Unicode request for Cyrillic letter *jeru* with connecting bar

Kirk Miller, kirkmiller, gmail   2025 April 4

Ҍ

This proposal is for the letter ⟨ы⟩ used in Cyrillic phonetic notation.

There are a number of non-Unicode phonetic symbols found in Cyrillic texts. Many appear to be idiosyncratic, and in addition it can often be difficult to tell whether a given letter is Latin or Cyrillic in a phonetic tradition that mixes the two. ⟨ы⟩ is found across multiple sources and is clearly Cyrillic. In some sources a glyph with an overstruck bar is found (Figure 1), but in most cases it is the tops of the two elements of the letter that are connected; boths forms may even be found in the same text, with a top bar illustrated in explanatory notes and an overstruck bar used in the majority of the text.

⟨ы⟩ is typically used with the value of Russian ⟨ы⟩ in IPA-based phonetic notation (Menovschikov 1975: 16). The connecting bar was presumably added because the *yeru* ⟨ы⟩ could be confused for the sequence ⟨ы⟩ – that is, the symbol for palatalization, ⟨ь⟩, which is found in many of these texts, followed by small-capital IPA ⟨ɪ⟩, which is also used. However, in some traditions ⟨ы⟩ and ⟨ы⟩ contrast with one another (Figure 1, Figure 3). These traditions do not use ⟨ь⟩ for palatalization, so there is no risk of confusion with ⟨ы⟩.

⟨ы⟩ is distinct from *yeru* with a macron, ⟨ӹ⟩, which is a long [ɨː] sound found i.a. in Tatar.

A casing pair of ⟨Ы ы⟩ with a cursive-style connecting stroke has been used for Wakhi [ISO wbl] in Tajikistan (Khan 2006). However, this does not appear to be general usage, and because there is no contrast, standard ⟨Ы ы⟩ could presumably be substituted. We therefore only request the lower case.

Characters

Ҍ   1C8B CYRILLIC SMALL LETTER YERU WITH CONNECTING BAR.

Properties

1C8B;CYRILLIC SMALL LETTER YERU WITH CONNECTING BAR;Ll;0;L;;;;;N;;;;;

References

Bičeldej (2001): К.А. Бичелдей, *Звуковой строй диалектов тувинского языка* (*Sound structure of the dialects of the Tuvan language*). Moscow.

Bondarko et al. (2004): Л. В. Бондарко, Л. А. Вербицкая & М. В. Гордина, *Основы общей фонетики* (*Basics of general phonetics*). 4th edition, Akademia Press, St. Petersburg.

Zulfiqar Ali Khan (2006) *Xikwor Rasmulxat.*
http://www.gojal.net/languages_wakhi_%20alphabets.htm

V.A. Kočergina (1991) *Введение в языковедение. Основы фонетики-фонологии. Грамматика: Учебное пособие.* University of Moscow Press.

Krejnovič (1961): Е.А. Крейнович, *О проекте единой фонетической транскрипции для палеоазиатских языков.* USSR Academy of Sciences, Institute of Linguistics.s

Menovschikov (1975): Г.А. Меновщиков, *Язык науканских эскимосов* (*Naukan Eskimo language*). Nauka Press.

Popova (1978): Я.Н. Попова, *Фонетические особенности лесного наречия ненецкого языка.* Nauka Press.

Chart

Cyrillic Extended-C

1C80   1C8F

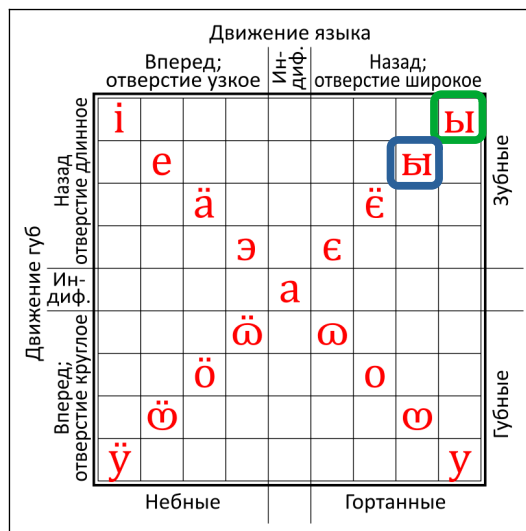|	1C8
0	ℓ
1	д
2	о
3	с
4	ꙟ
5	ɯ
6	ъ
7	ҍ
8	ꙑ
9	Ҍ
A	Ѣ
B	Ы
C	
D	
E	
F	

Figures

Figure 1. ⟨ы⟩ with connecting bar (blue) contrasts with ⟨ы⟩ without (green). The bar in ⟨ы⟩ crosses below the top of the letter, but it is more common for it to be at the top. At left are front vowels, at right back vowels; at top are spread vowels and at bottom rounded vowels. That is, ⟨y⟩ has its Cyrillic value [u].
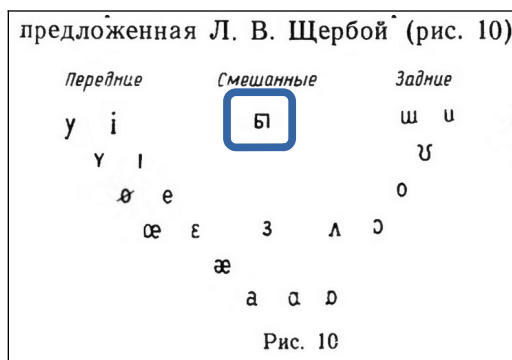

Figure 2. Kočergina (1991: 33). Here ⟨ы⟩ is used as the equivalent of Russian ⟨ы⟩ in IPA-based phonetic transcription. The connecting bar is at the top of the letter. Latin ⟨y⟩ is used with its IPA value [y].

Figure 3. Krejnovič (1961: tables 8 and 9). Cyrillic transcription on the left and the corresponding Latin transcription on the right. Both are mixed script. As in Figure 1, ⟨ы⟩ (blue) contrasts with ⟨ы⟩ (dark green), though in this case ⟨ы⟩ is the central vowel [ɨ] while ⟨ы⟩ is the back vowel [ɯ]. Also appearing are the Bulgarian vowel letter ⟨ъ⟩ for [ɤ] and the Azeri letter ⟨ə⟩ for [ä] (light green).

Figure 4. Krejnovič (1961: 13 and table 3). ⟨ы⟩ (blue) as a high unrounded back vowel [ɯ], vs ⟨ы⟩ (green) as a near-high central vowel [ɨ], in the transcription of Ket.

Figure 5. Bondarko et al. (2004: 27). Vowel chart in an introductory textbook on phonetics. The notation is IPA, but with ⟨ы⟩ substituted for IPA ⟨ɨ⟩. (The ⟨ʒ⟩ is a typo, perhaps for the Bulgarian vowel letter ⟨ъ⟩ seen in Figure 6; the IPA does not have a distinct letter for this sound.)

Figure 6. Bondarko et al. (2004: 142). A typewritten hack with an overstruck bar for the better-typeset symbol in the chart in Figure 6. Here transcriptional ⟨ы⟩ (blue) contrasts with orthographic ⟨ы⟩ (green). Note also the Bulgarian vowel letter ⟨ъ⟩.

Подъем	Р я д		
	передний	смешанный	задний
Высокий	i	ы	u
Средний	(e)		(o) –
Низкий	(ɛ̆)	a ⟶	

Figure 7. Menovschikov (1975: 17). The vowels of Naukan Yupik, with ⟨ы⟩ substituted for IPA ⟨ɨ⟩.

Figure 8. Menovschikov (1975: 20). A comparison of Naukan and Chaplino (top). The letter was typed with the ⟨ы⟩ that was available on the typewriter, then manually corrected to ⟨ы⟩ (blue). The surrounding Russian text uses plain ⟨ы⟩ (green).

Figure 9. Menovschikov (1975: 28). ⟨ы̂⟩ with a circumflex diacritic. The odd gamma-like letter is just the handwritten form of IPA ⟨ɣ⟩: e.g. [ы̂naɣaq] in line 3.

3-е, ед.	ыглыˌӷутāӄāӈа
мн.	ыглыˌӷутāӄāтӈа
ед.	ыглыˌӷутāӄāтын

Figure 10. Menovschikov (1975: 241). In Naukan orthography, plain ⟨ы⟩ is used, in contrast to the transcriptions in Figure 8 and Figure 9. In this table of inflections, the root is *ыглыˌӷутāӄ-*.

Figure 11. Bičeldej (2001: 23-24). The vowel /ы/ and its nasalized counterpart /ы̃ː/ in Tuvan phonology (blue). Note that a normal *jeru* ⟨ы⟩ (green) is used in the Russian description.

Figure 12. Bičeldej (2001: 40). The word 'year' is transcribed [ħʃˑ ы l], in a mixture of Latin and Cyrillic script; in Tuvan orthography, it's *чыл* with a normal ⟨ы⟩. ⟨ħ⟩ *tje* is the Serbian Cyrillic letter for [t͡ɕ], tied here to the IPA Latin letter ⟨ʃ⟩ *esh*.

в) по месту образования краткие и долгие гласные тувин-
ского языка целесообразно противопоставлять как твердорядные **a,
ы, ɔ, ʊ, a:,** ы: ɔ:, ʊ: и мягкорядные **ɛ, i, ɵ, ɣ,** ɛ:, i:, ɵ:, ɣ:; термины 'пе-
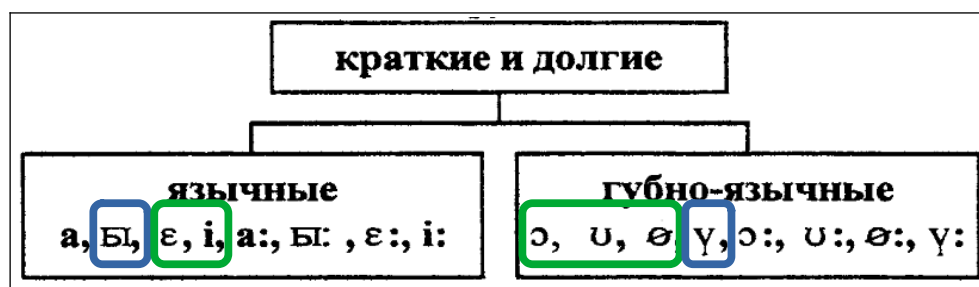
краткие и долгие	
язычные **a, ы, ɛ, i, a:,** ы: , ɛ:, i:	**губно-язычные** ɔ, ʊ, ɵ, ɣ, ɔ:, ʊ:, ɵ:, ɣ:

Figure 13. Bičeldej (2001: 44, 67). A mixture of IPA and Cyrillic script, with IPA ⟨ɔ ʊ ɛ i ɵ⟩ (green) augmented by Cyrillic ⟨ы ɣ⟩ (blue) for IPA ⟨ɨ y⟩. (IPA ⟨y⟩ could be confused with the Cyrillic letter for [u].) Bottom: harmonic sets of unrounded vowels (left) and the corresponding rounded vowels (right). /ы/ and /ʊ/ form a pair.

2) узкие огубленные краткие и долгие гласные употребляются
в любой позиции слова, например: ʊq 'носки', ʊqʊ '(*его*) носки',
ʊqʊm '(*мои*) носки', ы̆t 'собака', ы̆dы '(*его*) собака', ы̆dы m (*моя*)
собака' и т.д., а долгие назализованные и фарингализованные - со-
гласно их, дистрибутивной характеристике.

19] гласных в них, например: **at** 'имя' - ы**t** 'отправь', ɦʃ'**a:r** 'мазать' -
ɦʃ'ы:**r** 'собирать', **ã:t** 'качай' - ы̃:**t** 'шум, голос'; **ăt** 'лошадь' - ы̆**t**
'собака', ɦʃ'ɔrʊ 'иди' - ɦʃ'ʊrʊ 'рисуй', **xɵl** 'озеро' - **xɣl** 'зола', ɛ:ʃ'
'товарищ' - iʃ' 'дело, работа', **qăʃ'** 'сколько' - **qы̆ʃ'** 'зима' и т.д. В

Figure 14. Bičeldej (2001: 69, 72). The transcription is largely Latin or IPA, including the old Soviet Latin letter ⟨ɋ⟩ *gha* (red). The letter ⟨ы⟩ occurs long, short and nasalized, e.g. ⟨ы̃:⟩.

Figure 15. Popova (1978: 165). ⟨ы⟩ in a phonemic transcription of Nenets.

Figure 16. Popova (1978: 106). Several words with the vowel [ы] in Nenets dialects.

Figure 17. Popova (1978: 105). Dialect variation contrasting [ы] (blue) with [ɪ] and [ы] (green).

Figure 18v. Popova (1978: 130 [table 2]). ⟨ы⟩ contrasting with its graphic inverse, which I suspect is the Bulgarian vowel letter ⟨ъ⟩.

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 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 std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html.
See also std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html for latest *Roadmaps*.

A. Administrative

1. **Title:** *Cyrillic letter ЬІ with connecting bar*

2. Requester's name: *Kirk Miller*

3. Requester type (Member body/Liaison/Individual contribution): *individual*

4. Submission date: *2025 April 04*

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:

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: *yes*
        Name of the existing block: *Cyrillic Extended-C*

2. Number of characters in proposal: *1*

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) x
    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?
        Kirk Miller
    b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
        SIL open font license

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)? *no*

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 www.unicode.org for such information on other scripts. Also see Unicode Character Database (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.)? _____ *no* _____
   If YES, with whom? _____
   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: _____

4. The context of use for the proposed characters (type of use; common or rare) _ *phonetic transcription* _
   Reference: _____

5. Are the proposed characters in current use by the user community? _____ *yes* _____
   If YES, where? Reference: _____ *See figures* _____

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)? _ *NA* _

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 se- _____ *no* _____
    quences?
    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 characters? _____ *no* _____
    If YES, are the equivalent corresponding unified ideographic characters identified? _____
    If YES, reference: _____