Proposal to encode the Chorasmian script in Unicode

Anshuman Pandey pandey@umich.edu

July 26, 2018

1 Introduction

This proposal is a revision of the following:

- L2/17-054R: "Proposal to encode the Khwarezmian script in Unicode"
- L2/18-010R: "Proposal to encode the Khwarezmian script in Unicode"
- L2/18-164R: "Proposal to encode the Chorasmian script in Unicode"

It incorporates comments provided by the UTC Script Ad Hoc Committee in:

- L2/17-255: Recommendations to UTC #152 July-August 2017 on Script Proposals
- L2/18-039: Recommendations to UTC #154 January 2018 on Script Proposals
- L2/18-168: Recommendations to UTC #155 April-May 2018 on Script Proposals
- L2/18-241: Recommendations to UTC #156 July 2018 on Script Proposals

The major changes to the previous version (L2/18-164R), which incorporate editorial revisions and comments from the latest script ad hoc report, are as follows:

- Revised encoding for aleph, and change of the representative glyph for ALEPH
- · Additional details to support the separate encoding for WAW and YODH, and glyph enhancements
- Improvements to the representative glyph for ZAYIN
- Change of the representative glyph for NUN
- Change of the joining class for AYIN
- Corrections to the shaping properties of non-joining letters
- Details on vertical orientation of text (see § 8)

2 Background

The 'Chorasmian' script is one of two that were used for recording the Chorasmian language. These are: 1) the indigenous script discussed here, which is derived from Imperial Aramaic; and 2) the Arabic script.

The indigenous script developed from the Aramaic script, which was likely introduced to Chorasmia at the beginning of the 4th century BCE by Achaemenid scribes. It is classified by scholars into 'lapidary' and 'cursive' types (Pavel Lurje, personal communication, December 2017 and July 2018):

The 'lapidary' occurs, for instance, on silver bowls no. 1 and no. 2 from Isakovka (Исаковка), see fig. 33 here; and on a flask found in 2005 at Chirik-rabat (Чирик-рабата), described in Ivantchik and Lurje (2013: 286), see fig. 34 here. Similar types occur on ostraca from Koy-krylgan Kala (Кой-крылган-кала). The earliest lapidary inscriptions are dated to the Achaemenid period and appear in a script closely related to Imperial Aramaic. They are the earliest attestations of the Chorasmian written language (Livshits 2003: 147–148). The 'lapidary' types are all non-joining *abjad*-s (see fig. 35).

The 'cursive' Chorasmian script is a development of the lapidary style. It is fully attested in the records from Toprak Kala, 3rd century CE. It may be considered the 'normative' or 'national' Chorasmian script. It is attested on at least the following materials:

- Coinage with Chorasmian legends, which are the earliest attested records in the script, from the 2nd century CE onwards (see Vainberg 1977, Federov 2005). The coins have been classified by Vainberg and are referred to using the Cyrillic and Roman numeral designations BII–ΓVI. Facsimilies of these coins are shown in fig. 2–6 and tracings of inscriptions are shown in fig. 7–13. Some coins are bilingual: class B coins have inscriptions in Chorasmian and Greek, and class ΓV have transcriptions in Chorasmian and the cursive Sogdian script.
- Inscriptions on wooden items and leather from the palace at Торгак Kala (Топрак-кала), dated to the 3rd century се.
- Leather inscriptions and documents from a fort at Yakke Parsan (Якке парсан) dated to the 8th century CE (see fig. 14).
- Inscriptions on silver vessels dated between the 6th and 8th centuries CE. Reproductions of nearly all vessels and their inscriptions were published in Smirnov (1909) and republished in Azarpay (1969). Seven bowls and one pitcher, along with their inscriptions, are shown in fig. 16–23.
- Ossuary inscriptions at Tok Kala (Ток-кала), from the 7th and 8th centuries се. The script of these records represent a development of the style used in the Toprak Kala documents. There are around 100

of these inscriptions, of which nine were initially deciphered by Tolstov and Livshits (1964), shown here in fig. 24–32. Additional inscriptions were deciphered by Lurje (2013).

Of the 22 letters of the Aramaic alphabet, 19 are attested collectively across the relevant sources. Chorasmian analogues for Aramaic *teth*, *qoph*, and *sadhe* are not attested. Silver vessel and ossuary inscriptions contain all 19 letters, as well as numerical signs. Coins have a smaller subset of letters, and no numerical signs. The ossuary texts contain additional characters for marking grammatical features. A comparison of the repertoire and letterforms made by Vainberg (1977, plate VIII) is reproduced here in fig. 1. A list of signs used on the silver vessel inscriptions has been produced by Lurie (forthcoming), see fig. 15 here.

Although there is no formal evidence of a standard form of the script, the inscriptions on Tok Kala ossuaries indicate the existence of common scribal practices in terms of orthography and letterforms. The use of cursive practices for joining letters of a word gave the script distinctive graphical and structural features. The rules for connecting letters in order to maintain distinctions between letters with similar shapes.

The Chorasmian script is related to other Iranian scripts derived from Imperial Aramaic, such as Inscriptional Parthian; Inscriptional, Psalter, and Book Pahlavi; and the Old Sogdian of the 'Ancient Letters' and the later Sogdian 'formal' and 'cursive' scripts (see table 1). However, among these, Chorasmian was more conservative in its retention of older letterforms and it underwent considerably less change than its sister scripts (Tolstov and Livshits 1964: 234).

After continuous usage over 800 years, the Chorasmian script was replaced by the Arabic script by the turn of the 9th century. Some insight into the demise of the script, among other aspects of Chorasmian culture and society, is provided by the medieval scholar Abū Rayḥān Muḥammad ibn Aḥmad Al-Bīrūnī. Known more commonly as 'Al-Biruni', this native of Chorasmia is considered one of the greatest scholars of the medieval Islamic period. In his *Al-Āthār al-bāqiya 'an al-qurūn al-ḥāliya* (*The Remaining Signs of Past Centuries*), which was completed in 1000 CE, Al-Biruni wrote:

When Kutaiba ben Muslim had conquered Khwârizm a second time ... [he] had extinguished and ruined in every possible way all those who knew how to write and to read the Khwârizmî writing, who knew the history of the country and who studied their sciences. In consequence these things are involved in so much obscurity, that it is impossible to obtain an accurate knowledge of the history of the country since the time of Islam (not to speak of pre-Muhammadan times). (Sachau 1879: 41–42)

The change of orthography for the Chorasmian language was soon after compounded by a larger linguistic change. By the 14th century, the native Iranian language was replaced by Turkic languages. The script is no longer used by a native community. However, there has been active scholarship on Chorasmian studies since the early 20th century. The field was established by Russian scholars, who conducted archaeological excavations in the region, which yielded numerous epigraphical and inscriptional records. Since that time, scholarship on the language, script, and culture has continued to grow.

3 Script identifier

The native name of the script is unknown. Therefore, the identifier 'Chorasmian' is proposed for the script in Unicode. This name is used in the *Encyclopædia Iranica*, eg. MacKenzie (1991) and Humbach (1998). It also aligns with a scholarly periodization of the history of the language and script: Some experts apply the term 'Chorasmian' to the pre-Islamic period and 'Khwarezmian' to the post-Islamic period (Lurje, personal correspondence, April 2018). As the proposed script was used in the pre-Islamic period, the designation 'Chorasmian' is appropriate. Moreover, 'Chorasmian' has been used in English since the early 19th century, and will facilitate identification of the script within a global context. The variant name 'Khwarezmian' is also used in a generic sense in some sources, particularly in transcriptions of the Russian 'Хорезмийский'. For this reason, 'Khwarezmian' has been given as an alias for the script in the names list.

4 Script details

4.1 Structure

The Chorasmian script proposed for encoding is a cursive joining *abjad*. It is written from right to left, with lines that advance from top to bottom. Some inscriptions are written vertically with letters rotated 90° counter-clockwise with lines that advance from left to right (see § 8).

Letters are classified as dual-joining, right-joining, and non-joining. Dual-joining and right-joining letters have contextual shapes that are determined by adjacent letters. In some sources the left-side connection of a dual-joining letter is suspended (see § 6).

4.2 Unification

The encoding for Chorasmian is based upon the 'cursive' script styles attested in inscriptions on coins, silver vessels, and ossuaries. Although the style on these materials span several centuries, it is practical to consider them as developmental phases of a distinctive 'Chorasmian' script. The development culminates in the style used in the Tok Kala ossuaries. Unifying the various styles of the cursive script in a single encoding enables texts to be represented using the same underlying character set.

The encoding for the 'cursive' Chorasmian script does not encompass the 'lapidary' type described earlier. These non-joining scripts that may be unified with the Imperial Aramaic encoding, for purposes for representation in Unicode.

4.3 Representative glyphs

In general, the representative glyphs are based upon letterforms used in the style of Chorasmian in the Tok Kala ossuary inscriptions. However, on account of homography in this latest style of the script, the representative glyph for some letters are based upon forms in silver vessel inscriptions, where distinctions between letters are apparent (see details in § 5.2).

¹ Notably in the poem "Alastor; or, The Spirit of Solitude" (1816) by Percy Bysshe Shelley: "At length upon the lone Chorasmian shore / He paused, a wide and melancholy waste / Of putrid marshes. A strong impulse urged / His steps to the sea-shore. A swan was there, / Beside a sluggish stream among the reeds. / It rose as he approached, and with strong wings / Scaling the upward sky, bent its bright course / High over the immeasurable main. / His eyes pursued its flight." (lines 272–280)

The representative glyph is the isolated or independent form of a letter, ie. the shape of a letter in final position or unjoined to the preceding letter on account of a break in cursive joining. Only the isolated form of each letter is included in the proposed repertoire. Positional forms are to be maintained in a font and substitution of glyphs is to be performed by the shaping engine (see § 6). In some sources the positional forms of several letters have the same or similar shape (see § 5.1). The display of script style is to be managed by custom fonts.

4.4 Character repertoire

The proposed repertoire for Chorasmian contains 28 characters: 21 letters and 7 numbers. The code chart and names list follows p. 12. The encoded set may differ from traditional and scholarly inventories of script varieties that occur in written and inscriptional sources. Such differences naturally arise from the requirements for digitally representing a script in plain text and for preserving the semantics of characters.

Traditional names for Chorasmian letters are not attested. Therefore, Unicode character names are based upon those of 'Imperial Aramaic' characters. This convention has been followed for Unicode encodings of other Iranian scripts such as 'Inscriptional Parthian', etc. In this document names in italics refer to scholarly names for graphemes while names in small capitals refer to Unicode characters, eg. — is *aleph* and Chorasmian Letter aleph. For sake of brevity, the descriptor 'Chorasmian Letter' is dropped when refering to Chorasmian characters, eg. Chorasmian letter aleph is referred to as Aleph. Characters of other scripts are designated by their full Unicode names.

Latin transliteration of Chorasmian follows the current scholarly convention, with Aramaic heterograms given in uppercase letters. A hyphen is used in transliterations to indicate special a special letter or a rendering of a letter that differs from its normatively defined joining behavior.

4.4.1 Letters

Letters whose joining class is marked by an asterisk are exceptions to regular joining behavior (see § 6)

Character name	Glyph	Joining	Latin
CHORASMIAN LETTER ALEPH	_	dual	,
CHORASMIAN LETTER SMALL ALEPH	4	non	_)
CHORASMIAN LETTER BETH	د	dual	b
CHORASMIAN LETTER GIMEL	7	dual	g
CHORASMIAN LETTER DALETH	7	right	d
CHORASMIAN LETTER HE	7	right	h
CHORASMIAN LETTER WAW	1	right	w
CHORASMIAN LETTER CURLED WAW	,	non	-w

CHORASMIAN LETTER ZAYIN	1	dual*	z
CHORASMIAN LETTER HETH	п	right	ķ
CHORASMIAN LETTER YODH	1	right	у
CHORASMIAN LETTER KAPH	フ	dual	k
CHORASMIAN LETTER LAMEDH	>	dual	l
CHORASMIAN LETTER MEM	Þ	right	m
CHORASMIAN LETTER NUN	1	dual	n
CHORASMIAN LETTER SAMEKH	ھ	dual	S
CHORASMIAN LETTER AYIN	¥	non	C
CHORASMIAN LETTER PE	,	dual*	p
CHORASMIAN LETTER RESH	7	right	r
CHORASMIAN LETTER SHIN	W	right	š
CHORASMIAN LETTER TAW	æ	dual	t

4.4.2 Numbers

The numerical notation system is described in § 7.

Character name	Glyph	Joining	Value
CHORASMIAN NUMBER ONE	1	non	1
CHORASMIAN NUMBER TWO	11	non	2
CHORASMIAN NUMBER THREE	111	non	3
CHORASMIAN NUMBER FOUR	1111	non	4
CHORASMIAN NUMBER TEN	٥	right	10
CHORASMIAN NUMBER TWENTY	3	dual	20
CHORASMIAN NUMBER ONE HUNDRED	3	left	100

4.5 Punctuation

Spaces are commonly used for separating words in the ossuary inscriptions and on some silver bowls. There are no special signs for punctuation.

