_Bowed step keeps the dotted or dotless statuss.
			Marked								Ȯ	
			Feminine							O	U	
											O	
			Pos_Bow								Ȯ	
											O	
ü	ue	1861	Default					AÜ	AOI	OI	Ü	
n		1828	Default					N–	N	A	A	
			Onset							N		
			Devsger							A		
ŋ	ng	1862	Default					–AG–	–AG–	AG	AG3	
k		1863	Default					H–	H	H	Ñ	
			Mas_Ons						H	H		
			Feminine						G	G		
			Devsger							Ñ		
g		1864	Default					Ḣ–	Ḣ	Ḣ	–Ḣ–	
			Mas_Ons						Ḣ	Ḣ		
			Feminine						Ġ	Ġ		
h		1865	Default					Ḧ–	Ḧ	Ḧ	–Ḧ–	
			Mas_Ons						Ḧ	Ḧ		
			Feminine						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	Ḏ	Đ
			Mas_Ons						T	D	
			Feminine						Ṯ	Ḏ	
			Devsger							Đ	
d		1869	Default					Ṫ–	Ṫ	Ḋ	–Ḋ–
			Mas_Ons						Ṫ	Ḋ	
			Feminine						Ṫ	Ḋ	
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					Ḱ–	Ḱ	Ḱ	–Ḱ–
ḥ	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 (锡伯文) (version 0.1.0), rule.mongoltoli.cn/allchars-sibe.php.
- 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.
- 付勇, 郭公, 锋晖 (2017). 基于音素法的锡伯文字母划分与拼写的研究. 吉林师范大学学报 (人文社会科学版), 45 (05), pp. 35–44.

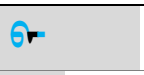
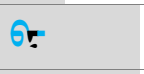
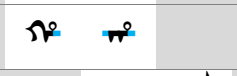
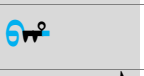
6 Manchu writing system 圈点满文

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			
À	Ah			
Ả	At			
Ă	Ai			
I				
Ï	Iy			
Î	Ic			
O				
Ò	Oh			
U				
Ù	Uh			
Ü	Ue			
N				
H				
Ḣ	Hh			
Ḥ	Hc			
Ḧ	Hx			
G				Unified: ‹G›.
Ġ	Gh			
Ġ̊	Gc			
G̈	Gx			Historical literature.

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

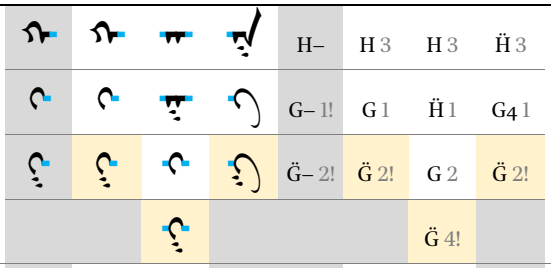
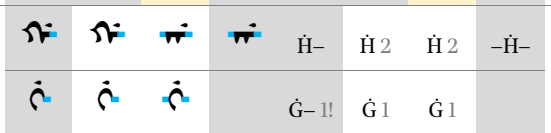
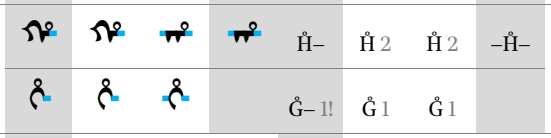
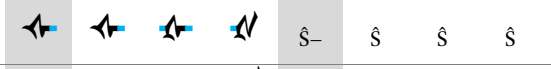
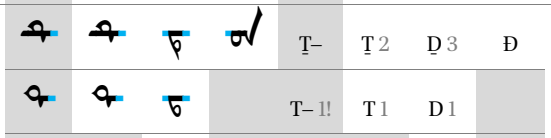
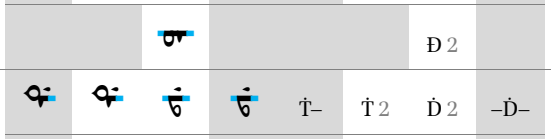
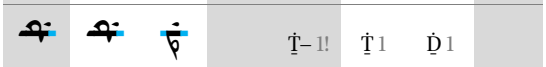
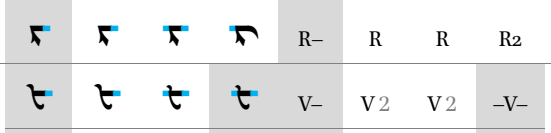
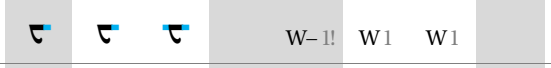
R̂	Rr								

6.1.2 Phonetic letters 音位字母

Table 19 Manchu phonetic letters

表 19 圈点满文音位字母

Phon. let. 音位字母		Unic. 码位	Written forms 书写形式								Comments 备注
ʔ	sbm	1807					–A–	–A–	A	–A–	A.medi: *kuiʔi*, 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*, ṬARĠI; *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 ẐÏ.
							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	Ů 4	O.medi, U.fina [F]: *tugi*, ṬOĠI; *albatu*, AALBAḊU. 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	A 2	A 2	A.medi [D]: *anda*, AAAḊA. N₂.fina [P, L]: *en*, AN₂; *han*, ḪAN₂.
									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		Ḥ–	Ḥ 2	Ḥ 2	–Ḥ–	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 differing form 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。由于本文倾向于分析成 *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	*ʔ 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`
	t d	**if** precedes *a*/*i*/*o*:	`Masculine_Onset`
		else if precedes *e*/*u*/*ü*:	`Feminine`
		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
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/Ṗ/K2/Ḱ/K̊:	Post_Bowed
		e u	**if** follows a (bowed) written unit B/Ṗ/G/Ġ/G̊/G̈:	Post_Bowed

Table 22 Manchu shaping: lookups

表 22 圈点满文成形过程：条件变形

Phon. let.		Unic.	Lookup	Written forms								Comments
ʔ	sbm	1807	Default					–A–	–A–	A	–A–	
a		1820	Default					AA	AA	A	A	
			Pos_Bow								Á	
e		185D	Default					A	A	Ȧ	Ȧ	Post_Bowed step keeps the dotted or dotless status.
			Feminine							A	A	
			Pos_Bow								Å	
											Á	
i		1873	Default					AI	AI	I	I	
			Marked								Ï	
			Particle					I				
			Devsger							II		
ï	ii	185F	Default					–Ă–	–Ă–	Ă	Ă	
o		1823	Default					AO	AO	O	U	
			Marked								O	
			Pos_Bow								O	
u		1860	Default					AȮ	AȮ	Ȯ	Ù	Feminine step removes the dots in written units. Post_Bowed step keeps
			Marked								Ȯ	

			Feminine							O	U	the dotted or dotless statuss.
											O	
			Pos_Bow								Ȯ	
											O	
ü	ue	1861	Default					AÜ	AOI	OI	Ü	
n		1828	Default					N–	N	A	A	
			Onset							N		
			Devsger							A		
ŋ	ng	1829	Default					–AG–	–AG–	AG	AG	
k		1874	Default					H–	H	H	Ḧ	
			Mas_Ons						H	H		
			Feminine						G	G	G4	
			Mas_Dev							Ḧ	Ḧ	
g		1864	Default					Ḣ–	Ḣ	Ḣ	–Ḣ–	
			Mas_Ons						Ḣ	Ḣ		
			Feminine						Ġ	Ġ		
h		1865	Default					Ḧ–	Ḧ	Ḧ	–Ḧ–	
			Mas_Ons						Ḧ	Ḧ		
			Feminine						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	
			Mas_Ons						T	D		
			Feminine						T̲	D̲		
			Devsger							Đ		
d		1869	Default					Ṫ–	Ṫ	Ḋ	–Ḋ–	

			Mas_Ons						Ṫ	Ḋ		
			Feminine						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–	
			Marked						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					Ǩ–	Ǩ	Ǩ	–Ǩ–	
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					Ǐ–	Ǐ	J̌	–J̌–	

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/allchars-man.php.
- 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 Alikali 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 Alikali written units

表 23 传统蒙古文阿礼嘎礼书写单位

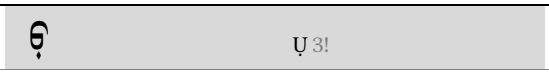
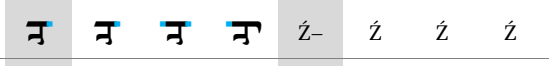
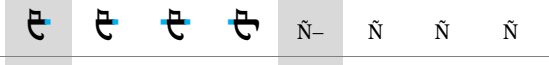
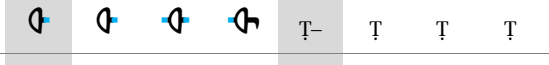
Written unit		Positional forms: .isol, .init, .medi, .fina				Sub-written-unit variants				Comments 备注
A										Unified:  ‹A2 (A2)›.
Á	Aa									
I										
O										
U										
G										
K										Unified: ‹K2 (K2)›.
N										
Ź	Zc									
C										
Z										
Ñ	Ny									
Ž	Zr									
Č	Cr									
Ḍ	Ds									
Ẉ	Wn									
Đ	Dv									
Ṭ	Tp									
Ď	Dq									
Ǩ	Bg									
Ǩ	Pg									
B										Unified:  ‹B2 (B2)›.

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

Phon. let. 音位字母		Unic. 码位	Written forms 书写形式								Comments 备注
a		1820					AA 3	AA 2	A	A 2	A.isol: *a* (Sanskrit *a*), A. Á.isol [C]: $t^{h}{}_{H}a{\cdot}a$ (Sanskrit $t^{h}\bar{a}$), ṬÁ Á. A.fina [P]: $c_{H}la$ (Sanskrit *cla*), ŹLA. Á.fina: $c_{H}a$ (Sanskrit *ca*), ŹÁ.
							A 1!	A 1	AA 1!	Á 1	
							Á 2				
a_{H}	a_{H}	1887								A2 1	A2.fina: $s\bar{w}_{H}a_{H}$ (Sanskrit *swa*), SW̄A2.
i_{H}	i_{H}	1888					AI4	AI	I	I4	
ü	ue	1826					AÜ	AOI 2	O 3	U 3	AU.isol: *ü* (Sanskrit *u*), AU. AOI.init: *ün* (Sanskrit *un*), AOIA. OI.medi: *nün* (Sanskrit *nun*), NOIA. U.fina: $c_{H}ü$ (Sanskrit *cu*), ŹU. O.fina [I, G]: *bü* (Sanskrit *bu*), BO.
							AU 1!	O 1	OI 1	Ü 1	
							U 2		AOI 2!	O 2	

							Ụ 3!				
é	ee	1827					AW	AW	W	W	
o		1823					AO	AO	O	U 2	AO.init: *ou* (Sanskrit *ū*), AOU. O.medi: *lowa* (Sanskrit *lo*), LOWA. U.fina: $k_{H}oo$ (Sanskrit *kau*), KOU. O.fina: $k_{H}y_{H}o$ (Tibetan *kyo*), GIO.
										O 1	
k_{H}	k_{H}	1889					G–	G	G	GV̀	GV̀.fina: $m_{H}y_{H}ak_{H}$ (Sanskrit *myak*), MIAGV̀.
k		183B					K–	K	K	K	
ḳ	k2	183A					K2–	K2	K2	K2	K2V̀.fina: $m_{H}oḳ$ (Sanskrit *mug*), MOK2V̀.
										K2V̀ 1!	
h		183E					AH–	AH	H	H	H.medi [P]: *bha* (Sanskrit $b^{h}a$), BHÁ.
ŋ	ng	1829					–AG–	–AG–	AG	AG	
$ŋ_{H}$	ng_{H}	188A					NG–	NG	NG	NG4	NG4.fina: $m_{H}aŋ_{H}$ (Sanskrit *mañ*), MANG4.
c_{H}	c_{H}	188B					Ź–	Ź	Ź	Ź	
c		183C					C–	C	C	C	
z		183D					Z–	Z	Z	Z	
$ñ_{H}$	ny_{H}	185B					Ñ–	Ñ	Ñ	Ñ	
$ṭ_{H}$	tr_{H}	188C					Ž–	Ž	Ž	Ž	Ž.fina: $m_{H}iṭ_{H}$ (Sanskrit *miṭ*), MIŽ.
$ṭ^{h}_{H}$	trh_{H}	188D					Č2–	Č2	Č2	–Č2–	
$ḍ_{H}$	dr_{H}	188E					Ḍ–	Ḍ	Ḍ	Ḍ	Ḍ.fina: $ṣ_{H}aḍ_{H}$ (Sanskrit *ṣaḍ*), ŠAḌ.
$ṇ_{H}$	nr_{H}	188F					�working–	Ṇ	Ṇ	Ṇ	Ṇ.fina: $d_{H}aṇ_{H}$ (Sanskrit *daṇ*), ĎAṆ.
t_{H}	t_{H}	1890					Ď–	Ď	Ď	Ď	Ď.fina: $ṭ_{H}at_{H}$ (Sanskrit *ṭat*), ŽAĎ.
t^{h}_{H}	th_{H}	1850					Ṭ–	Ṭ	Ṭ	Ṭ	
d_{H}	d_{H}	1891					Ď–	Ď	Ď	Ď	Ď.fina: od_{H} (Sanskrit *ud*), AOĎ.
n		1828					N–	N	A 2	A 2	
									N 1		
p_{H}	p_{H}	1892					Ǎ–	Ǎ	Ǎ	Ǎ	
p^{h}_{H}	ph_{H}	1893					Ǐ–	Ǐ	Ǐ	Ǐ	
b		182A					B–	B	B	B	
										B2 1!	

