L2/22-154

Proposal to encode 18 Cyrillic characters for old Bashkir

in ISO/IEC 10646 and The Unicode Standard

Nikita Manulov, pacancuvak15@gmail.com

2022-06-25

1.Introduction

In 1869, Russian linguist Mirsalikh Bekchurin published the first guide to Bashkir grammar in the book An Initial Guide to the Study of Arabic, Persian and Tatar Languages with the Adverbs of Bukhara, Bashkirs, Kyrgyz and Residents of Turkestan. The first Cyrillic Bashkir introductory book was published by Vasily Katarinsky in Orenburg in 1892, with his proposed alphabet excluding the letters ё, й, ө, and ѵ from the contemporaneous Cyrillic alphabet and including the additional characters of Ӓ, ґ, ҥ, Ӧ, ӳ. Another primer was prepared at the end of the 19th century by Nikolaï Katanov [de] using the umlaut (ӓ – / ə /, ӧ – / ө /, ӟ – / ʒ /, ӄ – / ҡ /, ӳ – / и /, etc.), however this work was never published.

In 1907, Alexander Bessonov published The Primer for the Bashkirs. This publication proposed that alphabet included all the letters of the Cyrillic alphabet of the time, except for ё and й, and added the characters ӓ, ґ, д [fr], ҥ, ӧ, ҫ [fr], and ӳ. Five years later(1912), Mstislav Kulaev (Mukhametkhan Kulaev) published The Basics of Onomatopoeia and the Alphabet for Bashkirs (reprinted in 1919), again making use of the Cyrillic alphabet in conjunction with new characters.

2.Chart

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
U+1E06x															𞹮	𞹯
U+1E07x	𞹰	𞹱	𞹲	𞹳	𞹴	𞹵	𞹶	𞹷	𞹸	𞹹	𞹺	𞹻	𞹼	𞹽	𞹾	𞹿

3.Characters

𞹮 U+1E06E Cyrillic capital letter script A

𞹯 U+1E06F Cyrillic small letter script A

𞹰 U+1E070 Cyrillic capital letter split by middle ring Ghe

𞹱 U+1E071 Cyrillic small letter split by middle ring Ghe

𞹲 U+1E072 Cyrillic capital letter Je with inverted belt

𞹳 U+1E073 Cyrillic small letter Je with inverted belt

𞹴 U+1E074 Cyrillic capital letter byelorussian-ukrainian I with curve at bottom

𞹵 U+1E075 Cyrillic small letter byelorussian-ukrainian I with curve at bottom

𞹶 U+1E076 Cyrillic capital letter open at bottom O

𞹷 U+1E077 Cyrillic small letter open at bottom O

𞺈 U+1E078 Cyrillic capital letter open at left O

𞺉 U+1E079 Cyrillic small letter open at left O

𞺊 U+1E07A Cyrillic capital letter long Es

𞺋 U+1E07B Cyrillic small letter long Es

→ 017F ſ latin small letter long es

𞺌 U+1E07C Cyrillic capital letter script U

𞺍 U+1E07D Cyrillic small letter script U

𞺎 U+1E07E Cyrillic capital letter bashkir Ha

𞺏 U+1E07F Cyrillic small letter bashkir Ha

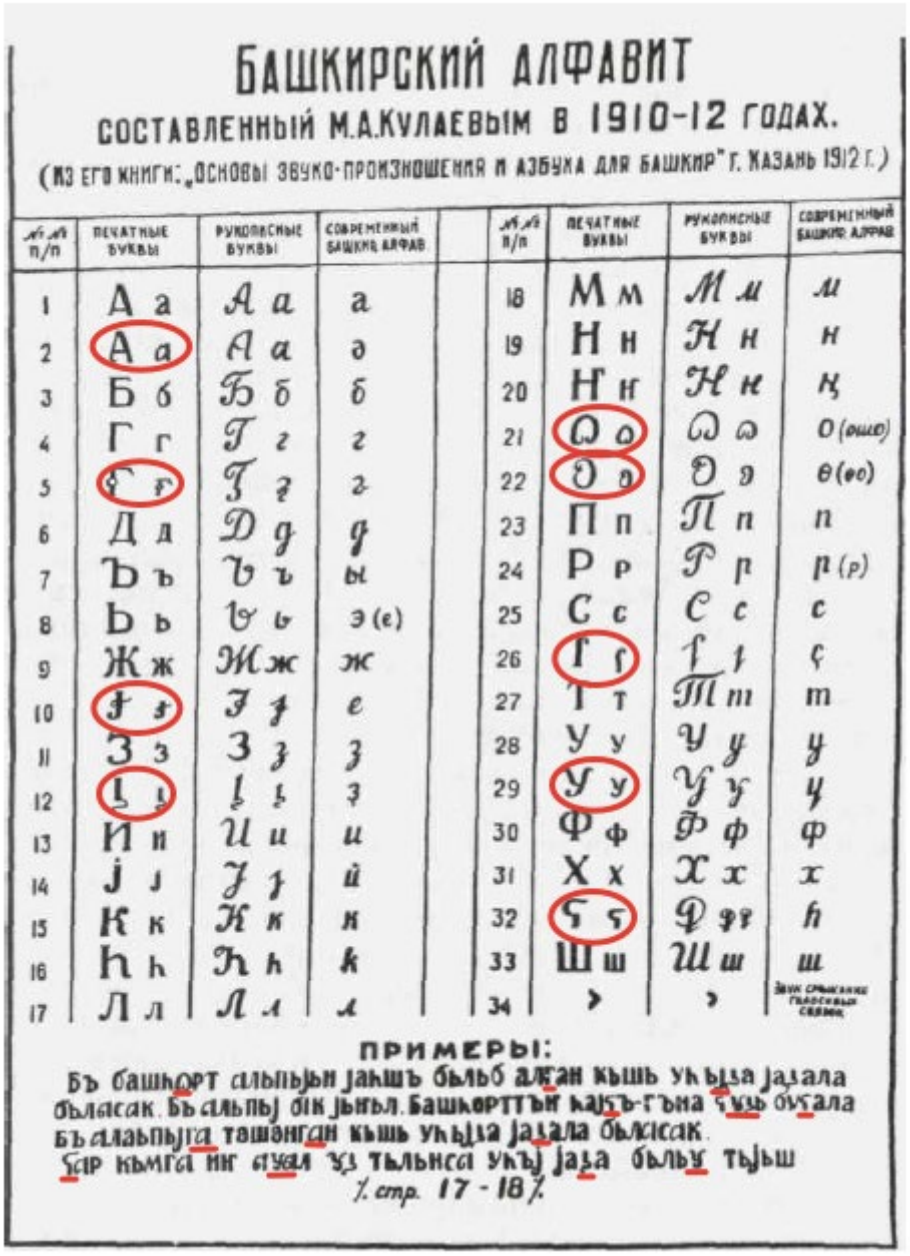
4.UnicodeData.txt

```
1E06E;CYRILLIC CAPITAL LETTER SCRIPT A;Lu;0;L;;;;;N;;;;1E06F;

1E06F;CYRILLIC SMALL LETTER SCRIPT A;Ll;0;L;;;;;N;;;1E06E;;1E06E

1E070;CYRILLIC CAPITAL LETTER SPLIT BY MIDDLE RING
GHE;Lu;0;L;;;;;N;;;;1E071;

1E071;CYRILLIC SMALL LETTER SPLIT BY MIDDLE RING
GHE;Ll;0;L;;;;;N;;;1E070;;1E070

1E072;CYRILLIC CAPITAL LETTER JE WITH INVERTED
BELT;Lu;0;L;;;;;N;;;;1E073;

1E073;CYRILLIC SMALL LETTER JE WITH INVERTED
BELT;Ll;0;L;;;;;N;;;1E072;;1E072

1E074;CYRILLIC CAPITAL LETTER BYELORUSSIAN-UKRAINIAN I WITH
CURVE AT BOTTOM;Lu;0;L;;;;;N;;;;1E075;

1E075;CYRILLIC SMALL LETTER BYELORUSSIAN-UKRAINIAN I WITH CURVE
AT BOTTOM;;Ll;0;L;;;;;N;;;1E074;;1E074

1E076;CYRILLIC CAPITAL LETTER OPEN AT BOTTOM
O;Lu;0;L;;;;;N;;;;1E077;

1E077;CYRILLIC SMALL LETTER OPEN AT BOTTOM
O;Ll;0;L;;;;;N;;;1E076;;1E076

1E078;CYRILLIC CAPITAL LETTER OPEN AT LEFT
O;Lu;0;L;;;;;N;;;;1E079;

1E079;CYRILLIC SMALL LETTER OPEN AT LEFT
O;Ll;0;L;;;;;N;;;1E078;;1E078

1E07A;CYRILLIC CAPITAL LETTER LONG ES;Lu;0;L;;;;;N;;;;1E07B;

1E07B;CYRILLIC SMALL LETTER LONG ES;Ll;0;L;;;;;N;;;1E07A;;1E07A

1E07C;CYRILLIC CAPITAL LETTER SCRIPT U;Lu;0;L;;;;;N;;;;1E07D;

1E07D;CYRILLIC SMALL LETTER SCRIPT U;Ll;0;L;;;;;N;;;1E07C;;1E07C

1E07E;CYRILLIC CAPITAL LETTER BASHKIR HA;Lu;0;L;;;;;N;;;;1E07F;

1E07F;CYRILLIC SMALL LETTER BASHKIR
HA;Ll;0;L;;;;;N;;;1E07E;;1E07E
```

5.Figures

Figure 1.

Cyrillic Bashkir alphabet of 1912 from the book by M.A. Kulaev "Basics of onomatopoeia and the alphabet for the Bashkirs" (republished in 1919)



БАШКИРСКИЙ АЛФАВИТ

СОСТАВЛЕННЫЙ М.А.КУЛАЕВЫМ В 1910-12 ГОДАХ.

(ИЗ ЕГО КНИГИ: „ОСНОВЫ ЗВУКО-ПРОИЗНОШЕНИЯ И АЗБУКА ДЛЯ БАШКИР" Г. КАЗАНЬ 1912 Г.)

№№ п/п	ПЕЧАТНЫЕ БУКВЫ	РУКОПИСНЫЕ БУКВЫ	СОВРЕМЕННЫЙ БАШКИР. АЛФАВ.		№№ п/п	ПЕЧАТНЫЕ БУКВЫ	РУКОПИСНЫЕ БУКВЫ	СОВРЕМЕННЫЙ БАШКИР. АЛФАВ.
1	А а	А а	а		18	М м	М м	м
2	Ӓ ӓ	Ӓ ӓ	ә		19	Н н	Н н	н
3	Б б	Б б	б		20	Ң ң	Ң ң	ң
4	Г г	Г г	г		21	Ө ө	Ө ө	ө (өшө)
5	Ғ ғ	Ғ ғ	ғ		22	Ө ө	Ө ө	ө (өө)
6	Д д	Д д	д		23	П п	П п	п
7	Ъ ъ	Ъ ъ	ы		24	Р р	Р р	р (р)
8	Ь ь	Ь ь	э (е)		25	С с	С с	с
9	Ж ж	Ж ж	ж		26	Ҫ ҫ	Ҫ ҫ	ҫ
10	Ҙ ҙ	Ҙ ҙ	е		27	Т т	Т т	т
11	З з	З з	з		28	У у	У у	у
12	Ҙ ҙ	Ҙ ҙ	ҙ		29	Ү ү	Ү ү	ү
13	И и	И и	и		30	Ф ф	Ф ф	ф
14	Й й	Й й	й		31	Х х	Х х	х
15	К к	К к	к		32	Ҫ ҫ	Ҫ ҫ	һ
16	Һ һ	Һ һ	һ		33	Ш ш	Ш ш	ш
17	Л л	Л л	л		34	ъ	ъ	Знак смягчения гласных звуков

ПРИМЕРЫ:

Бъ башҡорт алпыһын јаҡшъ быльб алған кьшь уқыјса јазала бълясак. Бъ алпыһ бік јьңыл. Башҡорттъң ҡајъ-гъна ғүмь оҙала бъ алдынъјға төшөнгән кьшь уқыјса јазала бълясак. Ҡар кьмға иң ябъал ҙз тьлънсә уқъј јаза быльу тьјьш

/. стр. 17 - 18 /.

Figure 2.

Handwritten letters of the Bashkir alphabet of Kulaev of 1912 from Хусаинова Л.М. Башкирское письмо монография. Стерлитамак Стерлитамакский филиал БашГУ, 2012 (Khusainova L.M. Bashkir letter monograph. Sterlitamak Sterlitamak branch of BashSU, 2012)

Figure 3.

Letters proposed by M.S. Kulaev

(from the book "Basics of Sound Pronunciation and the Alphabet for the Bashkirs"

Приложение № 18

Буквы, предложенные М.С. Кулаевым
(из книги «Основы звукопроизношения и азбука для башкир»)

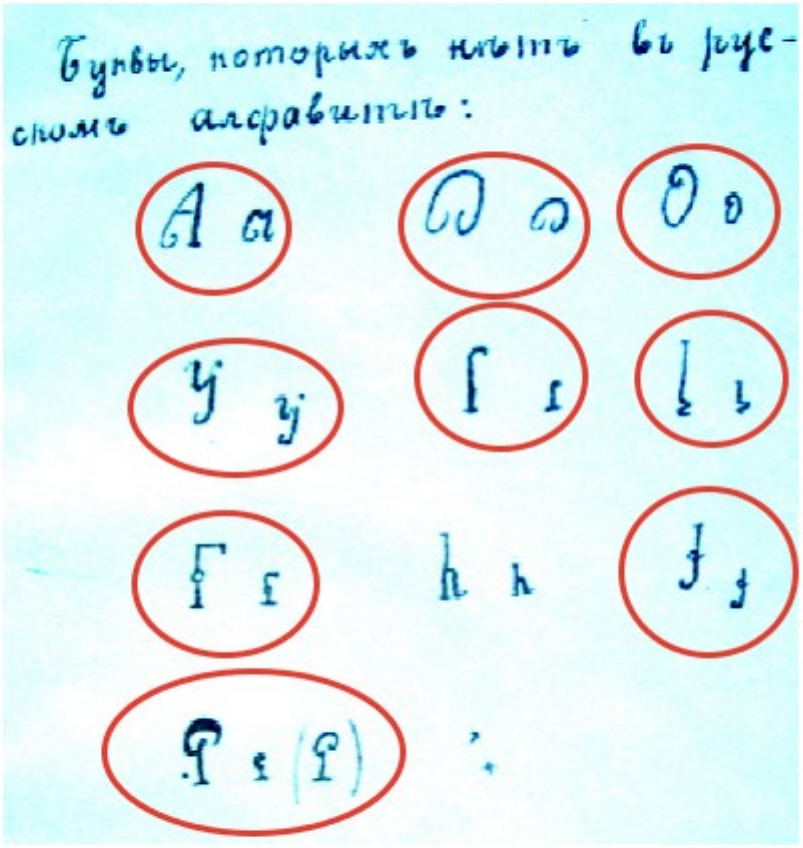


Figure 4.

Classification of sounds of the Bashkir language according to M.S. Kulaev

(from the book "Basics of Sound Pronunciation and the Alphabet for the Bashkirs")

Приложение № 21

Классификация звуков башкирского языка по М.С. Кулаеву
(из книги «Основы звукопроизношения и азбука для башкир»)

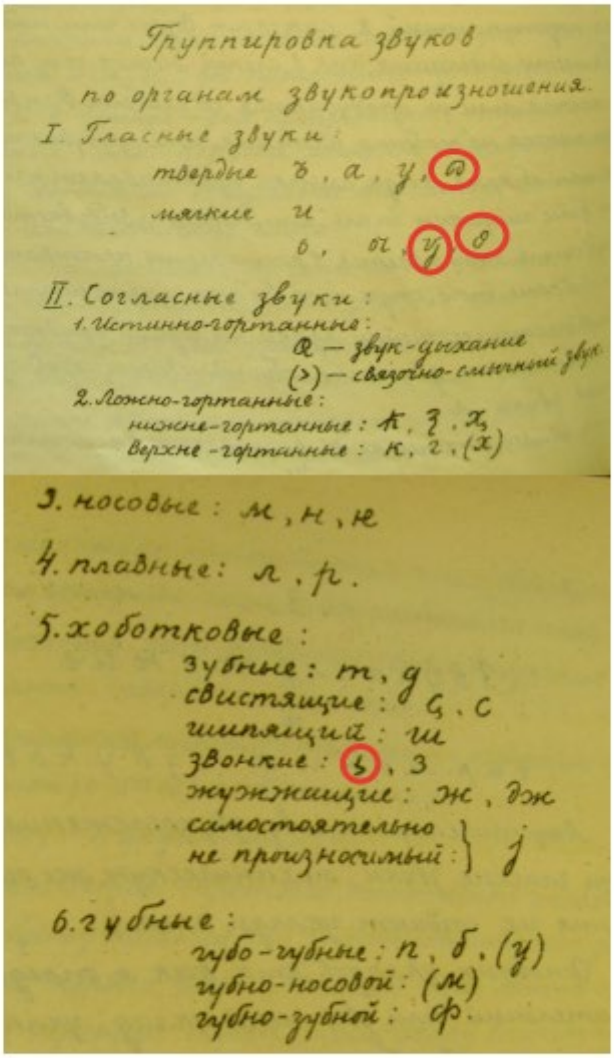
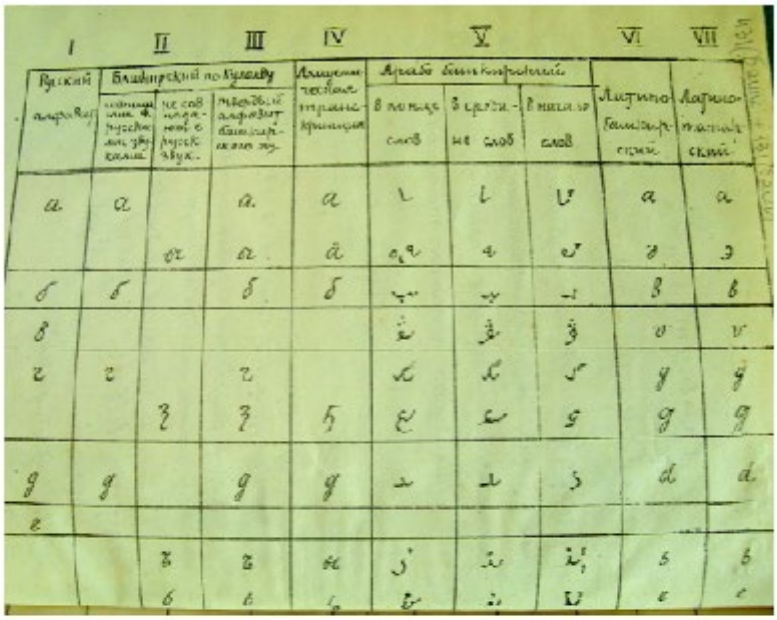


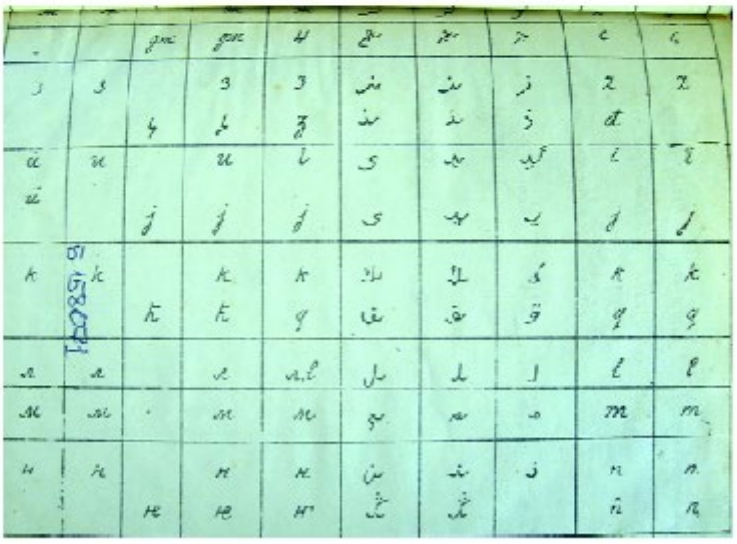
Figure 5.

Comparative table of alphabets compiled by M.S. Kulaev

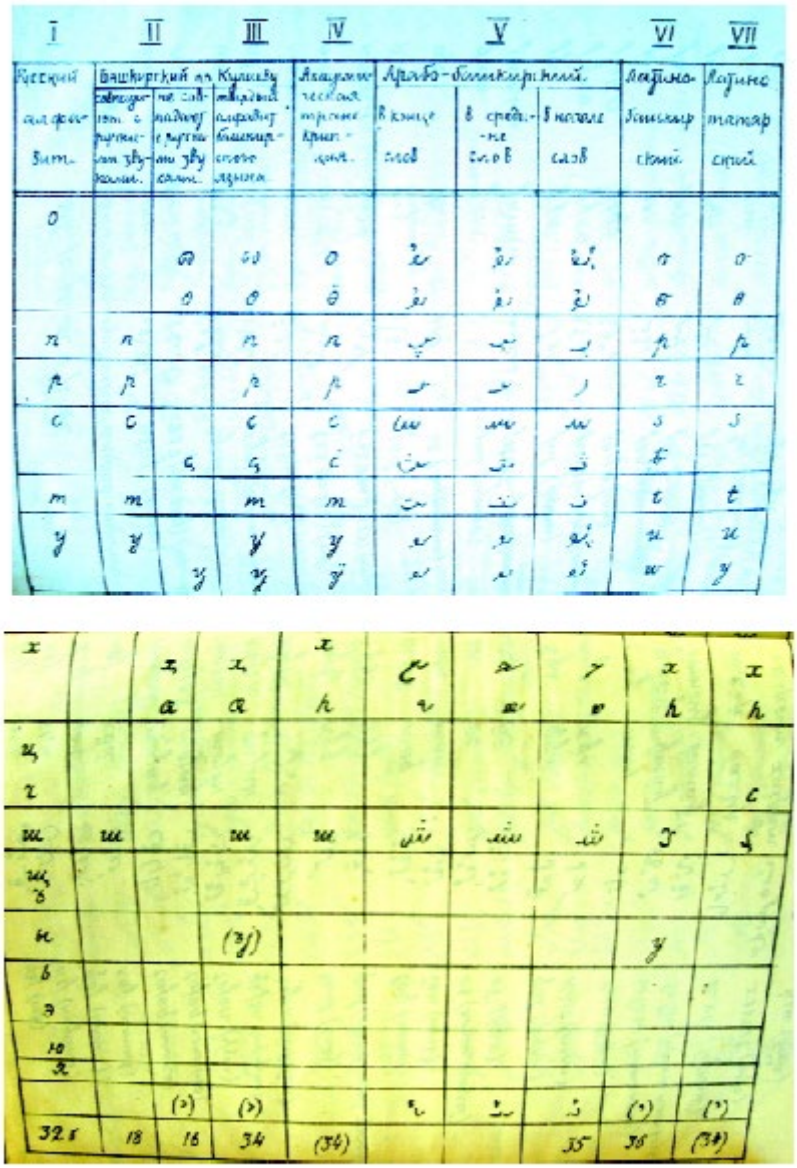
(from the book "Basics of Sound Pronunciation and the Alphabet for the Bashkirs")



I	II		III	IV	V			VI	VII
Русский алфавит	Башкирский по Кулеву			Лишенческая транскрипция	Арабо башкирский			Латино-Башкирский	Латино-татарский
	[illegible]	не совпадающие с русск. звук.	твердый алфавит башкирского яз.		в конце слов	в средине слов	в начале слов		
а	а		а	а	[illegible]	[illegible]	[illegible]	а	а
		ӕ	ӕ	ӓ	[illegible]	[illegible]	[illegible]	ӕ	ә
б	б		б	б	[illegible]	[illegible]	[illegible]	в	в
в					[illegible]	[illegible]	[illegible]	v	v
г	г		г		[illegible]	[illegible]	[illegible]	g	g
		ҙ	ҙ	ҕ	[illegible]	[illegible]	[illegible]	g	g
д	д		д	д	[illegible]	[illegible]	[illegible]	d	d
е									
		ҕ	ҕ	ҕ	[illegible]	[illegible]	[illegible]	ь	ь
		ь	ь	ь	[illegible]	[illegible]	[illegible]	e	e



ж	ж		ж	ж	[illegible]	[illegible]	[illegible]	ж	ж
з	з		з	з	[illegible]	[illegible]	[illegible]	z	z
		ҙ	ҙ	ҙ	[illegible]	[illegible]	[illegible]	ҙ	
и	и		и	i	[illegible]	[illegible]	[illegible]	i	i
ӣ		j	j	j	[illegible]	[illegible]	[illegible]	j	j
к	к		к	к	[illegible]	[illegible]	[illegible]	k	k
		ҡ	ҡ	q	[illegible]	[illegible]	[illegible]	q	q
л	л		л	л,l	[illegible]	[illegible]	[illegible]	l	l
м	м		м	м	[illegible]	[illegible]	[illegible]	m	m
н	н		н	н	[illegible]	[illegible]	[illegible]	n	n
		ҥ	ҥ	ҥ	[illegible]	[illegible]		ñ	ñ



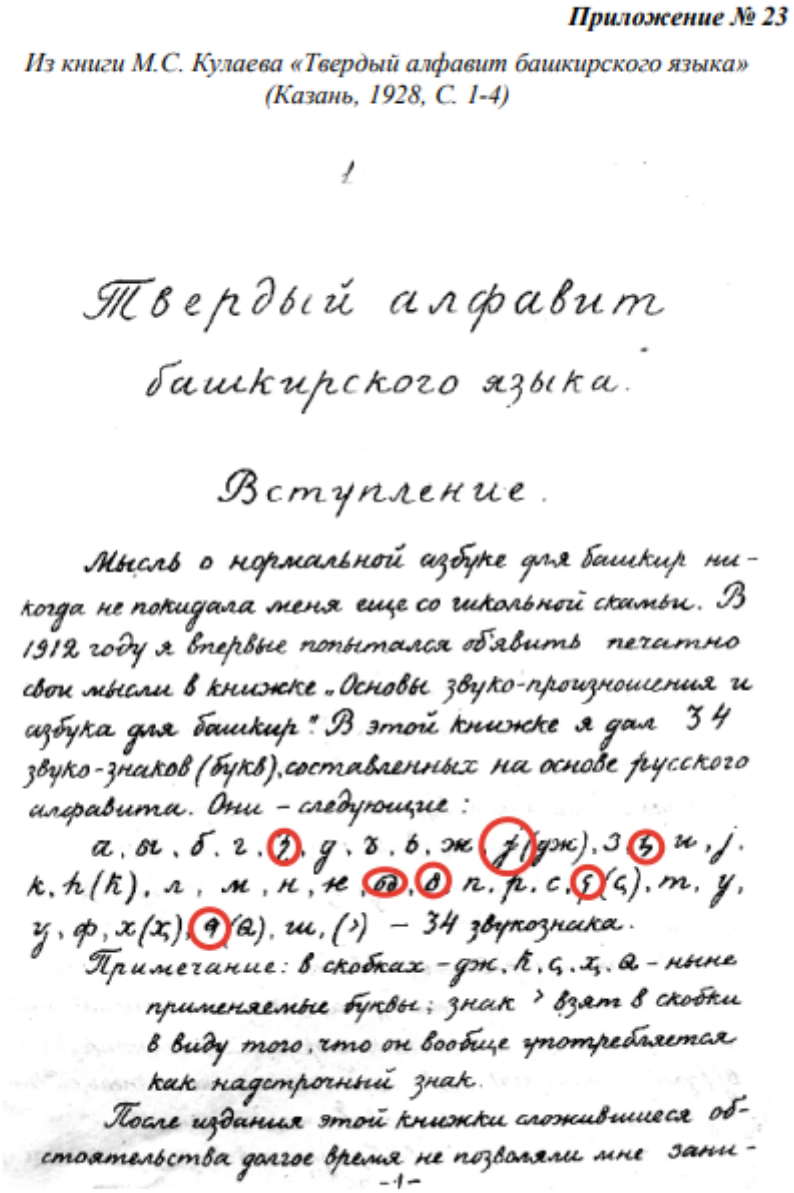
I	II		III	IV	V			VI	VII
Русский алфавит.	Башкирский по Куликову			Академическая транскрипция.	Арабо-башкирский.			Латино-башкирский.	Латино-татарский.
	совпадают с русскими звуками.	не совпадают с русскими звуками.	твердый алфавит башкирского языка.		в конце слов	в средине слов	в начале слов		
о									
		ѳ	ѳ	о	ـۇ	ـۇ	ـۇ	о	о
		ѳ	ѳ	ѳ̄	ـۇ	ـۇ	ـۇ	ѳ	ѳ
п	п		п	п	ـپ	ـپ	ـپ	р	р
р	р		р	р	ـر	ـر	ـر	г	г
с	с		с	с	ـس	ـس	ـس	s	s
		ç	ç	c̄	ـث	ـث	ـث	ѣ	
т	т		т	т	ـت	ـت	ـت	t	t
у	у		у	у	ـو	ـو	ـو	u	u
		ӱ	ӱ	ӱ	ـو	ـو	ـو	w	y

х		х	х	х	ـح	ـح	ـح	х	х
		а	а	h	ـه	ـه	ـه	h	h
ц									
ч									с
ш	ш		ш	ш	ـش	ـش	ـش	ʒ	ʃ
щ									
ъ									
ы			(ʒ)					у	
ь									
э									
ю									
я									
		(ʾ)	(ʾ)		ـٔ	ـٔ	ـٔ	(ʾ)	(ʾ)
32 б	18	16	34	(34)			35	36	(34)

Figure 6.

From the book by M.S. Kulaev "The Solid Alphabet of the Bashkir Language"

(Kazan, 1928, P. 1-4)



Приложение № 23

Из книги М.С. Кулаева «Твердый алфавит башкирского языка»
(Казань, 1928, С. 1-4)

1

Твердый алфавит
башкирского языка.

Вступление.

Мысль о нормальной азбуке для башкир никогда не покидала меня еще со школьной скамьи. В 1912 году я впервые попытался об'явить печатно свои мысли в книжке „Основы звуко-произношения и азбука для башкир". В этой книжке я дал 34 звуко-знаков (букв), составленных на основе русского алфавита. Они – следующие:

а, ы, б, г, ҡ, д, ъ, ь, ж, ҙ (дж), з, ҫ, и, ј, к, ћ (ћ), л, м, н, ҥ, ҭ, ү, п, р, с, ҫ (ҫ), т, у, ҷ, ф, х (х̣), ҩ (ҩ), ш, (›) – 34 звукознака.

Примечание: в скобках – дж, ћ, ҫ, х̣, ҩ – ныне применяемые буквы; знак › взят в скобки в виду того, что он вообще употребляется как надстрочный знак.

После издания этой книжки сложившиеся обстоятельства долгое время не позволяли мне зани-

-1-

113

6.References

1. Основы звукоподражания и азбука для башкир» 1912 (переиздана в 1919) Basics of Onomatopoeia and the Alphabet for the Bashkirs" 1912 (reprinted in 1919)

2. Хусаинова Л.М. Башкирское письмо монография. Стерлитамак Стерлитамакский филиал БашГУ, 2012 (L.M. Bashkir letter monograph. Sterlitamak Sterlitamak branch of BashSU, 2012)

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 18 Cyrillic characters for old Bashkir*
2. Requester's name: *Nikita Manulov <pacancuvak15@gmail.com>*
3. Requester type (Member body/Liaison/Individual contribution): *Individual contribution*
4. Submission date: *2022-06-25*
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: *YES*

B. Technical – General

1. Choose one of the following:
   - a. This proposal is for a new script (set of characters):
     - Proposed name of script:
   - b. The proposal is for addition of character(s) to an existing block: *YES*
     - Name of the existing block: *Cyrillic Extended-D*
2. Number of characters in proposal: *2*
3. Proposed category (select one from below - see section 2.2 of P&P document):

A-Contemporary	X	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? *Nikita Manulov*
   - b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.): *<pacancuvak15@gmail.com>*
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)?

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? *Examples*
   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? *NO*
   Reference:
4. The context of use for the proposed characters (type of use; common or rare) *YES*
   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, 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? *NO*
    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?
    If YES, are the equivalent corresponding unified ideographic characters identified? *NO*
    If YES, reference: