Proposal to encode 23 Cyrillic characters for old Uslar's Caucasian Alphabets

in ISO/IEC 10646 and The Unicode Standard

Nikita Manulov, pacancuvak15@gmail.com

2022-09-18

1.Introduction

In this document, it is proposed to add 23 additional Cyrillic letters that were in the pre-revolutionary Caucasian alphabets of Uslar (Abkhazian, Lak, Tabasaran, Avar, Chechen, Dargin, Lezgin, Tsakhur alphabets). Some of the letters are similar to Georgian and Armenian letters.

Note: The Tabasaran alphabet was not used due to the death of the author, it remained in the manuscript and was never used in practice.

2.Code Charts

Cyrillic Extended-D	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
U+1E08x	Ҧ	ҧ	Ҧ	ҧ	Ҕ	ҕ	ҕ	Ҟ	ҟ	ҹ	Ҕ	ҕ	Ҕ	ҕ	Ц	
U+1E09x	Ц	Ȟ	ȟ	ȟ	ŋ	ŋ	∩	∩								

U+1E080 CYRILLIC CAPITAL LETTER SHHA WITH HIGH RIGHT BREVE SERIF

U+1E081 CYRILLIC SMALL LETTER SHHA WITH HIGH RIGHT BREVE SERIF

U+1E082 CYRILLIC CAPITAL LETTER SHHA WITH CIL TOP

U+1E083 CYRILLIC SMALL LETTER SHHA WITH CIL TOP

U+1E084 CYRILLIC SMALL LETTER DJE WITH HIGH RIGHT BREVE SERIF

U+1E085 CYRILLIC CAPITAL LETTER SHHA WITH HOOK

U+1E086 CYRILLIC SMALL LETTER SHHA WITH HOOK

U+1E087 CYRILLIC CAPITAL LETTER KA WITH LOOP

U+1E088 CYRILLIC SMALL LETTER KA WITH LOOP

U+1E089 CYRILLIC CAPITAL LETTER ENLARGED SMALL QA

U+1E08A CYRILLIC CAPITAL LETTER CIL

→ ჿ U+10EC GEORGIAN LETTER CIL

U+1E08B CYRILLIC SMALL LETTER CIL

→ ჶ U+10EC GEORGIAN LETTER CIL

U+1E08C CYRILLIC CAPITAL LETTER CIL WITH BAR

U+1E08D CYRILLIC SMALL LETTER CIL WITH BAR

U+1E08E CYRILLIC CAPITAL LETTER TSE WITH LONG LEFT LEG

U+1E090 CYRILLIC SMALL LETTER TSE WITH LONG LEFT LEG

U+1E091 CYRILLIC CAPITAL LETTER CHAR

→ ჭ U+10ED GEORGIAN LETTER CHAR

U+1E092 CYRILLIC SMALL LETTER CHAR

→ ჭ U+10ED GEORGIAN LETTER CHAR

U+1E093 CYRILLIC SMALL LETTER CHAR WITH HIGH RIGHT BREVE SERIF

U+1E094 CYRILLIC CAPITAL LETTER GHAD

→ Ղ U+0552 ARMENIAN CAPITAL LETTER GHAD

U+1E095 CYRILLIC SMALL LETTER GHAD

→ ղ U+0572 ARMENIAN SMALL LETTER GHAD

U+1E096 CYRILLIC CAPITAL LETTER CURLED ROUNDTOP EL

U+1E097 CYRILLIC SMALL LETTER CURLED ROUNDTOP EL

3.Unicode Data

```
1E080; CYRILLIC CAPITAL LETTER SHHA WITH HIGH RIGHT BREVE
SERIF;Lu;0;L;;;;;N;;;;1E081;

1E081; CYRILLIC SMALL LETTER SHHA WITH HIGH RIGHT BREVE
SERIF;Ll;0;L;;;;;N;;;1E080;;1E080

1E082; CYRILLIC CAPITAL LETTER SHHA WITH CIL
TOP;Lu;0;L;;;;;N;;;;1E083;

1E083; CYRILLIC SMALL LETTER SHHA WITH CIL
TOP;Ll;0;L;;;;;N;;;1E082;;1E082

1E084; CYRILLIC SMALL LETTER DJE WITH HIGH RIGHT BREVE
SERIF;Lu;0;L;;;;;N;;;;;

1E085; CYRILLIC CAPITAL LETTER SHHA WITH
HOOK;Ll;0;L;;;;;N;;;1E086;

1E086; CYRILLIC SMALL LETTER SHHA WITH
HOOK;Lu;0;L;;;;;N;;;1E085;;1E085

1E087; CYRILLIC CAPITAL LETTER KA WITH
LOOP;;Ll;0;L;;;;;N;;;1E088;

1E088; CYRILLIC SMALL LETTER KA WITH LOOP;Lu;0;L;;;;;N;;;
;1E087;;1E087

1E089; CYRILLIC CAPITAL LETTER ENLARGED SMALL QA;Ll;0;L;;;;;N;;;;
```

```
1E08A;CYRILLIC CAPITAL LETTER CIL;Lu;0;L;;;;;N;;;;1E08B;

1E08B;CYRILLIC SMALL LETTER CIL;Ll;0;L;;;;;N;;;1E08A;;1E08A

1E08C;CYRILLIC CAPITAL LETTER CIL WITH
BAR;Lu;0;L;;;;;N;;;;1E08D;

1E08D;CYRILLIC SMALL LETTER CIL WITH
BAR;Ll;0;L;;;;;N;;;1E08C;;1E08C

1E08E;CYRILLIC CAPITAL LETTER TSE WITH LONG LEFT
LEG;Lu;0;L;;;;;N;;;;1E090;

1E090;CYRILLIC SMALL LETTER TSE WITH LONG LEFT
LEG;Ll;0;L;;;;;N;;;1E08E;;1E08E

1E091;CYRILLIC CAPITAL LETTER CHAR;Lu;0;L;;;;;N;;;;1E092;

1E092;CYRILLIC SMALL LETTER CHAR;Ll;0;L;;;;;N;;;1E091;;1E091

1E093;CYRILLIC SMALL LETTER CHAR WITH HIGH RIGHT BREVE
SERIF;Lu;0;L;;;;;N;;;;;

1E094;CYRILLIC CAPITAL LETTER GHAD;Ll;0;L;;;;;N;;;1E095;

1E095;CYRILLIC SMALL LETTER GHAD;Lu;0;L;;;;;N;;;1E094;;1E094

1E096;CYRILLIC CAPITAL LETTER CURLED ROUNDTOP
EL;Ll;0;L;;;;;N;;;1E097;

1E097;CYRILLIC SMALL LETTER CURLED ROUNDTOP
EL;Lu;0;L;;;;;N;;;1E096;;1E096
```

4.Figures

Figure 1. Sample from Russian-Avar Book «Книжка для чтения на аварском и русском языках. Казань, 1899»
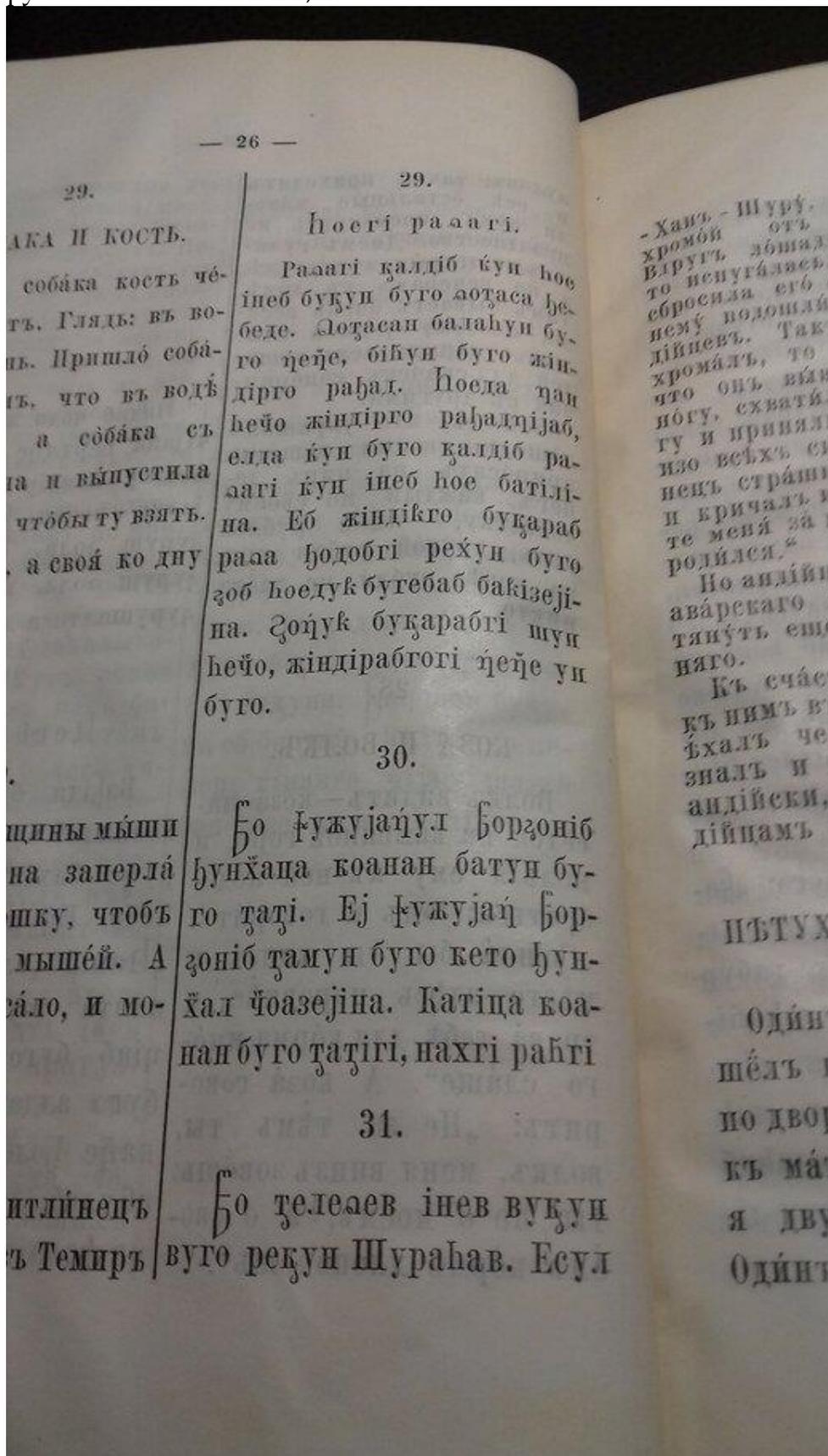
КНИЖКА

ДЛЯ ЧТЕНІЯ

НА РУССКОМЪ И АВАРСКОМЪ ЯЗЫКАХЪ.

Ђурус маџаҥгі маһарул маџаҥгі һалізе коараб жуз.

Изданіе Православнаго Миссіонерскаго Общества.

КАЗАНЬ.
Типо-литографія Императорскаго Университета.
1899.

Figure 2. Sample from «Кюринская азбука. Казанфар Зульфикаров. 1871 г.»

Figure 3. Sample from «Кюринская азбука и первая книга для чтения. Тифлис, издание Кавказского Учебного Округа, 1911»
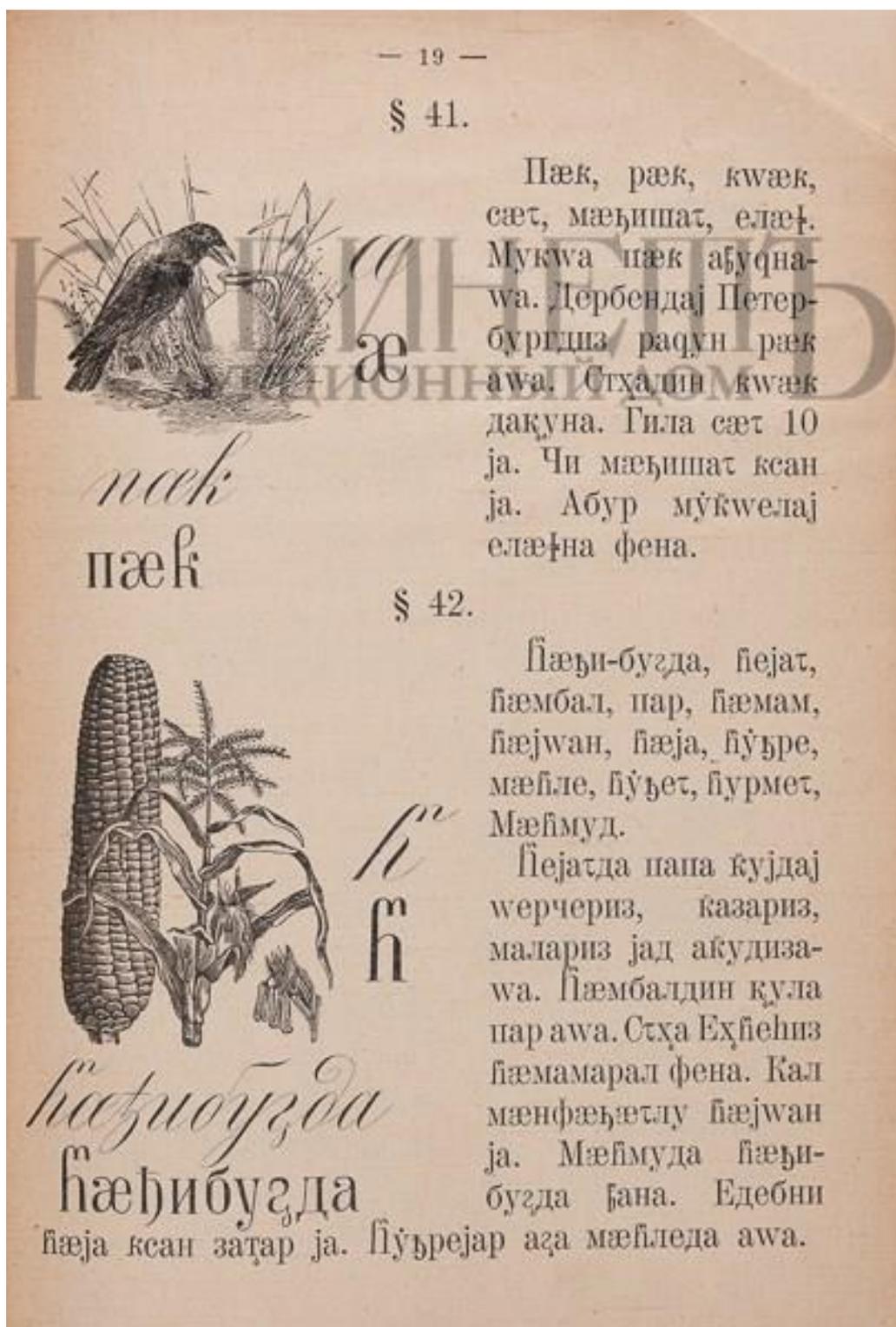
5.References

1. Книжка для чтения на аварском и русском языках. Казань, 1899

2. Кюринская азбука. Казанфар Зульфикаров. 1871 г.

3. Кюринская азбука и первая книга для чтения. Тифлис, издание Кавказского Учебного Округа, 1911

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://www.dkuug.dk/JTC1/SC2/WG/docs/principles.html **for guidelines and details before filling this form.**
Please ensure you are using the latest Form from http://www.dkuug.dk/JTC1/SC2/WG/docs/summaryform.html**.**
See also http://www.dkuug.dk/JTC1/SC2/WG/docs/roadmaps.html **for latest *Roadmaps.***

A. Administrative

1. **Title:** *Proposal to encode 23 Cyrillic characters for old Uslar's Caucasian Alphabets*
2. Requester's name: *Nikita Manulov*
3. Requester type (Member body/Liaison/Individual contribution): *Individual contribution*
4. Submission date: *2022-09-18*
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: *No*

B. Technical – General

1. Choose one of the following:
   a. This proposal is for a new script (set of characters): *No*
      Proposed name of script:
   b. The proposal is for addition of character(s) to an existing block:
      Name of the existing block: *Cyrillic Extended-D*
2. Number of characters in proposal:
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?
   b. Are the character shapes attached in a legible form suitable for review?
5. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?
   If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:
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 http://www.unicode.org/Public/UNIDATA/UCD.html and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

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? *Wikipedia editors*	
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 proposal*	
4. The context of use for the proposed characters (type of use; common or rare)	*Historic*
Reference: *See proposal*	
5. Are the proposed characters in current use by the user community?	No
If YES, where? Reference:	
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?	*No*
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)?	
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 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?	
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 character(s)?	*No*
If YES, is the equivalent corresponding unified ideographic character(s) identified?	
If YES, reference:	