m_H	m_H	184F					M–	M	M	M2	
y_H	y_H	1855					I–	I	I	I4	
r		1837					R–	R	R	R	
l		182F					L–	L	L	L	
w		1838					W–	W	W	W	
$\bar{w}_H$	wa_H	18A6					W̄–	W̄	W̄	–W̄–	
$\acute{s}$	sh	1831					Ś–	Ś 2	Ś 2	Ś	
$\acute{s}_H$	sh_H	1894					S̈–	S̈	S̈	S̈	
s		1830					S–	S	S	S	
h_H	h_H	1859					AH–	AH	AH	AH	AH.medi: $h_H h_H a$ (Sanskrit *hha*), AHAHÁ.
									H 1!		
$\check{c}_H$	cr_H	1854					J–	J	J	J	
$\acute{c}$	ch	1834					Ć–	Ć	Ć	Ć	
$\check{z}_H$	zr_H	1895					Ř2–	Ř2	Ř2	Ř2	
z_H	z_H	1896					Ż–	Ż	Ż	Ż	
$ʔ_H$	q_H	1897					Q–	Q	Q	–Q–	
$ɬ$	lh	1840					LH–	LH	LH	–LH–	

Final A2 in letter a_H. Final A2 in letter a_H is introduced in GB/T 36645—2018, where Tibetan /wa/ is analyzed to represent using $\bar{w}_H$ and a_H. Although all other forms recoded Tibetan /a/ are analyzed as corresponding to *a*, for compatibility with the results of the historical standard, the letter a_H is still considered to be used.

字母 a_H 的前连形 A2。字母 a_H 的前连形 A2 在 GB/T 36645—2018 中引入，其中藏语 /wa/ 被分析为 $\bar{w}_H a_H$。虽然藏语 /a/ 对应的其他字形均被分析为 *a*，为兼容历史版本的分析结果，本文仍认为藏语 /wa/ 中的 /a/ 对应 a_H。

Final AG in letter *ŋ*. Both final NG4 and final AG4 are used to record Tibetan /ŋ/ in 大藏全咒 (“Buddhist Mantras”). The final AG4 is considered as a variant of final AG4, and the two shapes correspond to the same written form.

字母 *ŋ* 的前连形 AG。《大藏全咒》中使用前连形 NG4 与前连形 AG4 记录藏语 /ŋ/。本文认为前连形 AG4 是前连形 AG 的书写变体，对应于同一个书写形式。

7.2 Shaping process 变形流程

Table 25 Hudum Alikali phonetic letter classes
表 25 传统蒙古文阿礼嘎礼音位字母类

Class	Members
vowel	*a* a_{H} i_{H} *ü é o*
consonant	k_{H} *k ķ h ŋ* $ŋ_{H}$ c_{H} *c z* $ñ_{H}$ $ţ_{H}$ $ţ^{h}_{H}$ $ḍ_{H}$ $ṇ_{H}$ t_{H} t^{h}_{H} d_{H} *n* p_{H} p^{h}_{H} *b* m_{H} y_{H} *r l w* $w̃_{H}$ *ś* $ş_{H}$ *s* h_{H} $č_{H}$ *ć* $ž_{H}$ z_{H} $ʔ_{H}$ *ł*

Table 26 Hudum Alikali 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	*ü*	**if** follows an initial *consonant* or an initial *consonant* cluster:	Marked
		a	**if** precedes an MVS that precedes an isolated *a*:	Chachlag_Onset
5.	Post-bowed	a_{H}	**if** follows $w̃_{H}$:	Post_Wa
		a o u	**if** follows a (bowed) written unit G/K/K2/Ǩ/P̌/B:	Post_Bowed

Table 27 Hudum Alikali shaping: lookups
表 27 传统蒙古文阿礼嘎礼成形过程：条件变形

Phon. let.	Unic.	Lookup	Written forms								Comments
a	1820	Default					AA	AA	A	A	
		Chachlag					Á				
		Pos_Bow								Á	
a_{H}	1887	Default								A2	
		Post_Wa								A2	
i_{H}	1888	Default					AI4	AI	I	I4	
ü	1826	Default					AÜ	AOI	O	U	
		Marked							OI		
		Pos_Bow								O	
é	1827	Default					AW	AW	W	W	
o	1823	Default					AO	AO	O	U	

			Pos_Bow								O
k_H	k_H	1889	Default					G–	G	G	GV̀
k		183B	Default					K–	K	K	K
ḳ	k2	183A	Default					K2–	K2	K2	K2
h		183E	Default					AĦ–	AĦ	Ħ	Ħ
ŋ	ng	1829	Default					–AG–	–AG–	AG	AG
$ŋ_H$	ng_H	188A	Default					NG–	NG	NG	NG4
c_H	c_H	188B	Default					Ź–	Ź	Ź	Ź
c		183C	Default					C–	C	C	C
z		183D	Default					Z–	Z	Z	Z
$ñ_H$	ny_H	185B	Default					Ñ–	Ñ	Ñ	Ñ
$ṭ_H$	tr_H	188C	Default					Ž–	Ž	Ž	Ž
$ṭ^h{}_H$	trh_H	188D	Default					Č2–	Č2	Č2	–Č2–
$ḍ_H$	dr_H	188E	Default					D̰–	D̰	D̰	D̰
$ṇ_H$	nr_H	188F	Default					Ẉ–	Ẉ	Ẉ	Ẉ
t_H	t_H	1890	Default					D̃–	D̃	D̃	D̃
$t^h{}_H$	th_H	1850	Default					Ṭ–	Ṭ	Ṭ	Ṭ
d_H	d_H	1891	Default					Ď–	Ď	Ď	Ď
n		1828	Default					N–	N	N	A
p_H	p_H	1892	Default					B̌–	B̌	B̌	B̌
$p^h{}_H$	ph_H	1893	Default					P̌–	P̌	P̌	P̌
b		182A	Default					B–	B	B	B
m_H	m_H	184F	Default					M–	M	M	M2
y_H	y_H	1855	Default					I–	I	I	I
r		1837	Default					R–	R	R	R
l		182F	Default					L–	L	L	L
w		1838	Default					W–	W	W	W
$\bar{w}_H$	wa_H	18A6	Default					W̄–	W̄	W̄	–W̄–

ś	sh	1831	Default					Ś–	Ś	Ś	Ś
ṣʜ	shʜ	1894	Default					S̈–	S̈	S̈	S̈
s		1830	Default					S–	S	S	S
hʜ	hʜ	1859	Default					ʜ–	ʜ	Aʜ	ʜ
čʜ	crʜ	1854	Default					J–	J	J	J
ć	ch	1834	Default					Ć–	Ć	Ć	Ć
žʜ	zrʜ	1895	Default					Řƨ–	Řƨ	Řƨ	Řƨ
zʜ	zʜ	1896	Default					Ż–	Ż	Ż	Ż
ʔʜ	qʜ	1897	Default					Q–	Q	Q	–Q–
ɬ	lh	1840	Default					LƬ–	LƬ	LƬ	–LƬ–

7.3 Reference　参考文献

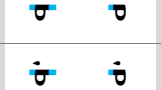
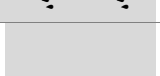
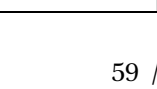
- (1929). 同文韻統. 上海涵芬楼影印清乾隆内府刻本.
- (1929). 漢滿蒙藏四體合璧阿禮嘎禮讀咒法. 上海涵芬楼影印清乾隆内府刻本.
- (1929). 漢滿蒙藏四體合璧大藏全咒. 上海涵芬楼影印清乾隆内府刻本.
- 乌日晗 (2019). 康熙版蒙古文《甘珠尔》字体复刻开发研究 / *Study on the revival font development of the Kangxi version of Mongolian Kanjur*. Master thesis, 内蒙古农业大学 / Inner Mongolia Agricultural University.
- 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 Alikali 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 Alikali 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			
Ả	At			
Ǎ	Ai			
I				
Ï	Iy			
Ỉ	Ic			
Ï	Ih			
O				
Ȯ	Oh			
U				
Ủ	Uh			
Ü	Ue			
K2	K2			
Ḱ	Kh			
Ḣ	Hh			Unified: ‹Ḣ2 (Hh2)›.
Ḫ	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												
J̡	Jq											
J̊	Jt											
J̡̊	Jc											
J̈	Jh											
Ć	Ch											
Ć̊	Cc											
Ẉ	Wn											
W												
T												
Ṫ	Th											
T̈	Tx											
T̲	Tb											
Ṫ̲	Tt											
T̲̈	Ts											
Ḍ	Dr											
Ḍ̡	Dy											
D												

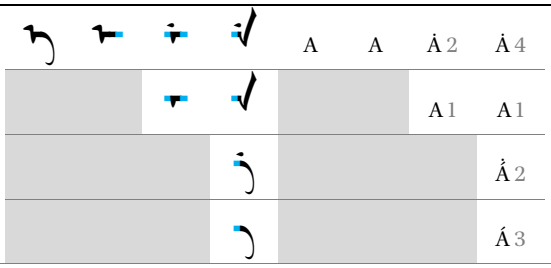
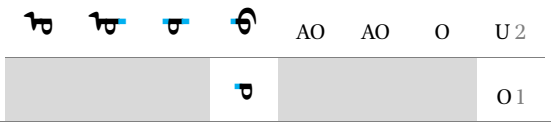
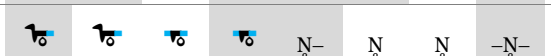
Ḋ	Dh										
Ḍ	Dx										
Ḍ	Db										
Ḍ	Dt										
Ḍ	Ds										
Đ	Dd										
B											
Ɓ	Bg										
Ḃ	Bc										
Ƥ	Pb									Unified: ‹Ƥ2 (Pb2)›.	
M											
Y											
R										Unified: ‹R3 (R3)›.	
L										Unified: ‹L2 (L2)›.	
Ĺ	Lc										
S										Unified: ‹S3 (S3)›.	
Ś	Sc										
Ŝ	Sp										
Ś	St									Unified: ‹Ś2 (St2)›.	
S̈	Sx										

8.1.2 Phonetic letters 音位字母

Table 29 Manchu Alikali phonetic letters

表 29 满文阿礼嘎礼音位字母

Phon. let. 音位字母		Unic. 码位	Written forms 书写形式								Comments 备注
a		1820					AA 3	AA 2	A	A 2	Á.fina [G]: *ba* (Sanskrit *ba*), BÁ.
										Á 1	
*ā*M	aaM	1887					AĀ	AA	A	Ā	

e		185D					A	A	Ȧ 2	Ȧ 4	A.medi, A.fina [F]: *tei* (Sanskrit *tʰai*), ṬAI; *te* (Sanskrit *tʰe*), ṬA. Å.fina, Á.fina [G]: *be* (Sanskrit *be*), BÅ; *ke* (Sanskrit *kʰe*), GÁ.
									A 1	A 1	
										Å 2	
										Á 3	
i		1873					AI 2	AI	I 3	I 3	Ï.fina [P]: *zi* (Sanskrit *ji*), ẐÏ.
							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]: *lo* (Sanskrit *lo*), LO; *bo* (Sanskrit *bo*), BO.
										O 1	
u		1860					AȮ	AȮ	Ȯ 2	Ů 4	O.medi, U.fina [F]: *tuu* (Sanskrit *tʰū*), ṬOU. Ů.fina: *niyu* (Sanskrit *ñu*), NIYŮ. Ȯ.fina [I, G]: *žu* (Sanskrit *ḍu*), ǰȮ; *bu* (Sanskrit *bu*), BȮ. O.fina [F, I, G]: *tu* (Sanskrit *tʰu*), ṬO; *ku* (Sanskrit *kʰu*), GO.
									O 1	U 1	
										Ȯ 2	
										O 3	
ü	ue	1861					AÜ	AOI	OI	Ü	
ġ	gh	186C					Ḱ–	Ḱ	Ḱ	Ḱ	
ķ	kh	183A					K₂–	K₂	K₂	–K₂–	
g		1864					Ḣ–	Ḣ 2	Ḣ 2	–Ḣ–	Ġ.init [F]: *gi* (Sanskrit *ki*), ĠI. Ḣ₂.medi [L]: *ʔgeks* (Tibetan *ⁿgegs*), A₂Ḣ₂AGS.
							Ġ– 1!	Ġ 1	Ġ 1		
									Ḣ₂ 3!		
k		1874					H–	H 3	H 3	Ḧ 3	G.init, G.medi, G₄.fina [F]: *ki* (Sanskrit *kʰi*), GI; *ʔgeks* (Tibetan *ⁿgegs*), A₂Ḣ₂AGS; *sek* (Tibetan *seg*), SÀG₄.
							G– 1!	G 1	Ḧ 1	G₄ 1	
							Ġ– 2!	Ġ 2!	G 2	Ġ 2!	
									Ġ 4!		
gʰ M	gh M	189A					Ñ–	Ñ	Ñ	–Ñ–	
									Ñ₂ 1!		
ŋ M	ng M	189B					Ṇ–	Ṇ	Ṇ	–Ṇ–	

