Title
Proposal to encode the Chorasmian script in Unicode

Author
Pandey, Anshuman

Publication Date
2019-07-26

Peer reviewed

L2/18-164R2
2018-07-26

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 proposed script was used between the 2nd century BCE and 8th-9th century CE for writing Chorasmian (ISO 639-3: xco), a now-extinct Eastern Iranian language. The script and language were used in Chorasmia, a region in Central Asia that was situated at the delta of the Oxus (Amu Darya) river, which is today spread across Uzbekistan, Kazakhstan, and Turkmenistan. The name of the territory was first mentioned in the *Avesta* as 𐬓𐬀𐬌𐬭𐬌𐬰𐬆𐬨 *hvâirizem* (*Yašt* 10.14). It occurs as 𐎢𐎺𐎠𐎼𐏀𐎷𐎡𐏁 *uvârazmiš* in the ‘Daiva Inscription’ (XPh, lines 21–22) of Xerxes I (486–465) at Persepolis. It is called خوارزم *xvārazm* in classical Persian, which is transcribed as ‘Khwarezm’ in English. The Greeks knew it as *Χορασμία*, and this hellenic form entered the English lexicon as ‘Chorasmia’.

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 БII–ГVI. 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 Б coins have inscriptions in Chorasmian and Greek, and class ГV have transcriptions in Chorasmian and the cursive Sogdian script.

- Inscriptions on wooden items and leather from the palace at Toprak Kala (Топрак-кала), dated to the 3rd century CE.

- 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 CE. 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 Lurje (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 ʾan al-qurūn al-ḫāliya* (*The Remaining Signs of Past Centuries*), which was completed in 1000 CE, Al-Biruni wrote:

> When Ḳutaiba 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,[1] 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. 𐾀 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	𐾀	dual	ʾ
CHORASMIAN LETTER SMALL ALEPH	𐾁	non	-ʾ
CHORASMIAN LETTER BETH	𐾂	dual	*b*
CHORASMIAN LETTER GIMEL	𐾃	dual	*g*
CHORASMIAN LETTER DALETH	𐾄	right	*d*
CHORASMIAN LETTER HE	𐾅	right	*h*
CHORASMIAN LETTER WAW	𐾆	right	*w*
CHORASMIAN LETTER CURLED WAW	𐾇	non	-*w*

CHORASMIAN LETTER ZAYIN	𐾖	dual*	*z*
CHORASMIAN LETTER HETH	𐾗	right	*ḥ*
CHORASMIAN LETTER YODH	𐾘	right	*y*
CHORASMIAN LETTER KAPH	𐾙	dual	*k*
CHORASMIAN LETTER LAMEDH	𐾚	dual	*l*
CHORASMIAN LETTER MEM	𐾛	right	*m*
CHORASMIAN LETTER NUN	𐾜	dual	*n*
CHORASMIAN LETTER SAMEKH	𐾝	dual	*s*
CHORASMIAN LETTER AYIN	𐾞	non	ʿ
CHORASMIAN LETTER PE	𐾟	dual*	*p*
CHORASMIAN LETTER RESH	𐾠	right	*r*
CHORASMIAN LETTER SHIN	𐾡	right	*š*
CHORASMIAN LETTER TAW	𐾢	dual	*t*

4.4.2 Numbers

The numerical notation system is described in § 7.

Character name	Glyph	Joining	Value
CHORASMIAN NUMBER ONE	𐾁	non	1
CHORASMIAN NUMBER TWO	𐾂	non	2
CHORASMIAN NUMBER THREE	𐾃	non	3
CHORASMIAN NUMBER FOUR	𐾄	non	4
CHORASMIAN NUMBER TEN	𐾅	right	10
CHORASMIAN NUMBER TWENTY	𐾆	dual	20
CHORASMIAN NUMBER ONE HUNDRED	𐾇	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:

𐾀 ALEPH < 𐾁 SMALL ALEPH < 𐾂 BETH < 𐾃 GIMEL < 𐾄 DALETH < 𐾅 HE < 𐾆 WAW <
𐾇 CURLED WAW < 𐾈 ZAYIN < 𐾉 HETH < 𐾊 YODH < 𐾋 KAPH < 𐾌 LAMEDH < 𐾍 MEM <
𐾎 NUN < 𐾏 SAMEKH < 𐾐 AYIN < 𐾑 PE < 𐾒 RESH < 𐾓 SHIN < 𐾔 TAW

5 Character Details

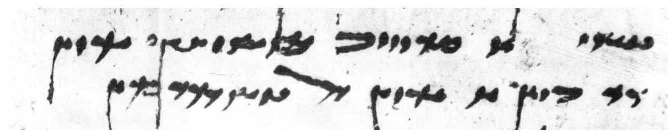
5.1 Homography

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 𐾃 in cursive-medial position. Their nominal forms are distinctive, and are distinguished by the degree of curvature of the primary stroke.

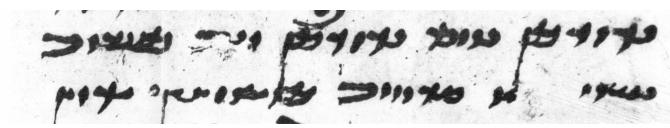
gimel, ***he***, ***kaph*** The letters 𐾃 *gimel*, 𐾅 *he*, 𐾋 *kaph* have the same basic structure: a horzontal stroke attached to a descending stroke. The *kaph* is identified by its broad, horizontal top stroke, and the elongated descender of its final form. The *gimel* is written with a 90° angle. 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 𐾄 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 𐾆 *waw* in 𐾓𐾆𐾔 *gwšt* and 𐾓𐾆𐾈𐾋 *hwnšk* (TK 52, fig. 24), and 𐾔𐾎𐾒𐾑𐾋 *tpnkwk* (TK 25, fig. 25).

daleth*, *ayin*, *resh The letters 𐾴 *daleth*, 𐿀 *ayin*, 𐿂 *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 𐿂 *resh* being longer than that of 𐾴 *daleth* as shown in 𐾳𐿂𐾴𐾽𐾰𐾾 *grdm'n* (TK no. 25, fig. 25 and excerpt below). Such a distinction appears to be carried over from Imperial Aramaic, where the original 𐡃 *daleth* and 𐡓 *resh* differ by the length of the primary stroke, and both differ from 𐡏 *ayin*. The Chorasmian 𐿀 *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. 𐿀𐾼 *'L* (below) and 𐿀𐾲𐾴𐿄 *'BDt*.

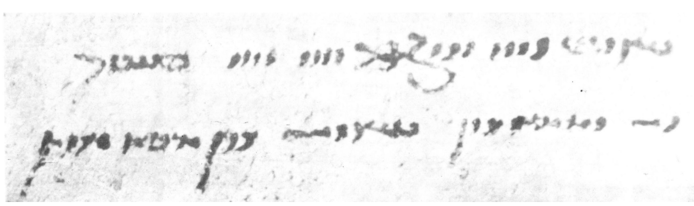


waw*, *yodh The shapes of *waw* and *yodh* appear to have merged in the hand of some scribes and engravers to such an extant that it is difficult to distinguish the two letters in some inscriptions. For example, they occur as 𐾶, 𐾷, 𐾺. The degree of similarity is increased by their identical joining behavior.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

For example, in Tok Kala no. 25, *waw* and *yodh* are clearly distinct in 𐾲𐾺𐾶𐾽 *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: 𐾶 for *waw* in 𐾲𐿂𐾶𐿂𐿄𐾾 *brwrtn*, 𐿄𐿁𐾾𐾻𐾶𐾻 *tpnkwk*, etc. and 𐾺 for *yodh* in 𐾰𐾺 *'y*. Of particular interest is 𐾿𐿂𐾺𐾶𐾺𐾻, which may be interpreted as either the name *srywyk* or *srwywk*, depending upon the scholar's preference. But the first *yodh* in the word resembles that in *BYWM* and *'y*.



The curved sign 𐾺 for *yodh* is also evident in TK 19 (fig. 26, and below). It is clearly observed in 𐾺𐾰 *y'* and less prominently in 𐾾𐿁𐿃𐾺𐾰 *NPŠY'*; compare these with the straight form 𐾶 for *waw* in 𐿄𐿁𐾾𐾻𐾶𐾻 *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 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:

- *Initial* When cursive-initial *aleph* is followed by a letter with a baseline, it connects to the following at the baseline as , eg. *ʾpbntn*. With other letters the left connecting stroke of is raised as in order to join to the midpoint of the main stroke of the following letter, eg. *ʾztyk*. Examination of the occurrences of *aleph* indicates that it consistently connects at the midpoint to letters with a single, vertical stroke at the right edge, eg. *waw*, *yodh*, cursive-final *nun*, and non-left-joining *zayin*. It also exhibits such behavior with the non-left-joining form of *pe*, and angled letters such as *gimel*, *he*, *resh*.

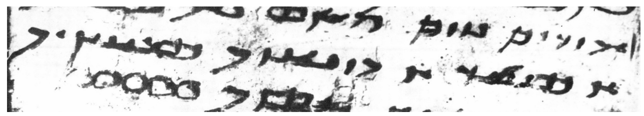
- *Medial* The connection of medial *aleph* follows the rule for the initial form; compare the attachment to a following letter at the midpoint in *kʾkʾn-w*, *tnbʾr*, *grdmʾn* with connections at the baseline in *prʾnyʾty*. In coin and silver vessel inscriptions, the archaic form of cursive medial *aleph* often occurs instead of ; 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. *MLK-ʾ* in coinage (eg. class Γ III, fig. 9); *ZWZN-ʾ* in silver vessels (#5, fig. 20); *YRḤʾ* in ossuaries (TK 25, fig. 25; also the excerpt below). In *MLKʾ* and *ZWZN-ʾ* 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 * 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. *tʾβʾnʾn-ʾ-k* (TK 52, fig. 24; excerpt below); *hyʾn-ʾ* (TK 25, 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 -ʾ 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.

TK 25 

TK 52 

gimel The ר *gimel* connects to the following letter at the top edge, eg. *gimel* + *waw* as ח in the name of a day, חושט *gwšt*; *gimel* + *resh* as רר in גרדמאן *grdmʾn* ‘paradise’. Sequences of *gimel* + *waw* or *yodh* may resemble the letter ח *heth*, but are distinguishable based upon context. The *gimel* has the archaic form ⲧ; which occurs, for example in silver vessel #6 (see fig. 21).

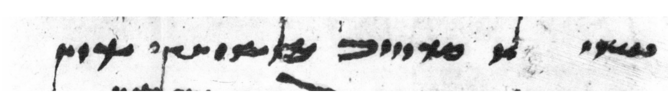
he The following points should be noted:

- *Shaping* The method by which ל *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. אהורים *ʾhwrym* (TK no. 52, fig. 24). When it follows a dual-joining letter that connects at the baseline, the ל *he* is rotated clock-wise as › such that its right-side point may effect a connection at the baseline, eg. זנה *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 ו waw is based upon the form used in Tok Kala no. 25, etc. This form is distinguished from י 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 ו *waw*, which is represented as the non-joining special sign ᵓ (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 *ḍamm*” and occurs in words, such as טישיאנאןᵓ *tyšyʾnʾn-w* in TK no. 25 and וחוונטאןᵓ *wḥwntʾn-w* in TK no. 69 (1965: 178). In both these cases, the ᵓ is attached to personal names that function as patronyms. The “*ḍamm*” 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 ᵓ 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 𐾹.

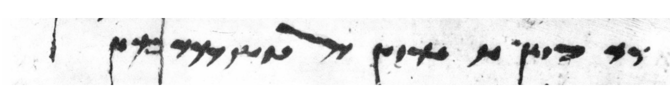
- *Joining* The joining behavior of *zayin* differs depending upon context. In Aramaic heterograms in the Tok Kala ossuaries it joins to the left, eg. 𐾹𐾿𐾵 *ZNH*, while in the silver vessel inscriptions it does not, eg. 𐾹𐾶𐾹𐾿 *ZWZN*. Given this varying behavior, *zayin* is defined as a dual-joining letter. 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 𐾹. When ZWNJ is placed after *zayin*, the letter will be rendered as 𐾹 (see § 6.3).

yodh The representative glyph for 𐾻 *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 𐾿 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 𐾿 in the ossuary inscriptions, in both right-joining and isolated contexts. In coinage and silver vessels, the word final form appears as 𐾿. 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, 𐾹𐾶𐾹𐾿𐾰 *ZWZN-* ʾ occurs on silver vessels, but would appear as 𐾹𐾶𐾹𐾿𐾰 in later styles of the script. A common example is *MN*, which occurs as 𐾾𐾿 in silver vessels and as 𐾾𐾿 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 𐾿 and 𐾿 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: 𐿁𐾲𐾴𐿅 *ʿBDt* 'done, made' (silver bowl #7, fig. 22) and 𐿁𐾽 *ʿL*, 'to' (silver bowls and ossuaries, ie. Tok Kala no. 25, see excerpt below). While it appears to join to *lamedh* in 𐿁𐾽 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.

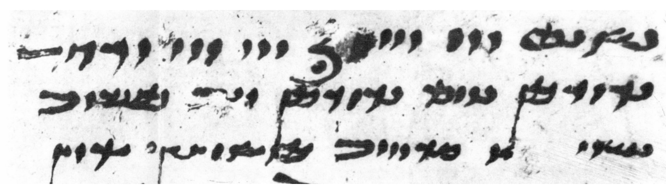
pe Although it defined is a default dual-joining letter, in some sources medial PE does not connect to the left, compare *ʾpbntn* with *ʾp-bntn*. The suspension of the left connection of *pe* occurs in the silver vessels before *beth*, perhaps in order to distinguish the two letters, as their dual-joining medial shape ( ) 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 is represented using the glyphic variant 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 for *taw* is based upon the Tok Kala form. This letter has the glyphic variant that has an open right stroke, which is to be managed using a font.

