ISO/IEC JTC1/SC2/WG2 N1947
1999-01-29

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: **Working Group Document**
Title: **Revised proposal for encoding the Old Permic script in the UCS**
Source: **Michael Everson, EGT (IE)**
Status: **Expert Contribution**
Action: **For consideration by JTC1/SC2/WG2 and UTC**
Date: **1999-01-29**
Replaces: **N1687**

This document is a revision of N1687 based on additional research, and contains the proposal summary.

A. Administrative

1. Title
Revised proposal for encoding the Old Permic script in the UCS.
2. Requester's name
Michael Everson, EGT (WG2 member for Ireland).
3. Requester type
Expert contribution.
4. Submission date
1999-01-29.
5. Requester's reference
6a. Completion
This is a complete proposal.
6b. More information to be provided?
No.

B. Technical -- General

1a. New script? Name?
Yes. Old Permic.
1b. Addition of characters to existing block? Name?
No.
2. Number of characters
38
3. Proposed category
Category D.
4. Proposed level of implementation and rationale
Level 1 noncombining.
5a. Character names included in proposal?
Yes.
5b. Character names in accordance with guidelines?
Yes.
5c. Character shapes reviewable?
Yes (see below).

6a. Who will provide computerized font?
Michael Everson.
6b. Font currently available?
Yes.
6c. Font format?
TrueType.
7a. Are references (to other character sets, dictionaries, descriptive texts, etc.) provided?
Yes, see bibliography below.
7b. Are published examples (such as samples from newspapers, magazines, or other sources) of use of proposed characters attached?
Yes, see below.
8. Does the proposal address other aspects of character data processing?
Yes; see Unicode properties below.

C. Technical -- Justification

1. Contact with the user community?
Yes: There is contact with a major centre for Uralic studies, Kotimaisten kielten tutkimuskeskus, Helsinki. I have no contact in the Komi Republic. The proposal is being circulated on uralist@helsinki.fi. The principle authors of some of the materials used to prepare the proposal (chiefly Лыткин 1952 and Stipa 1961), are deceased.

2. Information on the user community?
Scholars and Komi language enthusiasts.
3a. The context of use for the proposed characters?
Old Permic was used chiefly to write Komi texts, but a number of Russian texts are also found in Old Permic script.
3b. Reference
See bibliography.
4a. Proposed characters in current use?
The script is extinct. Old Permic does not appear ever to have been set in type before the present time. I do not know of any computer implementations. However, it can be said to have some use nevertheless.
4b. Where?
Scholars and Komi language enthusiasts (one presumes).
5a. Characters should be encoded entirely in BMP?
No.
5b. Rationale
Rarity and accordance with the Roadmap.
6. Should characters be kept in a continuous range?
Yes.
7a. Can the characters be considered a presentation form of an existing character or character sequence?
No; though this is a Greek-derived script like Coptic, Cyrillic, Armenian, Georgian, and Gothic.
7b. Where?
7c. Reference
8a. Can any of the characters be considered to be similar (in appearance or function) to an existing character?
No.
8b. Where?
8c. Reference

9a. Combining characters or use of composite sequences included?
No; but see issues listed below for possible null use of COMBINING DIAERESIS.
9b. List of composite sequences and their corresponding glyph images provided?
No.
10. Characters with any special properties such as control function, etc. included?
No.

E. Proposal

The Old Permic script was devised in the 14th century by Russian missionary Stefan of Perm (*fl.* 1373-1395), and was used to write Komi and Komi-Permyak languages. Its model was based on the Greek and Cyrillic alphabets (as can be seen from the order of the letters) but many of the glyphs were taken from the "Tamga signs", used in indigenous Komi religious practices, by Stefan as a way of linking traditional signs, which had a great deal to do with the self-identity of the Komi, with Christian traditions.

The term "Old Permic" is suggested as the name for this script. Stipa calls it "Permisch"; Pillu uses the terms "vanapermi", "древнепермский", and "Old-Permian" in Estonian, Russian, and English respectively. Haarmann calls it "Abur" and "Altsyrjänisch" (= 'Old Komi'); "Abur" is a name derived from the names of the first two characters AN and BUR (so ⸗∨-ϕ∩▽ is equivalent to αλφα–βητα and аз-буки).

Structure

Old Permic is a simple alphabetic script, read from left to right in horizontal lines running from top to bottom. In the manuscripts, Old Permic glyphs are often shown with a diaeresis, but this doesn't seem to be significant in the regular way it is for, for example, the German letters *u* and *ü*. Old Permic is caseless.