ŋ	ng	1829					–AG–	–AG–	AG	AG	
*c*M	cM	189C					Ç̊–	Ç̊	Ç̊	–Ç̊–	
c		186E					Ç–	Ç	Ç	–Ç–	
z		186F					Ẓ–	Ẓ	Ẓ	–Ẓ–	
*jʰ*M	jhM	189D					Ẓ̊–	Ẓ̊	Ẓ̊	–Ẓ̊–	
*ṭ*M	trM	189E					Ĭ–	Ĭ	J̊	J̊	
č	cr	1871					Č̊–	Č̊	Č̊	Č̊	
ž	zr	1877					Ĭ–	Ĭ	J̊	J̊	
*ḍʰ*M	drhM	189F					Ï–	Ï	J̈	J̈	
*ṇ*M	nrM	188F					Ṿ–	Ṿ	Ṿ	Ṿ	
*t*M	tM	18A0					Ḍ–	Ḍ 2	Ḍ 2	Ḍ	Ḍ.init, Ḍ.medi [F]: *t*M*e* (Sanskrit *te*), ḌA; *st*M*eŋ* (Tibetan *steŋ*), SḌAAG.
							Ḍ– 1!	Ḍ 1	Ḍ 1		
t		1868					Ṭ–	Ṭ 2	Ḍ 3	Đ	Ṭ.init, Ḍ.medi [F]: *te* (Sanskrit *tʰe*), ṬA; *ʔten* (Tibetan *ⁿtʰen*), A2ḌAA. Đ.medi, Đ.fina [D]: *oᵗʰburub*, AOĐBȮRȮB; *ʔoᵗʰ*, A2OĐ.
							T– 1!	T 1	D 1		
									Đ 2		
d		1869					Ṫ–	Ṫ 2	Ḋ 2	–Ḋ–	Ṭ.init, Ḍ.medi [F]: *de*, ṪA; *rdul*, RḊOL.
							Ṯ– 1!	Ṯ 1	Ḏ 1		
*dʰ*M	dhM	18A1					T̈–	T̈ 2	Ḋ 2	–Ḋ–	Ṭ.init, Ḍ.medi [F]: *dʰe*, T̈A.
							T̈– 1!	T̈ 1	Ḏ 1		
n		1828					N–	N	A 2	A 2	A.init [D]: *nku* (Tibetan *ⁿkʰu*), AGO. A.medi [D]: *kʰantʰ* (Uyghur), K2AAĐ. N2.fina [D]: *byan*, BYAN2.
								A 1	N 1	N2 1!	
*p*M	pM	1892					Ƀ–	Ƀ	Ƀ	Ƀ	
p		1866					Ƥ–	Ƥ	Ƥ	–Ƥ–	
							Ƥ2– 1!	Ƥ2 1!	Ƥ2 1!		
b		182A					B–	B	B	B	
*bʰ*M	bhM	18A8					B̊–	B̊	B̊	–B̊–	
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– 1!	R3 1!	R3 1!		
l		182F					L–	L	L	L	L2.init [Vocalic]: *li*, L2I.
							L2– 1!	L2 1!			
w		1838					W–	W	W	W	
ś	sh	1867					Ŝ–	Ŝ	Ŝ	Ŝ	
*ṣ*M	shM	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–	G̊.init, G̊.medi [F]: *hi*, G̊I.
							G̊– 1!	G̊ 1	G̊ 1		
*č*M	crM	18A3					I–	I	J	J	
ć	ch	1834					Ć–	Ć	Ć	Ć	
j		1835					I–	I	J	J	
*ẑ*M	zhM	18A4					Ŝ̊–	Ŝ̊	Ŝ̊	Ŝ̊	
							Ŝ̊2–1!	Ŝ̊2 1!	Ŝ̊2 1!		
*z*M	zM	18A5					Ŝ–	Ŝ	Ŝ	Ŝ	
ʔ	sbm	1807					A2–	A2	A	–A–	
*ɬ*M	lhM	18AA					L̊–	L̊	L̊	–L̊–	

8.2 Shaping process 变形流程

Table 30 Manchu Alikali 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	*ġ ķ g k gʰ*M *ŋ*M *ŋ cʜ*M *c z jʰ*M *ţ*M *č ž ḍʰ*M *ṇ*M *t*M *t d dʰ*M *n p*M *p b bʰ*M *m y r l w ś ṣ*M *s h č*M *ć j ẑ*M *z*M *ʔ ɬ*M

Table 31 Manchu Alikali 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 $c_{\mathrm{M}}/c/z$:	`Marked`
	e u	if follows $t_{\mathrm{M}}/t/d/d^{h}{}_{\mathrm{M}}/g/k/g^{h}{}_{\mathrm{M}}/h$:	`Feminine`
	e	if follows η_{M}:	`Feminine`
	n	if precedes a *vowel*:	`Onset`
		else if precedes a *consonant*:	`Devsger`
	g k 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_{M} *t d* $d^{h}{}_{\mathrm{M}}$	if precedes *a*/*i*/*o*:	`Masculine_Onset`
		else if precedes *e*/*u*/*ü*:	`Feminine`
	t	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 K̇/K2/B̆/P̧/B/B̊:	`Post_Bowed`
	e u	if follows a (bowed) written unit Ġ/G/G̊/Ḣ/Ḣ2/H̊/H̊2/N̥/B̆/P̧/B/B̊:	`Post_Bowed`

Table 32 Manchu Alikali shaping: lookups
表 32 满文阿礼嘎礼成形过程：条件变形

Phon. let.		Unic.	Lookup	Written forms								Comments
a		1820	`Default`					AA	AA	A	A	
			`Pos_Bow`								Á	
$\bar{a}_{\mathrm{M}}$	aa_{M}	1887	`Default`					AĀ	AA	A	Ā	
e		185D	`Default`					A	A	Ȧ	Ȧ	Post_Bowed step keeps the dotted or dotless status.
			`Feminine`							A	A	
			`Pos_Bow`								Ȧ́	
											Á	
i		1873	`Default`					AI	AI	I	I	
			`Marked`								Ï	

			Devsger							II		
ï	ii	185F	Default					–Ă–	–Ă–	Ă	Ă	
o		1823	Default					AO	AO	O	U	
			Marked								O	
			Pos_Bow								O	
u		1860	Default					AȮ	AȮ	Ȯ	Ů	Feminine step removes the dots in written units. Post_Bowed step keeps the dotted or dotless statuss.
			Marked								Ȯ	
			Feminine							O	U	
											O	
			Pos_Bow								Ȯ	
											O	
ü	ue	1861	Default					AÜ	AOI	OI	Ü	
ġ	gh	186C	Default					Ḳ–	Ḳ	Ḳ	Ḳ	
ḳ	kh	183A	Default					K2–	K2	K2	–K2–	
g		1864	Default					Ḣ–	Ḣ	Ḣ	–Ḣ–	
			Mas_Ons						Ḣ	Ḣ		
			Feminine						Ġ	Ġ		
k		1874	Default					H–	G	Ḧ	Ḧ	
			Feminine						G	G	G4	
			Mas_Dev							Ḧ	Ḧ	
$g^h{}_M$	gh_M	189A	Default					Ñ–	Ñ	Ñ	–Ñ–	
$ŋ_M$	ng_M	189B	Default					Ṇ–	Ṇ	Ṇ	–Ṇ–	
ŋ	ng	1829	Default					–AG–	–AG–	AG	AG	
c_M	c_M	189C	Default					Ç̊–	Ç̊	Ç̊	–Ç̊–	
c		186E	Default					Ç–	Ç	Ç	–Ç–	
z		186F	Default					Ẓ–	Ẓ	Ẓ	–Ẓ–	
$j^h{}_M$	jh_M	189D	Default					Ẓ̊–	Ẓ̊	Ẓ̊	–Ẓ̊–	
$ṭ_M$	tr_M	189E	Default					Å–	Å	Å	Å	

č	cr	1871	Default					Č–	Č	Č	Č
ž	zr	1877	Default					Ǐ–	Ǐ	J̌	J̌
*ḍʰ*M	drhM	189F	Default					Ï–	Ï	J̈	J̈
*ṇ*M	nrM	188F	Default					Ṷ–	Ṷ	Ṷ	Ṷ
*t*M	tM	18A0	Default					Ḍ–	Ḍ	Ḍ	Ḍ
			Mas_Ons						Ḍ	Ḍ	
			Feminine						Ḍ	Ḍ	
t		1868	Default					T̲–	T̲	D̲	Ð
			Mas_Ons						T	D	
			Feminine						T̲	D̲	
			Devsger							Ð	
d		1869	Default					Ṫ–	Ṫ	Ḋ	–Ḋ–
			Mas_Ons						Ṫ	Ḋ	
			Feminine						Ṯ	Ḍ	
*dʰ*M	dhM	18A1	Default					T̈–	T̈	D̈	–D̈–
			Mas_Ons						T̈	D̈	
			Feminine						T̤	D̤	
n		1828	Default					N–	N	A	A
			Onset							N	
			Devsger						A	A	
*p*M	pM	1892	Default					Ǎ–	Ǎ	Ǎ	Ǎ
p		1866	Default					Ṗ–	Ṗ	Ṗ	–Ṗ–
b		182A	Default					B–	B	B	B
*bʰ*M	bhM	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

w		1838	Default					W–	W	W	W
ś	sh	1867	Default					Ŝ–	Ŝ	Ŝ	Ŝ
ṣ$_{M}$	sh$_{M}$	18A2	Default					S̈–	S̈	S̈	S̈
s		1830	Default					S–	S	S	S
			Devsger								S$_{3}$
h		1865	Default					H̊–	H̊	H̊	–H̊–
			Mas_Ons						H̊	H̊	
			Feminine						G̊	G̊	
č$_{M}$	cr$_{M}$	18A3	Default					I–	I	J	J
ć	ch	1834	Default					Ć–	Ć	Ć	Ć
j		1835	Default					I–	I	J	J
ẑ$_{M}$	zh$_{M}$	18A4	Default					Š–	Š	Š	Š
z$_{M}$	z$_{M}$	18A5	Default					Š–	Š	Š	Š
ʔ	sbm	1807	Default					A2–	A2	A	–A–
ɬ$_{M}$	lh$_{M}$	18AA	Default					L̊–	L̊	L̊	–L̊–

8.3 Reference 参考文献

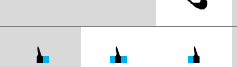
- (1929). 同文韻統. 上海涵芬楼影印清乾隆内府刻本.
- (1929). 漢滿蒙藏四體合璧大藏全咒. 上海涵芬楼影印清乾隆内府刻本.
- (1957). 御制五体清文鉴. 北京: 民族出版社.
- 郑维喆 / Weizhe ZHENG, 王奕桦 / Yihua WANG (2017). *Glyph corrections for Ali Gali final forms in the Mongolian block*. L2/17–115R.
- 郑维喆 / Weizhe ZHENG (2016). *Preliminary proposal to encode four letters 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.
- 格桑居冕 / [Tibetan] སྐལ་བཟང་འགྱུར་མེད་ (*skal bzaŋ ᵑgyur med*) (1991). “藏文字性法与古藏语音系”. In 民族语文, No. 6.