- *Ligation* In ossuary inscriptions, letters that follow *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. is written as to accommodate a following letter. For instance, *taw* + *yodh* is as in *ʾztyk* on TK no. 52 (fig. 24); *taw* + final *nun* as *tn* in *ʾpbntn*, eg. silver bowl #2 (fig. 17); *taw* + *waw* is *tw* in the name *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	𐾀	*𐾀	(𐾀) 𐾀	(𐾀) 𐾀
BETH	𐾁	𐾁	𐾁	𐾁
GIMEL	𐾂	𐾂	𐾂	𐾂
ZAYIN	𐾃	𐾃	𐾃	𐾃
KAPH	𐾄	𐾄	𐾄	𐾄
LAMEDH	𐾅	𐾅	𐾅	𐾅
NUN	𐾆	𐾆	𐾆	𐾆
SAMEKH	𐾇	𐾇	𐾇	𐾇
PE	𐾈	𐾈	𐾈	𐾈
TAW	𐾉	𐾉	𐾉	𐾉

Joining features of right-joining letters are shown below:

	X_n	X_r
DALETH	𐾊	𐾊
HE	𐾋	(𐾋), 𐾋
WAW	𐾌	𐾌
HETH	𐾍	𐾍
YODH	𐾎	𐾎
MEM	𐾏	𐾏
RESH	𐾐	𐾐
SHIN	𐾑	𐾑

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
HE	𐾵	(𐾵), 𐾵	𐾵
MEM	𐾽	𐾽	𐾽

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): 𐾲𐾺𐾶𐾽 *BYWM* and 𐾽𐾾 (𐾽𐾾) *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. 𐾵 and 𐾵. See, for example 𐾵𐾶 *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:

ʾgʾdk	𐾰𐾳𐾰𐾴𐾻	<𐾰 ALEPH, 𐾳 GIMEL, 𐾰 ALEPH, 𐾴 DALETH, 𐾻 KAPH>
ʾztyk	𐾰𐾸‌𐿇𐾺𐾻	<𐾰 ALEPH, 𐾸 ZAYIN, ‌ ZWNJ, 𐿇 TAW, 𐾺 YODH, 𐾻 KAPH>
ʾpbntn	𐾰𐿂𐾲𐾾𐿇𐾾	<𐾰 ALEPH, 𐿂 PE, 𐾲 BETH, 𐾾 NUN, 𐿇 TAW, 𐾾 NUN>
ʾrʾškrk	𐾰𐿅𐾰𐿆𐾻𐿅𐾻	<𐾰 ALEPH, 𐿅 RESH, 𐾰 ALEPH, 𐿆 SHIN, 𐾻 KAPH, 𐿅 RESH, 𐾻 KAPH>
ʾšpynšwk	𐾰𐿆𐿂𐾺𐾾𐿆𐿇𐾻	<𐾰 ALEPH, 𐿆 SHIN, 𐿂 PE, 𐾺 YODH, 𐾾 NUN, 𐿆 SHIN, 𐿇 TAW, 𐾻 KAPH>
BYWM	𐾲𐾺𐾶𐾽	<𐾲 BETH, 𐾺 YODH, 𐾶 WAW, 𐾽 MEM>
bntk	𐾲𐾾𐿇𐾻	<𐾲 BETH, 𐾾 NUN, 𐿇 TAW, 𐾻 KAPH>

BŠNT	𐾲𐿃𐾾𐿄	<𐾲 BETH, 𐿃 SHIN, 𐾾 NUN, 𐿄 TAW>
grdmʾn	𐾳𐿂𐾴𐾽𐾰𐾾	<𐾳 GIMEL, 𐿂 RESH, 𐾴 DALETH, 𐾽 MEM, 𐾰 ALEPH, 𐾾 NUN>
gwšt	𐾳𐾶𐿃𐿄	<𐾳 GIMEL, 𐾶 WAW, 𐿃 SHIN, 𐿄 TAW>
hwnšk	𐾵𐾶𐾾𐿃𐾻	<𐾵 HE, 𐾶 WAW, 𐾾 NUN, 𐿃 SHIN, 𐾻 KAPH>
hyʾn-ʾ	𐾵𐾺𐾰𐾾𐾱	<𐾵 HE, 𐾺 YODH, 𐾰 ALEPH, 𐾾 NUN, 𐾱 SMALL ALEPH>
wḥwntʾn-w	𐾶𐾹𐾶𐾾𐿄𐾰𐾾𐾷	<𐾶 WAW, 𐾹 HETH, 𐾶 WAW, 𐾾 NUN, 𐿄 TAW, 𐾰 ALEPH, 𐾾 NUN, 𐾷 CURLED WAW>
ZNH	𐾸𐾾𐾵	<𐾸 ZAYIN, 𐾾 NUN 𐾵 HE>
ḥwsrw	𐾹𐾶𐾿𐿂𐾺	<𐾹 HETH, 𐾶 WAW, 𐾿 SAMEKH, 𐿂 RESH, 𐾺 YODH>
ḥwpsk	𐾹𐾶𐿁𐾿𐾻	<𐾹 HETH, 𐾶 WAW, 𐿁 PE, 𐾿 SAMEKH, 𐾻 KAPH>
YRḤʾ	𐾺𐿂𐾹𐾰	<𐾺 YODH, 𐿂 RESH, 𐾹 HETH, 𐾰 ALEPH>
kʾkʾn-w	𐾻𐾰𐾻𐾰𐾾𐾷	<𐾻 KAPH, 𐾰 ALEPH, 𐾻 KAPH, 𐾰 ALEPH, 𐾾 NUN, 𐾷 CURLED WAW>
KSP	𐾻𐾿𐿁	<𐾻 KAPH, 𐾿 SAMEKH, 𐿁 PE>
mʾnyʾʾty	𐾽𐾰𐾾𐾺𐾰𐾰𐿄𐾺	<𐾽 MEM, 𐾰 ALEPH, 𐾾 NUN, 𐾺 YODH, 𐾰 ALEPH, 𐾰 ALEPH, 𐿄 TAW, 𐾺 YODH>
MRʾY	𐾽𐿂𐾰𐾺	<𐾽 MEM, 𐿂 RESH, 𐾰 ALEPH, 𐾺 YODH>
NPŠY	𐾾𐿁𐿃𐾺	<𐾾 NUN, 𐿁 PE, 𐿃 SHIN, 𐾺 YODH>
syʾwršprn	𐾿𐾺𐾰𐾶𐿂𐿃𐿁𐿂𐾾	<𐾿 SAMEKH, 𐾺 YODH, 𐾰 ALEPH, 𐾶 WAW, 𐿂 RESH, 𐿃 SHIN, 𐿁 PE, 𐿂 RESH, 𐾾 NUN>
ʿBDt	𐿀𐾲𐾴𐿄	<𐿀 AYIN, 𐾲 BETH, 𐾴 DALETH, 𐿄 TAW>
ʿL	𐿀𐾼	<𐿀 AYIN, 𐾼 LAMEDH>
prnxwnt	𐿁𐿂𐾰𐾾𐾶𐾾𐿄	<𐿁 PE, 𐿂 RESH, 𐾰 ALEPH, 𐾾 NUN, 𐾶 WAW, 𐾾 NUN, 𐿄 TAW>
twtwḥs	𐿄𐾶𐿄𐾶𐾹𐾿	<𐿄 TAW, 𐾶 WAW, 𐿄 TAW, 𐾶 WAW, 𐾹 HETH, 𐾿 SAMEKH>
tnbryk	𐿄𐾾𐾲𐿂𐾺𐾻	<𐿄 TAW, 𐾾 NUN, 𐾲 BETH, 𐿂 RESH, 𐾺 YODH, 𐾻 KAPH>
tpnkwk	𐿄𐿁𐾾𐾻𐾶𐾻	<𐿄 TAW, 𐿁 PE, 𐾾 NUN, 𐾻 KAPH, 𐾶 WAW, 𐾻 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 U+200C ZERO WIDTH NON-JOINER (abbreviated as ZWNJ). This character is placed after the letter whose left-side connection is to be suspended.

ʾpbntn	𐾰𐿁𐾲𐾾𐿄𐾾	<𐾰 ALEPH, 𐿁 PE, 𐾲 BETH, 𐾾 NUN, 𐿄 TAW, 𐾾 NUN>
ʾp-bntn	𐾰𐿁‌𐾲𐾾𐿄𐾾	<𐾰 ALEPH, 𐿁 PE, ‌ ZWNJ, 𐾲 BETH, 𐾾 NUN, 𐿄 TAW, 𐾾 NUN>
ZNH	𐾸𐾾𐾵	<𐾸 ZAYIN, 𐾾 NUN 𐾵 HE>
ZNH	𐾸‌𐾾𐾵	<𐾸 ZAYIN, ‌ ZWNJ, 𐾾 NUN 𐾵 HE>
ZWZN	𐾸‌𐾶𐾸‌𐾾	<𐾸 ZAYIN, ‌ ZWNJ, 𐾶 WAW, 𐾸 ZAYIN, ‌ ZWNJ, 𐾾 NUN>
ZWZN-ʾ	𐾸‌𐾶𐾸‌𐾾‌𐾰	<𐾸 ZAYIN, ‌ ZWNJ, 𐾶 WAW, 𐾸 ZAYIN, ‌ ZWNJ, 𐾾 NUN, ‌ ZWNJ, 𐾰 ALEPH>
MLKʾ	𐾽𐾼𐾻‌𐾰	<𐾽 MEM, 𐾼 LAMEDH, 𐾻 KAPH, ‌ ZWNJ, 𐾰 ALEPH>
prnxwnt	𐿁𐿂𐾰𐾾𐾶𐾾𐿄	<𐿁 PE, 𐿂 RESH, 𐾰 ALEPH, 𐾾 NUN, 𐾶 WAW, 𐾾 NUN, 𐿄 TAW>
prnxwnt	𐿁‌𐿂𐾰𐾾𐾶𐾾𐿄	<𐿁 PE, ‌ ZWNJ, 𐿂 RESH, 𐾰 ALEPH, 𐾾 NUN, 𐾶 WAW, 𐾾 NUN, 𐿄 TAW>

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 𐾾𐿁𐿃𐾺 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 𐾾𐿁𐿃𐾺 is *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, 𐿄𐿁𐾾𐾻𐾶𐾻 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 𐿄𐿁𐾾𐾻𐾶𐾻 as *tpnkwk*, which carries the same meaning as *tpnkwk* and adheres to rules of the script.

The word 𐾵𐾶 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 𐾻𐾰𐾻𐾰𐾾‌𐾶 *kʾkʾn-w* (TK no. 26, fig. 31). Following the behavior of *kaph* in the latter, reading 𐾵𐾶 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	𐾡	—	—	—	non
TWO	𐾢	—	—	—	non
THREE	𐾣	—	—	—	non
FOUR	𐾤	—	—	—	non
TEN	𐾥	𐾥	—	—	right
TWENTY	𐾦	𐾦	𐾦	𐾦	dual
ONE HUNDRED	𐾧	—	—	𐾧	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 𐾡 for 'one'. Numbers 2–9 are written using sequences of 𐾡 arranged in groups containing three or four instances of the sign. For example, '𐾣 𐾤' for 7 and '𐾤 𐾤' for 8 in TK no. 19 (fig. 26), as well as '𐾡 𐾣 𐾣' for 7 in TK no. 25 (fig. 25). Also, '𐾢 𐾣' for the number 5 in silver bowl #2 (fig. 17). The number 5 in silver bowl #5 appears as '𐾤𐾡' without a spaced grouping, but the extended terminal of the third 𐾡 suggests the intended grouping '𐾢 𐾣' despite lack of spacing (fig. 20).

1	𐾡	𐾡 ONE
2	𐾢	𐾢 TWO
3	𐾣	𐾣 THREE
4	𐾤	𐾤 FOUR
5	𐾢 𐾣	<𐾣 THREE, 𐾢 TWO>
6	𐾣 𐾣	<𐾣 THREE, 𐾣 THREE>
7	𐾣 𐾤	<𐾤 FOUR, 𐾣 THREE>

	𐾁 𐾂 𐾂	<𐾂 THREE, 𐾂 THREE, 𐾁 ONE>
8	𐾃 𐾃	<𐾃 FOUR, 𐾃 FOUR>
9	𐾂 𐾂 𐾂	<𐾂 THREE, 𐾂 THREE, 𐾂 THREE>

Ten The 𐾊 TEN resembles a vertically compressed 𐾉 LAMEDH.

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

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

10	𐾊	<𐾊 TEN>
20	𐾋	<𐾋 TWENTY>
30	𐾋𐾊	<𐾋 TWENTY, 𐾊 TEN>
40	𐾋𐾋	<𐾋 TWENTY, 𐾋 TWENTY>
50	𐾋𐾋𐾊	<𐾋 TWENTY, 𐾋 TWENTY, 𐾊 TEN>
60	𐾋𐾋𐾋	<𐾋 TWENTY, 𐾋 TWENTY, 𐾋 TWENTY>
70	𐾋𐾋𐾋𐾊	<𐾋 TWENTY, 𐾋 TWENTY, 𐾋 TWENTY, 𐾊 TEN>
80	𐾋𐾋𐾋𐾋	<𐾋 TWENTY, 𐾋 TWENTY, 𐾋 TWENTY, 𐾋 TWENTY>
90	𐾋𐾋𐾋𐾋𐾊	<𐾋 TWENTY, 𐾋 TWENTY, 𐾋 TWENTY, 𐾋 TWENTY, 𐾊 TEN>

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

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

100	𐾌	<𐾌 ONE HUNDRED>
200	𐾌𐾀	<𐾀 TWO, 𐾌 ONE HUNDRED>
300	𐾌𐾂	<𐾂 THREE, 𐾌 ONE HUNDRED>

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

570	𐿆𐿇𐿋𐿊𐿊𐿊𐿉	<𐿆 TWO, 𐿇 THREE, 𐿋 ONE HUNDRED, 𐿊 TWENTY, 𐿊 TWENTY, 𐿊 TWENTY, 𐿉 TEN>
678	𐿇𐿇𐿋𐿊𐿊𐿊𐿉𐿇𐿇𐿆	<𐿇 THREE, 𐿇 THREE, 𐿋 ONE HUNDRED, 𐿊 TWENTY, 𐿊 TWENTY, 𐿊 TWENTY, 𐿉 TEN, 𐿇 THREE, 𐿇 THREE, 𐿆 TWO>

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 `Vertical_Orientation` (`vo`) property for specifying the orientation of characters in vertical enviroments. For Chorasmian, this property would be defined as: `Vertical_Orientation=R` or `vo=R`, 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, A. I.; Lurje, P. B.]. 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, I. A.] 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 Khorezma*]. Москва: Наука, Глав. ред. восточной литераы

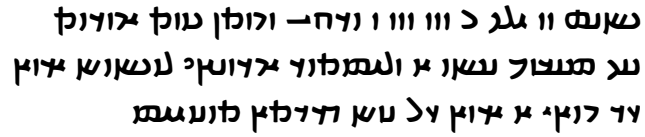
11 Acknowledgments

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.

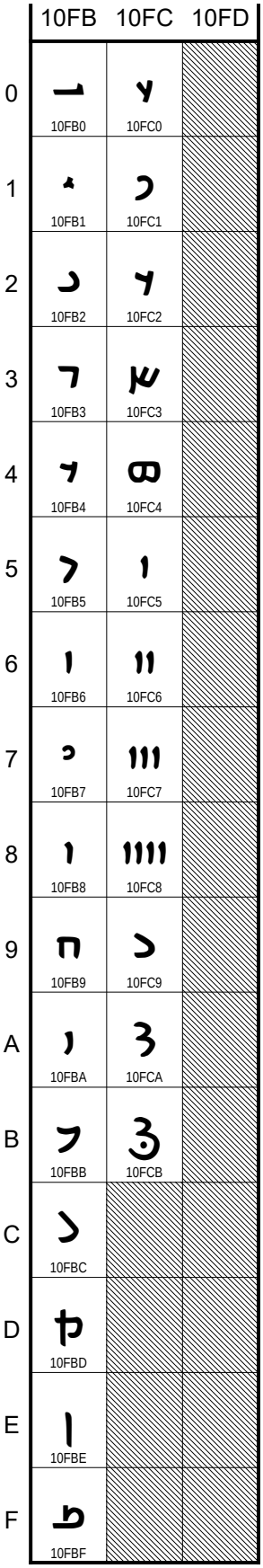
12 Dedication

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).



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



Also known as 'Khwarezmian'.

Letters

10FB0 𐾰 CHORASMIAN LETTER ALEPH
10FB1 𐾱 CHORASMIAN LETTER SMALL ALEPH
10FB2 𐾲 CHORASMIAN LETTER BETH
10FB3 𐾳 CHORASMIAN LETTER GIMEL
10FB4 𐾴 CHORASMIAN LETTER DALETH
10FB5 𐾵 CHORASMIAN LETTER HE
10FB6 𐾶 CHORASMIAN LETTER WAW
10FB7 𐾷 CHORASMIAN LETTER CURLED WAW
10FB8 𐾸 CHORASMIAN LETTER ZAYIN
10FB9 𐾹 CHORASMIAN LETTER HETH
10FBA 𐾺 CHORASMIAN LETTER YODH
10FBB 𐾻 CHORASMIAN LETTER KAPH
10FBC 𐾼 CHORASMIAN LETTER LAMEDH
10FBD 𐾽 CHORASMIAN LETTER MEM
10FBE 𐾾 CHORASMIAN LETTER NUN
10FBF 𐾿 CHORASMIAN LETTER SAMEKH
10FC0 𐿀 CHORASMIAN LETTER AYIN
10FC1 𐿁 CHORASMIAN LETTER PE
10FC2 𐿂 CHORASMIAN LETTER RESH
10FC3 𐿃 CHORASMIAN LETTER SHIN
10FC4 𐿄 CHORASMIAN LETTER TAW

Numbers

10FC5 𐿅 CHORASMIAN NUMBER ONE
10FC6 𐿆 CHORASMIAN NUMBER TWO
10FC7 𐿇 CHORASMIAN NUMBER THREE
10FC8 𐿈 CHORASMIAN NUMBER FOUR
10FC9 𐿉 CHORASMIAN NUMBER TEN
10FCA 𐿊 CHORASMIAN NUMBER TWENTY
10FCB 𐿋 CHORASMIAN NUMBER ONE HUNDRED

	Chorasmian	Old Sogdian	Inscriptional Pahlavi	Inscriptional Parthian	Imperial Aramaic
aleph	𐾀	𐼀	𐭠	𐭡	𐡀
beth	𐾁	𐼁	𐭡	𐭢	𐡁
gimel	𐾂	𐼂	𐭢	𐭣	𐡂
daleth	𐾃	(𐼀)	𐭣	𐭤	𐡃
he	𐾄	𐼃, 𐼄	𐭤	𐭥	𐡄
waw	𐾅	𐼅	𐭥	𐭦	𐡅
zayin	𐾆	𐼆	𐭦	𐭧	𐡆
heth	𐾇	𐼇	𐭧	𐭨	𐡇
teth	—	—	𐭨	𐭩	𐡈
yodh	𐾈	𐼈	𐭩	𐭪	𐡉
kaph	𐾉	𐼉	𐭪	𐭫	𐡊
lamedh	𐾊	𐼊	𐭫	𐭬	𐡋
mem	𐾋	𐼋	𐭬	𐭭	𐡌
nun	𐾌	𐼌	𐭭	𐭮	𐡍
samekh	𐾍	𐼍	𐭮	𐭯	𐡎
ayin	𐾎	𐼎, (𐼀)	(𐭥)	𐭰	𐡏
pe	𐾏	𐼏	𐭰	𐭱	𐡐
sadhe	—	𐼐	𐭲	𐭲	𐡑
qoph	—	—	(𐭬)	𐭳	𐡒
resh	𐾐	𐼑	(𐭥)	𐭴	𐡓
shin	𐾑	𐼒	𐭱	𐭵	𐡔
taw	𐾒	𐼓	𐭲	𐭶	𐡕

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	𐾡	𐼧	𐭡	𐭡	𐡡
TWO	𐾢	𐼨	𐭢	𐭢	𐡢
THREE	𐾣	𐼩	𐭣	𐭣	𐡣
FOUR	𐾤	𐼪	𐭤	𐭤	—
FIVE	—	𐼫	—	—	—
TEN	𐾥	𐼬	𐭥	𐭥	𐡥
TWENTY	𐾦	𐼭	𐭦	𐭦	𐡦
THIRTY	—	𐼮	—	—	—
ONE HUNDRED	𐾧	𐼯	𐭧	𐭧	𐡧
ONE THOUSAND	—	—	𐭨	𐭨	𐡨
TEN THOUSAND	—	—	—	—	𐡩
ONE HALF	—	𐼰	—	—	—

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

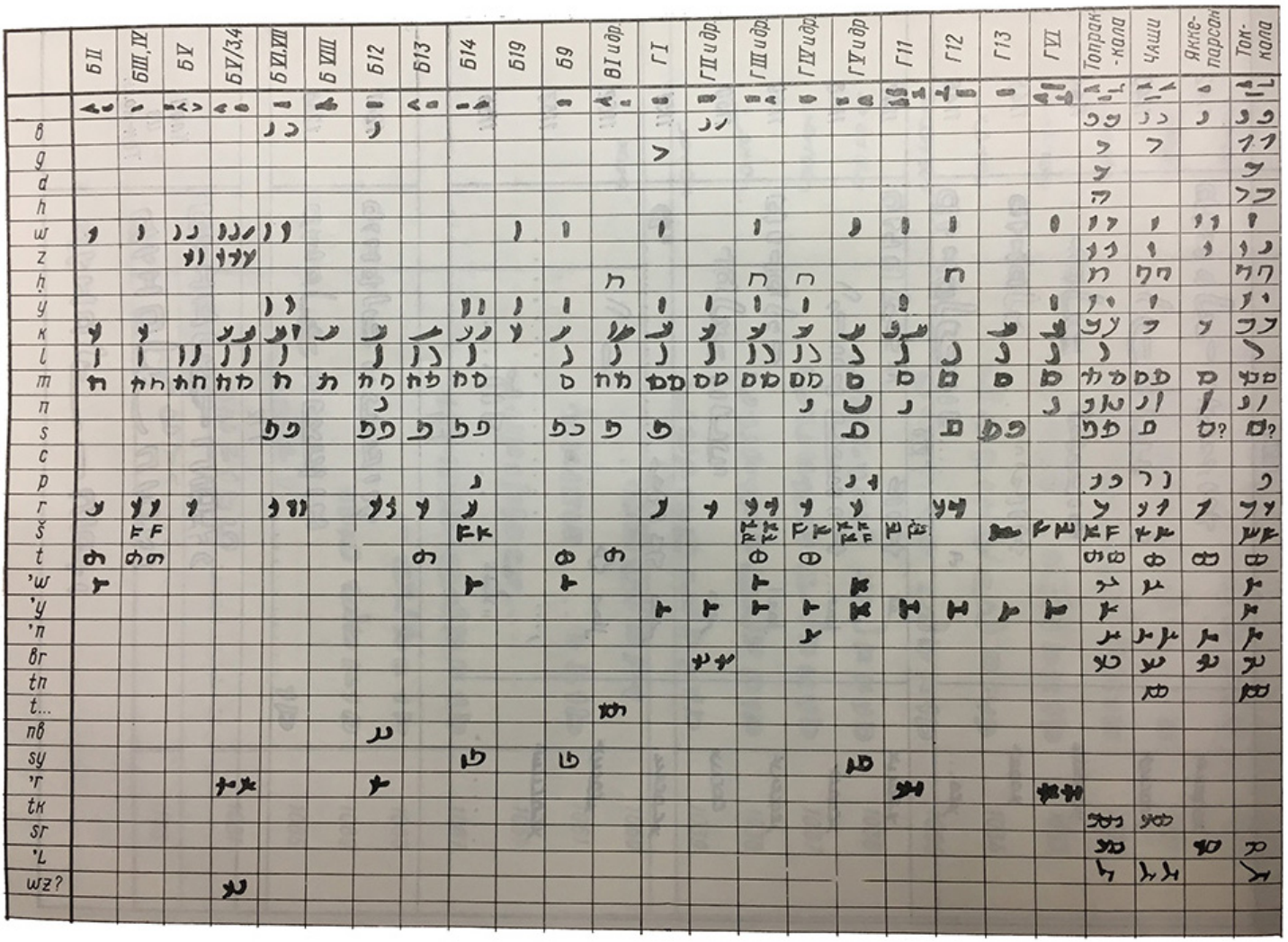


Figure 1: Inventory of characters on Chorasmian coins (БII–ГVI), Toprak Kala (Топрак-кала), Yakke Parsan (Якке парсан), Tok Kala (Ток-кала) (from Vainberg 1977: Table 8).

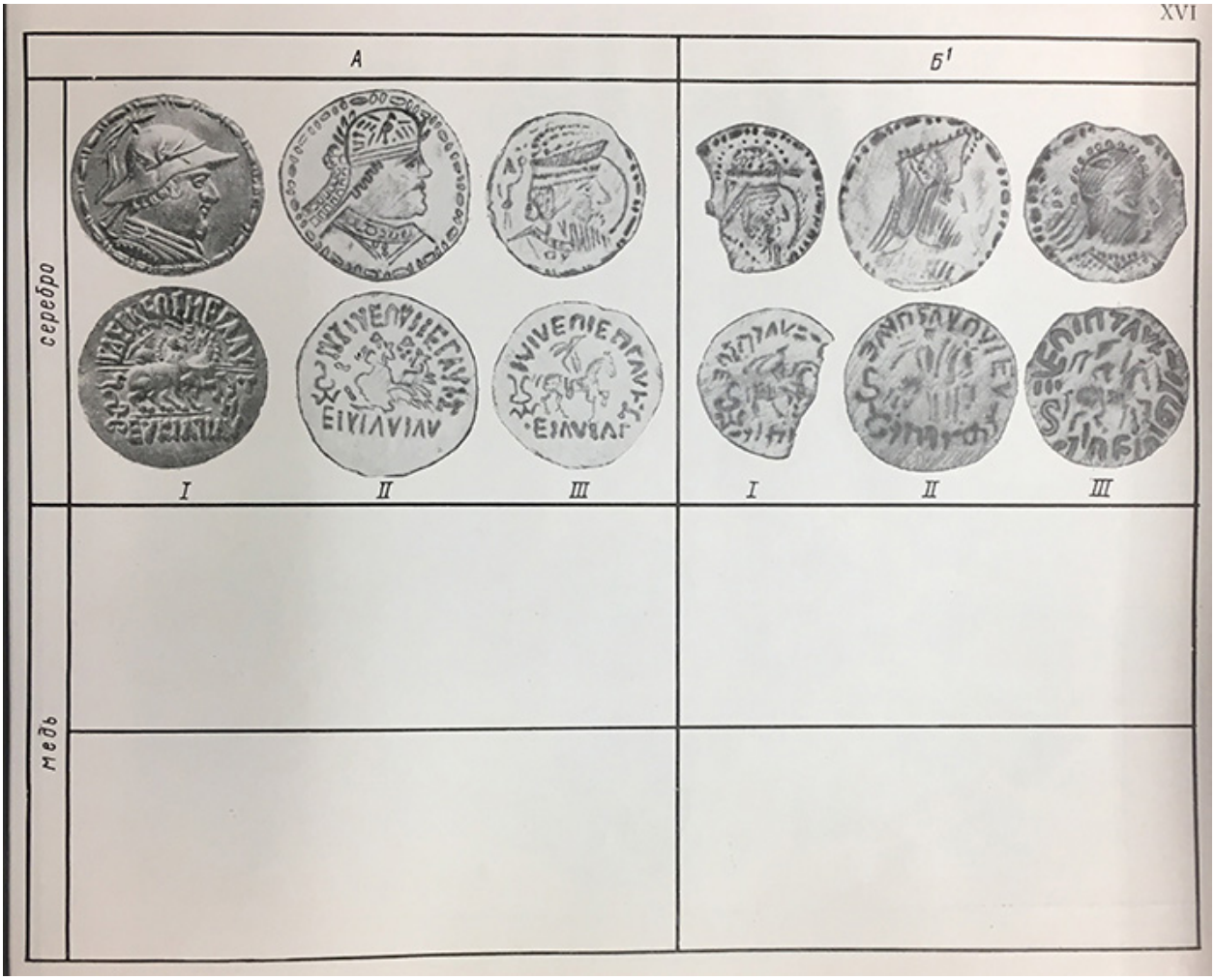


Figure 2: Chorasmian coins (from Vainberg 1977: Table 16).