4.6 Line-breaking

There are no formal rules for the breaking of words at the end of line. Moreover, the available sources do not contain text with words broken across lines. 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 occur after words.

4.7 Collation

The sort order of the letters follows the encoded order:

```
→ ALEPH < A SMALL ALEPH < J BETH < J GIMEL < J DALETH < J HE < I WAW <
CURLED WAW < I ZAYIN < THETH < J YOOH < J KAPH < J LAMEDH < THEM <
I NUN < J SAMEKH < J AYIN < J PE < J RESH < W SHIN < TO TAW
```

5 Character Details

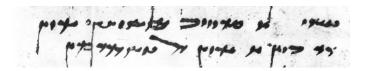
5.1 Homography

A peculiarity of Chorasmian is that the shapes of several letters resemble each other in different styles of the script. The reasons for this are unknown. It may be a natural development or it may be scribal carelessness. Moreover, the merger of letterforms is inconsistent across the sources. The merger of letter shapes occurs most often in cursive-medial position. While the nominal form of a letter is distinctive, its contextual form may be similar or identical to the contextual form of another. The shapes of letters such as waw and yodh, and daleth and resh, typically merged in Aramaic-based Iranian scripts; however, the sources show distinctions between some pairs in Chorasmian. These distinctions have been preserved in the proposed repertoire.

beth, nun, pe The letters > beth, nun, pe are often written using a shape similar to 1 in cursive-medial position. Their nominal forms are distinctive, and are distinguished by the degree of curvature of the primary stroke.

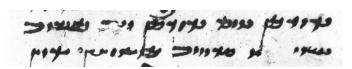
gimel, he, kaph The letters \neg gimel, \neg he, \neg kaph have the same basic structure: a horzontal stroke attached to a descending stroke. The kaph is identified by its broad, horizontal top stroke, and the elongated descender of its final form. The gimel is written with a 90° angle. The he consistently appears with a descending top stroke and angled bottom stroke; and when in cursive-final position following a dual-joining letter, may angled 45°, opening to the left. The distinctiveness of gimel is evidenced by the archaic form \neg used in silver vessels (see #6 in fig. 21). These three letters are further differentiated by their joining behaviors. The gimel and kaph are dual joining and he is right-joining, as indicated by their interactions with \mathbf{l} waw in gwšt and \mathbf{r} hwnšk (TK 52, fig. 24), and \mathbf{r} transfer the probability transfer to the same basic structure: a horzontal stroke attached to a descending top stroke, and the elongated descender of its final form. The gimel at the probability transfer to the same basic structure: a horzontal stroke attached to a descending top stroke, and the elongated descender of its final form. The gimel at the probability appears with a descending top stroke, and the elongated descender of its final form. The probability attached to a descending top stroke, and the elongated descender of its final form. The probability attached to a descending top stroke and the elongated descender of its final form. The probability attached to a descending top stroke, and the elongated descender of its final form. The probability attached to a descender of its final form. The probability attached to a descender of its final final form. The probability attached to a descender of its final form. The probability attached to a descender of its final final form. The probability attached to a descender of its final fi

daleth, ayin, resh The letters א daleth, א ayin, א resh have the same basic structure, but are differentiated by variations in the angle and length of strokes. The resh and daleth are differentiated by the length of the primary stroke, with that of א resh being longer than that of א daleth as shown in א grdm'n (TK no. 25, fig. 25 and excerpt below). Such a distinction appears to be carried over from Imperial Aramaic, where the original א daleth and א resh differ by the length of the primary stroke, and both differ from v ayin. The Chorasmian v ayin is written with a narrower angle at the left as compared to daleth and resh, and the left stroke connects at a lower point to the right stroke, eg. > v L (below) and vBDt.

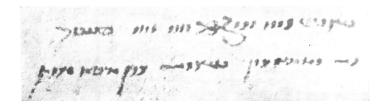


waw, yodh The shapes of waw and yodh appear to have merged in the hand of some scribes and engravers to such an extant that it is difficult to distinguish the two letters in some inscriptions. For example, they occur as 1, 1, 1. The degree of similarity is increased by their identical joining behavior. However, in some sources they are distinguished.

For example, in Tok Kala no. 25, waw and yodh are clearly distinct in **BYWM** (see fig. 25 and excerpt below). In this word, yodh is a curved stroke whose terminal extends below the baseline, while waw is written as a straight vertical stroke that stops at the baseline. This representation of waw and yodh is consistent in this source: I for waw in שרוול brwrtn, מרוול brwrtn, מר



The curved sign $\mathbf{1}$ for yodh is also evident in TK 19 (fig. 26, and below). It is clearly observed in $\rightarrow \mathbf{1}$ y' and less prominently in $\rightarrow \mathbf{1}$ $NP\check{S}Y'$; compare these with the straight form $\mathbf{1}$ for waw in tpnkwk:



While the similarity between waw and yodh may suggest unifying these letters as a single character, there is strong evidence from Tok Kala no. 25 and no. 19 for their distinctive representations. These sources illustrates that the scribes intentionally distinguished waw from yodh by using a more conversative stroke for the former, and a curved stroke for the latter. Therefore, it is practical to encode these two letters separately.

5.2 Letters

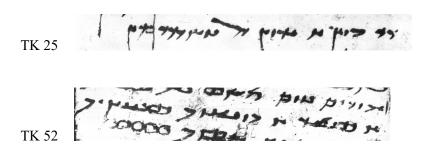
aleph The representative glyph \rightarrow for *aleph* is based upon the isolated form used at the end of words in coins, silver vessels, and ossuaries. The letter has special shaping behaviors:

- Medial The connection of medial aleph follows the rule for the initial form; compare the attachment to a following letter at the midpoint in מבאי k'k'n-w, שנאי tnb'r, מבאי grdm'n with connections at the baseline in pr'ny'ty. In coin and silver vessel inscriptions, the archaic form x of cursive medial aleph often occurs instead of x; this form joins at the baseline on both sides. It is a stylized form that is to be treated as a glyphic variant, and managed at the font level.
- Final Despite being a dual-joining letter, when aleph occurs in word final position, it is written using its nominal form —, which is disconnected from the preceding letter, even when the latter is dual-joining, eg. שלכי MLK-' in coinage (eg. class Γ III, fig. 9); מון ZWZN-' in silver vessels (#5, fig. 20); און ארץ ' in ossuaries (TK 25, fig. 25; also the excerpt below). In MLK' and ZWZN-' the aleph is disconnected from the preceding kaph and nun, respectively, which are dual-joining letters. Moreover, in both of these cases, the letter that precedes aleph is rendered using its final or isolate form. Based upon the evidence, the writing of word-final aleph in this manner is conventional behavior for the letter. There is no attestation for a form of cursive final aleph that joins to the preceding letter.

In order to produce this disconnected word-final *aleph*, the zwnj should be placed between a dual-joining letter and the following ALEPH. This would render the preceding letter in its final or isolated form, and would produce the isolated form of *aleph* (see § 6.3). This behavior of *aleph* creates a complication: as there is no attested form for a cursive final *aleph*, an artificial cursive joining glyph must be created for cases where word-final *aleph* is not preceded by zwnj. This artificial form *a is the same as the cursive initial and medial forms. It would connect to the preceding letter. Usage of this connected cursive-final *aleph* is strongly discouraged. If a final cursive-joining form of *aleph* is identified in the future, the artificial form may be replaced by the actual form.



'small' aleph A smaller and raised non-joining form • of aleph is used for indicating a possessive in patronyms in the Tok Kala ossuary inscriptions, eg. אָרָי הייחיי $t'\beta'n'n-k$ (TK 52, fig. 24; excerpt below); $hy^n-k(TK 52, fig. 25; excerpt below)$. It occurs only in the middle or at the end of words. It is an isolate character and lacks cursive joining properties. The 'small' aleph is transliterated as - in scholarly materials, which is also used for transliterating a final aleph that follows a dual-joining letter (transliteration certainly is imperfect). The 'small' aleph is not attested in sources before the Tok Kala inscriptions. It certainly reflects a distinctive development in the final stages of the script. Given the behavior of the • it is encoded as the distinctive character small Aleph. As there is no known native term or scholarly name for this letter, the character name is based upon its graphical attribute.



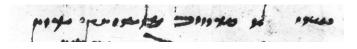
gimel The \neg gimel connects to the following letter at the top edge, eg. gimel + waw as \neg in the name of a day, ששח gwšt; gimel + resh as \neg in \neg grdm'n 'paradise'. Sequences of gimel + waw or yodh may resemble the letter \neg heth, but are distinguishable based upon context. The gimel has the archaic form \neg ; which occurs, for example in silver vessel #6 (see fig. 21).

he The following points should be noted:

- Shaping The method by which , he connects to a preceding letter is determined by the connection point of the latter. The aleph connects to the midpoint of the right stroke of he, eg. אויי אווינל, 'hwrym (TK no. 52, fig. 24). When it follows a dual-joining letter that connects at the baseline, the , he is rotated clock-wise as , such that its right-side point may effect a connection at the baseline, eg. ענ ZNH (see TK. no 52).
- Stylistic variation The Chorasmian he is a right-joining letter. It exhibits stylistic variation in word final and non-final positions in some sources (see § 6.1).

waw The representative glyph for I waw is based upon the form used in Tok Kala no. 25, etc. This form is distinguished from I YODH through its representation as a straight vertical stroke, while YODH is a curved stroke.

'curled' waw In Tok Kala ossuaries no. 25 and no. 52, a grammatical possessive or patronym is indicated using a special form of waw, which is represented as the non-joining special sign (transliterated as -w). It is an isolate character and lacks cursive joining properties. The 'curled' waw is described by Henning as a "vocalization mark", which is "a rounded form reminiscent of an Arabic damm" and occurs in words, such as "yöyn'n-w in TK no. 25 and whwnt'n-w in TK no. 69 (1965: 178). In both these cases, the is attached to personal names that function as patronyms. The "damm" to which Henning refers is the sign encoded as 'U+064F ARABIC DAMMA. Although Henning calls the character a 'sign', it behaves more as a 'letter' than a 'sign'. Moreover, it is a spacing character whose glyph has the same proportions as other letters. The is encoded as a distinctive letter named CURLED waw. As there is no native or scholarly term for this letter, a name based upon its shape has been selected.



zayin The following points about this letter should be noted:

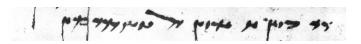
- Shape In the Tok Kala inscriptions zayin is distinguished from waw and yodh by a prominent head, which is curved or notched, depending upon the preference of the scribe. In the silver-vessel inscriptions zayin is distinguished from waw by the curvature of its stroke. These features have been preserved in the representative glyph 1.
- Joining The joining behavior of zayin differs depending upon context. In Aramaic heterograms in the Tok Kala ossuaries it joins to the left, eg. על ZNH, while in the silver vessel inscriptions it does not, eg. אווו ZWZN. Given this varying behavior, zayin is defined as a dual-joining letter. This will permit usage of zwnJ after the letter to break the connection when necessary. By default zayin will join the following letter at the baseline using the shape J. When zwnJ is placed after zayin, the letter will be rendered as 1 (see § 6.3).

yodh The representative glyph for **1** yodh is based upon the form used in Tok Kala no. 25. This form is distinguished from waw through its representation as a curved vertical stroke, while waw is a straight stroke.

mem The Chorasmian *mem* is a right-joining letter. It exhibits stylistic variation in word final and non-final positions in some sources (see § 6.1).

nun The representative glyph \mid for *nun* is based upon the isolated form that occurs in word-final position the Tok Kala inscriptions. Word final and cursive final *nun* is written as \mid in the ossuary inscriptions, in both right-joining and isolated contexts. In coinage and silver vessels, the word final form appears as \bot . The difference in the glyphs is the orientation of the tail: vertical or horizontal. The orientation may be guided by the physical constraints of the face upon which the inscription occurs: the tail is curved or elongated horizontally at the lip of a vessel or at the edge of a coin because there is insufficient space for a vertical tail. For example, \bot in ZWZN- occurs on silver vessels, but would appear as \bot in later styles of the script. A common example is MN, which occurs as \bot in silver vessels and as \bot in ossuary inscriptions. As the Tok Kala inscriptions are the latest records of the script, and as word final, cursive final, and isolated *nun* occurs in these sources with a vertical terminal, this shape has been selected as the representive form of *nun*. If a distinction between \mid and \bot is required in plain text, then the latter may be encoded as a separate letter as was done for the Sogdian forms of *nun*.

ayin This letter is attested only in word-initial position, and only in Aramaic heterograms, in the available sources: שנדים 'BDt 'done, made' (silver bowl #7, fig. 22) and אַ 'L, 'to' (silver bowls and ossuaries, ie. Tok Kala no. 25, see excerpt below). While it appears to join to lamedh in אַ in the ossuary inscriptions, the connection is likely a result of inadequate letter spacing — the stroke of the lamedh being broadened in word-final position — not a cursive property of the letter.



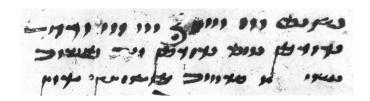
Given its structural similarity to *daleth* and *resh*, it may be practical on palaeographical grounds to define the Chorasmian *ayin* as a right-joining letter. However, as there is an absence of evidence for supporting this recommendation, *ayin* is defined as non-joining (although if joining properties were strictly assigned according to attestations then the ossuary inscriptions would support a left-joining property for *ayin*...). If additional evidence indicates that *ayin* is right-joining, then the joining property may be changed in the future.

Pe Although it defined is a default dual-joining letter, in some sources medial \mathbf{j} PE does not connect to the left, compare \mathbf{j} pbntn with \mathbf{j} p-bntn. The suspension of the left connection of pe occurs in the silver vessels before beth, perhaps in order to distinguish the two letters, as their dual-joining medial shape (\mathbf{j}) is similar in those inscriptions. In order to enable both the left joining and the suspension of the left connection, pe is defined as a dual joining letter. The zwnj is to be used for breaking the connection by placing it after pe (see § 6.3).

shin The letter ψ is represented using the glyphic variant ψ on some coins. This form is to be managed using a font.

taw The following points about this letter should be noted:

- Representative form The representative glyph **o** for taw is based upon the Tok Kala form. This letter has the glyphic variant **o** that has an open right stroke, which is to be managed using a font.
- Ligation In ossuary inscriptions, letters that follow **a** taw join to its left edge without any spacing or extension of the baseline, or are incorporated into the left edge of the glyph, eg. **a** is written as **b** to accommodate a following letter. For instance, taw + yodh is **b** as in אינדעל on TK no. 52 (fig. 24); taw + final nun as **b** tn in אינדעל pbntn, eg. silver bowl #2 (fig. 17); taw + waw is **b** tw in the name שחשם twtwxs on type BI coins (fig. 8).



6 Joining behavior

A summary of the joining properties of Chorasmian letters is given below:

right- & left-joining aleph, beth, gimel, zayin, kaph, lamedh, nun, samekh, pe, taw

right-joining daleth, he, waw, heth, yodh, mem, resh, shin

non-joining ayin

exceptions aleph, zayin, pe

In the tables below, ' X_n ' is nominal form of a letter. The labels ' X_l ', ' X_m ', ' X_r ' refer to the positional forms of a letter when it joins to the left, occurs in the middle, or joins to the right. Note the following annotations in the tables:

- A red dash indicates the connection point of a glyph.
- A vertical bar indicates that a connection occurs without an extension of the baseline.
- Parentheses indicate glyphic variants of contextual forms that occur before certain letters or when the letter occurs in certain positions within a word (see § 5.2 for details).
- An asterisk indicates an artificial form (see description of the given letter for details).

Joining features of dual-joining letters are shown below.

	X_n	X_{r}	X_{m}	X_l
ALEPH	_	*▲_	(~) <u>~</u>	(-)
ВЕТН	د	7	7	د
GIMEL	7	7	7	7
ZAYIN	1	L	_	د
КАРН	フ	ァ	7	ور
LAMEDH	>	7	7	۷
NUN	1	۲	_	ر
SAMEKH	ط	<u> </u>	<u>p</u>	ط
PE)) _	<u>.</u>	د
TAW	æ	Φ.	p _	B

Joining features of right-joining letters are shown below:

	X_n	X_{r}
DALETH	7	7.
НЕ	7	۲,), ۲
WAW	1	L
НЕТН	n	n.
YODH	,	1_
MEM	þ	₽
RESH	7	7
SHIN	W	Ju

6.1 Word-level glyph variants

In some sources, the following right-joining letters have a different form when they occur in non-final positions within a word:

	X_n	word final	word non-final
НЕ	7	(>), >	7
MEM	þ	₽	ct

The primary difference in the word-final and word-non-final forms of *he* and *mem* is the extension of the descending stroke below the baseline in the in the word-final forms. It is difficult to determine if these glyphic variations were practiced uniformly.

- mem Such variation for mem is seen in silver vessel inscriptions (eg. silver bowl #2; fig. 17): עול BYWM and לוב) MN.
- he The word non-final form of he varies stylistically from its word-final form in some sources: in the former its right stroke does not descend below the baseline as it does in the latter; cf. 7 and 7. See, for example 17 hw in TK no. 69 (fig. 28).

However, the Unicode cursive-joining algorithm, which is based upon the model for Arabic, is ignorant of word boundaries. As the Arabic model will be used for Chorasmian, it may not be possible to produce the word-level distinctions for Chorasmian *he* and *mem* using the encoding. The model will use the nominal form of these letters in cursive non-final positions. If a distinction between word final and word non-final glyphs is required, a variation selector may be needed.

6.2 Examples of encoded representations

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 Chorasmian records are given below along with their input strings:

```
<-- ALEPH, 7 GIMEL, -- ALEPH, 7 DALETH, プ KAPH>
'g'dk
                       <-- ALEPH, 1 ZAYIN, ZWNJ, OD TAW, 1 YODH, 7 KAPH>
'ztyk
pbntn
                      <-- ALEPH, ) PE, J BETH, | NUN, @ TAW, | NUN>
                wrma
'r'škrk
                       <- ALEPH, 7 RESH, - ALEPH, W SHIN, 7 KAPH, 7 RESH, 7 KAPH>
             ナミスト
<sup>°</sup>špynšwk
                       ALEPH, W SHIN, J PE, J YODH, I NUN, W SHIN, TAW, J KAPH>
             שמשול
BYWM
                 uid
                      SETH, J YODH, I WAW, † MEM>
bntk
                 <> אורן, NUN, ס TAW, א KAPH>
```

```
BŠNT
                യയ
                       <br/>
SHIN, ا NUN, م TAW> د>
grdm'n
               רדילא
                        <7 GIMEL, 7 RESH, 7 DALETH, 7 MEM, - ALEPH, 1 NUN>
gwšt
                מאש
                        <7 GIMEL, I WAW, W SHIN, OD TAW>
hwnšk
                        <> HE, I WAW, I NUN, W SHIN, J KAPH>
                דושוד
hy'n-'
                 4417
                        <¬ HE, J YODH, → ALEPH, | NUN, ^ SMALL ALEPH>
whwnt'n-w
              الاالكاد
                        < | WAW, | HETH, | WAW, | NUN, | TAW, | ALEPH, | NUN, | CURLED WAW>
ZNH
                        <1 ZAYIN, | NUN 7 HE>
                   יעכ
hwsrw
                חופצו
                        <⊓ HETH, I WAW, ⊅ SAMEKH, 7 RESH, J YODH>
hwpsk
               טווקב
                        <T HETH, I WAW, )PE, ら SAMEKH, ブ KAPH>
YRH
                        <) YODH, 7 RESH, 17 HETH, - ALEPH>
                一口ソリ
k'k'n-w
                محكاه
                        <¬ KAPH, ¬ ALEPH, ¬ KAPH, ¬ ALEPH, | NUN, ¬ CURLED WAW>
KSP
                  צפנ
                        < SAMEKH, و SAMEKH, و PE>
m'ny''ty
              duum
                        <₱ MEM, → ALEPH, ↑ NUN, ♪ YODH, → ALEPH, → ALEPH, ♠ TAW, ♪ YODH>
MR'Y
                 ナイヤ
                        <₱ MEM, ₱ RESH, ■ ALEPH, ₱ YODH>
NPŠY
                        < NUN, J PE, W SHIN, J YODH>
                 ונשונ
sy'wršprn
            מאדשנדן
                        SAMEKH, J YODH, → ALEPH, I WAW, 7 RESH, ₩ SHIN, 2 PE, 7 RESH, | NUN>
^{\circ}BDt
                        <y ayın, J beth, 7 daleth, a taw>
                צערם
\mathcal{L}
                       <y AYIN, > LAMEDH>
                   V
                        PE, 7 RESH, - ALEPH, | NUN, I WAW, | NUN, TAW>
prnxwnt
             כלתות
twtwhs
                        < TAW, I WAW, ™ TAW, I WAW, ™ HETH, ™ SAMEKH>
              ממחפ
tnbryk
               מענוכ
                        < TAW, I NUN, → BETH, → RESH, I YODH, → KAPH>
tpnkwk
               מחבוב
                        < TAW, ) PE, I NUN, J KAPH, I WAW, J KAPH>
```

6.3 Modifying joining behavior

As mentioned above, the letters *aleph*, *zayin*, and *pe* are exceptions to the regular joining behavior. Although they are defind as dual-joining letters, in some sources and contexts their left-side connection is suspended. This modification of regular behavior may be effected using the control character VUL-200C ZERO WIDTH NON-JOINER (abbreviated as ZWNJ). This character is placed after the letter whose left-side connection is to be suspended.

```
pbntn
                   <- ALEPH, ) PE, J BETH, | NUN, TAW, | NUN>
                   ALEPH, ) PE, ZWNJ, J BETH, I NUN, TAW, I NUN>
p-bntn
ZNH
                   <1 ZAYIN, 1 NUN 7 HE>
ZNH
                  <1 ZAYIN, ZWNJ, NUN 7 HE>
              וע
                   <1 ZAYIN, X ZWNJ, I WAW, I ZAYIN, X ZWNJ, I NUN>
ZWZN
                   <1 ZAYIN, [XV] ZWNJ, 1 WAW, 1 ZAYIN, [XV] ZWNJ, 1 NUN, [XV] ZWNJ, - ALEPH>
ZWZN-'
                   <→ MEM, → LAMEDH, → KAPH, W ZWNJ, → ALEPH>
MLK
                   PE, 7 RESH, — ALEPH, I NUN, I WAW, I NUN, TAW>
prnxwnt
         כגתות
          ררחונס <> PE, W ZWNJ, א RESH, → ALEPH, | NUN, I WAW, | NUN, Œ TAW>
prnxwnt
```

6.4 A note on joining behavior

The similarities between nominal and non-initial forms of letters led earlier scholars to suggest that joining rules may vary in particular sources. For such cases, it is be useful to consider Henning's advice:

It becomes then all the more important to observe, in the strictest manner, certain scribal conventions that arise from the material, in particular the rules of linking and separating letters. It seems to me that by refusing any license in such matters we can improve the security of reading [...] Attempts have been made from time to time to arrogate to oneself some license, so as to assert: "in this work W has been connected to the left"; in the long run they have invariably been rejected. (1965: 171)

For instance, the word in TK no. 69 (fig. 28) was interpreted as *nwšy* by Tolstov and Livshits (1964). The reading of the second letter as *waw* lead to this right-joining letter being analyzed as a letter that joins to the left in some Tok Kala texts and as right-joining in others. A more likely interpretation of *NPŠY*. Reading *pe* instead of *waw* not only provides an accurate reading, but also adheres to the spelling conventions of the script and eliminates the need to add complexity to the joining behavior of *waw*.

Similarly, Similarly was interpreted as *tnbryk* by Tolstov and Livshits. This reading required an analysis of medial *resh* as a potentially left-joining letter, contrary to all other evidence. A re-analysis of the fifth letter in this word as *kaph* instead of *resh* provided a more accurate reading of as *tpnkwk*, which carries the same meaning as *tpnkwk* and adheres to rules of the script.

The word איז was interpreted as kw by Tolstov and Livshitz, in which kaph was treated as a non-joining letter. This assumption did not align with the features of kaph in k^*n -w (TK no. 26, fig. 31). Following the behavior of kaph in the latter, reading 17 as hw provides for more a more accurate interpretation.

Such issues are typical in initial stages of script decipherment. At this point in time, the joining behaviors of Chorasmian letters are understood, with the possible exception of *ayin*.

7 Numerical notation

The ordering of numbers follows the right-to-left directionality of the script. The expression of numbers is additive. Compounds of different units are produced by placing larger units first. The exception is the usage of primary units for expressing multiples of hundred, which are placed before the character ONE HUNDRED.

The joining features for Chorasmian numbers are:

	X _n	X_{r}	X _m	X_l	Join
ONE	1	_	_	_	non
TWO	11		_		non
THREE	111	_	_	_	non
FOUR	1111	_	_	_	non
TEN	٥	>	_	_	right
TWENTY	3	})	3	dual
ONE HUNDRED	3	_	_	3	left

Primary units The primary numbers 1–4 are encoded as atomic characters. The numbers 5–9 are expressed using combinations of ONE .. FOUR. This model is based upon the encoding for 'Imperial Aramaic', and related scripts such as 'Inscriptional Parthian', 'Inscriptional Pahlavi', and 'Old Sogdian' (see table 2).

The primary units are expressed using repetitions of the sign 1 for 'one'. Numbers 2–9 are written using sequences of 1 arranged in groups containing three or four instances of the sign. For example, 'm m' for 7 and 'm m' for 8 in TK no. 19 (fig. 26), as well as '1 m m' for 7 in TK no. 25 (fig. 25). Also, 'n m' for the number 5 in silver bowl #2 (fig. 17). The number 5 in silver bowl #5 appears as 'mm' without a spaced grouping, but the extended terminal of the third 1 suggests the intended grouping 'n m' despite lack of spacing (fig. 20).