9 Todo Alikali 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 Alikali 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									
Ẉ	Wn									

D̃	Dv										
Ṭ	Tp										
Ḍ	Dp										
Ƀ	Bh										
Ƥ	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 Alikali phonetic letters

表 34 托忒文阿礼嘎礼音位字母

Phon. let. 音位字母	Unic. 码位		Written forms 书写形式						Comments 备注
a	1820	✓				AA 3	AA 2	A	A 2	

							A 1!		AA 1!	Á 1	
e		1844					AE	AE	E	E	
									AE 1!		
i		1845					AI3	AÎ	Î 2	I3 2	
								AI 1!	AÎ 1	Î 1	
									I 3!		
o		1846					AO̱	AO̱	O̱	O̱	
									AO̱ 1!		
ü	ue	1849					AU	AO	O	U 2	
							AO 1!		AO 1!	O 1	
ł	lvs	1843					Ł	Ł	Ł	Ł	
k_T	k_T	183A					K2–	K2	K2	–K2–	
k^h_T	kh_T	183B					K–	K	K	–K–	
h_T	h_T	183E					AĦ–	AĦ	Ħ	Ħ	Ħ.medi [P]: *bha* (Sanskrit b^ha), BĦÁ.
g_T	g_T	1889					G–	G	G	–G–	
g		184E					H̥–	H̥ 2	H̥ 3	Ĥ	
							G– 1!	G 1	Ĥ 1		
								Ĥ 3!	G 2		
ŋ	ng	184A					–AG–	–AG–	AG	AG2	
$ŋ_T$	ng_T	188A					NG–	NG	NG	–NG–	
j		1853					Ĉ–	Ĉ	Ĉ	–Ĉ–	
c^h_T	ch_T	183C					C–	C	C	–C–	
j_T	j_T	185C					Ź– 1!	Ź 1!	Ź 1!	–Ź–	
ñ	ny	185B					Ñ–	Ñ	Ñ	–Ñ–	
$ṭ_T$	tr_T	188C					Ž–	Ž	Ž	–Ž–	
$ṭ^h_T$	trh_T	188D					Č2–	Č2	Č2	–Č2–	
$ḍ_T$	dr_T	188E					D̰2– 1!	D̰2 1!	D̰2 1!	–D̰2–	
$ṇ_T$	nr_T	188F					W̥–	W̥	W̥	–W̥–	

*t*ᴛ	tᴛ	1890					Ď–	Ď	Ď	Ď
t		1850					Ṭ–	Ṭ	Ṭ	–Ṭ–
d		1851					Ḍ–	Ḍ	Ḍ	–Ḍ–
n		1828					N–	N	A 2	A 2
									N 1	
*p*ᴛ	pᴛ	1892					Ƀ– 1!	Ƀ 1!	Ƀ 1!	–Ƀ–
*pʰ*ᴛ	phᴛ	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–
*y*ᴛ	yᴛ	18A7					–Ɣ–	–Ɣ–	Ɣ	Ɣ
										Ī 1!
r		1837					R–	R	R	R
l		182F					L–	L	L	L
w		1856					Ẉ–	Ẉ	Ẉ	–Ẉ–
*w̄*ᴛ	waᴛ	18A6					–W̄	–W̄	–W̄	W̄
ś	sh	1831					Ś–	Ś 2	Ś 2	Ś
*ṣ*ᴛ	shᴛ	1894					S̈–	S̈	S̈	–S̈–
s		1830					S–	S	S	S
ħ	hh	1859					AH–	AH	AH	–AH–
									Ħ 1!	
c		1854					J–	J	J	–J–
z		1834					Ć–	Ć	Ć	–Ć–
*ž*ᴛ	zhᴛ	1899					Ŕ–	Ŕ	Ŕ	–Ŕ–
*z*ᴛ	zᴛ	1896					Ż–	Ż	Ż	–Ż–
*ʔ*ᴛ	qᴛ	1897					Q–	Q	Q	–Q–
ɬ	lh	1840					LH–	LH	LH	–LH–

9.2 Shaping process 变形流程

Table 35 Todo Alikali 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 ü*
neuter vowel	*i*
vowel	{*masculine vowel*} {*feminine vowel*} {*neuter vowel*}
consonant	k_T k^h_T h_T g_T *g ŋ* $ŋ_T$ *j* c^h_T j_T *ñ* t_T t^h_T d_T n_T t_T *t d n* p_T p^h_T *b m y* y_T *r l w* w_T *ś* $ş_T$ *s ħ c z* $ž_T$ z_T $ʔ_T$ *ł*

Table 36 Todo Alikali 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	*g*	**if** precedes a *masculine vowel*:	`Masculine_Onset`
		else if precedes a *feminine vowel* or *neuter vowel*:	`Feminine`
		else if follows a *vowel*:	`Masculine_Devsger`
5. Post-bowed	*a*	**if** follows an initial *consonant*:	`Post_Bowed`
	i o ü	**if** follows a (bowed) written unit K2/K/G/Ḃ/Ṕ/B:	`Post_Bowed`

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

Phon. let.	Unic.	Lookup	Written forms								Comments
a	1820	`Default`					A	AA	A	A	
		`Pos_Bow`								Á	
e	1844	`Default`					AE	AE	E	E	
i	1845	`Default`					AI3	AÎ	Î	I3	
		`Pos_Bow`								Î	
o	1846	`Default`					AǪ	AǪ	Ǫ	Ǫ	
ü	1849	`Default`					AU	AO	O	U	
		`Pos_Bow`								O	
ł	lvs 1843	`Default`					Ł	Ł	Ł	Ł	
k_T	k_T 183A	`Default`					K2–	K2	K2	–K2–	

*kʰ*T	khT	183B	Default					K–	K	K	–K–
*h*T	hT	183E	Default					AH–	AH	H	H
*g*T	gT	1889	Default					G–	G	G	–G–
g		184E	Default					Ḥ–	Ḥ	Ḥ	Ĥ
			Mas_Ons						Ḥ	Ḥ	
			Feminine						G	G	
			Mas_Dev							Ĥ	
ŋ	ng	184A	Default					Ḥ–	Ḥ	Ḥ	Ĥ
*ŋ*T	ngT	188A	Default					NG–	NG	NG	AG2
j		1853	Default					Ĉ–	Ĉ	Ĉ	–Ĉ–
*cʰ*T	chT	183C	Default					C–	C	C	–C–
*j*T	jT	185C	Default					Ź–	Ź	Ź	–Ź–
ñ	ny	185B	Default					Ñ–	Ñ	Ñ	–Ñ–
*ṭ*T	trT	188C	Default					Ž–	Ž	Ž	–Ž–
*ṭʰ*T	trhT	188D	Default					Č–	Č	Č	–Č–
*ḍ*T	drT	188E	Default					Ḍ–	Ḍ	Ḍ	–Ḍ–
*ṇ*T	nrT	188F	Default					Ṿ–	Ṿ	Ṿ	–Ṿ–
*t*T	tT	1890	Default					D–	D	D	D
t		1850	Default					Ṭ–	Ṭ	Ṭ	–Ṭ–
d		1851	Default					Ḍ–	Ḍ	Ḍ	–Ḍ–
n		1828	Default					N–	N	N	A
*p*T	pT	1892	Default					Ɓ–	Ɓ	Ɓ	–Ɓ–
*pʰ*T	phT	184C	Default					Ƥ–	Ƥ	Ƥ	–Ƥ–
b		184B	Default					B–	B	B	B2
m		184F	Default					M–	M	M	M2
y		1855	Default					I–	I	I	–I–
*y*T	yT	18A7	Default					–Ɣ–	–Ɣ–	Ɣ	Ɣ
r		1837	Default					R–	R	R	R

l		182F	Default					L–	L	L	L
w		1856	Default					Ẉ–	Ẉ	Ẉ	–Ẉ–
$\bar{w}_T$	wa_T	18A6	Default					–W2	–W2	–W2	W2
ś	sh	1831	Default					Ś–	Ś	Ś	Ś
$ṣ_T$	sh_T	1894	Default					S̈–	S̈	S̈	–S̈–
s		1830	Default					S–	S	S	S
ħ	hh	1859	Default					AH–	AH	AH	–H–
c		1854	Default					J–	J	J	–J–
z		1834	Default					Ć–	Ć	Ć	–Ć–
$ž_T$	zh_T	1899	Default					Ŕ–	Ŕ	Ŕ	–Ŕ–
z_T	z_T	1896	Default					Ż–	Ż	Ż	–Ż–
$ʔ_T$	q_T	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 SEPARATE 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—2023. This document contains all the written forms of all the letters in GB/T 25914—2023 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 *ɬ*. This presentation form consists of a chachlag and a long vowel sign, and is part of the ligatures like *baɬ*, *paɬ* 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 Alikali *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. GB/T 36645—2018 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.

10.5 Hudum Alikali writing system

Comparison with Ćoijuŋjab (2013). Ćoijuŋjab (2013) introduces MONGOLIAN LETTER ALI-GALI A SECOND ISOLATE FORM for the letter *a*. This presentation form is a stylistic toothed variant of the isolate A, so this written form is not included in the Hudum Alikali part of this document.

Ćoijuŋjab (2013) introduces MONGOLIAN LETTER ALI-GALI A FINAL FORWARD FORM for the letter a_{H}. This presentation form is used in the Hudum Alikali ligature, and is equivalent to the final A of letter *a*, so this written form is not included in the Hudum Alikali part of this document. For the same reason, both MONGOLIAN LETTER ALI-GALI NGA INITIAL SHORT FORM and MONGOLIAN LETTER ALI-GALI NGA MEDIAL SHORT FORM are not included either.

Comparison with GB/T 25914—2010. For the same reasons as in comparison with Ćoijuŋjab (2013) in section 10.5, the MONGOLIAN LETTER ALI-GALI A SECOND ISOLATE FORM, MONGOLIAN LETTER ALI-GALI A FINAL FORWARD FORM, MONGOLIAN LETTER ALI-GALI NGA INITIAL SHORT FORM, and MONGOLIAN LETTER ALI-GALI NGA MEDIAL SHORT FORM are not included in the Hudum Alikali part of this document.

Comparison with GB/T 25914—2023. For the same reasons as in comparison with Ćoijuŋjab (2013) in section 10.5, the MONGOLIAN LETTER ALI-GALI A SECOND ISOLATE FORM and MONGOLIAN LETTER ALI-GALI A THIRD FINAL FORM are not included in the Hudum Alikali part of this document.

Since the shaping step for MONGOLIAN LETTER ALI-GALI A SECOND FINAL FORM is quite clear, this document adds the shaping step for this form.

10.6 Manchu Alikali writing system

Comparison with Ćoijuŋjab (2013). Ćoijuŋjab (2013) does not add the feminine forms for the letters t_{M} and $d^{h}{}_{\mathrm{M}}$, this document adds the feminine forms for the letters t_{M} and $d^{h}{}_{\mathrm{M}}$.

Ćoijuŋjab (2013) introduces MONGOLIAN LETTER ALI-GALI I SECOND FINAL FORM. This presentation form is equivalent to the final ï of letter *i*, so this written form is not included in the Manchu Alikali part of this

document.

Comparison with GB/T 36645—2018. For the same reasons as in comparison with Ćoijuŋjab (2013) in section 10.6, the MONGOLIAN LETTER ALI-GALI I SECOND FINAL FORM is not included in the Manchu Alikali part of this document.

10.7 Todo Alikali writing system

Comparison with Ćoijuŋjab (2013). Ćoijuŋjab (2013) introduces MONGOLIAN LETTER TODO ALI-GALI I TA. This letter is not used in the Todo Alikali, thus not included in the Todo Alikali part of this document.

Comparison with GB/T 36649—2018. GB/T 36649—2018 introduces MONGOLIAN LETTER TODO GA SECOND FINAL FORM. This presentation form is used as the initial N of letter *n*, but since it is analyzed as the letter *g*, this document analyzes it as the initial ĥ of letter *g*.

GB/T 36649—2018 introduces MONGOLIAN LETTER TODO GA THIRD FINAL FORM. This presentation form is the stylistic vairant of the MONGOLIAN LETTER TODO GA FIRST FINAL FORM, and thus is not included in the Todo Alikali part of this document.

GB/T 36649—2018 introduces MONGOLIAN LETTER ALI-GALI PA THIRD INITIAL FORM and MONGOLIAN LETTER ALI-GALI PA THIRD MEDIAL FORM. These presentation forms record different phonemes with MONGOLIAN LETTER ALI-GALI PA SECOND INITIAL FORM and MONGOLIAN LETTER ALI-GALI PA SECOND MEDIAL FORM, and thus are analyzed as the variant of $p^{h}{}_{\mathrm{T}}$ in this document.

GB/T 36649—2018 introduces MONGOLIAN LETTER ALI-GALI HALF U SECOND MEDIAL FORM. This presentation form is not used in the Todo Alikali, thus not included in the Todo Alikali part of this document.

11 Acknowledgement

This document uses *Menksoft Vran Tig* for Hudum, *Mongolian E* for Todo, and *Abkai Xanyan* for Sibe and Manchu. The font for Hudum Alikali is based on *Mongolian Universal BaoHei*, with the author of this document supplementing the glyph set. The font for Manchu Alikali is based on *Abkai Xanyan*, with the author of this document supplementing the glyph set. The font for Todo Alikali is a revival font based on the design in Хөхийн Бямбажав (2019).