



BRILL

INNER ASIA 27 (2025) 277–313



brill.com/inas

The Digitisation Odyssey of the Mongolian Script in Unicode

Tergel Teneg | ORCID: 0000-0001-6518-6094

School of the Humanities and Social Sciences, University of Cambridge,
Cambridge, UK

tt475@cam.ac.uk

Received 1 November 2024 | Accepted 10 March 2025 |

Published online

Abstract

The Mongolian script, the world's only surviving vertical alphabetic system, is written from left to right with letters joined vertically. Used in China's Inner Mongolia, it became Mongolia's co-official script alongside Cyrillic in 2025. Despite its official status, it remains one of the few living language scripts that is digitally unsearchable. No search engines or web browsers support its vertical orientation, and there is no universal standard to ensure accurate letter joining or consistent display of compatible digital content. This article explores the ethnographic dimensions of the Mongolian script's digitisation within the Unicode standard, focusing on its social, historical and political implications for language use. Rather than engaging with the intricate technicalities of Unicode encoding, the analysis highlights three key factors: state involvement and support, methodological limitations or biases in Unicode-related research, and cross-national collaboration between Mongolia and China. Through this framework, the study traces the origins of the digitisation initiative in Inner Mongolia, the subsequent adoption of the current Unicode model for the Mongolian script, and the roles played by both the Mongolian and the Chinese governments. It further examines how the sociopolitical environments and language policies of each country have shaped the contemporary status and usage of the Mongolian script and language within their respective territories. The article concludes by offering recommendations to improve the visibility, accessibility and functional integration of the Mongolian script in digital spaces.

Keywords

Mongolian script – Unicode – digitisation of the vertical writing system – Mongolia – Inner Mongolia – China

1 Introduction

The Mongolian script is the longest surviving script used by the Mongols. It is valued not only for its representation of an early form of written Mongolian, but also, more importantly, for being adopted by Chinggis Khan as the sole writing system of the Mongol state established in 1206. The script was derived from Ancient Uyghur alphabets during the eleventh and twelfth centuries (Janhunen 2012).¹ Over time, the Mongols have employed various scripts. In the thirteenth century, the Phasgpa script, adapted from the Tibetan alphabet, was introduced, but it fell out of use with the decline of the Yuan Dynasty. In the sixteenth century, new letters, known as Ali-Gali letters, were added to the Mongolian script to transcribe foreign words, specifically from Tibetan and Sanskrit. By the mid-seventeenth century, the Todo script emerged among the Western Mongols (in modern-day Xinjiang and western Mongolia). In the late seventeenth century, the Soyombo script was created by Zanabazar (1635–1723), the first Bogd Gegeen and spiritual leader of the Khalkha Mongols, primarily for religious and liturgical purposes. Its use remained largely restricted to monastic contexts. Today, its most prominent application is found in the Mongolian national flag and state emblem, where it serves as a symbol of national identity and sovereignty. In the early twentieth century, the Vagindra script was created among the Buryats, Northern Mongols (present-day Republic of Buryatia in Russia, east of Lake Baikal). During the 1930s and 1940s, Mongolians in Mongolia, Kalmyks (Western Mongols residing in Russia west of the Volga River), and Buryats transitioned to the Cyrillic script, adopting distinct regional variations. None of these scripts have achieved widespread use among all Mongolic people, with the Mongolian script being the exception.

Today, Buryats, Kalmyks and Mongols in Mongolia predominantly use the Cyrillic script for their written languages. The Mongolian script serves as a universal writing system for all Mongols, irrespective of dialectal variations, as it is morphologically based rather than phonetic. Unlike phonetically devised

1 The Ancient Uyghur script, derived from the Sogdian script, which itself traces its origins to Aramaic, was initially written from right to left, following the pattern of most Semitic scripts.

scripts, it does not correspond to any specific Mongolic spoken variant or dialect, allowing for broader linguistic applicability.

However, the future of the Mongolian script in both Mongolia and China remains uncertain. In September 2020, China's central government enacted a New Bilingual Education Policy, 内蒙古双语教育新政策 *neimenggu shuangyu jiaoyu xinzhengce*, reducing the use of Mongolian as the medium of instruction in schools in Inner Mongolia (Inner Mongolia Department of Education 2020). As a result, the Mongolian script is no longer the primary educational language in Mongolian schools in Inner Mongolia. Conversely, Mongolia is advancing the National Mongolian Script Programme III (Montsame.mn 2020), with the aim of making the Mongolian script co-official alongside Cyrillic from 2025. Written Mongolian also remains a co-official language in the Inner Mongolia Autonomous Region of China, used alongside Chinese. Ironically, despite its official status, the Mongolian script has not achieved widespread use in everyday digital interactions. It still cannot be easily searched on the internet.

One of the primary barriers to the digital viability of the Mongolian script lies in the incomplete and inconsistent implementation of the Unicode Standard. A key challenge involves the establishment of a unified implementation standard within a single rendering model to ensure consistent deployment across digital platforms. The model that is currently recognised by Unicode for rendering the Mongolian script is the phonetic model, which was introduced in 1999. However, its formal implementation standard – referred to as its specification – was not officially consolidated until the publication of Unicode Technical Note #57 (UTN #57) by Jiang in 2024, a full 24 years after the model's initial introduction. UTN #57 provides an essential framework by detailing the encoding practices, character behaviours and structural characteristics specific to the Mongolian script.

If UTN #57 is effectively adopted, it holds the potential to significantly improve the script's digital usability, and to pave the way for broader and more consistent implementation across various platforms. However, two major challenges hinder its widespread adoption. First, the fragmented nature of the digital environment must be addressed by ensuring that all relevant stakeholders, including platform developers, software vendors and user communities, adopt the Unicode Standard's phonetic model as the sole rendering framework. Currently, this is far from being realised. Second, UTN #57 must be formally accepted as the authoritative specification for implementing the phonetic model. Importantly, this step is contingent upon the prior universal acceptance of the phonetic model itself, which continues to face considerable resistance.

These implementation challenges are largely rooted in more than two decades of under-specification of the phonetic model.² The initial version of the model was marred by numerous issues, ‘including typing and display errors, visual ambiguities leading to security concerns, positional mismatches, missing glyphs, and under-documented specifications’ (Anderson 2017: 2). These deficiencies have significantly delayed the integration of the Mongolian script into major computing platforms. For instance, according to Unicode documentation published in 2017, both Microsoft and Google postponed implementing fixes for Mongolian script rendering, citing the absence of ‘a stable and universally accepted encoding model’ (Anderson 2017: 1). Consequently, the practical application of the phonetic model has been undermined by inconsistent support across digital platforms.

The sluggish adoption of the phonetic model has given rise to two persistent challenges now hampering the implementation of UTN #57. The first issue concerns the existence of unofficial implementation standards for the phonetic model. Despite ongoing efforts to standardise the phonetic model, a wide range of proposals has emerged, each offering alternative ways of applying the phonetic model (e.g. Batjargal *et al.* 2010; Bolorsoft 2019; CITA 2019; Enkhdalai 2019; Glass 2019; Khaltarkhuu & Maeda 2015; Jiang 2023; Jiang 2024; Li 2020; Liang, H. 2019a; Liang, H. 2019b; Liang, H. 2019c; Liang, H. 2020; Liang, H. 2024; Liang, J.B 2018a; Liang, J.B 2018b; Liang, J. 2019; Ma 2020; Menghejiya 2018; Namujila 2020a; Namujila 2020b; Samr.gov.cn 2010; Samr.gov.cn 2018; Sanlig *et al.* 2018; Sanlig & Enkhtur 2018; Shen 2017; Shen 2018; Shen 2019a; Shen 2019b; Surgaltu *et al.* 2020; Umamaheswaran & WG2 Secretary 2019; Wang *et al.* 2016). These include ‘the refined phonetic model’ (Bolorsoft 2019),³ which was developed by experts in Mongolia, as well as a range of approaches proposed in China (Liang, H. 2018; Liang, H. 2019a; Samr.gov.cn 2010; Samr.gov.cn 2018;

2 A common instance of under-specification arises in the context of clarifying normative rules governing the alternation behaviour of vowels and consonants in non-initial positions. For example, when the final variants of the vowels [a] and [e] are preceded by the bowed consonants [b, p, f, k, g, h], they are expected to exhibit a left-falling stroke rather than a right-falling one. These subtle variations, which were previously unaccounted for, have been addressed in UTN #57 since 2024.

3 As reported by journalist B. Batchimeg from the Mongolian National News Agency, Montsame, the refined phonetic model, known as *Tunggmal*, ‘satisfies Unicode standard halfway’ (Batchimeg 2019: para. 2). To further assess this claim, I consulted a technical expert to evaluate the implementation of the refined phonetic model, which is available at <https://www.mngl.net/> (accessed 4 July 2025). The expert confirmed that this model does not implement the specifications outlined in Unicode Technical Note #57 (UTN #57), which defines the most up-to-date shaping and rendering behaviour for the Mongolian script within the Unicode framework.

Shen 2017) that seek to enhance the phonetic model. Collectively, the Chinese proposals may be referred to as ‘the improved phonetic model’. Such variations have caused inconsistency in the representation of the Mongolian script in digital environments. The successful implementation of the phonetic model depends on all users and vendors adopting UTN #57 as the single authoritative standard.

The second issue concerns the emergence of an alternative encoding model that is fundamentally incompatible with both the phonetic model and UTN #57. In response to the persistent challenges of the phonetic model, some researchers have proposed a graphic-based model, called the graphetic model. Although this approach has never been formally adopted by Unicode, or by authorities in either Mongolia or China, it has gained a degree of usage among users and platforms as an alternative solution to ongoing problems with the phonetic model. The existence of this alternative approach has further fragmented the digital landscape.

In summary, the concurrent use of the Unicode-based phonetic model and the non-Unicode graphic-based model, combined with competing standards for the phonetic model implementation, has led to widespread inconsistency in the digital representation of the Mongolian script. Some users do not employ Unicode at all, while others who do may not conform to UTN #57. This lack of uniformity results in significant variation in how the script is rendered across platforms and applications, limiting its accessibility and usability. Consequently, the development of compatible and sustainable digital content in Mongolian script has been severely restricted.

Moreover, this lack of digital compatibility has discouraged the Mongolian-language community from producing Unicode-encoded contents. In turn, web search engines struggle to index sufficient and consistent data pertaining to the Mongolian script, thereby hampering their ability to return relevant results to user queries. Even though major platforms such as Microsoft and Google have declared support for Unicode-based Mongolian script, the presence of alternative solutions and the absence of interoperability among them have hindered the full integration of the script into digital domains. This also complicates digital pedagogy, as the Mongolian script often fails to render correctly, impeding educational efforts.

This article examines the historical and contemporary digitisation journey of the Mongolian script within the Unicode framework, aiming to illuminate the sociopolitical and historical factors that influence its status in the digital domain. By tracing this trajectory, the analysis seeks to uncover the complexities surrounding the script's integration into digital technologies, and its implications for users and stakeholders.

2 The Mongolian Script Unicode

Unicode is an international character-encoding standard that ensures consistent coding for the world's writing systems. It serves as the standard for base computing platforms such as Windows, MacOS and Linux, as well as for applications such as input tools. If a script is not encoded in the Unicode Standard, it will not be possible to accurately interchange data on any digital platform. The journey of the Mongolian script within Unicode began in 1987 (Unicode 2024a). The Mongolian alphabet was formally registered in Unicode under the phonetic model in 1999 (ISO 10646) (Figure 1).⁴ This model was developed at the Mongolian Studies Centre of Inner Mongolia University by Quejingzhabu, and it was mutually endorsed by Mongolian authorities through the Users Implementation Agreement on Mongolian Encoding, 蒙古文编码系统实现用户协定 *mengguwen bian ma yonghu xieding* (Quejingzhabu 1999).