1	1	1 ONE
2	11	11 TWO
3	111	111 THREE
4	1111	1111 FOUR
5	11 111	<111 THREE, 11 TWO>
6	111 111	<111 THREE, 111 THREE>
7	111 1111	<1111 FOUR, 111 THREE>

Ten The **>** TEN resembles a vertically compressed **>** LAMEDH.

Twenty The sign for **3** TWENTY is derived palaeographically from a vertical stack of two instances of **5** TEN. It is treated as an atomic character.

Multiples of ten are written using sequences of **>** TEN and **3** TWENTY. Even multiples are expressed with repetitions of TWENTY. Odd multiples are produced by attaching TEN at the end.

```
10
              <> TEN>
20
              <3 TWENTY>
30
         3
              <3 TWENTY, > TEN>
              <3 TWENTY, 3 TWENTY>
40
         <del>}</del>}
        >>>
              <3 TWENTY, 3 TWENTY, 5 TEN>
50
60
        <del>}}</del>}
              <3 TWENTY, 3 TWENTY, 3 TWENTY>
70
      >}}}
              <3 TWENTY, 3 TWENTY, 3 TWENTY, 5 TEN>
80
              <3 TWENTY, 3 TWENTY, 3 TWENTY, 3 TWENTY>
      <del>}}}</del>}
90
              <3 TWENTY, 3 TWENTY, 3 TWENTY, 3 TWENTY, 3 TEN>
      ><del>}}}</del>}
```

Hundreds The **3** ONE HUNDRED represents the value 100 and functions as a unit mark.

Multiples of the hundreds are represented using 3 ONE HUNDRED in conjunction with the primary units. The primary units are placed before ONE HUNDRED in the input sequence.

```
100 3 <3 ONE HUNDRED>

200 311 <11 TWO, 3 ONE HUNDRED>

300 3111 <111 THREE, 3 ONE HUNDRED>
```

Composite numbers found in the sources are given below along with their encoded representations:

Higher orders There are no distinctive signs for orders larger than the hundreds.

8 Vertical text

Chorasmian text is oriented horizontally by default. However, according to Pavel Lurje, some Tok Kala ossuary inscriptions may have been inscribed vertically (personal communication, July 2018). In vertical environments, Chorasmian letters are rotated 90° counter-clockwise from their regular upright shapes. The glyphs are ordered from top to bottom with lines that advance from left to right. The orientation is identical to the Upper Indus Inscriptions in Old Sogdian (see L2/16-312R).

By default, Chorasmian is to be oriented horizontally in plain text representations. Vertical orientation is to be considered a formatting preference, and is not to be managed at the encoding level. The "Unicode Technical Report #50: Unicode Vertical Text Layout" describes the <code>Vertical_Orientation</code> (vo) property for specifying the orientation of characters in vertical environments. For Chorasmian, this property would be defined as: <code>Vertical_Orientation=R</code> or <code>vo=R</code>, where the value 'R' indicates that the glyphs are rotated in vertical layout. The rotation is 90° counter-clockwise.

9 Character Properties

9.1 Core data: UnicodeData.txt

```
10FB0; CHORASMIAN LETTER ALEPH; Lo; 0; R;;;;; N;;;;;
10fB1;CHORASMIAN LETTER SMALL ALEPH;Lo;0;R;;;;;N;;;;;
10FB2; CHORASMIAN LETTER BETH; Lo; 0; R;;;;; N;;;;;
10FB3; CHORASMIAN LETTER GIMEL; Lo; 0; R;;;;; N;;;;;
10FB4; CHORASMIAN LETTER DALETH; Lo; 0; R;;;;; N;;;;;
10FB5; CHORASMIAN LETTER HE; Lo; 0; R;;;; N;;;;
10FB6; CHORASMIAN LETTER WAW; Lo; 0; R;;;;; N;;;;;
10FB7; CHORASMIAN LETTER CURLED WAW; Lo; 0; R;;;;; N;;;;;
10FB8; CHORASMIAN LETTER ZAYIN; Lo; 0; R;;;;; N;;;;;
10FB9; CHORASMIAN LETTER HETH; Lo; 0; R;;;;; N;;;;;
10FBA; CHORASMIAN LETTER YODH; Lo; 0; R;;;;; N;;;;
10FBB; CHORASMIAN LETTER KAPH; Lo; 0; R;;;;; N;;;;;
10FBC; CHORASMIAN LETTER LAMEDH; Lo; 0; R;;;;; N;;;;;
10FBD; CHORASMIAN LETTER MEM; Lo; 0; R;;;;; N;;;;
10FBE; CHORASMIAN LETTER NUN; Lo; 0; R;;;;; N;;;;;
10FBF; CHORASMIAN LETTER SAMEKH; Lo; 0; R;;;;; N;;;;;
10FC0; CHORASMIAN LETTER AYIN; Lo; 0; R;;;;; N;;;;;
10FC1; CHORASMIAN LETTER PE; Lo; 0; R;;;;; N;;;;;
10FC2; CHORASMIAN LETTER RESH; Lo; 0; R;;;;; N;;;;
10FC3; CHORASMIAN LETTER SHIN; Lo; 0; R;;;;; N;;;;;
10FC4; CHORASMIAN LETTER TAW; Lo; 0; R;;;;; N;;;;;
10FC5; CHORASMIAN NUMBER ONE; No; 0; R;;;; 1; N;;;;
```

```
10FC6; CHORASMIAN NUMBER TWO; No; 0; R;;;; 2; N;;;;;
10FC7; CHORASMIAN NUMBER THREE; No; 0; R;;;; 3; N;;;;;
10FC8; CHORASMIAN NUMBER FOUR; No; 0; R;;;; 4; N;;;;
10FC9; CHORASMIAN NUMBER TEN; No; 0; R;;;; 10; N;;;;
10FCA; CHORASMIAN NUMBER TWENTY; No; 0; R;;;; 20; N;;;;
10FCB; CHORASMIAN NUMBER ONE HUNDRED; No; 0; R;;;; 100; N;;;;
```

9.2 Linebreak data: LineBreak.txt

```
10FB0..10FC4;AL  # Lo [21] CHORASMIAN LETTER ALEPH..CHORASMIAN LETTER TAW 10FC5..10FCb;AL  # No [7] CHORASMIAN NUMBER ONE..CHORASMIAN NUMBER ONE HUNDRED
```

9.3 Shaping properties: ArabicShaping.txt

```
10FB0; CHORASMIAN ALEPH; D; No Joining Group
10FB1; CHORASMIAN SMALL ALEPH; U; No Joining Group
10FB2; CHORASMIAN BETH; D; No_Joining_Group
10FB3; CHORASMIAN GIMEL; D; No_Joining_Group
10FB4; CHORASMIAN DALETH; R; No_Joining_Group
10FB5; CHORASMIAN HE; R; No_Joining_Group
10FB6; CHORASMIAN WAW-YODH; R; No Joining Group
10FB7; CHORASMIAN CURLED WAW; U; No Joining Group
10FB8; CHORASMIAN ZAYIN; D; No Joining Group
10FB9; CHORASMIAN HETH; R; No Joining Group
10FBA; CHORASMIAN YODH; R; No_Joining_Group
10FBB; CHORASMIAN KAPH; D; No Joining Group
10FBC; CHORASMIAN LAMEDH; D; No_Joining_Group
10FBD; CHORASMIAN MEM; R; No Joining Group
10FBE; CHORASMIAN NUN; D; No Joining Group
10FBF; CHORASMIAN SAMEKH; D; No Joining Group
10FC0; CHORASMIAN AYIN; U; No Joining Group
10FC1; CHORASMIAN PE; D; No_Joining_Group
10FC2; CHORASMIAN RESH; R; No Joining Group
10FC3; CHORASMIAN SHIN; R; No_Joining_Group
10FC4; CHORASMIAN TAW; D; No Joining Group
10FC5; CHORASMIAN ONE; U; No Joining Group
10FC6; CHORASMIAN TWO; U; No Joining Group
10FC7; CHORASMIAN THREE; U; No Joining Group
10FC8; CHORASMIAN FOUR; U; No Joining Group
10FC9; CHORASMIAN TEN; R; No Joining Group
10FCA; CHORASMIAN TWENTY; D; No Joining Group
10FCB; CHORASMIAN ONE HUNDRED; L; No Joining Group
```

10 References

Anderson, Deborah, et. al. 2017 "Recommendations to UTC #152 July-August 2017 on Script Proposals". L2/17-255. https://www.unicode.org/L2/L2017/17255-script-ad-hoc.pdf

```
———. 2018a. "Recommendations to UTC #154 January 2018 on Script Proposals". L2/18-039. https://www.unicode.org/L2/L2018/18039-script-adhoc-rec.pdf
```

——. 2018b. "Recommendations to UTC #155 April-May 2018 on Script Proposals". L2/18-168. https://www.unicode.org/L2/L2018/18168-script-rec.pdf

- ——. 2018c. "Recommendations to UTC #156 July 2018 on Script Proposals". L2/18-241. https://www.unicode.org/L2/L2018/18241-script-ad-hoc.pdf
- Azarpay, Guitty 1969. "Nine Inscribed Choresmian Bowls". Artibus Asiae, vol. 31, no. 2/3, pp. 185–203.
- "Coins of Central Asia". http://www.sogdcoins.narod.ru/english/khwarezm/coins.html
- Durkin-Meisterernst, Desmond. 2009. "Khwarezmian'. In *The Iranian Languages*, Gernot Windfuhr [ed], pp. 336–376. New York: Routledge.
- Federov, Michael 2005. "On Some Articles in the Recent Issues (Nos. V, VI, VII) of the Нумизматика Центральной Азии (Numismatics of Central Asia)". *Central Asiatic Journal*, vol. 49, no. 2 (2005), pp. 175–203.
- ———. 2006. "First Reported Find of an Early Medieval Khwarezmian Drachm in the Kyrgyz Republic". *American Journal of Numismatics*, vol. 18 (2006), pp. 123–130.
- Fedorov, Michael and Kuznetsov, Andrew. 2008. "A hoard of early mediaeval Khwarezmian drachms from the Kuiuk-kala hill fort". *The Numismatic Chronicle*, vol. 168 (2008), pp. 446–451.
- ——. 2011. "On some previously unknown Khwarazmian drachms and the names of rulers on them". *Iran*, vol. 49 (2011), pp. 79–88.
- ———. 2012. "Returning to the Kuiuk-Kala hoard of early Mediaeval Khwarezmian drachms: Khorezm, eastern delta of the Amu Darya, Kuiuk-kala hill fort, 2008". *The Numismatic Chronicle*, vol. 172 (2012), pp. 335–341.
- Henning, Walter B. 1965. "The Choresmian Documents". Asia Major, vol. 11, no. 2, pp. 166–79.
- Humbach, Helmut. 2011 [1998]. "Epigraphy i. Old Persian and Middle Iranian epigraphy". *Encyclopædia Iranica*, vol. VIII, fasc. 5, pp. 478–488. http://www.iranicaonline.org/articles/epigraphy-i
- Иванчик, А. И.; Лурье, П. Б. [Ivantchik, А. І.; Lurje, Р. В.]. 2013. "Две надписи из Чирик-рабата". *Commentationes Iranicae*. Сборник статей к 90-летию Владимира Ароновича Лившица, Под редакцией С. Р. Тохтасьева и П. Б. Лурье, pp. 286–294. Санкт-Петербург: Нестор-История.
- Livshits, Vladimir A. 1968. "The Khwarezmian Calendar and the Eras of Ancient Chorasmia", *Acta Antiqua Academiae Scientiarum Hungaricae*, vol. 16, pp. 434–446.
- ——. 2003. "Three silver bowls from the Isakovka burial-ground no. 1 with Khwarezmian and Parthian inscriptions". *Ancient Civilizations from Scythia to Siberia*, vol. 9, no. 1–2, pp. 147–172.
- Лурье, П. Б. [Lurje, Pavel B.] 2013. "Несколько неизданных хорезмийских надписей из Ток-Калы". Scripta Antiqua: Вопросы древней истории, филологии, искусства и материальной культуры [= Ancient history, Philology, Arts and Material Culture], Том третий [= Volume three], К юбилею Эдварда Васильевича Ртвеладзе [= E. v. Rtveladze Felicitation volume], с. 729–740. Москва: Собрание.
- ——. forthcoming. "Some New Readings of Chorasmian Inscriptions on Silver Vessels And Their Relevance to Chorasmian Era". *Ancient Civilizations from Scythia to Siberia*.

- MacKenzie, D. N. 2011 [1991]. "Chorasmia iii. The Chorasmian Language." *Encyclopædia Iranica*, vol. V, pp. 517–520. http://www.iranicaonline.org/articles/chorasmia-iii
- Minardi, Michelle 2013. "A Four-Armed Goddess from Ancient Chorasmia: History, Iconography and Style of an Ancient Chorasmian Icon". *Journal of the British Institute of Persian Studies*, vol. 51, no. 1, pp. 111–143.
- Pandey, Anshuman. 2016. "Proposal to encode the Old Sogdian script in Unicode" (L2/16-312R). http://www.unicode.org/L2/L2016/16312r-old-sogdian.pdf
- ——. 2017. "Proposal to encode the Khwarezmian script in Unicode" (L2/17-054R). https://www.unicode.org/L2/L2017/17054r-khwarezmian.pdf
- ——. 2018. "Proposal to encode the Khwarezmian script in Unicode" (L2/18-010R). https://www.unicode.org/L2/L2018/18010r-khwarezmian.pdf
- Sachau, C. Eduard [tr]. 1897. The Chronology of Ancient Nations: An English Version of the Arabic Text of the Athâr-ul-Bâkiya of Albîrûnî, or 'Vestiges of the Past', Collected and Reduced to Writing By the Author in A. H. 39–1, A. D. 1000. London: Oriental Translation Fund of Great Britain & Ireland.
- Skjærvø, Prods Oktor. 1996. "Aramaic Scripts for Iranian Languages." *The World's Writing Systems*, edited by Peter T. Daniels and W. Bright, pp. 515–535. New York and Oxford: Oxford University Press.
- Смирнов, Я. И. [Smirnov, І. А.] 1909. *Восточное серебро* [= *Vostochnoe Serebro*]. Атлас древней серебряной и золотой посуды восточного происхождения, найденной преимущественно в пределах Российской империи.
- Tolstov, S. P. and Livshitz, V. A. 1964. "Decipherment and Interpretation of the Khwarezmian Inscriptions from Tok Kala", *Acta Antiqua Academiae Scientiarum Hungaricae*, vol. 12, pp. 231–251.
- Вайнберг, Бэлла Ильинична [Vainberg, B. I.]. 1977. *Монеты Древнего Хорезма* [= *Monety Drevnego Кhorezma*]. Москва: Наука, Глав. ред. восточной литераы

11 Acknowledgments

I would like to thank Pavel Lurje (State Hermitage Museum, St. Petersburg) for sharing information and materials on the Chorasmian script, and for granting me permission to use a preprint of his forthcoming article on the silver vessels. I am also grateful to Judith Lerner (Institute for Study of the Ancient World, New York University, New York City) for initial discussions on the script. Martin Schwartz (University of California, Berkeley) offered comments on the script name. David Corbett and Eduardo Marin Silva provided feedback on the description of the numerical notation system and general comments.

The project to encode Chorasmian 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.

12 Dedication

This proposal to encode the Chorasmian script in Unicode is dedicated to the patriarch of Chorasmian studies, the late Professor Aronovich Vladimir Livshits, who passed away in 2017. The epitaph given below was designed by the proposal author at the request of Pavel Lurje, who prepared the text, for engraving on the headstone for Livshits (see image and details in fig. 36).

הר בנאי א אוא גל חא נגרבוא בנחאם תמוח אוא מל מחבוב תאו א ומשבונ אבוחל מחבר תאו א ווו וו ונטר ונובן מוב אוגוב

BŠNT ii-'LP X iii iii i YRX' whwmn BYWM 'hwrym ZNH tpnkwk NPŠY 'y wl'tymyr 'hrwn'n-w lybšyš 'rw'n 'D hy'n-' 'y 'rw'n 'L nwš grdm'n myny''ty

In the year 2017, the month Wahman, the day Ahurim This chest is the property of the soul of Vladimir, son of Aharon, Livshits Let their souls remain in eternal paradise

	10FB	10FC	10FD
0	1	Y	
	10FB0	10FC0	
1	*)	
	10FB1	10FC1	
2	١	7	
	10FB2	10FC2	
3	7	W	
	10FB3	10FC3	
4	7	Œ	
	10FB4	10FC4	
5	7	1	
	10FB5	10FC5	
6	1	11	
	10FB6	10FC6	
7	2	111	
	10FB7	10FC7	
8	1	1111	
	10FB8	10FC8	
9	П	>	
	10FB9	10FC9	
Α)	3	
	10FBA	10FCA	
В	フ	3	
	10FBB	10FCB	
С	3		
	10FBC		
D	ヤ		
	10FBD		
Е			
	10FBE		
	.		
F	פ		
	10FBF		

Also known as 'Khwarezmian'.

Letters

10FB0 - CHORASMIAN LETTER ALEPH 10FB1 • CHORASMIAN LETTER SMALL ALEPH 10FB2 **CHORASMIAN LETTER BETH** 10FB3 **7** CHORASMIAN LETTER GIMEL 10FB4 7 CHORASMIAN LETTER DALETH 10FB5 **,** CHORASMIAN LETTER HE 10FB6 I CHORASMIAN LETTER WAW 10FB7 • CHORASMIAN LETTER CURLED WAW 10FB8 1 CHORASMIAN LETTER ZAYIN 10FB9 n CHORASMIAN LETTER HETH 10FBA , CHORASMIAN LETTER YODH 10FBB **7** CHORASMIAN LETTER KAPH 10FBC **3** CHORASMIAN LETTER LAMEDH 10FBD **†** CHORASMIAN LETTER MEM 10FBE | CHORASMIAN LETTER NUN 10FBF ع CHORASMIAN LETTER SAMEKH 10FC0 y CHORASMIAN LETTER AYIN 10FC1 **)** CHORASMIAN LETTER PE 10FC2 7 CHORASMIAN LETTER RESH 10FC3 W CHORASMIAN LETTER SHIN 10FC4 CHORASMIAN LETTER TAW

Numbers

10FC5	1	CHORASMIAN NUMBER ONE
10FC6	11	CHORASMIAN NUMBER TWO
10FC7	111	CHORASMIAN NUMBER THREE
10FC8	1111	CHORASMIAN NUMBER FOUR
10FC9	>	CHORASMIAN NUMBER TEN
10FCA	3	CHORASMIAN NUMBER TWENTY
10FCB	3	CHORASMIAN NUMBER ONE HUNDRED

	Chorasmian	Old Sogdian	Inscriptional Pahlavi	Inscriptional Parthian	Imperial Aramaic
aleph	_	×	П	77	*
beth	د	3	ل	ے	,
gimel	7	и	٦	J	4
daleth	7	(y)	3	Ž	,
he	7	ਮ , 亡	ಆ	H	7)
waw	1	2	2	2	•
zayin	1	J	s	ſ	1
heth	n	N	A.	N.	"
teth	_	_	2	לל	G
yodh	1	4	2	J	4
kaph	フ	У	1	9	y
lamedh	>	7	}	5	L
mem	þ	%	ঠ	Я	3
nun	1	J	١	ユ	5
samekh	ھ	n	n	מ	,
ayin	¥	(۲) کی	(2)	5	v
pe)	9	٩	>	,
sadhe	_	2	٤	_^	r
qoph	_	_	(প্র)	מ	7
resh	7	У	(2)	9	•
shin	W	n	22	¥	v
taw	σ	מ	2	Эn	٢

Table 1: Comparison of Chorasmian letters with those in Unicode blocks for related Iranian scripts and Aramaic. Parenthesis indicate that a letter has been unified with another in the respective encoding. In Inscriptional Pahlavi, *ayin* and *resh* are unified with *waw*, and *qoph* with *mem*. For Old Sogdian, *daleth* and regular *ayin* are unified with *resh*.

	Chorasmian	Old Sogdian	Inscriptional Pahlavi	Inscriptional Parthian	Imperial Aramaic
ONE	1	J	1	J	1
TWO	11	n	n	IJ	V
THREE	111	m	m	JJJ	\//
FOUR	1111	mn	m	וווו	_
FIVE	_	mn	_	_	_
TEN	>	2	٦	٧	~
TWENTY	3	3	3	9	3,
THIRTY	_	ŧ	_	_	_
ONE HUNDRED	3	4	ģ	ح	ئ
ONE THOUSAND	_	_	f	3	X
TEN THOUSAND	_	_	_	_	**
ONE HALF	_	p	_	_	

Table 2: Comparison of Chorasmian numerical signs with those in Unicode blocks for related Iranian scripts and Aramaic.

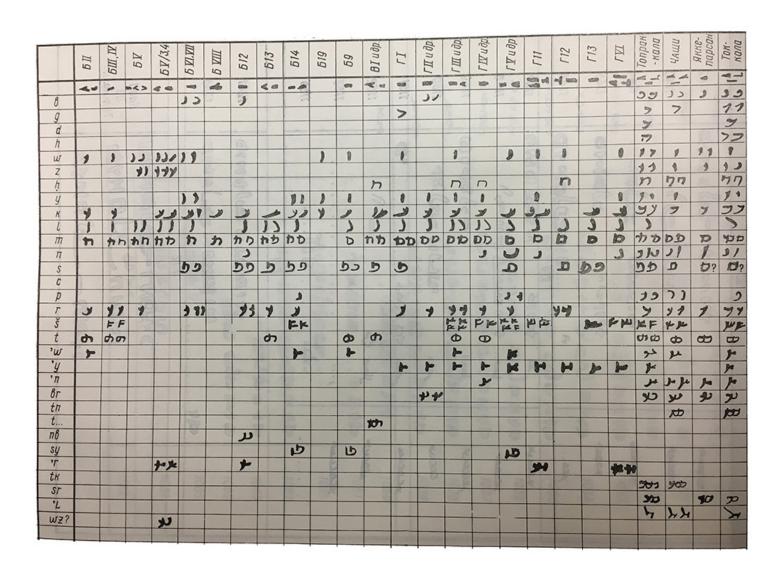


Figure 1: Inventory of characters on Chorasmian coins (БІІ–ГVІ), Toprak Kala (Топрак-кала), Yakke Parsan (Якке парсан), Tok Kala (Ток-кала) (from Vainberg 1977: Table 8).

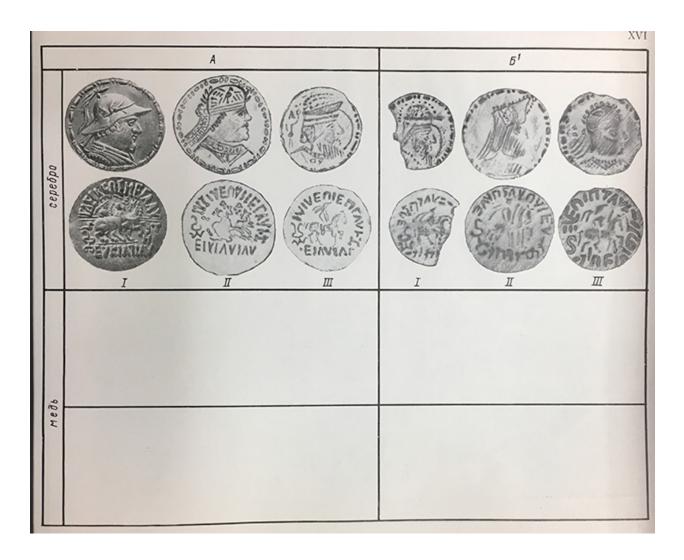


Figure 2: Chorasmian coins (from Vainberg 1977: Table 16).

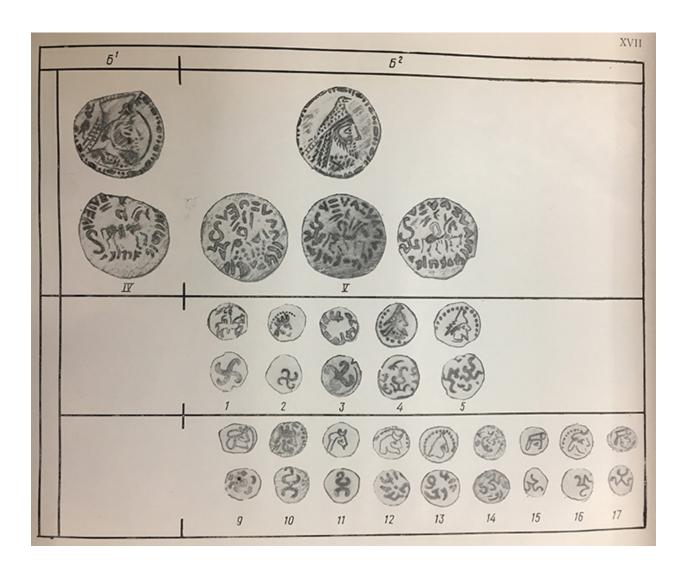


Figure 3: Chorasmian coins (from Vainberg 1977: Table 17).

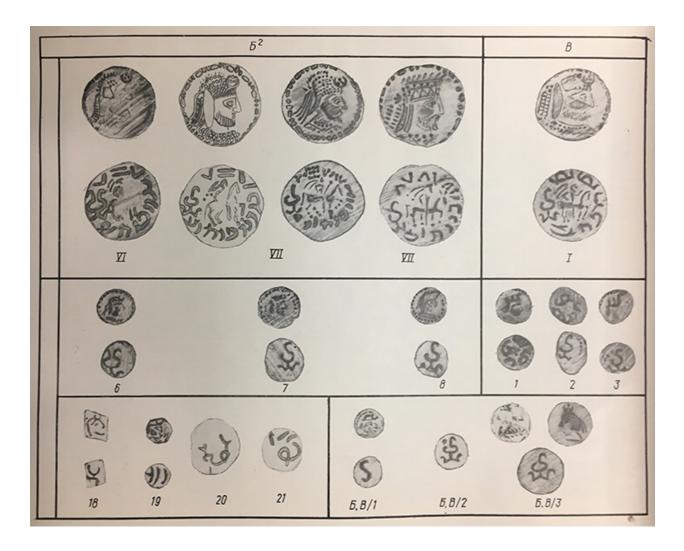


Figure 4: Chorasmian coins (from Vainberg 1977: Table 18).

