

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. Naarmann calls it “Abur” and “Altsyryänisch” (= ‘Old Komi’); “Abur” is a name derived from the names of the first two characters AN and BUR (so $\mathfrak{z}\mathfrak{v}\text{-}\mathfrak{d}\mathfrak{r}\mathfrak{v}$ is equivalent to $\alpha\lambda\phi\alpha\text{-}\beta\eta\tau\alpha$ and азбуки).

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 \mathfrak{r} and SHCHOOI \mathfrak{w}), 0367/0368 (YRY \mathfrak{r} and YERU \mathfrak{w}), 0369/036A (O \mathfrak{r} and OO \mathfrak{w}), and 0374/0375 (YA \mathfrak{r} and IA \mathfrak{w}). 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, \mathfrak{w} 036B (EF) and \mathfrak{r} 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: \mathfrak{h} is represented by 035F PEEI \mathfrak{h} + ¨ and \mathfrak{z} is represented by 03FA KOKE \mathfrak{z} + ¨ . Note the similarity of \mathfrak{w} and \mathfrak{r} to Greek ω 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 \mathfrak{u} 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.)
- Lepetchin, I. I. 1790. “Histoire de l’Académie Imperiale des Sciences année MDCCLXXXVIII [1787]”, in *Nova Acta Academiae 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. *Finnisch-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 епископах 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, епископе Пермском”, in *Памятники старинной русской литературы* 4: 119-71. С. Петербург.

Proposal for the Old Permic Script

	035	036	037
0	𐀀	𐀁	𐀂
1	𐀃	𐀄	𐀅
2	𐀆	𐀇	𐀈
3	𐀉	𐀊	𐀋
4	𐀌	𐀍	𐀎
5	𐀏	𐀐	𐀑
6	𐀒	𐀓	
7	𐀔	𐀕	
8	𐀖	𐀗	
9	𐀘	𐀙	
A	𐀚	𐀛	
B	𐀜	𐀝	
C	𐀞	𐀟	
D	𐀠	𐀡	
E	𐀢	𐀣	
F	𐀤	𐀥	

G = 00
P = 01

Наше воспроизведение*

сэс'⁵ вѣлмѳсс'ис⁶ йѣн аврамлу
мамбри тупу дорын, пукыс'лу
сылу ѳз'ѳс¹⁰ вѳз'ын ч'ѳмлѳн⁶
сылѳн луншѳрѳ.

виѳ'ис авраам ас с'инйасѳн и
адѳ'ис ас с'инмѳн: вѳз'з'а 3 вѳрѳс
сулалѳ мадыд сылыс';

адѳ'ымыс' пышйис вѳз'з'а нылу
ас ч'ѳмлѳн ѳйыс¹⁷ дорыс' и
йурбыртис муѳз'.

и шиис: мз (мѳзѳсѳ) йѣнмѳ
бара-ѳ адз'им кѳ мѳдалсѳ тѳ
вѳз'ад, ас вѳрид вылти ин мун.

мѳд вайс'ас ва мыс'кыныс
пѳдѳс тийан и сайкѳдч'адныд ту-
пу дынын и вайс'аснис²⁰ н'ан'йас²²,
мѳд с'ѳйадныд сэс' мунадныд ту
туйѳ, кодудлыс' кѳжидныд ас вѳр
ордѳ. и шиисныс: сиз' уѳзал, куз'ик
шинн.

и пышйис аврам сара ордѳ
ч'ѳмйѳ и шиис: тѳрмѳдс'и, лой
3 тас'ти сѳстѳм пыс' и уѳзал
3 кѳлдѳм н'ан'.

пышйи авраам мѳсйѳс ордѳ,
кутис ч'ѳл ѳшѳс, бурѳс, и с'ѳтис
сийѳ вѳрлы, и рѳгыд дас'тис.

и бос'тис вый, ма, йѳл, кук,
кѳдѳс дас'тис и пугтис ны вѳз'ын;
и с'ѳисныс; ач'ыс виз'ис ны вѳз'ын
тупу дынын.

Наше чтение*

1* Сэ(ш) вѣлмос(ш)ис⁷ иѣн
аурамлу⁸ мамбри тупу дорын
пукышлу⁹ сылу ѳ(з'о)с у(ѳ)з'ын
чѳмлѳн сылон луншо(ро)¹¹/виѳ'ис
ѳураѳм¹² ѳ(с) (ш)ин(и)а(сон) й¹³
ѳ(д)з'ис ѳс (ш)инмон (уѳ)з'з'а¹⁴

2. Г̄ верос сулало (м)атыд сы-
лыш/ѳдз'ы мыш¹⁵ пышиис уѳз'з'а
нылу⁹ ѳс¹⁶ чѳмлѳн ѳйис д(о)рыш
и¹³ иврбыртис⁸ муѳ(з')/й¹³ шиис
м̄з иѣнмо (ба)ра о ѳдз'им ко мѳ-
дал (сѳ)¹⁸ тѳ

3. уѳз'ад ѳ(с) уѳрид¹⁹ уылти
йн мун/мѳд уаиша(с) у(а) (мыш'к)ы-
ныс п(ѳ)дос тииан. й¹³ сай(к)од-
(ча)дныд²¹ (т)упу ды(ны)н (и)
у(а)ишаснис¹⁹ н(а)ниас м(ѳ)д шѳи-
адныд сэш мунад(н)ыд

4. ту²³ туио (кѳ)дудлыш
кѳжидн(ы)д²⁴ а(с) уѳр ордо й²⁵
шийсныс сиз' уѳзал куз'(ик)²⁶
(ш)н(й)н/и пы(ш)иис ѳураѳм сара
ордо (ч)ѳмио и шийс тѳрмѳдши²⁷
лои

5. г̄ ташти (со)стом пыш²⁸ й
уѳзал Г̄ колдом ѳан /пышии
ѳураѳм мосиос ѳрдо (к)утис чѳл
ѳшѳс бурѳс й шѳтис сиио уѳрлу⁹
й рѳгыд²⁹ даштис/й¹⁰

6. бо(ш)ти(с) у(ыи) м(а) иол
кук к(ѳ)дос(с) д(ашт)ис и пугтис
ны уѳз'ын й шѳисныс (ачыс)
в(из')ис ны уѳз'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.