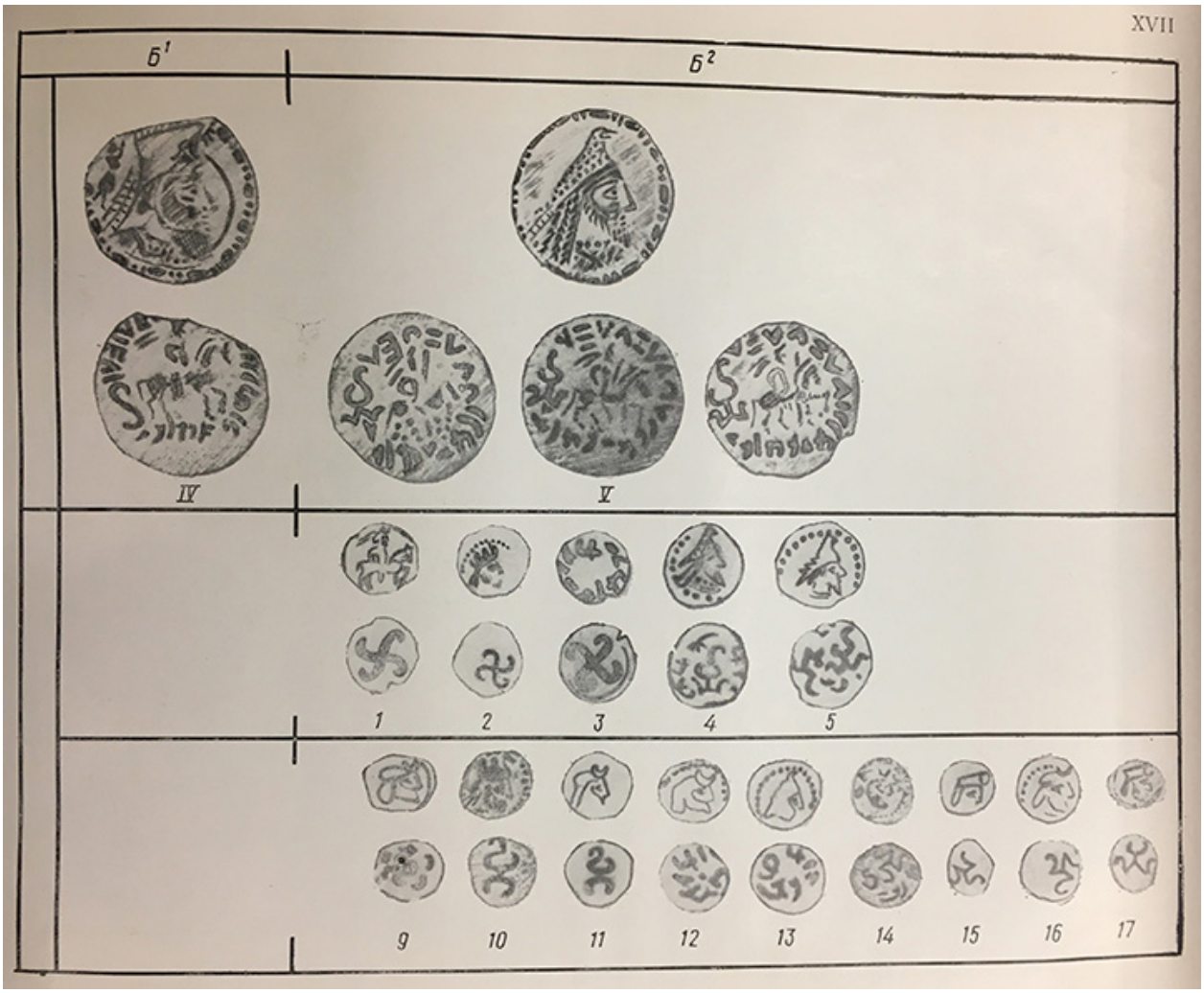


Figure 3: Chorasmian coins (from Vainberg 1977: Table 17).

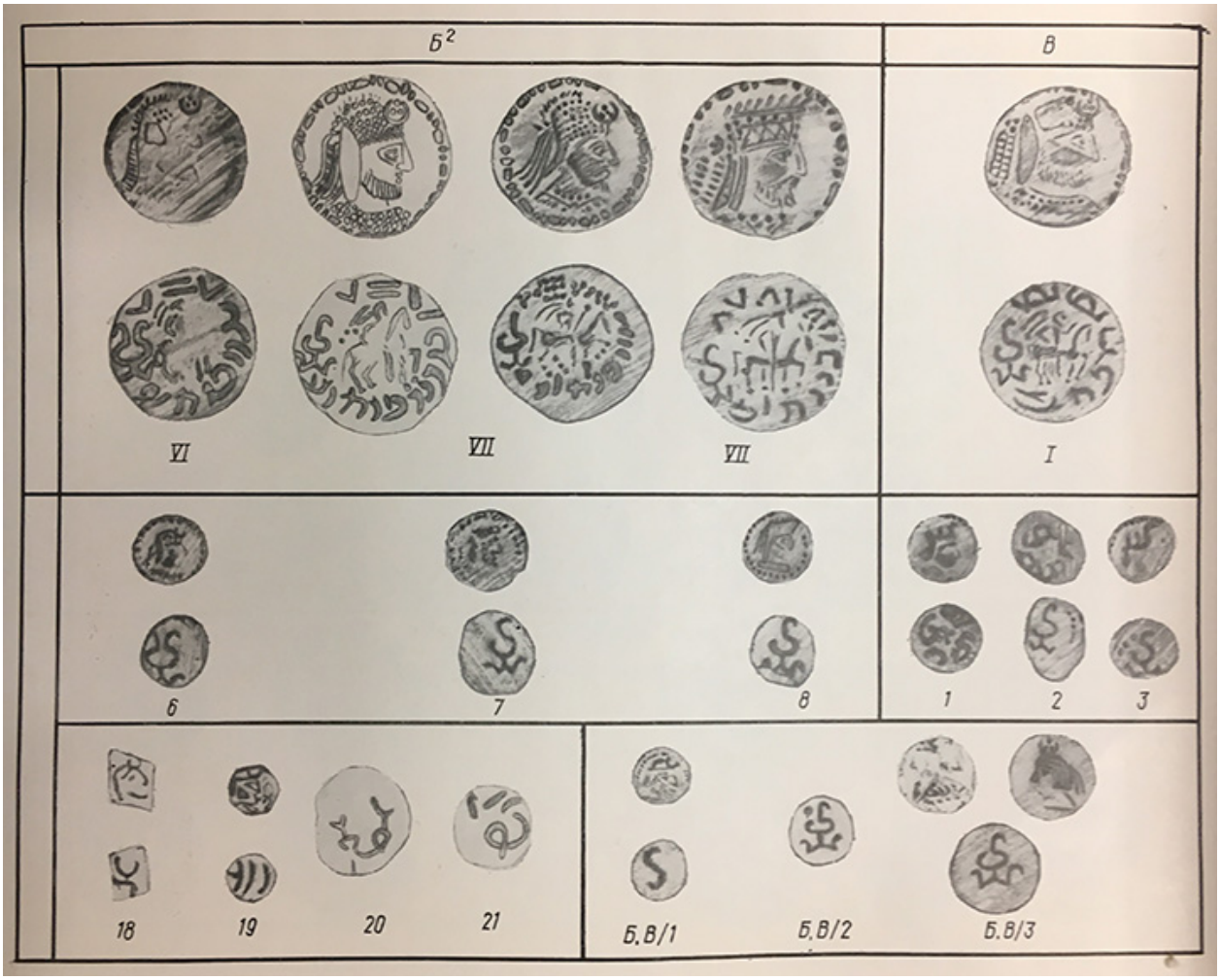


Figure 4: Chorasmian coins (from Vainberg 1977: Table 18).

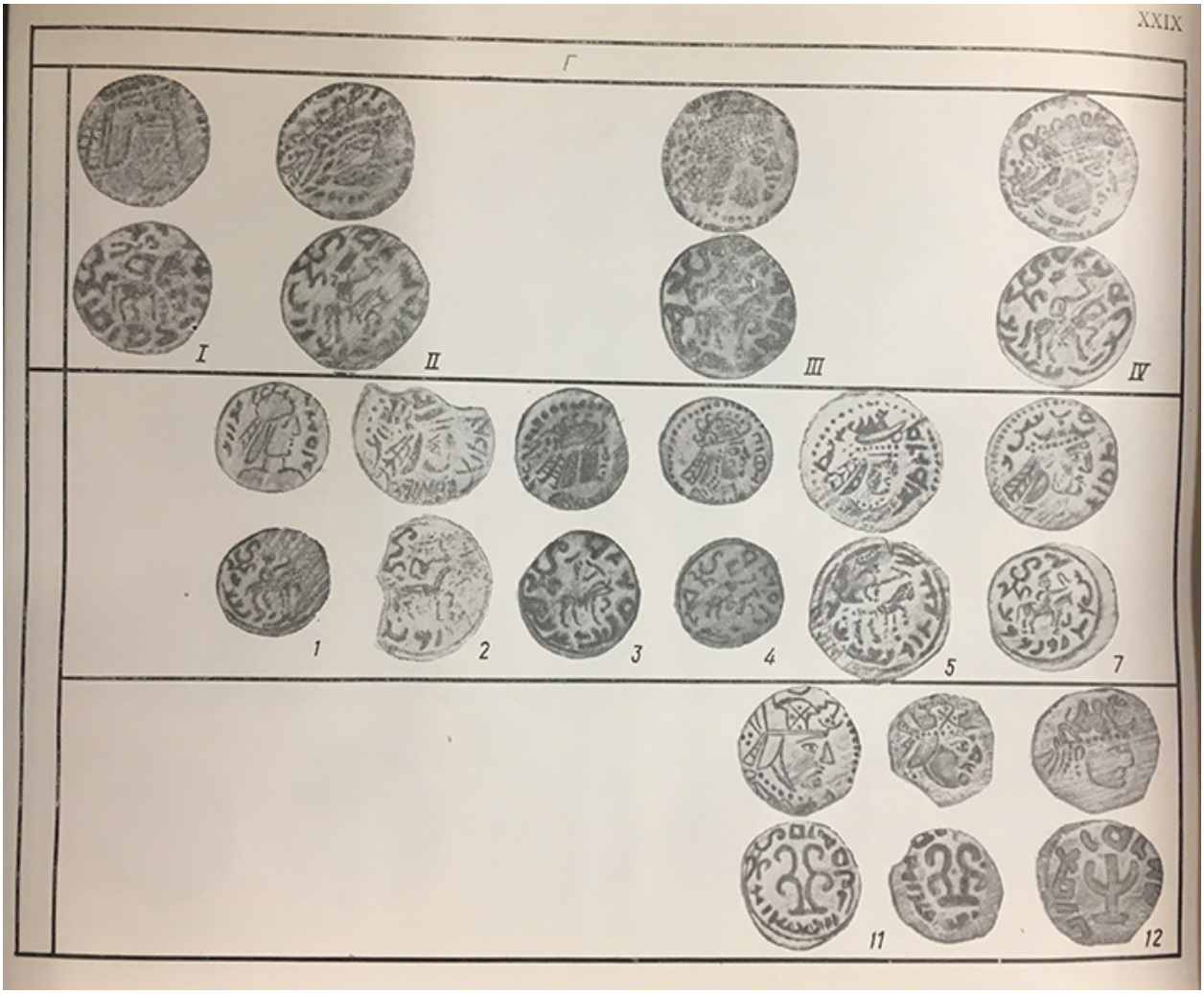


Figure 5: Chorasmian coins (from Vainberg 1977: Table 19).

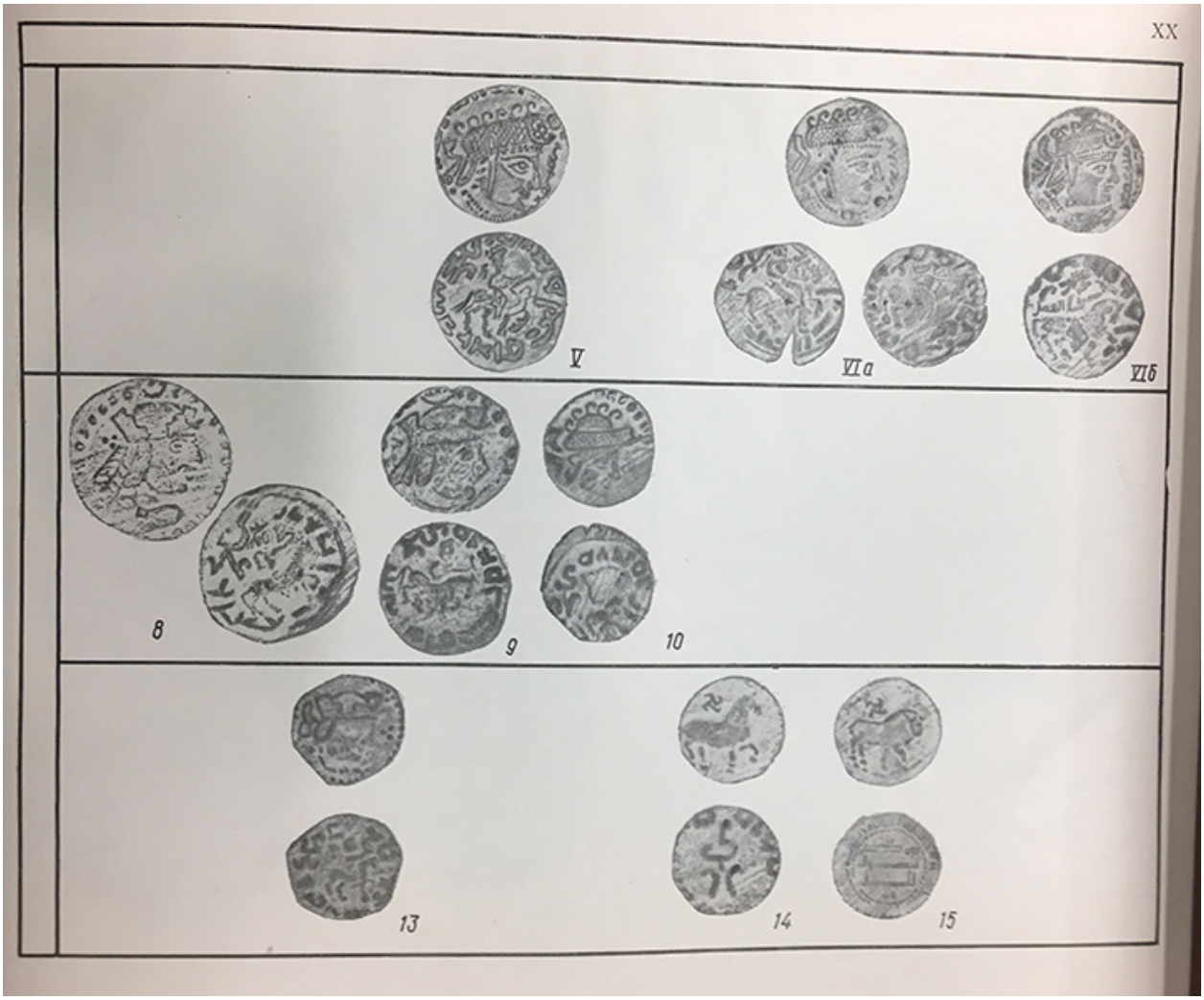


Figure 6: Chorasmian coins (from Vainberg 1977: Table 20).

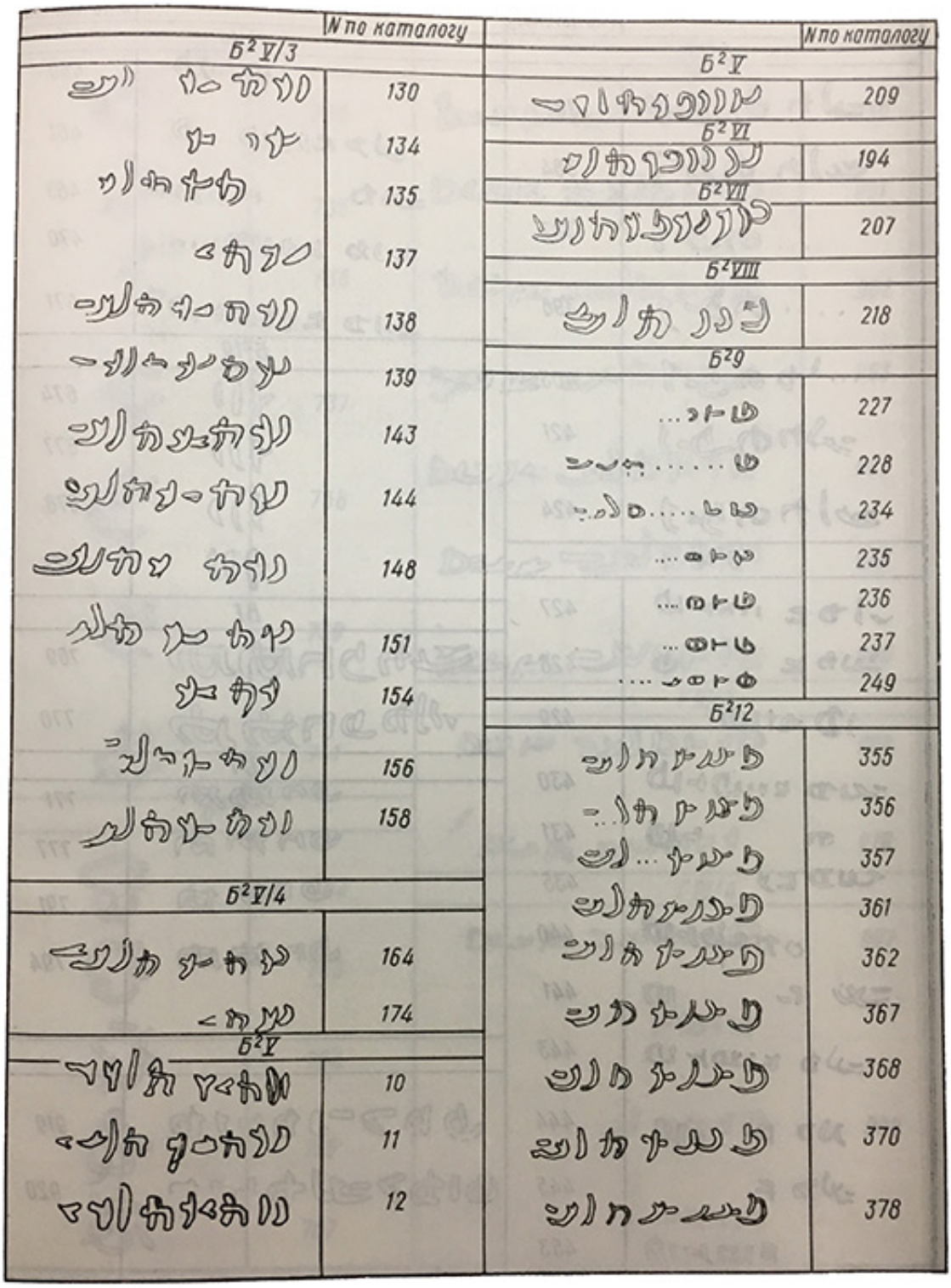


Figure 7: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 1).

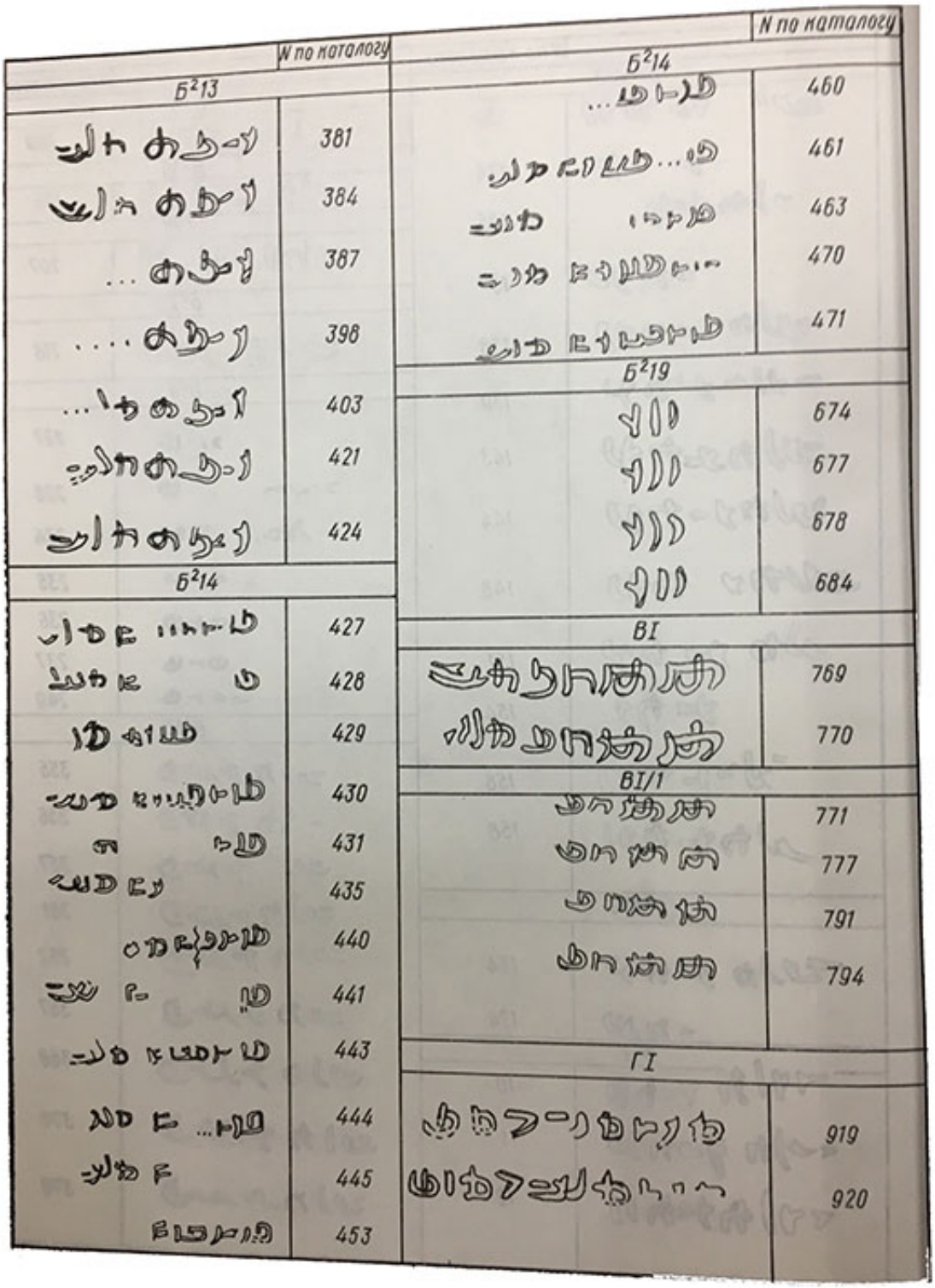


Figure 8: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 2).

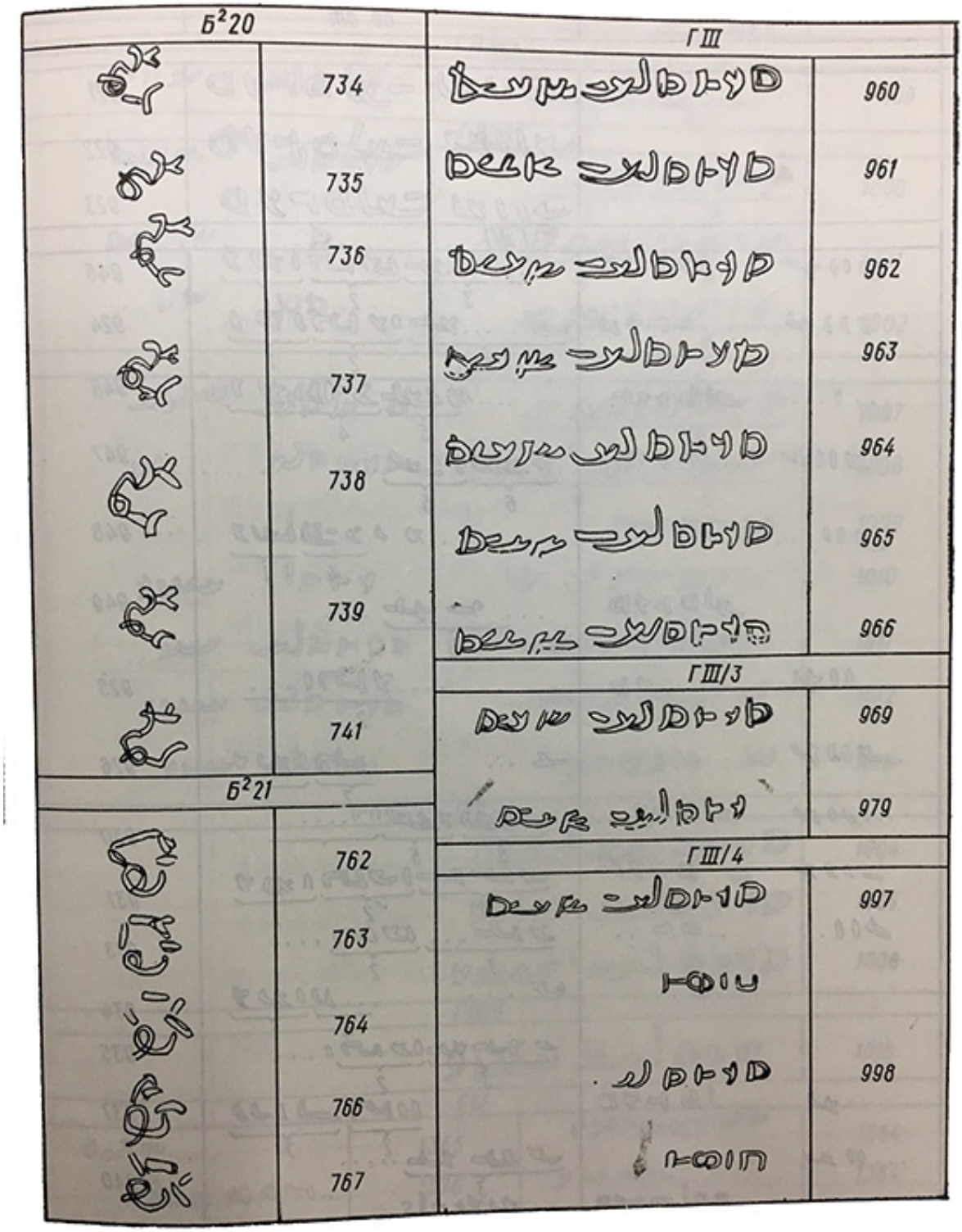


Figure 9: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 3).

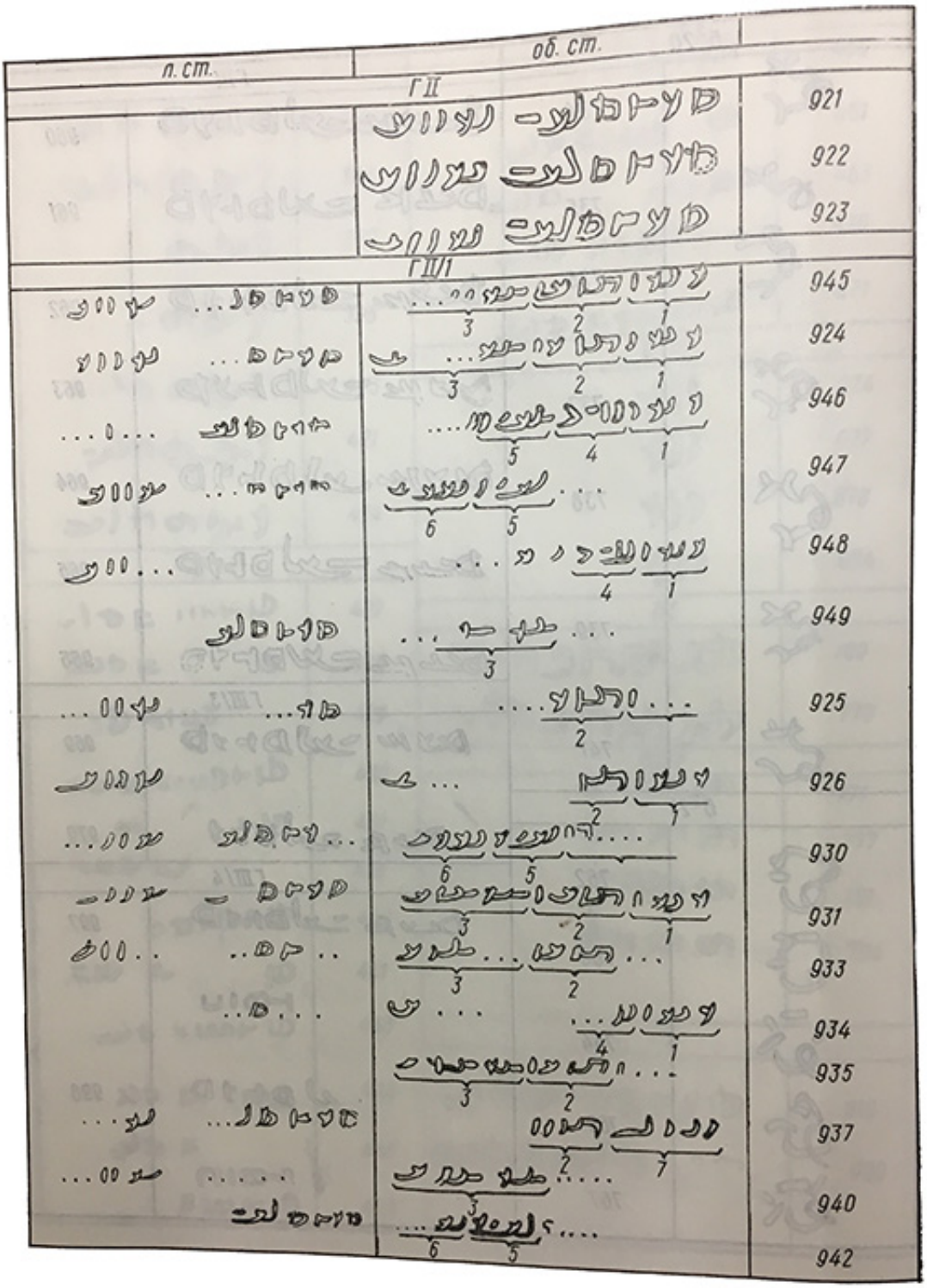


Figure 10: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 4).

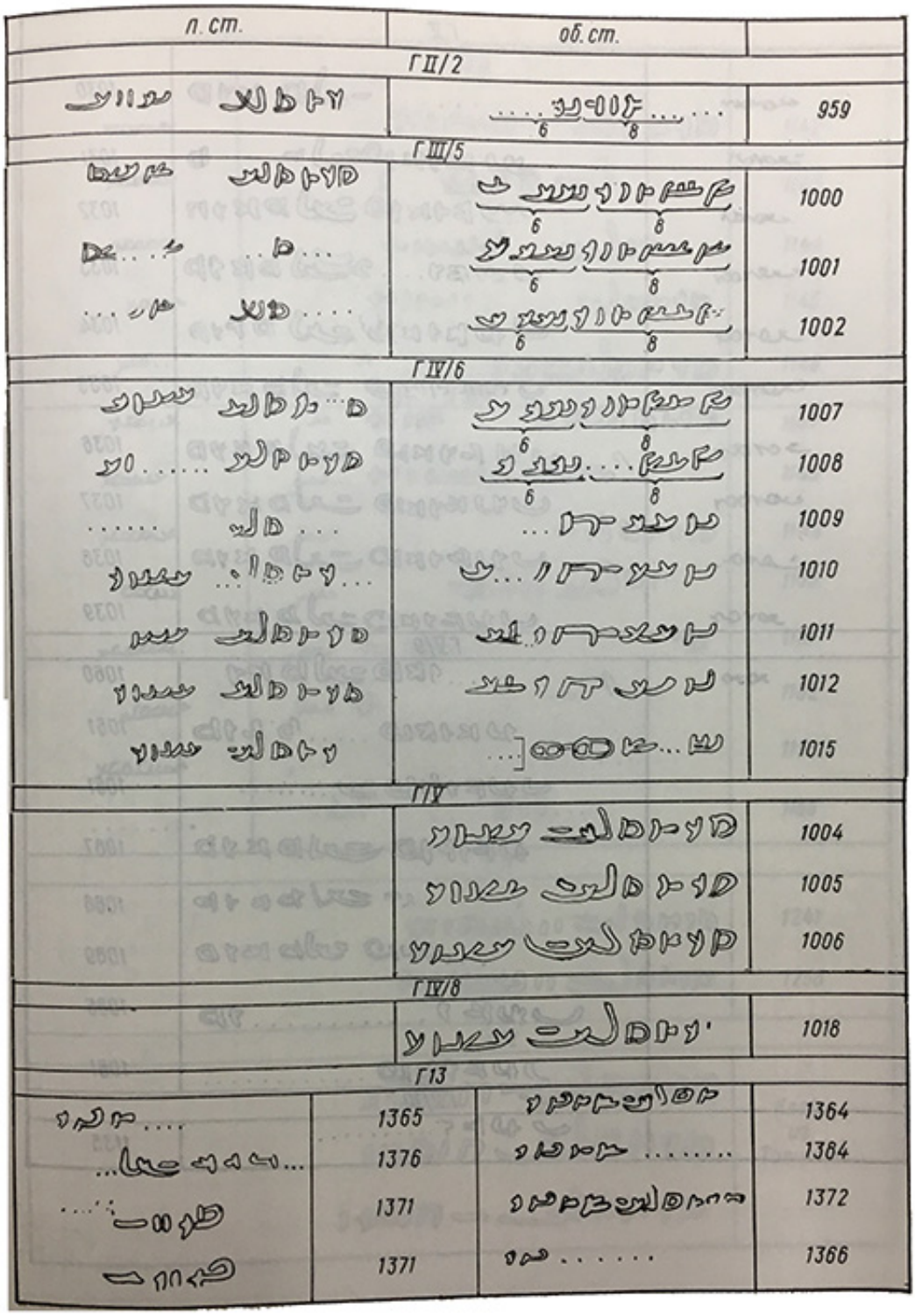


Figure 11: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 5).

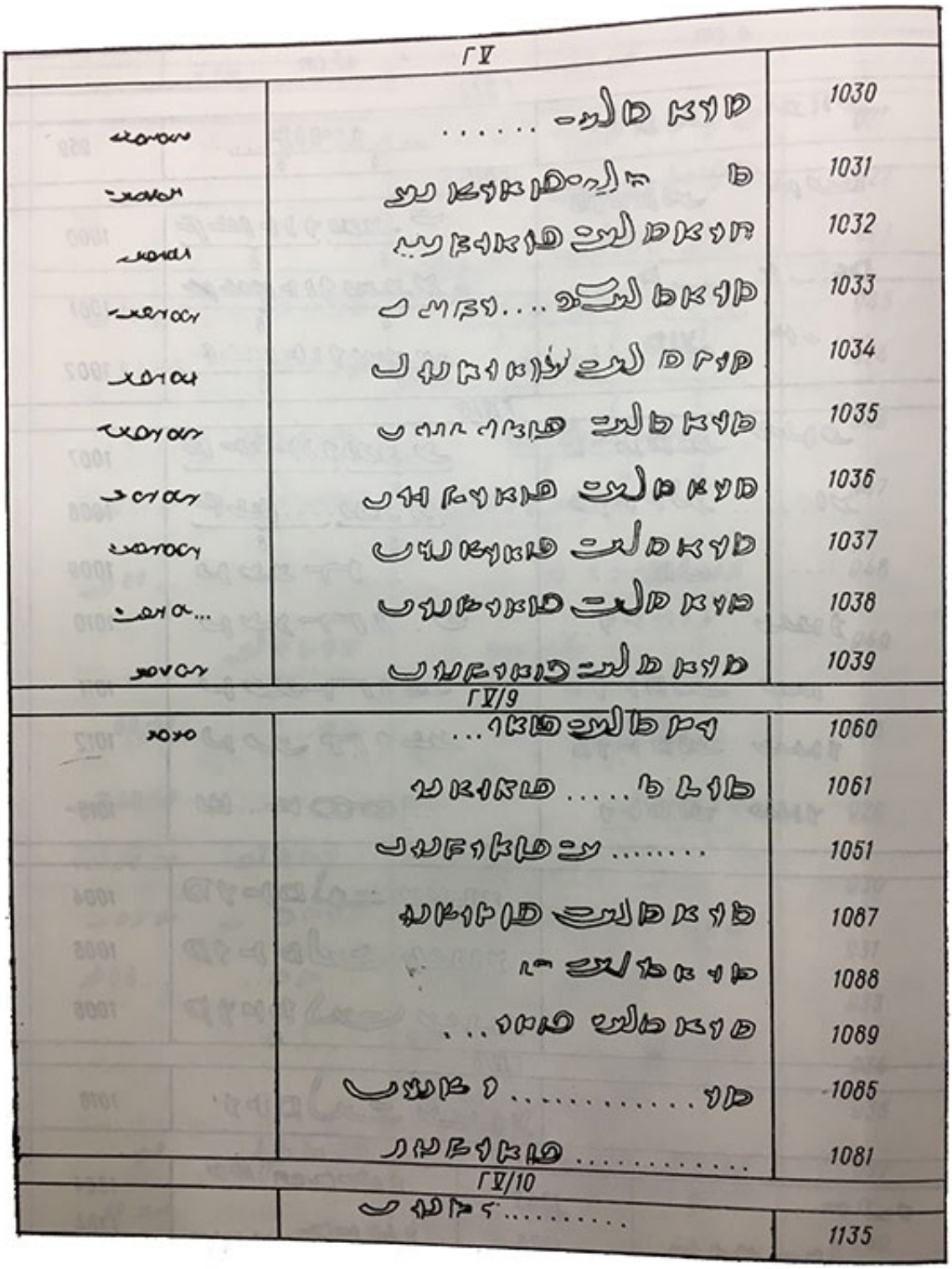


Figure 12: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 6).

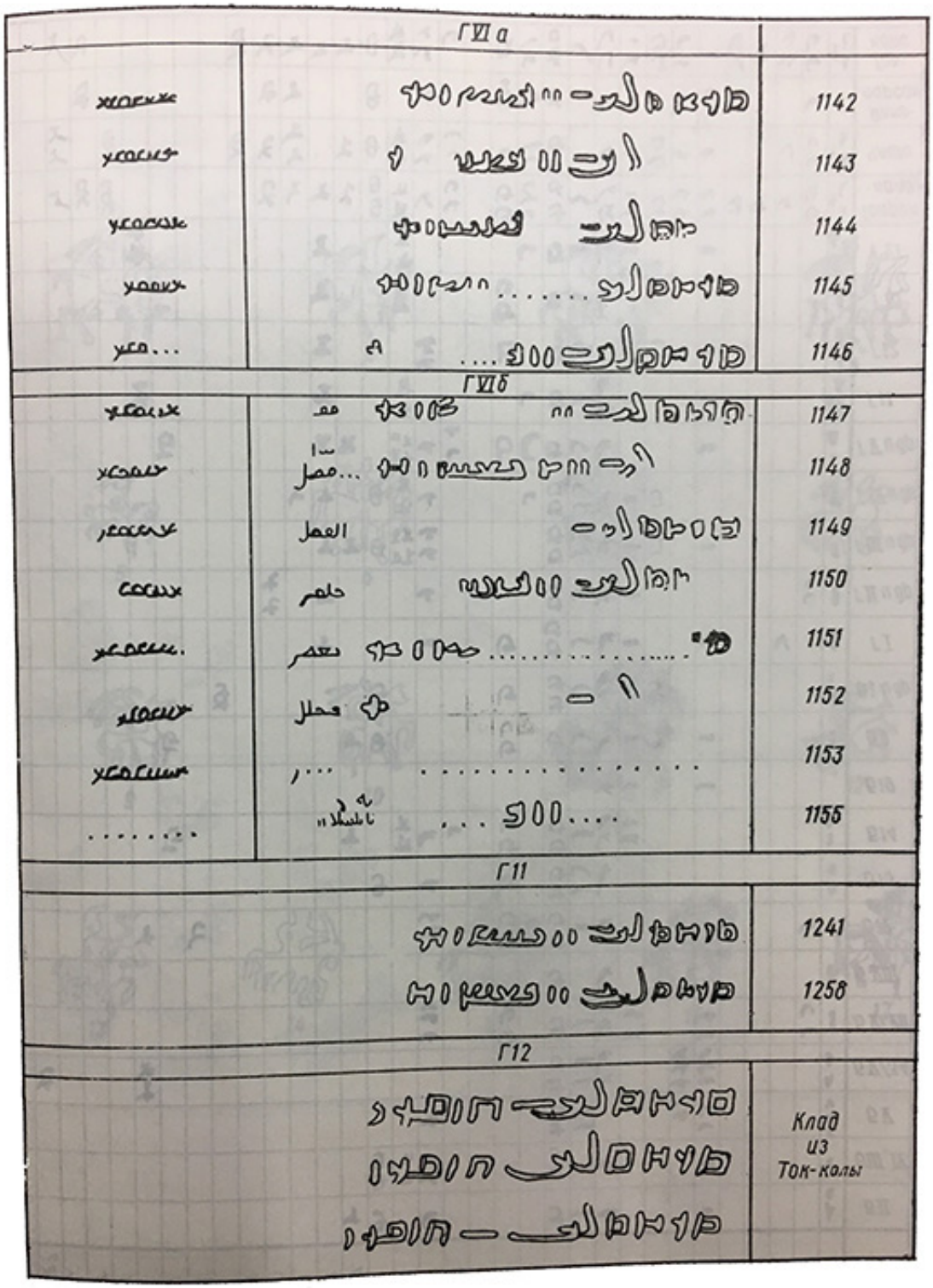


Figure 13: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 7).

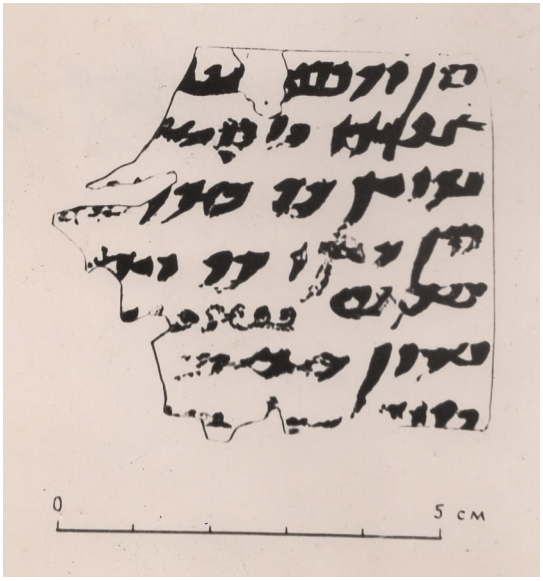


Figure 14: Fragment of a leather document with Chorasmian inscription from Yakke Parsan. Image courtesy of Lurje.

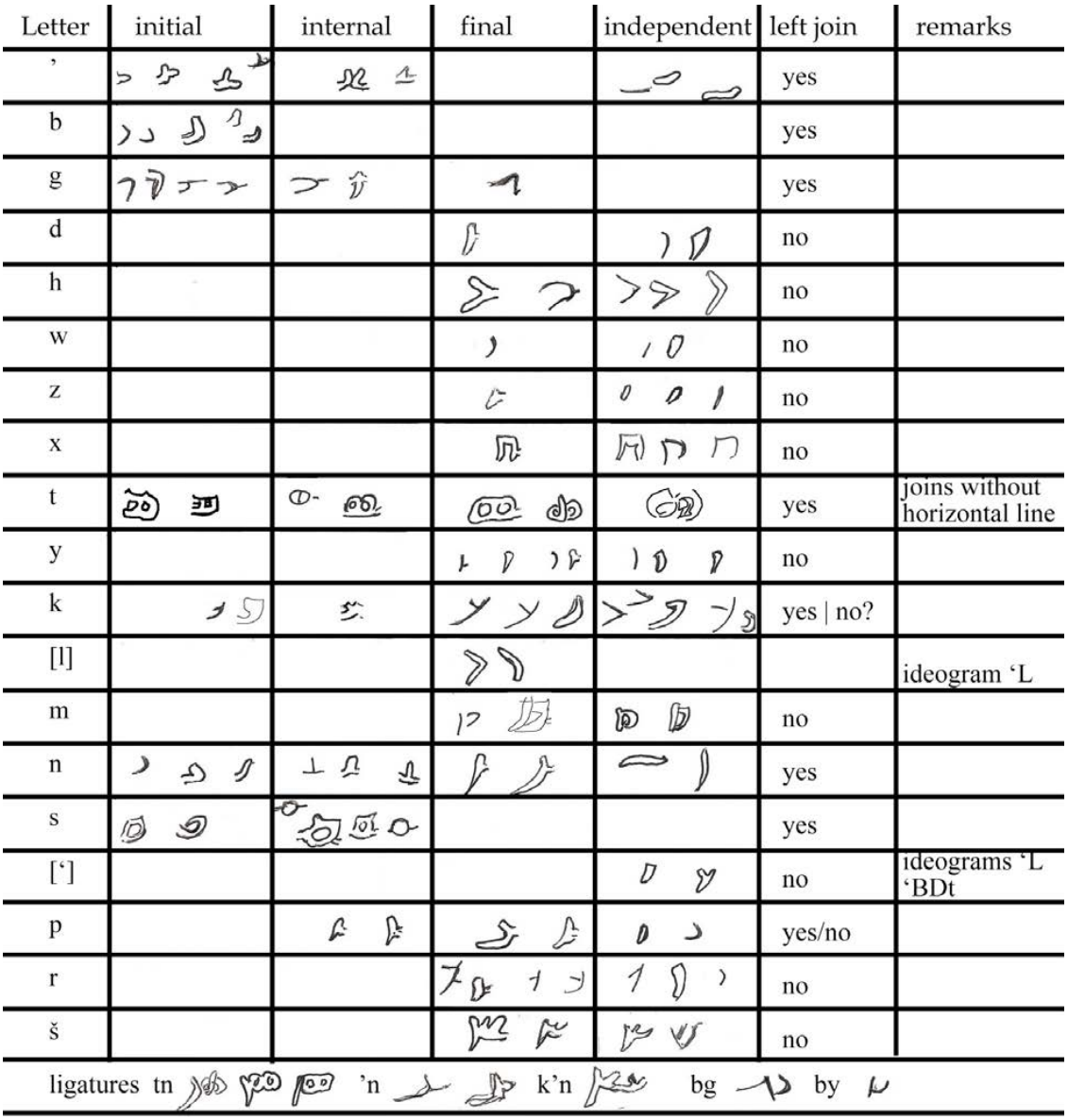
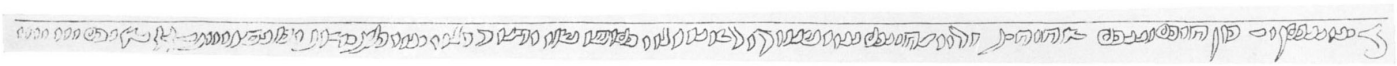


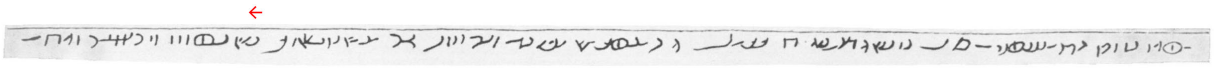
Figure 15: Nominal and positional forms of letters in silver vessel inscriptions (from Lurje 2017).



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

Pl. 1:b *Choresmian No. 1:* A.D. 658. Silver phiale in the British Museum, see pl. 1:a. Diam. 12.7 cm. Photo courtesy the 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.



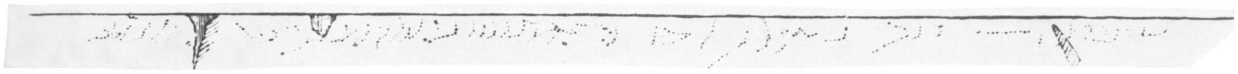
Pl. 3:a *Choresmian No. 2:* A.D. 538 (probably 638). Inscription from silver phiale in the Hermitage Museum, Leningrad, Smirnov, *VS*, pl. XIX: 42.

BŠNT 3 2 100 20 20 20 10 YRXʾ ʾtry BYWM bgy ʾpbntn yʾ MN bwnšrgk-š x(r)kʾn ʿL ʾtršnky bg ZNH ZWZN 20 10 ʾšpynšwk

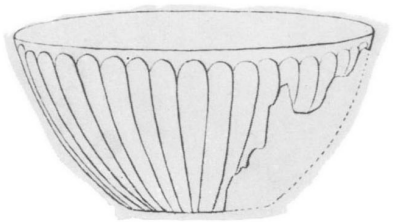
Pl. 3:b *Choresmian No. 2:* A.D. 538 (probably 638). Silver phiale in the Hermitage Museum, Leningrad, see pl. 3:a. Diam. 10.6 cm.

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

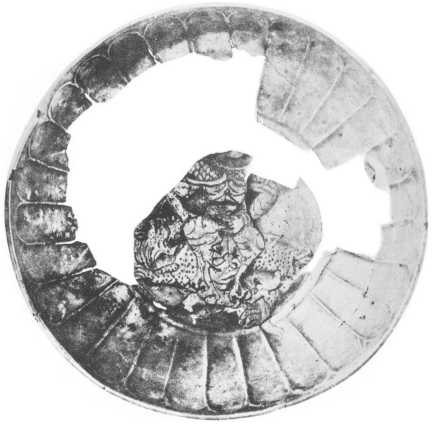
Figure 17: Silver bowl #2: either 538 or 638 CE (from Azarpay 1969: Plate 3:a, b, c). Silver philae in the Hermitage Museum (St. Petersburg). Original from Smirnov 1909, plate XIX:42 and XVIII: 42. Transliteration from Lurje (forthcoming).



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



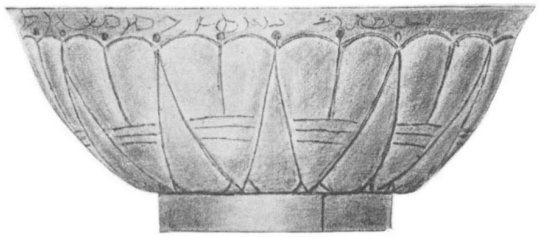
Pl. 5:b *Choresmian No. 3:* silver phiale in the Hermitage Museum, Leningrad, see pl. 5:a, diam. 12.5 cm. Smirnov, *VS*, pl. XVIII: 44.



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.

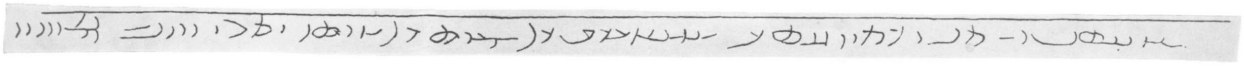
Pl. 8:a
Choresmian No. 4: inscription from silver phiale in the Hermitage Museum, Leningrad, Smirnov, *VS*, pl. XIX:45.

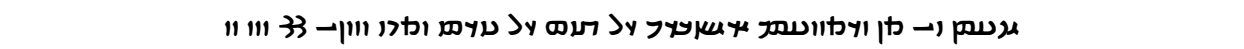


Pl. 8:b *Choresmian No. 4:* silver phiale in the Hermitage Museum, Leningrad, see pl. 8:a. Diam. 12.7 cm.

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.





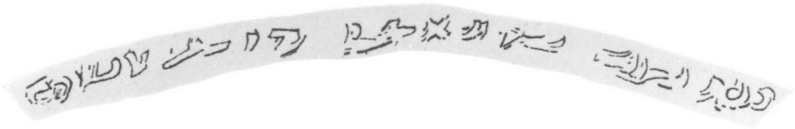
ʾpbntn yʾ MN wrmwzbntk ʾrʾškrk ʿL (g)nyt ʿL byrty zmhy ZWZN-ʾ 20 20 3 2

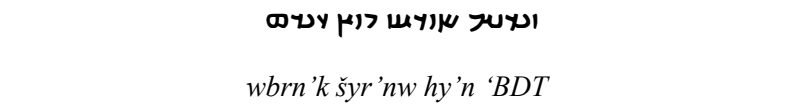
Pl. 11:a, b, c *Choresmian No. 5:* silver phiale in the Hermitage Museum, Leningrad, Smirnov, *VS*, pl. XIX:47. Diam. 13 cm.

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).

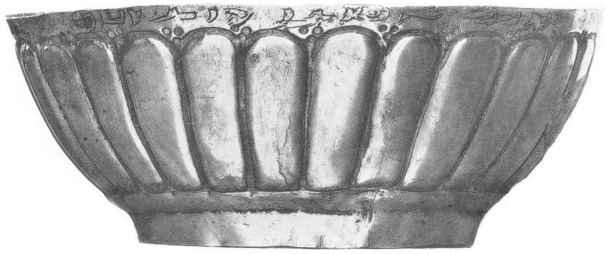
gty (xwpsk | xw ksp) ZWZN-ʾ 4 10 20 20 20 7

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





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.

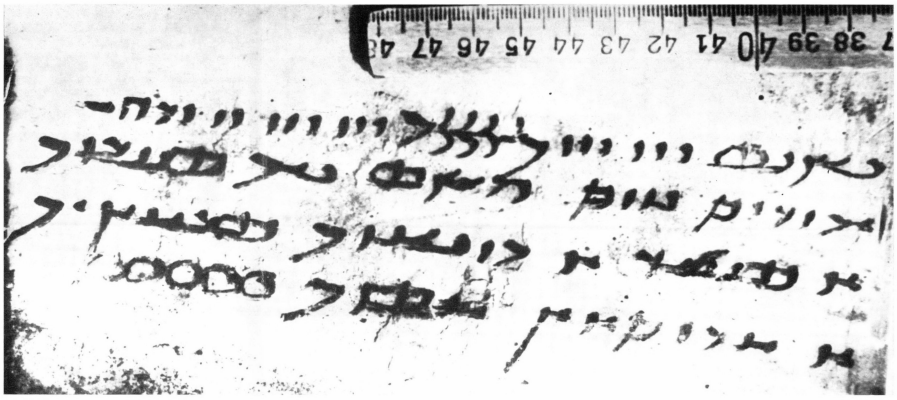


Fig. 1. Inscription No. 52.

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 X III III II YRḤ ›
›hwrym BYWM γwšt ZNH tpnkwk
›y tnβ›r ›y hwnšk (?) t›β›n›n-›-k (?)
›y ›yrw|zm›w›n ›ztyk ○○○○

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

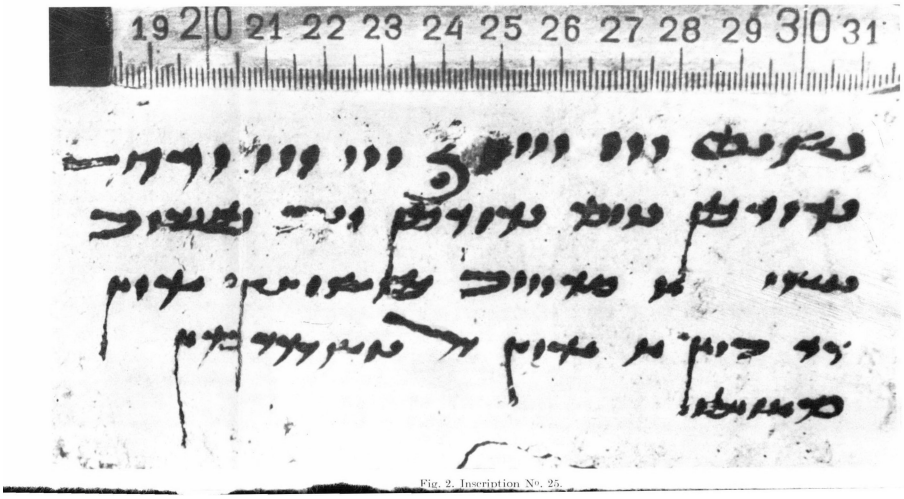


Fig. 2. Inscription No. 25.

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 YRḤ'
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-yōk*, the son of *Tiš-yān*. 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.

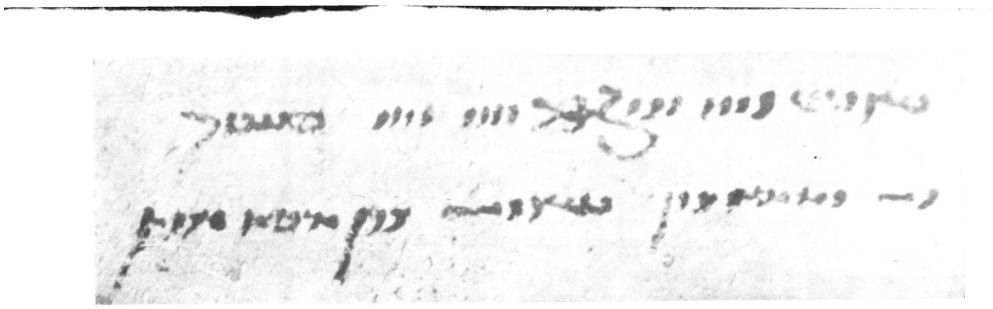


Fig. 3. Inscription N°. 19.

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.

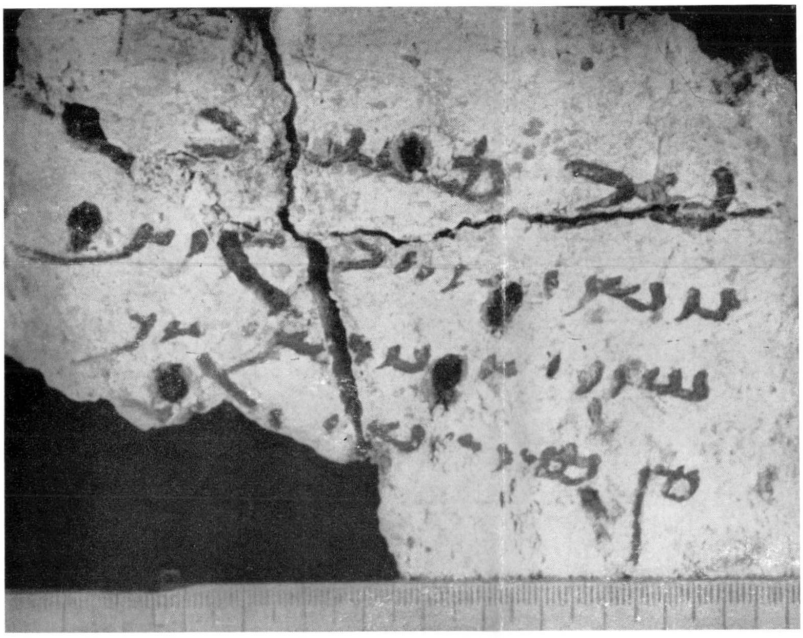


Fig. 4. Inscription No. 8.

ZNH tn[b]ʳrˡyk
’(?)ynšy ʳšhˡk ’.wnʳyˡ.
…’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.

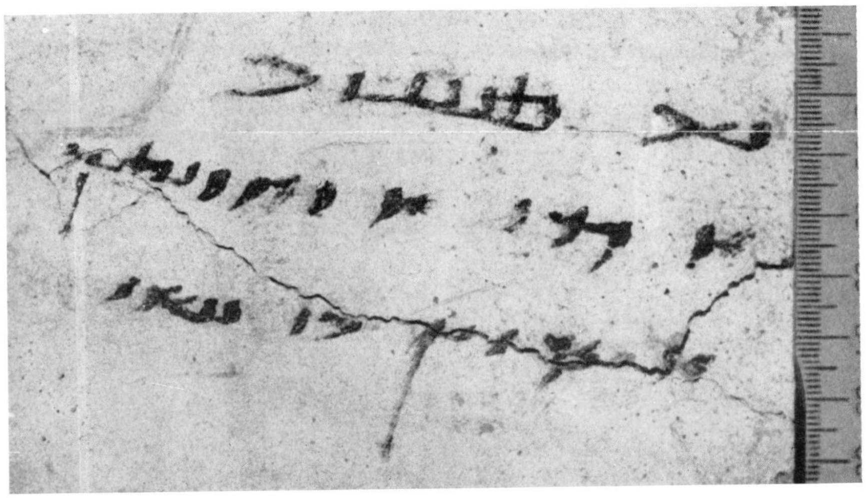


Fig. 5. Inscription No. 69.

𐾸𐾾𐾵 𐿄𐿁𐾾𐾻𐾶𐾻
𐾰𐾺 𐾳𐿂𐾺 𐾰𐾺 𐾶𐾹𐾶𐾾𐿄𐾰𐾾𐾶
𐾰𐾺 𐾰𐿂𐾶𐾰𐾾 𐾵𐾶 𐾾𐿁𐿃𐾺

Tolstov and Livshitz 1964	Reanalysis based upon Henning 1965
ZNH tnbryk *’y gry ’y wḥwntk* *’y ’rw’n kw nwšy*	*ZNH tpnkwk* *ʾy γry ʾy wḥwnt’n-w* *ʾy ʾrwʾn hw NPŠY*
This ossuary contains the body of wnwnxk Soul [his may be sent] to the beautiful [Paradise].	

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

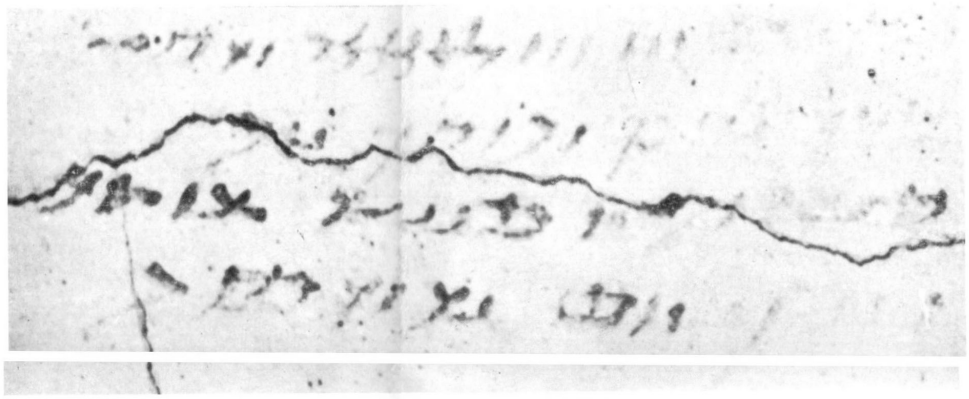


Fig. 6. Inscription N°. 39.

ʳBŠNTˡ III III C XX XX XX X YRḤ’
mʳtry?ˡ ʳBYWˡM whwmn ZNH
tnʳbrˡyk ’y tnb’r ’rw’zd
w … nʳyˡ 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.

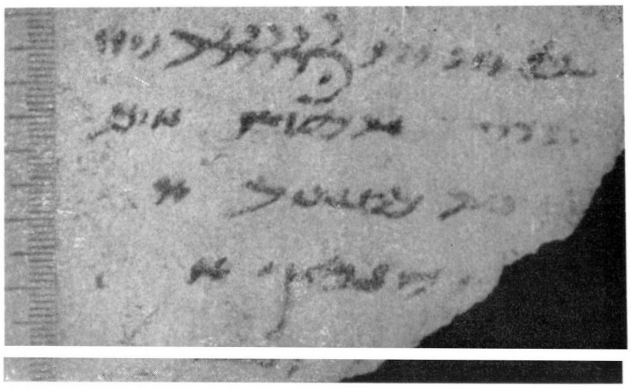
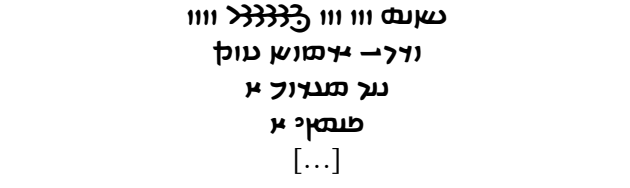


Fig. 7. Inscription No. 12



[…]

ʳBŠʳNT III III C XX XX XX XX X IIII
YRḤ’ ’rtwyš BYWM
]. ZNH tnbryk ’y
]. s|ḫnt’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.

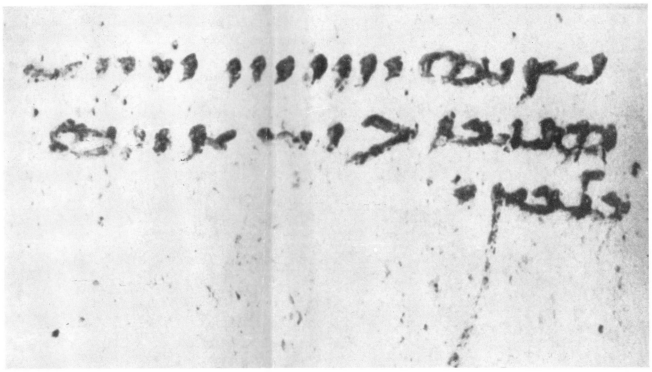
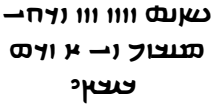


Fig. 8. Inscription No. 26.



Tolstov and Livshitz 1964

BŠNT III IIII YRḤ'
tnbryk y' 'y wrt
k'k'ny

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

Reanalysis based upon Henning 1965

BŠNT III IIII YRḤ'
tpnkwk y' 'y wrt
k'k'n-w

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 N°. 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.

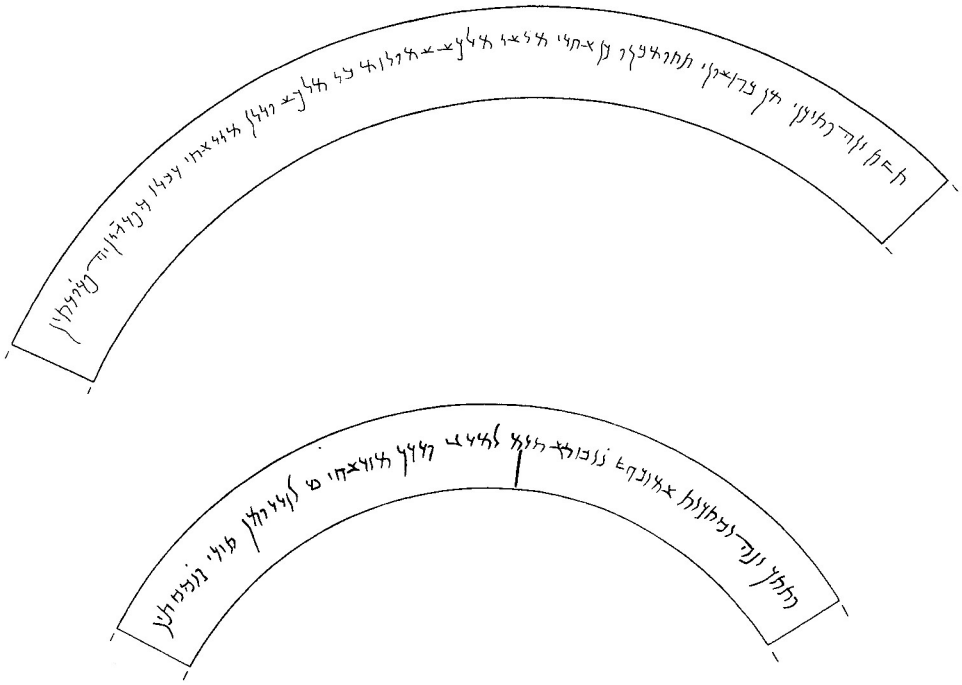
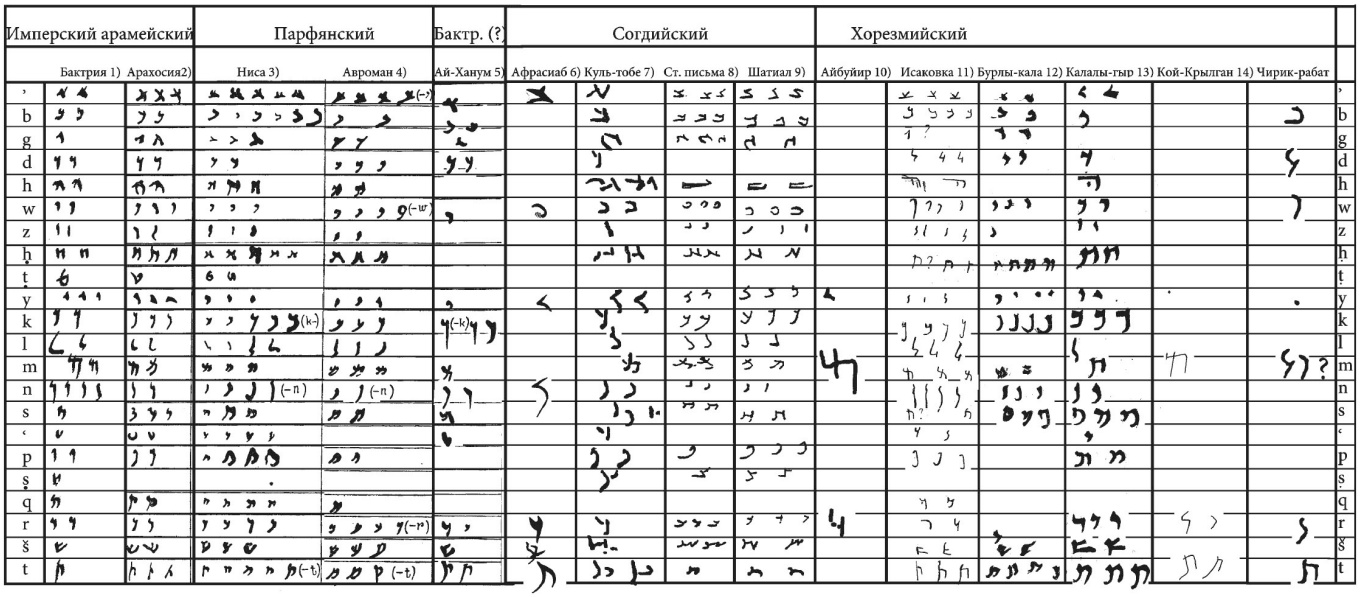


Figure 33: Archaic Chorasmian inscriptions on silver bowls no. 1 (top) and no. 2 (bottom) from Isakovka from the Achaemenid period (from Livshits 2003: 152, 163). This type of script is out of scope for the present encoding, and may be unified with Imperial Aramaic.

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.



1 — *Shaked, Naveh* 2012, Doc. A1;
2–5 В.А. Лившиц по *Расторгуева, Молчанова* 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 10646[1]
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 Unicode*
2. Requester's name: *Anshuman Pandey <pandey@umich.edu>*
3. Requester type (Member body/Liaison/Individual contribution): *Expert contribution*
4. Submission date: *2018-07-26*
5. Requester's reference (if applicable):
6. Choose one of the following:
   This is a complete proposal: *Yes*
   (or) More information will be provided later:

B. Technical – General

1. Choose one of the following:
   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 collection)	
C-Major extinct		D-Attested extinct		E-Minor extinct	
F-Archaic Hieroglyphic or Ideographic				G-Obscure or questionable usage 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 publishing the standard?
   Anshuman Pandey
   b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
   Anshuman Pandey
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 sources) of proposed characters attached? *Yes*
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)? *Yes*

8. Additional Information:

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see Unicode Character Database (http://www.unicode.org/reports/tr44/) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

[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

1. 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, user groups of the script or characters, other experts, etc.)? *Yes*
   If YES, with whom? *Pavel Lurje <pavlvslvra@gmail.com>*
   If YES, available relevant documents:
3. 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 Asian studies*
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely 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)? *Yes*
8. 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:
9. 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) to, or could be confused with, an existing character? *No*
    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) provided? *N/A*
    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: