

Unicode request for South Slavic Cyrillic modifier letters

Kirk Miller, kirkmiller, gmail.com

2026 September 02

We request 6 additional characters for Cyrillic Extended-D to support South Slavic.

Proposals [L2/21-107](#), [L2/21-142](#) and [L2/22-010](#) requested modifier letters that became the Cyrillic Extended-D block. There are gaps in coverage, particularly for South Slavic languages, due to gaps in attestation in those proposals:

Gaps in super- and sub-script letters of the national Slavic Cyrillic alphabets. (Requested characters in red.)

А	Б	В	Г	Г	Д	Ђ	Е	Є	Ж	З	Ѕ	И	І	Ј	К	Л	Љ	М	Н	Њ	О	П	Р	С	Т	Ћ	У	Ф	Х	Ц	Ч	Џ	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я
а	б	в	г	г	д	ђ	е	е	ж	з	з	и	і	ј	к	л	љ	м	н	њ	о	п	р	с	т	ћ	у	ф	х	ц	ч	џ	ш	ш	ъ	ы	ь	э	ю	я
а	б	в	г	г	д	—	е	—	ж	з	ѕ	и	і	—	к	л	—	—	—	—	о	п	—	с	—	—	у	ф	х	ц	ч	џ	ш	—	ъ	ы	ь	—	—	—

Here we show that the South Slavic letters are attested for more than a century. Modern usage includes both new text and reproduction of older material (e.g. Figure 9).

Thanks to Uroš Stefanović for providing several of the figures.

Characters

Љ 1E06E MODIFIER LETTER CYRILLIC SMALL LJE.
 Њ 1E06F MODIFIER LETTER CYRILLIC SMALL NJE.
 Ѓ 1E070 MODIFIER LETTER CYRILLIC SMALL DJE.
 Ћ 1E071 MODIFIER LETTER CYRILLIC SMALL TSHE.
 Џ 1E072 MODIFIER LETTER CYRILLIC SMALL DZHE.
 Ъ 1E073 CYRILLIC SUBSCRIPT SMALL LETTER SOFT SIGN.

Properties

1E06E;MODIFIER LETTER CYRILLIC SMALL LJE;Lm;0;L;<super> 0459;;;N;;;;;
 1E06F;MODIFIER LETTER CYRILLIC SMALL NJE;Lm;0;L;<super> 045A;;;N;;;;;
 1E070;MODIFIER LETTER CYRILLIC SMALL DJE;Lm;0;L;<super> 0452;;;N;;;;;
 1E071;MODIFIER LETTER CYRILLIC SMALL TSHE;Lm;0;L;<super> 045B;;;N;;;;;
 1E072;MODIFIER LETTER CYRILLIC SMALL DZHE;Lm;0;L;<super> 045F;;;N;;;;;
 1E073;CYRILLIC SUBSCRIPT SMALL LETTER SOFT SIGN;Lm;0;L;<sub> 044C;;;N;;;;;

Charts

1E030 Cyrillic Extended-D 1E08F

	1E03	1E04	1E05	1E06	1E07	1E08
0	а	т	l	ф	ђ	
1	б	у	а	х	ћ	
2	в	ф	б	ц	џ	
3	г	х	в	ч	ь	
4	д	ц	г	ш		
5	е	ч	д	ъ		
6	ж	ш	е	ы		
7	з	ы	ж	г		
8	и	э	з	і		
9	к	ю	и	ѕ		
A	л	ђ	к	џ		
B	м	э	л	џ		
C	о	і	о	ы		
D	п	ј	п	џ		
E	р	е	с	љ		
F	с	џ	у	њ		і̇

References

- Žarko Bošnjaković (2021) Funkcije i značenja nominativa u govoru sela Binač i Klokot u Gornjoj Moravi. *Zbornik Matice Srpske* LXIV/1. Novi Sad. <https://www.maticasrpska.org>.
- Pavle Ivić (1986) Govor ličkih jekavaca. *Srpski dijalektološki zbornik* 32. Belgrade: Srpska akademija nauka i umetnosti. <https://dais.sanu.ac.rs/>.
- Aleksandar Belić (1913) Review of V. Bogorodickiy [1912] 'Dialektologičeskija zamětki IX.' *Južnoslovenski filolog* 1.1-2. Belgrade. <https://dais.sanu.ac>.
- Dragomir Kozomara (2022). Fonetske osobine govora Bastasa. *Srpski dijalektološki zbornik* 69. Belgrade: Srpska akademija nauka i umetnosti. <https://dais.sanu.ac.rs/>.
- Tanja Milosavljević & Radivoje Mladenović (2025) Rečnik prizrenskog govora. *Srpski dijalektološki zbornik* 72. Belgrade: Srpska akademija nauka i umetnosti.
- Radivoje Mladenović (2023) Govor sela Vrbičane kraj Prizrena. *Srpski dijalektološki zbornik* 70/1. Belgrade: Srpska akademija nauka i umetnosti.
- Milivoj Pavlović (1939) Govor sretečke župe. *Srpski dijalektološki zbornik* 8. Belgrade: Srpska kraljevska akademija. <https://dais.sanu.ac.rs/>.
- Trajković (2024) Govori Kosovskog Pomoravlja. *Srpski dijalektološki zbornik* 71. Belgrade: Srpska akademija nauka i umetnosti.

Figures

од оба гласа, врло велика. При изговору рускога љ и њ имамо у првом делу тих непчаника затворено (са високим властитим тоном) л и н, а тек у последњем делу праве умекшане сугласнике, дакле неко **л^а, н^а**. На српско ухо руски гласови чине утисак полумеких сугласника, док су у српском језику ти сугласници од почетка до краја меки, дакле **љ^а, њ^а**. Ја сам поредио и артикулације тих гласова у руском и српском и нашао сам да оне потпуно одговарају акустичком утиску. Руско се љ и њ отпо-

Figure 1. Belić (1913: 129). Modifier *љ* and *њ* indicating palatalization of both alveolars and palatals. The text is reviewing a publication from the previous year.

Различит је инвентар семикопулативних глагола уз које се, у функцији допуне, појављује номинатив (он ради́о ко суди́ја у Г^ни́лане Кл. Ё, м^н и́рске ви́камо. Ви́ и́рснице, м^н и́рске Би).

Figure 2. Bošnjaković (2021: 91). Modifier ɸ indicating the local pronunciation of the city of Gnjilane, Kosovo, which sounds more like “Njilane.”

(Пл.), *пóвноћу* (Др., Ср.), *у пóвноћи*, *до пóвноћи* (Г. С., Ср., Др., Н., Ж., Л.), *сúноћ'ке* (Ср), *сúнојк'а* (Др.) **нош^ћно** *време*, *нош'но* (Ср., Г. С.), *занóћуваш* (М.), *плáћа* (Р., Ср.), *плеће* — *ја у*

Figure 3. Pavlović (1939: 28). Modifier *h* indicating epenthesis.

кућа у градњи), *зграђа* (Др.), *зграђ'е* (Ср., Г. С.), *заграђа* (Р.), *заграђено* (Р.) он *доцађа* (Ср.), *крађа* (Ср.), *међ'а* (Р.), *међа* (Др.), *међе* (Ср., Г. С.), *пређа* (Ср., Г. С., Др.), *рађа*, *оне ра-*

Figure 4. Pavlović (1939: 29). Modifier *h* indicating reduced articulation.

„У току се времена језик јужно-моравске долине изменио у том правцу што је *l* изменио у *lʷ* и даље у *lu*¹). То развијање ла-

Figure 5. Pavlović (1939: 78). Subscript soft sign ь. As in the Russian tradition, subscript and superscript phonetic symbols contrast.

Као и у говору ијекаваца Баније и Кордуна,¹²⁷ и овдје разнолики ликови облика (-)iод доприносе чешћем јављању африката *ђ* и *ћ*. Исп.:
ђе iđђ било Би, *који iđђ* Би Ке, *оклен iđђ* Ке, *ђе iđђ^h Вр, кад iđђ^h Ке,*
ђе iđђ Вр ЗК Ке, чим је iđђ мѡiа Бе, шио iđђ Ви, шиоiоћ Би;

Figure 6. Ivić (1986: 100). Modifier *ђ* indicating a lack of word-final devoicing.

в) Облик императива: *наиди* Бе, забиљежио сам само једном. Обично је: *наиђи* Бе Вр Ке Ми ПС, *наиђ^hи* Га,

Figure 7. Ivić (1986: 168). Modifier *ћ* indicating anticipatory devoicing.

Сэс пѣпоў турал’е јѣно врѣме.
Цѣђ^h прáјл’е. Отсѣјоў пѣпоу жѣне сэс рѣщѣто, сэ[c] сýто, сэ[c] што имáл’е. Отсѣјоў па тýриў на верíгу кóтоў. Ё́, кэд врíје вóда, óни тýриў онáј пѣпоў по двѣ, трí шáке дóбре. Нал’ѣгне се јýтре дѣн. Тóј турал’е.
 И л’ýди мíл’е и кóсе. Јѣз де. Па немáло. А сэге нѣ, сáг је дрýго. Сэге *цѣђ^h* нѣ.

Figure 8. Mladenović (2019: 1033). The schwa, assuming Cyrillic is desired rather than Latin, is already encoded for Kazakh.

И дòшла jâ зá њга. Дà, ко *гòђ^h* дòђѣ, кàжѣ: „Цýјо, бòк т-ўбијò, вèдйш да ће Стéво ўмрати.“ Ако ўмрѣ, реко, нíсам ш њíм лѣжàла, нíсам нйшта, jâ сам цýра, нйшта jâ òде свòйје прàвцом.

Figure 9. Kozomara (2022: 82). Modifier *ђ* indicating anticipatory voicing. Digitization of a text from 1923.

док се óпере. Ѐ у зíма, ѝ у л'это. Нé-било лóко за жéне. Бóгме. Па нéмало ни прáшкови ка сáг. Ос **цéћ^h** Тóј от пéпел. Тóури у кóтал и грéј, па врí да се óмијев ствáри. А íмало мáло и сáпун, вáј сáпун што смо прáјл'е од мрс, от

Figure 10. Mladenović (2023: 341).

испóсни* свр оїере од масноће, одсїрани масноћу. – Изнése судóве у двóр и свé ги исїóсни. – Нáпре исїóсни^ф судóве сьс **цéћ^h** – Свáке годíне

Figure 11. Milosavljević & Mladenović (2025: 352).

1) У више примјера потврђена је општепозната замјена $m > ч$ у: *јèднōч*, *јèднōч* ѝшō, ***јèднōч^u*** би мòгло, по *јèднōч* ћемо, ако *јèднōч* ѝ није, *нијèднōч* нíсам бíла, *одјèднōч* скòчй, мòре стàти *одјèднōч*.

Figure 12. Kozomara (2022: 56). Modifier μ for anticipatory voicing.

йма чкóлу Кр; јá сам ѝшла у Билáчу Д; на тòј вéнче Д; влáчимо у влaчáру Д; Светè врáче Д; тэг чéкаш **гòч^u** Д; фрљимо на кўйче Д; како плóча Д; пўшка йўче Д; нéко шéлче Д; у чарáйе Д; чíс лéб нeмáло Д; па не пратíли у чкóлу Д;

Figure 13. Trajković (2024: 81). Modifier μ .

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹.

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:	<i>South Slavic Cyrillic modifiers</i>
2. Requester's name:	<i>Kirk Miller</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>individual</i>
4. Submission date:	<i>2026 September 02</i>
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	<i>yes</i>
(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):			
Proposed name of script:			
b. The proposal is for addition of character(s) to an existing block:			
Name of the existing block: <i>Cyrillic Extended-D</i>			
2. Number of characters in proposal: <i>6</i>			
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? <i>yes</i>			
a. If YES, are the names in accordance with the “character naming guidelines” in Annex L of P&P document? <i>yes</i>			
b. Are the character shapes attached in a legible form suitable for review? <i>yes</i>			
5. Fonts related:			
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard? <i>Kirk Miller</i>			
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.): <i>SIL (Gentium Release)</i>			
6. References:			
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? <i>yes</i>			
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? <i>yes</i>			
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)? <i>yes</i>			

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.)?	yes
If YES, with whom? Uroš Stefanović.	
If YES, available relevant documents: emails on request	
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?	
Reference:	
4. The context of use for the proposed characters (type of use; common or rare)	phonetic
Reference:	
5. Are the proposed characters in current use by the user community?	yes
If YES, where? Reference: see illustrations	
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)?	no
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?	no
If YES, are the equivalent corresponding unified ideographic characters identified?	
If YES, reference:	

UnicodeData.txt

```
1 1E06E;MODIFIER LETTER CYRILLIC SMALL LJE;Lm;0;L;<super> 0459;;;N;;;;  
2 1E06F;MODIFIER LETTER CYRILLIC SMALL NJE;Lm;0;L;<super> 045A;;;N;;;;  
3 1E070;MODIFIER LETTER CYRILLIC SMALL DJE;Lm;0;L;<super> 0452;;;N;;;;  
4 1E071;MODIFIER LETTER CYRILLIC SMALL TSHE;Lm;0;L;<super> 045B;;;N;;;;  
5 1E072;MODIFIER LETTER CYRILLIC SMALL DZHE;Lm;0;L;<super> 045F;;;N;;;;  
6 1E073;CYRILLIC SUBSCRIPT SMALL LETTER SOFT SIGN;Lm;0;L;<sub> 044C;;;N;;;;
```

+6