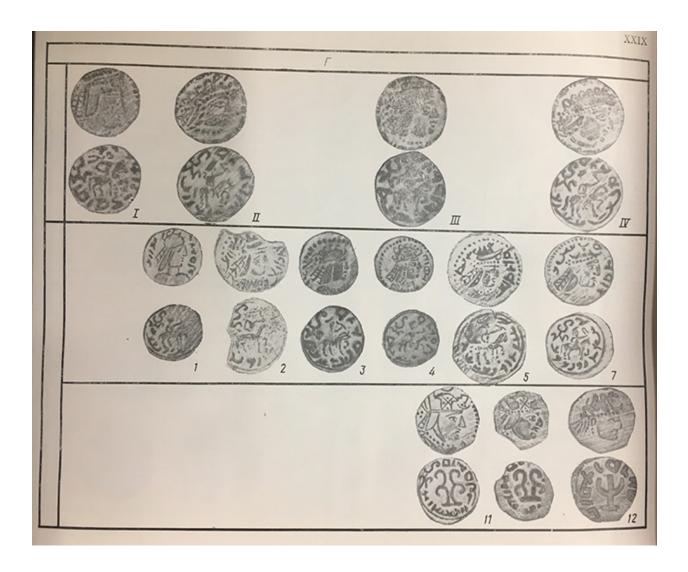


Figure 5: Chorasmian coins (from Vainberg 1977: Table 19).

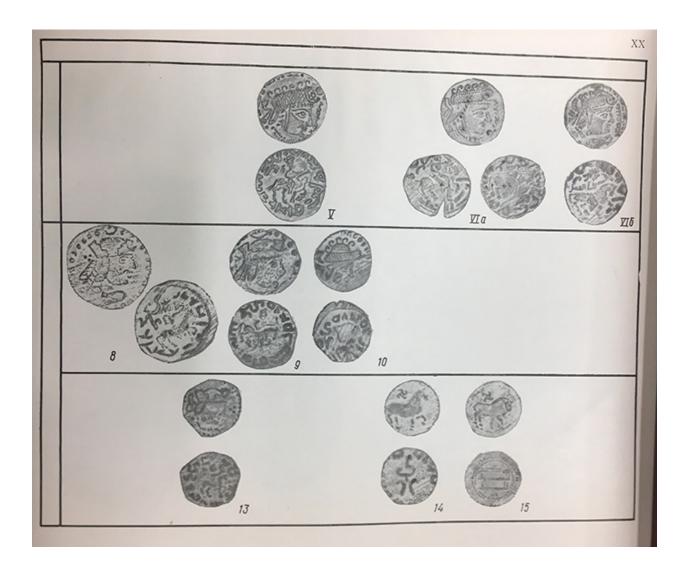


Figure 6: Chorasmian coins (from Vainberg 1977: Table 20).

W по каталогу		W по каталогу		
5° ¥/3		<i>5</i> ² <i>Y</i>		
का) विकास	130	~ 7 1 PS 9 9 9 9 9 P	209	
D= 15	134	21 to 1211 12	194	
的品品	135	52 VII		
3 th 50	137	all states with	207	
0000	137	6 ² VIII		
न्यों के रीन का रीत	138	3/20 100	218	
-414 JUBY	139	<i>5</i> ² <i>9</i>	1 198	
यो क्राक्रक्री	143	>1-1D	227	
N O SI	100	٥٠٠٠٠٠١١١١١١١١١١١١١١١١١١١١١١١١١١١١١١١١	228	
श्रीकार-कार्य	144	00060	234	
প্রাক্তির প্রথা	148	0 - 63	235	
The same of the sa		arb	236	
अधि भी व्यक्ति	151	OH6	237	
D= \$79	154	→ ∞ → ⊕ 5 ² 12	249	
30.000	CT CILCON	A AND DESCRIPTION OF THE PARTY	-	
20 co el e fe	156	פינות מלתב	355	
मिक कि कि वि	158	- ११ ० १८०	356	
111		day le	357	
δ ² Ψ/4		and warden	361	
SOUTH SO A COLOR	164	काष्य के भारत	362	
< 10 MD 527	174	रामिक लिए	367	
-2418 24800	10	والمراء والد	368	
40)to 90th 30	11	201107205	370	
The second secon	and and	200000	1	
22) 4348 M	12	פינת מו מוני	378	
	and the state of t	Tops Wilding		

Figure 7: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 1).

TW	TO KOTOAOZU		N по каталогу	
W по каталогу		11 114	460	
5 ² 13		الراء الله	400	
الاصطلا الد	381	. N > ED LAD 15	461	
win ord	384	व्यवसा क्षाट=	463	
051	387	= 100 BOMDe	470	
A Dog	398	المعادية والعامل	471	
		5 ² 19		
\+ 00 5=1	403	200	674	
l-G-wales	421	400	677	
with on by s)	424	200	678	
5 ² 14		500	684	
JOE HAND	427	BI	Alla I	
more o	428	स्मिनार्टिक	769	
1D 41 LLD	429	心のからからからからから	770	
מושפתויום פרנבי	430	RT/1		
990		Do the land	771	
a PD	431	Son per par	777	
CAD EA	435	S man gh	791	
वान्त्र त	440	Dram	794	
200 F 100	441	William Colors		
الما عرصه ما هرا	443		Con Commence	
NO E 10	444	ゆかつつりもいりも	919	
当時下	445	ゆうかかまからっ~	920	
BL5 p19	453	75000500	100	

Figure 8: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 2).

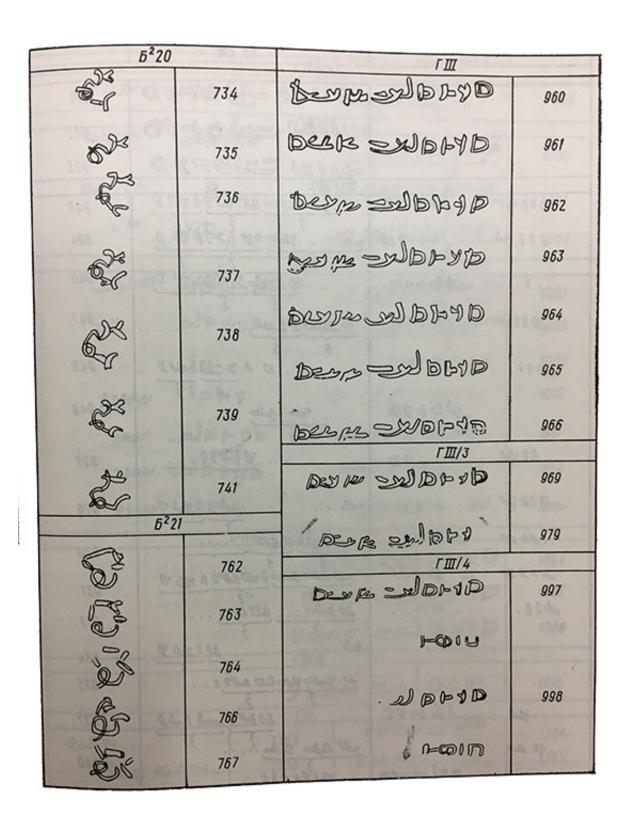


Figure 9: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 3).

-	cm.	οδ. cm.	60
11.	, III.	TIL BA COMPO	921
	Call Call Care	والمطورات والمدادة	921
098	00000	1 - 200	922
		والاعا ها لسع مدروس	322
190		- DAM (1)	923
		SIN ENDERND	
		ΓШ1 A (\$1.1 A \$1.1) \$2	945
300 p	@ C & C & C & C & C & C & C & C & C & C	(द ति । कि कि कि कि कि कि । । । । । । । । ।	945
		त का विक विक विक विकास	924
80000	6000	20 BIR 1000 D	28
596		7	946
	-10 0 0 0 cm	10 END D-000 20 2	340
	ED & C.II.	5 4 1	0/7
200000	no po no	2 3120 0 2 30	947
2000		3	
201110	1000	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	948
\$00		2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	
1		4 1	0000
	@ الما ه الور	2 22	949
	3740000	and the same	
00.40	. F. T.	9150	925
00 49	10		223
		2 7	
20000		الم	926
		2	
0000	ADPT	2000 Desp 072	930
	a. ITTI	6 5 000	930
-0000	- DODD	1000 100 00 00 00 00 00 00 00 00 00 00 0	071
103 92	MINNE TO LET	3 2 1	931
200.	₽₽	coa	933
	UIGH!	3 2	
		B B CB CB	934
		00000000000000000!	9.0
		- Was 12 80 11	935
	1 b 000	2000	A STATE OF THE STA
-		0000 2000	937
صت ٥٥	Comment of	200 pls ? 1	200
	عرام م العد	3	940
	المراح ها محمد	martens	000
		6 5	942

Figure 10: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 4).

л. ст.	11	0δ. cm.	
	ΓΙΙ/2		
اطالع مواالا	>M	35-60\$=	959
1500	ΓIII/5	and the	VO.E.T
all a ma	-AD	- massell	1000
De	Se ve	22 20 20 0 bors bo	1001
000 3000.	Call was	~ 2000 210 ars 4.	1002
7 0 0	ΓIV/6		1110
20000 0000		क्र यक्ष विष्ठिक कि	1007
٥٩ لعر ٥٧	-DD	Dan Ere	1008
ساله .	and the second	ರಾಖ್ಯಾರಾ	1009
ع ماه معدود	D	5 11 Des 20 10 11 Co	1010
وهلم سو	-DD	M O POSSOR	1011
ه هالت معدده	- AP	MAN TO MAN	1012
some enter	D]@-60 ts M	1015
1001	rix		7100
1001 1001	CONCESSOR CONCESSOR	هم موالات معد	1004
1801 of t 49 di		20 all a 10	1005
		विषय किल्ल	1006
500 EUG	FIY/8	en endors.	1018
1805	N L	D D D D D D D D D D D D D D D D D D D	7010
9.100	1365	वळात्रवाव	1364
120 - 20 - 20 - 20 - 20 - 20 - 20 - 20 -	1376	2000	1384
-00 AB	1371	ल्याय क्षाय क्षाय व्याय	1372
= 10 A	1371	٥٥٥	1366

Figure 11: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 5).

	ſ ¥	
عدماها	वाध्य व्यक्ति	1030
Movare	का भ्यार किल्या स्वास्त्र ति	1031
Appell	धार शव किल वर्ग भारत करा भारत	1032
-127001	य महरा श्री कार नि	1033
W X MON	والمع والمعدد الخالف المالية ا	1034
crovar	विषय विषय व्याधार १११०	1035
30100	वर्ष वर्षा वर्ष का मान	1036
שמיסכץ	Chamber and the	1037
٠ ٥ ١٠٠٠	वश्य वर्षाः व्यायक्षाक	1038
په ۲۵۷مور سالم	वारत व कि ट्रायर अपा	1039
ж	१४७ व्याधिक व्याधाः	1060
THE RESERVE TO SERVE	4124 B21514	1061
6101	anether	1051
1001 CT 8 =0	at a a les and that	1087
0001 25-3.	1- 27/16 13 4 16	1088
2001 0 0.00	व राज्य व्योष्ट व्यापर	1089
9101 1010	OND 5	-1085
-	೨೪೬೪೬೩	1081
1000 000	Γ <u>Ψ/10</u>	
and a second	-C 700 CA C C	1135

Figure 12: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 6).

	ΓΣΙα	10000
MUENTE	ara ala-massoa	1142
yeacus	4 mm 00 m	1143
усоска	लहारिक हिल्ला	1144
אממטא	विश्ववादि १८५१ विश्ववादि	1145
уса	a 200 = 10	1146
Arcorix	क काराह्य प्राथम का काराह्य के	1147
услагу	ارده ۱۱۱ م معدد ۱ معدد ۱ معدد	1148
reacense	ाष्ट्र हाउन्नविश्लि ् । । । । । । । । । । । । । । । । । ।	1149
cocnx	الله ١٥٠٠ مم	1150
receu.	العمر المعمر المعمر	1151
*TOEW	ا ا ا ا ا ا ا ا ا ا ا ا ا ا ا ا ا ا ا	1152
yencum.	· ····	1153
	500	1155
	rii ang	
2 4 5	41 Em 10 20 pm 10	1241
	Calendo Como 1000	1258
	Γ12	
3	DENGE TIERS	Клад из
	arala Mari	Ток-коль
	ا ماد الله الله المام الله المام الله المام الله الله	

Figure 13: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 7).

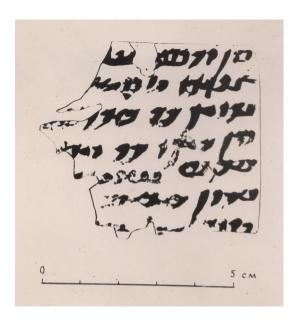


Figure 14: Fragment of a leather document with Chorasmian inscription from Yakke Parsan. Image courtesy of Lurje.

Letter	initial	internal	final	independent	left join	remarks
,	> 5 5	502 1		_0	yes	
b	و کو در				yes	
g	フラテァ	フ ŷ	1		yes	
d			D	10	no	
h			87	>> >	no	
W)	10	no	
z			C	001	no	
X			Ė	にに	no	
t	# 100	^ന - <u>ആ</u>	60 1	(Ga)	yes	joins without horizontal line
У			1 1 16) D P	no	
k	1 5	2	ソソ0	>> 2 /s	yes no?	
[1]			86			ideogram 'L
m			12 1	d a	no	
n	1 9 1	丁辽 夺	S J		yes	
S	9	合回の			yes	
[,]				D A	no	ideograms 'L 'BDt
p		a a	\frac{\(\)}{\(\)}	د ه	yes/no	
r			70 1 1	10)	no	
š			m &	12 V	no	
ligatu	res tn JB 72	'n 🔌	- Jo k'n	bg -	1) by 1/2	

Figure 15: Nominal and positional forms of letters in silver vessel inscriptions (from Lurje 2017).

्रक्तराहित वा प्राव्यक्तराह न्याया निवास कार्य वित्रकार हित्य वित्रकार कार्य वित्रकार कार्य वित्रकार कार्य वित्रकार वित्

Pl. 1:a Choresmian No. 1: A.D. 658. Inscription from silver phiale in the British Museum, Smirnov, VS, pl. XIX:43.



Trustees of the British Museum.

Figure 16: Silver vessel #1: 658 CE (from Azarpay 1969: Plate 1:a, b). Silver philae in the British museum. Original from Smirnov 1909, plate XIX: 43.

- יין טופן לתישמי - ם - נושאראשי ה ביר ל לדסה א היה ותנונל אל היוחאוף מין שווו וילאויך ווח-

Pl. 3:a Choresmian No. 2: A.D. 538 (probably 638). Inscription from silver phiale in the Hermitage Museum, Leningrad, Smirnov, VS, pl. XIX: 42.

BŠNT 3 2 100 20 20 20 10 YRX ' 'try BYWM bgy 'pbntn y' MN bwnšrgk-š x(r)k'n 'L 'tršnky bg ZNH ZWZN 20 10 'špynšwk



Pl. 3:b Choresmian No. 2: A.D. 538 (probably 638). Silver phiale in the Hermitage Museum, Leningrad, see pl. 3:a. Diam. 10.6 cm.



Pl. 3:c Choremian No. 2: A.D. 538 (probably 638). Silver phiale in the Hermitage Museum, Leningrad, see pl. 3:b. Smirnov, V-S, pl. XVIII:42.

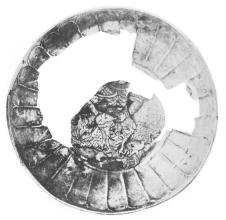
Figure 17: Silver bowl #2: either 538 or 638 CE (from Azarpay 1969: Plate 3:a, b, c). Silver philae in the Hermitage Museum (St. Petersburg). Original from Smirnov 1909, plate XIX:42 and XVIII: 42. Transliteration from Lurje (forthcoming).



Pl. 5:a Choresmian No. 3: inscription from silver phiale in the Hermitage Museum, Leningrad, Smirnov, VS, pl. XIX:44.



Pl. 5:b Choresmian No. 3: silver phiale in the Hermitage Museum, Leningrad, see pl. 5:a, diam. 12.5 cm. Smirnov, VS, pl. XVIII: 44.



Pl. 5: c Choresmian No. 3: silver phiale in the Hermitage Museum, Leningrad, see pl. 5: b.

Figure 18: Silver vessel #3 (from Azarpay 1969: Plate 5:a, b, c). Silver philae in the Hermitage Museum (St. Petersburg). Original from Smirnov 1909, plate XIX: 44 and XVIII: 44.



Pl. 8:a

Choresmian No. 4: inscription from silver phiale in the Hermitage

Museum, Leningrad, Smirnov, VS, pl. XIX:45.



Pl. 8:b *Choresmian No. 4:* silver phiale in the Hermitage Museum, Leningrad, see pl. 8:a. Diam. 12.7 cm.



Pl. 8:c Choresmian No. 4: silver phiale in the Hermitage Museum, Leningrad, see pl. 8:b. Smirnov, VS, pl. XVIII:45.

Figure 19: Silver vessel #4 (from Azarpay 1969: Plate 8:a, b, c). Silver philae in the Hermitage Museum (St. Petersburg). Original from Smirnov 1909, plate XIX: 45 and XVIII: 45.

リリーカーにはいけいはいとしてのとしてはないにはしいしてはいい

אנשם ו- לן ודלוונשת אשפר ול דעם ול עדש ולכו ווון- 3 ווו וו

'pbntn y' MN wrmwzbntk 'r'škrk 'L (g)nyt 'L byrty zmhy ZWZN-' 20 20 3 2



Pl. 11:a,b,c Choresmian No. 5: silver phiale in the Hermitage Museum, Leningrad, Smirnov, VS, pl. XIX:47. Diam. 13 cm.

Figure 20: Silver vessel #5 (from Azarpay 1969: Plate 11:a, b, c). Silver philae in the Hermitage Museum (St. Petersburg). Original from Smirnov 1909, plate XIX: 47. Transliteration from Lurje (forthcoming).

רש חומל וווו 2333 וווווו

gty (xwpsk | xw ksp) ZWZN-, 4 10 20 20 20 7



Figure 21: Silver vessel #6. Original from Smirnov 1909, plate L: 84. Transliteration from Lurje (forthcoming).



וטרטל אוגאו נוא אתם

wbrn'k šyr'nw hy'n 'BDT



Pl. 9:a Choresmian 7: silver phiale in the Hermitage Museum, Leningrad, Smirnov, VS, pl. XX:46.



Pl. 9:b Choresmian No. 7: silver phiale in the Hermitage Museum, Leningrad, see pl. 9:a.

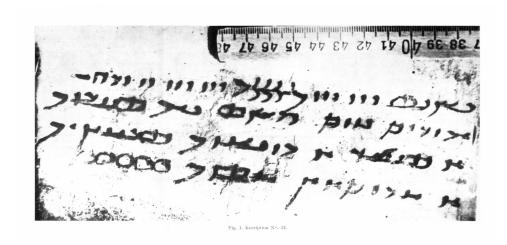
Figure 22: Silver bowl #7 (from Azarpay 1969: Plate 9:a, b, c). Silver philae in the Hermitage Museum (St. Petersburg). Original from Smirnov 1909, plate XX: 46. Transliteration from Lurje (forthcoming).





Pl. 10:a, b, c Choresmian No. 8: silver phiale in the Hermitage Museum, Leningrad, Smirnov, VS, pl. CXIV: 286.

Figure 23: Silver bowl #8 (from Azarpay 1969: Plate 10:a, b, c). Silver philae in the Hermitage Museum (St. Petersburg). Original from Smirnov 1909, plate CXIV: 286.



א ארוקאא אשב ٥٥٥٥ א שמא א נושל שבמאיב אוגול מול נואם מל שהוב מאש ווו ווו \$\$\$\$? איים ווו ווו \$\$\$\$?

Tolstov and Livshitz 1964

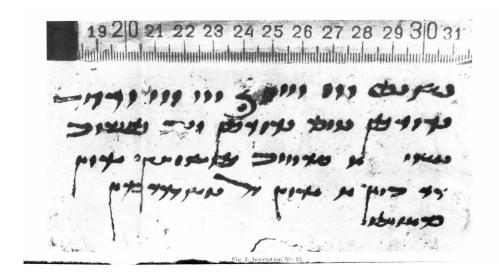
BŠNT III III C XX XX XX X III III II YRḤ'
'hwrym BYWM gwšt ZNH tnbryk
'y tnb'r 'y hwnšk (?) t'b'n'n'k (?)
'y 'yrw|zm'w'n 'ztyk

Year 678. Month Ahurem, day Gost. This ossuary contains the body of hwnsk t'b'n'n'k, son of 'yrw|zm'w'n

Reanalysis based upon Henning 1965

BŠNT III III C XX XX XX XIII III II YRḤ '
'hwrym BYWM γwšt ZNH tpnkwk
'y tnβ'r 'y hwnšk (?) t'β'n'n-'-k (?)
'y 'yrw|zm'w'n 'ztyk 0000

Figure 24: Tok Kala no. 52, ossuary inscription (from Tolstov and Livshitz 1964: Figure 1). Representation in Chorasmian script based upon the reanalysis.



קשמש האו א קבוווב שאומאל אבוא האו א קבוווב שאומאל אבוא האם חון ווו ו € ווו ווו וב⊔ר האש ווו וו €

Tolstov and Livshitz 1964

BŠNT III III I C III III YRḤ' brwrtn BYWM brwrtn ZN[H] tnbryk nwšy (?) 'y srywyk tyšy'n'ny 'rw'n GD kw'n[y] 'y 'rw'n 'L nwš grdm'n pr'ny'ty

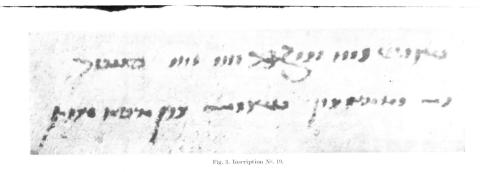
Year 705. Month Rawacina, day Rawacina. This is the ossuary of srywyk [son] of tysy'n, soul [whose] [possesses] kayan farrah. Soul [his] may be sent to the beautiful Paradise.

Henning 1965

BŠNT vii C vi YRḤ² βrwrtn BYWM βrwrtn ZNH tpnkwk NPŠY 'y srwywk tyšy'n'n-w'rw'n 'D hw'n-' 'y 'rw'n 'L nwš γrδm'n m'ny'(')ty

In the year 706, on the 19th day of the first month. This chest is the property of the soul of $Sraw-y\bar{o}k$, the son of $Ti\check{s}-y\bar{a}n$. May their souls rest in the eternal Paradise.

Figure 25: Tok Kala no. 25, ossuary inscription (image from Tolstov and Livshitz 1964: Figure 2). Words highlighted in red in the revised interpretation by Henning (1965) indicate differences from Tolstov and Livshitz. The representation in Chorasmian script is based upon the revised reading.



ני ואימגון חאני בגל אמא אוא מאתם וווו ווו נאל וווו וווו מחבוב

Tolstov and Livshitz 1964

BŠNT IIII III C XX X IIII IIII tnbryk y' w'z'sw|ydyn nwšy' ?grn 'rt'w 'rw'n

Year 738. This is the ossuary of w'z'swdyn (?) [May] in the beautiful Paradise [be sent his] true soul.

Reanalysis based upon Henning 1965

BŠNT IIII III C XX X IIII IIII tpnkwk y' w'z'sw|ydyn NPŠY' ?grn 'rt'w 'rw'n

Figure 26: Tok Kala no. 19, ossuary inscription (from Tolstov and Livshitz 1964: Figure 3). Transliterations from same; but may be erroneous or outdated.

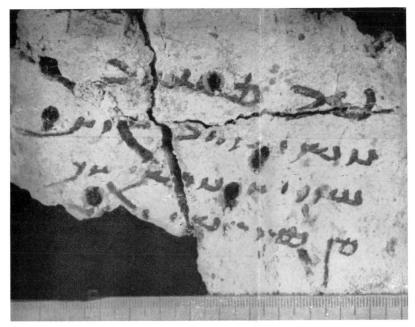


Fig. 4. Inscription No. 8.

ZNH tn[b]'r¹yk
'(?)ynšy 'šh¹k '.wn'y¹.
...'y nykšy ? 'YK
MN ty'zhwndy 'L ... [

This is the ossuary of woman (? shk, daughter of '.w ... May [soul her be sent] from the [world] of full danger to (the world of safety?).

Figure 27: Tok Kala no. 8, ossuary inscription (from Tolstov and Livshitz 1964: Figure 4). Transliterations from same; but may be erroneous or outdated.

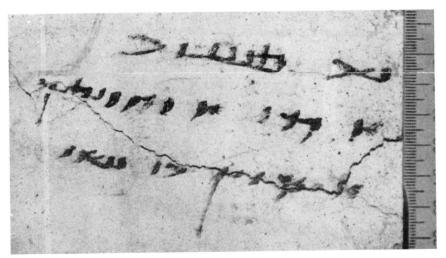


Fig. 5. Inscription No. 69.

א הגוא נו תאנ א בגו א וטותסב תל שתבוב

Tolstov and Livshitz 1964

ZNH tnbryk 'y gry 'y wḥwntk 'y 'rw'n kw nwšy

This ossuary contains the body of wnwnxk Soul [his may be sent] to the beautiful [Paradise].

Reanalysis based upon Henning 1965

ZNH tpnkwk
'y yry 'y wḥwnt'n-w
'y 'rw'n hw NPŠY

Figure 28: Tok Kala no. 69, ossuary inscription (from Tolstov and Livshitz 1964: Figure 5). Representation in Chorasmian script based upon reanalysis.

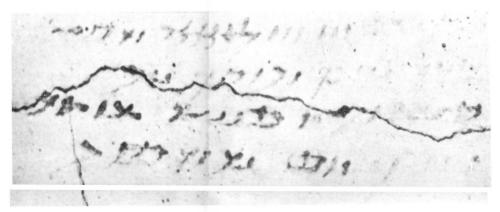


Fig. 6. Inscription No. 39.

「BŠNT¹ III III C XX XX XX X YRḤ' m^rtry?¹ 「BYW¹M whwmn ZNH tn^rbr¹yk 'y tnb'r 'rw'zd w ... n^ry¹ zyt brwrtyk

Year 690, month of Miri, day of Ahumen. This ossuary holds the body of 'rw'zd w...n, son of Hravardik.

Figure 29: Tok Kala no. 39, ossuary inscription (from Tolstov and Livshitz 1964: Figure 6). Transliterations from same; but may be erroneous or outdated.

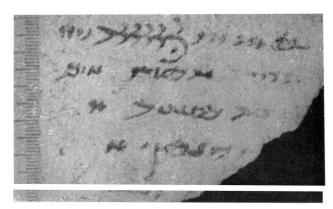


Fig. 7. Inscription No. 12

קרטאל א הרל שהלוב א הל שהלוב א הלר אשוא מול האתם ווו ווו \$

^гBйNT III III С XX XX XX XX X IIII YRḤ' 'rtwyš BYWM]. ZNH tnbryk 'y]. s|ḥnt'ny 'y]. y'

Year 694, month of Ardwis, day [] This ossuary ... [of son] of ...s|hnt...

Figure 30: Tok Kala no. 12, ossuary inscription (from Tolstov and Livshitz 1964: Figure 7). Transliterations from same; but may be erroneous or outdated.

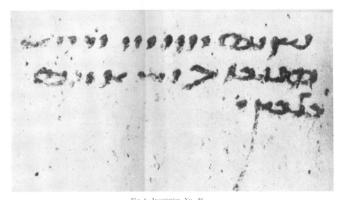


Fig. 8. Inscription No. 26.

הראל. מודוב ו− א וגט האתם וווו ווו נגטר

Tolstov and Livshitz 1964

BŠNT III IIII YRḤ' tnbryk y' 'y wrt k'k'ny Reanalysis based upon Henning 1965

BŠNT III IIII YRḤʾ tpnkwk yʾ ʾy wrt kʾkʾn-w

Year 7[00]. Month. This is the ossuary of wrt, [of son] of k'k.

Figure 31: Tok Kala no. 26, ossuary inscription (from Tolstov and Livshitz 1964: Figure 8). Representation in Chorasmian script based upon reanalysis by the proposal author.



Fig. 9. Inscription No. 21.

BŠN'T' [
YR'H'' [
whwmn' [ZHN tnbryk 'y ḥw'r'n [
wḥnwy(?) [

Figure 32: Tok Kala no. 21, ossuary inscription (from Tolstov and Livshitz 1964: Figure 9). Transliterations from same; but may be erroneous or outdated.

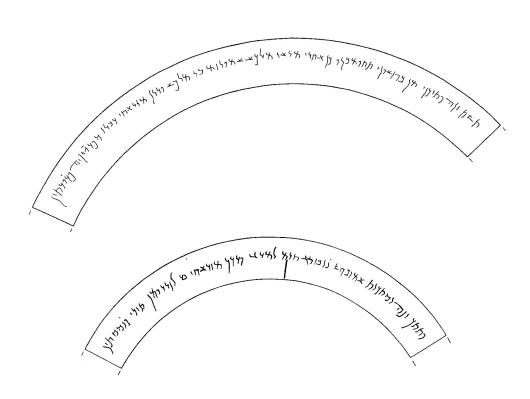


Figure 33: Archaic Chorasmian inscriptions on silver bowls no. 1 (top) and no. 2 (bottom) from Isakovka from the Achaemenid period (from Livshits 2003: 152, 163). This type of script is out of scope for the present encoding, and may be unified with Imperial Aramaic.



Figure 34: A lapidary Chorasmian inscription from Chirik-rabat, likely dated between the 2nd and 5th century BCE (from Ivantchik and Lurje 2013: 286). The likely reading is *tyrybwdy*. This type of script is out of scope for the present encoding, and may be unified with Imperial Aramaic.

Имп	ерский ара	амейский	Парфя	янский	Бактр. (?)		Согди	йский		Хоре	змийский					
	Бактрия 1)	Арахосия2)	Ниса 3)	Авроман 4)	Ай-Ханум 5)	Афрасиаб 6) Куль-тобе 7)	Ст. письма 8) Шатиал 9)	Айбуйир 10) Исаковка 11)	Бурлы-кала 12)	Калалы-гыр 13)) Кой-Крылган 14) Чирик-рабат	Ш
,	4 4	FKK	***	× × × × (-)	44	X	×	☆	212		* * *	-K -A	4 4			,
b	9 7	75	2 , 2 , 3 5				7	ב ב ב	7 7 7		5 6 6 G	.5 5	3		ر د	Ь
g	4	1 1	× > 1	77	7		2	スカる	4 4		1 /	11				g
d	7 7	47	7 7	, , ,	y_y.		7		-		7 4 4	> >	7		ς	d
h	71	117	21 74 H	20 25			マナス				Mr. w		-1)	,		h
w	1 1	737	, , ,	9 , 9 9(-w)	•	3	2 2	219	2 2 2		7771	33.3	77)	w
z	11	11	1 1 3	, ,	•		1	٠ ٠	ノリリ		31 1 4	١	11			z
ķ	11 11	カカガ	KK KK KK	***			1-11	NN	<i>N N</i>		D? A +	* 17 27 21	1111			þ
t	6	V	6 U								· · / · /					ţ
у	4 4 1	110	7 1 4	111	,	4	くく	33		4	5 1 5	7 1 14	3.4	•		у
k	1 1	777	ソ フ ブ フ ブ (k-)	9 9 9	4(-k)4-)-	,	١, ١	ソソ	y j j		_1 977_	7777	355			k
1	4	66	1116	111	1-1-7-		_2	27	7 7		[4 4 7]		(-160		1
m	77 41	カグ	% » n	מי פיצי ש	%	,	- 47	يع.ولا	カカ	4	7 7 8	W ==	(h	71	<u> </u>	m
n	1115	1 1	/ / (-n)) 1 (-n)	77		7	ر د	3 1	_ !_	1115	111	1 5			n
s	ħ	3 4 5	מ לל ה	ח מ	4	/	12 1.	אל בע	HH		μ.; , ν	Dy7_	מ מ כ	*		s
٠	v	UV	y y y y		\		7		- 1		4 5		7			ŕ
p	1 1	27	* A A A	10 1			22	و	9))		3 🔰]_		מוכ			р
ş	ħ		•),	-3	7 7							ş
q	ħ	7 70	ות חב ת יו	y						ļ. —	7 7		I			q
r	7 7	- /	7 7 7 7	9 9 9 7(-1)	9,	7	7	272	9 + >	4	7 4	4	779	<i>ነ</i>	<u> </u>	r
š	r	マヤ	V Y &	y y D	E	3	بالم	سو سد	74 14		FE	* \$	上下	Lns-		š
t	P	hhh	トックック(-t)	10 P (-t)	ין ינ	「ルノ」	در در	310	7		h h h	7 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	חתת	ンソノ	7_	Įt į

- 1 Shaked, Naveh 2012, Doc. A1;
- 2–5 В.А. Лившиц по $\it Pacmopryeea$, $\it Moлчанова$ 1981;
- 6 Grenet 2006;
- 7 Sims-Williams, Grenet, 2006;
- 8, 9 *Исхаков* 2008. Табл. XI, XV;

- 10 В.А. Лившиц по Мамбетуллаев 1979;
- 11 Лившиц 2002;
- 12 Лившиц, Мамбетуллаев 1985;
- 13 Лившиц 2004;
- 14 Толстов, Вайнберг 1967. С. 220.

Илл. 2. Знаки чирик-рабатской надписи в сравнении с другими письменностями древней Средней Азии

Figure 35: Comparison of early Iranian lapidary script types derived from Imperial Aramaic (from Ivantchik and Lurje 2013: 290).



The concept and realization of the monument: Boris Elkin

Architecture: Vasilisa Belova Text of inscription: Pavel Lurje

Typesetting of inscription: Anshuman Pandey Calculation of the date: François de Blois

Figure 36: Headstone of Professor Aronovich Vladimir Livshits inscribed in Chorasmian.

ISO/IEC JTC 1/SC 2/WG 2 PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 106461

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from http://std.dkuug.dk/JTC1/SC2/WG2/docs/principles.html for guidelines and details before filling this form.

Please ensure you are using the latest Form from http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html.

See also http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html for latest Roadmaps.

A. Administrative

1. Title: Proposal to encode the Chorasmian script in Unico	ode
2. Requester's name: Anshuman Pandey <pandey @umich.edu=""></pandey>	
3. Requester type (Member body/Liaison/Individual contribution): Expert contribution	
4. Submission date: 2018-07-2	26
5. Requester's reference (if applicable):	
6. Choose one of the following:	V
This is a complete proposal: (or) More information will be provided later:	Yes
B. Technical – General	
1. Choose one of the following:	V
a. This proposal is for a new script (set of characters):	Yes
Proposed name of script: Chorasmian b. The proposal is for addition of character(s) to an existing block:	
Name of the existing block:	
2. Number of characters in proposal:	28
3. Proposed category (select one from below - see section 2.2 of P&P document):	
A-Contemporary B.1-Specialized (small collection) B.2-Specialized (large of	collection)
C-Major extinct D-Attested extinct E-Minor extinct	
F-Archaic Hieroglyphic or Ideographic G-Obscure or questionable usa	ige symbols
4. Is a repertoire including character names provided?	Yes
 a. If YES, are the names in accordance with the "character naming guidelines" 	
in Annex L of P&P document?	Yes
b. Are the character shapes attached in a legible form suitable for review?	Yes
5. Fonts related:	
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for pu	blishing the
standard?	
Anshuman Pandey b. Identify the party granting a license for use of the font by the editors (include address, e	moil ftp cito etc.):
Anshuman Pandey	-maii, rip-site, etc.).
6. References:	
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	Yes
b. Are published examples of use (such as samples from newspapers, magazines, or other	er sources)
of proposed characters attached?	,
7. Special encoding issues:	
Does the proposal address other aspects of character data processing (if applicable) such	as input,
presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)	
8. Additional Information:	
Submitters are invited to provide any additional information about Properties of the proposed Ch	aracter(s) or Script
that will assist in correct understanding of and correct linguistic processing of the proposed char-	
Examples of such properties are: Casing information, Numeric information, Currency information	
information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Direction	
Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicod	
related information. See the Unicode standard at http://www.unicode.org for such information or	
see Unicode Character Database (http://www.unicode.org/reports/tr44/) and associated Unicode for information needed for consideration by the Unicode Technical Committee for inclusion in the	

¹ Form number: N4502-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

C. Technical - Justification

Has this proposal for addition of character(s) been submitted before?	No
If YES explain	
2. Has contact been made to members of the user community (for example: National Body,	V
user groups of the script or characters, other experts, etc.)?	Yes
If YES, with whom? Pavel Lurje <pavlvslvra@gmail.com></pavlvslvra@gmail.com>	
If YES, available relevant documents:	
Information on the user community for the proposed characters (for example:	
size, demographics, information technology use, or publishing use) is included?	Yes
Reference: See text of proposal	
4. The context of use for the proposed characters (type of use; common or rare)	Common
Reference: See text of proposal	
5. Are the proposed characters in current use by the user community?	Yes;
If YES, where? Reference: Currently used by scholars of Iranian and Central Asia	n studies
6. After giving due considerations to the principles in the P&P document must the proposed character	
in the BMP?	N/A
If YES, is a rationale provided?	
If YES, reference:	
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered	d)? Yes
Can any of the proposed characters be considered a presentation form of an existing	
character or character sequence?	No
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
S. Can any of the proposed characters be encoded using a composed character sequence of either	
existing characters or other proposed characters?	No
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
10. Can any of the proposed character(s) be considered to be similar (in appearance or function)	No
to, or could be confused with, an existing character?	740
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
11. Does the proposal include use of combining characters and/or use of composite sequences?	No
If YES, is a rationale for such use provided?	
If YES, reference:	
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provid	led? <i>N/A</i>
If YES, reference:	
12. Does the proposal contain characters with any special properties such as	
control function or similar semantics?	No
If YES, describe in detail (include attachment if necessary)	
13. Does the proposal contain any Ideographic compatibility characters?	No
If YES, are the equivalent corresponding unified ideographic characters identified?	
If YES, reference:	
·	