Ordering

Alphabetical ordering is a little bit erratic in the sources, especially toward the end of the alphabet. This proposal follows the order 1–34 given in Лыткин 1952, where 1–24 are primary letters and 25–34 are supplementary (see the reproduction of Лыткин's tables below). Four characters are considered to be subordinate in the order to another character, as seen in the pairs 0365/0366 (SHOOI ⸗ and SHCHOOI ш), 0367/0368 (YRY ъ and YERU ы), 0369/036A (O ∩ and OO ω), and 0374/0375(YA ъ and IA ⸗). Of these only SHCHOOI hasn't (yet) been identified as corresponding to its Cyrillic equivalent 0449, though it appears in a number of sources. The name SHCHOOI is not attested. The others correspond to 044B, 0460, and 0467.

Use of combining characters

In the table, ϕ 036B (EF) and ⸗ 036C (HA) are given with glyphs which are not the most commonly found ones. More common representation of the corresponding phonetic entities makes use of combining characters: ϕ̈ is represented by 035F PEEI ϕ + ¨ and ⸗̈ is represented by 03FA KOKE ⸗ + ¨. Note the similarity of ϕ and ⸗ to Greek ϖ and χ. Five combining characters may be found with Old Permic characters: COMBINING DIAERESIS, COMBINING DOT ABOVE, COMBINING BREVE, COMBINING COMMA ABOVE, and COMBINING THREE DOTS ABOVE (the last makes us ask whether 20DB can be used with letters in Unicode). The use of diacritical marks is peculiar in Old Permic. Sometimes it appears that the diacritic is semantically zero, analogous to the use of a breve in German handwriting (especially the angular handwritten script Fraktur) over lower-case *u* to differentiate it from *n* (that is, where *ŭ* is not really supposed to be *u* + *breve* and would not be so encoded in text). Sometimes, particularly in word-initial position, vowel quality or consonant palatalization may be marked, apparently indifferently, by these diacritics. There may be more information on this in Лыткин 1952 but the Russian is slow going for me. Characters like COMBINING CYRILLIC TITLO may be used, e.g. to indicate abbreviation or numeric use of a character (see sample texts below).

Code positions proposed in the table are based on my Roadmap to the BMP, version 3.0, 1999-01-26.

Unicode Character Properties

Spacing letters, category “Lo”, bidi category “L” (strong left to right)

0001 0350 - 0001 0375

Bibliography

Haarmann, Harald. 1990. *Universalgeschichte der Schrift*. Frankfurt/Main; New York: Campus. ISBN 3-593-34346-0

Hämäläinen, A. 1909. “Èpifanij Viisan tiedot Pyhästä Tapanista ja syrjääneistä”, in *Suomi* 4,6,2: 1-27. (Tr. of Эпифаний Премудры. 1862.)

Lepechin, I. I. 1790. “Histoire de l’Académie Imperiale des Sciences année MDCCLXXXVIII [1787]”, in *Nova Acta Acadamiae Scientiarum Imperiales Petropolitanae* 6: 1-34. Petropoli.

Lytkin, V. I. 1962. “Über die Geschichte der syrjänischen Literatur”, in *Ural-Altäische Jahrbücher* 34: 214-22.

Pillu, Kaljo. 1990. *Tallinna Kunstiülikooli kümme soome-ugri uurimisreisi*. Tallinn: Olion. ISBN 5-450-00582-2

Stipa, Günter Johannes. 1961. “Der Ursprung der permischen Schrift”, in *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 110: 342-64. Wiesbaden.

Stipa, Günter Johannes. 1963. “Der Ursprung der permischen Schrift”, in *Congressus internationalis Fenno-ugristarum Budapestini habitus 20-24*. Pp. 281-87. Budapest.

Stipa, Günter Johannes. 1990. *Finnishc-ugrische Sprachforschung von der Renaissance bis zum Neupositivismus*. (Mémoires de la Société Finno-ougrienne; 206) Helsinki: Suomalais-ugrilainen Seura. ISBN 951-9403-35-3

Кузнецова, З. И. 1958. “Обзор памятников коми письменности XVIII в.”, in *Историко-филологический сборник* 4: 213-41. Сыктывкар.

Лыткин, В. И. 1952. *Древнепермский язык: чтение текстов, грамматика, словарь*. Москва.

Лыткин, В. И. 1957. *Историческая грамматика коми языка*. 1: Введение. Фонетика. Сыктывкар.

Лыткин, Г. С. 1889. *Зырянский Край при эпископах пермских и зырянский язык*. Санктпетербург.

Сидоров, А. С. 1962. “Новые памятники древнекоми письменности”. Ed. В. И. Лыткин. In *Вопросы* 1962: 178-211.

Эпифаний Премудрый. 1862. “Повесть о Степане, эпископе Пермском”, in *Памятники старинной русской литературы* 4: 119-71. С. Петербург.

Proposal for the Old Permic Script

	035	036	037
0			
1			
2			
3			
4			
5			
6			
7			
8			
9			
A			
B			
C			
D			
E			
F			

G = 00
P = 01

TABLE XXX - Row 03: OLD PERMIC

dec	hex	Name	dec	hex	Name
	50	OLD PERMIC LETTER AN			
	51	OLD PERMIC LETTER BUR			
	52	OLD PERMIC LETTER GAI			
	53	OLD PERMIC LETTER DOI			
	54	OLD PERMIC LETTER E			
	55	OLD PERMIC LETTER ZHOI			
	56	OLD PERMIC LETTER DZHOI			
	57	OLD PERMIC LETTER ZATA			
	58	OLD PERMIC LETTER DZITA			
	59	OLD PERMIC LETTER I			
	5A	OLD PERMIC LETTER KOKE			
	5B	OLD PERMIC LETTER LEI			
	5C	OLD PERMIC LETTER MENOE			
	5D	OLD PERMIC LETTER NENOE			
	5E	OLD PERMIC LETTER VOOI			
	5F	OLD PERMIC LETTER PEEI			
	60	OLD PERMIC LETTER REI			
	61	OLD PERMIC LETTER SII			
	62	OLD PERMIC LETTER TAI			
	63	OLD PERMIC LETTER U			
	64	OLD PERMIC LETTER CHERY			
	65	OLD PERMIC LETTER SHOOI			
	66	OLD PERMIC LETTER SHCHOOI			
	67	OLD PERMIC LETTER YRY			
	68	OLD PERMIC LETTER YERU			
	69	OLD PERMIC LETTER O			
	6A	OLD PERMIC LETTER OO			
	6B	OLD PERMIC LETTER EF			
	6C	OLD PERMIC LETTER HA			
	6D	OLD PERMIC LETTER TSIU			
	6E	OLD PERMIC LETTER VER			
	6F	OLD PERMIC LETTER YER			
	70	OLD PERMIC LETTER YERI			
	71	OLD PERMIC LETTER YAT			
	72	OLD PERMIC LETTER IE			
	73	OLD PERMIC LETTER YU			
	74	OLD PERMIC LETTER YA			
	75	OLD PERMIC LETTER IA			
	76	(This position shall not be used)			
	77	(This position shall not be used)			
	78	(This position shall not be used)			
	79	(This position shall not be used)			
	7A	(This position shall not be used)			
	7B	(This position shall not be used)			
	7C	(This position shall not be used)			
	7D	(This position shall not be used)			
	7E	(This position shall not be used)			
	7F	(This position shall not be used)			

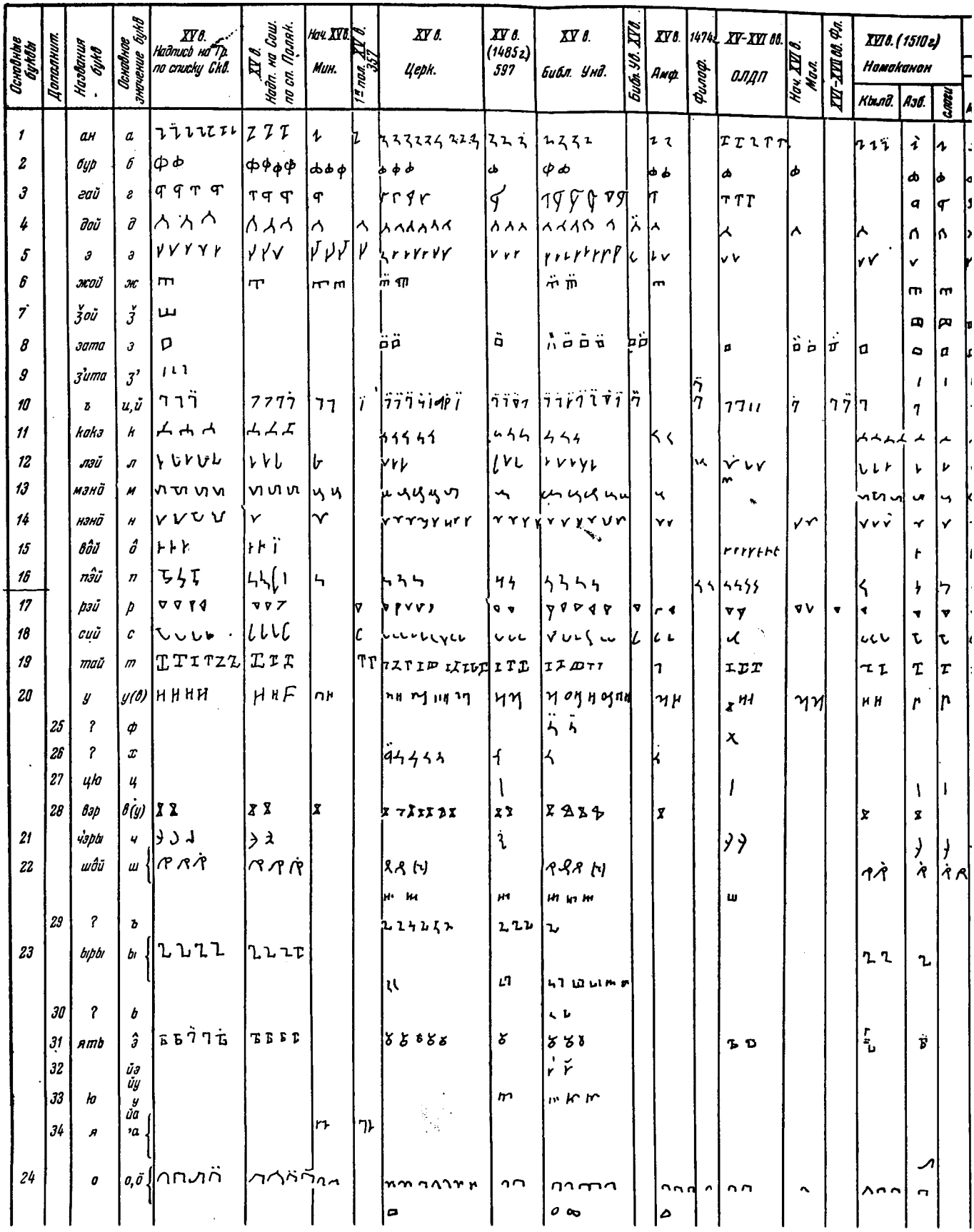

Figure 1. Part 1 of Table from Лыткин 1952, showing glyph variants of Old Permic characters. Note columns 1 and 2 which give ordering of primary (1–24) and secondary (25–34) characters.

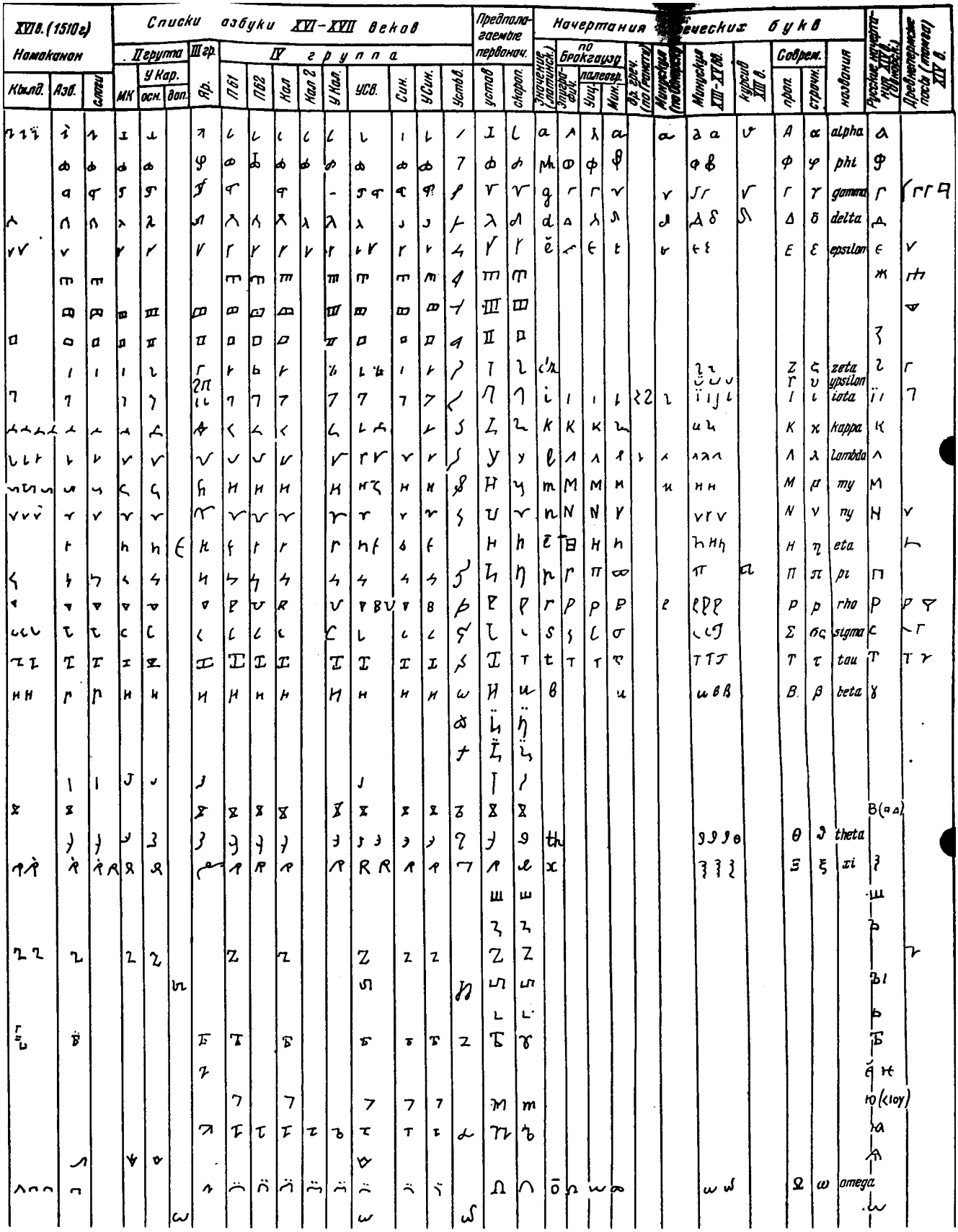

Figure 2. Part 2 of Table from Лыткин 1952, showing glyph variants of Old Permic characters. Note that the first three columns from the Nomokanon 1510 are repeated from Part 1 of the table.