Figure 1 illustrates the Mongolian Block, the current Unicode Standard for Mongolian script. Each Mongolian letter is assigned a unique codepoint, reflecting its phonetic value, as illustrated. According to Unicode, the phonetic model for the Mongolian script 'is somewhat different from that for any other script within Unicode, and in many respects, it is the most complicated' (Unicode 2006: 534). The complexity of this model arises from the nature of the Mongolian script, where letters are vertically connected to syllables with variations in their initial, medial and isolated forms. For instance, the letter [a] has four variants: initial 'ᠠ' [a], medial 'ᠡ' [a], final 'ᠢ' [a] or 'ᠣ' [a], and isolated 'ᠣ' [a], depending on its position within a word. Many letters and their variants are graphically identical, such as the letters 'ᠢ' [ɔ] and 'ᠣ' [ɔ], which share the same shape, and the final variant of 'ᠢ' [a], which is identical to the final form of 'ᠣ' [e]. As a result, multiple codepoints correspond to the same shape under the phonetic model. Therefore, establishing standard shaping phases and steps (shaping rules) for the Mongolian script in Unicode is essential for the universal implementation of the phonetic model. UTN #57 is specifically designed to define and standardise these shaping rules for Mongolian.

UTN #57 is based on two pivotal documents that form the foundation of China's national standard for rendering the Mongolian script. These documents are:

- (1) *Text Representation and Shaping Manual of the Mongolian Script* (Jiang 2023)

4 The complete Mongolian Range is available at: <https://www.unicode.org/charts/PDF/U1800.pdf> (accessed 4 July 2025).

- (2) *China's National Standard GB/T 25914-2023: Traditional Mongolian Nominal Characters, Presentation Characters and Use Rules of Controlling Characters* (Samr.gov.cn 2024).

According to Jiang, 'UTN #57 will formally define modular, hierarchical, unambiguous text shaping rules, implement the requirements of national standards, and provide font generation tools and font testing tools' (Jiang 2024: 1). This initiative aims to create a robust framework for the accurate representation of Mongolian in Unicode.

UTN #57 represents a fully developed model for Mongolian text shaping, with complete coding and thorough testing implemented. The testing has utilised two distinct fonts: Noto Sans and Mongolian Baiti. The Mongolian Baiti font is currently undergoing a review process, while the re-engineering of Noto Sans Mongolian has been completed, and all associated pull requests have been addressed.⁵ Further work remains to be completed on UTN #57, specifically in the area of documenting the specification for numbers and punctuation. However, the toolchain for font generation is fully functional.

2.1 *Challenges in Implementing UTN #57*

As previously discussed, the prolonged delay in finalising a universal standard for the shaping rules of the Mongolian script within the Unicode framework has compelled major software vendors to develop their own divergent solutions. This situation has resulted in the emergence of independently devised shaping rules, and the adoption of an alternative rendering model known as the graphetic model, which functions outside the scope of Unicode.

Current efforts aim to standardise the phonetic model by advocating for Unicode UTN #57 as the definitive reference for implementation, particularly among users who have adopted the phonetic model but who continue to rely on their own shaping rules. In contrast, facilitating the adoption of UTN #57 among users of the graphetic model poses significantly greater challenges, on which the following section will focus.

While many vendors initially attempted to render the Mongolian script within the Unicode system, the graphetic model gradually gained prominence as a non-Unicode encoding paradigm. Unlike the Unicode-compliant phonetic model, which encodes letters based on their phonological values, the graphetic model encodes characters based on their visual shapes. This structural difference results in a single codepoint being used to represent multiple phonetic values, thereby eliminating duplicated codepoints for graphically

⁵ The underlying code utilised for both fonts is available under open access at the following repository: <https://github.com/kushim-jiang/Mongolian/releases> (accessed 4 July 2025).

1800

Mongolian

18AF

	180	181	182	183	184	185	186	187	188	189	18A
0	ᠠ	ᠡ	ᠢ	ᠣ	ᠤ	ᠥ	ᠦ	ᠨ	ᠨ	ᠢ	ᠣ
1	ᠠᠠ	ᠡᠡ	ᠢᠢ	ᠣᠣ	ᠤᠤ	ᠥᠥ	ᠦᠦ	ᠨᠨ	ᠨᠨ	ᠢᠢ	ᠣᠣ
2	ᠠᠢ	ᠡᠢ	ᠢᠣ	ᠣᠣ	ᠤᠣ	ᠥᠣ	ᠦᠣ	ᠨᠣ	ᠨᠣ	ᠢᠣ	ᠣᠣ
3	ᠠᠣ	ᠡᠣ	ᠢᠣ	ᠣᠣ	ᠤᠣ	ᠥᠣ	ᠦᠣ	ᠨᠣ	ᠨᠣ	ᠢᠣ	ᠣᠣ
4	ᠠᠣ	ᠡᠣ	ᠢᠣ	ᠣᠣ	ᠤᠣ	ᠥᠣ	ᠦᠣ	ᠨᠣ	ᠨᠣ	ᠢᠣ	ᠣᠣ
5	ᠠᠣ	ᠡᠣ	ᠢᠣ	ᠣᠣ	ᠤᠣ	ᠥᠣ	ᠦᠣ	ᠨᠣ	ᠨᠣ	ᠢᠣ	ᠣᠣ
6	ᠠᠣ	ᠡᠣ	ᠢᠣ	ᠣᠣ	ᠤᠣ	ᠥᠣ	ᠦᠣ	ᠨᠣ	ᠨᠣ	ᠢᠣ	ᠣᠣ
7	ᠠᠣ	ᠡᠣ	ᠢᠣ	ᠣᠣ	ᠤᠣ	ᠥᠣ	ᠦᠣ	ᠨᠣ	ᠨᠣ	ᠢᠣ	ᠣᠣ
8	ᠠᠣ	ᠡᠣ	ᠢᠣ	ᠣᠣ	ᠤᠣ	ᠥᠣ	ᠦᠣ	ᠨᠣ	ᠨᠣ	ᠢᠣ	ᠣᠣ
9	ᠠᠣ	ᠡᠣ	ᠢᠣ	ᠣᠣ	ᠤᠣ	ᠥᠣ	ᠦᠣ	ᠨᠣ	ᠨᠣ	ᠢᠣ	ᠣᠣ
A	ᠠᠣ	ᠡᠣ	ᠢᠣ	ᠣᠣ	ᠤᠣ	ᠥᠣ	ᠦᠣ	ᠨᠣ	ᠨᠣ	ᠢᠣ	ᠣᠣ
B	ᠠᠣ	ᠡᠣ	ᠢᠣ	ᠣᠣ	ᠤᠣ	ᠥᠣ	ᠦᠣ	ᠨᠣ	ᠨᠣ	ᠢᠣ	ᠣᠣ
C	ᠠᠣ	ᠡᠣ	ᠢᠣ	ᠣᠣ	ᠤᠣ	ᠥᠣ	ᠦᠣ	ᠨᠣ	ᠨᠣ	ᠢᠣ	ᠣᠣ
D	ᠠᠣ	ᠡᠣ	ᠢᠣ	ᠣᠣ	ᠤᠣ	ᠥᠣ	ᠦᠣ	ᠨᠣ	ᠨᠣ	ᠢᠣ	ᠣᠣ
E	ᠠᠣ	ᠡᠣ	ᠢᠣ	ᠣᠣ	ᠤᠣ	ᠥᠣ	ᠦᠣ	ᠨᠣ	ᠨᠣ	ᠢᠣ	ᠣᠣ
F	ᠠᠣ	ᠡᠣ	ᠢᠣ	ᠣᠣ	ᠤᠣ	ᠥᠣ	ᠦᠣ	ᠨᠣ	ᠨᠣ	ᠢᠣ	ᠣᠣ

The Unicode Standard 12.1, Copyright © 1991-2019 Unicode, Inc. All rights reserved.

FIGURE 1 Mongolian Block: 1800-18AF The Unicode Standard, Version 16.0 (Unicode 2024c)

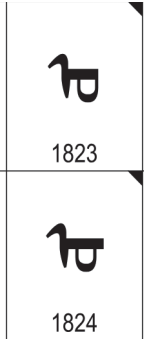


FIGURE 2
 ᠮ [ᠵ]1823 and ᠮ [ᠮ]1824 (Unicode 2024c)

identical characters. For instance, the characters representing the sounds ᠮ [ᠵ] and ᠮ [ᠮ], which are assigned distinct codepoints in Unicode (U+1823 and U+1824 respectively), are given a single codepoint in the graphetic system (see Figure 2), thereby linking one shape to multiple phonemes.

However, this model presents substantial obstacles to the implementation of Unicode's phonetic model and the uptake of UTN #57. The coexistence of competing encoding systems and divergent rendering practices has led to systemic incompatibility across platforms. As noted by Liang, 'various vendors have since struggled to make their own senses of the characters, while users have been suffering from both the poor interoperability between vendor implementations and the consequent lack of native support on major platforms' (Liang, H. 2019c: 1). This fragmentation has had substantial negative consequences for user experience, resulting in inconsistent rendering and reduced accessibility.

Historically, the graphetic model predates the phonetic model (News.MN 2017). It originated in the 1930s as a practical Mongolian printing technique (metal font type for the Mongolian typewriter), and it was later adapted in the 1980s for use in computer systems, including early DOS-based applications such as the Hua Guang and Feng Zheng publishing systems. All of these early implementations treated font glyphs as encoding units, a practice later termed 'graphetic encoding' (Jirimutu 2018: 6). The earliest formal effort to encode Mongolian script within the Unicode framework was made in 1987, through the proposal of the 'Mongolian 7-bit and 8-bit Coded Graphic Character Sets' led by Huang Wei Min and the Chinese Centre for Computer and Information Research and Development (Huang Wei Min & China 1987).⁶ While this

6 China's Centre for Computer and Information Research and Development (CCID) 中国电子信息产业发展研究院 (*zhongguo dianzi xinsi chanye fazhan yanjiuyuan*) is affiliated under China's Ministry of Industry and Information Technology (MIIT) 工业和信息化部 (*gongye he xinxihua bu*).

proposal marked an initial engagement with the graphetic approach, it did not result in formal standardisation.

Contemporary systems such as Saiyin and Z Mongol continue to implement graphetic encoding.⁷ However, despite its technical viability and broad use, the graphetic model has never received official endorsement from either the Chinese government or international standardisation bodies such as Unicode. Instead, it has primarily been supported by individual scholars and technical experts, including Liang Hai, Jirimutu and Chuck Namujila, who presented arguments in favour of the graphetic model in various professional contexts between 2017 and 2020 (Jirumutu 2018; Liang, H. 2018; Liang, J.B. 2018b; Namujila 2020a; Namujila 2020b; Script Ad Hoc Group Recommendations 2017; Shen 2018).⁸ Their opinions are best understood as personal or academic perspectives, and they do not reflect an official policy stance by the Chinese government.

When the graphetic model was introduced to the Unicode Consortium during the 2018 WG2 Ad Hoc Meeting in Hohhot, technical experts emphasised its clarity and user-friendliness, positioning it as a model that avoided the ambiguity and instability associated with the phonetic approach (Liang, H. 2018; Sanlig & Togoobat 2018a).⁹ Jirimutu (2018: 1) asserted that, from a purely encoding perspective, the graphetic model was ‘absolutely correct’.

One particular domain where the graphetic model demonstrates a clear technical advantage is optical character recognition (OCR). Pan *et al.* (2023) report that Unicode’s phonetic encoding struggles to represent distinct glyphs effectively, thereby hindering handwritten script recognition (Pan *et al.* 2023). In contrast, their study shows that a 51-element grapheme code set, based on graphetic principles, significantly improves OCR performance for handwritten Mongolian script (Pan *et al.* 2023).¹⁰

7 Another widely used piece of software, MenkSoft, which is currently employed in Mongolian research-related departments at Chinese universities, local government offices and some press agencies, previously relied on the graphetic model, but it has since transitioned to the phonetic model.

8 The technical term for Chuck Namujila’s proposed approach is the glyphic model, a divergent form of the graphetic model, both of which are grounded in grapheme-based encoding principles (Namujila 2020b).

9 The graphetic approach offers a ‘cleaner, unambiguous representation of text and a simpler user experience’ (Sanlig & Togoobat 2018a: 8). It also avoids the ‘ambiguity and uncertainty inherent to the phonetic model’ (Liang, H. 2018: 3).

10 ‘Because Mongolian Unicode encoding cannot represent unique glyphs, a Mongolian grapheme code set containing 51 elements was proposed by Fan *et al.* Compared to Unicode, grapheme codes have shown better performance on Mongolian handwritten recognition tasks’ (Pan *et al.* 2023: 3–4).

Nonetheless, determining the superiority of either model is not straightforward. The phonetic model assigns multiple codepoints to a single visual form (based on pronunciation), while the graphetic model assigns a single codepoint to multiple phonemes (based on visual form). While the graphetic approach may perform better in visual recognition tasks, it is less effective in phonological representation and parsing.

Despite the graphetic model's strengths, it has faced considerable opposition from representatives of Mongolia and members of the Mongolian Working Group (MWG). These stakeholders, who favoured the phonetic model, contended that the graphetic model failed to adequately represent the phonological structure and syllable-joining rules intrinsic to the Mongolian script (Anderson 2017; Anderson 2020; Moore 2018; Moore 2019).

The Unicode Technical Committee (UTC) has also reaffirmed its commitment to the phonetic model. It emphasised that this model had been unanimously agreed upon during MWG meetings in 2017 and 2018 (Anderson 2020). Furthermore, the UTC invoked the Unicode Encoding Stability Policy, which prohibits the removal or redefinition of existing characters (Anderson 2020), thereby preserving the integrity of the Unicode standard. These procedural safeguards effectively preclude any retroactive transition to the graphetic model, regardless of its technical merits or subsequent scholarly support.

2.2 *Current Status of the Mongolian Script Unicode*

A broad consensus has been established among Unicode and key stakeholders, including representatives from Mongolia, Inner Mongolia (China), the Chinese National Body, global Unicode experts, software implementers and end users, that the phonetic model will remain the standard framework for encoding the Mongolian script (Anderson 2020; Moore 2019). During the UTC meeting in 2018, participants collectively agreed that rather than discarding the phonetic model, efforts should be directed toward creating a stable and unified implementation environment. It was recognised that consistent text representation and encoding would only be achievable if the same set of standards and rules were adopted uniformly across Mongolia, Inner Mongolia and the global digital community (Moore 2019).¹¹ The decision to refine and stabilise the existing phonetic model reflects Unicode's aim to deliver a durable, consensus-based

11 'Agreement was reached that the phonetic model should not be abandoned, but the focus needed to be on creating a stable implementation environment. In order to have stable encoding and text representation for Mongolian, all participants in the working group agreed that the same set of standards and rules should be used in Mongolia, Inner Mongolia of China, and the rest of the world' (Moore 2019: 1).

solution for Mongolian script implementation (Anderson 2020).¹² Looking ahead, the central challenge lies in securing an international consensus on the practical adoption of the phonetic model and its implementation standard, UTN #57, to ensure long-term consistency and cross-platform compatibility.

UTN #57 currently serves as the national standard of China, and it has also been formally integrated into the Unicode Standard, forming the foundational basis for Chapter 13.5: Mongolian in the Unicode Core Specification (Unicode 2024d). This chapter offers a comprehensive articulation of the glyph set and shaping rules, not only for the Mongolian script, but also for six other vertically written scripts: Todo, Sibe, Manchu, Hudum Alikali, Manchu Alikali and Todo Alikali (Unicode 2024d). Through its integration into the Unicode core specification, UTN #57 has acquired authoritative status in global standardisation efforts for these scripts.

Importantly, the Unicode Consortium maintains a policy of neutrality regarding political considerations. The question of whether UTN #57 aligns more closely with China's national standard or with that of Mongolia falls outside the Consortium's remit. Rather, UTN #57 is regarded solely as a technical specification that is valued for providing the most current and functionally viable solution for the rendering and implementation of the Mongolian script within the Unicode framework.

A notable fact is that Mongolia has yet to issue an official national standard for the Mongolian script. This places the country at a critical juncture: it may either adopt standards aligned with those of China, or it may pursue the creation of an independent standard, which would necessitate bilateral negotiations to achieve a mutually acceptable framework. Consequently, despite the formal recognition of the Mongolian script as a co-official state script as of 1 January 2025, and the shared agreement between Mongolia and China to uphold the phonetic model, the practical implementation of UTN #57 remains both problematic and unresolved within the broader Unicode ecosystem.

This article examines the sociopolitical factors underlying this implementation stalemate. The analysis contends that the core challenges are not primarily technical in nature, but stem from deficiencies in interdisciplinary collaboration, institutional alignment, and the availability of financial and

12 'The Unicode Technical Committee has made the decision to improve the phonetic model, based on response from the standards bodies in China, Inner Mongolia of China, and Mongolia, the local community of Mongolian users, and various publishers, though Unicode had earlier voiced a preference for a graphetic approach. The goal of Unicode is to provide a stable solution for implementing Mongolian that has been reached by consensus from all the stakeholders' (Anderson 2020: 2).

intellectual resources. These structural and relational dynamics constitute the foundation of the ongoing difficulties, and they will be explored in greater detail in the following sections.

The article identifies four key factors that have negatively impacted the digitisation of the Mongolian script:

- (1) The unstable language ideologies in Mongolia, which have undermined public confidence in the Mongolian script. The introduction of Cyrillic in the 1940s has perpetuated attitudes that devalue the Mongolian script, affecting its reintroduction in Mongolia.
- (2) The central language policies of both China and Mongolia are not primarily or exclusively oriented towards ensuring the sustainability of the Mongolian language in China or the Mongolian script in Mongolia, respectively. In particular, the legal enforceability and practical efficacy of Mongolia's language policy remain limited, impeded by a range of persistent implementation challenges at the grassroots level.
- (3) Local cultures of China and Mongolia stimulate biases among researchers and organisations who lead Mongolian script Unicode research. Thus, the relevant research is not on the same path to seek its development.
- (4) The collaboration between Mongolia and Inner Mongolia has not been consistent, due to various financial and non-financial factors.

These analyses aim to foster a better understanding among stakeholders of the Mongolian Unicode situation and its fundamental role in supporting the Mongolian language on digital platforms. Hopefully, this article may be of some use to encourage a consensus in establishing the universal standard for using the Mongolian script. Implementing a universal standard for rendering the Mongolian script will ensure a consistent appearance across platforms, thereby enhancing compatibility and cross-platform support. This initiative is expected to facilitate the development of fully fledged font and keyboard support for the script, ultimately establishing mature OpenType fonts that are accessible and user-friendly for a global audience.

3 First Factor – Unstable Language Ideologies

The first problem is historical. The instability of Mongolia's language ideologies,¹³ alongside the fluctuation of political interest projected on to the

13 Language ideologies are defined as 'the cultural system of ideas about social and linguistic relationships, together with their loading of moral and political interests' (Irvine 1989: 250).

Mongolian script, has resulted in the coexistence of two writing systems for the same language in Mongolia. In examining the state language ideologies related to the Mongolian script, the author focuses on how the literary forms of the Mongolian language (Mongolian and Cyrillic scripts) intersect with social values in Mongolia.

The linguistic instability is evident in the devaluation and revaluation of the Mongolian script within a span of fifty years, from the 1940s to the 1990s, in Mongolia. Following the introduction of the Cyrillic script in 1941, the Mongolian script was rapidly devalued. Cyrillic soon gained dominance, being referred to as the New Script – modern and progressive – fully reflecting the political climate of Mongolia post-communist revolution: the People's Revolution of 1921. The Mongolian script came to be known as the Old Script, connoting backwardness and evoking memories of Mongolia's feudal past under Manchu control. With the USSR's involvement in Mongolia's domestic affairs, 'during Communism, the Mongolian script was banned and those who tried to promote it were seen as the enemies of the USSR' (Dulam 2021: para. 11). Consequently, the Mongolian script was largely eradicated in Mongolia. Despite the dissolution of the USSR in 1991 sparking a nationalist enthusiasm for reviving the Mongolian script, most Mongolians today prefer the Cyrillic script. The derogatory representations of the Mongolian script persist, serving as a form of resistance, and hindering its use in contemporary Mongolia.

This issue was further elucidated through my fieldwork in Mongolia in 2023, where some participants disparaged the script as 'old' by politicising it in reference to Inner Mongolia. For instance, one interviewee stated: 'Inner Mongolians have been using the Mongolian script for over 800 years, and the script is a part of their identity as Chinese citizens. We do not wish to appear as Inner Mongolians.' This individual's rationale was that adopting the Mongolian script as an official script of Mongolia would render Mongolia into the Outer Mongolia Autonomous Region of China. Such narratives, echoed by a few participants, suggest that the promotion of the Mongolian script is perceived by some as a revival of 'federalism' reminiscent of the Ming and Qing Dynasties' suppression of Mongol independence. While these extreme views may not represent the majority opinion, they highlight the ongoing controversy surrounding the reintroduction of the Mongolian script under Mongolia's National Mongolian Script Programme III. This controversy is exacerbated by the current digital status of the Mongolian script. Suboptimal user experiences with search engines and web browsers further reinforce the perception of the Mongolian script as 'old'. Many interviewees considered that the Mongolian

script fails to represent advanced science and knowledge, thereby impeding Mongolia's cultural advancement.

4 Second Factor – Lack of Support from State Authorities

The status of the Mongolian script and language policies in both Mongolia and China reflects distinct yet parallel challenges. In both contexts, the Mongolian script does not occupy a central place in national language-planning agendas, and the prospect of the Mongolian script and language becoming the dominant means for communication remains uncertain.

4.1 *Mongolia's Language Policy*

A significant concern in Mongolia is the absence of enforceable rules or regulations to address Mongolian Unicode issues within a policy-driven framework. In Mongolia, there have been three state-sponsored initiatives aimed at restoring the Mongolian script: the National Mongolian Script Programme I (1995–2005), the National Mongolian Script Programme II (2008–15) and the National Mongolian Script Programme III (2020–4). Programmes I and II initiated the instruction of the Mongolian script in schools and its use in government administration. However, these efforts were largely unsuccessful due to the prevailing reliance on the Cyrillic script, lack of political will and insufficient funding (Billé 2010; Grivelet 1995). The primary objective of Programme III was to establish the Mongolian script as a co-state script alongside Cyrillic for use in government administration. Beyond this immediate goal, the programme also aspires, in the long term, to 'incorporate the Mongolian language and script into national values and foster its proficiency and use by every citizen' (Cabinet.gov.mn 2020: 3).

The governance and legal validity of language policy play a critical role in a bi-script environment, particularly when one writing system is institutionally privileged over another, as is the case in Mongolia. Programme III functions as an 'action plan' (ажиллагааны хөтөлбөр/*ajillagaany khotolbor*), operating within the framework of existing legislation. According to Article 7.2 of the Mongolian Language Law, government administration adopted both Cyrillic and Mongolian scripts from January 2025 (Parliament of Mongolia 2015).¹⁴ Programme III serves as the detailed action plan associated with Article 7.2,

14 '7.2.Энэ хуулийн 11.1.2-т заасан үндэсний бичгийн хөтөлбөрийг үе шаттай хэрэгжүүлж, зохих бэлтгэл хангасны үндсэн дээр төрийн болон нутгийн өөрөө

and it should be implemented by the Language Council and the Ministry of Digital Development and Communications of Mongolia.

However, the implementation of Programme III does not appear to have the highest probability of nationwide enforcement. As Sodnum *et al.* (2008: 43) have explained, in Mongolia, ‘only laws, parliamentary resolutions, and cabinet resolutions may be implemented nationwide, all other normative acts are restricted to the sector for which a given ministry or agency is responsible’. Although Article 7.2 is likely to fall under the category of formal ‘laws’, Programme III is classified under ‘normative acts’. Rather than imposing obligations on individuals directly, it serves as a guiding framework for the adoption and use of the Mongolian script across various administrative and institutional domains. Accordingly, the binding nature of such normative acts is inherently weaker than that of formal legislation.

This limitation in enforceability is further illustrated by the provisions concerning penalties under the Mongolian Language Law. As clarified by Member of Parliament M. Batchimeg, the law includes only a single clause related to the imposition of fines, which applies solely to legal entities. As such, there is no legal basis within the current framework for imposing penalties on individual citizens to ensure their compliance (Enkhmaa 2015: para. 24).¹⁵

Taken together, this legal context suggests that the enforcement of both the Mongolian Language Law, specifically Article 7.2 concerning the use of the Mongolian script, and the operational framework of Programme III, remain contingent upon voluntary compliance. Given the low cost and minimal punitive measures associated with non-compliance, the successful realisation of these language policies ultimately relies on cooperative engagement, rather than on coercive enforcement mechanisms.

This issue is particularly relevant to Mongolian Unicode, as Programme III explicitly mandates efforts to digitise the Mongolian script. For example, Article 3.1.4 of Programme III states, ‘Introduce the technology of using Mongolian script in social networks and multiply the types and forms of

удирдах байгууллага албан хэргээ кирил болон үндэсний хос бичгээр хөтлөн явуулна’ (Parliament of Mongolia 2015: Article 7.2).

15 ‘Misinterpretation 4: The law provisions [imposes] high fines if not adhered [to]. The only Mongolian language law has only one provision on fines to be imposed on legal entities equal to 10–15 times minimal [minimum] wage ([MNT] 1,920,000–2,880,000). The law itself is not hard to adhere to and does not require any monetary funds. There is no provision that states fine a citizen. It only states that every citizen is to adhere to the law’ (Enkhmaa 2015: para. 24).

Mongolian script content' (Montsame 2020: n.p.).¹⁶ However, in practice, the implementation of this directive is contingent upon how the Language Council and the Ministry of Digital Development and Communications choose to allocate responsibilities and determine timelines, depending on their respective capacities and resources.

4.2 *China's Language Policy*

In China, the digitisation of the Mongolian script and the maintenance of the Mongolian language must be understood in a nuanced manner. It is a recognised fact that China's current national standard GB/T 25914-2023 is the only enforceable language policy pointedly aimed at resolving the Mongolian script Unicode issues. However, within the broader context of China, state-level language support is provided only to Mandarin Chinese. The status of the Mongolian script has significantly declined in Mongolian-speaking areas in China following the implementation of China's New Bilingual Education Policy in September 2020. This policy transitioned the use of the Mongolian script and its spoken language to Chinese in all Mongolian schools by altering the language of instruction and course materials for literacy subjects to Chinese.

The policy has faced considerable criticism for its human rights implications and allegations of ethnic assimilation, and it has triggered protests inside and outside China (BBC 2020; Davidson 2020; Xiao, E. 2020). This policy has led to two significant changes. First, the removal of the Mongolian-language coursebooks and instructions for three subjects: Chinese, history and politics. As a result, Chinese is now taught as the primary language, enjoying the same status as the Mongolian language in Mongolian education. Furthermore, Chinese has replaced Mongolian as the language of instruction for history and politics. This shift effectively establishes Chinese as the dominant language of instruction in Mongolian schools, thereby eliminating the use of the Mongolian script and its spoken forms as the medium of instruction – the primary language of literacy – for Mongolian students. The primary language of literacy is a vital indicator of a writing system's vitality, influencing how speakers develop literacy and cognitive skills across various domains. As more Mongolians in China acquire literacy in Chinese, the necessity for the Mongolian script in digital and sophisticated contexts, such as knowledge acquisition and dissemination, naturally diminishes. Second, China's central government has applied mitigation policies to suppress the protests. Research associated with

16 3.1.4. Нийгмийн сүлжээнд монгол бичгээ хэрэглэх технологийг нэвтрүүлж, контентын төрөл, хэлбэрийг олшруулах (Montsame 2020).

Mongolian language use, including Mongolian script Unicode research, has been facing the risk of being identified as non-patriotic studies that disagree with the implementation of the New Bilingual Education Policy. Since 2019, the field of Mongolian Unicode studies has experienced significant challenges within Inner Mongolian academia, resulting in some intellectual decline within the community of Unicode experts. This was evident from the author's fieldwork, as a number of prominent Unicode scholars and institutions in Inner Mongolia initially expressed willingness to collaborate in 2022. However, they subsequently reversed their decision and gradually ceased communication with the author.

5 Third Factor – Researchers' Bias on Research Approach

The third problem impacting the digitisation of the Mongolian script is the restricted research in the field of Mongolian script in Unicode, which is influenced by researcher bias favouring the phonetic model over the graphetic model. This bias manifests differently in Mongolia and Inner Mongolia.

5.1 *Research Bias in Mongolia*

In Mongolia, opposition to the graphetic model is associated with anti-Chinese sentiment, often referred to as Sinophobia, even though the Chinese government has never officially endorsed this model. This sentiment can be traced to historical tensions and concerns regarding China's influence in the region. During the Unicode Consortium meetings in 2018, the involvement of the introducers of a graphetic approach from China, namely Liang Hai and Shen Yilei (Liang, H. 2018; Shen, Y.L. 2018), intensified these sentiments among Mongolians.¹⁷ The Unicode Consortium is an international organisation comprising contributors from various countries. Liang and Shen are recognised contributors to Unicode, particularly in light of actions taken by the Unicode Technical Committee to enhance the documentation and implementation of the existing Unicode Model for Mongolian, as well as to make recommendations for its improvement.

During the meeting, Mongolia's experts strongly disapproved of the graphetic model, arguing that it did not adequately represent the phonetic

17 Liang Hai is a Vice-Chair of the Unicode Editorial Working Group, and Shen Yi Lei is an individual expert whom the Unicode Committee invited. There is also an Inner Mongolian expert called Jirimutu (Jirimutu 2018), who is affiliated with a private firm, and who made a detailed analysis of the graphetic model.

features of the Mongolian script and failed to consider traditional Mongolian syllable-joining rules (Anderson 2017; Moore 2019). Liang criticised this stance, stating that, ‘This way of thinking is not a rational analysis of technical pros and cons of the graphetic model, but rather it is pure emotional resistance’ (Liang, H. 2018: 8). Liang’s critique of this emotional resistance highlights a divergence in perspectives regarding the relative merits of technical arguments versus the cultural implications of the graphetic model in Mongolia.

The emotional weight of this issue is accentuated by the way in which Mongolia’s experts characterised their resistance. Following the 2018 Unicode meeting, two of the three Mongolian experts who attended the meeting were interviewed by a prominent local media outlet in Ulaanbaatar. The two experts framed their discussion as a ‘war’ and, likening the meeting to the Treaty of Kyakhta (1915),¹⁸ they articulated a profound commitment to preserving their linguistic heritage via committing to the phonetic model.¹⁹ They noted that when faced with the debate between the two models, Inner Mongolians proposed pursuing the development of both. Their rationale may have been a sense of urgency to reach a decision, or a diminishing ‘faith’ in the phonetic model.²⁰ The graphetic model was not merely regarded as a technical solution proposed by Unicode experts, irrespective of their nationalities or professional affiliations, to address the representation of the Mongolian script within Unicode. Rather, the identity of the proponents appeared to play a significant role in shaping perceptions. The model was interpreted not as a neutral or collaborative proposal, but as an initiative originating from China and associated with Chinese individuals, in their own words, as ‘the graphetic model that the Chinese developed’ (Sanlig & Togoobat 2018b: para. 6).²¹ It is important to note

18 The Treaty of Kyakhta (1915) reflected conflicting interests: Mongolia sought full independence, while China aimed to limit Mongolian autonomy. As a result, the Mongols viewed the treaty as a failure, as it denied their aspirations for true sovereignty (Soni 2024).

19 As the original quotation states, ‘Some who understand the significance of this meeting are right when they say, “this was like the modern Khiagt agreement”. This is a matter of whether Mongolian script will exist for the next five years, 500, or 5000’ (Sanlig & Togoobat 2018b: para. 22).

20 As the original quotation states, ‘Inner Mongolians, on the other hand, suggested to develop both models, maybe they are in a rush to decide on some solutions to prevent further drag out, or they lost faith in the phonetic model’ (Sanlig & Togoobat 2018b: para. 15).

21 As the original quotation states, ‘Had we approved the Unicode standard for Mongolian script as the graphetic model that the Chinese developed, it would’ve then been discussed and approved at the ISO international meeting, and everything could have been over. Because these two meetings were scheduled right after another, we tried very hard to not take it to the ISO meeting. That Unicode Technical Commission meeting was almost like war’ (Sanlig & Togoobat 2018b: para. 6).

that a graphetic idea was, in fact, proposed and advocated by individuals from diverse backgrounds. According to members of the Unicode Consortium, it was not a uniquely *Chinese idea*, nor did it reflect a singular national agenda.

Nevertheless, the introduction of the graphetic model has been racialised and perceived as a form of Chinese imposition, seen by some as undermining Mongolia's linguistic sovereignty. The fact that the Chinese language is encoded graphetically in Unicode has led to concerns that adopting a similar approach for Mongolian would symbolically align it with Chinese characters, thus contributing to fears of cultural assimilation. As articulated by Mongolian representatives, 'How could we have the Mongolian script encoded graphically, like Chinese characters on our watch?' (Sanlig & Togoobat 2018b: para. 15).²² According to Sanlig and Togoobat (2018b), their successful appeal to the Unicode Consortium not to adopt the graphetic model underscores the vital importance of preserving Mongolia's linguistic identity.²³ As they emphasised, 'In this last meeting, we have achieved our objective for the past few years and defended our heritage' (Sanlig & Togoobat 2018b: para. 27).

While their discourse surrounding the graphetic model encapsulates a multifaceted interplay of cultural identity and geopolitical sensitivities, another possible explanation for the two experts' opposition to the graphetic model may be a potential conflict of interests, given their affiliation with a technology company, Bolorsoft. This company was preparing to launch an application based on the refined phonetic model in early 2019 (Bolorsoft 2018, 2019). Although the current level of adoption of the refined phonetic model in Mongolia remains uncertain, a 2019 article published by the Mongolian National News Agency, titled 'Typing in Mongolian script becomes available in cyberspace', cites the president of Bolorsoft stating that their version of the phonetic model 'will be used in cyberspace for typing in Mongolian script and further we will create other products that support it' (Batchimeg 2019: para. 2).²⁴

22 The graphetic features of the Mongolian script are discussed in the 'Final remarks' section of this article.

23 'If we had not given a presentation at this last meeting and changed the commission's understanding, our Mongolian script really would have been encoded graphetically' (Sanlig & Togoobat 2018b: para. 21).

24 As a related observation, the news article appears to contain potentially misleading information. Reporter B. Batchimeg claims that the refined phonetic model for typing the Mongolian script, developed by Bolorsoft, 'was approved by Unicode technical committee experts' (Batchimeg 2019: para. 2). However, the author was unable to verify this statement in any official documentation listed under the Unicode Topical Document List: Mongolian (Unicode 2024a). In particular, no reference to such approval could be found in the Mongolian Working Group Meeting 2 Report (Moore 2018). Furthermore, in April 2025, the author consulted three anonymous members of the Unicode Technical

Considering the government's efforts to reintroduce the Mongolian script into official domains such as education and public administration, the anticipated market potential of this application may have shaped the experts' rejection of the graphetic model. Their preference for the phonetic approach may thus reflect not only technical or theoretical concerns, but also strategic interests tied to the commercial viability of their affiliated project.

5.2 *Research Bias in Inner Mongolia*

In contrast, Inner Mongolian scholars exhibit their own biases. The phonetic model was established as the Unicode standard model for Mongolian in 1999, largely due to the efforts of prominent Inner Mongolian linguists. This decision occurred within a historical context marked by a significant lack of mutual understanding between the Unicode Consortium and the proponents of phonetic models. At the time Mongolian was added to Unicode (1999), members of the Unicode Consortium were reportedly not sufficiently aware of best practices for encoding, nor of the unique characteristics of the Mongolian script (Anonymous 2024a). It has also been suggested that the script itself was poorly understood by Unicode encoders, just as the Unicode standard was not well understood by those who initially proposed the Mongolian encoding scheme (Anonymous 2024b).²⁵

As a result, the encoding decisions made by Inner Mongolian linguists, who were responsible for China's national standard GB18030, were adopted *wholesale* by the Unicode Consortium. This model has since maintained dominance in Inner Mongolian academic circles, resulting in a substantial body of Unicode-related publications. However, the ongoing debates between the phonetic model and the graphetic model have not ignited meaningful scholarly engagement within Inner Mongolian academia. The author identified 23 peer-reviewed articles published in China between 1994 and 2024 that appear to address Unicode-related issues concerning the Mongolian script, written in either Chinese or the Mongolian script. It seems that only three of these briefly refer to the graphetic model as a potential encoding method (Garudi & Jirimutu n.d.; Si, 2009; Suyaltu 2001). To the best of the author's knowledge, no peer-reviewed academic publications appear to have

Committee who had participated in the meeting in question. None of these individuals confirmed that the refined phonetic model had received formal approval from the committee.

25 'From what I understand, the script was not well-understood by the Unicode encoders at the time, and I believe Unicode was not well-understood by the proposers of the original Mongolian script' (Anonymous 2024b).

conducted an in-depth analytical or comparative examination of the graphetic model in relation to the phonetic model.

The graphetic model, proposed by independent experts (Jirumutu 2018; Liang, H. 2018; Namujila 2020a; Namujila 2020b; Script Ad Hoc Group Recommendations 2017), lacks institutional academic backing in both Inner Mongolia and Mongolia. Consequently, literature addressing the graphetic model is sparse and largely confined to technical documentation from the Unicode meetings held in 2017 and 2018 (report no. MWG/2-N4R; MWG/2-N5R; MWG/2-N8; MWG/2-N2R; MWG/2-N3R) (Unicode 2024a).

One contributing factor to this research bias may be the predominantly descriptive and non-critical academic culture frequently observed in China. Academic writing in the Chinese context is often characterised by a preference for descriptive reporting over critical analysis (Chilton *et al.* 2012; Clark & Gieve 2006; Normile 2023; Tan 2017). Critiquing the work of established or well-regarded scholars poses significant challenges. As a result, scholars in Inner Mongolia are generally less inclined, or may perceive little need, to critically examine the phonetic model, which was developed by widely respected figures in the field.

An anonymous Unicode expert from Inner Mongolia has observed that although the graphetic model may provide certain technical advantages in the digitisation of the Mongolian script, its practical implementation is constrained by the entrenched status of the phonetic model within academic communities. According to this expert, transitioning to the graphetic model would necessitate a substantial reinvestment of time and resources, which is viewed as impractical given the current infrastructure developed around the phonetic standard.²⁶ This academic environment limits objective evaluation and impedes innovation in addressing the technical challenges associated with the digital representation of the Mongolian script.

26 An insider who attended the 2017 and 2018 meetings recalled that, despite the practical risks associated with adopting a new model, most experts from Inner Mongolia and Mongolia remained committed to the phonetic model: 'This is possible, but my memory was that the overwhelming view from most Mongolian and Inner Mongolian experts was to retain the phonetic model, though there were some issues that need to be fixed in that model' (Anonymous 2024a).

6 Collaboration between Inner Mongolia and Mongolia

Mongolia played a minimal role in the initial stages of developing the phonetic model, resulting in inconsistent collaboration with Inner Mongolia. This limited engagement led to the development of the Mongolian Unicode Block (Figure 1) being shaped predominantly by the language digitisation priorities of China. The design of the block reflects broader considerations for various vertically written scripts in China, rather than focusing exclusively on Mongolia’s national script. This is evident in the inclusion of characters unrelated to Mongolic languages, such as Manchu and Sibe.

The early development of the Mongolian Block was primarily driven by scholars based in Inner Mongolia, who exercised near-exclusive authority over its structural decisions. This centralised control contributed to the implementation of certain character sets in a manner that later proved controversial; however, under the Unicode Stability Policy, such implementations cannot be easily revised once published.

Efforts from Mongolia to address issues within the Mongolian script Unicode framework have not matched those from China, including both Han-Chinese and Inner Mongolian scholars. Mongolia has continued to face difficulties initiating equitable dialogue and collaboration to address these challenges.

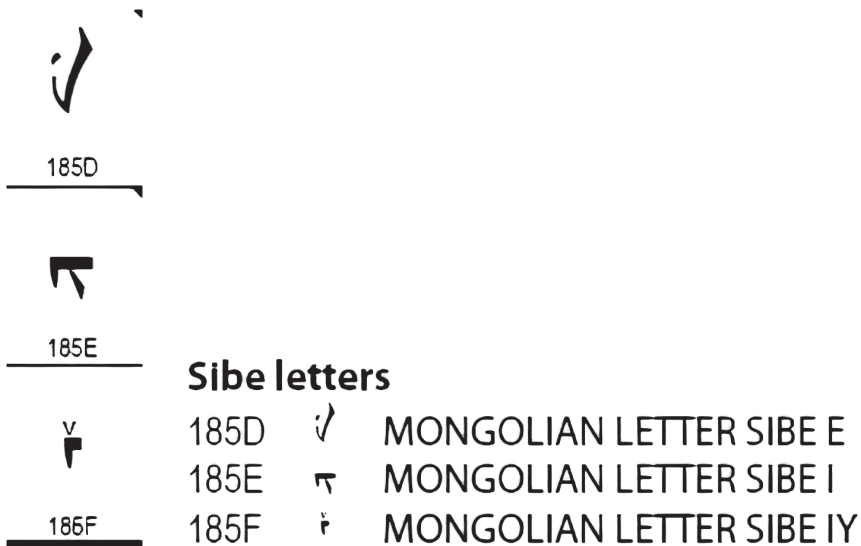


FIGURE 3 Three examples of Manchu letters out of 18 Manchu letters in the Mongolian Range (Unicode 2024c: 162)

Manchu letters

- 1873 𐎢 MONGOLIAN LETTER MANCHU I
- 1874 𐎣 MONGOLIAN LETTER MANCHU KA
- 1875 𐎤 MONGOLIAN LETTER MANCHU RA

FIGURE 4 Three examples of Manchu letters out of 18 Manchu letters in the Mongolian Range (Unicode 2024c: 162)

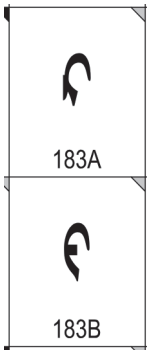


FIGURE 5 183A 𐎣 Mongolian letter 'ka' 183B 𐎤 Mongolian letter 'kha' (Unicode 2024c)

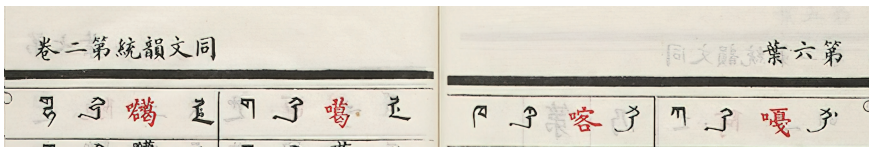


FIGURE 6 The examples of Ali-Gali Mongolian (Qing Government 1929: 6a–6b)

6.1 Monopoly of Knowledge in Inner Mongolia

While acknowledging that the Unicode approach is independent of national politics – focusing solely on the graphic elements necessary for digital interchange – the structure of the Mongolian Block reveals a distinctly China-centric orientation. The block accommodates multiple vertical writing systems, not solely the Mongolian script. As stated in Unicode (2006: 532), ‘the Mongolian block unifies the traditional writing system for the Mongolian language and the three derivative writing systems, Todo, Manchu and Sibe’ (see Figures 3 and 4). These languages are officially recognised as ethnic minority languages within China (Ministry of Education of China 2021).

In addition to Todo, Sibe and Manchu, the Mongolian Block also includes Ali-Gali letters, which extend the Mongolian script. The Ali-Gali script, developed in 1587 by Ayuush Güüsh under the inspiration of the Third Dalai Lama, was created to transcribe Buddhist texts from Tibetan and Sanskrit (Kao 2022). Today, Ali-Gali letters are primarily used to transcribe foreign terms, such as

表 A. 28 U+183A (᠎) 蒙古文字母 KA/ MONGOLIAN LETTER KA 变形
显现形式及自由变体选择符使用规则

词里位置	变形显现形式			中 / 英文名称	自由变体选择符	位置默认变体	语法默认变体
	序号	编号	图形				
独立	1	00D2	᠎	蒙字母.kā 独立变体 ml. ka isolate form		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
词首	1	00D2	᠎	蒙字母.kā 词首变体 ml. ka initial form		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
词中	1	00D2	᠎	蒙字母.kā 词中变体 ml. ka medial form		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
词末	1	0035	᠎	蒙字母.kā 词末变体 ml. ka final form		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

FIGURE 7 Mongolian letter 'ka' (Samr.gov.cn 2024: 40)

GB/T 25914—2023

表 A. 29 U+183B (᠎) 蒙古文字母 KHA/ MONGOLIAN LETTER KHA 变形
显现形式及自由变体选择符使用规则

词里位置	变形显现形式			中 / 英文名称	自由变体选择符	位置默认变体	语法默认变体
	序号	编号	图形				
独立	1	00D3	᠎	蒙字母.khā 独立变体 ml. kha isolate form		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
词首	1	00D3	᠎	蒙字母.khā 词首变体 ml. kha initial form		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
词中	1	00D3	᠎	蒙字母.khā 词中变体 ml. kha medial form		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
词末	1	0036	᠎	蒙字母.khā 词末变体 ml. kha final form		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

FIGURE 8 Mongolian letter 'kha' (Samr.gov.cn 2024: 41)

ᠬᠠᠭᠠᠨᠠᠨᠠᠭᠠᠨ 'computer'. However, the modern implementation standard adopted by China for two Ali-Gali codepoints remains controversial, particularly in educational contexts across both China and Mongolia.

Figure 5 illustrates codepoints 183A (᠎, Mongolian letter Ka) and 183B (᠎, Mongolian letter Kha) (Unicode 2024c: 161). These characters correspond to distinct Tibetan letters, and they reflect aspirated and unaspirated phonemes respectively, consistent with the phonetic basis of the script (see Figure 6).²⁷ According to China's National Standard GB/T 25914-2023 (Samr.gov.cn. 2024), which informs UTN #57, 183A and 183B are implemented differently in Mongolia

27 Mongolia's representatives acknowledged that 183B is for non-native Mongolian words: 'There were other consonant letters to write foreign words. Those are: ᠮ fa, ᠮ ca, ᠮ za, ᠮ ka, ᠮ ha, ᠮ Zhi, ᠮ Chi' (Sanlig & Togoobat 2018a: 3).

and Inner Mongolia (see Figure 7 and Figure 8), with each region adopting a distinct interpretation of the characters in their school curricula.

In Mongolia's standard coursebook, 183B (ᠬ) is taught as 'ka' rather than 'kha' (Ministry of Education and Science of Mongolia 2018: 90). Conversely, in Inner Mongolia's materials, 183A (ᠬ) is identified as 'ka' (Inner Mongolia Education Press 2016: 16). This bifurcation assigns different characters to represent the same sound in the two regions, despite historical evidence contradicting this distinction.

Consultations with native Mongolian linguists from both Mongolia and Inner Mongolia confirm that historically there has been no differentiation between a Mongolian 'ka' and an Inner Mongolian 'ka'. Initially developed to represent different Tibetan sounds, the two letters have become stylistic variants over time as knowledge of their phonetic distinctions diminished. The notion of distinct regional variants of 'ka' only emerged in recent decades, coinciding with the implementation of Mongolian Unicode standards.

No formal documentation has been found to explain the decision to designate 183B as Mongolia's exclusive 'ka' and 183A as Inner Mongolia's. One possible hypothesis is that some words in Khalkha, the standard spoken form of Mongolian in Mongolia, feature an aspirated pronunciation of the 'ka' phoneme, potentially influencing the decision to adopt 183B in Mongolia. However, this justification appears insufficient, as not all occurrences of 'ka' in Khalkha are aspirated. Furthermore, phonological variation exists across many Mongolic languages and dialects, not just Khalkha. The decision to exclusively associate 183B with Mongolia likely reflects the intentions of code proposers to tailor the standard specifically to Mongolia's needs, rather than reflecting broader linguistic realities.

There is also a potential concern regarding duplicated encoding. Some may argue that 183A and 183B represent the same phoneme in practice, thus violating the Unicode Standard's principle of 'avoiding duplicated characters across scripts and languages' (Unicode 2024b: sec. 7). Theoretically, the Mongolian Block is phonetic, distinguishing characters based on sound. Therefore, 183A and 183B, which denote different phonemes, are not considered duplicates. However, in actual usage, the two codepoints often represent the same phoneme, 'ka'. Unicode's focus remains on encoding rather than on practical application, and educational implications typically fall outside its scope.

The inclusion of 183A and 183B is appropriate given the Mongolian Block's broader inclusion of non-Mongolic scripts. Nonetheless, the implementation standard issued by China (Samr.gov.cn. 2024) has significant implications for education and script standardisation. The exclusive designation of 183B for Mongolia raises practical questions. Should keyboards in Mongolia accommodate both characters? Should educational materials acknowledge the Inner

Mongolian ‘ka’ (183A)? How should spell-checkers and translation software account for these distinctions across regions?

The differing approaches to using ‘ka’ risk deepening the divide between Mongolian and Inner Mongolian users, who may remain unaware of alternative characters due to their absence in formal education. These disparities in implementation have substantial consequences for the teaching, standardisation and digital application of the Mongolian script, and they warrant further consideration.

6.2 *Collaboration Issues and Organisational Challenges in Mongolia*

Mongolia’s limited involvement in the development of the Mongolian script Unicode has had a lasting impact on its collaboration with Inner Mongolia – a situation that persists to this day. This limited participation can be partially attributed to structural and administrative challenges within Mongolian government institutions.

Quantitative data highlight Mongolia’s minimal engagement in international dialogues concerning the Mongolian script Unicode. For instance, among more than 200 proposals listed in the *Topical Document List: Mongolian* (Unicode 2024a) (excluding joint submissions and meeting reports), only approximately 24 were submitted by Mongolian delegates or experts, whereas around 84 proposals originated exclusively from China or Han-Chinese/Inner Mongolian scholars and specialists. Additionally, during the 2017 Working Group Meeting (a pivotal forum for discussing the graphetic model among Unicode experts) only 4 of the 32 attendees represented Mongolia. The remaining 28 included 11 from Inner Mongolia, 5 from China, and 12 from international organisations such as the Unicode Consortium, Adobe, Google, Microsoft and SIL (Anderson 2017). In the following year, at the 2018 Working Group Meeting, only 3 out of 27 attendees were from Mongolia, while the remaining 24 comprised 7 from China (4 from Inner Mongolia and 3 from China), and 17 were from various organisations or affiliations outside China and Mongolia (Moore 2018).

The evolution of Mongolia’s institutional framework has significantly influenced its capacity to participate in Unicode-related initiatives. In 2017, responsibility for Unicode matters shifted from the Mongolian Agency for Standardization and Metrology to the Communications and Information Technology Authority of Mongolia (CITA).²⁸ Following the 2017 meeting, President Kh. Battulga issued directives for CITA, in cooperation with the

28 CITA was dissolved in 2022, and the responsibilities for Mongolian Unicode were subsequently transferred to the Ministry of Digital Development and Communications. For more information, see: <https://montsame.mn/en/read/286665> (accessed 16 July 2025).

Agency for Standardization and Metrology, to further develop the phonetic model (Sanlig & Togoobat 2018b). Despite this instruction, no tangible progress was made, primarily due to financial constraints. As reported, ‘they did nothing because they didn’t have a budget’ (Sanlig & Togoobat 2018b: para. 9). Compounding this issue was ineffective coordination between governmental bodies, which hindered the establishment of a dedicated entity to address matters concerning Mongolian Unicode. Sanlig and Togoobat aptly describe the situation: ‘The Mongolian script including improvement issue was bounced between state organisations like a tennis ball’ (Sanlig & Togoobat 2018b: para. 10).

These organisational limitations also reveal deeper issues regarding internal collaboration and national representation. At the 2018 Working Group Meeting, three experts attended from Mongolia: a representative from CITA, a software architect, and a typography specialist affiliated with a local technology company. The latter two, recognised for their professional expertise, had previously participated in the 2017 meeting, and they were once again invited by the Unicode Consortium in 2018.

Despite their qualifications, the Mongolian government did not officially designate them as state representatives, expressing reservations about including individuals from private companies in a governmental working group. As stated, ‘it’s not suitable for private company representatives to be included in a state working group that’s private party interest’ (Sanlig & Togoobat 2018b: para. 22). The government even took the step of sending a formal letter to the Unicode Committee in an effort to prevent these two experts from acting on behalf of the state (Sanlig & Togoobat 2018b).²⁹ This decision effectively resulted in Mongolia having only one officially sanctioned delegate among the 27 attendees at the 2018 meeting.

As a result, the three Mongolian participants operated as an ad hoc delegation during the 2018 meeting. This interim team faced internal disagreements and failed to present a cohesive national stance. The CITA representative, hindered by limited experience, struggled to articulate Mongolia’s official position. Notably, just one week before the meeting, the group had not yet agreed

29 ‘A letter was sent to the Unicode technical committee stating these two men do not have the right to represent the state and that the person named Enkhdalai today will be coming. Those two men [two experts from a private firm] have just returned from participating in that meeting. But political suggestions and conclusion do not affect the Unicode technical commission meeting’ (Sanlig & Togoobat 2018b: para. 18). This information is accurate, as confirmed by a Unicode staff member who facilitated the meeting in 2018.

on who would deliver the country's formal statement (Sanlig & Togoobat 2018b). The performance of the CITA expert during the meeting was criticised by his two colleagues, who claimed that he lacked the necessary experience and was manipulated by 'Chinese experts' (Liang Hai, Shen Yi Lei) into supporting the graphetic model (Sanlig & Togoobat 2018b: para. 18). Although both of these Chinese experts are ethnically Han-Chinese and one is affiliated with the Unicode Consortium as an individual member, neither represents the official delegation or position of the Chinese government.³⁰ The CITA expert's tentative and 'few unsure words' while presenting the phonetic model led other participants to advocate for the superiority of the graphetic model. Following his statement, the Chinese experts stood up and declared, 'This is exactly why there should be a graphetic model' (Sanlig & Togoobat, 2018b: para. 18).

In 2019, Mongolia, through CITA, hosted a Mongolian Working Group Meeting in Ulaanbaatar with the aim of addressing the shortcomings of the phonetic model. A member of the Unicode Technical Committee noted that opportunities for Mongolia to participate in Unicode discussions have consistently been available, and that such engagement is not only welcomed but actively encouraged (Anonymous 2024a). However, since 2020, Mongolia has not submitted any additional documents to the Unicode Consortium related to the unification of Mongolian-specific shaping rules. This continued hiatus remained in effect at the time of submission of this article in May 2025.

The absence of input from Mongolia suggests a lack of consensus on shaping rules between Mongolia and China. This incomplete collaboration impedes progress on related initiatives, such as the development of standard-conforming OpenType tables for the Mongolian script, resulting in inconsistent representation across platforms, particularly in Mongolia and China.

In summary, sociopolitical and historical factors have contributed to China's stronger decision-making position in this context. Mongolia's internal organisational issues have undermined the development of a cohesive research team capable of effectively representing Mongolia in international discussions about the Mongolian script Unicode. It is hard to argue that the minority opinions of Mongolia were not drowned out by the opinions of the majority. The two countries' collaboration was unequal, and it continues to exhibit disparity.

30 See footnote 17. As mentioned, Liang Hai is an individual member of the Unicode Consortium, while Shen Yilei is an independent expert.

7 Final Remarks

This article highlights a rare case where the national script of a sovereign state, which is also an official writing system in two countries, remains inadequately digitised in the twenty-first century. Despite the recent efforts articulated in UTN #57, which have addressed the under-specification issues of the phonetic model, it remains evident that this model has not yet fulfilled the fundamental Unicode Design Principles of ‘universality’ (Unicode 2023: sec. 7). The phonetic model has not been universally adopted by all vendors (Batjargal *et al.* 2011; Jirimutu 2018), and it still has not achieved widespread implementation. While the technical challenges associated with digitising a vertical script within horizontal systems are acknowledged, the non-technical issues, including sociopolitical, historical, economic and intellectual factors, continue to significantly influence the digitisation process. More collaborative research efforts are essential to address the remaining challenges associated with the phonetic model implementation and to advance the use of the Mongolian script in digital devices.

The remit of this article is confined to a brief discussion of the debates between the phonetic model and the graphetic model, alongside an ethnographic analysis of how the phonetic model has been established as the Unicode Standard for the Mongolian script. An important question remains unresolved regarding the contexts in which the Mongolian script should be interpreted phonetically versus graphically in digital environments, particularly in relation to applications such as voice recognition and optical character recognition (OCR). The underlying principles of the Mongolian script indicate that it cannot be exclusively understood as a phonetic language; rather, ‘it may be concluded that there are two levels at which the Mongol script can be segmented: a surface level, which may be called alphabetic, and a deep level, which may be called glyphic’ (Janhunnen 2003: 37). The distinct vertical features of the Mongolian script, which derive and inflect words along a vertical line using graphic shapes culturally referred to as ‘tooth’, ‘horn’, ‘hair’, ‘belly’ and ‘tail’ illustrate that this script is also composed graphetically. Further research is necessary to explore methodologies that promote a flexible and unbiased approach, enabling the effective use of the Mongolian script while addressing the challenges posed by its inherent phonetic and graphetic characteristics.

Acknowledgements

The author would like to extend my sincere gratitude to my colleagues from the Script Encoding Initiative at UC Berkeley, Microsoft, the Academy of Sciences

of China, and the two anonymous reviewers for their invaluable insights and contributions to this article.

References

- Anderson, D. 2017. Mongolian Ad Hoc Report (Hohhot, Inner Mongolia). *Topical Document List: Mongolian*. <https://www.unicode.org/wg2/docs/n4893-MongolianMeetingSummary.pdf> (accessed 11 June 2024).
- Anderson, D. 2020. Notes on Informal Meeting on Mongolian with Chuck Namujila. *Topical Document List: Mongolian*. <https://www.unicode.org/mwg/mwg4docs/mwg4-8NotesOnInformalMeeting23April20-Anderson.pdf> (accessed 23 September 2024).
- Anonymous. 2024a. Interviewed by Tergel. 8 July 2024.
- Anonymous. 2024b. Interviewed by Tergel. 18 October 2024.
- Batchimeg, B. 2019. Typing in Mongolian script becomes available in cyberspace. MONTSAME News Agency. <https://www.montsame.mn/en/read/180087> (accessed 14 April 2025).
- Batjargal, B., F. Kimura & A. Maeda. 2010. A Survey on Rendering Traditional Mongolian Script. 'International Conference on Asian Language Processing', 3–6. <https://doi.org/10.1109/IALP.2010.16>.
- Batjargal, B., G. Khaltarkhuu, F. Kimura & A. Maeda. 2011. A study of traditional Mongolian script encodings and rendering: use of Unicode in OpenType fonts. *International Journal on Asian Language Processing* 21(1): 23–43. <https://www.colips.org/journals/volume21/21.1.3-Biligsaikhan.pdf> (accessed 11 June 2024).
- BBC. 2020. 从内蒙古双语教学争议看中国打造少数民族“大熔炉” (cong neimenggu shuangyu jiaoxue zhengyi kan zhongguo dazao shaoshu minzu 'daronglu') [From the controversy of bilingual teaching in Inner Mongolia to see China's building a 'melting pot' for minorities]. *BBC News*. <https://www.bbc.com/zhongwen/simp/chinese-news-54346824> (accessed 11 June 2024).
- Billé, F. 2010. Sounds and scripts of modernity: language ideologies and practices in contemporary Mongolia. *Inner Asia* 12(2): 231–52. <https://doi.org/10.1163/000000010794983522>.
- Bolorsoft. 2018. *The Mongol Script Encoding – 2018 Meeting Summary*. Bolorsoft. <http://blog.bolorsoft.com/wp-content/uploads/2018/12/Meeting-Summary-En.pdf> (accessed 05 November 2024).
- Bolorsoft. 2019. The Refined Phonetic Model: the most convenient solution for recovering the current Mongolian encoding. *Topical Document List: Mongolian*. <https://www.unicode.org/mwg/mwg3docs/mwg3-7RefinedPhoneticModelR.pdf> (accessed 23 September 2024).

- Cabinet.gov.mn. 2020. Vision-2050 Long-Term Development Policy of Mongolia. *Cabinet*. https://en.iss.gov.mn/wp-content/uploads/2021/10/2050_VISION_LONG-TERM-DEVELOPMENT-POLICY.pdf (accessed 29 August 2023).
- Chen, Z. 2017. Draft of Mongolian Shaping Rules by Prof. Quejingzhabu and His Colleagues. *Topical Document List: Mongolian*. <https://www.unicode.org/L2/L2017/17251-mong-shaping.pdf> (accessed 01 April 2025).
- Chilton, P., H. Tian & R. Wodak. 2012. Reflections on discourse and critique in China and the West, in P. Chilton, H. Tian & R. Wodak (eds), *Discourse and Socio-Political Transformations in Contemporary China*: 1–18. Amsterdam: John Benjamins. <https://doi.org/10.1075/bct.42.02chi>.
- CITA. 2019. Improvement of Mongolian Unicode Block. *Topical Document List: Mongolian*. https://www.unicode.org/mwg/mwg3docs/mwg3-5UB_Meeting_CITA-WG_Presentation.pdf (accessed 23 September 2024).
- Clark, R. & S.N. Gieve. 2006. On the discursive construction of ‘The Chinese Learner’. *Language, Culture and Curriculum* 19: 54–73. <https://doi.org/10.1080/07908310608668754>.
- Davidson, H. 2020. Inner Mongolia protests at China’s plans to bring in Mandarin-only lessons. *The Guardian*, 1 September. <https://www.theguardian.com/world/2020/sep/01/inner-mongolia-protests-china-mandarin-schools-language> (accessed 11 June 2024).
- Dulam, S. 2021. Are we ready to use national script by 2025? *The UB Post*, 24 May. <https://www.pressreader.com/mongolia/the-ub-post/20210524/281522228992935> (accessed 11 June 2024).
- Enkhdalai, B. 2019. Insufficiency of the Existing Description of Variant Selectors and Variant Sequences in Mongolian Unicode Character Table. *Topical Document List: Mongolian*. <https://www.unicode.org/mwg/mwg3docs/mwg3-4ModificationDescriptionOfCodeChartsR.pdf> (accessed 23 September 2024).
- Enkhmaa, E. 2015. Mongolian language law is effective from July 1st. *Gogo Mongolia*. <https://mongolia.gogo.mn/r/146942> (accessed 17 April 2025).
- Garidi. & Jirimutu. n.d. Problems for Preparing Character Code Standard of Mongol. No place: no publisher.
- Glass, A. 2019. Proposal to Create Open Source OpenType Tables for Modern Mongolian. *Topical Document List: Mongolian*. <https://www.unicode.org/mwg/mwg3docs/mwg3-14OpenSourceOpenTypeR.pdf> (accessed 23 September 2024).
- Grivelet, S. 1995. Reintroducing the Uighur-Mongolian script in Mongolia today. *Mongolian Studies* 18: 49–60. <http://www.jstor.org/stable/43193217>.
- Huang Wei Min & China. 1987. Mongolian 7-bit and 8-bit Coded Graphic Character Sets. *Topical Document List: Mongolian*. <https://unicode.org/L2/topical/mongolian/> (accessed 04 April 2015).

- Inner Mongolia Department of Education. 2020. 秋季学期起我区民族语言授课学校小学一年级和初中一年级使用国家统编语文教材 [Starting from the autumn semester, the first grade of primary school and the first grade of junior high school of Inner Mongolia will use the national unified Chinese textbooks]. *Inner Mongolia Department of Education*. https://jyt.nmg.gov.cn/slh/tpxw_25439/202008/t20200826_687130.html (accessed 11 June 2024).
- Inner Mongolia Education Press. 2016. *Mongolian*. Hohhot: Inner Mongolia Education Press.
- Irvine, J.T. 1989. When talk isn't cheap: language and political economy. *American Ethnologist* 16(2): 248–67. <https://doi.org/10.1525/ae.1989.16.2.02a00040>.
- Janhunen, J.A. (eds) 2003. *The Mongolic Languages*. London: Routledge.
- Janhunen, J.A. 2012. *Mongolian*. Amsterdam: John Benjamins.
- Jiang, K. 2023. 蒙古文字的文本表记与成形手册 [Text representation and shaping manual of the Mongolian script]. *Topical Document List: Mongolian*. <https://www.unicode.org/L2/L2023/23101-mongolian-draft-utn.pdf> (accessed 11 June 2024).
- Jiang, K. 2024. Representing Mongolian in Unicode. *Topical Document List: Mongolian*. <https://www.unicode.org/notes/tn57/> (accessed 23 September 2024).
- Jirimutu. 2018. Opinions on 'A Graphetic Approach for the Mongolian Encoding Model'. *Topical Document List: Mongolian*. <https://www.unicode.org/mwg/mwg2docs/mwg2-2Opinions-Jirimutu.pdf> (accessed 11 June 2024).
- Kao, H.-T. 2022. From the Sanskrit-Tibetan transliteration system to the Mongolian and the Manchu Ali-Gali script, in O. Corff (ed.), *Religion and State in the Altaic World*. <https://www.degruyter.com/document/doi/10.1515/9783110730562-006/html> (accessed 23 September 2024).
- Khaltarkhuu, G. & A. Maeda. 2015. Retrieval in Texts with Traditional Mongolian Script Realizing Unicoded Traditional Mongolian Digital Library. *Ritsumeikan University* 1–8. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=193933f95272bdb2adaea5ec47499359c9ede357> (accessed 11 June 2024).
- Li, B.W. 2020. Allocating Code Points for “Feminine and Masculine Characters” in the Mongolian Encoding. *Topical Document List: Mongolian*. <https://www.unicode.org/mwg/mwg4docs/mwg4-6CodePointsForFeminine-MasculineCharacters.pdf> (accessed 11 September 2024).
- Liang, H. 2018. 改进的蒙古文编码字形模型 [An improved graphetic model for the Mongolian encoding]. *Topical Document List: Mongolian*. <https://www.unicode.org/mwg/mwg2docs/mwg2-4Animprovedgrapheticmodel-LiangHai.pdf> (accessed 11 June 2024).
- Liang, H. 2019a. Towards a Well-Formed Mongolian Specification That Allows Interoperable Implementations. *Topical Document List: Mongolian*. <https://www.unicode.org/mwg/mwg3docs/mwg3-8MongolianSpecificationR.pdf> (accessed 23 September 2024).

- Liang, H. 2019b. Preliminary Analysis of L2/19-221 (WG2 N5073), China National Body's Comments on Mongolian. *Topical Document List: Mongolian*. <https://www.unicode.org/L2/L2019/19297-prelim-analysis-mong.pdf> (accessed 23 September 2024).
- Liang, H. 2019c. Text Representation and Shaping Specification of the Mongolian Script. *Topical Document List: Mongolian*. <https://www.unicode.org/L2/L2019/19368-draft-utn-mongolian.pdf> (accessed 23 September 2024).
- Liang, H. 2020. Proposal to Add Mongolian Free Variation Selector Four. *Topical Document List: Mongolian*. <https://www.unicode.org/L2/L2020/20057-mongolian-fvs4.pdf> (accessed 23 September 2024).
- Liang, H. 2024. Proposal to Refer to UTN #57 for Implementing the Mongolian Script. *Topical Document List: Mongolian*. <https://www.unicode.org/L2/L2024/24180-mongolian-utn.pdf> (accessed 18 April 2025).
- Liang, J.B. 2018a. Information technology – The transferring rules of traditional Mongolian nominal form to variant form 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. Hohhot: Ethnic Affairs Committee of the Inner Mongolia Autonomous Region and Inner Mongolia University Press.
- Liang, J.B. 2018b. A Contrastive Study on Phonetic Code Program and Graphic Code Program. *Topical Document List: Mongolian*. <https://www.unicode.org/mwg/mwg2docs/mwg2-8Comparisonofgrapheticandphonetic-LiangJinbao.pdf> (accessed 11 June 2024).
- Liang, J.B. 2019. An Introduction to the Execution of our Solution with Minimal Modifications – for achieving a stable and unified Mongolian encoding as soon as possible. *Topical Document List: Mongolian*. <https://www.unicode.org/mwg/mwg3docs/mwg3-11IntroductionToSolutionWithMinimalModificationsV1R.pdf> (accessed 9 April 2025).
- Ma, F.W. 2020. A Proposal to Add Characters to Mongolian Coding. *Topical Document List: Mongolian*. https://www.unicode.org/mwg/mwg4docs/mwg4-7Mongolian_MaFuquan.pdf (accessed 23 September 2024).
- Menghejiya. 2018. Revised Proposal for Minimizing the Current Phonetic Code. <https://www.unicode.org/L2/L2018/18103-mwg2-7-updates-phonetic.pdf> (accessed 23 September 2024).
- Ministry of Education and Science of Mongolia. 2018. *Үндэсний бичиг VI* [The National Script VI]. Ulaanbaatar: Ministry of Education and Science of Mongolia.
- Ministry of Education of China. 2021. 中国语言文字概况 [Languages of China]. Ministry of Education of China. http://www.moe.gov.cn/jyb_sjzl/wenzi/202108/t20210827_554992.html (accessed 11 June 2024).
- Montsame. 2020. Монгол бичгийн үндэсний хөтөлбөр III [National Mongolian Script Programme III]. Montsame.Mn. <https://montsame.mn/mn/read/224500> (accessed 11 June 2024).

- Moore, L. 2018. Mongolian Working Group Meeting 2 Report. *Topical Document List: Mongolian*. <https://www.unicode.org/mwg/mwg2docs/mwg2-14Meetingreport-LisaMoore.pdf> (accessed 11 June 2024).
- Moore, L. 2019. Summary of MWG2 Outcomes and Goals for MWG3 Meeting. *Topical Document List: Mongolian*. <https://www.unicode.org/mwg/mwg3docs/mwg3-13SummaryOfMWG2GoalsForMWG3R3.pdf> (accessed 11 June 2024).
- Namujila, C. 2020a. Preliminary Proposal for Encoding the Mongolian Usgiin Solution in the SMP of the UCS. *Topical Document List: Mongolian*. https://www.unicode.org/mwg/mwg4docs/mwg4-4PreliminaryProposal-ChuckN_0214.pdf (accessed 1 April 2025).
- Namujila, C. 2020b. The Simplified Mongolian Solution – Unicode. *Topical Document List: Mongolian*. <https://www.unicode.org/mwg/mwg4docs/mwg4-5Slides-ChuckN.pdf> (accessed 1 April 2025).
- News.MN. 2017. Төв Архивт Хадгалдаг Хосгүй Үнэт Баримтуудыг Дэлгэлээ [Unparalleled valuable documents stored in the central archives were revealed]. *News.MN*. <https://news.mn/r/765299/> (accessed 11 June 2024).
- Normile, I.H. 2023. Exploring criticality in Chinese philosophy: refuting generalisations and supporting critical thinking. *Studies in Philosophy and Education* 42: 123–41. <https://doi.org/10.1007/s11217-022-09855-3>.
- Pan, Y., D. Fan, H. Wu & D. Teng. 2023. A new dataset for Mongolian online handwritten recognition. *Scientific Reports* 13(26): 1–16. <https://doi.org/10.1038/s41598-022-27267-8>.
- Parliament of Mongolia. 2015. МОНГОЛ УЛСЫН ХУУЛЬ - МОНГОЛ ХЭЛНИЙ ТУХАЙ [The law of Mongolia – about the Mongolian language]. <https://legalinfo.mn/mn/detail/10932> (accessed 11 June 2024).
- Qing Government. 1929. 同文韻統 [*Tong Wen Yun Tong*]. Beijing: Place Museum.
- Quejingzhabu. 1999. 蒙古文编码 [Mongolian Script Encoding]. Hohhot: Inner Mongolia University Press.
- Samr.gov.cn. 2010. GB/T 25914–2010 Information Technology – Mongolian Presentation Forms Character Set and Use Rules of Controlling Characters. <https://std.samr.gov.cn/gb/search/gbDetailed?id=71F772D7F8C3D3A7E05397BE0A0AB82A> (accessed 1 June 2023).
- Samr.gov.cn. 2018. GB/T 36331-2018 Information technology – Uigur-Mongolian Characters, Presentation Characters and Use Rules of Controlling Characters. <https://openstd.samr.gov.cn/bz/gk/gb/newGbInfo?hcno=DFE87CC79EA67F8BF8B37C9C41CF9348> (accessed 9 April 2025).
- Samr.gov.cn. 2024. GB/T 25914-2023 Information Technology – Traditional Mongolian Nominal Characters, Presentation Characters and Use Rules of Controlling Characters. <https://openstd.samr.gov.cn/bz/gk/gb/newGbInfo?hcno=BD6429DE5A7FC782FAAE13938A07166E> (accessed 23 September 2024).

- Sanlig, B. & M.U. Enkhtur 2018. Solution for NNBS P Issues. *Topical Document List: Mongolian*. <https://www.unicode.org/L2/L2018/18293-nnbsp-solution.pdf> (accessed 11 June 2024).
- Sanlig, B. & J. Togoobat. 2018a. Analysis of the Graphetic Model and Improvements to the Current Model. *Topical Document List: Mongolian*. <https://www.unicode.org/L2/L2018/18101-mwg2-3-graphetic-analysis.pdf> (accessed 11 June 2024).
- Sanlig, B. & J. Togoobat. 2018b. Bolorsoft CEO and consultant discuss Unicode standard of Unicode. Interviewed by M. Oyungerel. *The UB Post*, 27 April. <https://www.pressreader.com/mongolia/the-ub-post/20180427/281522226698447> (accessed 12 February 2025).
- Sanlig, B., J. Togoobat & M.U. Enkhtur. 2018. Proposal to Encode Two Mongolian Letters. *Topical Document List: Mongolian*. <https://www.unicode.org/L2/L2018/18294-two-mongolian-ltrs.pdf> (accessed 11 June 2024).
- Script Ad Hoc Group Recommendations. 2017. A Graphetic Approach for the Mongolian Encoding Model. *Topical Document List: Mongolian*. <https://www.unicode.org/L2/L2017/17335-n4889-graphetic.pdf> (accessed 1 April 2024).
- Shen, Y.L. 2017. Positional Mismatches in Mongolian Encoding 蒙系文字编码中的位置失配. *Topical Document List: Mongolian*. <https://www.unicode.org/L2/L2017/17332-mong-pos-matches.pdf> (accessed 9 April 2025).
- Shen, Y.L. 2018. G2P 排序—形码蒙文的自动化自然排序法 [G2P sorting: an automated natural sorting method for graphetically encoded Mongolian]. *Topical Document List: Mongolian*. <https://www.unicode.org/mwg/mwg2docs/mwg2-5G2Psorting-ShenYilei.pdf> (accessed 11 June 2024).
- Shen, Y.L. 2019a. Filling the Gap Between Orthographic Patterns and Shaping Specs: A Case Study of Hudum g. *Topical Document List: Mongolian*. <https://www.unicode.org/mwg/mwg3docs/mwg3-9FillingTheGapR4.pdf> (accessed 23 September 2024).
- Shen, Y.L. 2019b. On Mongolian Punctuation Marks. *Topical Document List: Mongolian*. <https://www.unicode.org/mwg/mwg3docs/mwg3-10MongolianPunctuationMarksR2.pdf> (accessed 11 June 2024).
- Si, L. 2009. 蒙古文编码转换通用算法研究 [Research on the general algorithm of Mongolian encoding conversion]. *Journal of Inner Mongolia University* 41 (2): 133–136.
- Sodnom, D., P. Bayar & B. Gunbileg. 2008. An introduction to modern legal system of Mongolia: problems and solutions. *Asian Legal Studies* 39: 9–61. https://www.klri.re.kr/cmm/fms/FileDown.do?atchFileId=FILE_00000000008295LDuvR&fileSn=0&stat=journal¶m_gubun=A&data_no=8 (accessed 26 May 2025).
- Soni, S.K. 2024. Inner Asia in the ‘geopolitical game’ of the early 20th century impact on the trajectory of Mongol-seeking independence, in M.R. Debata and S. Hanova (eds), *India and Inner Asia Commerce, Culture and Connectivity*: 44–57. New York (NY): Routledge.

- Surgaltu, Unenmongke, Erhim & Orlog. 2020. A Character Encoding Model for Preserving the Mongolian Alphabetical System. <https://www.unicode.org/mwg/mwg4docs/mwg4-2PreservingMongolianAlphabeticalSystem.pdf> (accessed 24 September 2024).
- Suyalatu, S. 2001. Coding study of Mongolian whole words. *Journal of Chinese Information Processing* 15: 57–64. https://caod.oriprobe.com/articles/3242319/Coding_Study_of_Mongolian_Whole_Words.htm (accessed 17 April 2025).
- Tan, C. 2017. A Confucian conception of critical thinking. *Journal of Philosophy of Education* 51(1): 331–43. <https://doi.org/10.1111/1467-9752.12228>.
- Umamaheswaran, V.S. & WG2 Secretary. 2019. Recommendations from WG 2 Meeting 68 [See M68.16 re Mongolian.]. *Topical Document List: Mongolian*. <https://www.unicode.org/L2/L2019/19247-n5054-wg2-m68-recs.pdf> (accessed 11 June 2024).
- Unicode. 2006. Chapter 13.5 Mongolian, in Unicode 5.0.0. Unicode Consortium. <https://unicode.org/versions/Unicode5.0.0/> (accessed 11 June 2024).
- Unicode. 2023. The Unicode® Standard: A technical introduction. Unicode Consortium. <https://unicode.org/standard/principles.html> (accessed 11 June 2024).
- Unicode. 2024a. Topical Document List: Mongolian. Unicode Consortium. <https://unicode.org/L2/topical/mongolian/> (accessed 11 June 2024).
- Unicode. 2024b. Principles of the Unicode Standard, in *The Unicode® Standard: A technical introduction*. Unicode Consortium. <https://www.unicode.org/standard/principles.html> (accessed 11 June 2024).
- Unicode. 2024c. The Unicode Standard, Version 16.0 Mongolian Range: 1800–18AF. *Unicode Consortium* [dataset]. <https://www.unicode.org/charts/PDF/U1800.pdf> (accessed 7 October 2024).
- Unicode. 2024d. 13.5 Mongolian, in *The Unicode Standard, Version 16.0 – Core Specification*. San Francisco: Unicode Consortium.
- Wang, B., X. Shi & Y. Chen. 2016. Coping with problems of Unicoded traditional Mongolian, in M. Sun, X. Huang, H. Lin, Z. Liu & Y. Liu (eds), *Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data*. https://doi.org/10.1007/978-3-319-47674-2_11.
- Xiao, E. 2020. China cracks down on Mongols who say their culture is being snuffed out. *Wall Street Journal*, 4 September. <https://www.wsj.com/articles/china-clamps-down-on-inner-mongolians-protesting-new-mandarin-language-rules-11599132973> (accessed 11 June 2024).