L2/18-164
2018-04-29

Proposal to encode the Chorasmian script in Unicode

Anshuman Pandey
pandey@umich.edu

April 29, 2018

1 Introduction

This proposal is a revision of the following:

- L2/18-010R: "Proposal to encode the Khwarezmian 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

Major changes from L2/18-010R include:

- Change of script identifier from 'Khwarezmian' to 'Chorasmian'
- Renaming of DETACHED ALEPH to SMALL ALEPH
- Renaming of VOCALIZATION SIGN to CURLED WAW
- Improvements to the glyphs for *waw*, *zayin*, and *yodh*
- Additional details on joining properties and shaping behaviors

2 Background

The proposed script was used between the 2nd and 9th centuries CE for writing Chorasmian (ISO 639-3: xco), a now-extinct Eastern Iranian language. The script and languge was used in Chorasmia, a region in Central Asia that was situated at the Oxus (Amu Darya) river delta, which is today spread across Uzbekistan, Kazakhstan, and Turkmenistan. The territory is known in the *Avesta* as 𐬓𐬂𐬌𐬭𐬌𐬰𐬆𐬨 *hvâirizem* (*Yašt* 10.5.14). The Achaemenids referred to it as 𐎢𐎺𐎠𐎼𐏀𐎷𐎡𐏁 *uvârazmiš* (XPh inscription, lines 21–22). 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 three that were used for recording the Chorasmian language. These are: 1) the Imperial Aramaic script; 2) the indigenous script discussed here, which is derived from Imperial

Aramaic ; and 3) the Arabic script. The indigenous script is classified by scholars into 'archaic', 'lapidary', and 'cursive' types (Pavel Lurje, personal communication, December 2017):

- The 'archaic' occurs, for instance, on silver bowls no. 1 and no. 2 from Isakovka (Исаковка), see fig. 33 here. These 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). This type is a non-joining *abjad* (see fig. 35).

- The 'lapidary' is represented, for example, 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-калы). This type is also a non-joining *abjad* (see fig. 35).

The 'cursive' Chorasmian script is a development of the lapidary style. While it has the same repertoire of letters, the use of cursive practices for joining letters of a word gave the script distinctive graphical and structural features, and rules for connecting letters in order to maintain distinctions between letters with similar shapes. The 'cursive' script 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 of Chorasmian text in the 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).

The Chorasmian script of these sources 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 10th century. Some insight into the demise of the script, among other aspects of Chorasmian culture and society, is provided by the medieval scholar named 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 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 proposed identifier for the script in Unicode is ‘Chorasmian’. 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 use ‘Chorasmian’ to refer to the pre-Islamic period and ‘Khwarezmian’ for 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 in the names list.

4 Approach to the encoding

The encoding is based upon the cursive Chorasmian script as attested by inscriptions on coinage, silver vessels, and ossuaries. This script differs from the archaic and lapidary type described above, which are non-joining scripts that may be unified with the Imperial Aramaic encoding.

Of the 22 letters of the Aramaic alphabet, 19 are attested collectively across the relevant sources. Analogues for *teth*, *qoph*, and *sadhe* do not exist. 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.

It is practical to consider the styles used on coins, ossuaries, and other materials as developmental phases of a distinctive ‘Chorasmian’ script. For purposes of character encoding the varieties should be unified as a single

[1] Most 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).

'Chorasmian' script in Unicode. This approach enables texts to be represented using the same underlying character set, using normative glyphs. The display of the script style of particular records would be managed by custom fonts.

The representative glyphs are based upon the letterforms in the Tok Kala ossuaries. The style used in these sources are the latest development of the script and reflect its distinctiveness. A peculiarity of Chorasmian is that several letters resemble each other in the latest stage of the script. 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 such pairs in Chorasmian. These distinctions have been preserved in the proposed repertoire.

5 Proposed repertoire

The proposed repertoire for Chorasmian contains 29 characters: 22 letters and 7 numbers. The code chart and names list follows p. 11. 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. 𐡀 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.

5.1 Letters

Character name	Glyph	Joining	Latin
CHORASMIAN LETTER ALEPH	𐡀	dual	ʾ
CHORASMIAN LETTER SMALL ALEPH	𐡁	non	-ʾ
CHORASMIAN LETTER FINAL ALEPH	𐡂	non	ʾ
CHORASMIAN LETTER BETH	𐡃	dual	*b*
CHORASMIAN LETTER GIMEL	𐡄	dual	*g*
CHORASMIAN LETTER DALETH	𐡅	right	*d*
CHORASMIAN LETTER HE	𐡆	right	*h*
CHORASMIAN LETTER WAW	𐡇	right	*w*

CHORASMIAN LETTER CURLED WAW	𐾷	non	*-w*
CHORASMIAN LETTER ZAYIN	𐾸	dual*	*z*
CHORASMIAN LETTER HETH	𐾹	right	*ḥ*
CHORASMIAN LETTER YODH	𐾺	right	*y*
CHORASMIAN LETTER KAPH	𐾻	dual	*k*
CHORASMIAN LETTER LAMEDH	𐾼	dual	*l*
CHORASMIAN LETTER MEM	𐾽	right	*m*
CHORASMIAN LETTER NUN	𐾾	dual	*n*
CHORASMIAN LETTER SAMEKH	𐿀	dual	*s*
CHORASMIAN LETTER AYIN	𐿁	right	ʿ
CHORASMIAN LETTER PE	𐿂	dual*	*p*
CHORASMIAN LETTER RESH	𐿃	right	*r*
CHORASMIAN LETTER SHIN	𐿄	right	*š*
CHORASMIAN LETTER TAW	𐿅	dual	*t*

aleph The *aleph* is rendered in various ways:

- Initial When word-initial *aleph* is followed by a letter with a baseline, it connects at the baseline, eg. 𐾰𐿂𐾲𐾾𐿅𐾿 *ʾpbntn*. With other letters it may connect to the midpoint of the right edge. Examination of the occurrences of *aleph* indicates that it consistently connects in this manner to letters with a single, vertical stroke, eg. *waw*, *yodh*, final *nun*. It also exhibits such behavior with the non-connecting variants of *zayin* and *pe*, as well as angled letters such as *gimel* and *he*, eg. 𐾰𐾸𐿅𐾺𐾻 *ʾztyk*.

- Medial Word-medial *aleph* generally connects at the baseline, eg. 𐾻𐾰𐾻𐾰𐾾𐾷 *kʾkʾn-w*, 𐿂𐿃𐾰𐾾𐾺𐾰𐿅𐾺 *prʾnyʾty*. It connects to following letters according to the rules for the initial form, eg. 𐿅𐾾𐾲𐾰𐿃 *tnbʾr*, 𐾳𐿃𐾴𐾽𐾰𐾿 *grdm ʾn*. Medial *aleph* occurs as 𐾰 in some coins and vessels instead of as 𐾰. This 𐾰 is to be treated as a glyphic variant.

- Final Word-final *aleph* is written as an elongated stroke – that does not connect to the preceding letter, eg. 𐾺𐿃𐾹– *YRḤʾ* (TK 25, fig. 25). Moreover, final *aleph* triggers the rendering of a preceding letter in its final form. For example, *nun* is written as 𐾿 before – in 𐾸𐾶𐾸𐾿– *ZWZN-ʾ*. Given the behavior of – for representing word-final *aleph*, it is encoded as the separate letter FINAL ALEPH. There is no attestation for a connected final *aleph*; however, an artificial glyph for this contextual form may be needed in order to prevent issues with rendering the occurrence of the dual-joining ALEPH instead of

the non-joining FINAL ALEPH. The glyph ʌ may be used as the final form of ALEPH.

- ‘Small’ *aleph* A smaller and raised non-joining form ˑ of *aleph* is used for representing a possessive. It is attested in the ossuary inscriptions. This form occurs in medial and final positions. eg. טכנאןˑך *tknpʾn-ʾ-k* (TK 52, fig. 24); ˑויאן *hyʾn-ʾ* (TK 26). Given the behavior of ˑ it is encoded as the separate letter SMALL ALEPH. As there is no native or scholarly term for this letter, a name based upon its shape has been selected.

beth, ***nun***, ***pe*** The letters ב *beth*, נ *nun*, פ *pe* are often written using a shape similar to נ in medial position. Their initial forms are distinguished by the degree of curvature of the primary stroke.

gimel The ד *gimel* connects to the left at the top, eg. *gimel* + *waw* as ח in the name of a day, חושט *gwšt*; *gimel* + *resh* as רר in ררדמאן *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 ד in silver vessels.

gimel, ***he***, ***kaph*** The letters ד *gimel*, ה *he*, כ *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, while the *he* consistently appears with a descending top stroke and angled bottom stroke. The distinctiveness of *gimel* is evidenced by the archaic form ד used in silver vessels. They 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 ו *waw* in חושט *gwšt* and הונשך *hwnšk* (TK 52, fig. 24), and כואני *kwʾn-y* (TK 25, fig. 25).

daleth, ***ayin***, ***resh*** The letters ד *daleth*, ע *ayin*, ר *resh* have the same basic structure, but are differentiated in terms of their shapes. The *daleth* has a shorter primary stroke than *resh* and a wider top angle than *ayin*. The *resh* and *daleth* are differentiated by the length of the primary stroke, with that of ר *resh* being longer than that of ד *daleth* as shown in ררדמאן *grdm’n* (TK no. 25, fig. 25). Such a distinction appears to be carried over from Imperial Aramaic, where ד *daleth* and ר *resh* differ by the length of the primary stroke. The *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 on the right stroke, eg. עבדט *ʿBDt*.

waw The *waw* has two representations:

- Conventional *waw* The letters ו *waw* and י *yodh* are difficult to distinguish in various sources and they have the same joining properties. However, *yodh* has a notched head, while *waw* is typically a vertical stroke, either straight or slightly curved. The two may be written similar to ז *zayin*, but the latter is often distinguished by a curved terminal and its behavior of joining to the left in heterograms that occur in certain sources.

- ‘Curled’ *waw* In TK no. 25 and no. 52, when *waw* appears in word-final position and indicates a possessive, it is written using the non-joining sign ᵓ (transliterated as *-w*). It is described by Henning as a “vocalization mark”, which is “a rounded form reminiscent of an Arabic *ḍamm*” and occurs in words, such as טישיאנאנᵓ *tyšyʾnʾn-w* in TK no. 25 and וחונטאנᵓ *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 separate letter CURLED WAW. As there is no native or scholarly term for this letter, a name based upon its shape has been selected.

zayin The joining behavior of *zayin* differs depending upon context. In some Aramaic heterograms it joins to the left, eg. 𐾸𐾾𐾵 *ZNH*, while in others it does not, 𐾸𐾶𐾸𐾾 *ZWZN*. In Chorasmian words, it does not join to the left. For this reason, *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 𐾸. When ZWNJ is used, the letter will be rendered as 𐾸.

nun Word-final 𐾾 *nun* is written as 𐾾 in the ossuary inscriptions. In coinage and silver vessels, the final form appears as 𐾾. The orientation of the tail may be guided by the vertical constraints of the inscription, ie. being placed on the lip of a vessel or at the edge of a coin. For example, 𐾸𐾶𐾸𐾾𐾰 *ZWZN-* ʾ occurs on silver vessels, and would be normalized as 𐾸𐾶𐾸𐾾𐾰.

ayin In the available sources this letter occurs in 𐿀𐾲𐾴𐿄 ʿ*BDt* ʻdone, made' (silver bowl #7, fig. 22) and 𐿀𐾼 ʿ*L*, an Aramaic heterogram for 'to' (various silver bowls and ossuaries). It is only attested in word-initial position. While it appears to join to *lamedh* in 𐿀𐾼, the connection is likely a result of letter spacing, not a cursive property of the letter. Given its similar structure to *daleth* and *resh*, it is likely that *ayin* is right-joining. If evidence indicates that it is non-joining, then its property may be modified in the future.

pe Although it is a default dual-joining letter, in some sources medial 𐿁 PE does not connect to the left, compare 𐾰𐿁𐾲𐾾𐿄𐾾 ʾ*pbntn* with 𐾰𐿁‌𐾲𐾾𐿄𐾾 ʾ*p-bntn*. Such usage is unpredictable and appears to be a scribal convention. As there is no feasible means for selecting optional connections for a letter, it is necessary to define *pe* as a dual-joining letter. The ZWNJ may be used to break the connection by placing it after *pe*.

shin The letter 𐿃 is represented using the glyphic variant 𐿃 on some coins.

taw In ossuary inscriptions letters that follow 𐿄 *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. 𐿄 is written as 𐿄 to accommodate a following letter. For instance, *taw* + *yodh* is 𐿄𐾺 as in 𐾰𐾸𐿄𐾺𐾻 ʾ*ztyk* on TK no. 52 (fig. 24); and *taw* + final *nun* as 𐿄𐾾 in 𐾰𐿁𐾲𐾾𐿄𐾾 ʾ*pbntn*. In coins, the point of connection occurs at the head, where the top-stroke of *taw* is extended into that of the following letter, eg. *taw* + *waw* is written 𐿄𐾶 *tw*, as in the name 𐿄𐾶𐿄𐾶𐾹𐾿 *twtwxs* on type BI coins (fig. 8). The *taw* has a glyphic variant form 𐿄 that has an open right stroke.

5.2 Numbers

Character name	Glyph	Joining	Value
CHORASMIAN NUMBER ONE	𐿅	non	1
CHORASMIAN NUMBER TWO	𐿆	non	2
CHORASMIAN NUMBER THREE	𐿇	non	3
CHORASMIAN NUMBER FOUR	𐿈	non	4
CHORASMIAN NUMBER TEN	𐿉	right	10
CHORASMIAN NUMBER TWENTY	𐿊	dual	20
CHORASMIAN NUMBER ONE HUNDRED	𐿋	left	100

Primary units The primary units are expressed using repetitions of the sign 𐾁, which is a non-joining character. The numbers 5–9 are written using sequences of ONE arranged in groups containing three or four instances of 𐾁. See, for example, '𐾂 𐾂𐾁' for 7 and '𐾂𐾁 𐾂𐾁' for 8 in TK no. 19 (fig. 26), as well as '𐾁 𐾂 𐾂' for 7 in TK no. 25 (fig. 25). Also, '𐾁 𐾂' for the number 5 in silver bowl #2 (fig. 17). The number 5 in silver bowl #5 appears as '𐾂𐾁' without a spaced grouping, but the extended terminal of the third 𐾁 suggests the intended grouping '𐾁 𐾂' despite lack of spacing (fig. 20). Given the grouping behavior of 𐾁, the numbers 𐾁 ONE .. 𐾂𐾁 FOUR are encoded atomically. This follows the encoding for 'Imperial Aramaic'.

Ten The 𐾃 TEN resembles a vertically compressed 𐾄 LAMEDH. It is a right-joining character.

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

Hundreds The number 100 is written using 𐾆 ONE HUNDRED. The ONE HUNDRED also functions as a unit mark for the hundreds. Multiples of hundred are indicated using primary numbers placed before ONE HUNDRED. It is a left-joining character.

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

6 Script details

6.1 Structure

Chorasmian is a cursive joining *abjad*. It is written from right to left, with lines that advance from top to bottom. Letters have nominal shapes when they occur in isolation and contextual shapes when they occur in initial, medial, or final position. All letters are defined as dual joining. In some sources the connection between letters is suspended (see § 7). This feature may be supported by usage of the control character ZWNJ U+200C ZERO WIDTH NON-JOINER (abbreviated as ZWNJ).

6.2 Punctuation

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

6.3 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.

6.4 Collation

The sort order of the letters follows the encoded order:

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

7 Joining behavior

The shaping requirements for Chorasmian are similar to that of Arabic. A summary of the joining properties of the 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*, *ayin*, *resh*, *shin*
exceptions	*aleph*, *zayin*, *pe*

When a letter is described as joining another letter to the left, it is implied that the joining occurs only if the following letter is right-joining. Equally, a letter described as joining to the right implies that the preceding letter is left-joining. Otherwise, no connections are made between the letters. In the tables below, the nominal form of a letter is given in the 'X_n' column. The labels 'X_i', 'X_m', 'X_f' refer to their contextual initial, medial, and final forms. A red dash indicates the location on a glyph were connections should occur, while a vertical bar indicates that a connection occurs without an extension of the baseline.

Joining features of dual-joining letters are shown below:

	X_n	X_f	X_m	X_i
ALEPH	𐾀	[𐾀]	(𐾁) 𐾀	(𐾁) 𐾀
BETH	𐾃	𐾃	𐾃	𐾃
GIMEL	𐾄	𐾄	𐾄	𐾄
ZAYIN	𐾉	𐾉	𐾉 , 𐾇	𐾉 , 𐾇
KAPH	𐾌	𐾌	𐾃	𐾃
LAMEDH	𐾍	𐾍	𐾍	𐾍
NUN	𐾏	𐾏	𐾏	𐾏
SAMEKH	𐾐	𐾐	𐾐	𐾐
PE	𐾒	𐾃 , 𐾇	𐾇 , 𐾃	𐾃
TAW	𐾕	𐾕	𐾕	𐾕

Joining features of right-joining letters are shown below. Note that the letters *he* and *mem* have word-medial forms that differ from the final forms.

	X_n	X_f	X_m	X_i
DALETH	𐾙	𐾙	—	𐾙
HE	𐾚	𐾚	𐾚	𐾚
WAW	𐾛	𐾛	—	𐾛
HETH	𐾜	𐾜	—	𐾜
YODH	𐾝	𐾝	—	𐾝
MEM	𐾞	𐾞	𐾞	𐾞
AYIN	𐾟	𐾟	—	𐾟
RESH	𐾠	𐾠	—	𐾠
SHIN	𐾡	𐾡	—	𐾡

The left-side connection of the letters *zayin* and *pe* may be suspended in some cases. The default joining behavior may be modified using the generic control character ZWNJ U+200C ZERO WIDTH NON-JOINER (abbreviated as ZWNJ).

The joining features for the numbers are:

	X_n	X_f	X_m	X_i	Join
ONE	𐾩	—	—	—	non
TWO	𐾪	—	—	—	non
THREE	𐾫	—	—	—	non
FOUR	𐾬	—	—	—	non
TEN	𐾭	𐾭	—	𐾭	right
TWENTY	𐾮	𐾮	𐾮	𐾮	dual
ONE HUNDRED	𐾯	𐾯	—	𐾯	left

7.1 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). This reading of the second letter as *waw* forced an analysis of the right-joining letter as being potentially left-joining in some cases in the Tok Kala texts. 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 convetions of the script and eliminates the need to complicate the behavior of *waw*.

Similarly, the word 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 reanalysis of the fifth word as medial *kaph* instead of *resh* presented a better option. This approach allowed a more accurate reading of the word as *tpnkwk*. Yet another issue was the joining behaviors of *he* and *kaph*.

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 *k'k'n-w* (TK no. 26, fig. 31). Following the behavior of *kaph* in the latter, reading 𐾵𐾶 as *hw* provides for more accurate results.

Such issues are typical in initial attempts at decipherment. At this point in time, the joining behaviors of Chorasmian letters have been determined.

8 Encoded representations

8.1 Examples of usage

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:

'g'dk	𐾰𐾳𐾰𐾴𐾻	<𐾰 ALEPH, 𐾳 GIMEL, 𐾰 ALEPH, 𐾴 DALETH, 𐾻 KAPH>
'pbntn	𐾰𐿁𐾲𐾾𐿄𐾾	<𐾰 ALEPH, 𐿁 PE, 𐾲 BETH, 𐾾 NUN, 𐿄 TAW, 𐾾 NUN>
'p-bntn	𐾰𐿁‌𐾲𐾾𐿄𐾾	<𐾰 ALEPH, 𐿁 PE, ZWNJ, 𐾲 BETH, 𐾾 NUN, 𐿄 TAW, 𐾾 NUN>
'ztyk	𐾰𐾸𐿄𐾺𐾻	<𐾰 ALEPH, 𐾸 ZAYIN, 𐿄 TAW, 𐾺 YODH, 𐾻 KAPH>
'r'škrk	𐾰𐿂𐾰𐿃𐾻𐿂𐾻	<𐾰 ALEPH, 𐿂 RESH, 𐾰 ALEPH, 𐿃 SHIN, 𐾻 KAPH, 𐿂 RESH, 𐾻 KAPH>
'špynšwk	𐾰𐿃𐿁𐾺𐾾𐿃𐿄𐾻	<𐾰 ALEPH, 𐿃 SHIN, 𐿁 PE, 𐾺 YODH, 𐾾 NUN, 𐿃 SHIN, 𐿄 TAW, 𐾻 KAPH>

bntk	𐾲𐾾𐿄𐾻	<𐾲 BETH, 𐾾 NUN, 𐿄 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, ‌ ZWNJ, 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>
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, 𐾰 FINAL ALEPH>
ḥwsrw	𐾹𐾶𐾿𐿂𐾶	<𐾹 HETH, 𐾶 WAW, 𐾿 SAMEKH, 𐿂 RESH, 𐾶 WAW>
ḥwpsk	𐾹𐾶𐿁𐾿𐾻	<𐾹 HETH, 𐾶 WAW, 𐿁 PE, 𐾿 SAMEKH, 𐾻 KAPH>
YRḤʾ	𐾺𐿂𐾹𐾰	<𐾺 YODH, 𐿂 RESH, 𐾹 HETH, 𐾰 FINAL ALEPH>
kʾkʾn-w	𐾻𐾰𐾻𐾾𐾷	<𐾻 KAPH, 𐾰 ALEPH, 𐾻 KAPH, 𐾾 NUN, 𐾷 CURLED WAW>
KSP	𐾻𐾿𐿁	<𐾻 KAPH, 𐾿 SAMEKH, 𐿁 PE>
MLKʾ	𐾽𐾼𐾻𐾰	<𐾽 MEM, 𐾼 LAMEDH, 𐾻 KAPH, 𐾰 FINAL ALEPH>
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>

8.2 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 numbers 5–9 may be represented as shown below. Some numbers have more than one representations, as attested in the available sources:

5	𐾪 𐾩	<𐾩 THREE, 𐾪 TWO>
6	𐾩 𐾩	<𐾩 THREE, 𐾩 THREE>
7	𐾩 𐾨	<𐾨 FOUR, 𐾩 THREE>
	𐾧 𐾩 𐾩	<𐾩 THREE, 𐾩 THREE, 𐾧 ONE>
8	𐾨 𐾨	<𐾨 FOUR, 𐾨 FOUR>
9	𐾩 𐾩 𐾩	<𐾩 THREE, 𐾩 THREE, 𐾩 THREE>

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>

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	𐾭𐾪	<𐾪 TWO, 𐾭 ONE HUNDRED>

300 𐾳𐾰 <𐾰 THREE, 𐾳 ONE HUNDRED>

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

570 𐾲𐾲𐾲𐾳 𐾱 𐾰 <𐾱 TWO, 𐾰 THREE, 𐾳 ONE HUNDRED, 𐾲 TWENTY, 𐾲 TWENTY, 𐾲 TWENTY, 𐾴 TEN>

678 𐾱 𐾰 𐾰 𐾲𐾲𐾲𐾳 𐾰 𐾰 <𐾰 THREE, 𐾰 THREE, 𐾳 ONE HUNDRED, 𐾲 TWENTY, 𐾲 TWENTY, 𐾲 TWENTY, 𐾴 TEN, 𐾱 TWO, 𐾰 THREE, 𐾰 THREE>

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 FINAL ALEPH;Lo;0;R;;;;N;;;;;
10FB3;CHORASMIAN LETTER BETH;Lo;0;R;;;;N;;;;;
10FB4;CHORASMIAN LETTER GIMEL;Lo;0;R;;;;N;;;;;
10FB5;CHORASMIAN LETTER DALETH;Lo;0;R;;;;N;;;;;
10FB6;CHORASMIAN LETTER HE;Lo;0;R;;;;N;;;;;
10FB7;CHORASMIAN LETTER WAW;Lo;0;R;;;;N;;;;;
10FB8;CHORASMIAN LETTER CURLED WAW;Lo;0;R;;;;N;;;;;
10FB9;CHORASMIAN LETTER ZAYIN;Lo;0;R;;;;N;;;;;
10FBA;CHORASMIAN LETTER HETH;Lo;0;R;;;;N;;;;;
10FBB;CHORASMIAN LETTER YODH;Lo;0;R;;;;N;;;;;
10FBC;CHORASMIAN LETTER KAPH;Lo;0;R;;;;N;;;;;
10FBD;CHORASMIAN LETTER LAMEDH;Lo;0;R;;;;N;;;;;
10FBE;CHORASMIAN LETTER MEM;Lo;0;R;;;;N;;;;;
10FBF;CHORASMIAN LETTER NUN;Lo;0;R;;;;N;;;;;
10FC0;CHORASMIAN LETTER SAMEKH;Lo;0;R;;;;N;;;;;
10FC1;CHORASMIAN LETTER AYIN;Lo;0;R;;;;N;;;;;
10FC2;CHORASMIAN LETTER PE;Lo;0;R;;;;N;;;;;
10FC3;CHORASMIAN LETTER RESH;Lo;0;R;;;;N;;;;;
10FC4;CHORASMIAN LETTER SHIN;Lo;0;R;;;;N;;;;;
10FC5;CHORASMIAN LETTER TAW;Lo;0;R;;;;N;;;;;
10FC6;CHORASMIAN NUMBER ONE;No;0;R;;;;1;N;;;;;
10FC7;CHORASMIAN NUMBER TWO;No;0;R;;;;2;N;;;;;
10FC8;CHORASMIAN NUMBER THREE;No;0;R;;;;3;N;;;;;
10FC9;CHORASMIAN NUMBER FOUR;No;0;R;;;;4;N;;;;;
10FCA;CHORASMIAN NUMBER TEN;No;0;R;;;;10;N;;;;;
10FCB;CHORASMIAN NUMBER TWENTY;No;0;R;;;;20;N;;;;;
10FCC;CHORASMIAN NUMBER ONE HUNDRED;No;0;R;;;;100;N;;;;;
```

9.2 Linebreak data: `LineBreak.txt`

```
10FB0..10FC5;AL   # Lo  [22] CHORASMIAN LETTER ALEPH..CHORASMIAN LETTER TAW
10FC6..10FCC;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; N; No_Joining_Group
10FB2; CHORASMIAN FINAL ALEPH; N; No_Joining_Group
10FB3; CHORASMIAN BETH; D; No_Joining_Group
10FB4; CHORASMIAN GIMEL; D; No_Joining_Group
10FB5; CHORASMIAN DALETH; R; No_Joining_Group
10FB6; CHORASMIAN HE; R; No_Joining_Group
10FB7; CHORASMIAN WAW; R; No_Joining_Group
10FB8; CHORASMIAN CURLED WAW; N; No_Joining_Group
10FB9; CHORASMIAN ZAYIN; R; No_Joining_Group
10FBA; CHORASMIAN HETH; R; No_Joining_Group
10FBB; CHORASMIAN YODH; R; No_Joining_Group
10FBC; CHORASMIAN KAPH; D; No_Joining_Group
10FBD; CHORASMIAN LAMEDH; D; No_Joining_Group
10FBE; CHORASMIAN MEM; R; No_Joining_Group
10FBF; CHORASMIAN NUN; D; No_Joining_Group
10FC0; CHORASMIAN SAMEKH; D; No_Joining_Group
10FC1; CHORASMIAN AYIN; R; No_Joining_Group
10FC2; CHORASMIAN PE; D; No_Joining_Group
10FC3; CHORASMIAN RESH; R; No_Joining_Group
10FC4; CHORASMIAN SHIN; R; No_Joining_Group
10FC5; CHORASMIAN TAW; D; No_Joining_Group
10FC6; CHORASMIAN ONE; N; No_Joining_Group
10FC7; CHORASMIAN TWO; N; No_Joining_Group
10FC8; CHORASMIAN THREE; N; No_Joining_Group
10FC9; CHORASMIAN FOUR; N; No_Joining_Group
10FCA; CHORASMIAN TEN; R; No_Joining_Group
10FCB; CHORASMIAN TWENTY; D; No_Joining_Group
10FCC; 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`

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

Federov, 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. 2003. "Three silver bowls from the Isakovka burial-ground no. 1 with Khwarezmian and Parthian inscriptions". *Ancient Civilizations*, 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".

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

———. 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.

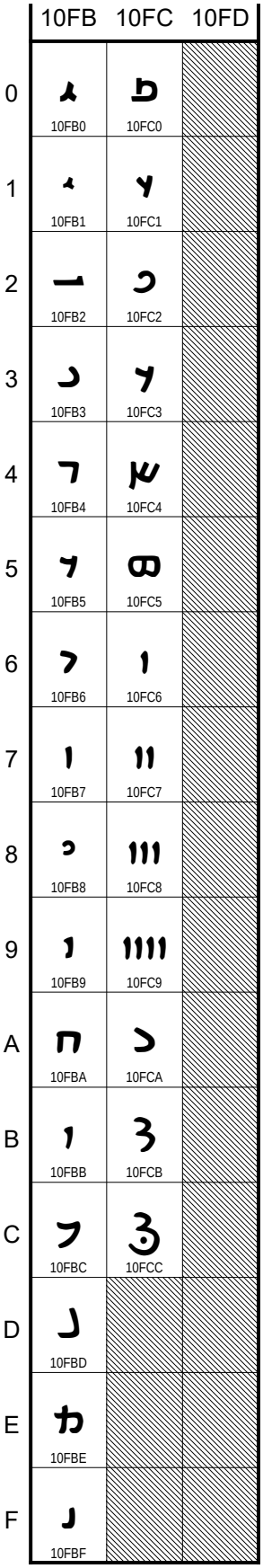
Tolstov, S. P. and Livshitz, V. A. 1964. "Decipherment and Interpretation of the Khwarezmian Inscriptions from Tok Kala", *Acta Antiqua Academiae Scientiarum Hungaricae*, vol. 12, pp. 231–251.

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.



Also known as 'Khwarezmian'.

Letters

10FB0 CHORASMIAN LETTER ALEPH
10FB1 CHORASMIAN LETTER SMALL ALEPH
10FB2 CHORASMIAN LETTER FINAL ALEPH
10FB3 CHORASMIAN LETTER BETH
10FB4 CHORASMIAN LETTER GIMEL
10FB5 CHORASMIAN LETTER DALETH
10FB6 CHORASMIAN LETTER HE
10FB7 CHORASMIAN LETTER WAW
10FB8 CHORASMIAN LETTER CURLED WAW
10FB9 CHORASMIAN LETTER ZAYIN
10FBA CHORASMIAN LETTER HETH
10FBB CHORASMIAN LETTER YODH
10FBC CHORASMIAN LETTER KAPH
10FBD CHORASMIAN LETTER LAMEDH
10FBE CHORASMIAN LETTER MEM
10FBF CHORASMIAN LETTER NUN
10FC0 CHORASMIAN LETTER SAMEKH
10FC1 CHORASMIAN LETTER AYIN
10FC2 CHORASMIAN LETTER PE
10FC3 CHORASMIAN LETTER RESH
10FC4 CHORASMIAN LETTER SHIN
10FC5 CHORASMIAN LETTER TAW

Numbers

10FC6 CHORASMIAN NUMBER ONE
10FC7 CHORASMIAN NUMBER TWO
10FC8 CHORASMIAN NUMBER THREE
10FC9 CHORASMIAN NUMBER FOUR
10FCA CHORASMIAN NUMBER TEN
10FCB CHORASMIAN NUMBER TWENTY
10FCC CHORASMIAN NUMBER ONE HUNDRED

	Chorasmian	Old Sogdian	Inscriptional Pahlavi	Inscriptional Parthian	Imperial Aramaic
aleph	𐾡	𐼁	𐭠	𐭡	𐡀
beth	𐾢	𐼂	𐭡	𐭢	𐡁
gimel	𐾣	𐼃	𐭢	𐭣	𐡂
daleth	𐾤	(𐼁)	𐭣	𐭤	𐡃
he	𐾥	𐼄, 𐼅	𐭥	𐭥	𐡄
waw	𐾦	𐼆	𐭦	𐭦	𐡅
zayin	𐾧	𐼇	𐭧	𐭧	𐡆
heth	𐾨	𐼈	𐭨	𐭨	𐡇
teth	—	—	𐭩	𐭩	𐡈
yodh	𐾩	𐼉	𐭪	𐭪	𐡉
kaph	𐾪	𐼊	𐭫	𐭫	𐡊
lamedh	𐾫	𐼋	𐭬	𐭬	𐡋
mem	𐾬	𐼌	𐭭	𐭭	𐡌
nun	𐾭	𐼍	𐭮	𐭮	𐡍
samekh	𐾮	𐼎	𐭯	𐭯	𐡎
ayin	𐾯	𐼏, (𐼁)	(𐭦)	𐭰	𐡏
pe	𐾰	𐼐	𐭱	𐭱	𐡐
sadhe	—	𐼑	𐭲	𐭲	𐡑
qoph	—	—	(𐭭)	𐭳	𐡒
resh	𐾱	𐼒	(𐭦)	𐭴	𐡓
shin	𐾲	𐼓	𐭶	𐭵	𐡔
taw	𐾳	𐼔	𐭷	𐭶	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.

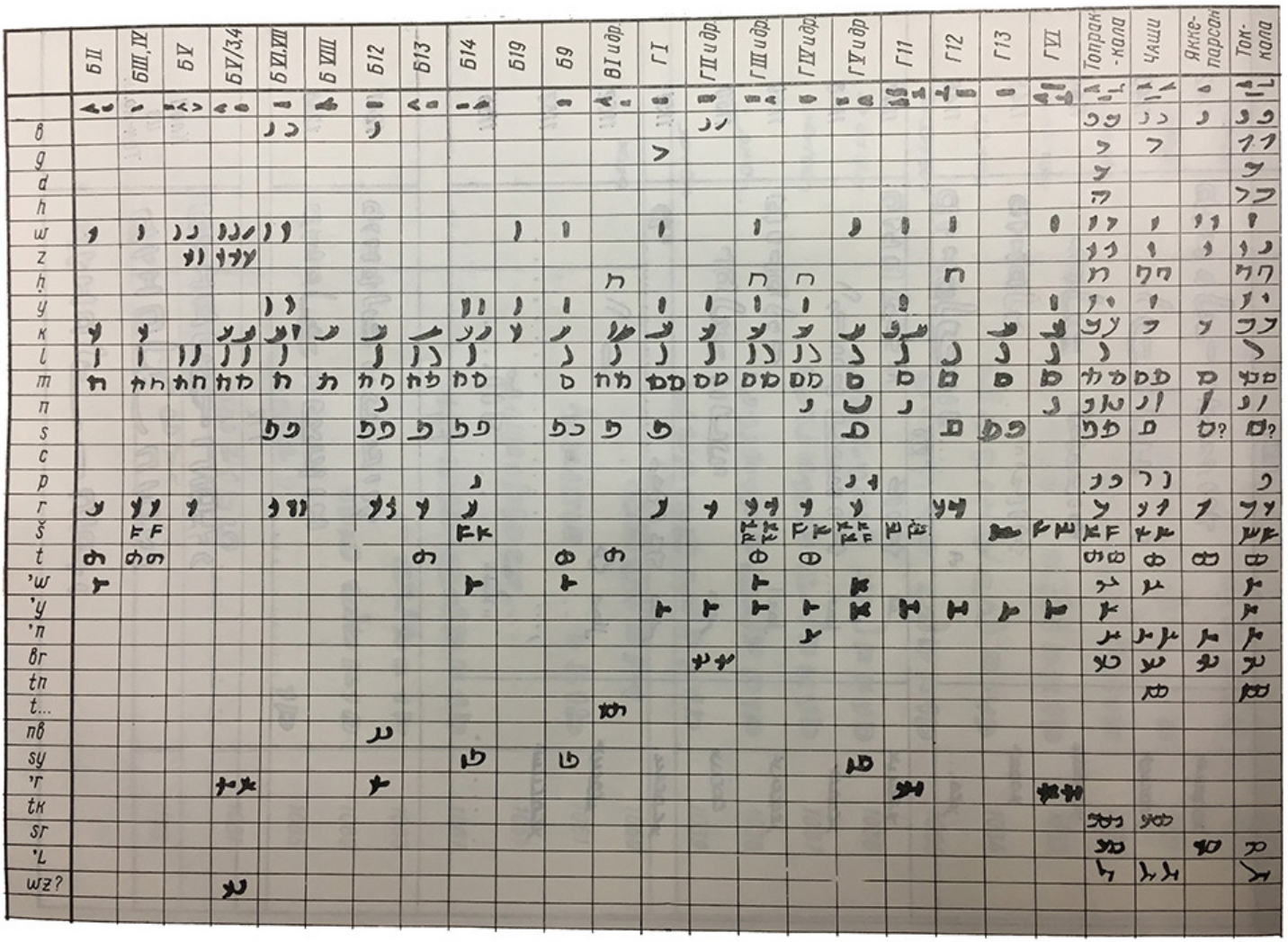


Figure 1: Inventory of characters on Chorasmian coins (БII–ГVI), Toprak Kala (Топрак-кала), Yakke Parsan (Якке парсан), Tok Kala (Ток-кала) (from Vainberg 1977: Table 8).

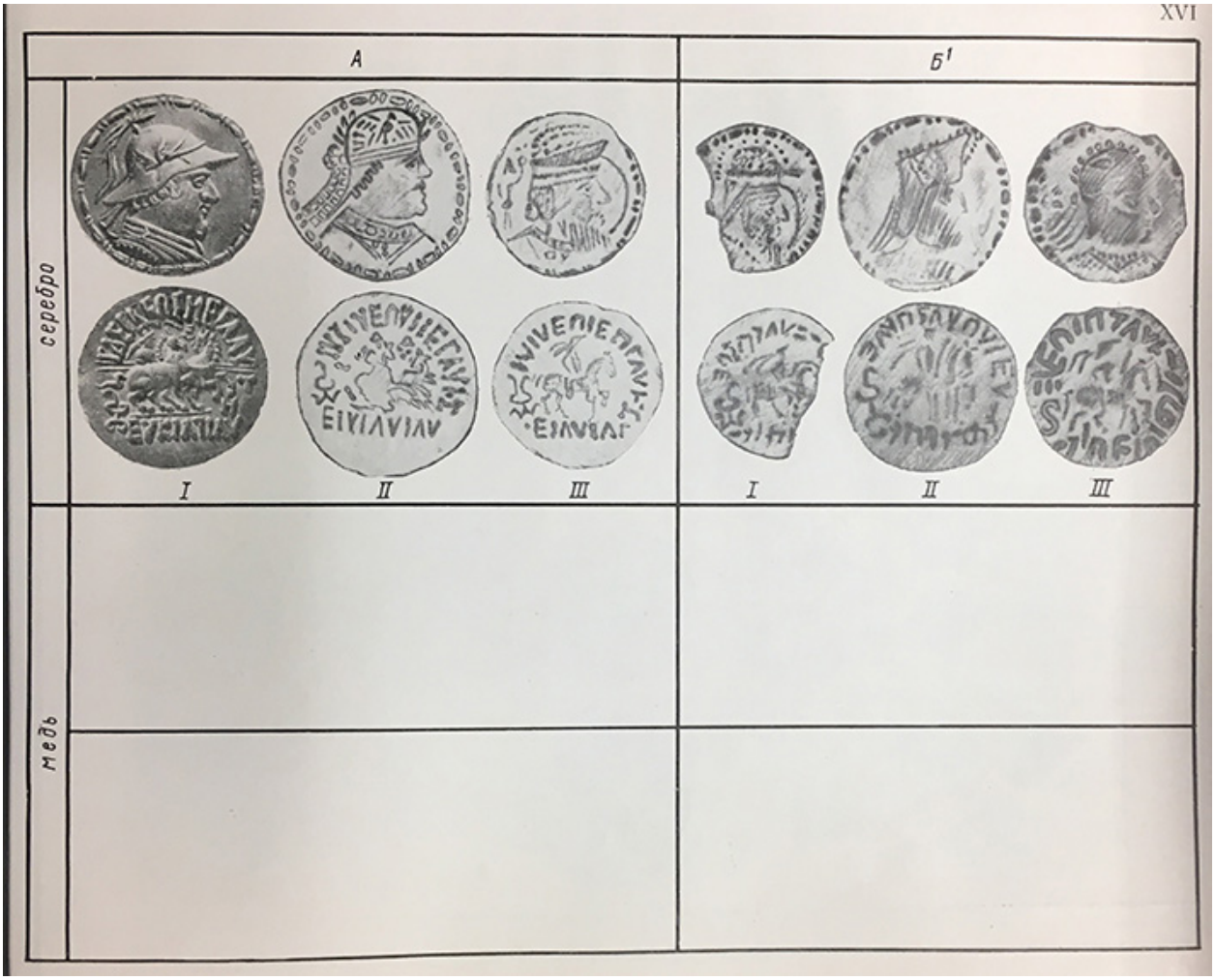


Figure 2: Chorasmian coins (from Vainberg 1977: Table 16).

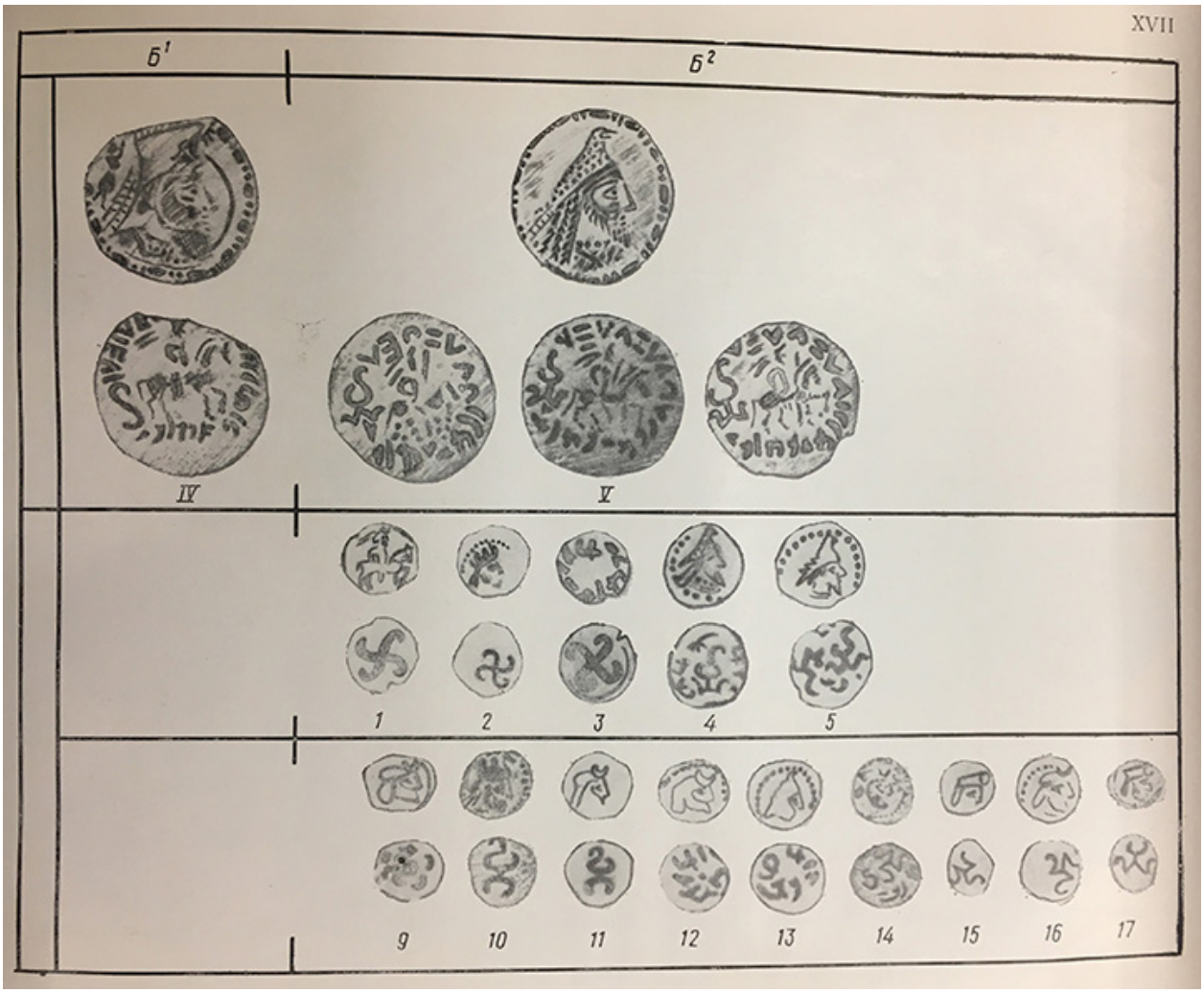


Figure 3: Chorasmian coins (from Vainberg 1977: Table 17).

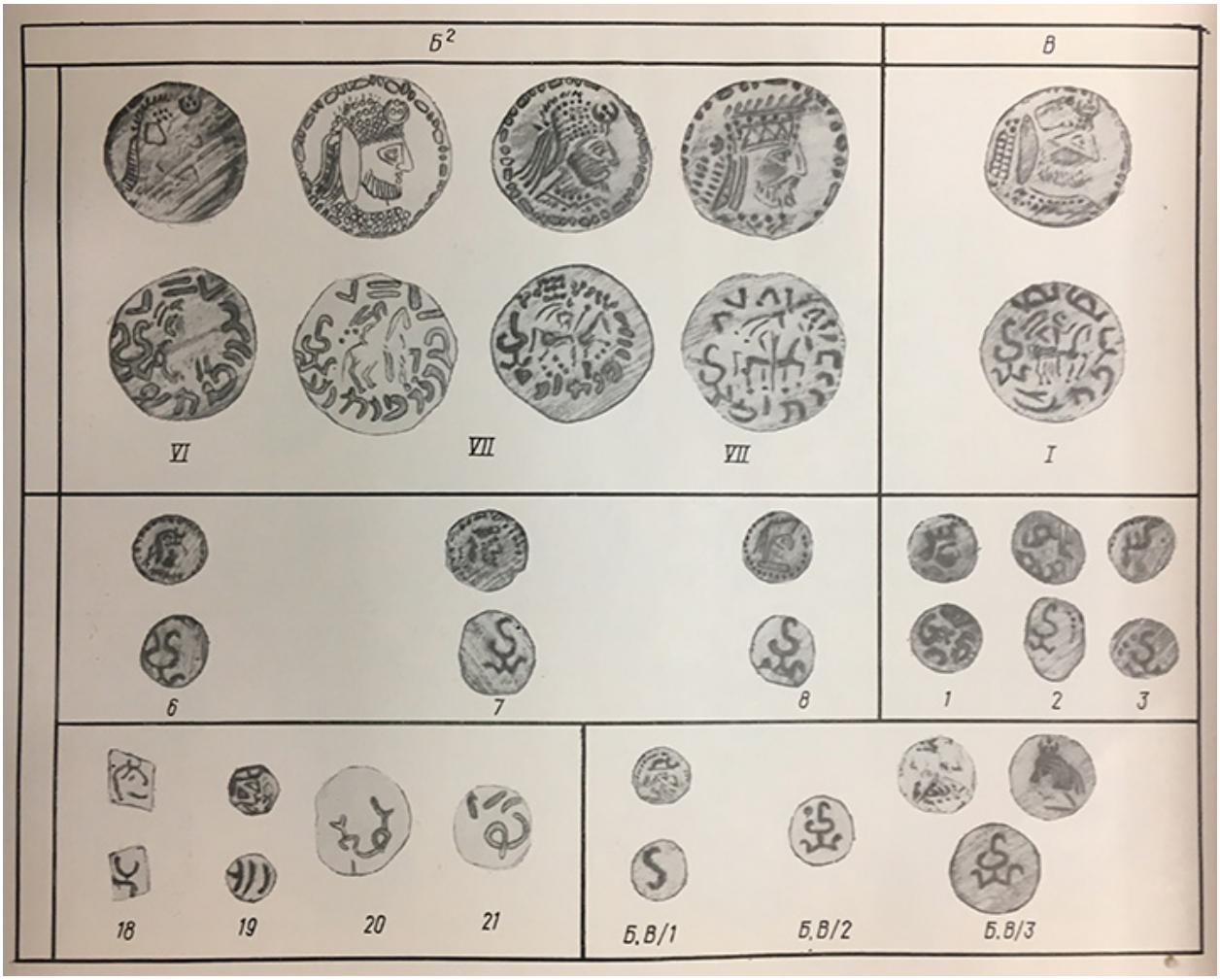


Figure 4: Chorasmian coins (from Vainberg 1977: Table 18).

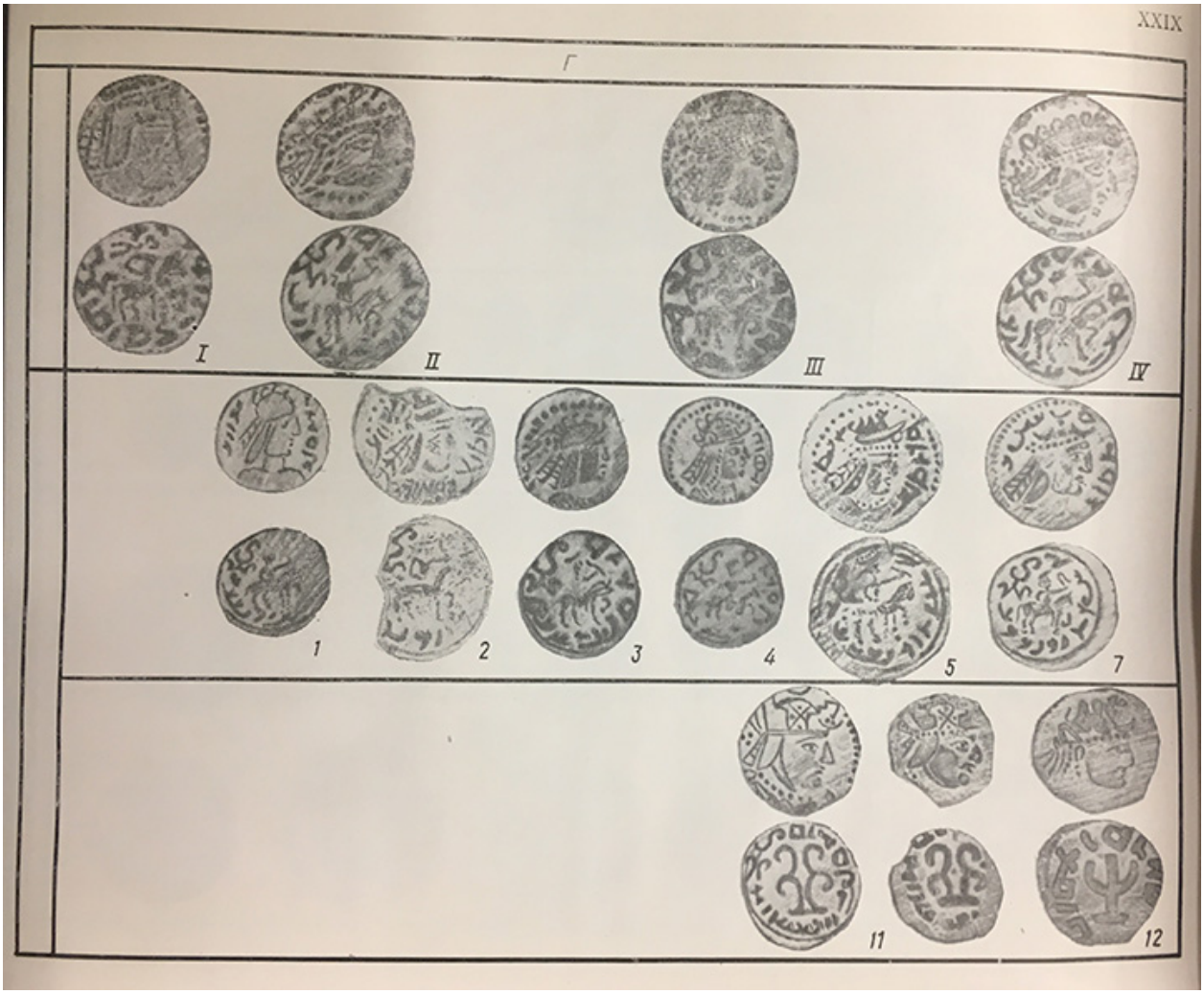


Figure 5: Chorasmian coins (from Vainberg 1977: Table 19).

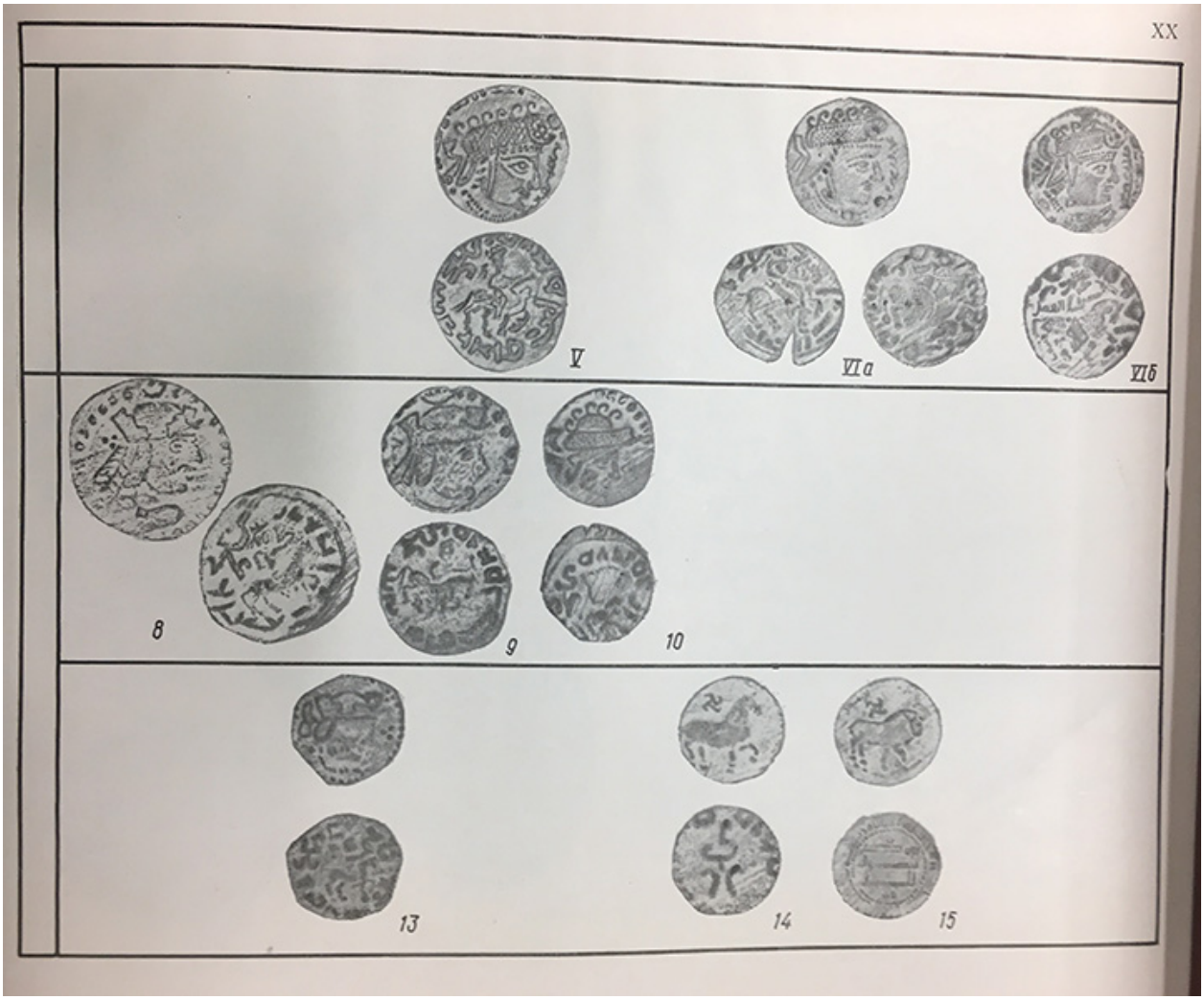


Figure 6: Chorasmian coins (from Vainberg 1977: Table 20).

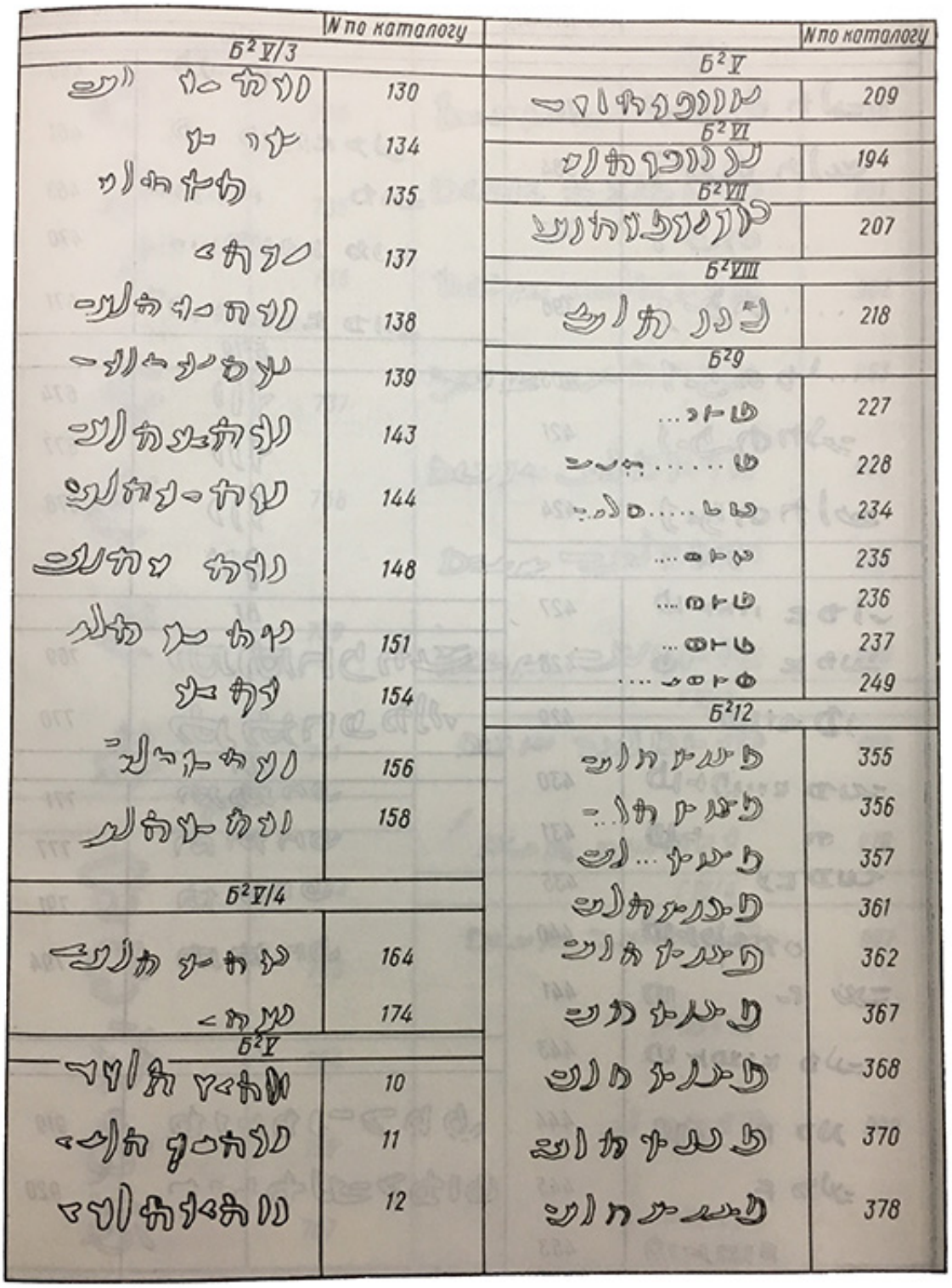


Figure 7: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 1).

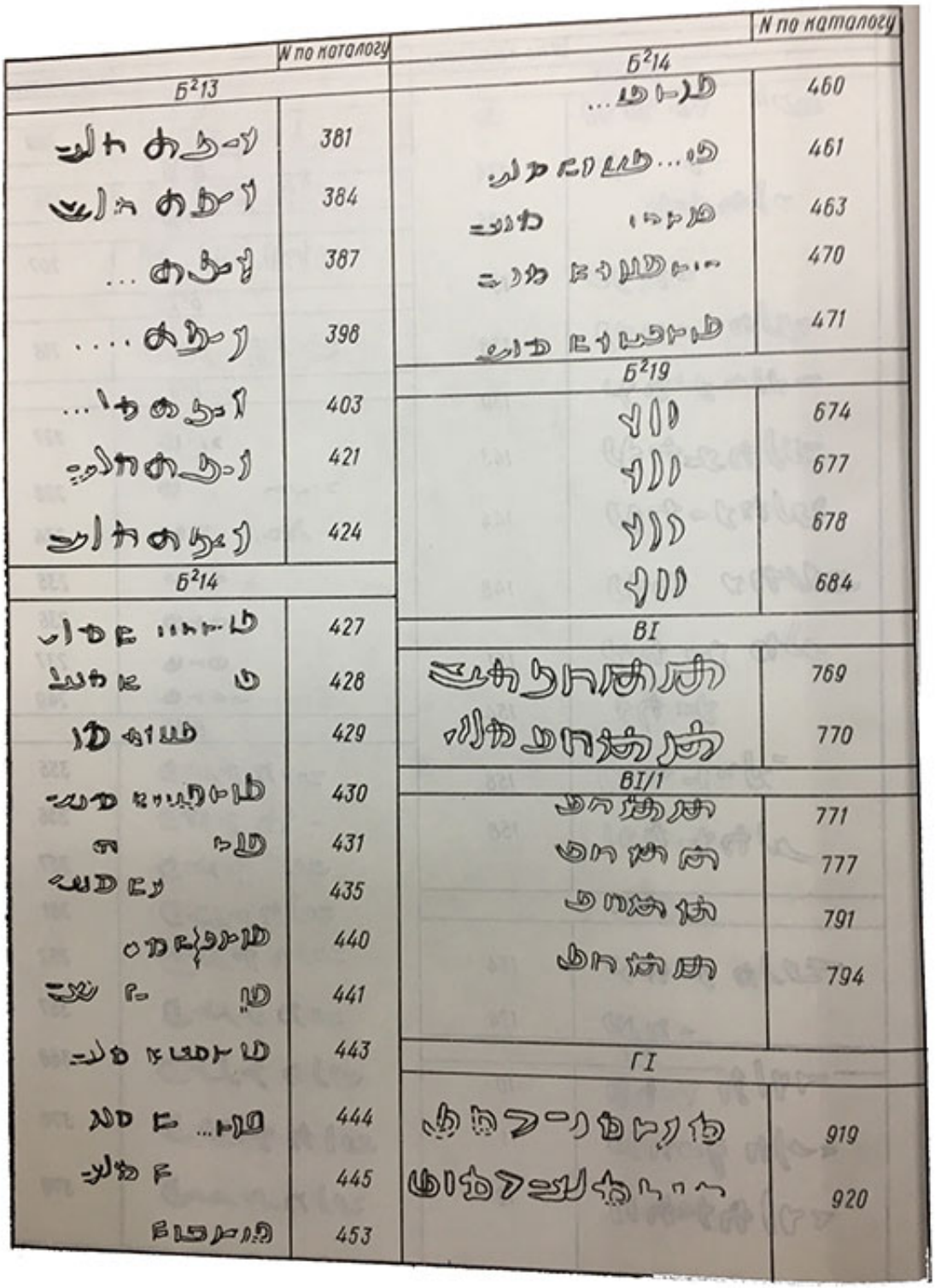


Figure 8: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 2).

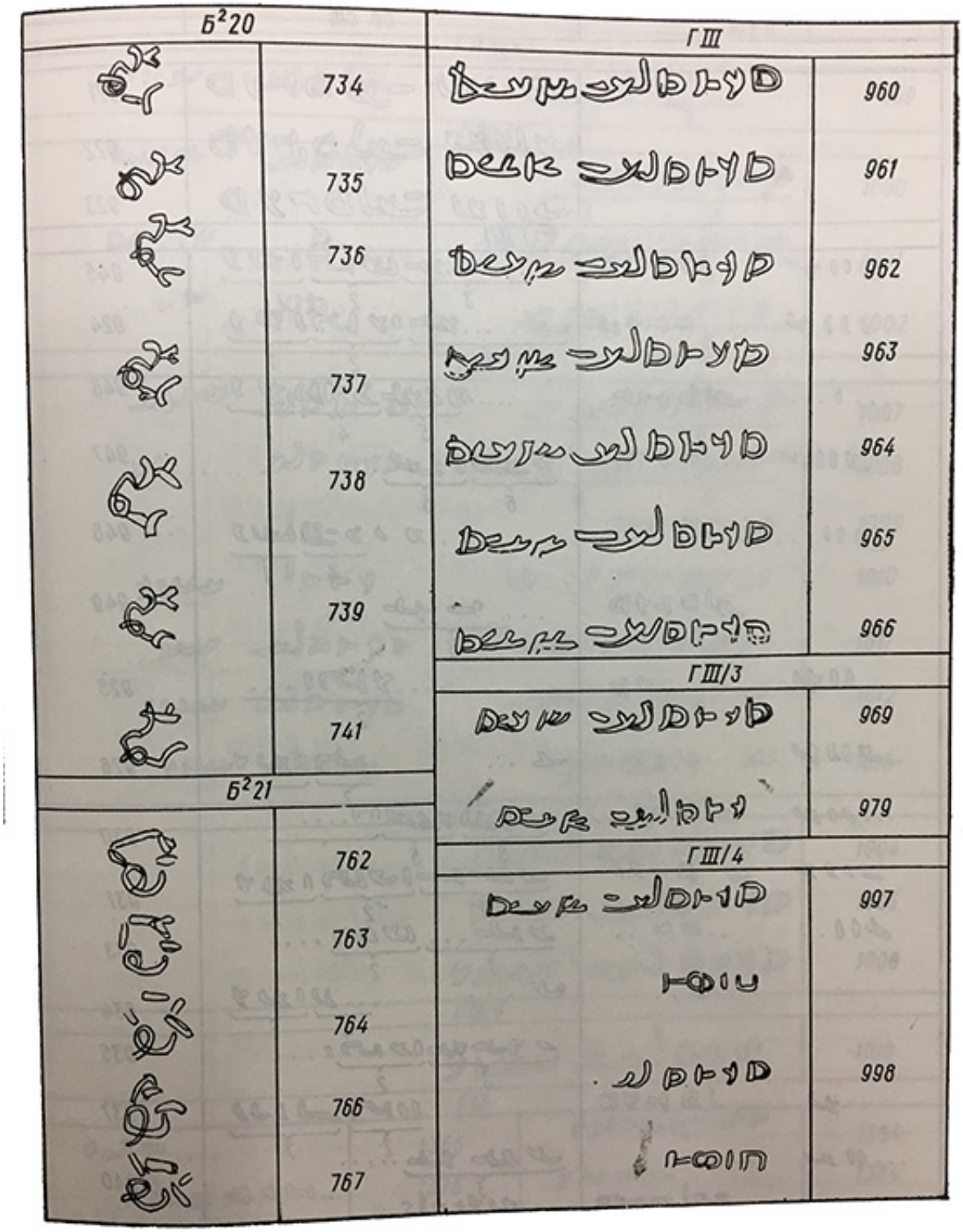


Figure 9: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 3).

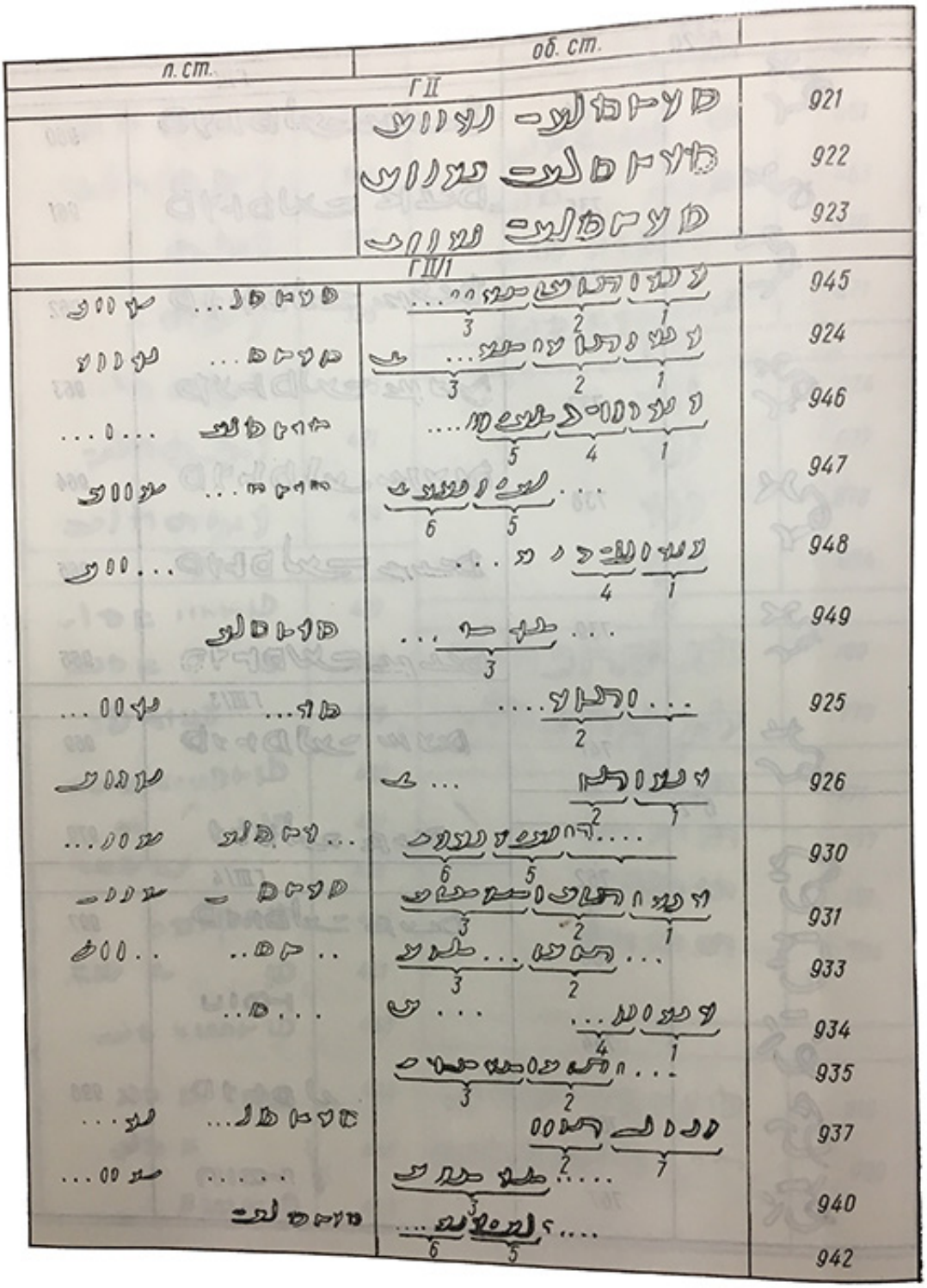


Figure 10: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 4).

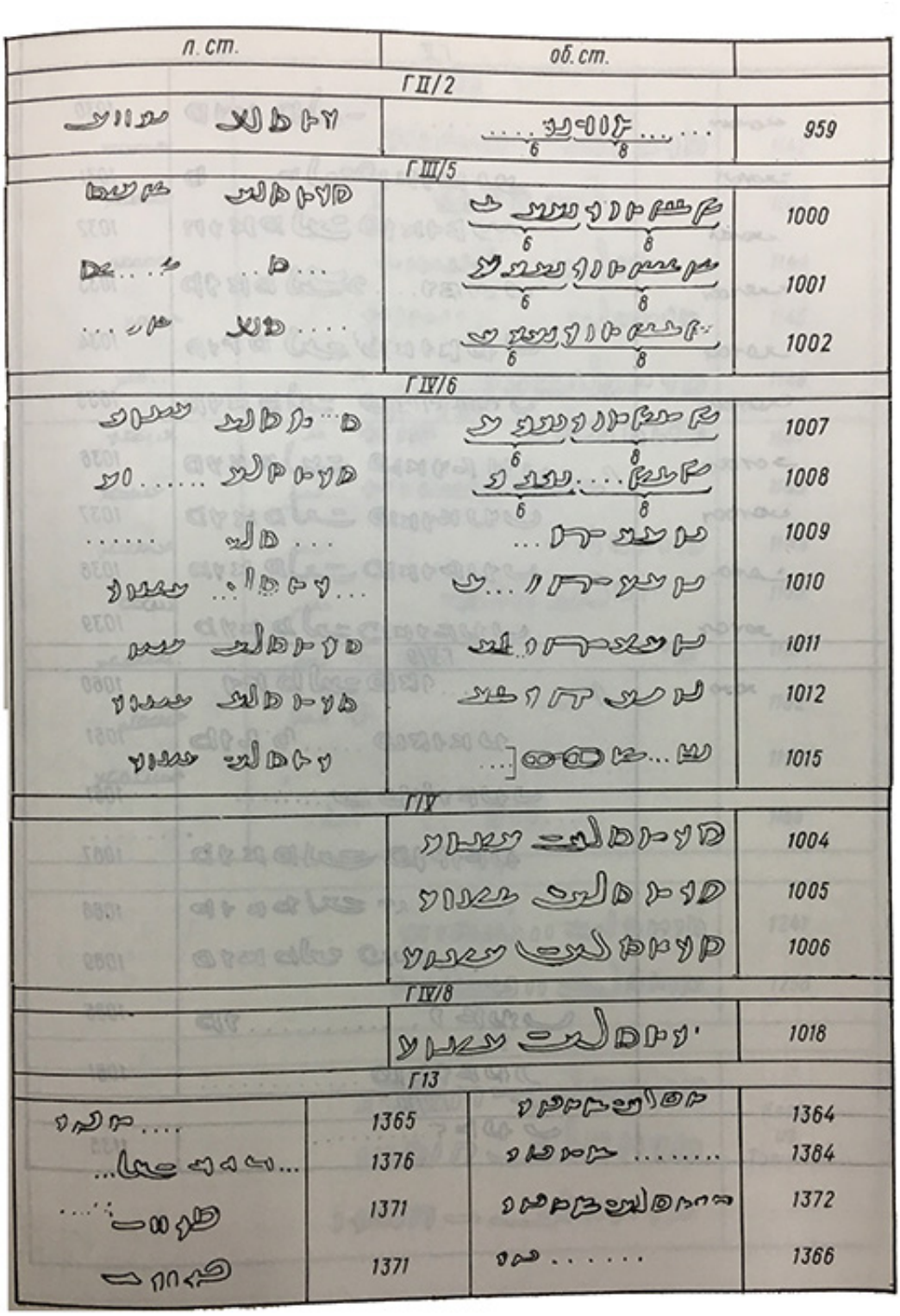


Figure 11: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 5).

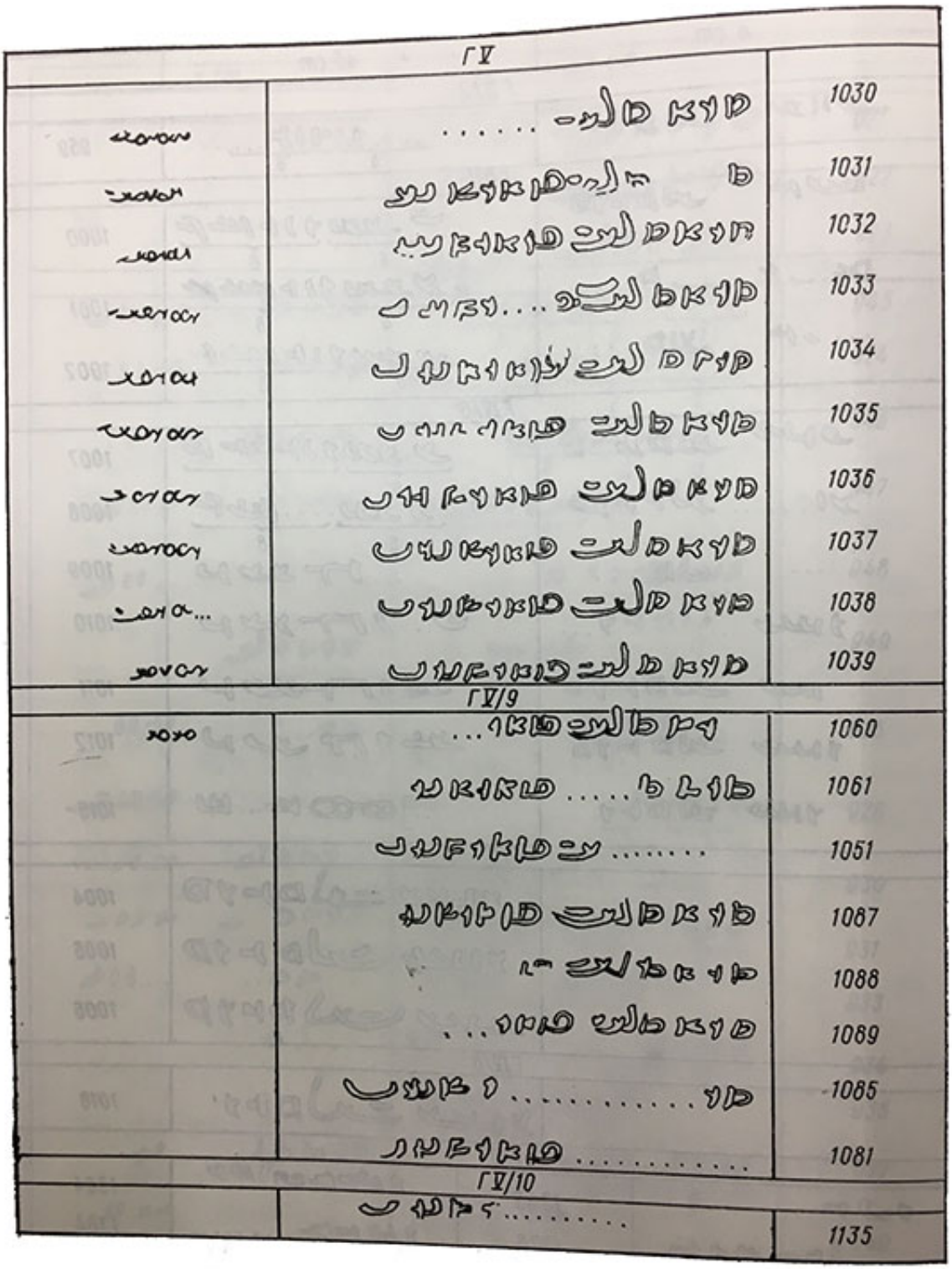


Figure 12: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 6).

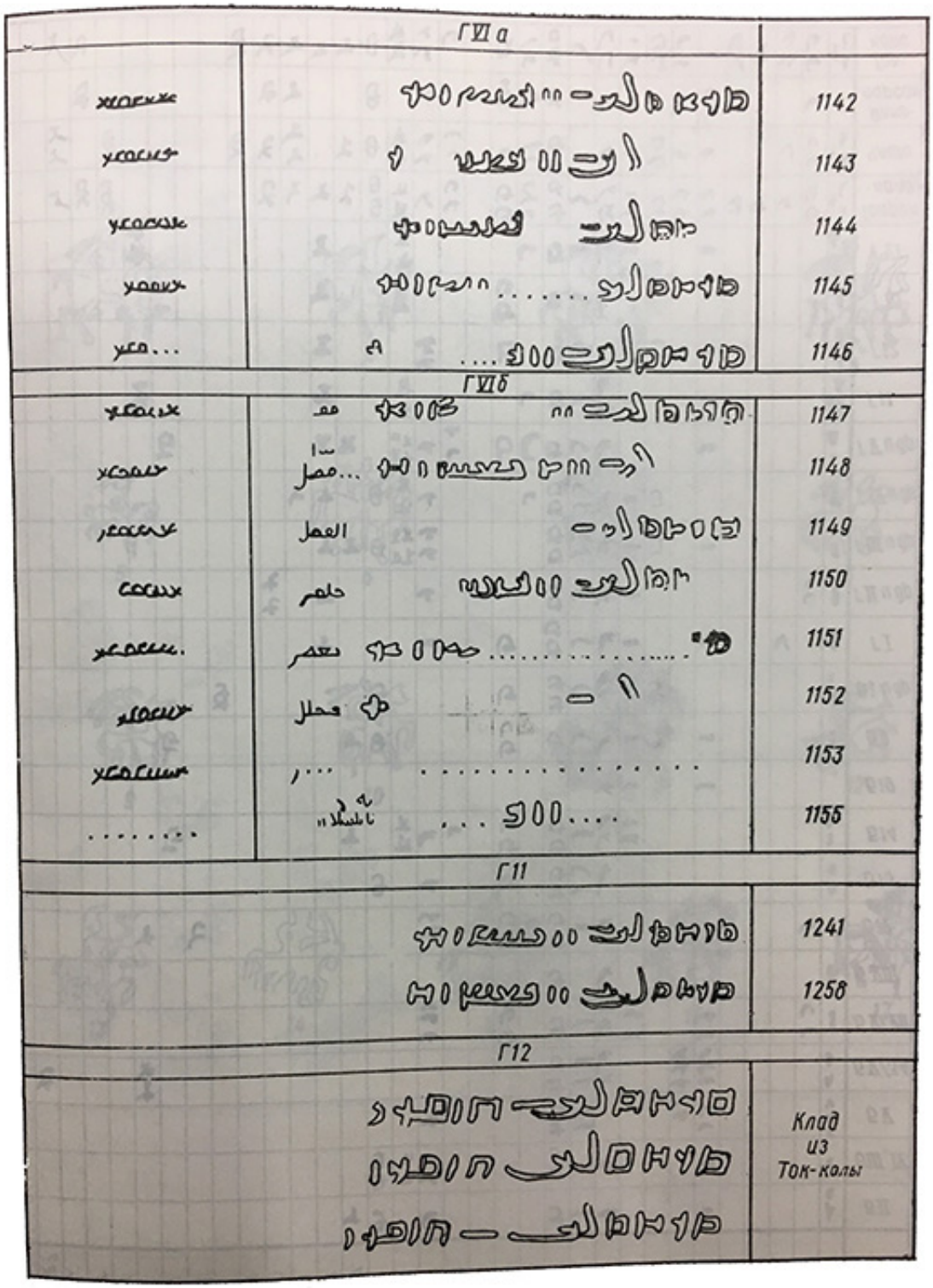


Figure 13: Inscriptions on Chorasmian coins (from Vainberg 1977: Table 7).

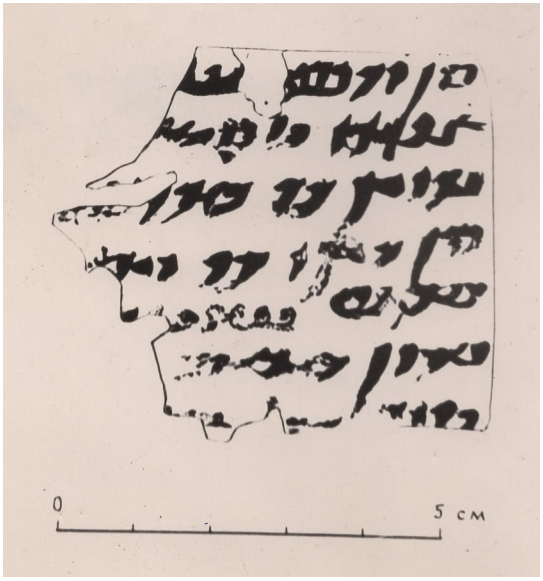


Figure 14: Fragment of a leather document with Chorasmian inscription from Yakke Parsan. Image courtesy of Lurje.

