

Proposal to encode Latin Capital and Small letter I with Bowl

in ISO/IEC 10646 and The Unicode Standard

Nikita Manulov, pacancuvak15@gmail.com

2022-05-20

This proposal is to add the additional Latin letter I with Bowl of Soviet languages

1.Introduction

Ь ь (I with bowl) is an additional letter of the Latin alphabet. It was introduced in 1928 into the reformed Yañalif, and later into other alphabets for Soviet minority languages. The letter was designed specifically to represent the non-front close vowel sounds IPA: [ɨ] and IPA: [ɯ].[1] Thus, this letter corresponds to the letter ⟨I ı⟩ in modern Turkic alphabets

The letter was originally included in the Yañalif, and later also included in the alphabets of the Kurdish, Abazin, Sami, Komi, Tsakhur, Azerbaijani and Bashkir languages, as well as in the draft reform of the Udmurt alphabet. During the project of the Latinization of the Russian language, this letter corresponded to the Cyrillic letter ⟨Ы ы⟩.

In alphabets that used this letter, the lowercase B was replaced by a small capital ⟨Ь⟩ so that there would be no confusion between ⟨b⟩ U+0185 and ⟨Ь⟩ U+044C.

2.Examples of use

Figure. 1

New Turkic alphabet (Yañalif)

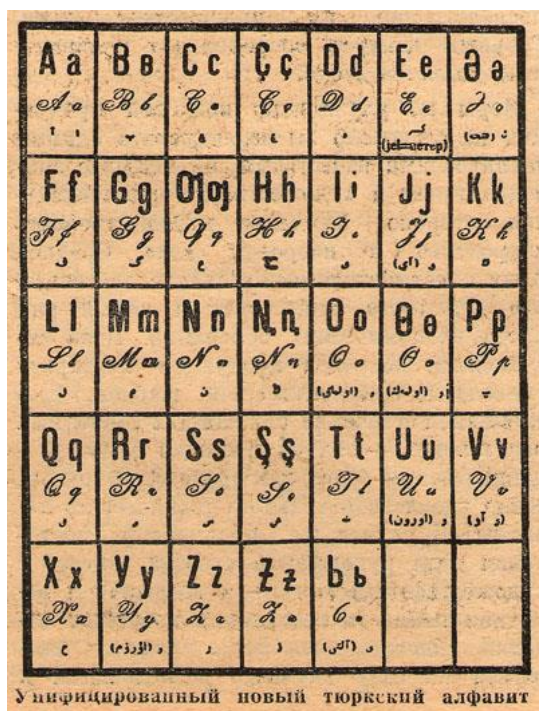


Figure. 2

The Latin-based Unified Northern Alphabet

Алфавит Народов Севера
Печатные формы букв

Aa	Bb	Cc	€€	Dd	Зз
Зз	Ee	Əə	Ææ	Ff	Gg
Hh	ḥḥ	Ii	Ьь	Jj	Kk
Ll	Łł	Mm	Nn	Ŋŋ	Oo
Өө	Pp	Qq	Rr	Rɾ	Ss
Œs	Tt	Uu	Vv	Ww	Xx
Yy	Zz	Ʒz			

Figure. 3

Kurdish alphabet of 1929

Әльфваје тәзә курманҗи
НОВЫЙ КУРДСКИЙ АЛФАВИТ

Aa <i>Aa</i> А	Bb <i>Bb</i> Б	Cc <i>Cc</i> Ч	€€ <i>€€</i> ТЧ сәрт (кожа)	Çç <i>Çç</i> ДЖ сег (кувшин)	Dd <i>Dd</i> Д
Ee <i>Ee</i> Е	Əə <i>Əə</i> МЯГКОЕ А әм (мы)	Ff <i>Ff</i> Ф	Gg <i>Gg</i> Г	Ŋŋ <i>Ŋŋ</i> Г МЯГКОЕ тәсвәр (ладья)	Hh <i>Hh</i> ХЫ нәсп (лошадь)
Ii <i>Ii</i> И	Ьь <i>Ьь</i> Ы ды (сердце)	Jj <i>Jj</i> Й	Kk <i>Kk</i> К	Ққ <i>Ққ</i> ТВЕРДОЕ К қер (нож)	Qq <i>Qq</i> КЫ қаз (гусь)
Ll <i>Ll</i> Л	Mm <i>Mm</i> М	Nn <i>Nn</i> Н	Oo <i>Oo</i> О	Өө <i>Өө</i> О МЯГКОЕ хоҫә (кусно)	Pp <i>Pp</i> П
Pp <i>Pp</i> Р рҫа (мост)	Rr <i>Rr</i> Р	Ss <i>Ss</i> С	Şş <i>Şş</i> Ш	Tt <i>Tt</i> Т	ƧƧ <i>ƧƧ</i> ТВЕРДОЕ тер (хурджин)
Uu <i>Uu</i> У	Yy <i>Yy</i> МЯГКОЕ У духьп (кушан)	Vv <i>Vv</i> В	Ww <i>Ww</i> W wörb (веревка)	Xx <i>Xx</i> Х	
Zz <i>Zz</i> З	Ʒz <i>Ʒz</i> Ж				

Figure 4.

Komi alphabet of 1934



Figure 5.

A Kazakh newspaper Socialist Kazakhstan (Социалды Қазағыстан) in Latin script (1937)



3.Characters to be added

Character	Code point	Name
-----------	------------	------

Ь	U+A7CC	LATIN CAPITAL LETTER I WITH BOWL
---	--------	----------------------------------

→ 042B cyrillic capital letter yeru

→ 042C cyrillic capital letter soft sign

→ 0184 latin capital letter tone six

ь	U+A7CD	LATIN SMALL LETTER I WITH BOWL
---	--------	--------------------------------

→ 0131 latin small letter dotless i

4.UnicodeData.txt

```
A7CC; LATIN CAPITAL LETTER I WITH BOWL; Lu;0; L;;;; N;;; A7CD;  
A7CD; LATIN SMALL LETTER I WITH BOWL; Ll;0; L;;;; N;;; A7CC;; A7CC
```

5.References

W. K. Matthews (2013). Languages of the USSR. Cambridge: Cambridge University Press. p. 71. ISBN 978-1-107-62355-2.

Tatar (tatarça / татарча / تاتارچا)

Azerbaijani (آذربایجانجا ديلي / Azərbaycan dili / Азәрбајҹан дили)

Karaim (къарай тили, Karay dili, לישון קדר)

Khakas (Хакас тили / Khakas tîli)

Kumuk (Къумукъ тил / Qumuq til)

Tuvan (Тыва дыл / Tuva dyl)