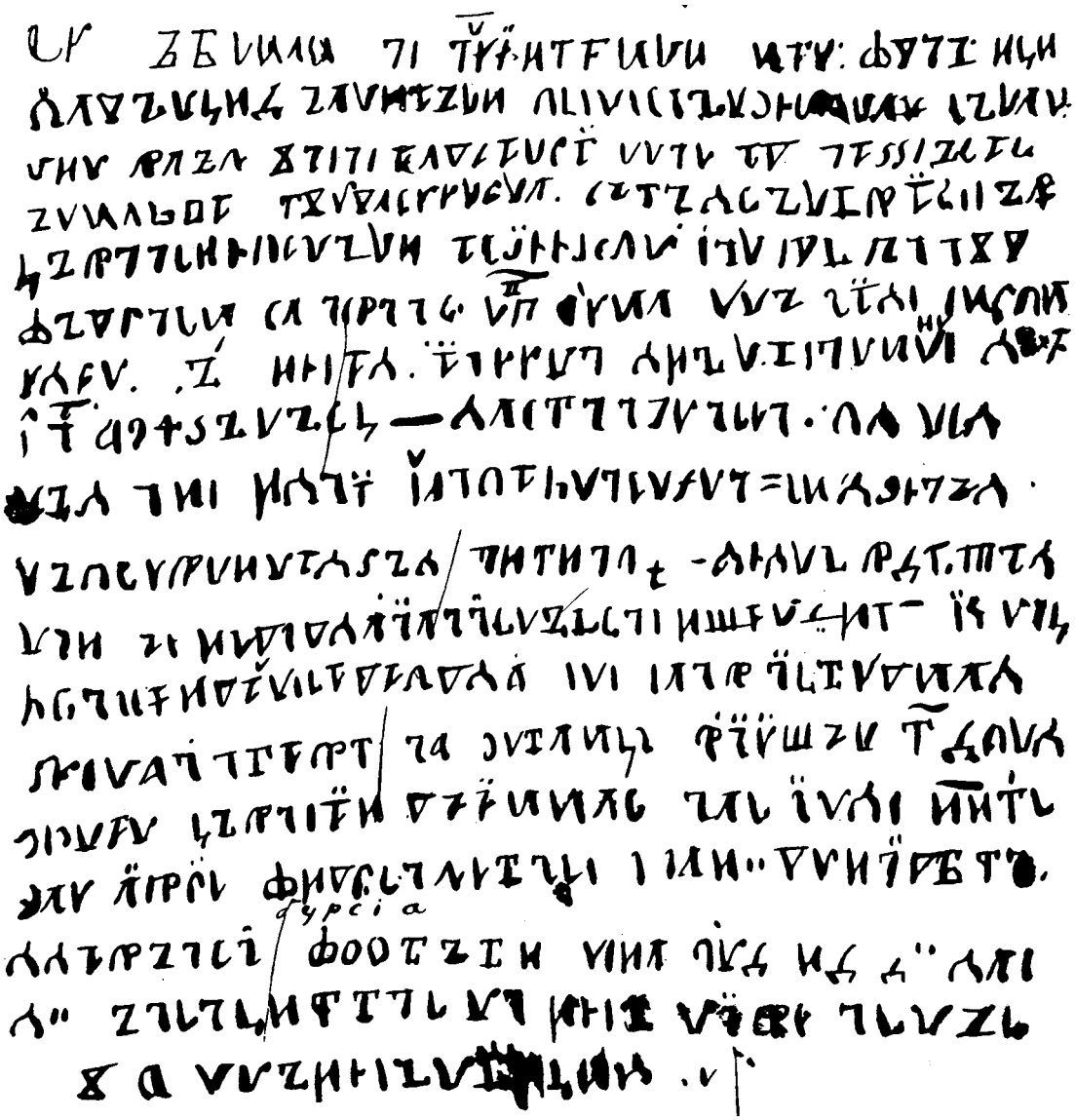

Figure 3. Old Permic inscription on the icon “Troitsy”. The text is the beginning of a story about Abraham and Sarah (Genesis 18:1–8), from Лыткин 1952.

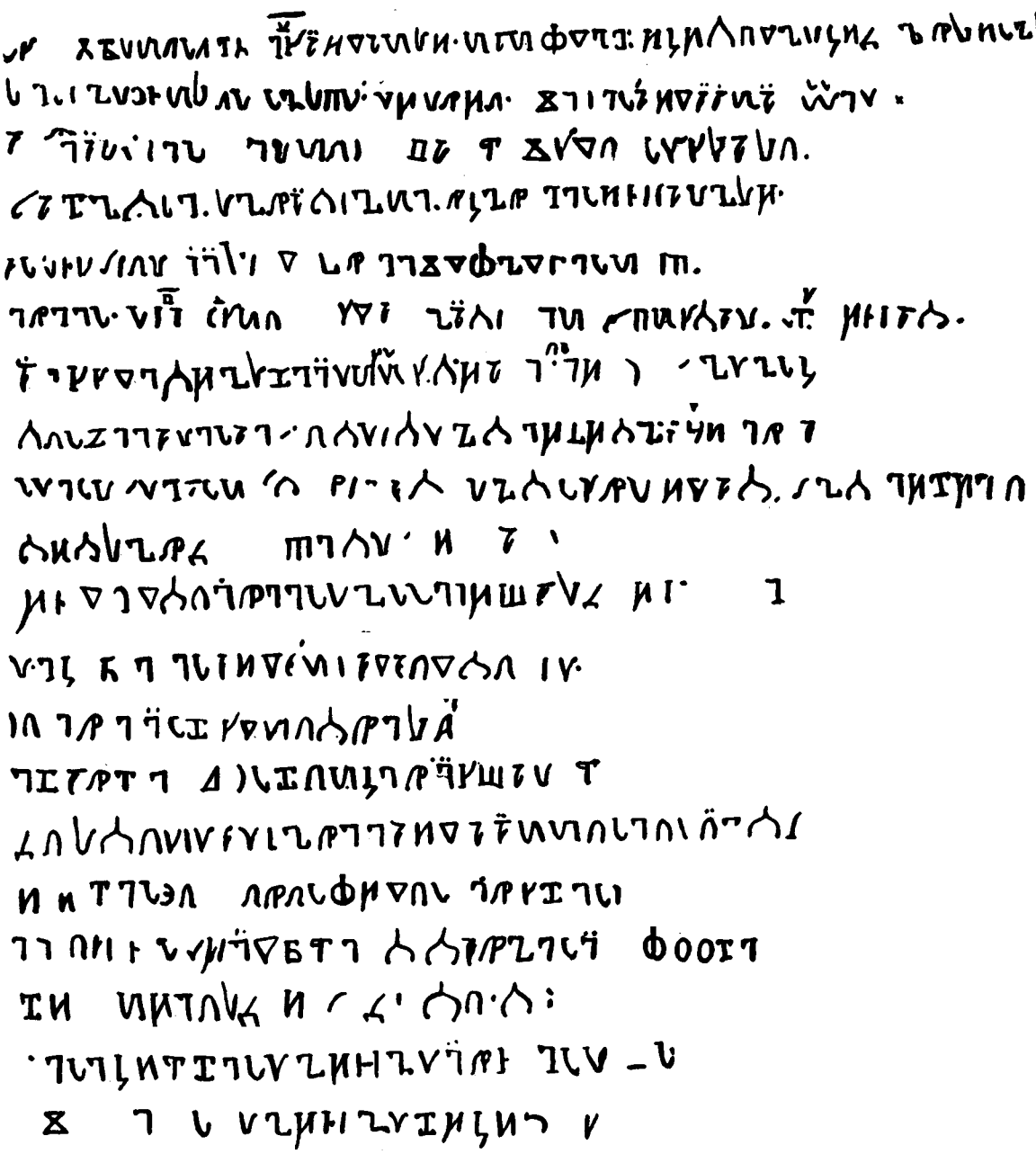

Figure 4. Скворцовский's transcription of the Old Permic inscription on the icon "Troitsy" from Лыткин 1952. Linebreaks here were to get a 40cm by 5cm text to fit on an A4 sheet.

Наше воспроизведение*

сэс'[5] вэ̂лмöсс'ис[6] йэн аврамлу мамбри тупу дорын, пукыс'лу сылу öз'öс[10] вôз'ын ч'ôмлöн[6] сылöн луншöрö.

виз'ис авраам ас с'инйасöн и адз'ис ас с'инмöн: вôз'з'а 3 вэрöс сулалö мадыд сылыс';

адз'ымыс' пышйис вôз'з'а нылу ас ч'ôмлöн öйыс[17] дорыс' и йурбыртис муöз'.

и шиис: мз (мэзöсö) йэнмö бара-ö адз'им кö мэдалсö тэ вôз'ад, ас вэрид вылти ин мун.

мэд вайс'ас ва мыс'кыныс пôдöс тийан и сайкöдч'адныд тупу дынын и вайс'аснис[20] н'ан'йас[22], мэд с'ôйадныд сэс' мунадныд ту туйö, кодудлыс' кэ̂жидныд ас вэр ордö. и шиисныс: сиз' уз̆ал, куз̆'ик шиин.

и пышйис аврам сара ордö ч'ôмйö и шиис: тэрмöдс'и, лой 3 тас'ти сöстöм пыс' и уз̆ал 3 кöлдöм н'ан'.

пышйи авраам мöсйöс ордö, кутис ч'öл öшöс, бурöс, и с'этис сийö вэрлы, и рэ̂гыд дас'тис.

и бос'тис вый, ма, йöл, кук, кôдöс дас'тис и пугтис ны вôз'ын; и с'ôисныс; ач'ыс виз'ис ны вôз'ын тупу дынын.

Наше чтение*

1* Сэ(ш) вѣлмос(ш)ис[7] иэн аурамлу[8] мамбри тупу дорын пукышлу[9] сылу ö(з'о)с у(ô)з̆'ын чôмлон сылон луншо(ро)[11]/виз'ис а̇ураäм[12] ä(с) (ш)ин(и)а(сон) й[13] ä(д)з'ис äс (ш)инмон (уô)з̆'з'а[14]

2. г̄ верос сулало (м)атыд сылыш/äдз'ы мыш[15] пышиис уôз'з'а нылу[9] а̇с[16] чôмлон öйс д(о)рыш и[13] иврбыртис[8] муо(з̆')/й[13] шиис мз̄ иэнмо (ба)ра о äдз'им ко мэдал (сö)[18] тэ

3. уôз̆'ад ä(с) уэрид[19] уылти йн мун/мэд уаиша(с) у(а) (мыш'к)ыныс п(ô)дос тииан. й[13] сай(к)од(ча)дныд[21] (т)упу ды(ны)н (и) у(а)ишаснис[19] н(а)ниас м(э)д шôиадныд сэш мунад(н)ыд

4. ту[23] туио (кô)дудлыш кѣжидн(ы)д[24] а(с) уэр ордо й[25] шийсныс сиз' уз̆ал куз'(ик)[26] (ш)и(й)н/и пы(ш)иис а̇ура̇м сара ордо (ч)ôмио и шийс тэрмодши[27] лои

5. г таштн (со)стом пыш[28] й уз̆ал г̄ колдом н̇ан/пышии äураäм мосиос öрдо (к)утис чол öшöс бурос й шэтис сиио уэрлу[9] й рѣгыд[29] даштис/й[10]

6. бо(ш)ти(с) у(ыи) м(а) иол кук к(ô)дос(с) д(аштс)ис и пугтис ны уôз'ын й шôисныс (а̇чыс) в(из')ис ны уôз̆'ын тупу д(ы)н(ын).

Figure 5. Лыткин's Cyrillic transcription of the Old Permic inscription on the icon "Troitsy". On the left ("our reproduction") is the normalized Komi text; on the right ("our reading") is the literal transcription, with additions and corrections in parentheses.

Figure 6. My typeset version of Лыткин's Cyrillic literal transcription of the Old Permic inscription on the icon "Troitsy". The linebreaks in the text above are the same as those of the text in Figure 3 (hyphens have been added at line-final word breaks). Note the use of COMBINING DOT ABOVE, COMBINING DIAERESIS, COMBINING COMMA ABOVE, and CYRILLIC COMBINING TITLO.

18 ісѵѐ ѯѣѵѵпісѐпі ѿѵѵ ѯѯѹѯѵѵір ѵѯѵѡѹѹ ѿірір ѧпѹіѵ,
ірдіѐір ісіір ӥіӥі рѡіісѵ ѭѡѵіӥѵ ісіір ірѵѧѐӥѹӥ.

2ѯгігі ѯѯѹѯѯѵ ѯі ѐгѵѿѯіӥѵ г ѯѧігі ѯі ѐгѵѵѵӥѵ: ѯѡііѯ 3 ѯѵѹӥі
ірѵѯѵӥ ѵѯѧіѧ ісіѵіѐ; ѯѧіісѵіѐ ѵіѧѐѿгі ѯѡііѯ ѵіір ѯі ѭѡѵіѵӥѵ
ӥӥіі ѧпѹіѐ г ѿірѹѡіѹѿігі ѵірӥі.

3г ѧгігі: “ѵѵ (ѵѵ҃ӥіі) ѿѵѵѵӥ ѡѯѹѯ-ӥ ѯѧігѵ ѯӥ ѵѵѧѯѵіӥ ѿѵ
ѯѡіѯѧ ѯі ѯѵѹгѧ ѯіѵѿігі ѵірѵ.

4"ѵѵѧ ѯѯѿѐѯі ѯѯ ѵіѐѯіѐіѵіі ѵѡѧӥі ѿгѿѯѵ г іѯѿіѐіѭіѯѧіѵіѧ
ірір ѧіѵіѵ г ѯѯѿѐѯіѵгі ѵѯѵѿѯі, ѵѵѧ ѐѡѿѯѧіѵіѧ іѵѐ
ѵірѵѯѧіѵіѧ ір ірѿӥ, ідѧірѧѵіѐ ідѣгігѧіѵіѧ ѯі ѯѵѹ пѹѧӥ.” г
ѧгігіѵіі: “ігі рѭѯѵ, ідірігд ѧгігѵ.”

6г ѵіѐѧѿгі ѯѯѹѯѵ іѯѹѯ пѹѧӥ ѭѡѵѿӥ г ѧгігі: “ірѵѹѵӥѧѐг, ѵпѿ 3
ірѯѐірі іӥіірӥѵ ѵіѐ г рѭѯѵ 3 ідӥѵѧӥѵ ѵѯѵ!”

7ѵіѐѧѿг ѯѯѹѯѯѵ ѵӥіѿӥі пѹѧӥ, ідрірі ѭӥѵ ӥѧӥі, ѡрѹӥі, г ѐѵірі
іг҃ӥ ѯѵѹѵіі, г ѹѣтіѧ ѧѯѐірі.

8г ѡпѐірі ѯіѿ, ѵѯ, ѿӥѵ, ідід, ідѧӥі ѧѯѐірі г ѵіртірі ѵіі
рѡіісѵ; г ѐѡгіѵіі; ѯѭіі ѯігі ѵіі ѯѡіісѵ ірѵір ѧіѵіѵ.

Figure 7. My typeset version of Лыткин’s Cyrillic normalized transcription of the Old Permic inscription on the icon “Troitsy”. It is possible that this is the first time that a text in this script has ever been typeset with fonts. Linebreaks, punctuation, and numbering follow normal biblical editorial practice. Note the use of COMBINING DOT ABOVE, COMBINING DIAERESIS, and COMBINING BREVE here.