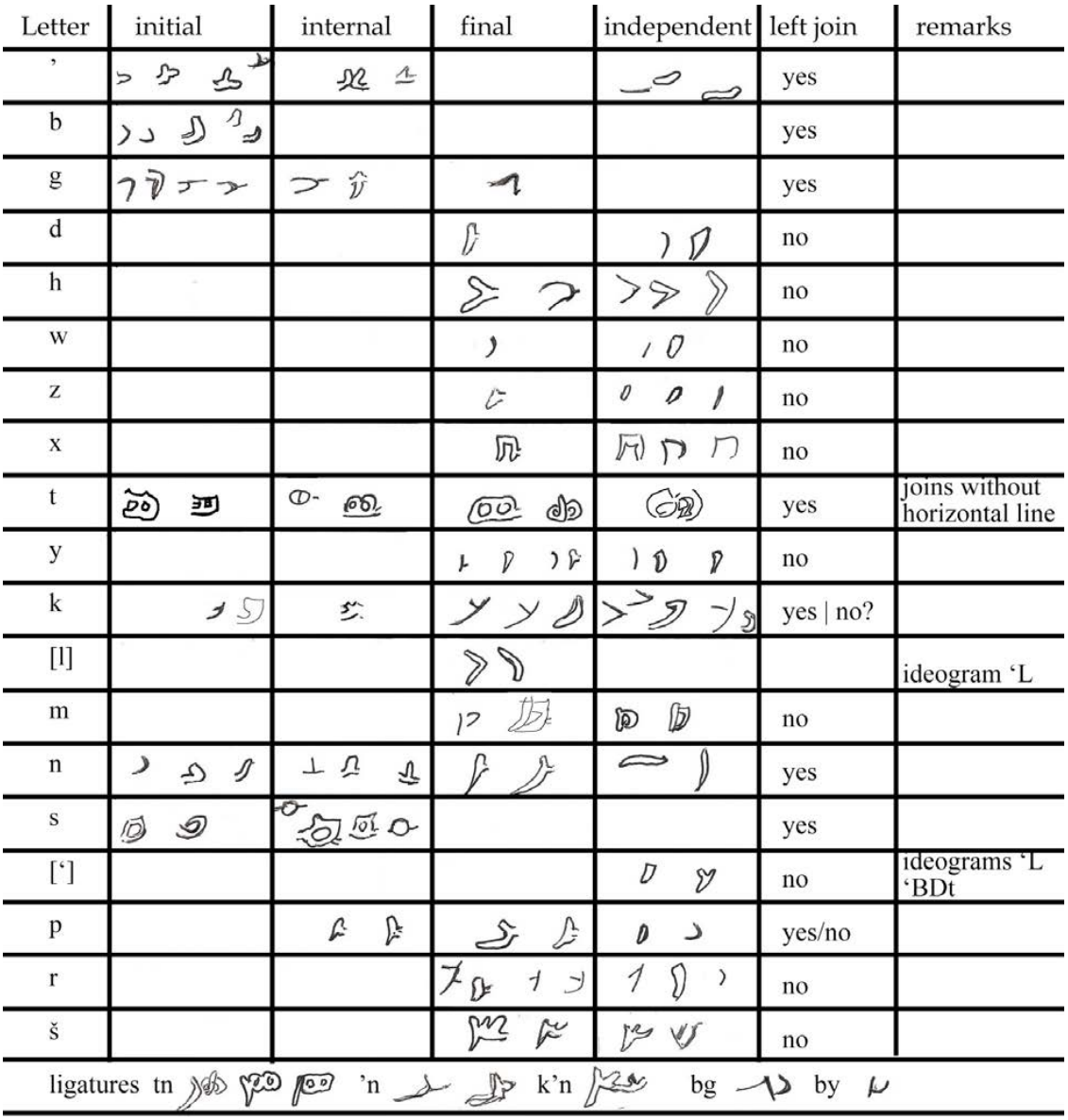
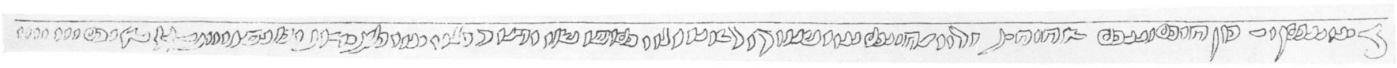


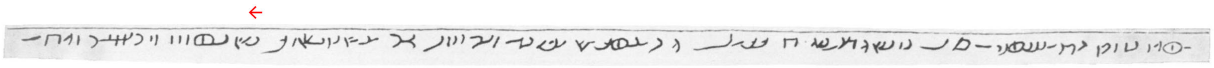
Figure 15: Nominal and positional forms of letters in silver vessel inscriptions (from Lurje 2017).



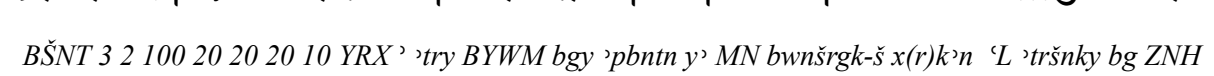
Pl. 1:a *Choresmian No. 1:* A.D. 658. Inscription from silver phiale in the British Museum, Smirnov, *VS*, pl. XIX:43.

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.

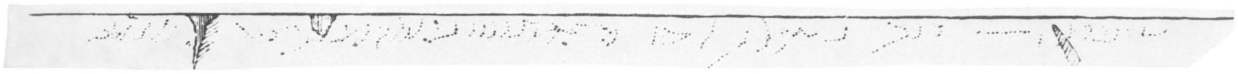


BŠNT 3 2 100 20 20 20 10 YRXʾ ʾtry BYWM bgy ʾpbntn yʾ MN bwnšrgk-š x(r)kʾn ʿL ʾtršnky bg ZNH ZWZN 20 10 ʾšpynšwk

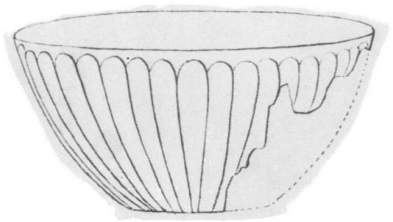
Pl. 3:b *Choresmian No. 2:* A.D. 538 (probably 638). Silver phiale in the Hermitage Museum, Leningrad, see pl. 3:a. Diam. 10.6 cm.

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.

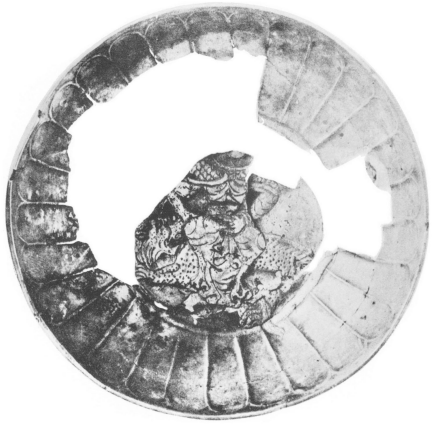
Figure 17: Silver bowl #2: either 538 or 638 CE (from Azarpay 1969: Plate 3:a, b, c). Silver philae in the Hermitage Museum (St. Petersburg). Original from Smirnov 1909, plate XIX:42 and XVIII: 42. Transliteration from Lurje (forthcoming).



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



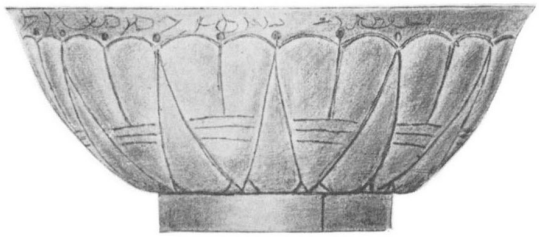
Pl. 5:b *Choresmian No. 3:* silver phiale in the Hermitage Museum, Leningrad, see pl. 5:a, diam. 12.5 cm. Smirnov, *VS*, pl. XVIII: 44.



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.

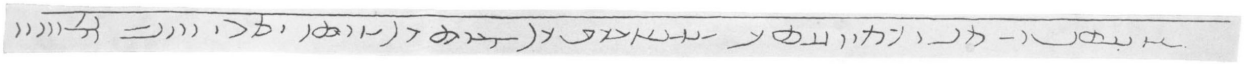
Pl. 8:a
Choresmian No. 4: inscription from silver phiale in the Hermitage Museum, Leningrad, Smirnov, *VS*, pl. XIX:45.

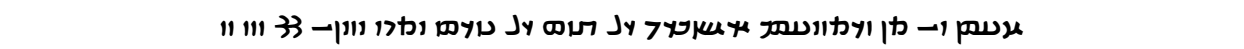


Pl. 8:b *Choresmian No. 4:* silver phiale in the Hermitage Museum, Leningrad, see pl. 8:a. Diam. 12.7 cm.

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

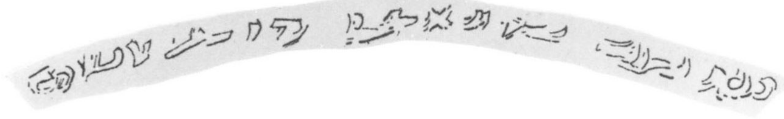
Pl. 11:a, b, c *Choresmian No. 5:* silver phiale in the Hermitage Museum, Leningrad, Smirnov, *VS*, pl. XIX:47. Diam. 13 cm.

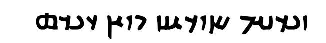
Figure 20: Silver vessel #5 (from Azarpay 1969: Plate 11:a, b, c). Silver philae in the Hermitage Museum (St. Petersburg). Original from Smirnov 1909, plate XIX: 47. Transliteration from Lurje (forthcoming).

𐾳𐿄𐾺 𐾹𐾶𐿁𐾿𐾻 𐾸𐾶𐾸𐾾𐾰 𐿈 𐿉 𐿊𐿊𐿊 𐿈𐿇

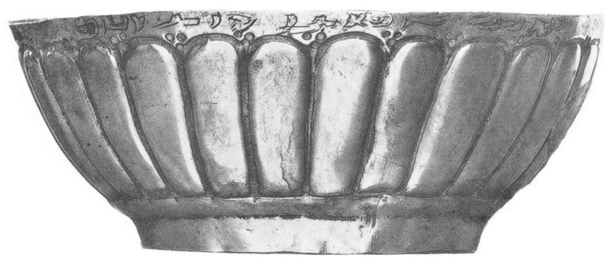
gty (xwpsk | xw ksp) ZWZN-ʾ 4 10 20 20 20 7

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.

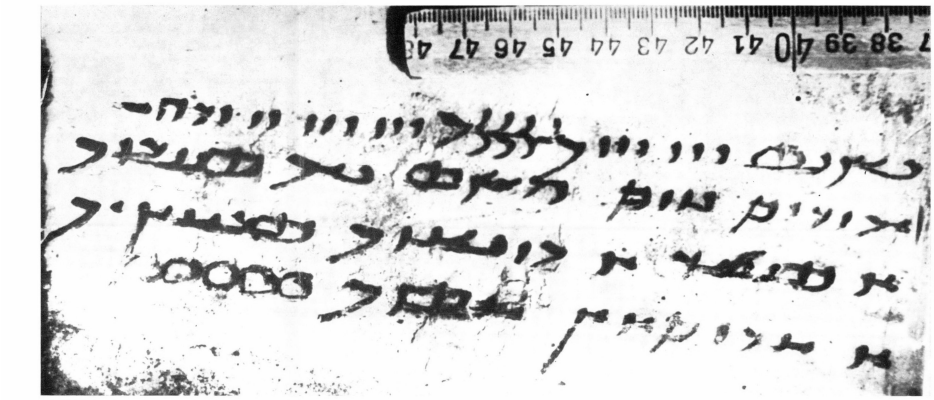


Fig. 1. Inscription No. 52.

BŠNT III III C XX XX XX X III III II YRḤʾ
ʾhwrym BYWM gwšt ZNH tnbryk
ʾy tnbʾr ʾy hwnšk (?) tʾbʾnʾnʾk (?)
ʾy ʾyrw|zmʾwʾn ʾztyk ○○○○

Year 678. Month Ahurem, day Gost. This
ossuary contains the body of hwnsk
tʾbʾnʾnʾk, son of ʾyrw|zmʾwʾn

Figure 24: Tok Kala no. 52, ossuary inscription (from Tolstov and Livshitz 1964: Figure 1). Transliterations from same; but may be erroneous or outdated.

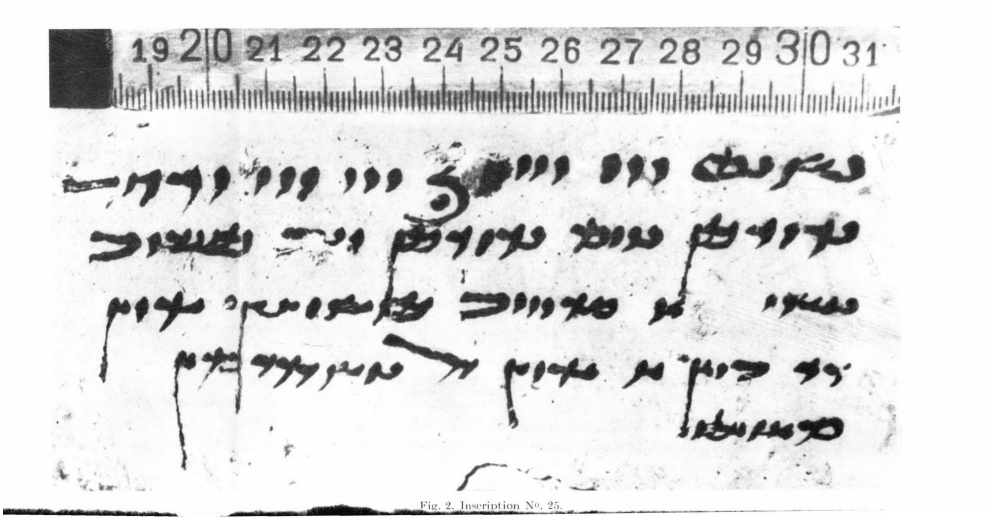


Fig. 2. Inscription No. 25.

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 ʾy srwywk tyšyʾnʾn-w ʾrwʾn
ʿD hwʾn-ʾ ʾy ʾrwʾn ʿL nwš γrδmʾn
mʾnyʾ(ʾ)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). Representation in Chorasmian script based upon the reanalysis by Henning.

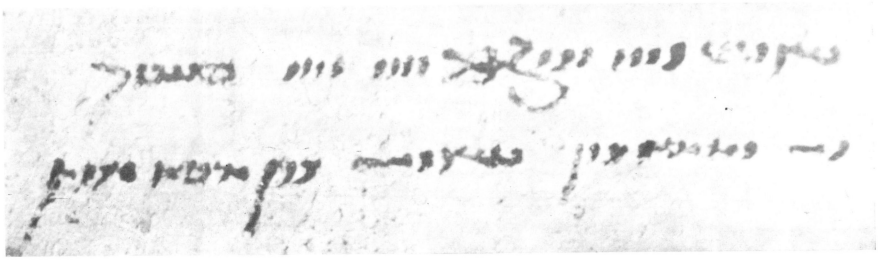


Fig. 3. Inscription N°. 19.

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.

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

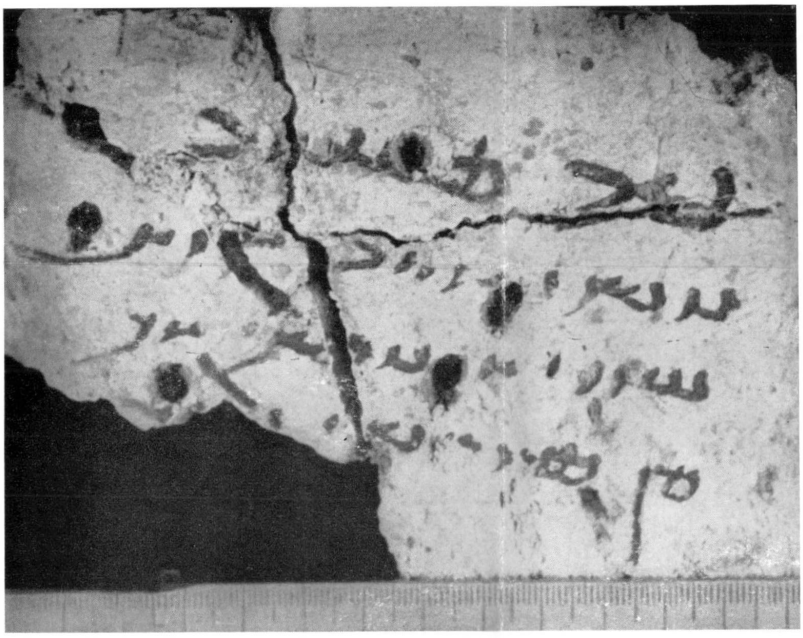


Fig. 4. Inscription No. 8.

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.

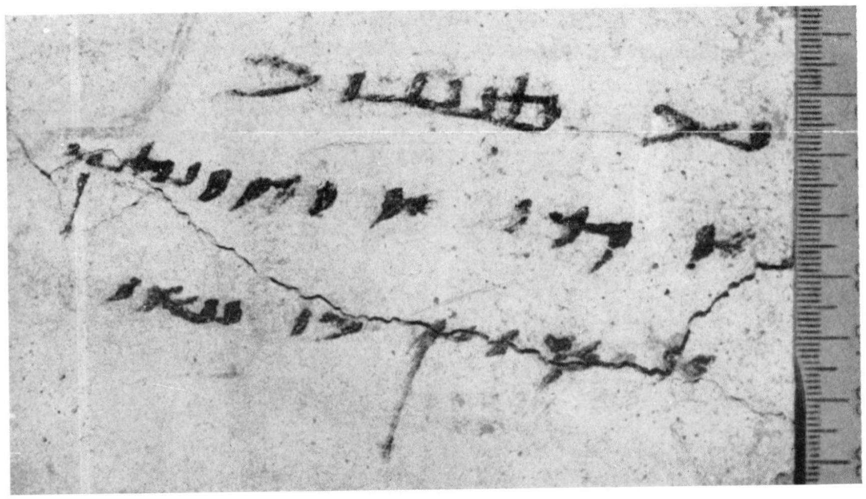
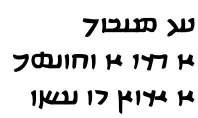


Fig. 5. Inscription N°. 69.



Tolstov and Livshitz 1964	Reanalysis by proposal author
ZNH tnbryk	*ZNH tpnkwk*
’y gry ’y wḥwntk	*ʾy gry ʾy wḥwnt’n-w*
’y ’rw’n kw nwšy	*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 by the proposal author.

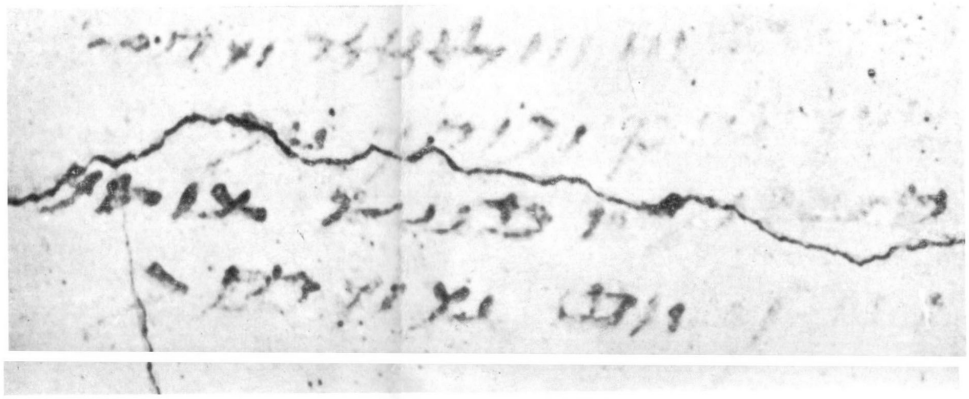


Fig. 6. Inscription N°. 39.

ˠBŠNTˡ III III C XX XX XX X YRḤ’
mˠtry?ˡ ˠBYWˡM whwmn ZNH
tnˠbrˡyk ’y tnb’r ’rw’zd
w … nˠyˡ 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.

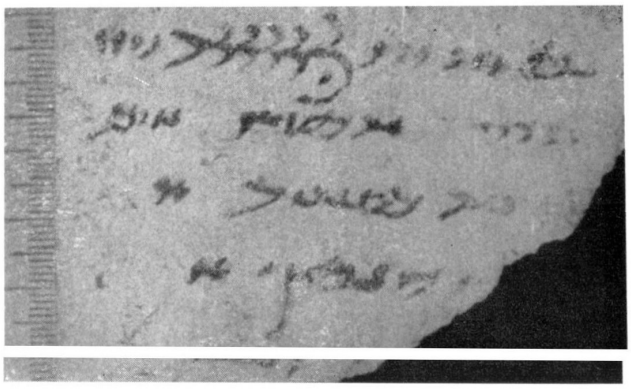
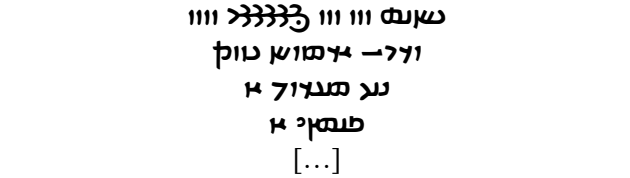


Fig. 7. Inscription No. 12



[…]

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

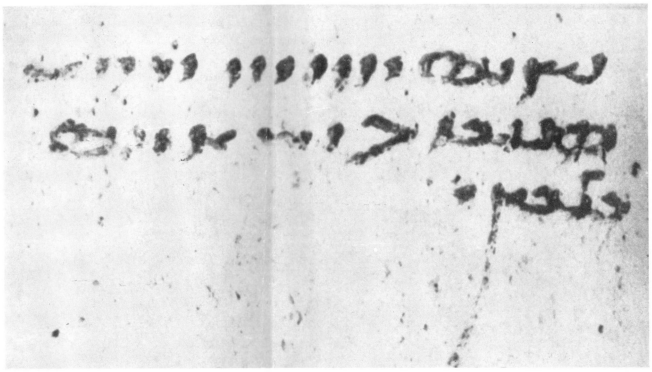
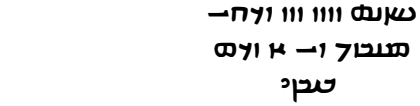


Fig. 8. Inscription No. 26.



Tolstov and Livshitz 1964	Reanalysis based upon Henning 1965
BŠNT III IIII YRḤ' *tnbryk y' 'y wrt* *k'k'ny*	*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.

BŠNᶠTᵎ [
YRᶠH’ᵎ [
whwmn’ [ZHN tnbryk ’y ḫw’r’n [
wḫnwy(?) [

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

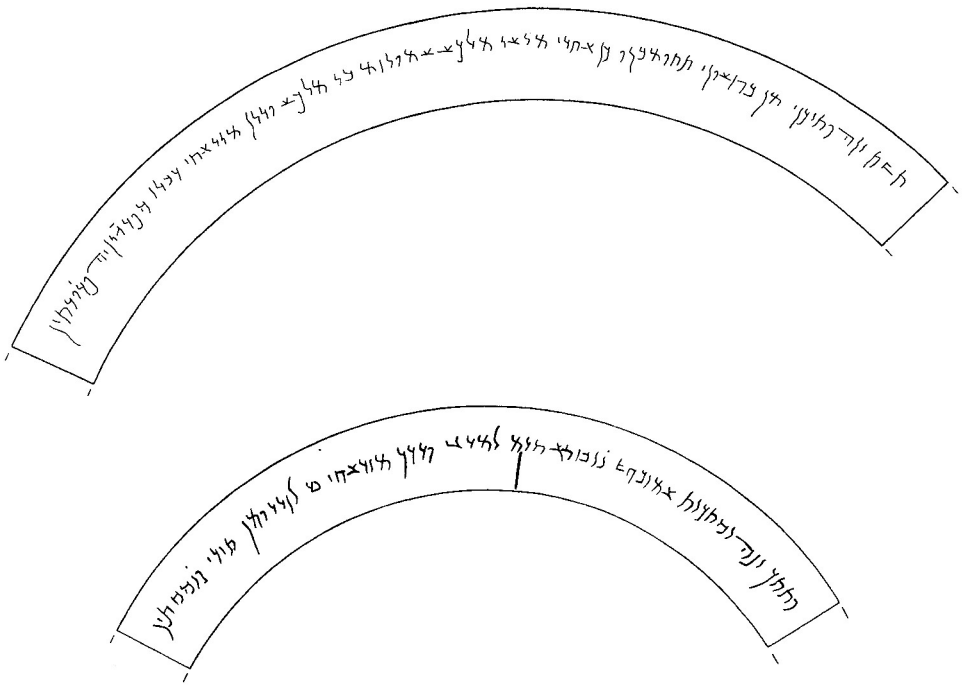
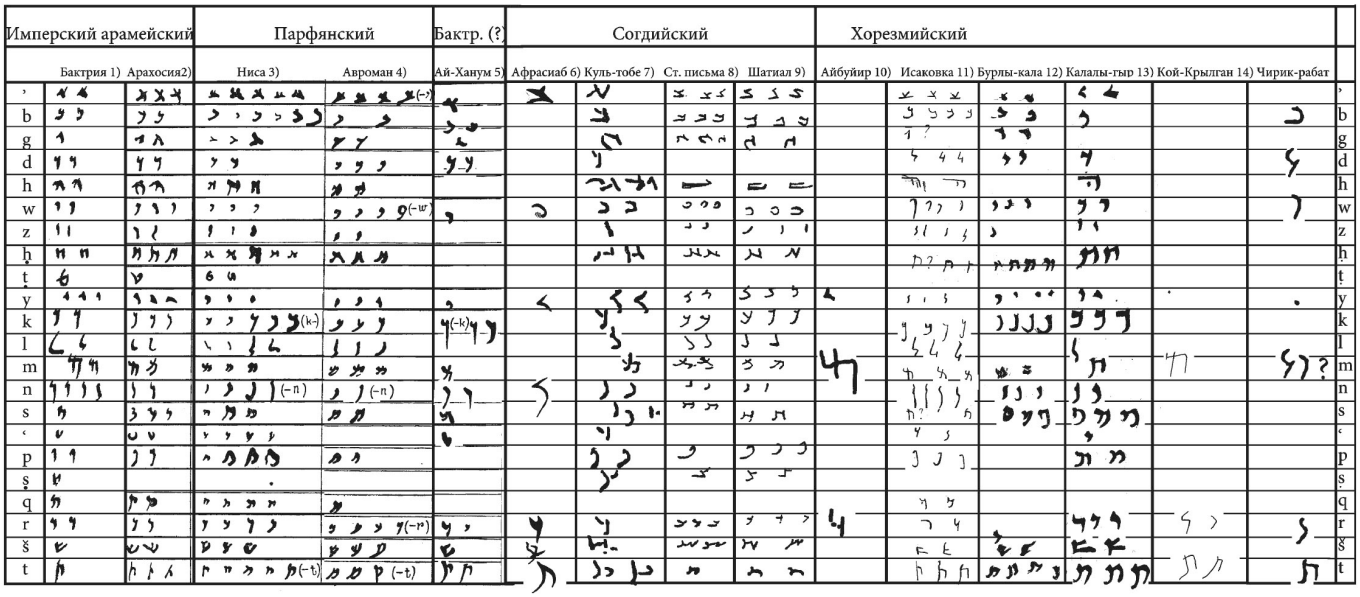


Figure 33: Archaic Chorasmian inscriptions on silver bowls no. 1 (top) and no. 2 (bottom) from Isakovka from the Achaemenid period (from Livshits 2003: 152, 163). This type of script is out of scope for the present encoding, and may be unified with Imperial Aramaic.

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).

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: *2018-04-29*
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: *29*
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: