

Proposal to encode Old Uyghur in Unicode

Anshuman Pandey
pandey@umich.edu

November 30, 2018

1 Introduction

This proposal is a revision of the following:

- L2/18-126: “Preliminary proposal to encode Old Uyghur in Unicode”

It incorporates comments and recommendations made by the UTC Script Ad Hoc Committee in:

- L2/18-168: “Recommendations to UTC #155 April-May 2018 on Script Proposals”

The major changes are as follows:

- Inclusion of a distinctive letter for *shin*
- Inclusion of an alternate form of independent *aleph*
- Addition of three new combining signs
- Addition of an editorial sign
- Improvements to the representative glyphs for *gimel* and *zayin*

This proposal has been reviewed by the following expert:

- Dai Matsui (Graduate School of Letters, Osaka University)

2 Background

The ‘Old Uyghur’ script was used between the 8th and 17th centuries primarily in the Tarim Basin of Central Asia, located in present-day Xinjiang, China. It is a cursive-joining alphabet with features of an *abjad*, and is characterized by its vertical orientation. The script was part of a vibrant orthographic and textual culture across Central Asia. Old Uyghur was used for recording religious, literary and administrative documents in Turkic languages, as well as Chinese, Mongolian, Sanskrit, Sogdian, and Tibetan.

In addition to its multilingual breadth, Old Uyghur was used alongside other scripts. There are numerous documents in the Old Uyghur script with intralinear Han characters, and Chinese manuscripts with Turkic translations in Old Uyghur script. Several manuscripts contain the Old Uyghur script with interlinear Sanskrit

annotations in ‘Turkestani’ or Central Asian styles of Brahmi. The Old Uyghur script also occurs in records containing the Phags-pa script, and in annotations accompanying the Khitan large script in a manuscript fragment. Documents containing text in both the Old Uyghur and the Arabic scripts are also extant.

The scribal tradition of Old Uyghur may be divided into three styles based upon analysis of letterforms in various records: ‘cursive’, ‘formal’, and ‘standard’. It may also be periodized into ‘early’, ‘proper’, and ‘late’. The ‘cursive’ style occurs in numerous civil and administrative documents from the 9th through 15th centuries (see fig. 32–33). By the 10th century, scribal refinements resulted in a style that may be called ‘formal’ Old Uyghur, which was used for religious and literary manuscripts (see fig. 23–30). The script was developed further through printing. The introduction of block printing for producing texts gave rise to what may be considered a ‘standard’ Old Uyghur script in the 12th century. Numerous folios and fragments of block-printed books have been preserved (see fig. 12–20 for specimens). This ‘standard’ block-print style is similar to the formal inscriptional type, which appears on the stone walls of the Cloud Platform at Juyong Guan, Beijing, erected in the 14th century (see fig. 34).

Old Uyghur represents a prominent development of a script continuum that originates in the Old Sogdian script of the ‘Ancient Letters’ and extends to modern Mongolian. It developed from the ‘cursive’ style of the Sogdian script during the 8th–9th century (Kara 1996: 539). Just as Turkic peoples adopted the Sogdian script for writing their language, speakers of other languages in Central Asia adapted the Old Uyghur script for their own writing practices. A popular narrative states that in the 13th century, during the reign of Genghis Khan, the scholar Tata Tonga developed an orthography for writing the Mongolian language using the Old Uyghur script. The Uyghur-based Mongolian script developed into a distinctive script with its own scribal and print culture. The Old Uyghur script flourished through the 15th century and was replaced by a new orthography based upon the Arabic script in the 16th century. It appears that Old Uyghur was used in Gansu through the 17th century.

At the beginning of the 20th century, German and Russian scholars adapted the Old Uyghur script for modern typesetting. Buddhist texts in the Old Uyghur script were edited and published by V. V. Radlov and others (see fig. 35, 36). The metal types used in these editions appear to have been cut to match the letterforms found in Old Uyghur block-prints.

There has been active scholarship on the Old Uyghur script and manuscripts since the early 20th century. It was during this time that European expeditions to Turfan unearthed vast amounts of materials in Old Uyghur and other scripts. The past century has witnessed increasing growth of interest in Old Uyghur sources of the 8th through 15th centuries within studies of cultures, peoples, and polities of the Silk Road. Various institutions that obtained materials from Turfan and other sites have digitized their collections or are in the process of doing so, such as the Berlin-Brandenburgische Akademie der Wissenschaften (BBAW), British Library, and other institutions associated with the International Dunhuang Project (IDP).

3 Script identifier

The proposed Unicode identifier for the script is ‘Old Uyghur’, which is a scholarly designation. The name applies specifically to the script within the context of Unicode, and it does not apply to any language, culture, or community. The script is also known generically as ‘Uyghur’, without the descriptor ‘Old’. The term ‘Uyghur’ has variant transliterations / transcriptions / spellings in English, such as ‘Uighur’, ‘Uigur’, ‘Uygur’, ‘Uyğur’, as well as ‘Ouïgour’, etc’ in French and other languages.

To be sure, neither ‘Uyghur’ nor ‘Old Uyghur’ is an entirely accurate designation for the script. The renowned Turkologist, Gerard Clauson notes that the “name is probably as anachronistic as that name when

applied to the language” (1962: 100). The script had been in use in Central Asia before the Uyghur language became prominent in the 8th century (1962: 43). However, Clauson concludes that “no useful purpose would be served by suggesting some other name” (1962: 100–101). This proposal abides by Clauson’s conclusion. For purposes of identifying the script in Unicode, the adjective ‘Old’ is appended to ‘Uyghur’ in order to distinguish the script from the later Arabic orthography used for writing the modern Uyghur language, which is not directly related to the Uyghur language of the 8th century. Given the polysemia of ‘Uyghur’, the term ‘Old Uyghur’ has become common for referring to the script, even if it is imprecise.

4 Encoding history

4.1 Previous proposals

A proposal for Old Uyghur was previously submitted to the Unicode Technical Committee (UTC) by Omarjan Osman in 2013. Osman’s “Proposal to Encode the Uyghur Script in ISO/IEC 10646” (L2/13-071) provides valuable background on the history and usage of the script, and details about the representation of letterforms and orientations of the script in different manuscripts. Based upon the provenance and attributes of two important sources, Osman identified two major variations of the script along a geographic basis. He describes the ‘western’ form as being written horizontally from right to left, and an ‘eastern’ form that is written vertically from top to bottom (p. 11). Osman thought it necessary to accommodate both orientations of the script at the character level. Therefore, he proposed a repertoire that contains upright glyphs for the horizontal form (for right-to-left display), and the same glyphs rotated 90 degrees counter-clockwise for the vertical form.

The model presented in L2/13-071 is ambitious, but it is not practical for purposes of character encoding. It is also incompatible with the Unicode character-glyph model. The encoding of separate characters for horizontal and vertical orientations of a letter results in a model that establishes separate semantic values for glyphic variants of a given letter. Such a repertoire is redundant and prone to complications, for example, errors caused by usage of a horizontal letter in a string of vertical characters, etc. It would be more appropriate to consider such glyphs as directional variants instead of separate characters. Moreover, instead of attempting to accommodate orientations of the script at the character level, it would be practical to use mark-up and layout to achieve the desired display. Nonetheless, Osman’s proposal is a useful resource for further investigating the requirements for encoding Old Uyghur. His proposed repertoire includes digits and several diacritics (whose exact provenance is not given), which must be investigated in order to determine a complete character repertoire for representing Old Uyghur texts.

4.2 Existing standards

There are no existing formal standards for the Old Uyghur script. The closest related digital standard for the script is the Unicode encoding for Mongolian. Recently, the government of China published a standard known as “GB/T 36331-2018 ‘Information technology – Uigur-Mongolian characters, presentation characters and use rules of controlling characters’”. According to Liang Hai, GB/T 36331-2018 is a subset of GB/T 26226-2010, which is China’s standard for encoding Mongolian — based upon the complete Unicode encoding for the script — and equivalent to Mongolia’s MNS 4932: 2000. Another subset of GB/T 26226-2010 is GB/T 25914-2010, which provides a standard for the modern writing system for the Mongolian language. Given the reference to “Uigur-Mongolian”, it is apparent that the standard is intended for the representation of the early stages of the Mongolian script, using the phonemic model of the Unicode encoding and similar glyphs. However, it is not specifically a standard for Old Uyghur.


5 Approach to the Encoding

The proposed encoding for Old Uyghur aligns with Unicode principles and is justified according to the following factors:

- *Standardization* The development of a block-printing standard established a distinctive Old Uyghur script. There is a need to represent content in this script in plain text. In particular, it is necessary to contrast texts in Old Uyghur from Sogdian and Mongolian in order to distinguish texts in these scripts for text processing and digitization on the basis of character identity and semantics.
- *Stylistic variation* Over the course of its development the Old Uyghur script developed its own stylistic variants. A distinctive Old Uyghur block provides a means for managing the taxonomy of these script variants by unifying the ‘cursive’, ‘formal’, and print ‘standard’ forms. The representative form of the script is based upon the block print style. Other styles may be displayed using fonts.
- *Character repertoire and semantics* The repertoire, order, and names of Old Uyghur letters is based upon that of Sogdian. The proposed encoding for Old Uyghur retains these attributes. The Mongolian encoding uses different names and ordering for letters, which reflect Mongolian preferences and pronunciations. Mongolian letter names do not correspond to Old Uyghur values.
- *Glyphic distinction* Following from the above, a separate encoding preserves the glyphic distinctions of formal Old Uyghur in multilingual contexts that include Sogdian and Mongolian text. In particular, Mongolian glyphs do not adequately transmit the aesthetic and orthographic features of Old Uyghur.
- *Encoding model* A practical encoding for Old Uyghur is based upon a graphetic model. The base letters of the script would be encoded, while letters with diacritics would be decomposed into a base and a combining sign. The graphetic model for Old Uyghur facilitates implementation. The current model for Mongolian, which has a phonetic basis, presents several issues and is unsuitable for Old Uyghur. The proposed model for Uyghur offers a practical implementation for a vertical script that avoids the complications of the Mongolian model.
- *Orientation* The default orientation for Old Uyghur is vertical. On the other hand, Sogdian is defined as a horizontal script. Mongolian is defined as a vertical script, but it differs from Old Uyghur in its representation in horizontal layouts.

6 Script details

6.1 Structure


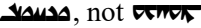

The Old Uyghur script is a cursive joining alphabet. Letters have nominal shapes when they occur in isolation and contextual shapes when they occur in initial, medial, or final position. All letters are defined as dual joining. In some sources the connection between letters is suspended (see § 8). This feature may be supported by usage of the control character  U+200C ZERO WIDTH NON-JOINER (abbreviated as ZWNJ). Although derived from an *abjad*, Old Uyghur possesses the features of an ‘alphabet’. Short vowels are generally indicated, and diagraphs and trigraphs are used for denoting the rich vowel repertoire of Turkic languages. Diacritics are used for diambiguating letters with similar appearances and for representing sounds for which distinctive letters do not exist.

6.2 Directionality

The standard writing direction for Old Uyghur is vertical, from top to bottom in columns that run from left to right. This orientation is ascertained by the majority of extant Old Uyghur documents, particularly those produced using block printing. The vertical orientation of Old Uyghur is confirmed especially in biscriptal documents containing Han characters and Central Asian Brahmi. In the documents from the 14th century, which may be considered the last phase of usage, the script is written horizontally. However, printed editions of Old Uyghur texts express fidelity to the standard vertical orientation (see fig. 35, 36).

When rendering Old Uyghur text in a system that does not support vertical layout, the text should be oriented in horizontal lines that run from right to left, and from top to bottom. This orientation is identical to conventional right-to-left scripts such as Sogdian and Arabic. In such cases, the glyphs of Old Uyghur letters are to be rotated 90 degrees clockwise with respect to their orientation in the code chart.

The above orientation is preferred over that specified for Mongolian in non-vertical environments. Mongolian is also a vertical script that runs from right to left, top to bottom. The Unicode core specification suggests that in non-vertical contexts Mongolian text should be laid out in horizontal lines running from left to right with glyphs as represented in the code chart. Mongolian glyphs are shown in the code chart rotated 90 degrees clockwise with respect to their conventional vertical orientation. In Latin or other left-to-right contexts, this glyph orientation results in Mongolian text being rendered upside down in relation to the letters of the surrounding script.


Accordingly, the Old Uyghur word  should appear in horizontal contexts as , not . However, as specified above, when Old Uyghur occurs in horizontal layouts, a right-to-left orientation with glyphs upright as in Latin is practical for purposes of legibility. This orientation was preferred by scholars, especially because it was convenient to reference Old Uyghur words and phrases in multilingual contexts that also contain Arabic, Cyrillic, Devanagari, Tibetan, and other scripts (see fig. 42). Given the global distribution of scholars of Old Uyghur and Turkic studies, it is likely that these users will prefer to read the script with glyphs oriented upright when it appears in non-vertical layouts.

Following upon the above, throughout this document the proposed Old Uyghur characters are presented in their conventional vertical forms, while references to letters and signs within a Latin-script environment are given in a right-to-left orientation.

7 Proposed repertoire

Scope The proposed Old Uyghur repertoire contains the full alphabet and a set of diacritics that occur commonly across the primary sources. This repertoire is sufficient for representing the majority of Old Uyghur texts. There are other diacritics, punctuation, digits, and other symbols, as such those shown in Omarjan’s proposal, that require additional research before a determination may be made regarding their suitability for encoding. These characters may be added to the repertoire in the future.









Character inventory The proposed repertoire for Old Uyghur contains 35 characters: 18 basic letters, 1 alternate letter, 7 combining signs, 6 punctuation signs, 1 stem-extending sign, and 1 editorial sign. The code chart and names list follows p. 12. The encoded set may differ from traditional and scholarly inventories of the script. Such differences naturally arise from the requirements for digitally representing a script in plain text and for preserving the semantics of characters.










Character names In the scholarly literature, Uyghur letters are known by the names of the original Aramaic letters from which they are derived. This nomenclature follows the scholarly convention for letters of the Sogdian scripts. The same is followed here. Throughout this proposal, italics are used for scholarly names for graphemes, while small capitals indicate Unicode character names, eg.  is *aleph* and OLD UYGHUR LETTER ALEPH. For brevity, the descriptor ‘OLD UYGHUR’ may be dropped, eg. OLD UYGHUR LETTER ALEPH is truncated to ALEPH. Characters of other scripts are designated by their full Unicode names. Latin transliteration of Old Uyghur follows the current scholarly convention.





Representative glyphs Representative glyphs and calligraphic features are based upon the script used in block prints, exemplified by the manuscript in fig. 12, and those in 13–20. The block-print style is the most suitable representative of the Old Uyghur script because it may be considered a ‘standard’. In general, preparing a script for printing requires making conscientious decisions about letterforms, calligraphic features, and general aesthetics of a script. Moreover, the block prints exemplify an Old Uyghur script that is distinctive. The font used here was designed by the proposal author.







7.1 Letters








The following letters are a superset of those used in Old Uyghur documents. The representative glyphs are based upon independent or isolated forms of letters. An inventory of letters appears in the margin of the manuscript U 40 (see fig. 26).



Character name	Glyph	Joining	Latin
OLD UYGHUR LETTER ALEPH		dual	ʾ
OLD UYGHUR LETTER INDEPENDENT ALEPH WITH VERTICAL TAIL		non	-ʾ
OLD UYGHUR LETTER BETH		dual	β
OLD UYGHUR LETTER GIMEL		dual	γ
OLD UYGHUR LETTER WAW		dual	w
OLD UYGHUR LETTER ZAYIN		dual	z, ž
OLD UYGHUR LETTER HETH		dual	x, q
OLD UYGHUR LETTER YODH		dual	y
OLD UYGHUR LETTER KAPH		dual	k
OLD UYGHUR LETTER LAMEDH		dual	δ






OLD UYGHUR LETTER MEM		dual	m
OLD UYGHUR LETTER NUN		dual	n
OLD UYGHUR LETTER SAMEKH		dual	s, š
OLD UYGHUR LETTER PE		dual	p
OLD UYGHUR LETTER SADHE		dual	c
OLD UYGHUR LETTER RESH		dual	r
OLD UYGHUR LETTER SHIN		dual	š
OLD UYGHUR LETTER TAW		dual	t
OLD UYGHUR LETTER LESH		dual	l







aleph In some manuscripts a word-final *aleph* is written detached from the previous letter, regardless of joining behavior, using the alternate form , which is distinguished from the representative  in terms of the vertical orientation of the terminal (see fig. 8). In some cases, the letters are used for distinguishing between final *a* () and *ä/e* () , see fig. 2; also see forms used for *-a* in fig. 5. Experts, such as Dai Matsui, state that there is a practical need to distinguish between these forms of *aleph* for purposes of research: when studying Old Uyghur manuscripts, there is a requirement to reproduce the orientation of the tail in order to faithfully document fragmented and illegible letters for aiding future decipherment.






gimel and heth In isolated and final positions these letters are distinguished using the glyphs  and , respectively; however, they have the same initial and medial forms. The sound *q* is represented by writing the diacritic  above  *heth* (see § 9.2). In initial and medial positions the same diacritic is used for distinguishing *gimel* from *heth*, eg. medial  *gimel* and medial  *heth*. When representing these letters in these positions, the correct letter should be used in order to preserve the underlying semantics of the text.

zayin The representative form  of *zayin* is based upon the ‘sawtooth’ form that occurs in the majority of hand-written Old Uyghur documents. This form differs from the glyph  used in block prints and certain inscriptions. The selection of the ‘sawtooth’ form is based upon the requirement to provide the most representative glyph for *zayin*. Usage of the ‘sawtooth’  may be controlled using fonts. In addition to variations in appearance, *zayin* also has varying joining behavior. In some sources  *zayin* does not connect to a following letter; however, it does in other sources. It is defined as a dual-joining letter in order to enable the left-joining feature. The control character  ZWNJ may be placed after ZAYIN to prevent joining with the following letter. In some sources *zayin* is distinguished using diacritics as  and  (see § 9.2).


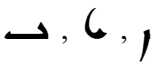




yodh In some sources, the  *yodh* does not connect to a following letter. The control character  ZWNJ may be placed after YODH to prevent joining with the following letter.

nun The  *nun* may be written with a dot above  to distinguish it from  *aleph* in non-final positions, eg. medial  *nun* and medial  *aleph* (see § 9.2).


samekh and **shin** The letters  *samekh* (/s/) and  *shin* (/š/) are palaeographically distinctive letters in the script (Clauson 1962: 109). By the 13th or 14th century, however, they merged into a single letter, eg.  (see fig. 6). The two letters are distinguished by the fact that *shin* is written using two strokes (with a right-downward ductus before the leftward curve of the second), while *samekh* is a single stroke (see fig. 9). In late documents in which /s/ and /š/ are written using the same letter, the latter sound is made explicit by applying the diacritic  to  *samekh*, ie.  (see § 9.2).

lesh The letter  is known as ‘hooked r’ and represents the sound [l]. It is derived from  U+10F44 SOGDIAN LETTER LESH, which likely evolved from the practice of indicating [l] by placing a subscript *resh* ( U+10F4F SOGDIAN COMBINING RESH BELOW) beneath  U+10F40 SOGDIAN LETTER RESH-AYIN to produce . The letter has been assigned the name ‘LESH’ based upon the name of the corresponding letter in the Sogdian block. This is not a historical name, but one suggested by modern scholars as a neologism that aligns with the Aramaic naming convention. The name ‘hooked r’ has been specified as an alias in the names list.

Note on final forms The terminal stroke of the final form of the following letters may be written in different directions, even within the same source.




	conventional	alternate
final ALEPH		
final KAPH		
final NUN		

The orientation of terminals vary according to the whim of the scribe or the space available on a page. Terminal variation occurs most often at the end of a line for filling space or for compensating for lack of space at a margin. These stroke variations are stylistic and there is no semantic difference between final forms with different terminals.

Of these, only  has been proposed for separate encoding, as INDEPENDENT ALEPH WITH VERTICAL TAIL. The other alternate final forms may be considered stylistic variants and should be controlled through fonts.

7.2 Combining signs

The following combining signs are used for disambiguation and representation of new sounds (see § 9.2):

Character name	Glyph
OLD UYGHUR COMBINING DOT RIGHT	
OLD UYGHUR COMBINING TWO DOTS RIGHT	
OLD UYGHUR COMBINING THREE DOTS RIGHT	

OLD UYGHUR COMBINING DOT LEFT	◌◌̇
OLD UYGHUR COMBINING TWO DOTS LEFT	◌◌̈̈
OLD UYGHUR COMBINING THREE DOTS LEFT	◌◌̉̉̉
OLD UYGHUR COMBINING HAMZA LEFT	◌◌̋

These combining signs have a function similar to the ‘nukta’ diacritic, which is used in other scripts for creating new letters. While it may be possible to encode combinations of base letter + combining sign as atomic letters, it is practical to avoid such an approach. The recommendation against atomic letter is that the number of combining signs for Old Uyghur is likely to grow. As a result, it will be necessary to encode new sets of atomic letters for each every base letter + combining sign combination when a new combining sign is added to the repertoire. The proposed approach of using combining signs follows the model for Sogdian, from which Old Uyghur is derived.

The signs ◌̇, ◌̈̈, etc. are analogous to Sogdian diacritics, eg. ◌̇ U+10F46 SOGDIAN COMBINING DOT BELOW and ◌̈̈ U+10F47 SOGDIAN COMBINING TWO DOTS BELOW. In Old Uyghur block prints dot diacritics appear as elongated strokes, which are reflective of the stylistic aesthetic of the print standard for the script. In other manuscripts, such diacritics are written as true dots. Despite the representative shape, it is appropriate to reference to these characters as ‘dot’ or ‘dots’.

The descriptors ‘right’ and ‘left’ refer to the placement of these signs with respect to the base letter in the traditional vertical orientation of the script. When Old Uyghur occurs in non-vertical contexts, the signs labeled ‘right’ would be placed below the base, and the signs labeled ‘left’ would occur above the base letter. For example, in horizontal layout the ◌̇ COMBINING DOT RIGHT would appear as ◌̇, a ‘below-base’ sign.

The signs ◌̋, ◌̌, and ◌̍ are used for representing non-Turkic sounds, particular when writing words of Arabic origin (see fig. 10).

There are other signs, such as ◌◌̋ ring below, which are used in some documents for transcription. Erdal (1984) describes some diacritic signs used for diambiguation and transliteration of Arabic in administrative documents in the Old Uyghur script of the 11th century from Yarkand. Clark (2010) also describes some signs used in the Old Uyghur manuscript of the *Kutadgu Bilig*, an 11th century Karakhanid work by Yusūf Khāṣṣ Ḥājīb. Further research is required to determine the complete set of these signs and the method for encoding them.

7.3 Punctuation signs

Spacing is used for separating words. The following signs are used for punctuation:

Character name	Glyph
OLD UYGHUR PUNCTUATION BAR	◌̋
OLD UYGHUR PUNCTUATION TWO BARS	◌̌

OLD UYGHUR PUNCTUATION TWO DOTS	⋮
OLD UYGHUR PUNCTUATION FOUR DOTS	❖
OLD UYGHUR PUNCTUATION FIVE DOTS	❖❖
OLD UYGHUR SECTION MARK	ⵛ

The signs ⵛ and ❖❖ are common forms of punctuation. They are used for delimiting text segments of various lengths, such as sentences. When these two signs are used together, ⵛ indicates smaller segments, while ❖❖ marks longer sections (see fig. 14, 19). The sign ⵛ is also used as a general delimiter. When it occurs in documents where ⵛ is used, it represents short segments of text and may function as a comma or semi-colon (see fig. 20).

The sign ❖ is used for indicating the end of a section. While this sign is similar to the ⋮ U+2058 FOUR DOT PUNCTUATION already encoded in Unicode, the Old Uyghur ❖ is used in a vertical environment and is, therefore, proposed for encoding as a script-specific character.

Similarly, the ❖❖ is used as a general sign of punctuation and decoration, for example in fig. 16, is also found in Sogdian documents, but is encoded as a script-specific sign because of directional considerations.

The ⵛ is used in the Juyong Pass inscription as a section mark.

7.4 Stem extender

The following character is used for extending the baseline (see § 9.3 for details). It is used as a typographic filler and also for indicating a suffix that is separated from the stem.

Character name	Glyph
OLD UYGHUR STEM EXTENDER	ⵀ

7.5 Editorial signs

The following editorial signs are used in manuscripts:

Character name	Glyph
OLD UYGHUR DELETION MARK	ⵀ

When written beneath a word or letter, this sign indicates that the respective text is an error and is to be omitted. The correct word is generally written after the misspelled word (see fig. 11).

7.6 Line-breaking

There are no formal rules for the breaking of Old Uyghur text at the end of line. Moreover, the available sources do not contain text with line-breaks for words. It may be assumed that words were not split at line boundaries. There are no indications of hyphens or other continuation marks. In digital layouts, line-breaks should occur after words.

7.7 Collation

The sort order for Old Uyghur follows the encoded order:











𐰀 ALEPH < 𐰁 BETH < 𐰂 GIMEL < 𐰃 WAW < 𐰄 ZAYIN < 𐰅 HETH < 𐰆 YODH <
 𐰇 KAPH < 𐰈 LAMEDH < 𐰉 MEM < 𐰊 NUN < 𐰋 SAMEKH < 𐰌 PE < 𐰍 SADHE <
 𐰎 RESH < 𐰏 TAW < 𐰐 LESH







8 Joining behavior

The nominal ('X_n'), final ('X_f') medial ('X_m'), and initial ('X_i') forms of letters are shown below.


	X _n	X _f	X _m	X _i
ALEPH	ﺍ	ﺍ	ﺍ	ﺍ
BETH	ﺏ	ﺏ	ﺏ	ﺏ
GIMEL	ﻎ	ﻎ	ﻎ	ﻎ
WAW	ﻭ	ﻭ	ﻭ	ﻭ
ZAYIN	ﺯ	ﺯ	ﺯ	ﺯ
HETH	ﻩ	ﻩ	ﻩ	ﻩ
YODH	ﻱ	ﻱ	ﻱ	ﻱ
KAPH	ﻙ	ﻙ	ﻙ	ﻙ
LAMEDH	ﻝ	ﻝ	ﻝ	ﻝ
MEM	ﻡ	ﻡ	ﻡ	ﻡ
NUN	ﻥ	ﻥ	ﻥ	ﻥ
SAMEKH	ﺱ	ﺱ	ﺱ	ﺱ
PE	ﭗ	ﭗ	ﭗ	ﭗ
SADHE	ﺹ	ﺹ	ﺹ	ﺹ
RESH	ﺭ	ﺭ	ﺭ	ﺭ
SHIN	ﺶ	ﺶ	ﺶ	ﺶ
TAW	ﺕ	ﺕ	ﺕ	ﺕ
LESH	ﻝ	ﻝ	ﻝ	ﻝ







The shaping engine substitutes the nominal glyph for each letter in the input with the appropriate positional glyph to produce the expected joined output. In order to illustrate the joining properties of letters, representations of words from Old Uyghur records are given below along with their input strings:








<i>ʔltwn</i> 'altun'		ʔ ALEPH, ʔ ALEPH, ʔ LESH, ʔ TAW, ʔ WAW, ʔ NUN
<i>bʔmsʔn</i> 'vamsan'		ʔ BETH, ʔ ALEPH, ʔ MEM, ʔ SAMEKH, ʔ ALEPH, ʔ NUN
<i>wyʔwr</i> 'üigür'		ʔ WAW, ʔ YODH, ʔ GIMEL, ʔ WAW, ʔ RESH
<i>qwtlwy</i> 'qutlug'		ʔ HETH, ʔ COMBINING TWO DOTS LEFT, ʔ TAW, ʔ LESH, ʔ GIMEL
<i>mncwšry</i> 'mancusari'		ʔ MEM, ʔ NUN, ʔ SADHE, ʔ WAW, ʔ SAMEKH, ʔ COMBINING TWO DOTS RIGHT, ʔ RESH, ʔ YODH
<i>swδwr</i> 'sudur'		ʔ SAMEKH, ʔ WAW, ʔ LAMEDH, ʔ WAW, ʔ RESH
<i>pwδystb</i> 'bodisatav'		ʔ PE, ʔ WAW, ʔ LAMEDH, ʔ YODH, ʔ SAMEKH, ʔ TAW ʔ BETH,
<i>pwryʔn</i> 'burxan'		ʔ PE, ʔ WAW, ʔ RESH, ʔ GIMEL, ʔ ALEPH, ʔ NUN
<i>pylyk</i> 'bilig'		ʔ PE, ʔ YODH, ʔ LESH, ʔ YODH, ʔ KAPH
<i>twykʔl</i> 'tükäl'		ʔ TAW, ʔ WAW, ʔ YODH, ʔ KAPH, ʔ ALEPH, ʔ LESH

tnkry   TAW,  NUN,  KAPH,  RESH,  YODH
‘tängri’









8.1 Modification of cursive joining

In some texts certain letters do not join to a following letter in order to distinguish between letters that have similar appearances. The  U+200C ZERO WIDTH NON-JOINER (ZWNJ) is to be used for modifying cursive joining. The ZWNJ is placed after the letter whose connection is suspended. The letter is rendered using its final form and the following letter appears in its initial form.

’wyz’   ALEPH,  WAW,  YODH,  ZAYIN,  ALEPH
‘üzä’

’wyz’   ALEPH,  WAW,  YODH,  ZAYIN,  ZWNJ,  ALEPH
‘üz-ä’




pwlzwn   PE,  WAW,  LESH,  ZAYIN,  WAW,  NUN
‘bolzun’





pwlz-wn   PE,  WAW,  LESH,  ZAYIN,  ZWNJ,  WAW,  NUN
‘bolz-un’

8.2 Glyph interactions



The following letters have special behaviors when they interact with other letters.



8.2.1 *waw*

In initial and medial position, the tails of *kaph* and *pe* attach below the baseline of the following letter, eg.  <KAPH, NUN> and  <PE, NUN>. When these letters are followed by  *waw*, their tails curve into the body of the *waw* to produce a ligature:





Character sequence	Ligated	Unligated
<KAPH, WAW>		
<PE, WAW>		





8.2.2 *mem*

The extender of *mem* extends below the baseline in initial  and medial  positions. The extender of medial *mem* is written at an angle that slopes downward. The shaping of a word containing *mem* depends upon the position of the letter within the word:

- *Following a word-initial letter:* When a word-initial letter is followed by *mem*, the letter is enlarged and its baseline connects to the extender of *mem*, while the letter that follows *mem* joins to the body, eg.  <ALEPH, MEM, WAW>.
- *Following a non-initial letter:* When following after a non-word-initial letter, it is shifted towards the baseline and the preceding letter is angled downward in order connect to its extender. In such cases, the following letter is shifted away from the baseline, eg.  <ALEPH, ALEPH, MEM, WAW>.




8.2.3 *lesh*















When  *lesh* follows letters with elements that extend below the baseline, the hook is detached from *lesh* and placed beneath the extension of the previous letter:  <KAPH, LESH>,  <MEM, LESH>,  <PE, LESH>. Even if *lesh* does not immediately follow *kaph*, *mem*, or *pe*, its hook attaches to the terminal of the latter for aesthetic considerations:

	shifted hook	static hook
<i>pylyk</i> ‘bilig’		
<i>k’lm’dwk</i> ‘käl Mädük’		

9 Encoded representations

9.1 Vowels

Vowels are indicated using  *aleph*,  *waw*, and  *yodh*, and combinations of these letters in digraphs and trigraphs. Vowels are represented as follows:

	Initial	Medial
<i>ä</i>	  ALEPH	  ALEPH
<i>a, e</i>	  ALEPH,  ALEPH	  ALEPH
<i>i, i</i>	  ALEPH,  YODH	  YODH

\bar{i}, \bar{i}		⌈ ALEPH, ⌈ YODH, ⌈ YODH		⌈ YODH, ⌈ YODH
o, u		⌈ ALEPH, 9 WAW		9 WAW
\bar{o}, \bar{u}		⌈ ALEPH, 9 WAW, ⌈ YODH		9 WAW
\bar{o}, \bar{u}		9 WAW, ⌈ YODH		9 WAW, ⌈ YODH
$\bar{o}, \bar{o}, \bar{u}, \bar{u}$		⌈ ALEPH, 9 WAW, 9 WAW		9 WAW, 9 WAW


9.2 Disambiguation and extension of letters

The combining signs enumerated in § 7.2 are written with letters to disambiguate consonants or to represent consonants for which distinctive letters do not exist. The following forms are attested. Combining signs are placed after a letter in encoded text:

	X_n	X_f	X_m	X_i		
dotted <i>gimel</i>	γ					GIMEL, ◌ COMBINING DOT LEFT
two-dotted <i>gimel</i>	γ					GIMEL, ◌ COMBINING TWO DOTS LEFT
dotted <i>zayin</i>	ζ					ZAYIN, ◌ COMBINING DOT RIGHT
two-dotted <i>zayin</i>	ζ					ZAYIN, ◌ COMBINING TWO DOTS RIGHT
dotted <i>heth</i>	q					HETH, ◌ COMBINING DOT LEFT
two-dotted <i>heth</i>	q					HETH, ◌ COMBINING TWO DOTS LEFT
dotted <i>nun</i>	n					NUN, ◌ COMBINING DOT LEFT
two-dotted <i>samekh</i>	\check{s}					SAMEKH, ◌ COMBINING TWO DOTS RIGHT

9.3 Stem extension

In some texts, a space and a short extension of the baseline is used for indicating suffixes. For such cases the **STEM EXTENDER** may be used:

<p><i>tinly lr-r</i> 'tinlag-lar-r'</p>		<p> 𐰀 TAW, 𐰁 YODH, 𐰂 NUN, 𐰃 LESH, 𐰄 GIMEL, SP SPACE, 𐰃 LESH, 𐰅 RESH, SP SPACE, 𐰆 STEM EXTENDER, 𐰅 RESH </p>
---	---	--

If there is a need to indicate explicitly that the suffix belongs to the preceding word in encoded text, then ZWJ ZWNJ may be used before the STEM EXTENDER instead of a space.

10 Character Properties

10.1 Core data: UnicodeData.txt

```

10F70;OLD UYGHUR LETTER ALEPH;Lo;0;AL;;;;N;;;;;
10F71;OLD UYGHUR LETTER INDEPENDENT ALEPH WITH VERTICAL TAIL;Lo;0;AL;;;;N;;;;;
10F72;OLD UYGHUR LETTER BETH;Lo;0;AL;;;;N;;;;;
10F73;OLD UYGHUR LETTER GIMEL;Lo;0;AL;;;;N;;;;;
10F74;OLD UYGHUR LETTER WAW;Lo;0;AL;;;;N;;;;;
10F75;OLD UYGHUR LETTER ZAYIN;Lo;0;AL;;;;N;;;;;
10F76;OLD UYGHUR LETTER HETH;Lo;0;AL;;;;N;;;;;
10F77;OLD UYGHUR LETTER YODH;Lo;0;AL;;;;N;;;;;
10F78;OLD UYGHUR LETTER KAPH;Lo;0;AL;;;;N;;;;;
10F79;OLD UYGHUR LETTER LAMEDH;Lo;0;AL;;;;N;;;;;
10F7A;OLD UYGHUR LETTER MEM;Lo;0;AL;;;;N;;;;;
10F7B;OLD UYGHUR LETTER NUN;Lo;0;AL;;;;N;;;;;
10F7C;OLD UYGHUR LETTER SAMEKH;Lo;0;AL;;;;N;;;;;
10F7D;OLD UYGHUR LETTER PE;Lo;0;AL;;;;N;;;;;
10F7E;OLD UYGHUR LETTER SADHE;Lo;0;AL;;;;N;;;;;
10F7F;OLD UYGHUR LETTER RESH;Lo;0;AL;;;;N;;;;;
10F80;OLD UYGHUR LETTER SHIN;Lo;0;AL;;;;N;;;;;
10F81;OLD UYGHUR LETTER TAW;Lo;0;AL;;;;N;;;;;
10F82;OLD UYGHUR LETTER LESH;Lo;0;AL;;;;N;;;;;
10F83;OLD UYGHUR COMBINING DOT RIGHT;Mn;220;NSM;;;;N;;;;;
10F84;OLD UYGHUR COMBINING TWO DOTS RIGHT;Mn;220;NSM;;;;N;;;;;
10F85;OLD UYGHUR COMBINING THREE DOTS RIGHT;Mn;220;NSM;;;;N;;;;;
10F86;OLD UYGHUR COMBINING DOT LEFT;Mn;230;NSM;;;;N;;;;;
10F87;OLD UYGHUR COMBINING TWO DOTS LEFT;Mn;230;NSM;;;;N;;;;;
10F88;OLD UYGHUR COMBINING THREE DOTS LEFT;Mn;230;NSM;;;;N;;;;;
10F89;OLD UYGHUR COMBINING HAMZA RIGHT;Mn;220;NSM;;;;N;;;;;
10F8A;OLD UYGHUR PUNCTUATION BAR;Po;0;AL;;;;N;;;;;
10F8B;OLD UYGHUR PUNCTUATION TWO BARS;Po;0;AL;;;;N;;;;;
10F8C;OLD UYGHUR PUNCTUATION TWO DOTS;Po;0;AL;;;;N;;;;;
10F8D;OLD UYGHUR PUNCTUATION FOUR DOTS;Po;0;AL;;;;N;;;;;
10F8E;OLD UYGHUR PUNCTUATION FIVE DOTS;Po;0;AL;;;;N;;;;;
10F8F;OLD UYGHUR SECTION MARK;Po;0;AL;;;;N;;;;;
10F90;OLD UYGHUR STEM EXTENDER;Po;0;AL;;;;N;;;;;
10F91;OLD UYGHUR DELETION MARK;Mn;220;NSM;;;;N;;;;;
    
```

10.2 Linebreak data: LineBreak.txt

```

10F70..10F82;AL # Lo [19] OLD UYGHUR LETTER ALEPH..OLD UYGHUR LETTER LESH
10F82..10F89;CM # Mn [7] OLD UYGHUR COMBINING DOT RIGHT..
                        OLD UYGHUR COMBINING HAMSA RIGHT
10F8A..10F8A;AL # Po [6] OLD UYGHUR PUNCTUATION BAR..OLD UYGHUR SECTION MARK
10F90;AL # Po OLD UYGHUR STEM EXTENDER
10F91;CM # Mn OLD UYGHUR DELETION MARK

```

10.3 Property list: PropList.txt

```

10F91 ; Extender # Po OLD UYGHUR STEM EXTENDER

```

10.4 Shaping properties: ArabicShaping.txt

```

10F70; OLD UYGHUR ALEPH; D; No_Joining_Group
10F71; OLD UYGHUR INDEPENDENT ALEPH WITH VERTICAL TAIL; N; No_Joining_Group
10F72; OLD UYGHUR BETH; D; No_Joining_Group
10F73; OLD UYGHUR GIMEL; D; No_Joining_Group
10F74; OLD UYGHUR WAW; D; No_Joining_Group
10F75; OLD UYGHUR ZAYIN; D; No_Joining_Group
10F76; OLD UYGHUR HETH; D; No_Joining_Group
10F77; OLD UYGHUR YODH; D; No_Joining_Group
10F78; OLD UYGHUR KAPH; D; No_Joining_Group
10F79; OLD UYGHUR LAMEDH; D; No_Joining_Group
10F7A; OLD UYGHUR MEM; D; No_Joining_Group
10F7B; OLD UYGHUR NUN; D; No_Joining_Group
10F7C; OLD UYGHUR SAMEKH; D; No_Joining_Group
10F7D; OLD UYGHUR PE; D; No_Joining_Group
10F7E; OLD UYGHUR SADHE; D; No_Joining_Group
10F7F; OLD UYGHUR RESH; D; No_Joining_Group
10F80; OLD UYGHUR SHIN; D; No_Joining_Group
10F81; OLD UYGHUR TAW; D; No_Joining_Group
10F82; OLD UYGHUR LESH; D; No_Joining_Group

```

11 References

- Anderson, Deborah, et. al. 2018. “Recommendations to UTC #155 April-May 2018 on Script Proposals” (L2/18-168). <https://www.unicode.org/L2/L2018/18168-script-rec.pdf>
- China. 2018. “GB/T 36331-2018 ”Information technology – Uigur-Mongolian characters, presentation characters and use rules of controlling characters”.
<http://c.gb688.cn/bzgk/gb/showGb?type=online&hcno=DFE87CC79EA67F8BF8B37C9C41CF9348>
- Clark, Larry. 2010. “The Turkic script and the *Kutadgu Bilig*”. *Turcology in Mainz*, Turcologica, Band 82, Hendrik E Boeschoten and Julian Rentzsch (ed.), pp. 89–106. Wiesbaden: Harrassowitz Verlag.
- Clauson, Gerard. 1962. *Studies in Turkic and Mongolic Linguistics*. London: Royal Asiatic Society of Great Britain and Ireland.
- Coulmas, Florian. 1996. *The Blackwell Encyclopedia of Writing Systems*. Oxford: Blackwell Publishers.
- Erdal, Marcel. 1984. “The Turkish Yarkand Documents”. *Bulletin of the School of Oriental and African Studies, University of London*, vol. 47, no. 2, pp. 260–301.



































- Hamilton J. 1986. *Manuscripts ouïgours de IXe–Xe siècle de Touen-houang*, no. 19. Paris)
- Israpil, Dilara and Yüsüp, Israpil. 2014. “Two Old Uighur Account Documents from Toqquzsaray Ruins in Maralbeşi”. *Studies on the Inner Asian Languages*, vol. 29, pp. 137–156.
- Kara, György. 1996. “Aramaic Scripts for Altaic Languages”. *The World’s Writing Systems*, Peter T. Daniels and William Bright (ed.), pp. 536–558. New York and Oxford: Oxford University Press.
- Müller, Friedrich Wilhelm Karl. 1910. *Uigurica*, vol. II. Berlin: Verlag der Königlichen Akademie der Wissenschaften.
- Nadeliaev, V. M.; Nasilov, D. M.; Scherbak, A. M.; Tenishev, E. R. 1969. *Drevnetiurkskii slovar*. Akademiia nauk SSSR. Institut iazykoznaniiia. Leningrad: Izd-vo ”Nauka” Leningradskoe otd-nie.
- Olmez, Mehmet. 2016. “Compared transcription system for Old Uyghur Alphabet”. Lecture at École des Hautes Études en Sciences Sociales (EHESS), Paris, May 2016. http://www.academia.edu/24939281/Lectures_at_EHESS_1_Compared_transcription_system_for_Old_Uighur_Alphabet
- Osman, Omarjan. 2013. “Proposal to Encode the Uyghur Script in ISO/IEC 10646”. L2/13-071. <http://www.unicode.org/L2/L2013/13071-uyghur.pdf>
- Pandey, Anshuman. 2016. “Revised proposal to encode the Sogdian script in Unicode”. L2/16-371R. <http://www.unicode.org/L2/L2016/16371r2-sogdian.pdf>
- Radloff, W. 1910. *Ṭiṣastvustik: ein in Türkischer Sprache bearbeitetes Buddhistisches Sūtra*. I. Transcription and Übersetzung; II. Bemerkungen zu den Brähmīglossen des Ṭiṣastvustik-Manuscripts (Mus. As. Kr. VII) von Baron A. von Staël-Holstein. Bibliotheca Buddhica, XII. St.-Pétersbourg.
- Radlov, Vasilii Vasil'evich; Malov, S. Efimovich. 1913. *Suvarṇaprabhāsa*. (Sutra zolotogo bleska); tekst ujugurskogo redakcii. Bibliotheca Buddhica, XVII. Sanktpeterburg: Imper. Akad. Nauk.
- Shōgaito, Masahiro. 1988. “Passages from Abhidharma-nyāyānusāra-śāstra. Quoted in the Uighur Text Or. 8212-75B, British Library”. *Studies on the Inner Asian Languages*, vol. 3, pp. 159–207.
- Sims-Williams, Nicholas. 1981. “The Sogdian sound-system and the origins of the Uyghur script”. *Journal Asiatique*, pp. 347–360.
- von Gabain, Annemarie. 1950. *Altürkische Grammatik*. Mit Bibliographie, Lesestücken und Wörterverzeichnis, auch Neutürkisch. Mir vier Schrifttafeln und sieben Schriftproben. Porta linguarum orientarium. no. 23. 1. verbesserte Auflage. Leipzig: Otto Harrassowitz.
- West, Andrew. 2006. “Phags-pa Script : Old Uyghur Script”. <http://www.babelstone.co.uk/Phags-pa/Uighur.html>
- . 2011a. “Khitan Miscellanea 1: Oh, How the Gods Mock Us!”. <http://babelstone.blogspot.com/2011/10/khitan-miscellanea-1.html>
- . 2011b. “Phags-pa Uyghur Seals”. <http://babelstone.blogspot.com/2011/11/phags-pa-uyghur-seals.html>

12 Acknowledgments




















I express my gratitude to Dai Matsui (Graduate School of Letters, Osaka University) for sharing information on the Old Uyghur script, and for patiently providing answers to my numerous questions. I thank Liang Hai for providing extensive comments on previous versions of this proposal. The present document owes its shape to his diligent review. I also would like to thank Andrew West for providing reproductions and transcriptions of the Old Uyghur inscriptions on the walls of the Cloud Platform at Juyong Guan (fig. 34).

This project has been made possible in part by funding from the Adopt-A-Character program of the Unicode Consortium, and has been supervised by Deborah Anderson and Rick McGowan.








It was also made possible in part by a grant from the U.S. National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at UC Berkeley). Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment of the Humanities.

	10F7	10F8	10F9	10FA
0	 10F70	 10F80	 10F90	
1	 10F71	 10F81	 10F91	
2	 10F72	 10F82		
3	 10F73	 10F83		
4	 10F74	 10F84		
5	 10F75	 10F85		
6	 10F76	 10F86		
7	 10F77	 10F87		
8	 10F78	 10F88		
9	 10F79	 10F89		
A	 10F7A	 10F8A		
B	 10F7B	 10F8B		
C	 10F7C	 10F8C		
D	 10F7D	 10F8D		
E	 10F7E	 10F8E		
F	 10F7F	 10F8F		







Letters

10F70		OLD UYGHUR LETTER ALEPH
10F71		OLD UYGHUR LETTER INDEPENDENT ALEPH WITH VERTICAL TAIL
10F72		OLD UYGHUR LETTER BETH
10F73		OLD UYGHUR LETTER GIMEL
10F74		OLD UYGHUR LETTER WAW
10F75		OLD UYGHUR LETTER ZAYIN
10F76		OLD UYGHUR LETTER HETH
10F77		OLD UYGHUR LETTER YODH
10F78		OLD UYGHUR LETTER KAPH
10F79		OLD UYGHUR LETTER LAMEDH
10F7A		OLD UYGHUR LETTER MEM
10F7B		OLD UYGHUR LETTER NUN
10F7C		OLD UYGHUR LETTER SAMEKH
10F7D		OLD UYGHUR LETTER PE
10F7E		OLD UYGHUR LETTER SADHE
10F7F		OLD UYGHUR LETTER RESH
10F80		OLD UYGHUR LETTER SHIN
10F81		OLD UYGHUR LETTER TAW
10F82		OLD UYGHUR LETTER LESH • hooked r

Combining signs

10F83		OLD UYGHUR COMBINING DOT RIGHT
10F84		OLD UYGHUR COMBINING TWO DOTS RIGHT
10F85		OLD UYGHUR COMBINING THREE DOTS RIGHT
10F86		OLD UYGHUR COMBINING DOT LEFT
10F87		OLD UYGHUR COMBINING TWO DOTS LEFT
10F88		OLD UYGHUR COMBINING THREE DOTS LEFT
10F89		OLD UYGHUR COMBINING HAMZA LEFT

Punctuation

10F8A		OLD UYGHUR PUNCTUATION BAR
10F8B		OLD UYGHUR PUNCTUATION TWO BARS
10F8C		OLD UYGHUR PUNCTUATION TWO DOTS
10F8D		OLD UYGHUR PUNCTUATION FOUR DOTS
10F8E		OLD UYGHUR PUNCTUATION FIVE DOTS
10F8F		OLD UYGHUR SECTION MARK

Stem extender

10F90		OLD UYGHUR STEM EXTENDER
-------	---	--------------------------

Editorial mark

10F91		OLD UYGHUR DELETION MARK
-------	---	--------------------------

ウイグル字母表

母音		1	2	3											
字の名		āleph	yod	waw											
音價		a, ä	ī, i	o, u, ö, ü											
語頭形		𐰀 𐰁 (a) (ä)	𐰂	𐰃 𐰄 (o, u) (ö, ü)											
語中形		𐰅	𐰆	𐰇 𐰈 = ö, ü (第一級)											
語尾形		𐰉 𐰊 𐰋 𐰌 𐰍		𐰎 𐰏	𐰐										
子音	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
字の名	ḥēth	kāph	yod	reš	lāmedh	taw	dāleth	šādhe	šin		nūn	pē	bēth	mēm	
音價	ḥ, q, ḫ	k, g	y	r	l	t, d	d, t	č	s, š	ś	z	n	b, p	w	m
語頭形	𐰑 𐰒	𐰓 𐰔	𐰕 𐰖	𐰗 𐰘	𐰙 𐰚	𐰛 𐰜	𐰝 𐰞	𐰟 𐰠	𐰡 𐰢	𐰣	𐰤 𐰥	𐰦 𐰧	𐰨 𐰩	𐰪 𐰫	𐰬 𐰭
語中形	𐰑 𐰒	𐰓 𐰔	𐰕 𐰖	𐰗 𐰘	𐰙 𐰚	𐰛 𐰜	𐰝 𐰞	𐰟 𐰠	𐰡 𐰢	𐰣	𐰤 𐰥	𐰦 𐰧	𐰨 𐰩	𐰪 𐰫	𐰬 𐰭
語尾形	𐰑 𐰒	𐰓 𐰔	𐰕 𐰖	𐰗 𐰘	𐰙 𐰚	𐰛 𐰜	𐰝 𐰞	𐰟 𐰠	𐰡 𐰢	𐰣	𐰤 𐰥	𐰦 𐰧	𐰨 𐰩	𐰪 𐰫	𐰬 𐰭
符號															
									句讀點						
									らめくさ	/	//	///			
									𐰮	=	◆	◆◆			

Figure 1: Table of Old Uyghur characters used in the Uyghur inscription in the multi-script Yuan dynasty inscriptions at Juyong Guan 居庸關 pass at the Great Wall northwest of Beijing (from Chū-Yung-Kuan 居庸關, “The Buddhist Arch of the Fourteenth Century A.D. at the Pass of the Great Wall Northwest of Peking”, vol. 1, p. 165; reproduced from West 2006). See photograph containing an excerpt of the inscription in fig. 34.

— XV —

	Буквы алфавита ДТС	Орхоно-енисейские знаки	Арабские знаки	Уйгурские знаки
1	a	𐰇 𐰈	آ آ	𐰇 𐰈 𐰉
2	ā	—	آ	—
3	ä	𐰇 𐰈	آ	𐰇 𐰈 𐰉
4	ā	—	—	—
5	b	𐰇 𐰈 𐰉	ب	𐰇 𐰈 𐰉
6	č	𐰇 𐰈	ج ج	𐰇 𐰈 𐰉
7	d	𐰇 𐰈 𐰉	د (ض)	𐰇 𐰈 𐰉
8	ḍ	—	—	𐰇 𐰈 𐰉
9	ḍ	—	ذ	—
10	e	𐰇 𐰈	ايه	𐰇 𐰈 𐰉
11	e	𐰇 𐰈 𐰉	آ	𐰇 𐰈 𐰉
12	ē	—	اي آ	—
13	f	—	ف	𐰇 𐰈 𐰉
14	g	𐰇	ك	𐰇 𐰈 𐰉
15	γ	𐰇 𐰈 𐰉	غ	𐰇 𐰈 𐰉 𐰊 𐰋
16	h	—	ه	—
17	ḥ	—	ح	—
18	i	𐰇 𐰈	ايه	𐰇 𐰈 𐰉
19	ī	—	ايه	𐰇 𐰈 𐰉
20	ī	𐰇 𐰈	ايه	𐰇 𐰈 𐰉
21	ī	—	ايه	𐰇 𐰈 𐰉
22	j	𐰇 𐰈	ي	𐰇 𐰈 𐰉
23	ī	𐰇 𐰈	—	—
24	k	𐰇 𐰈 𐰉 𐰊	ك	𐰇 𐰈 𐰉
25	l	𐰇 𐰈	ل	𐰇 𐰈 𐰉
26	m	𐰇 𐰈	م	𐰇 𐰈 𐰉

Figure 2: Representation of Old Turkic sounds in the Orkhon, Arabic, and Old Uyghur scripts (from Nadeliaev, et al. 1969: xv). Continued in fig. 3.

— XVI —

	Буквы алфавита ДТС	Орхоно-енисейские знаки	Арабские знаки	Уйгурские знаки
27	n	𐰇 𐰈 𐰉	ن	𐰇 𐰈 𐰉
28	ŋ	𐰇 𐰈	نك	𐰇 𐰈
29	o	𐰇	او - و	𐰇 𐰈
30	ō	—	—	𐰇 𐰈
31	ö	𐰇 𐰈	او - و	𐰇 𐰈
32	ō	—	—	𐰇 𐰈
33	p	1	پ ب	𐰇 𐰈
34	q	𐰇 𐰈 𐰉	ق	𐰇 𐰈 𐰉
35	r	𐰇 𐰈	ر	𐰇 𐰈
36	s	𐰇 𐰈	س ص	𐰇 𐰈
37	š	𐰇 𐰈	—	𐰇
38	š	𐰇 𐰈 𐰉	ش	𐰇 𐰈
39	š	𐰇 𐰈	—	—
40	t	𐰇 𐰈 𐰉 𐰊	ة ط ت	𐰇 𐰈 𐰉
41	t	—	—	𐰇 𐰈
42	ʈ	—	ت	—
43	u	𐰇	او - و	𐰇 𐰈
44	ū	—	—	𐰇 𐰈
45	ü	𐰇 𐰈	او - و	𐰇 𐰈
46	ū	—	—	𐰇 𐰈
47	v	—	ف و ف	𐰇 𐰈
48	w	см. 47	см. 47	см. 47
49	ʒ	—	خ	𐰇 𐰈 𐰉 𐰊
50	z	𐰇 𐰈 𐰉	ض ز ظ	𐰇
51	ž	—	—	𐰇
52	ž	—	ژ	𐰇
53	ž	—	—	𐰇
54	ž	—	ح	𐰇 𐰈
55	ʻ	—	ء	—
56	ʻ	—	ع	—

Figure 3: Representation of Old Turkic sounds in the Orkhon, Arabic, and Old Uyghur scripts (from Nadeliaev, et al. 1969: xvi). Continued from fig. 2.

TABLE 49.2: *Uyghur Script*^a

Name ^b	Uyghur	Initial	Medial	Final	Separate	Ligatures	Uyghur
'aleph	e/vowel initial	◀	◀	◀	◀		ka/e
	a/e	◀	◀	◀	◀	◀	pa/e
beth	w/v	◀	◀	◀	◀	◀	
gimel	γ	◀	◀	◀	◀		
waw	o/u	◀	◀	◀	◀		
waw+yodh	ö/ü	◀	◀	◀	◀	◀	ko/u/ö/ü
	o/u/ö/ü ^c	◀	◀	◀	◀	◀	po/uö/ü
zain	z	◀	◀	◀	◀		
marked z	ž	◀	◀	◀	◀		
heth	x	◀	◀	◀	◀		
2-dotted	q	◀	◀	◀	◀		
yodh	y	◀	◀	◀	◀	◀	ki/ï
		◀	◀	◀	◀	◀	pi/ï
kaph	k/g	◀	◀	◀	◀		
lamedh	d/ð	◀	◀	◀	◀		
mem	m	◀	◀	◀	◀	◀	ml
nun	n	◀	◀	◀	◀		
pe	b/p	◀	◀	◀	◀		
tsadi	č	◀	◀	◀	◀		
resh	r	◀	◀	◀	◀		
shin	s	◀	◀	◀	◀		
marked s	š	◀	◀	◀	◀		
tau	t	◀	◀	◀	◀		
hooked r	l	◀	◀	◀	◀		

a. Diacritics are often omitted. Some Uyghur alphabets have shin for samekh before pe; marked z, final m, and final q are added after hooked resh.
 b. Hebrew name for the ancestral Aramaic letter.
 c. In syllables other than the first.

Figure 4: Table showing letters of the Old Uyghur script (from Kara 1996: 540). See table of Mongolian letters from the same source in fig. 43.

Uigurische, sogdische und manichäische Schrift

Uigurisch	Sogdisch	Manichäisch	Uigurisch	Sogdisch	Manichäisch
a			χ = γ	= γ	
-a		= -a	h = g	= g	
-a			k = g	= g	
ä		= -a	l		
-ä	= -a	= -a	m		
-ä	= -a	= -a	-m		
i, i; i, i			n		
-i, -i		= -i	-n		
i, i			ŋ		
o, u			p = b	= b	
-o, -u		= -o	q		
-o, -u			-q		
ö, ü			-r		
-ö, -ü			-r		
-ö, -ü = -o	= -o	= -o	s		
b			-s		
-b			s = s		
c, ç			t		
d, ð			-t		
ð			v = -o	= -o	= -o
f = w	= w		w		
γ			y		
-γ			z		
g			z		
-g			z		
			Zeilenfüller		

Figure 5: Comparison of Old Uyghur, Sogdian, and Manichaean letters (from von Gabain 1950: 17). For clearer examples of Old Uyghur letterforms referenced by von Gabain see the three Old Uyghur manuscripts, two in the formal script and the third in the cursive script, illustrated and transcribed in her work, reproduced here in fig. 37–41.

Uighur writing

Transliteration	10th C.	10th C.	1072	13th – 14th C.
1 '	𐰀	𐰀	𐰀	𐰀
2 β	𐰁	𐰁	𐰁	𐰁
3 γ	𐰂	𐰂	𐰂	𐰂
4 w	𐰃	𐰃	𐰃	𐰃
5 z	𐰄	𐰄	𐰄	𐰄
6 x	𐰅	𐰅	𐰅	𐰅
7 y	𐰆	𐰆	𐰆	𐰆
8 k	𐰇	𐰇	𐰇	𐰇
9 d(δ)	𐰈	𐰈	𐰈	𐰈
10 m	𐰉	𐰉	𐰉	𐰉
11 n	𐰊	𐰊	𐰊	𐰊
12 s	𐰋	𐰋	𐰋	𐰋
13 p	𐰌	𐰌	𐰌	𐰌
14 č	𐰍	𐰍	𐰍	𐰍
15 r	𐰎	𐰎	𐰎	𐰎
16 š	𐰏	𐰏	𐰏	𐰏
17 t	𐰐	𐰐	𐰐	𐰐
18 l	𐰑	𐰑	𐰑	𐰑
19 ž	𐰒	𐰒	𐰒	𐰒
20 -m	𐰓	𐰓	𐰓	𐰓
21 ġ	𐰔	𐰔	𐰔	𐰔

Table 2 Various forms of the Old Uighur alphabet from texts dating between the fourteenth and the tenth centuries BCE
Source: adapted from Zieme 1991

Figure 6: Comparison of Old Uyghur letterforms (from Coulmas 1996: 526).

Compared transcription system for Old Uyghur Alphabet

	Berliner Transkription system	Turkey	transcription at <i>Uigurisches Wörterbuch</i>	transliteration at <i>Uigurisches Wörterbuch</i>
𐰀	a, ʌ	a, ʌ	a	'' / 'ʼ
𐰁	b	b	b	P
𐰂	č	ç	č	Č
𐰃	d, ɖ	d, ɖ	d, ɖ	D, T
𐰄	ä, 'ä	e, 'e	ä	'
𐰅	[e] i	ê / i	e	Y / 'Y
𐰆	g	g	g	K
𐰇	γ / ʒ	g / ğ	g	Q, Q̇, Q̈
𐰈	h / χ, x, ẖ	h / ħ, ḥ	h	H / X
𐰉	ī	ı	ı	Y, Y
𐰊	i	i	i	Y, 'Y
𐰋	ž, ʒ	j	ž, ʒ	Ž, ʒ, Z
𐰌	k	k	k	K
𐰍	[k] q, q̇, q̈	k / k̆	k	K / Q, Q̇, Q̈
𐰎	l	l	l	L
𐰏	m	m	m	M
𐰐	n, ñ	n, ñ	n	N, Ñ
𐰑	ng, ñ, ŋ	ñg, ng, ñ	ŋ	NK
𐰒	o	o	o	W / 'W
𐰓	ö, ʊ	ö, ʊ	ö	W / WY / 'WY
𐰔	p	p	p	P
𐰕	r	r	r	R
𐰖	s, ʒ	s, ʒ	s, ʒ	S, Z
𐰗	š	ş	š	Ş, Ş
𐰘	t, ɖ	t, ɖ	t, ɖ	T, D
𐰙	u	u	u	W / 'W
𐰚	ü, ʉ	ü, ʉ	ü	W / WY / 'WY
𐰛	v	v	v	V
𐰜	y	y	y	Y
𐰝	z, ʒ	z, ʒ	z, ʒ	Z, S

Figure 7: Comparison of transliteration schemes for Old Uyghur (from Olmez 2016).

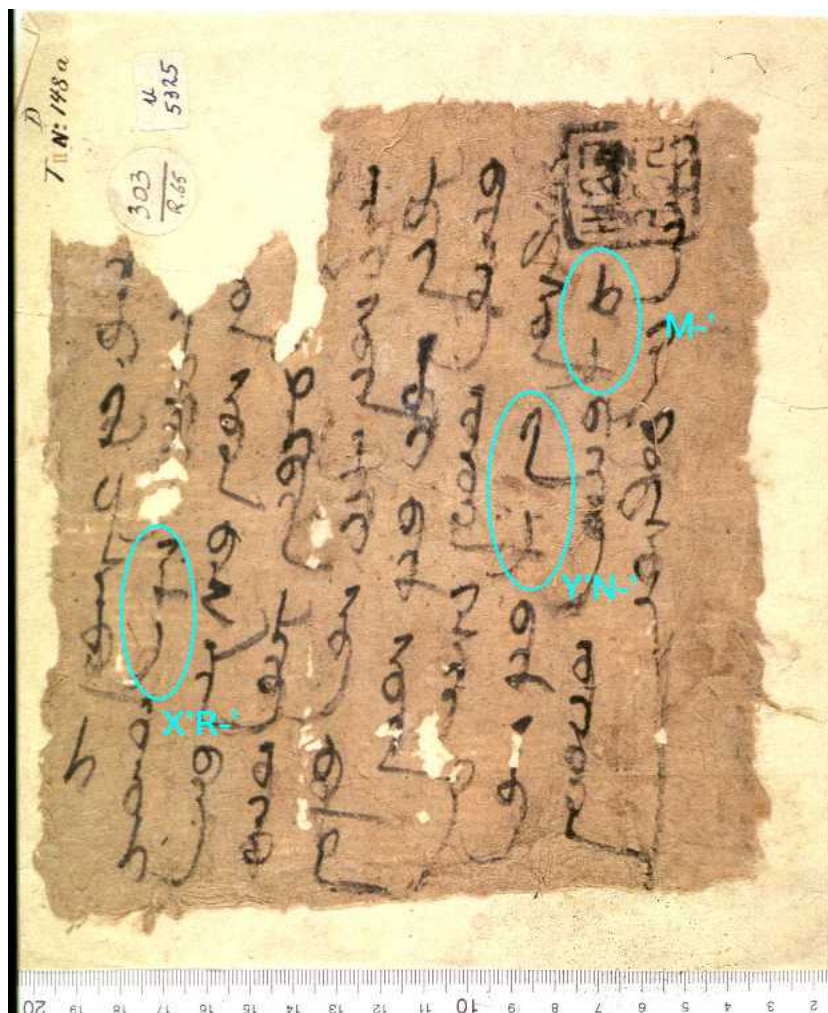


Figure 8: Usage of \curvearrowright INDEPENDENT ALEPH WITH VERTICAL TAIL after *mem* and *resh* in U 5325. Annotations produced by Dai Matsui, November 2018.

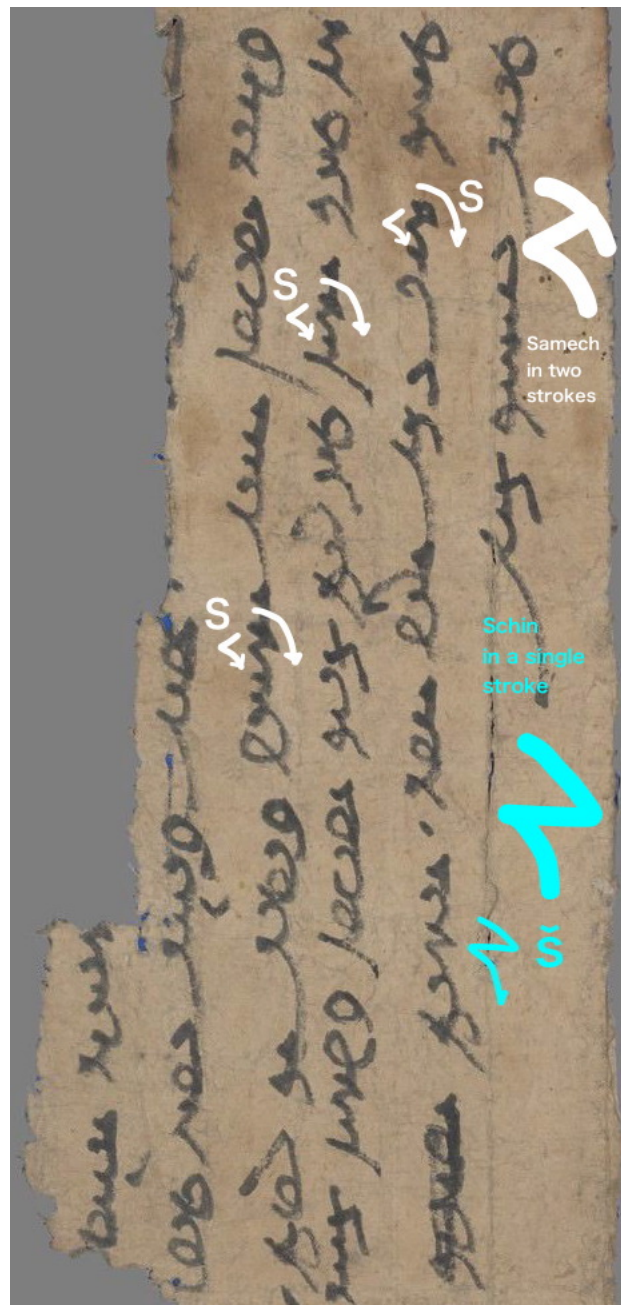


Figure 9: Excerpt from Pelliot Uighur 5 (9th–10th c.) showing the distinction between **Samekh** and **Shin** (from Hamilton 1986). Annotations produced by Dai Matsui, August 2018.

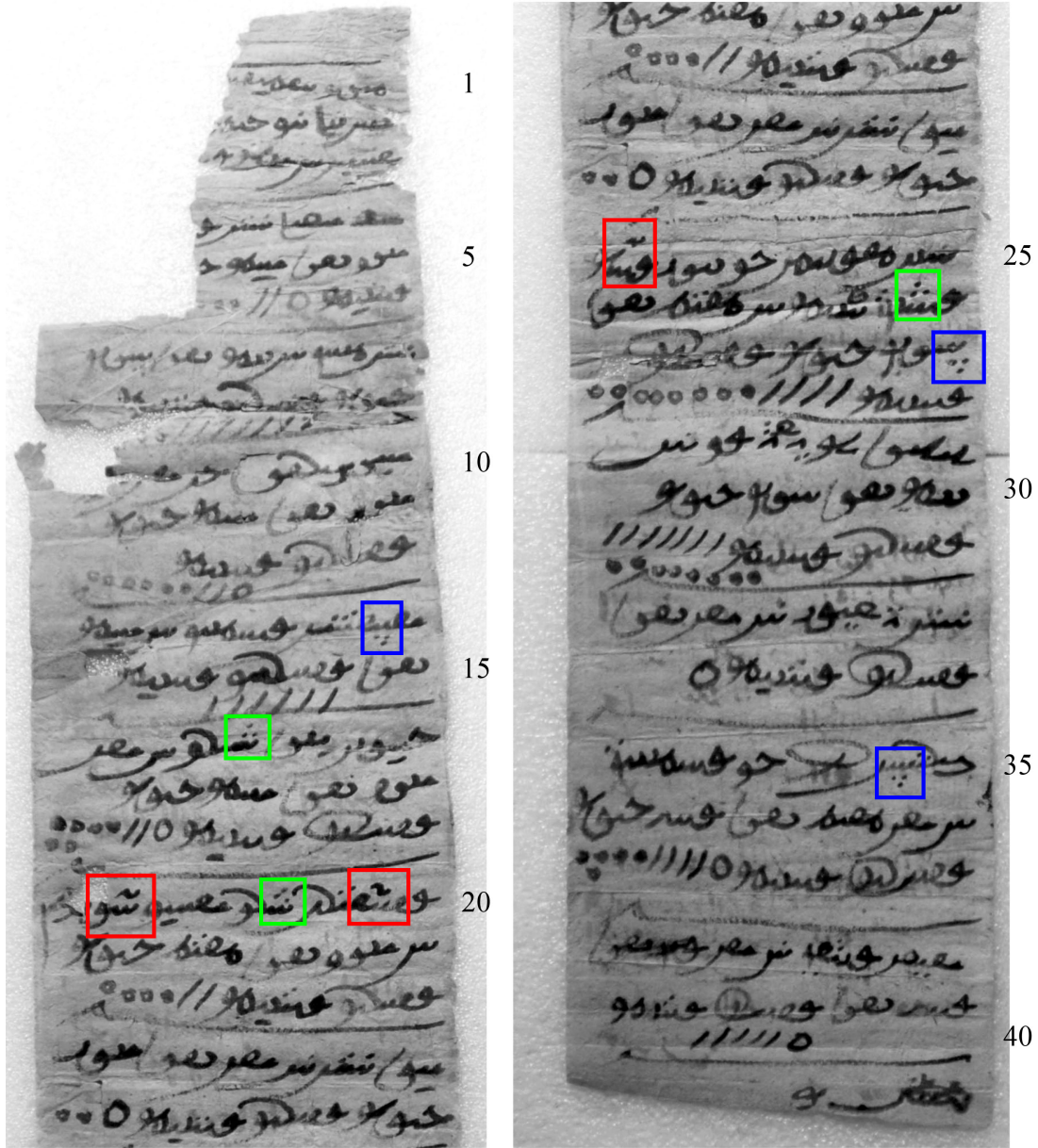


Figure 10: Usage of ۞ (blue), ۞ (green), and ۞ (red) for transcribing Arabic in a Old Uyghur administrative document (from Israpil 2014: plate I).

Handwritten Old Uyghur script with two red boxes highlighting specific marks. The first box highlights a mark resembling a circle with a cross inside (○✚), and the second box highlights a similar mark. The text is written in a cursive style with some intralinear annotations in Han characters.

大英図書館所蔵 Or. 8212-75B 頁76B-77A (本文テキスト219-239)

Figure 11: Usage of the ○✚ deletion mark for indicating error correction in Or. 8212/75, an Old Uyghur manuscript containing passages of the of the Buddhist text *Abhidharma-nyāyānūsāra-sāstra* (from Shōgaito 1988: 207). Note the intralinear text in Han characters.

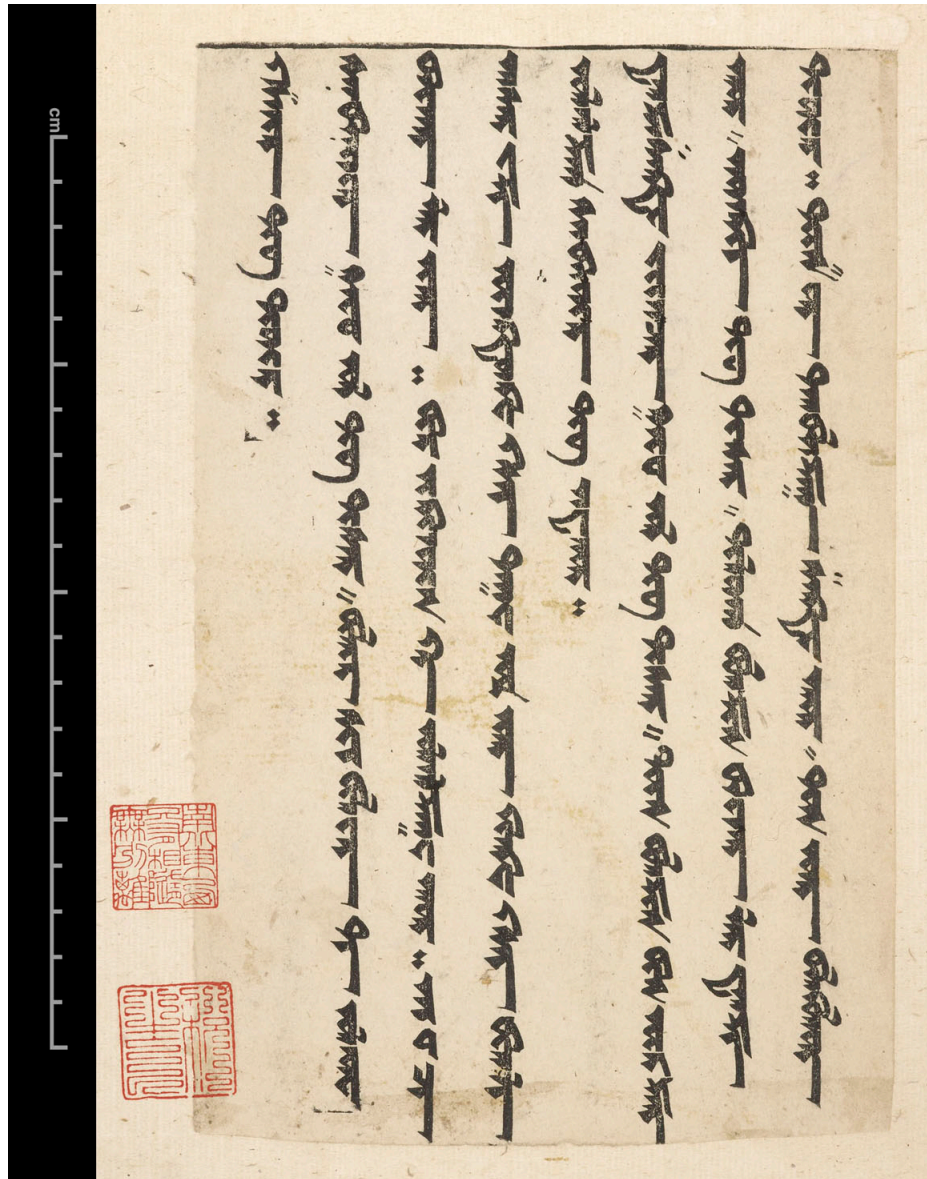


Figure 12: Princeton East Asian Library, PEALD 6a, recto. Block print.

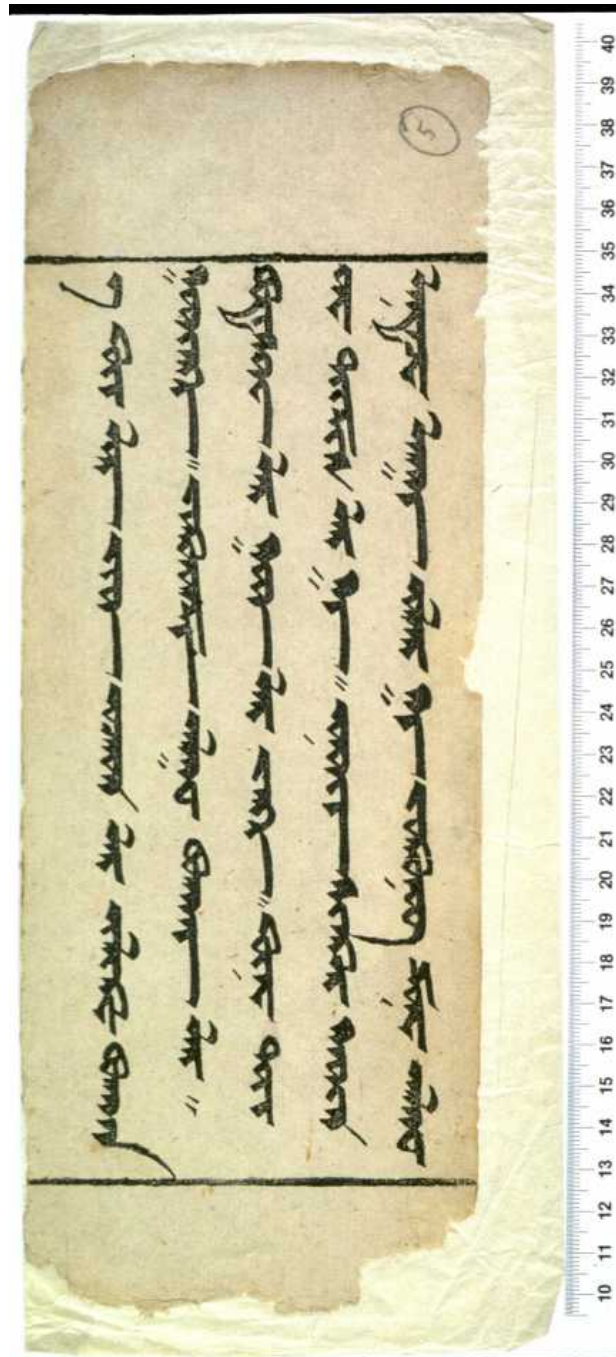


Figure 13: BBAW, U 0387, recto. Block print.



Figure 14: BBAW, U 4960, folio 1, recto. Block print. Seal in Han characters.

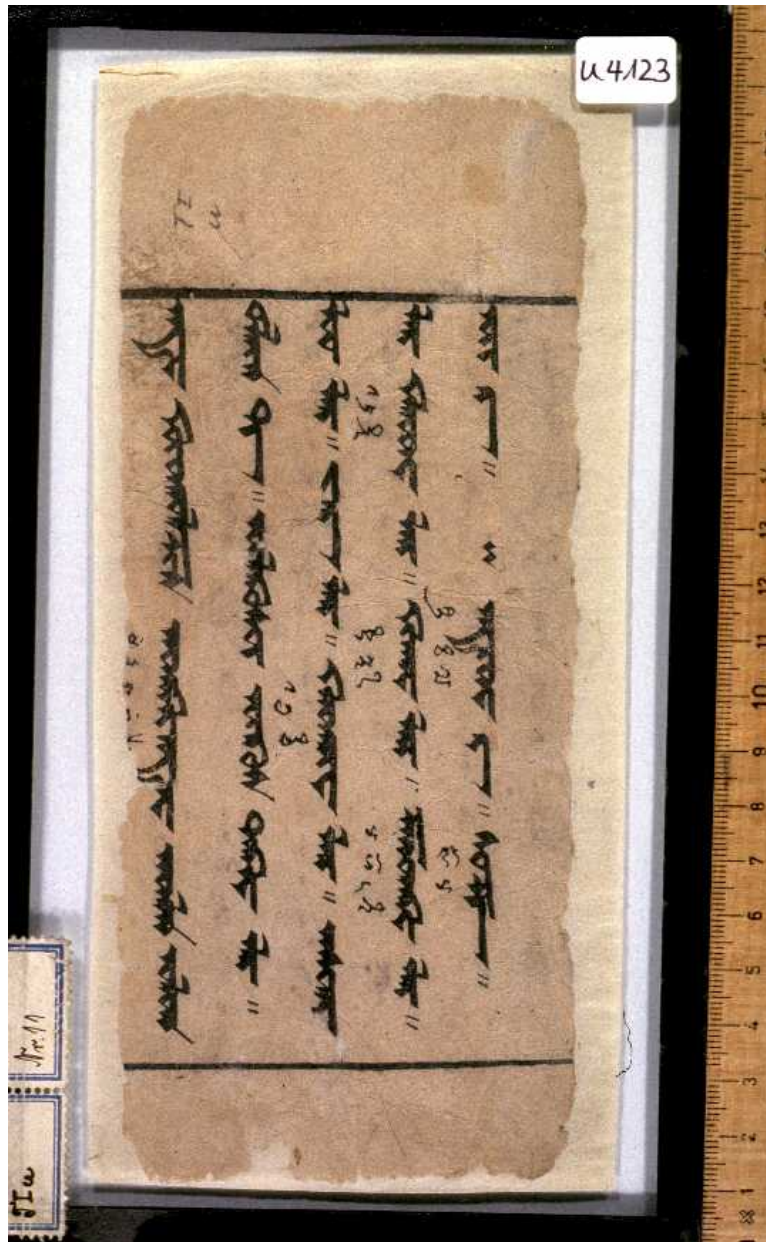


Figure 15: BBAW, U 4123. Block print.



Figure 16: BBAW, U 4124. Block print.

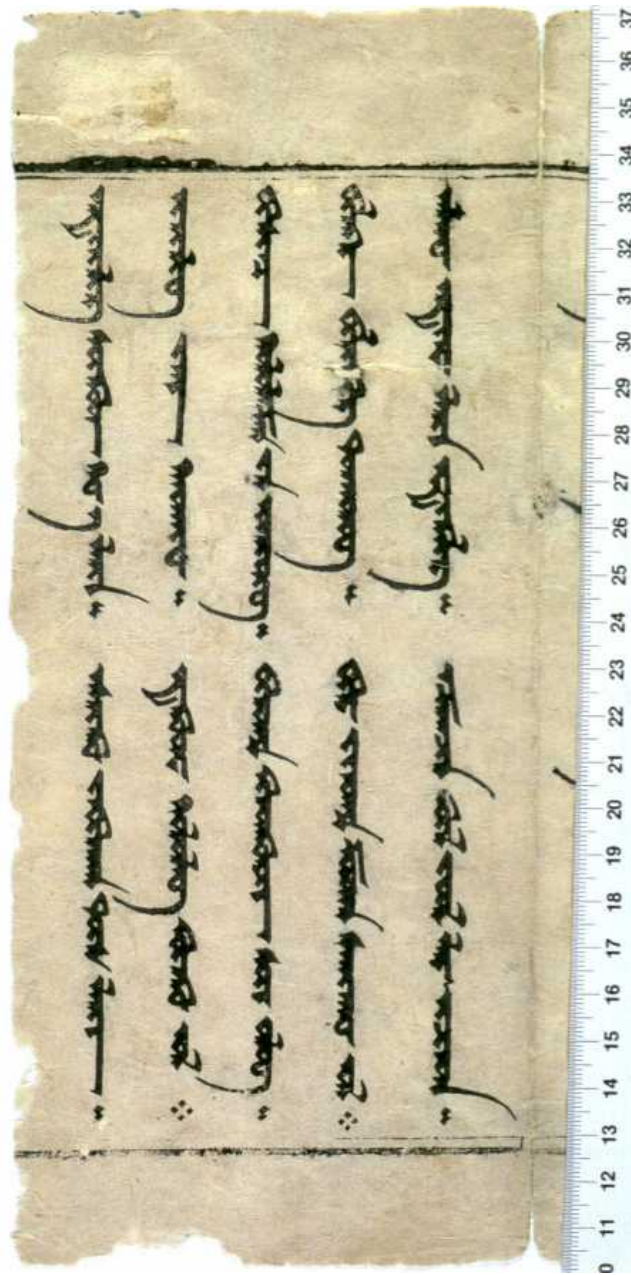


Figure 17: BBAW, U 0343, folio 1, recto. Block print.

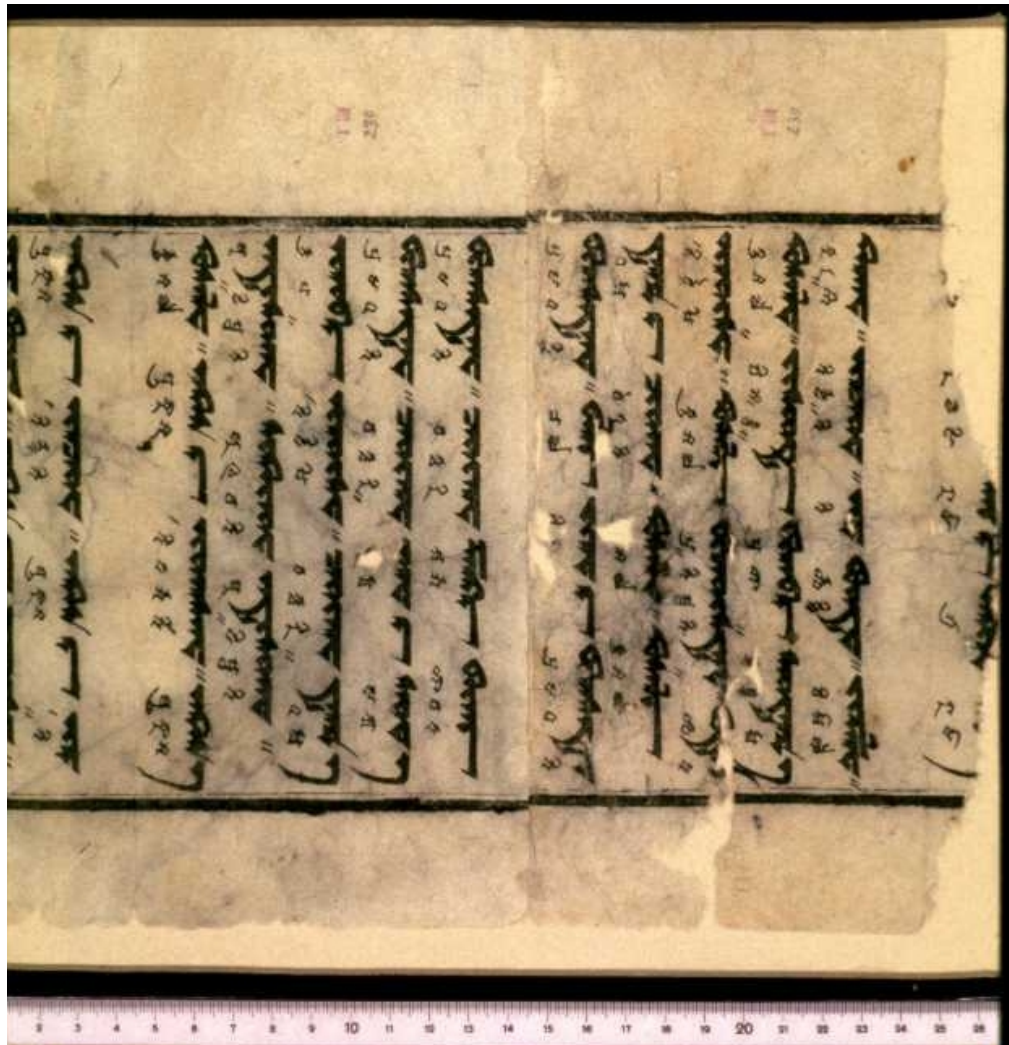


Figure 18: BBAW, Mainz 0801, middle portion. Block print. Annotations in Central Asian Brahmi.

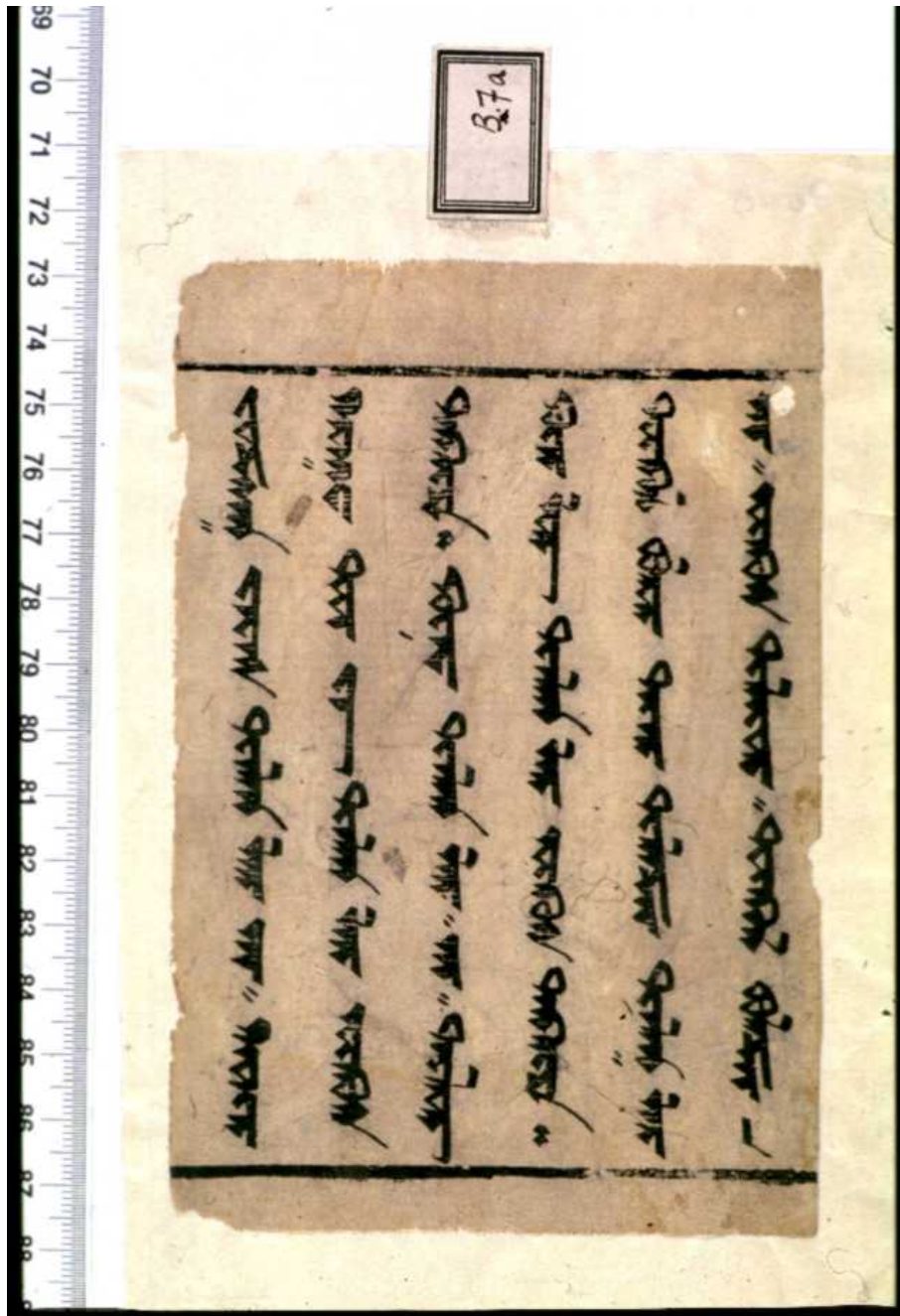


Figure 19: BBAW, U 7008. Block print.

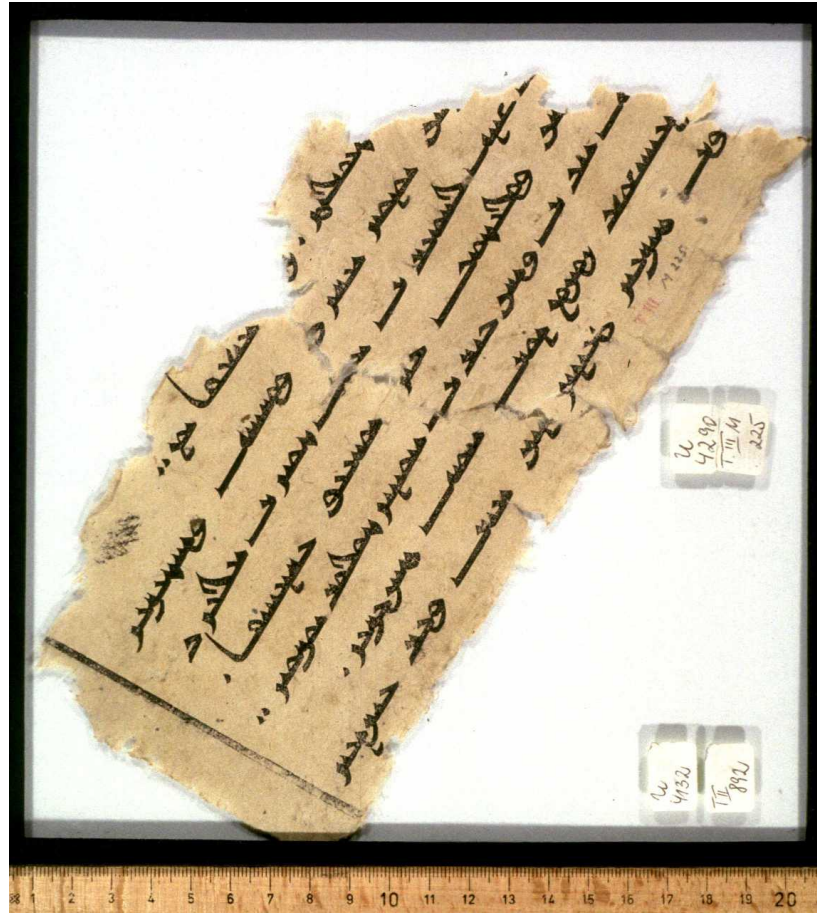


Figure 20: BBAW, U 4132 + 4290, folio 1. Block print.

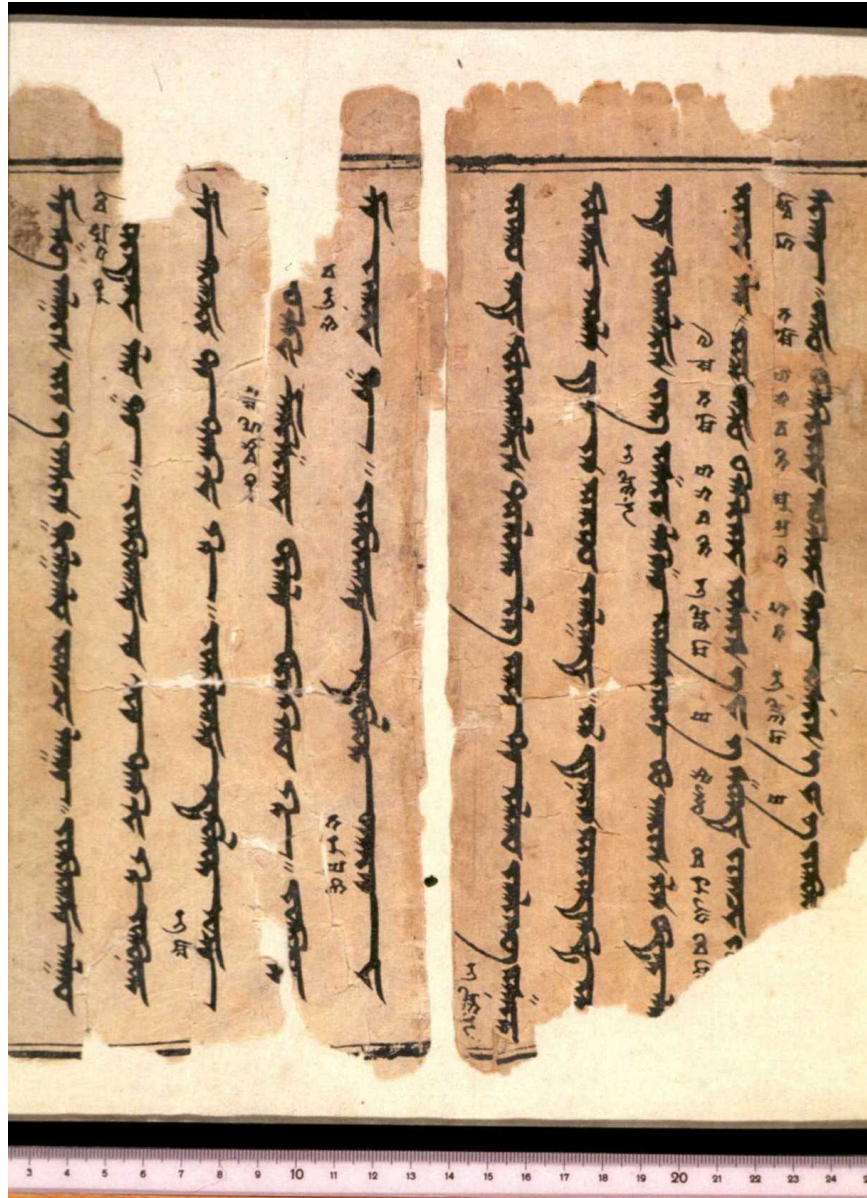


Figure 21: BBAW, Mainz 0764, middle. Formal script. Annotations in Central Asian Brahmi.



Figure 22: BBAW, U 3832, folio 1. Formal script.

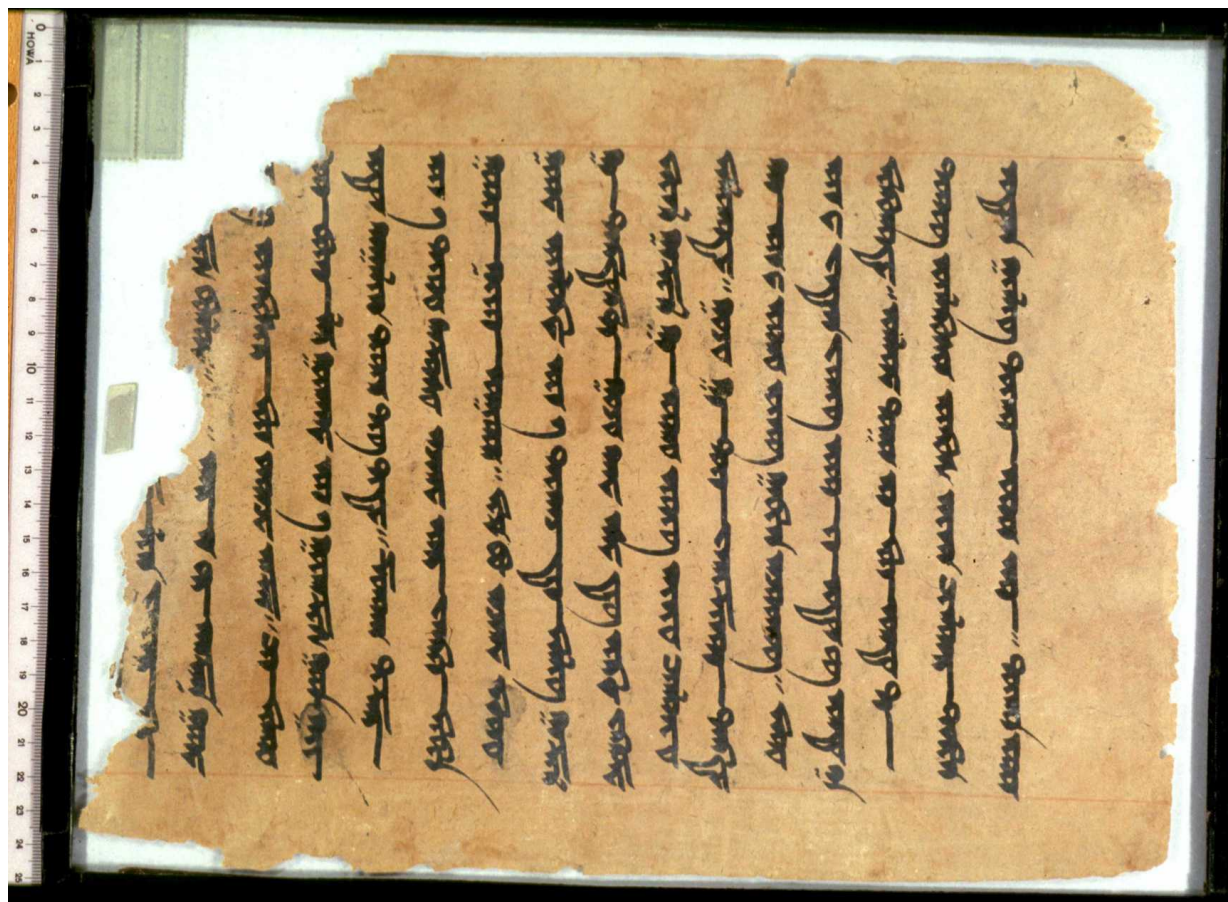


Figure 23: BBAW, Mainz 0841, folio 2.



Figure 24: BBAW, U 0924, folio 2.



Figure 25: PEALD 6r, recto.



Figure 26: BBAW, U 0040, recto. Note the inventory of Old Uyghur letters at the bottom of the folio. The letters **ا** *aleph* through **ت** *taw* are legible. The four letters following *taw* are unclear.



Figure 27: BBAW, U 0320, folio 1.

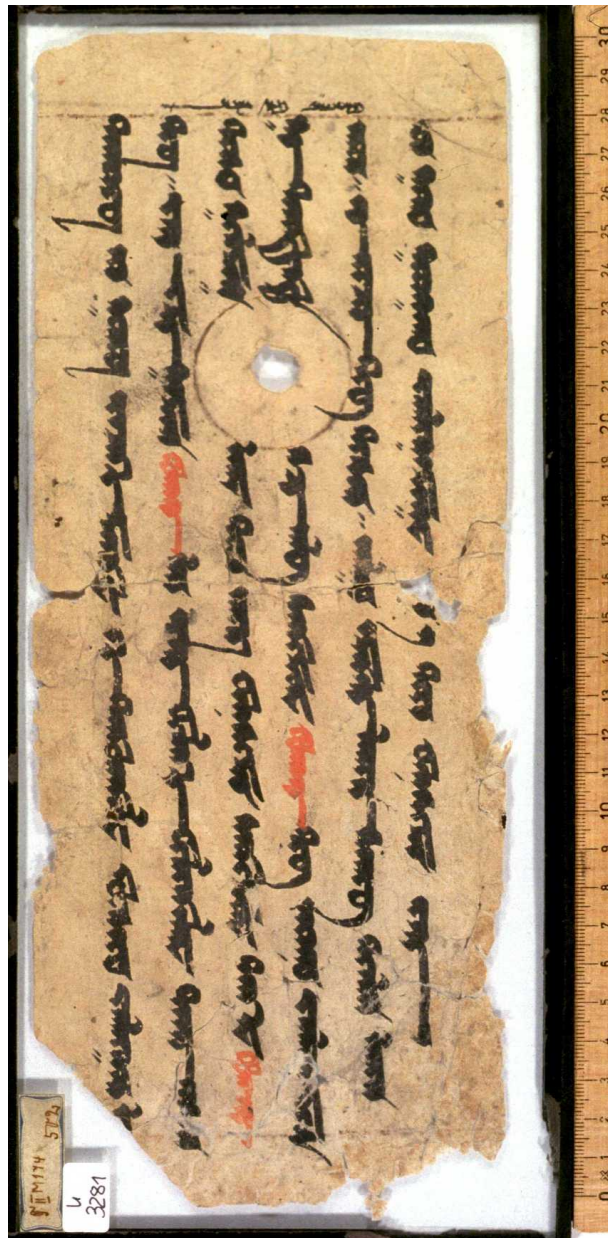


Figure 28: BBAW, U 3281, folio 1.

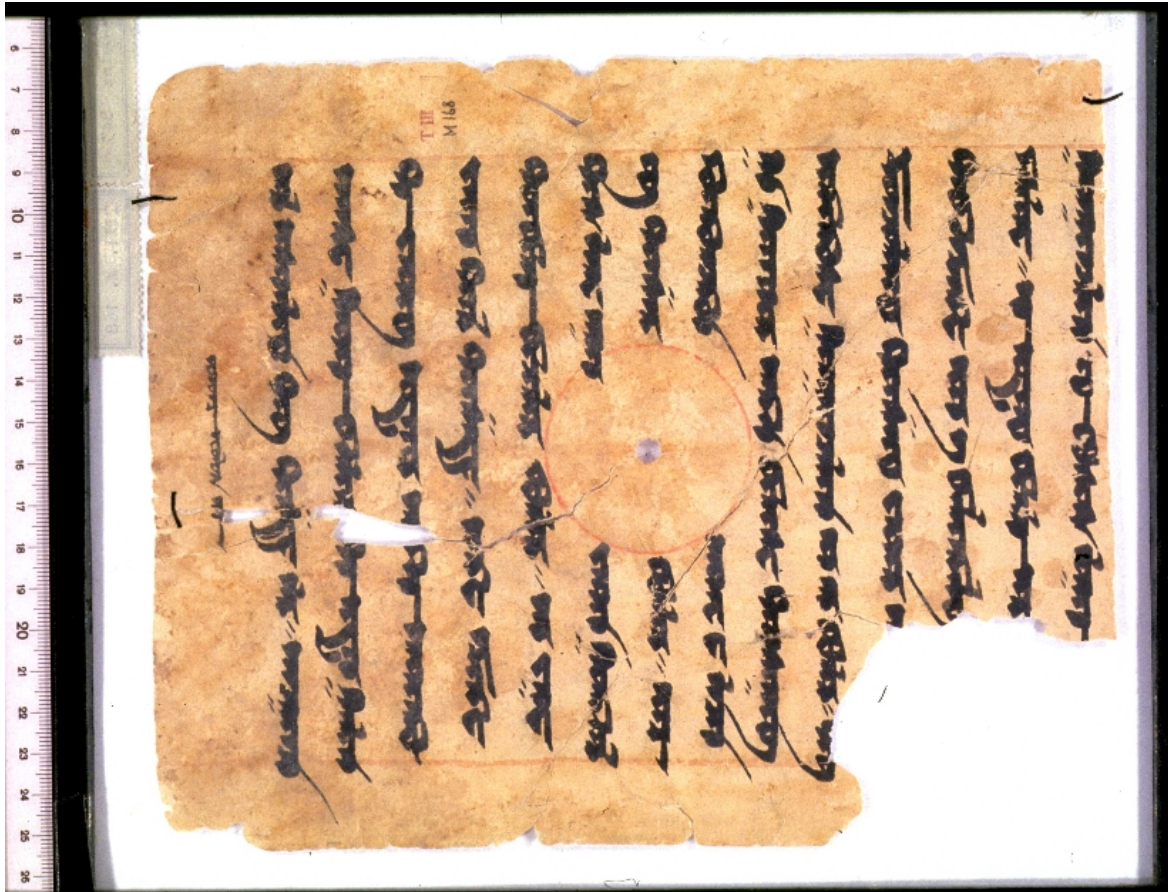


Figure 29: BBAW, Mainz 0843, folio 2.



Figure 30: BBAW, U 7123, recto.



Figure 31: BL / BBAW, Ch 5555, recto. *Ekottaragamasutra* / 增一阿含經 *Zeng yi e han jing*.

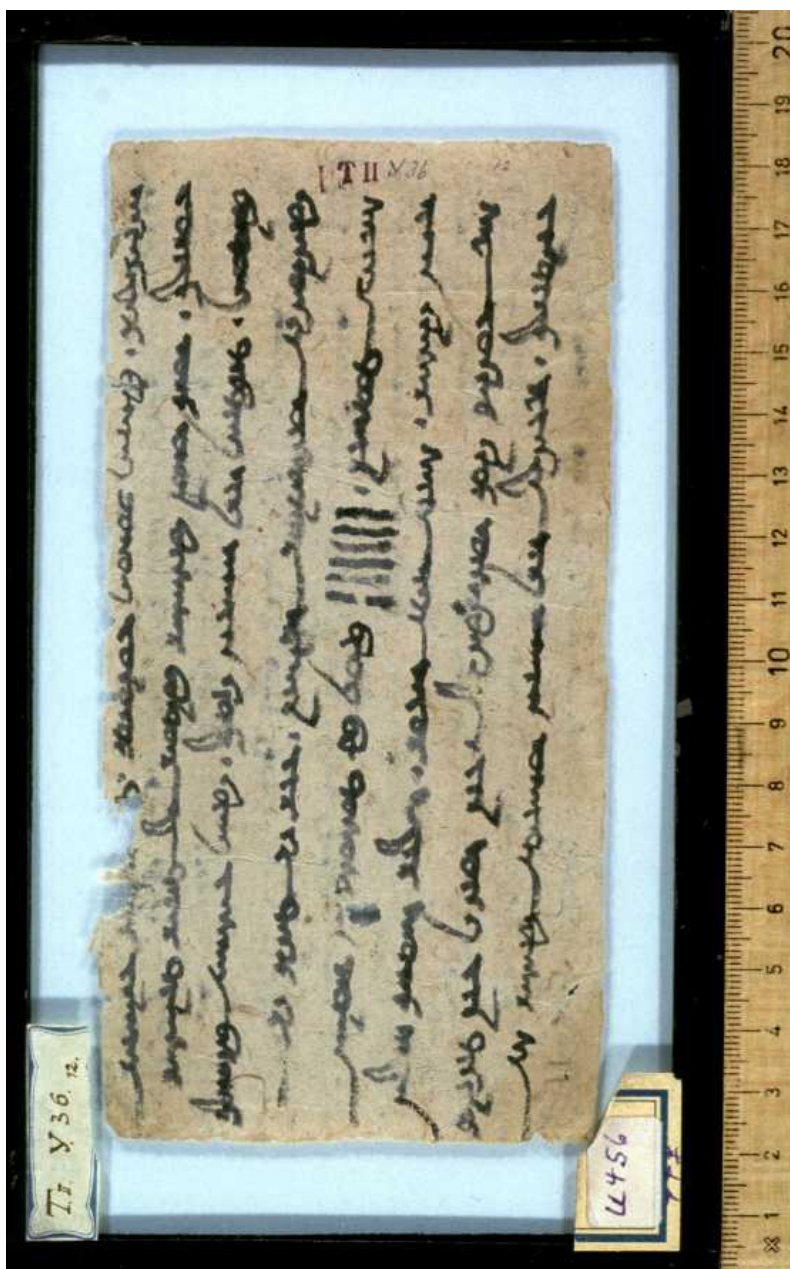


Figure 32: BBAW, U 0456, folio 1.

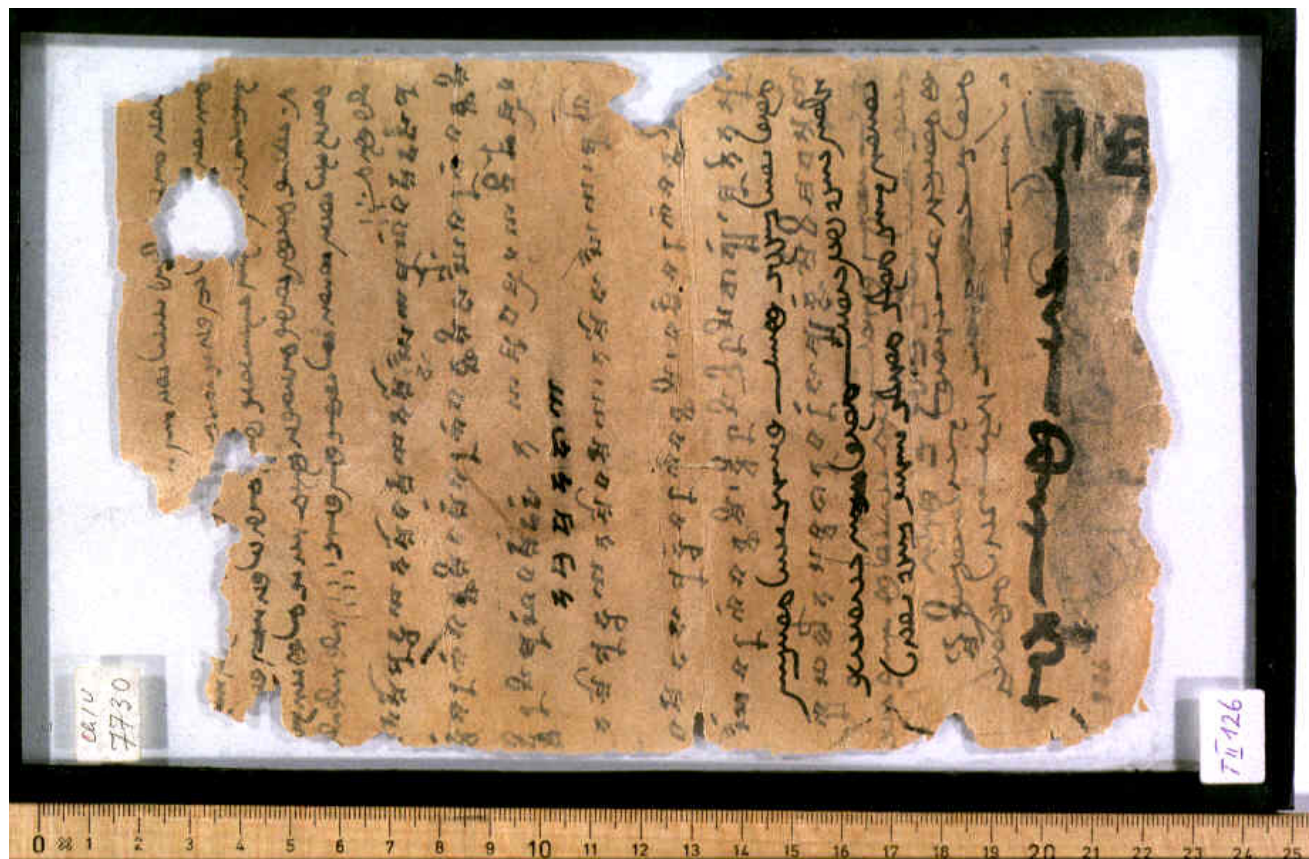


Figure 33: BBAW, Ch/U 7730, verso.



Figure 34: Detail of the Old Uyghur text of the multi-script Yuan dynasty Buddhist inscriptions on the west wall of the Cloud Platform at Juyong Guan 居庸關 pass at the Great Wall northwest of Beijing. Photograph by Andrew West, 2011.

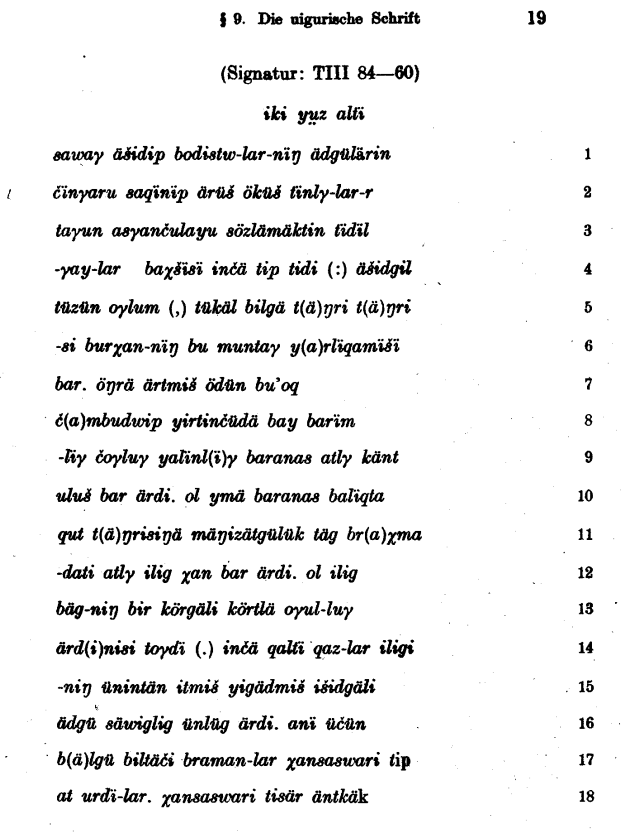
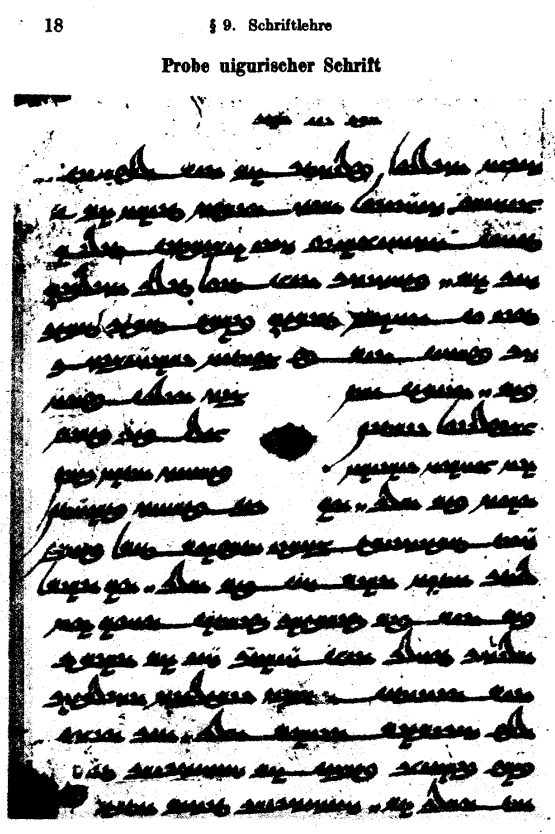


Figure 37: Transcription of an Old Uyghur manuscript (from von Gabain 1950: 18–19). Continued in fig. 38.

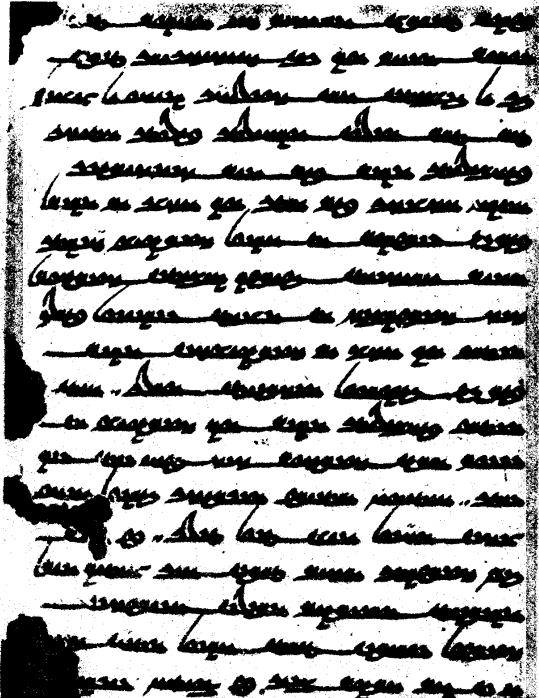
<p>20 § 9. Schriftlehre</p> 	<p>§ 9. Die uigurische Schrift 21</p> <p>bolur (.) türkâ awirsâr qaz ânlag tip 19</p> <p>yorug önâr (.) ol ymâ xansawari tigin 20</p> <p>-ki-â kiçmâtin ara suwdaqî linxua çâçâk 21</p> <p>tâg t(â)rk ödân ulıyadî bādâti (.) atasi 22</p> <p>br(a)ımadati ilig bâg-niıy şyawşagriw (lies swayşagriw?) 23</p> <p>atly aşçisi bar ârti (.) ol aşci âr ilig 24</p> <p>bagkâ yigulak ât atip söglünçü qılı (.) 25</p> <p>anııy arasinta koıul-lânmatin söışüıg 26</p> <p>şis söıgülmis ât icintü yilinip bardi (.) 27</p> <p>ötrü ol aşci âr söıglünçüsin ilig 28</p> <p>bagkâ kâlürüp üskintü urdı. anta 29</p> <p>ötrü br(a)ımadati ilig ol söıglünçü ât 30</p> <p>yııyür ârkân söışüıg şis boyzinta yil 31</p> <p>-inti . anta' oıq ârtiıü öwkâsi kâlıp ayru 32</p> <p>-çisin oıqıp inçâ tip tidi. bu âllig 33</p> <p>kim söıgülli ârsâr t(â)rkın anı çantal-niıy 34</p> <p>iliginü uruıylar (.) ikidin âniıyasın 35</p> <p>söküp yırâkin tarta atip otqa söıü 36</p> <p>-lân-lâr (.) nâıg(â)lâıg mini bu muntay yawlaııı 37</p>
--	--

Figure 38: Transcription of an Old Uyghur manuscript in a grammar of Old Turkic (from von Gabain 1950: 20–21). Continued from fig. 37.

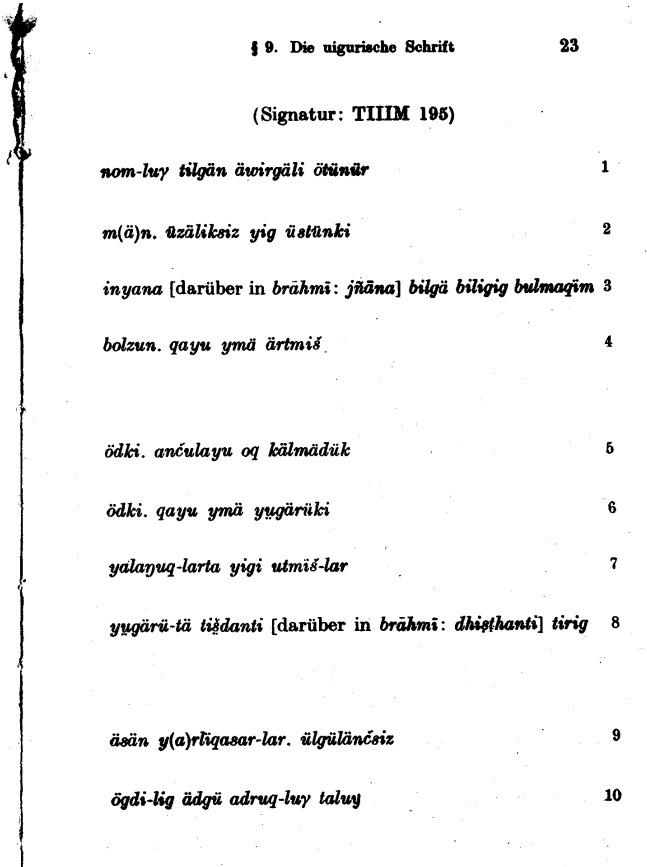
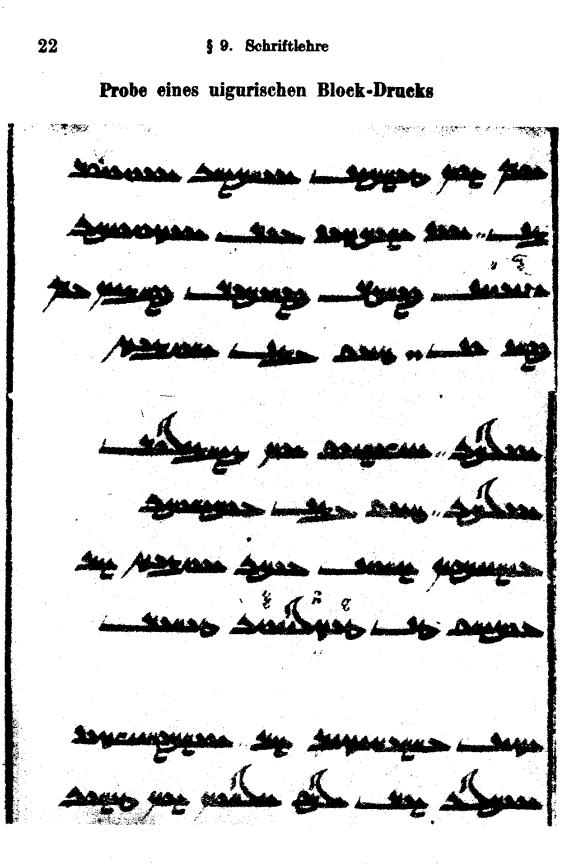
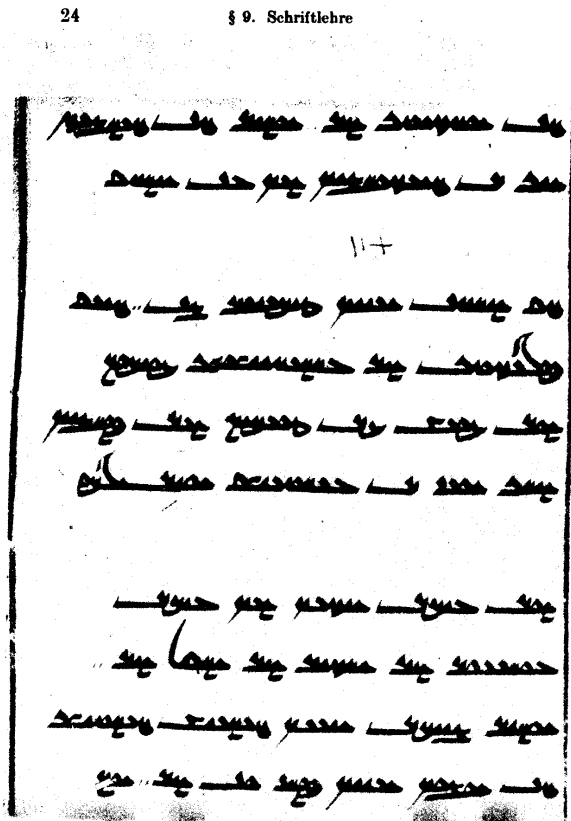


Figure 39: Transcription of an Old Uyghur manuscript in a grammar of Old Turkic (from von Gabain 1950: 22–23). Continued in fig. 40.



§ 9. Die uigurische Schrift 25

-qa oybatı-lar, ular-qa qılmıs	11
-aya qawturmaq-lıy-in alqu	12
[chin. Blattzählung:] 12	
qu ¹ -larqa inay tuginür m(ä)n. qayu	13
bodis(a)t(a)w-lar y(a)rtıqančuı konyul	14
-lüg küc-kä tükül-lig bolmaq	15
-lari üzä yirtinçü-nüy ädgü	16
-lüg-iñü asıy-lıy-ıñı	17
yoriyur-lar ärsär-lär alp-lar.	18
olar manı ayıy qilinç qiltac.	19
-qa umuy inay bolzun-lar. ol	20

¹) Dittographie

Figure 40: Transcription of an Old Uyghur manuscript in a grammar of Old Turkic (from von Gabain 1950: 24–25). Continued from fig. 39.

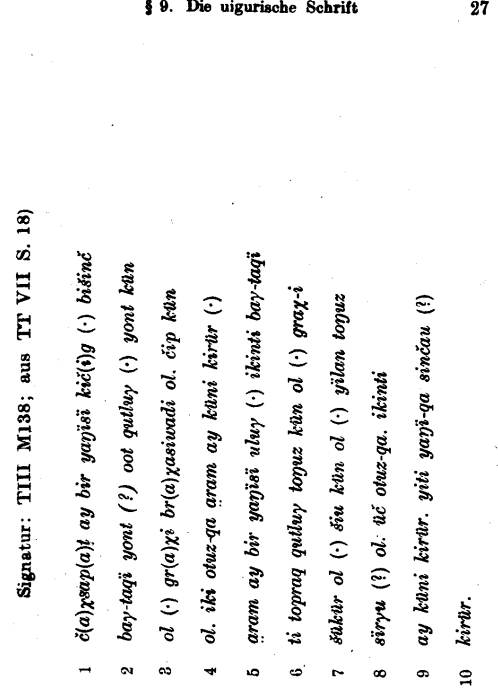
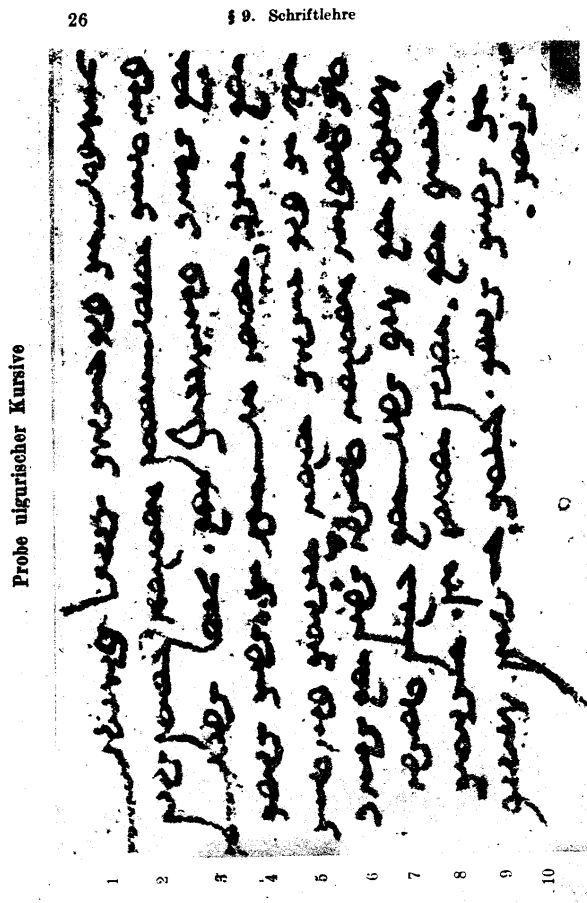


Figure 41: Transcription of an Old Uyghur manuscript in a grammar of Old Turkic (from von Gabain 1950: 26–27).

Uigurica II.

93

S. 46. *čaidan* stammt vielleicht aus dem chinesischen 齋壇 *čai-t'an* (alte, aus der Intonation zu erschließende Form: *čai-dan*), wörtlich »Fasten-Platz« oder »Fasten-Halle« (Giles, Lex., gibt die Bedeutung »altars of abstinence«, — »Taoist temples or halls«).

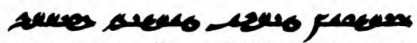
S. 48. Zu dem Ausdruck *ymki* »sitzen« (*ohur-*) sind die chinesisch-buddhistischen, mit 坐 *tso* »sitzen« zusammengesetzten Ausdrücke zu vergleichen:

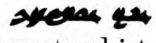
坐臘 to retreat during the twelfth moon, W. Williams, Dict;

打坐恭禪 to meditate in a retreat, ebenda;

坐安居 rester en retraite (St.-Julien, Ex. prat. S. 169) sc. retraite religieuse, ebenda;

坐夏 être sédentaire dans la retraite d'été, ebenda S. 191.

Ebenda. *tngrim* war tatsächlich eine Titulatur, denn unter den Fresken der Turfanexpedition II (A. von Le Coq) befindet sich die Abbildung einer uigurischen Prinzessin mit der Beischrift  *ögrünč tigin tngrim körki* = das Bild der Prinzessin Ögrünč (Freude). Vgl. auch den Titel *tngrilär* im Bekenntnis der Üträt, S. 80 Z. 64.

Ebenda.  [nach Radloff *El ökäsi* und ihm zufolge »Volksmutter« zu übersetzen] ist nicht *Il ögäsi* auszusprechen, sondern *Il ügäsi*, wie die chinesische Umschreibung beweist. Auf einem Fragmente des Kara Balgassun-Denkmalms findet sich nämlich der Titel

內宰相頡于伽思,

aus dem Schlegel (Chinesische Inschrift auf dem uigurischen Denkmal in Kara Balgassun S. 11) einen »inneren Minister, Kit-kan ka-su« oder »Kirkhan-kaš« (ebenda S. 11, 12) herausliest. Schlegel hat eigenmächtig 于 (*i*) in 干 (*kan*) verändert, da nach seiner Meinung die Bücher der T'ang-Dynastie maßgebend seien, nicht die Steininschriften! Umgekehrt vielmehr sind die durch Büchertradition überlieferten Titel

大相頡于迦斯 und 內宰相頡于伽思

in 大相頡于迦斯 und 內宰相頡于伽思

der Premierminister Il ü- gä- si der innere Minister Il ü- gä- si

zu restituieren. *Il ügäsi* »Ruhm des Reichs« (ungefähr الملك *فخر*) wird (wie الملك *نظام*) ein Titel gewesen sein, nicht ein Name. Damit entfällt auch die sachliche Schwierigkeit, den *Il ügäsi*, der schon a. 781 erster Minister war, noch 60 Jahre später, a. 841, fast am Ende der Glanzzeit des

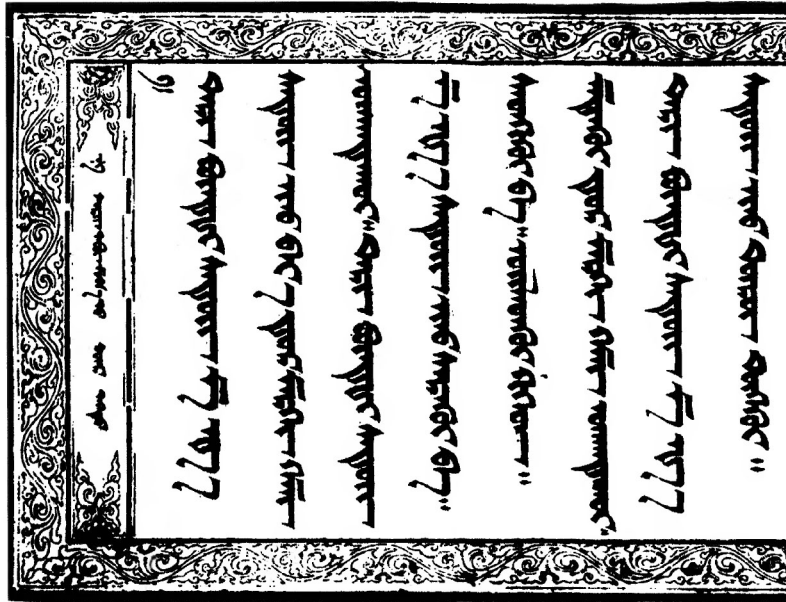
Figure 42: Excerpt from Müller's *Uigurica* showing Old Uyghur text in a horizontal layout (1910: 93). Note the orientation of the glyphs, turned 90 degrees clockwise in relation to their appearance in the code chart.

TABLE 49.4: The Mongolian Script

Mongol. Value	Initial	Medial	Final	Separate	Miscellaneous	Mongol. Value
a	ᠠ		ᠡ	ᠢ		
e	ᠡ	ᠢ	ᠣ	ᠤ	ᠤ	ba/e
i (yodh)	ᠢ	ᠣ	ᠤ	ᠥ	ᠥ	k/ga/e
o/u (waw)	ᠣ	ᠣ	ᠣ	ᠣ	ᠣ	bi
ö/ü=waw+yodh	ᠣ	ᠣ	ᠣ	ᠣ	ᠣ	k/gi
in non-1st syll.		ᠣ	ᠣ			
n before vowel	ᠨ	ᠨ	ᠨ		ᠨ	bo/u
n syll./wd. final	ᠨ	ᠨ	ᠨ	ᠨ		k/go/u
q	ᠬ	ᠬ	ᠬ	ᠬ		
γ before vowel	ᠬ	ᠬ	ᠬ			
γ syll./wd. final	ᠬ	ᠬ	ᠬ	ᠬ		
b	ᠪ	ᠪ	ᠪ			
s	ᠰ	ᠰ	ᠰ			
š	ᠰ	ᠰ	ᠰ			
s final (Uyg. z)	ᠰ	ᠰ	ᠰ			
t/d (taw)	ᠲ	ᠲ	ᠲ	ᠲ		
d/t (lamedh)	ᠲ	ᠲ	ᠲ	ᠲ		
l	ᠯ	ᠯ	ᠯ			Mongyol
m	ᠮ	ᠮ	ᠮ			
č	ᠴ	ᠴ	ᠴ			
j/y (medial: top, j; bottom, y)	ᠵ	ᠵ	ᠵ			ml
k/g	ᠬ	ᠬ	ᠬ			ja
r	ᠷ	ᠷ	ᠷ			
w/v	ᠸ	ᠸ	ᠸ			
h	ᠬ	ᠬ	ᠬ			
p	ᠮ	ᠮ	ᠮ			

Figure 43: Table showing letters of the Mongolian script (from Kara 1996: 545). See table of Old Uyghur letters from the same source in fig. 4.

SAMPLE OF MONGOLIAN



1. Transliteration: tʰr pwβʰɣy sʰɔwβʰ mʰ hʰ /sʰɔwβʰ ʰynw pʰy ʰdwr mʰrkʰn
 2. Normalization: tere bōdhi-saduva ma-hā-saduva inu bey-e-dür mergen
 3. Gloss: that bodhisattva mahāsattva 3POSS body-DAT wise
1. kʰmʰn / ʰwqʰqɔʰqwy : tʰr pwβʰɣy sʰɔwβʰ mʰ hʰ /sʰɔwβʰ ʰynw sereküi
 2. kemen / uqaydaqui tere bōdhi-saduva ma-hā-saduva inu sereküi
 3. saying should.know that bodhisattva mahāsattva 3POSS waking
1. ba : / sʰtkykwy pʰ : ʰwylʰtkwy kykʰt : / mʰɔʰkwy dwr mʰrkʰn kʰmʰn
 2. ba : / sedkiküi ba üiledküi kiged medeküi-dür mergen kemen
 3. and thinking and acting as.well knowing-DAT wise saying
1. ʰwqʰqɔʰqwy : / tʰr pwβʰɣy sʰɔwβʰ mʰ hʰ /sʰɔwβʰ ʰynw twyrwn twykʰkwy :
 2. uqaydaqui tere bōdhi-saduva ma-hā-saduva inu törön tügeküi
 3. should.know that bodhisattva mahāsattva 3POSS born spreading

‘You should know: that bodhisattva and mahasattva is wise in (the knowledge of) body. You should know: that bodhisattva and mahasattva is wise in watchfulness, thinking, acting as well as perceiving. That bodhisattva (is wise in the knowledge of) the sense organs and sense objects (lit. what is being generated and what is spreading.’

– From the printed Mongol Kanjur, vol. 49, folio 2A. Text without diacritics.
 Early 18th century blockprint.)

Figure 44: Sample Mongolian text (from Kara 1996: 546). Compare the Mongolian block print with the Old Uyghur block print in fig. 12.