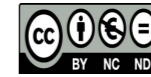


Sebastian Kempgen, Univ. of Bamberg, Germany



Unicode (v.13/14) and Slavic Philology Status Quo, Updates, Open Problems



Unicode and Slavic Philology

- **Previous papers & presentations:**

Sofia 2005: Unicode 4.1. [Paper](#) • [Presentation](#)

Ohrid 2008: Unicode 5.1. [Paper](#) • [Presentation](#)

Ohrid 2008: [Character Set Standardization for Early Cyrillic Writing after Unicode 5.1 \(D. Birnbaum, R. Cleminson, S. Kempgen, K. Ribarov\)](#) > **contra “Belgrade Proposal”**

Belgrade 2018: planned, but unable to attend (Unicode 11)



Unicode and Slavic Philology

Today:

1) Update on Latin Slavic Unicode

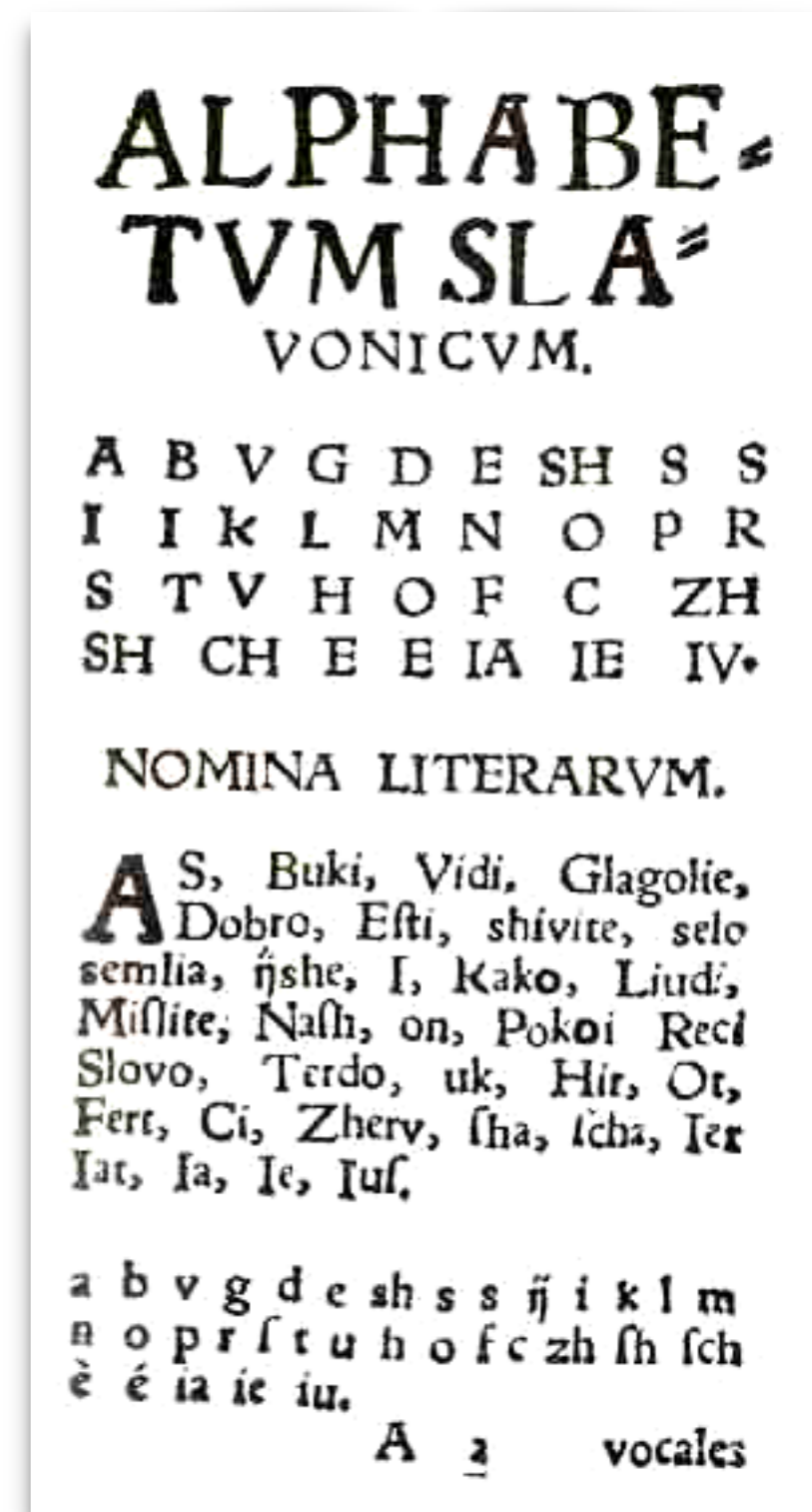
2) Cyrillic: Status Quo and Open Problems

3) Glagolitic: Status Quo and Open Problems

Unicode:

2020: Unicode v. 13;

2021: Unicode v. 14 (Sept.)



1. Update on Latin Unicode

Not realised for various reasons (examples)


Štokavian Accents

á é í ó ú ř • Latin 1
à è ì ò ù _ • Latin 1 ✗
â ê î ô û ř • Extended B
à è ì ò ù ã • Extended B ✗
ĩ missing from UC!

Long/Short Vowels

ā ē ī ō ū ŷ – long vowels ✓
ǎ ě ě ō ŷ _ – short vowels ✓
ÿ – not precomposed in Latin, available in Cyrillic (for Byelorussian) ✗

Russian Phonetics

1		2		3
nach API	nach Avanesov	nach API	nach Avanesov	
[i]	[и ^е]	/e/ /a/ /o/	/e/ /a/ /o/	American usage  1D7C
[i]	[ь]	/e/ /o/	/e/ /o/	
dotless i, not iota!				
[i] ←	[ы]	/i/	/и/	
[i] ←	[ы ^е]	/e/	/e/	
[Λ]	[Λ]	/a/	/a/	
		/o/	/o/	
		/e/	/e/	
		/a/	/a/	
		/o/	/o/	

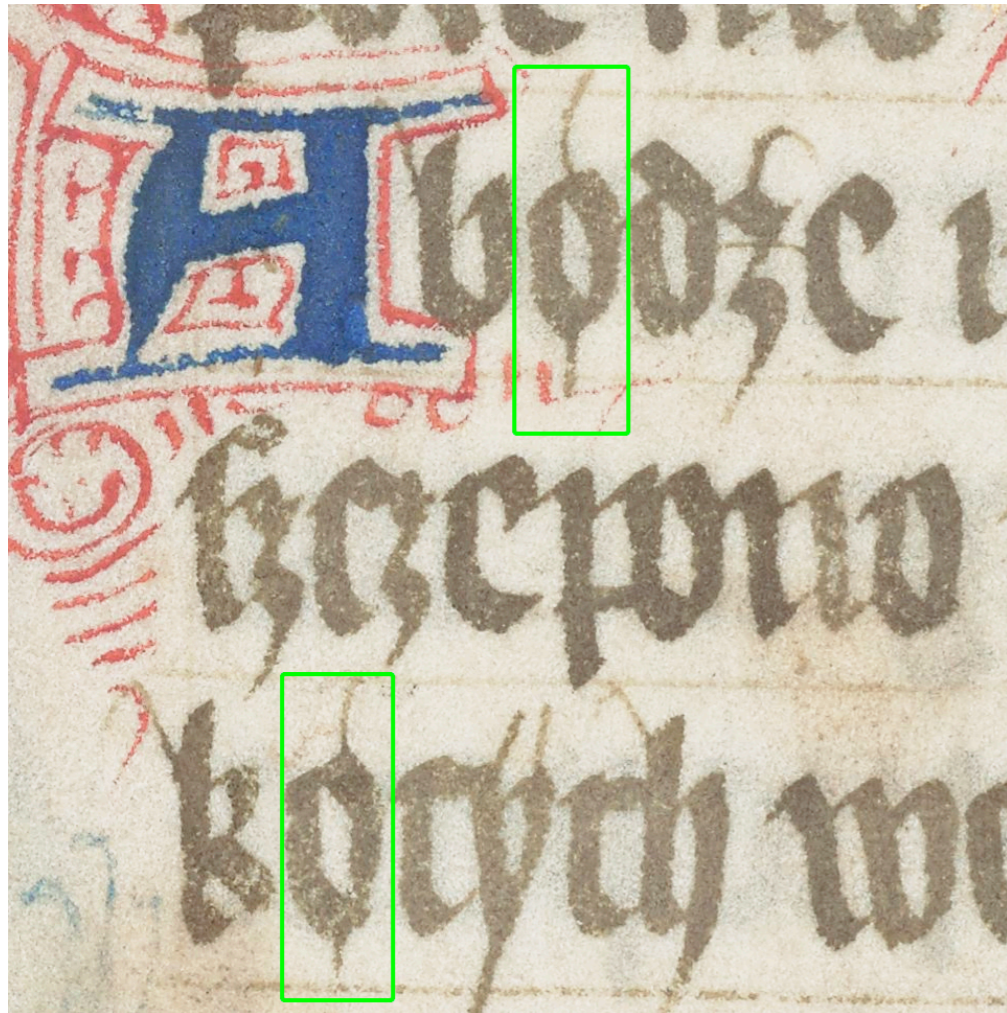
вЫдан, кнИга, пуля
на год, лѐтом, мѐрем

K. Gabka (ed.), Einführung in das Studium der russischen Sprache. Phonetik und Phonologie. Düsseldorf 1975, 34 & 164.

1. Update on Latin Unicode

Latin alphabets: Old Polish nasal o added

Proposal: Daniel Bunčić



Sankt Florian psalter

lniej przemawia za tym przypuszczeniem zwyczaj ozna-
ek wspólnym znakiem o, który się zaczyna u schyłku
powszechnia w XIV w. Oto wybrane z psalmu 56 *Psalte-*
wraży z nosówką: nademno (= nade mną), czó (= cię), bodo
= potęgę), depczócze (= depcące), prawdo (= prawdę),
svoio (= swoje), duszó (= duszę), moio (= moje), szczenót (= szczeniąt),
zamóczon (= zamacon), zoby (= zęby), yozik (= język), zemó (= ziemię),
iamó (= jamę), w no (= w nią), gózly (= gęśli), wstanó (= wstanie), rzekó
(= rzekę). Także w *Kazaniach świętokrzyskich* spotykamy jeden wspólny
znak dla obu odziedziczonych nosówek, np. ydeho (= idziechą), so (= się),
namodlitó (= na modlitwę), poczpho (= poczęchą), stoppiu (= stąpiw), vza-
stopy (= w zastępy), poc (= pięć), pote (= piąte). Później, w XV w.,

Z. Klemensiewicz, T. Lehr-Splawiński, S. Urbańczyk: Gramatyka historyczna języka polskiego. Warszawa 1965, 103

1. Update on Latin Unicode

Latin alphabets: Old Polish nasal o added



Font: Breitkopf Fraktur
by Sebastian Kempgen

emawia za tym przypuszczeniem zwyczaj ozna-
lnym znakiem o, który się zaczyna u schyłku
nia w XIV w. Oto wybrane z psalmu 56 *Psalte-*
rza florińskiego wyrazy z nosówką: nademno (= nade mną), czó (= cię), bodo
(= będę), potopo (= potęgę), depczocze (= depcące), prawdo (= prawdę),
svoio (= swoje), duszó (= duszę), moio (= moje), szczenót (= szczeniąt),
zamoczón (= zamacon), zoby (= zęby), yozik (= język), zemó (= ziemię),
iamó (= jamę), w no (= w nią), gózly (= gęśli), wstanó (= wstanę), rzekó
(= rzekę). Także w *Kazaniach świętokrzyskich* spotykamy jeden wspólny
znak dla obu odziedziczonych nosówek, np. ydeho (= idziechą), so (= się),
namodlitó (= na modlitwę), poczohó (= poczęchą), stoppiu (= stąpiw), vza-
stopy (= w zastępy), poc (= pięć), pote (= piąte). Później, w XV w.,

Z. Klemensiewicz, T. Lehr-Splawiński, S. Urbańczyk: Gramatyka historyczna języka polskiego. Warszawa 1965, 103

2. Cyrillic Unicode: Status Quo

0400	0401	0402	0403	0404	0405	0406	0407	0408	0409	040A	040B	040C	040D	040E	040F												
È	Ë	Ђ	Ѓ	Є	Ѕ	І	Ї	Ј	Љ	Њ	Ћ	Ќ	Ў	Ў	Џ												
0410	0411	0412	0413	0414	0415	0416	0417	0418	0419	041A	041B	041C	041D	041E	041F												
А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П												
0420	0421	0422	0423	0424	0425	A640	A641	A642	A643	A644	A645	A646	A647	A648	A649	A64A	A64B	A64C	A64D	A64E	A64F						
Р	С	Т	У	Ф	Х	З	з	Ѕ	ѕ	Љ	љ	Њ	њ	Ѓ	ѓ	Ѕ	ѕ	Ѓ	ѓ	Ђ	ђ						
0430	0431	0432	0433	0434	0435	A650	A651	A652	A653	A654	A655	A656	A657	A658	A659	A65A	A65B	A65C	A65D	A65E	A65F						
а	б	в	г	д	е	Ы	ы	Ї	й	О	о	Ў	у	А	а	Ж	ж	Ў	у	Ѓ	ѓ						
0440	0441	0442	0443	0444	0445	A660	A661	A662	A663	A664	A665	A666	A667	A668	A669	A66A	A66B	A66C	A66D	A66E	A66F						
р	с	т	у	ф	х	Ц	ц	Д	д	Л	л	М	м	О	о	О	о	О	о	О	о	О	о	О	о	О	о
0450	0451	0452	0453	0454	0455	A670	A671	A672	A673	A674	A675	A676	A677	A678	A679	A67A	A67B	A67C	A67D	A67E	A67F						
è	ë	ђ	ѓ	є	ѕ	⌘	⌘	⌘	⌘	⌘	⌘	⌘	⌘	⌘	⌘	⌘	⌘	⌘	⌘	⌘	⌘	⌘	⌘	⌘	⌘	⌘	
0460	0461	0462	0463	0464	0465	0466	0467	0468	0469	046A	046B	046C	046D	046E	046F												
Ѓ	ѓ	Ђ	ђ	Є	Ѕ	А	а	Ў	у	Ж	ж	Ў	у	Ѓ	ѓ												
0470	0471	0472	0473	0474	0475	0476	0477	0478	0479	047A	047B	047C	047D	047E	047F												
Ψ	ψ	Θ	θ	ϐ	ϑ	ϐ	ϑ	Оу	оу	О	о	Ѓ	ѓ	Ѓ	ѓ												
0480	0481	0482	0483	0484	0485	0486	0487	0488	0489	048A	048B	048C	048D	048E	048F												
С	с	ѣ	ѣ	ѣ	ѣ	ѣ	ѣ	ѣ	ѣ	Й	й	Ђ	ђ	Р	р												

- 1) Inline characters incl. OCS
- 2) Superscripts

2DE0	2DE1	2DE2	2DE3	2DE4	2DE5	2DE6	2DE7	2DE8	2DE9	2DEA	2DEB	2DEC	2DED	2DEE	2DEF
Б	В	Г	Д	Ж	З	К	Л	М	Н	О	П	р	с	Т	Х
2DF0	2DF1	2DF2	2DF3	2DF4	2DF5	2DF6	2DF7	2DF8	2DF9	2DFA	2DFB	2DFC	2DFD	2DFE	2DFF
Ц	Ч	Ш	Щ	Ѡ	ѡ	Ѣ	ѣ	Ѥ	ѥ	Ѧ	ѧ	Ѩ	ѩ	Ѫ	ѫ

2. Cyrillic Unicode: Coming up

superscripts and subscripts

	...0	...1	...2	...3	...4	...5	...6	...7	...8	...9	...A	...B	...C	...D	...E	...F
Cyrillic Extended-D																
U+??x	а	б	в	г	д	е	ж	з	и	к	л	м	о	п	р	с
	т	у	ф	х	ц	ч	ш	щ	ы	э	ю	я	ѓ	ә	і	ј
	ө	ү	1													
	а	б	в	г	д	е	ж	з	и	к	л	м	н	о	п	р
	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я	г
	і	ѕ	џ													
	г	ѕ	џ	ө	ү	1	ћ	ђ	љ	њ	є	ћ	ђ	љ	њ	є

Cyrillic modifier letters
Proposal: Kirk Miller

Cyrillic 1: Latest Additions (Cyrillic Extended-C)

1C80	1C81	1C82	1C83	1C84	1C85	1C86	1C87	1C88	1C89	1C8A	1C8B	1C8C	1C8D	1C8E	1C8F
ѳ	Ѥ	Ѧ	Ѩ	Ѭ	Ѯ	Ѱ	Ѳ	Ѵ							

Andreev, Ju. Šardt, N. Simmons (Ponomar Project)
Proposal L2/13-153 (2013)

Cyrillic Extended-C

1C8F

Historic letter variants

1C80	ѳ	CYRILLIC SMALL LETTER ROUNDED VE
1C81	Ѥ	CYRILLIC SMALL LETTER LONG-LEGGED DE
1C82	Ѧ	CYRILLIC SMALL LETTER NARROW O
1C83	Ѩ	CYRILLIC SMALL LETTER WIDE ES
1C84	Ѭ	CYRILLIC SMALL LETTER TALL TE
1C85	Ѯ	CYRILLIC SMALL LETTER THREE-LEGGED TE
1C86	Ѱ	CYRILLIC SMALL LETTER TALL HARD SIGN
1C87	Ѳ	CYRILLIC SMALL LETTER TALL YAT
1C88	Ѵ	CYRILLIC SMALL LETTER UNBLENDED UK

- lowercase only - no capital letters?
- inline characters only - no superscripts?

Cyrillic 1: Latest Additions (Cyrillic Extended-C)

1C80	1C81	1C82	1C83	1C84	1C85	1C86	1C87	1C88	1C89	1C8A	1C8B	1C8C	1C8D	1C8E	1C8F
Ѻ	Ѡ	Ѣ	Ѥ	Ѧ	Ѩ	Ѭ	Ѯ	Ѱ							

Andreev, Ju. Šardt, N. Simmons (Ponomar Project)
Proposal L2/13-153 (2013)

“For these three reasons, we are unable to endorse any argument for registration that depends entirely on the mere coexistence in a single document of different letterforms with comparable orthographic and linguistic function.”

“For the reasons described above, we are unable to endorse any proposal for the registration of additional early Cyrillic characters in Unicode that should more properly be addressed on the font level, rather than the character set level.”

(Birnbaum–Cleminson–Kempgen–Ribarov:
Unicode White Paper 2008, 4 and 6 resp.)

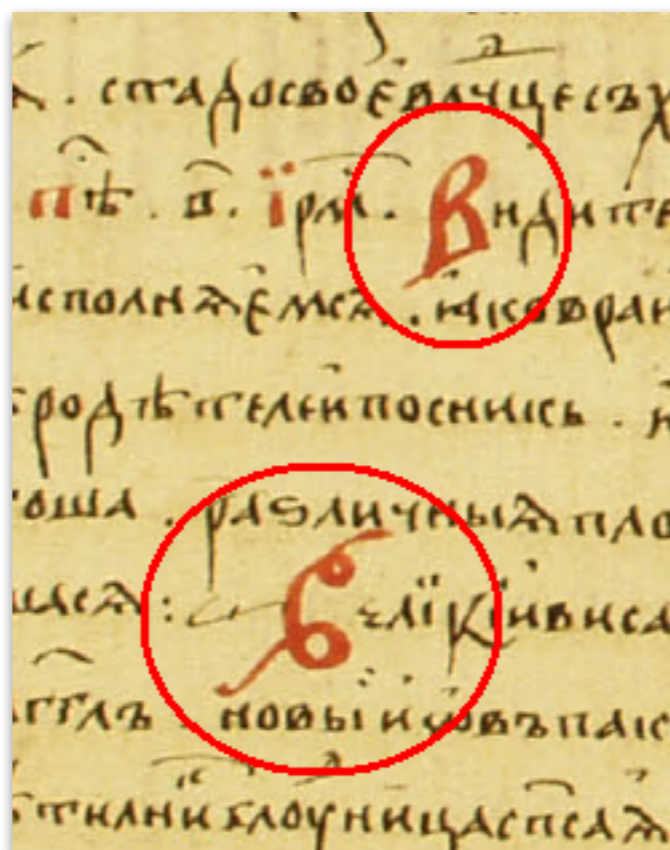


two letters in Greek,
two variants in OCS

Cyrillic Extended-C

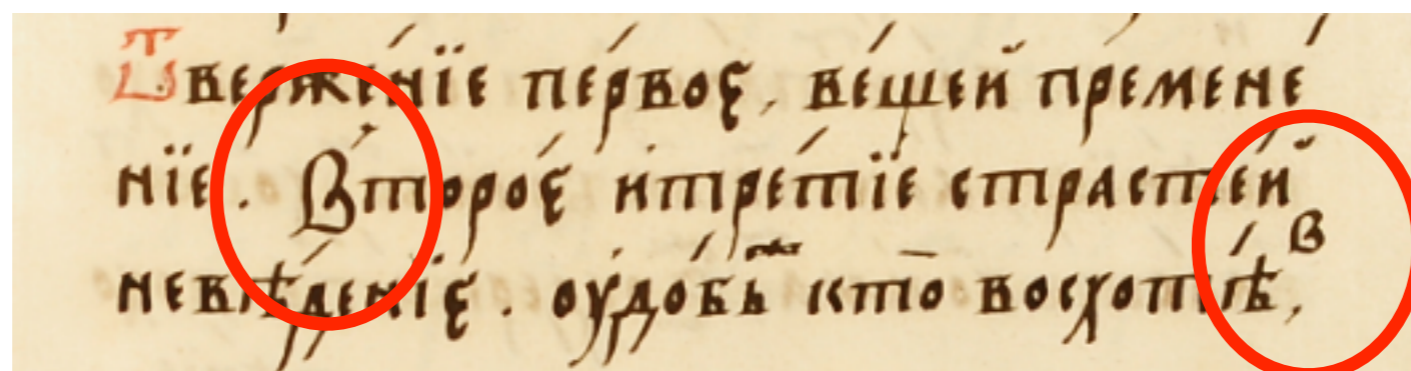
1C80	1C81	1C82	1C83	1C84	1C85	1C86	1C87	1C88	1C89	1C8A	1C8B	1C8C	1C8D	1C8E	1C8F
Ѡ	ѡ	Ѣ	ѣ	Ѥ	ѥ	Ѧ	ѧ	Ѩ							

round v



Buslaevskaja psaltyr' 253

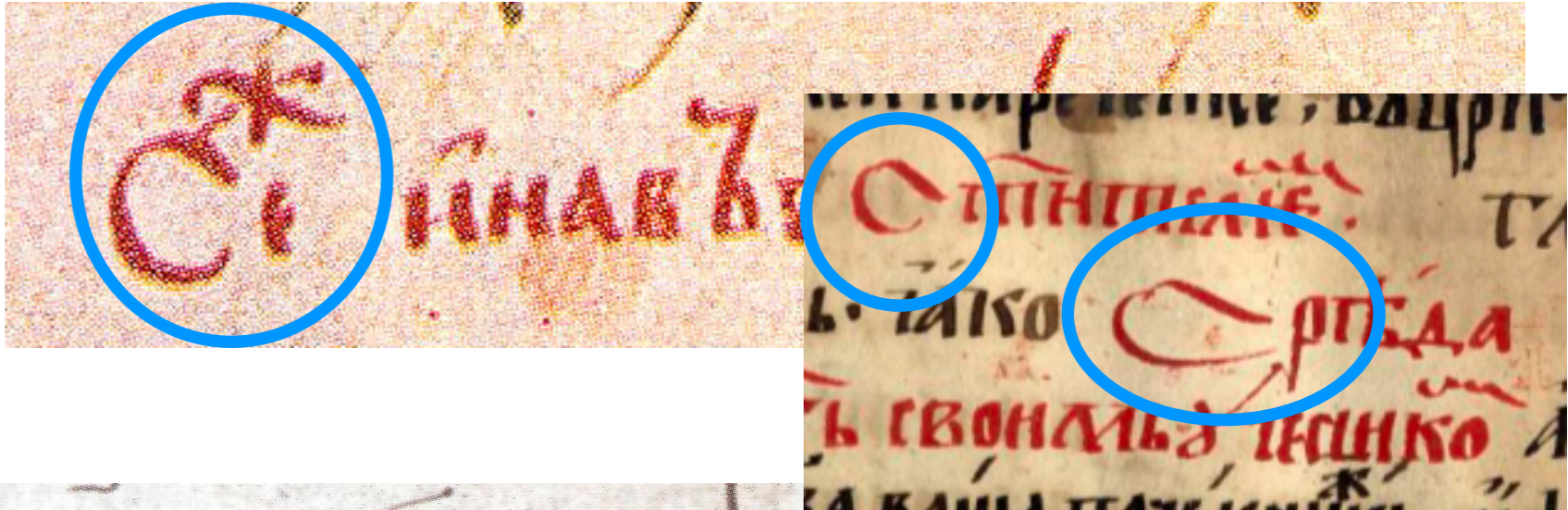
CAPITAL!
superscript!



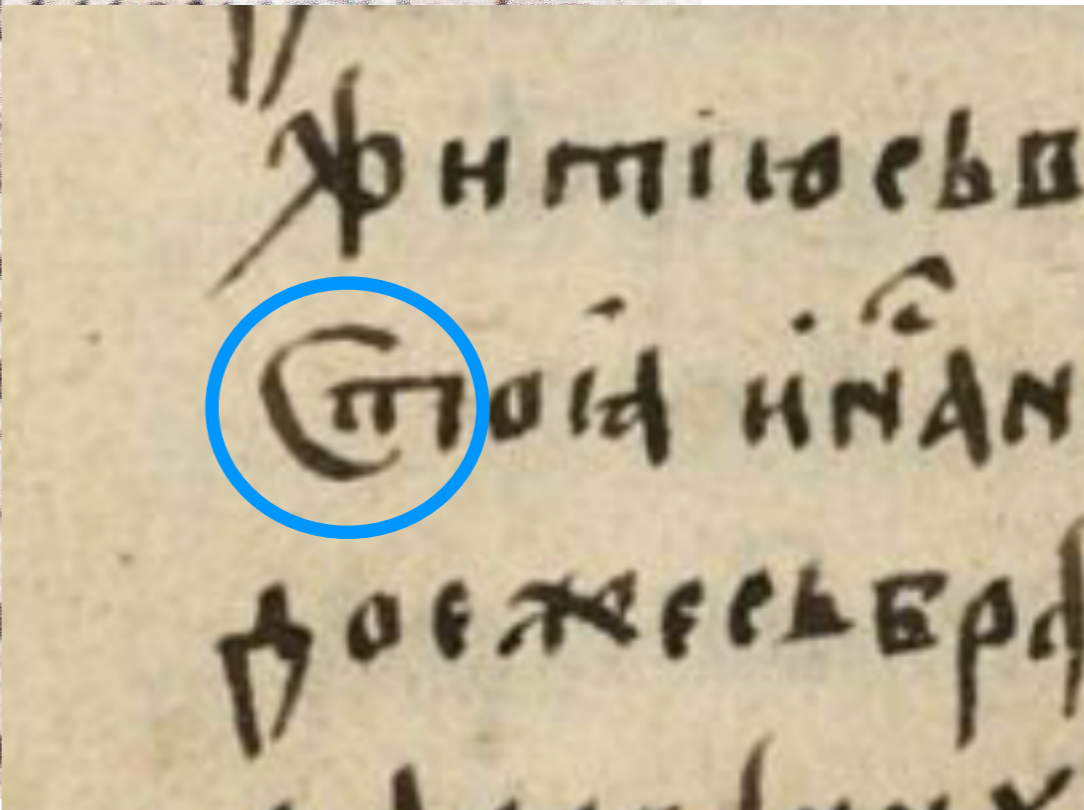
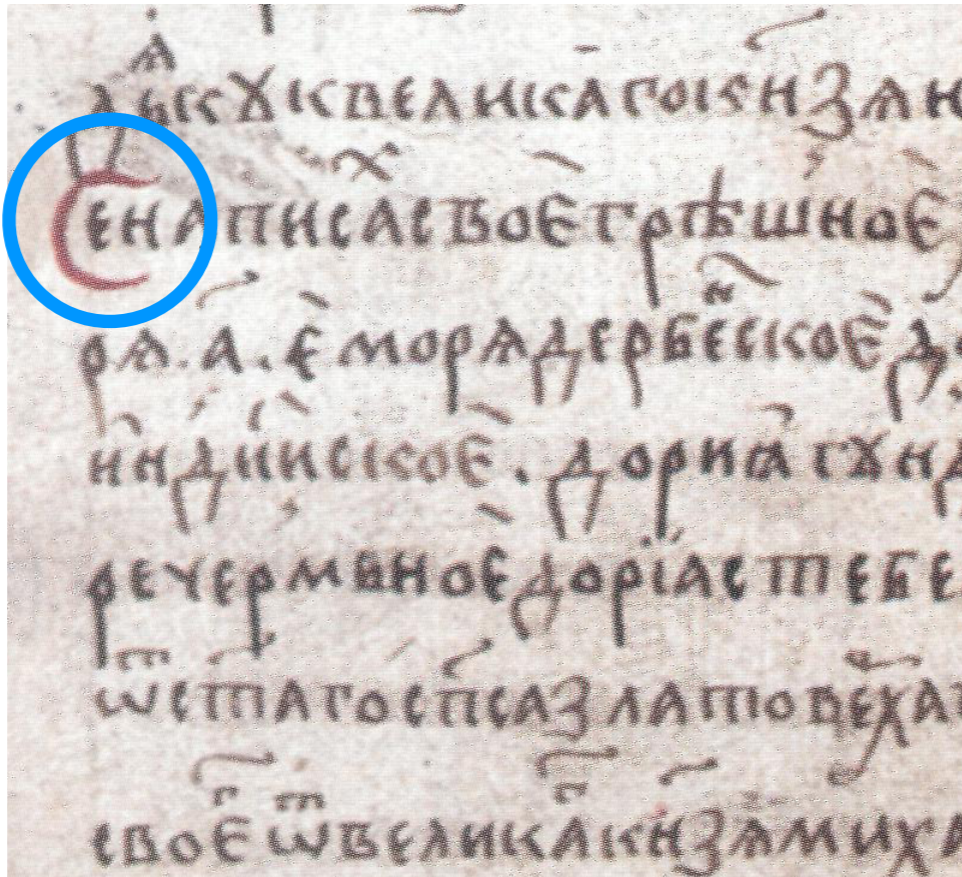
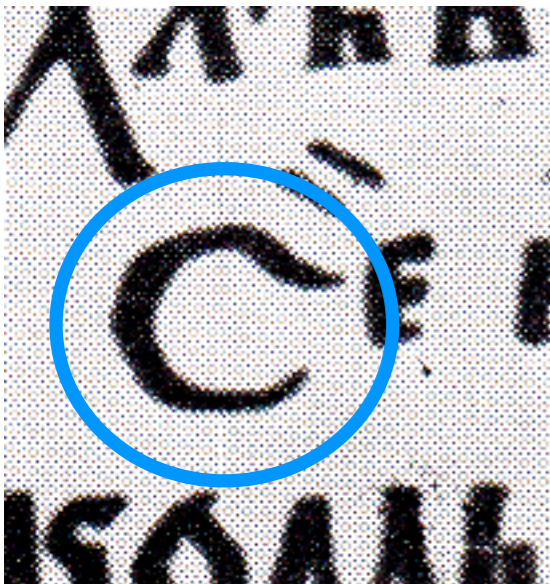
Cyrillic Extended-C

1C80	1C81	1C82	1C83	1C84	1C85	1C86	1C87	1C88	1C89	1C8A	1C8B	1C8C	1C8D	1C8E	1C8F
Ѡ	ѡ	Ѣ	ѣ	Ѥ	ѥ	Ѧ	ѧ	Ѩ							

broad c



CAPITAL!



Cyrillic Extended-C

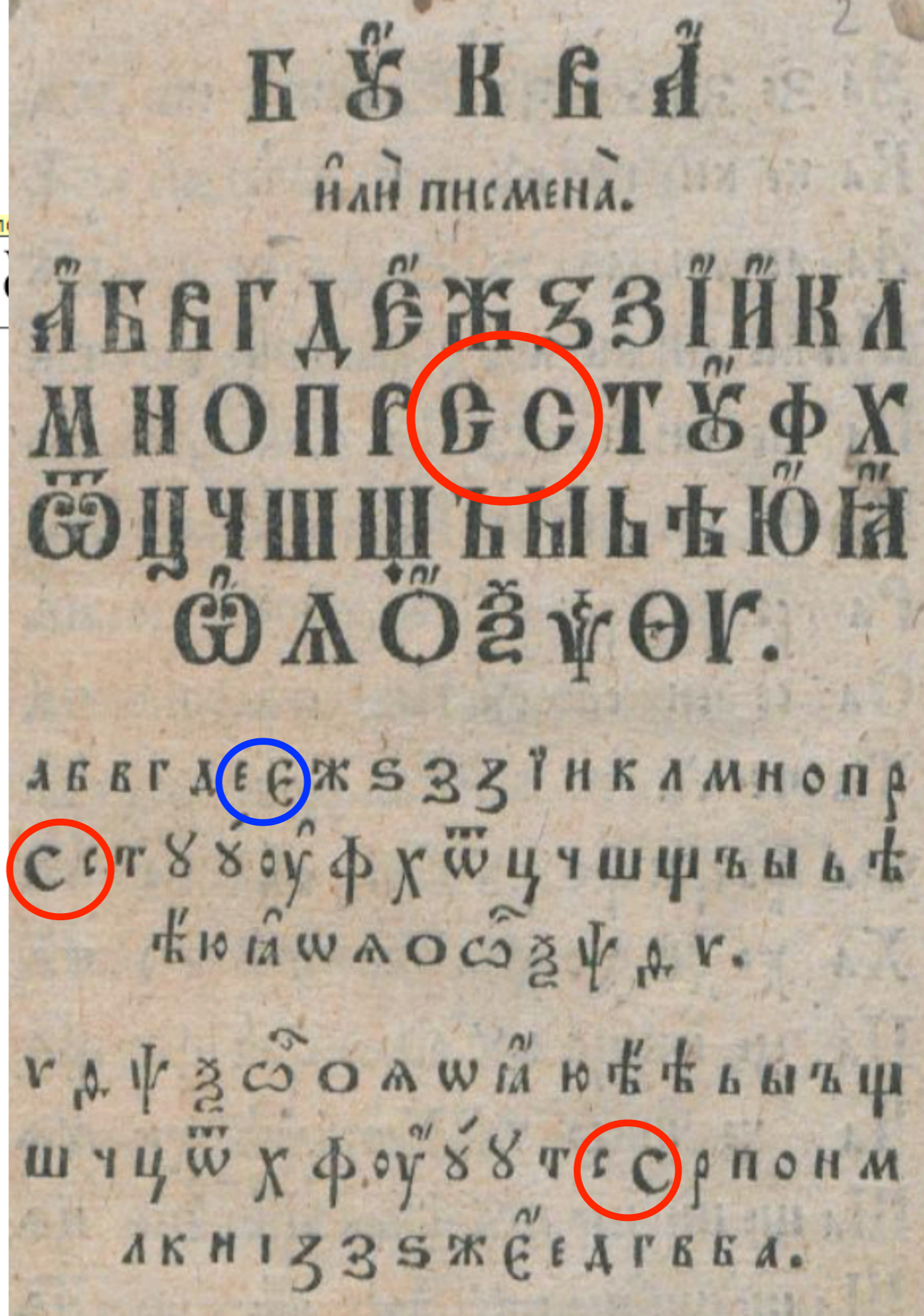
1C80	1C81	1C82	1C83	1C84	1C85	1C86	1C87	1C88
Ѡ	ѡ	Ѣ	ѣ	Ѥ	ѥ	Ѧ	ѧ	Ѩ

broad c

CAPITAL!

two 's' distinct elements of alphabet

Букварь языка славянского [Текст]. - Вильна :
Тип. Троицкая, 1782. - [32] л. : 1... [описание](#)

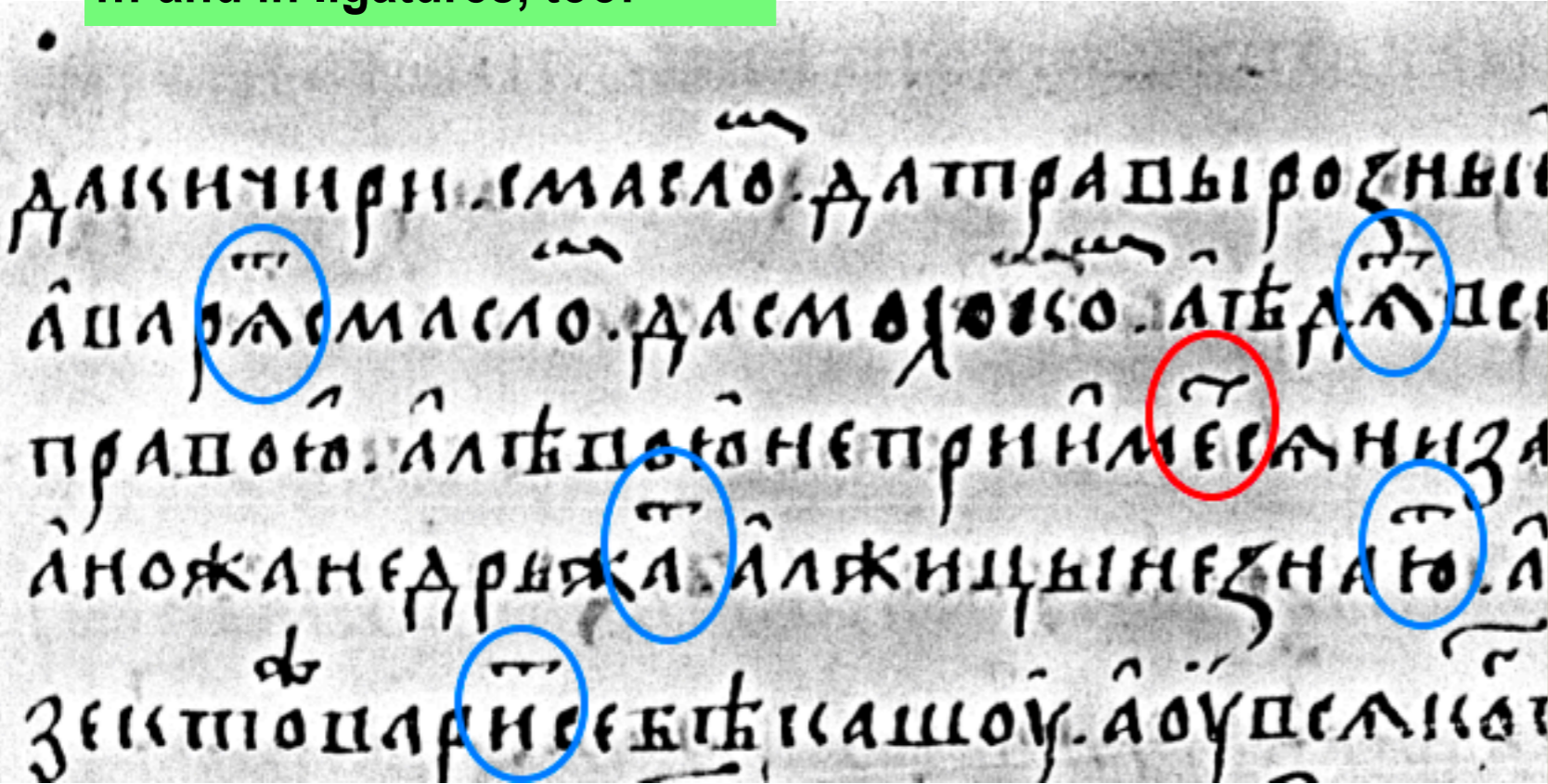


Cyrillic Extended-C

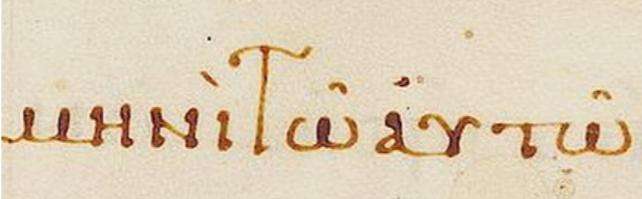
1C80	1C81	1C82	1C83	1C84	1C85	1C86	1C87	1C88	1C89	1C8A	1C8B	1C8C	1C8D	1C8E	1C8F
Ѡ	ѡ	Ѣ	ѣ	Ѥ	ѥ	Ѧ	ѧ	Ѩ							

tall t

superscript!
... and in ligatures, too!



Greek blueprint



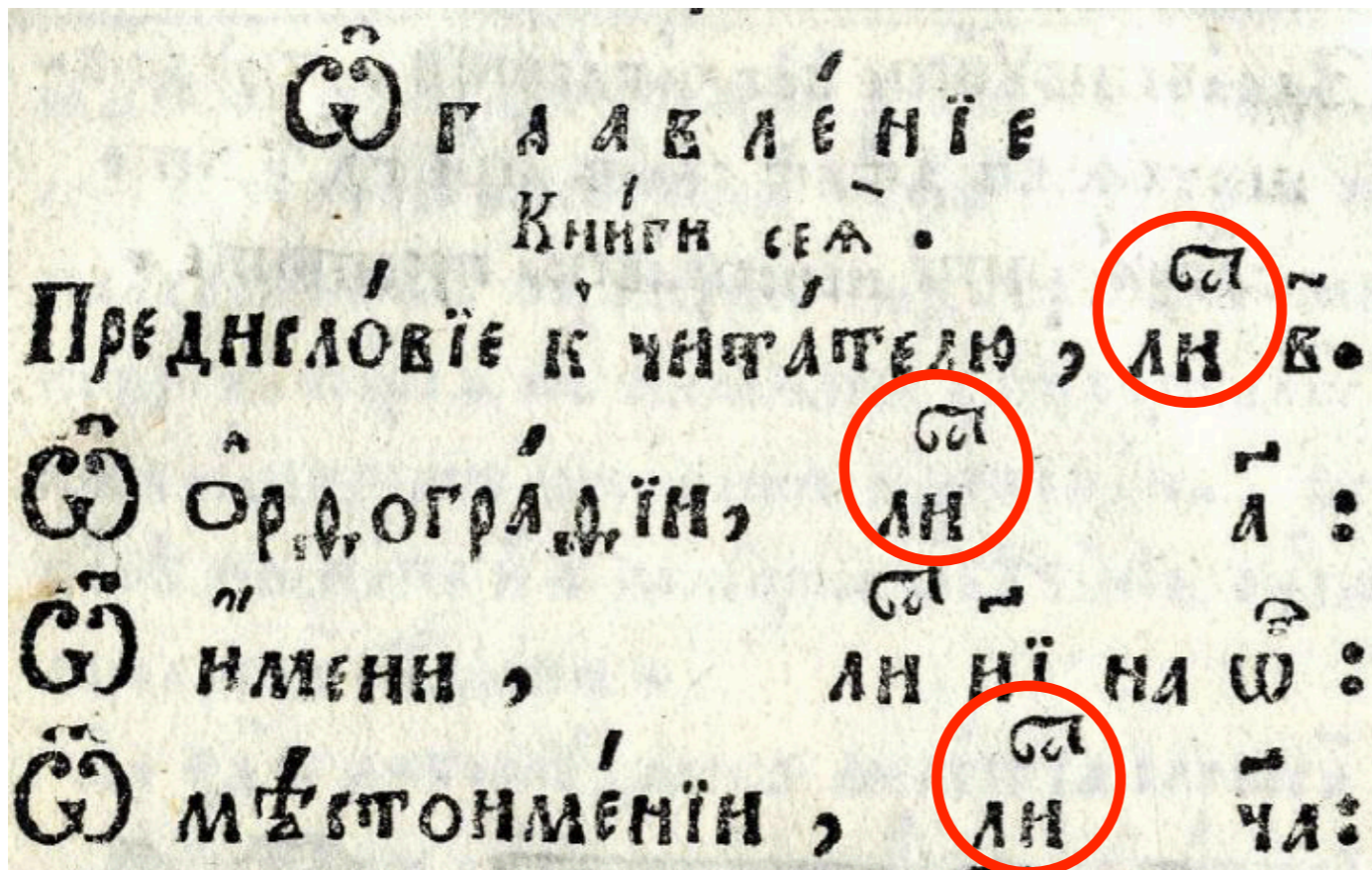
Cyrillic Extended-C

1C80	1C81	1C82	1C83	1C84	1C85	1C86	1C87	1C88	1C89	1C8A	1C8B	1C8C	1C8D	1C8E	1C8F
Ѣ	Ѥ	Ѧ	Ѩ	Ѭ	Ѯ	Ѱ	Ѳ	Ѵ							

tall t

superscript
even in printing!

Ѵ РѢѢИ ѢтарожнѣѴ



Гдѣ на землю справн и
х дѣнь чотх стый своего
Ѵправн, прѣѣмх

Titov p. 31 and p. 32

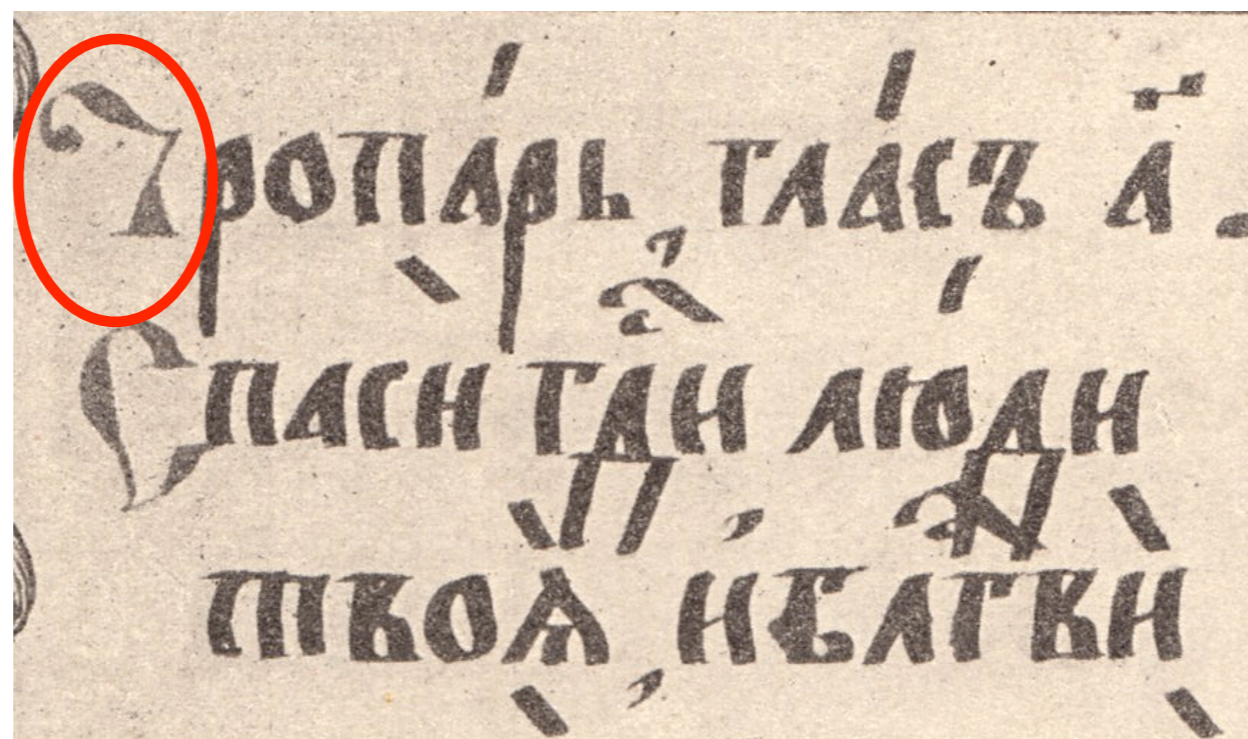
Smotrickij, Grammatika 1755

Cyrillic Extended-C

1C80	1C81	1C82	1C83	1C84	1C85	1C86	1C87	1C88	1C89	1C8A	1C8B	1C8C	1C8D	1C8E	1C8F
Ѹ	ѹ	Ѻ	ѻ	Ѽ	ѽ	Ѿ	ѿ	Ѡ							

tall t

CAPITAL!



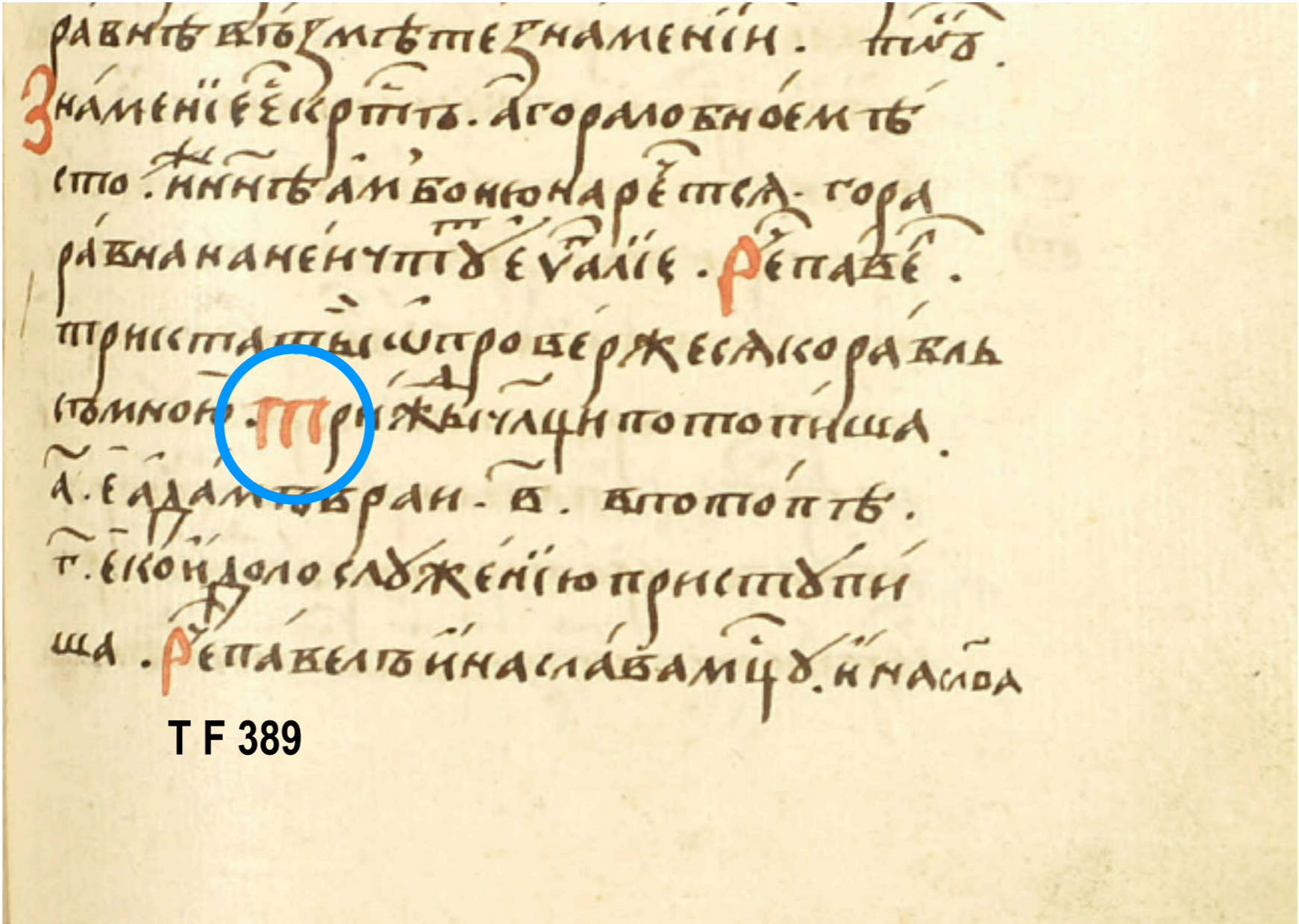
Sobolevskij 1905, 55: Kniga o muzax 1672

Cyrillic Extended-C

1C80	1C81	1C82	1C83	1C84	1C85	1C86	1C87	1C88	1C89	1C8A	1C8B	1C8C	1C8D	1C8E	1C8F
Ѡ	ѡ	Ѣ	ѣ	Ѥ	ѥ	Ѧ	ѧ	Ѩ							

three-legged t

CAPITAL!



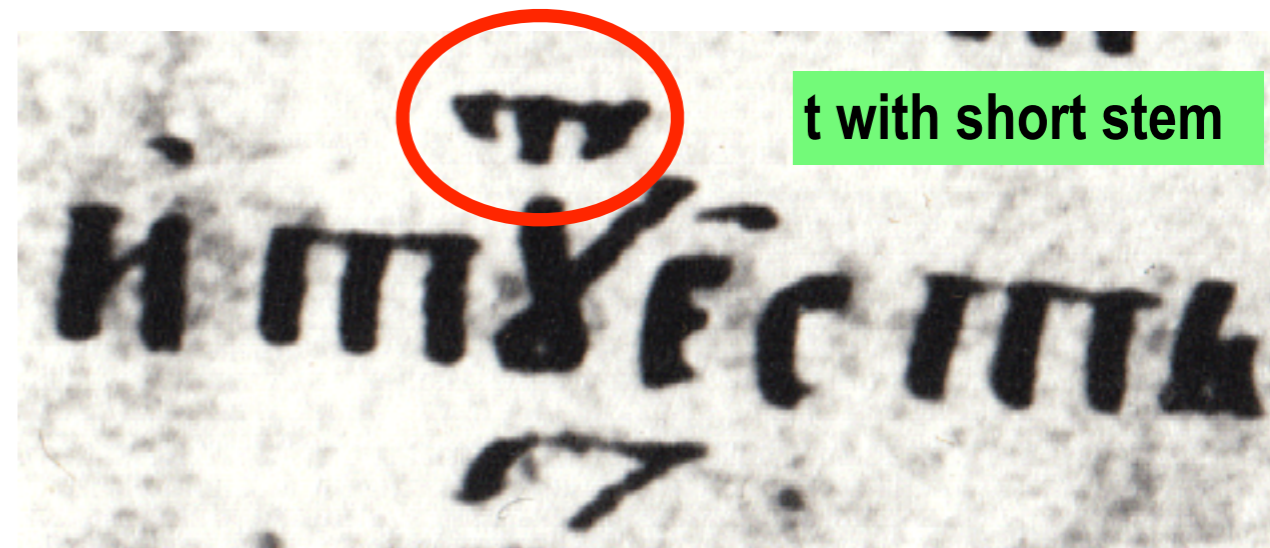
T F 389

spotted by William Veder!

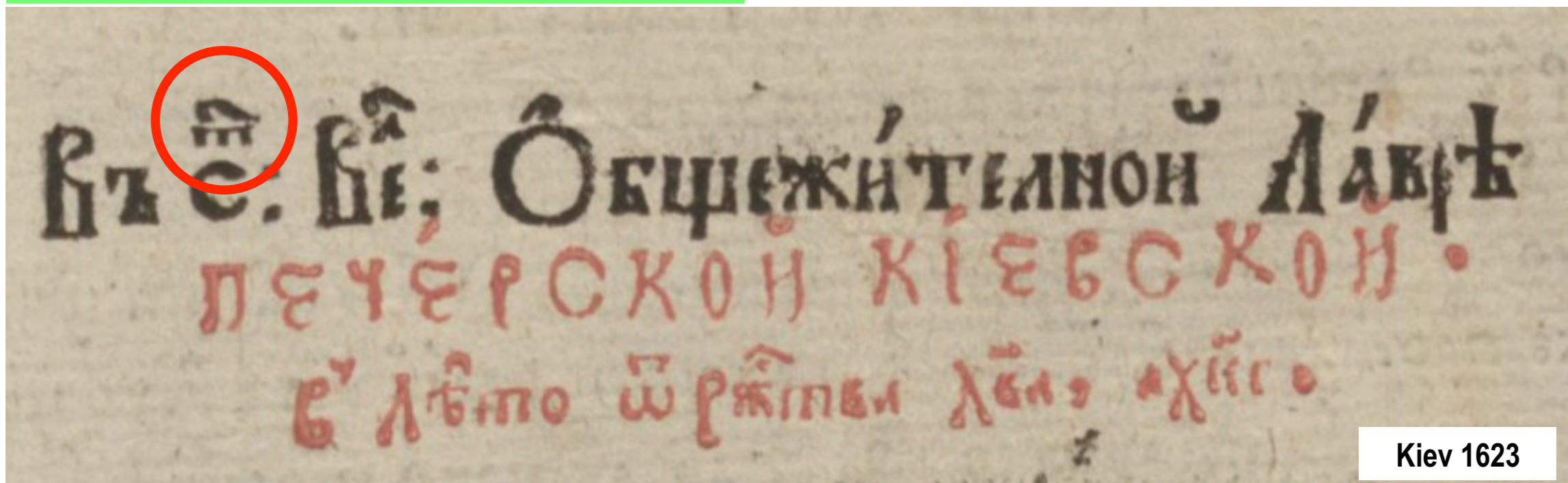
Cyrillic Extended-C

1C80	1C81	1C82	1C83	1C84	1C85	1C86	1C87	1C88	1C89	1C8A	1C8B	1C8C	1C8D	1C8E	1C8F
Ѡ	ѡ	Ѣ	ѣ	Ѥ	ѥ	Ѧ	ѧ	Ѩ							

three-legged t



three-legged <t> or a standard <t> with short stem?



Kiev 1623

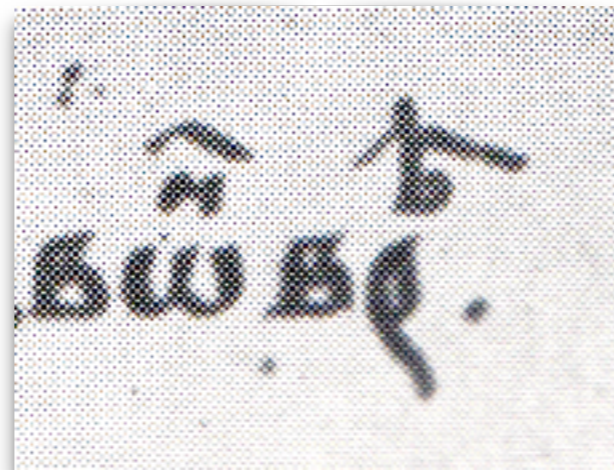
Cyrillic Extended-C

1C80	1C81	1C82	1C83	1C84	1C85	1C86	1C87	1C88	1C89	1C8A	1C8B	1C8C	1C8D	1C8E	1C8F
Ѻ	ѻ	Ѽ	ѽ	Ѿ	ѿ	Ѡ	ѡ	Ѣ							

high yat'

standard yat' or
high yat'?

Amalgam
of superscript and
pokrytie?



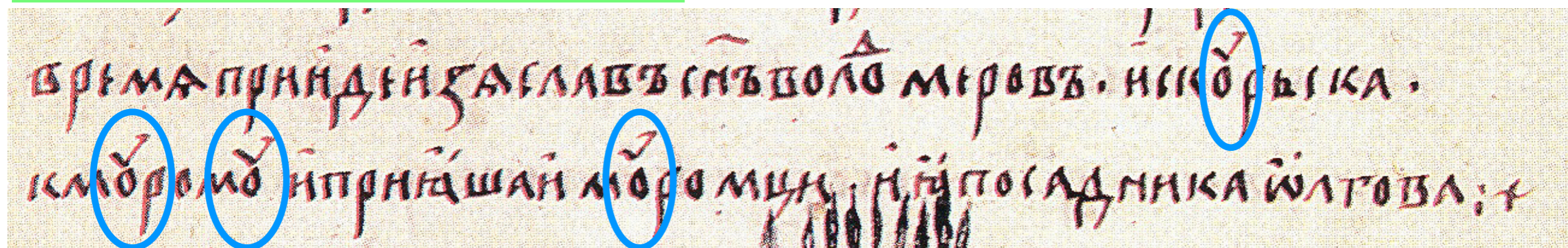
Cyrillic Extended-C

1C80	1C81	1C82	1C83	1C84	1C85	1C86	1C87	1C88	1C89	1C8A	1C8B	1C8C	1C8D	1C8E	1C8F
Ѡ	ѡ	Ѣ	ѣ	Ѥ	ѥ	Ѧ	ѧ	Ѩ							

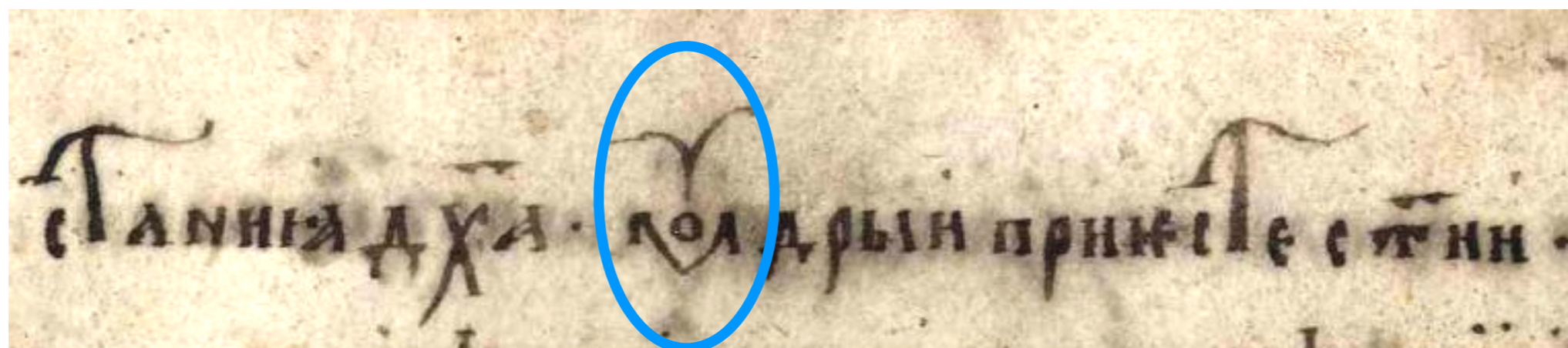
unblended oy

... long before printing in manuscripts:

Radzivillovskaja letopis', fol. 32 (XV c.)



superscript



Cyrillic Extended-C

1C80	1C81	1C82	1C83	1C84	1C85	1C86	1C87	1C88	1C89	1C8A	1C8B	1C8C	1C8D	1C8E	1C8F
Ѡ	ѡ	Ѣ	ѣ	Ѥ	ѥ	Ѧ	ѧ	Ѩ							

unblended oy

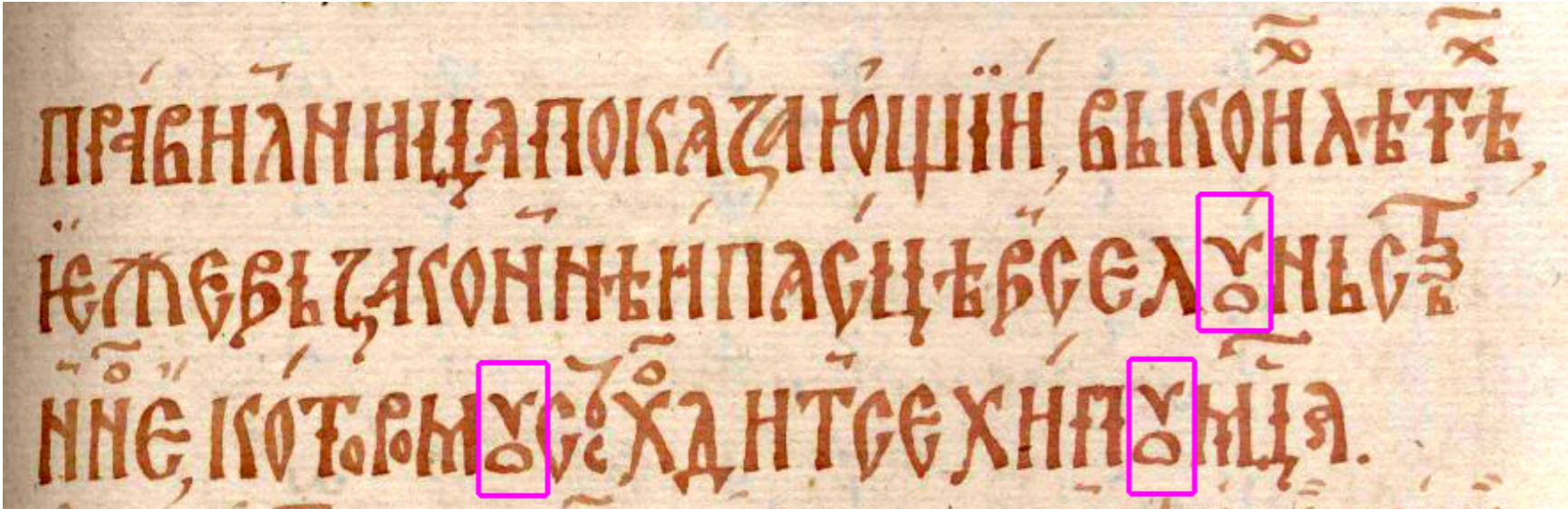
CAPITAL!

НѢРАЗОРИТЪ СЕ.
 СЪДЕЩОУЖЕ ЕСИМ
 Ѡ НА ГОРѢ ЕЛИФЪНЬ
 СУЪН ПРНСТѢ Ѡ ПН

lowercase

Ѡ ѡ Ѣ

Moščin, Al'bom... p. 127, Bosansko ev., XIII. v.



Синтагма Матије Властара са антилатинским списима Г. Паламе и др., 401

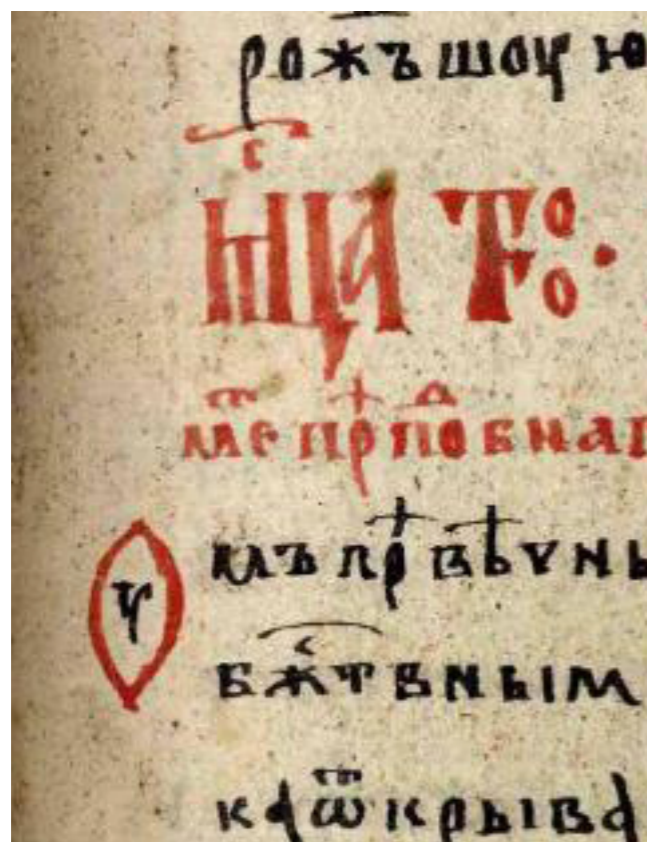
Cyrillic Extended-C

1C80	1C81	1C82	1C83	1C84	1C85	1C86	1C87	1C88	1C89	1C8A	1C8B	1C8C	1C8D	1C8E	1C8F
Ѻ	Ѡ	Ѣ	Ѥ	Ѧ	Ѩ	Ѭ	Ѯ	Ѱ							

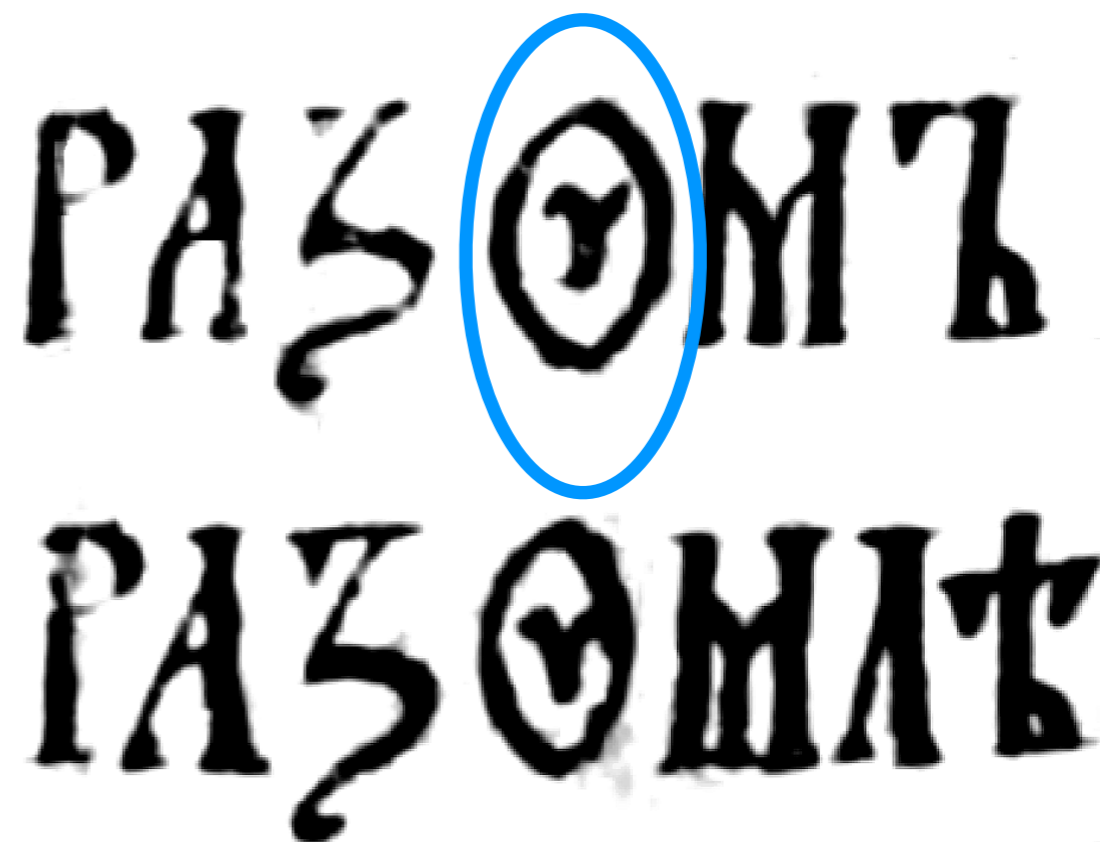
unblended oy

CAPITAL!

completely different solutions:



Minei za jul (13??), p. 49



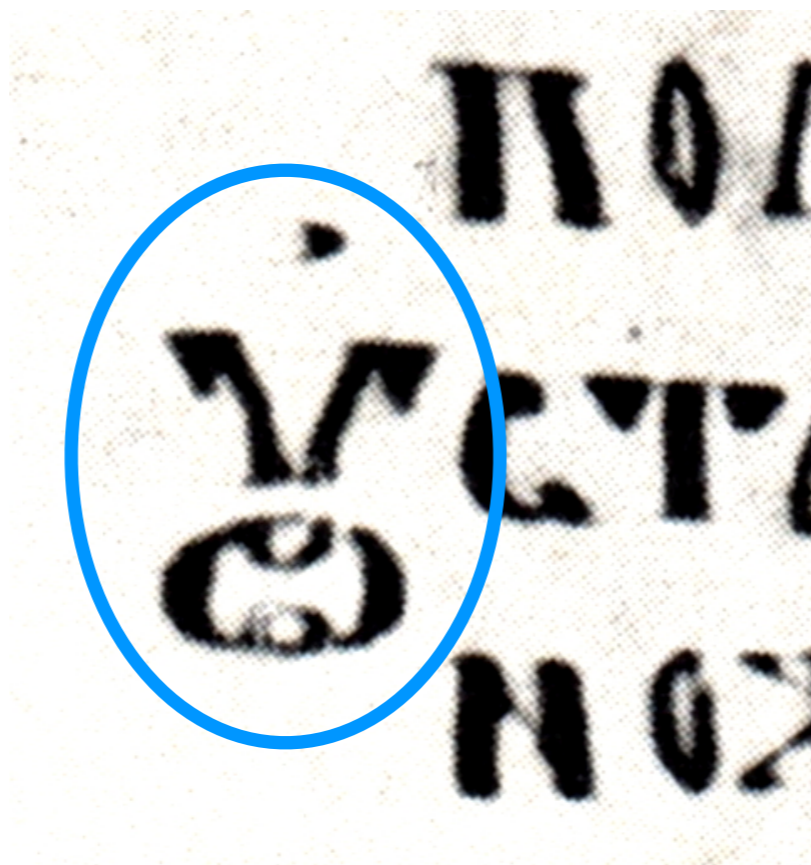
Codicis Sofiensis fol. 158 [recte 160]
from Jagić, Psalterium Bononiense, 1907, appendix pl. XVI.

Cyrillic Extended-C

1C80	1C81	1C82	1C83	1C84	1C85	1C86	1C87	1C88	1C89	1C8A	1C8B	1C8C	1C8D	1C8E	1C8F
Ѹ	ѹ	Ѻ	ѻ	Ѽ	ѽ	Ѿ	ѿ	Ѡ							

unblended oy

lowercase + CAPITAL



also used with

⦿ [U+047A resp. U+047B] as lower half

⦿ [U+A668 resp. U+A669] as lower half

Bolonski psaltir fol. 83v = p. 166 resp.
fol. 99 = p. 197

> should have been encoded as a sequence of base character (⦿⦿ ⦿⦿ ⦿⦿) plus combining izhica (Ѵѵ) in the first place > combining izhica should have been added.

Cyrillic Extended-C

> implementation incomplete / inconsistent

	1C80	1C81	1C82	1C83	1C84	1C85	1C86	1C87	1C88
	ѳ	Д	о	с	Т	ш	Ъ	ѣ	Ѧ
CAPS	✓			✓	✓	✓			✓
super	✓				✓	(✓)		(✓)	✓

not true for all characters!

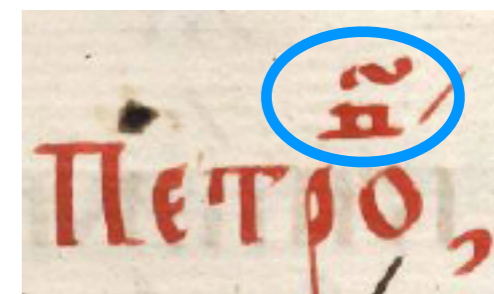
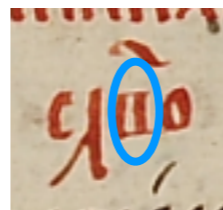
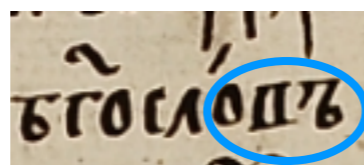
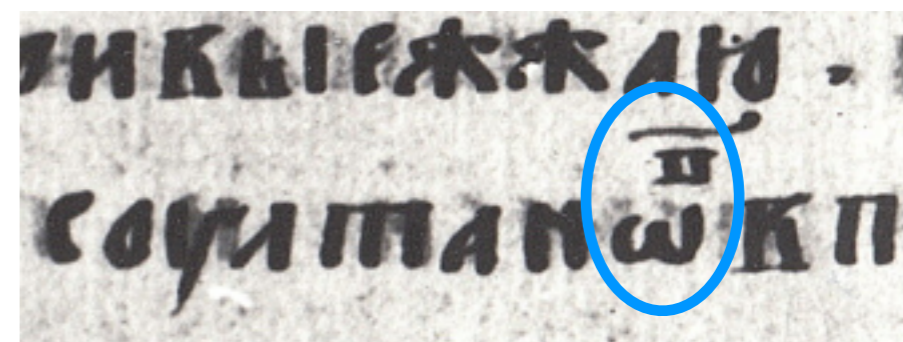
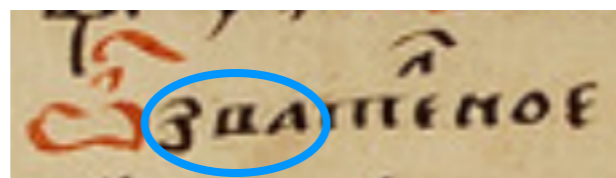
“The characters under consideration presently do not have extant uppercase forms. ... we do not wish to construct non-existent Cyrillic graphemes. Rather than constructing such uppercase forms artificially, we propose that only the lower-case forms be encoded and that the archaic characters uppercase to the upper-case forms of the modern character.” (Proposal N4607)

Cyrillic Extended-C: floodgates open?

1C80	1C81	1C82	1C83	1C84	1C85	1C86	1C87	1C88	1C89	1C8A	1C8B	1C8C	1C8D	1C8E	1C8F
Ѡ	ѡ	Ѣ	ѣ	Ѥ	ѥ	Ѧ	ѧ	Ѩ							

square v

inline
superscript
CAPITAL

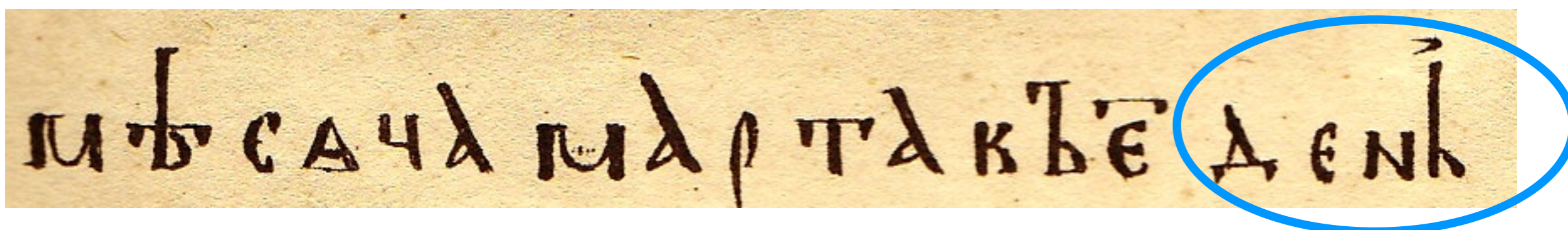
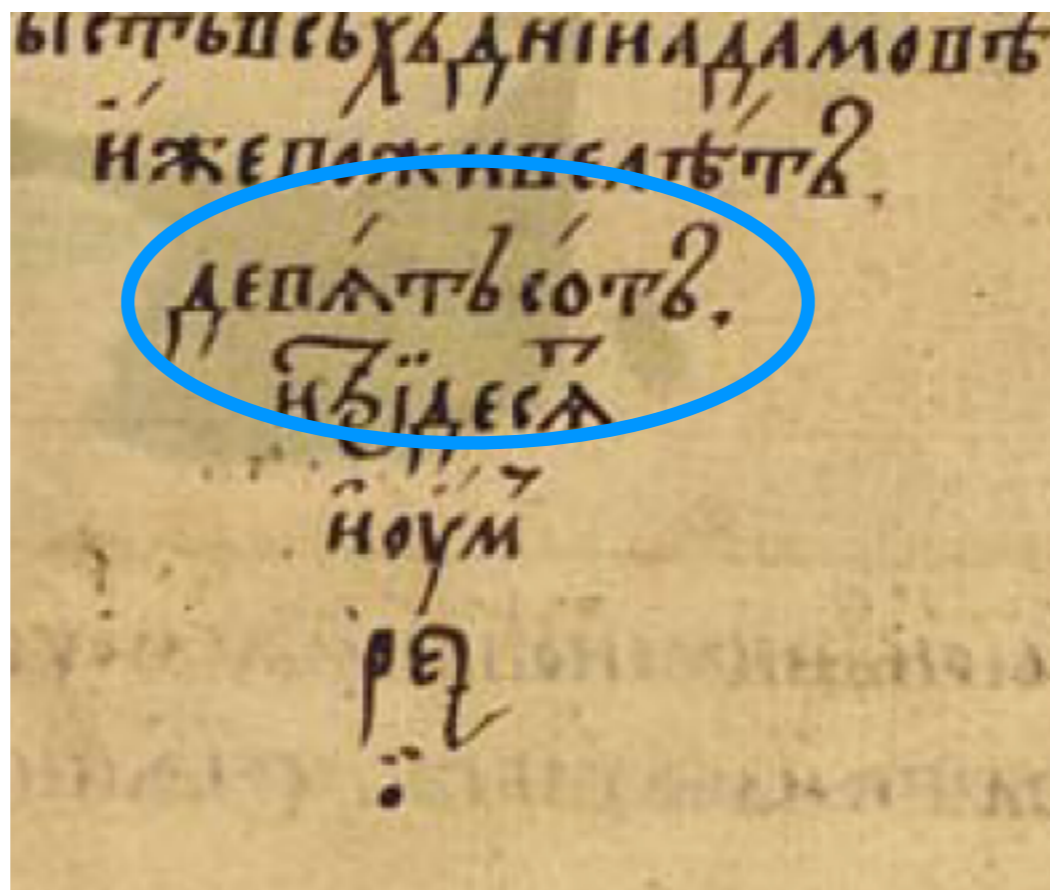


Cyrillic Extended-C: floodgates open?

1C80	1C81	1C82	1C83	1C84	1C85	1C86	1C87	1C88	1C89	1C8A	1C8B	1C8C	1C8D	1C8E	1C8F
Ѡ	ѡ	Ѣ	ѣ	Ѥ	ѥ	Ѧ	ѧ	Ѩ							

high yer

inline



Cyrillic Extended-C: no more additions?

1C80	1C81	1C82	1C83	1C84	1C85	1C86	1C87	1C88	1C89	1C8A	1C8B	1C8C	1C8D	1C8E	1C8F
ѐ	А	о	с	Т	ш	Ъ	ѣ	Ѧ							

“A representative of the Old Believers participated in several meetings of the Unicode Technical Committee and was able to get their proposals approved. [...] In hindsight we probably should have sought out more opposing views, so it wasn't so one sided.”

Deborah Wanders, UTC technical director, personal communication

Cyrillic 2: Kievan Printing, 17th c.

uses Greek Sigma (and Greek N)



Cyrillic 2: Kievan Printing, 17th c.

Greek-style old N
Greek-style Y for izhica
uppercase-only



Перша сторінка Синопису.

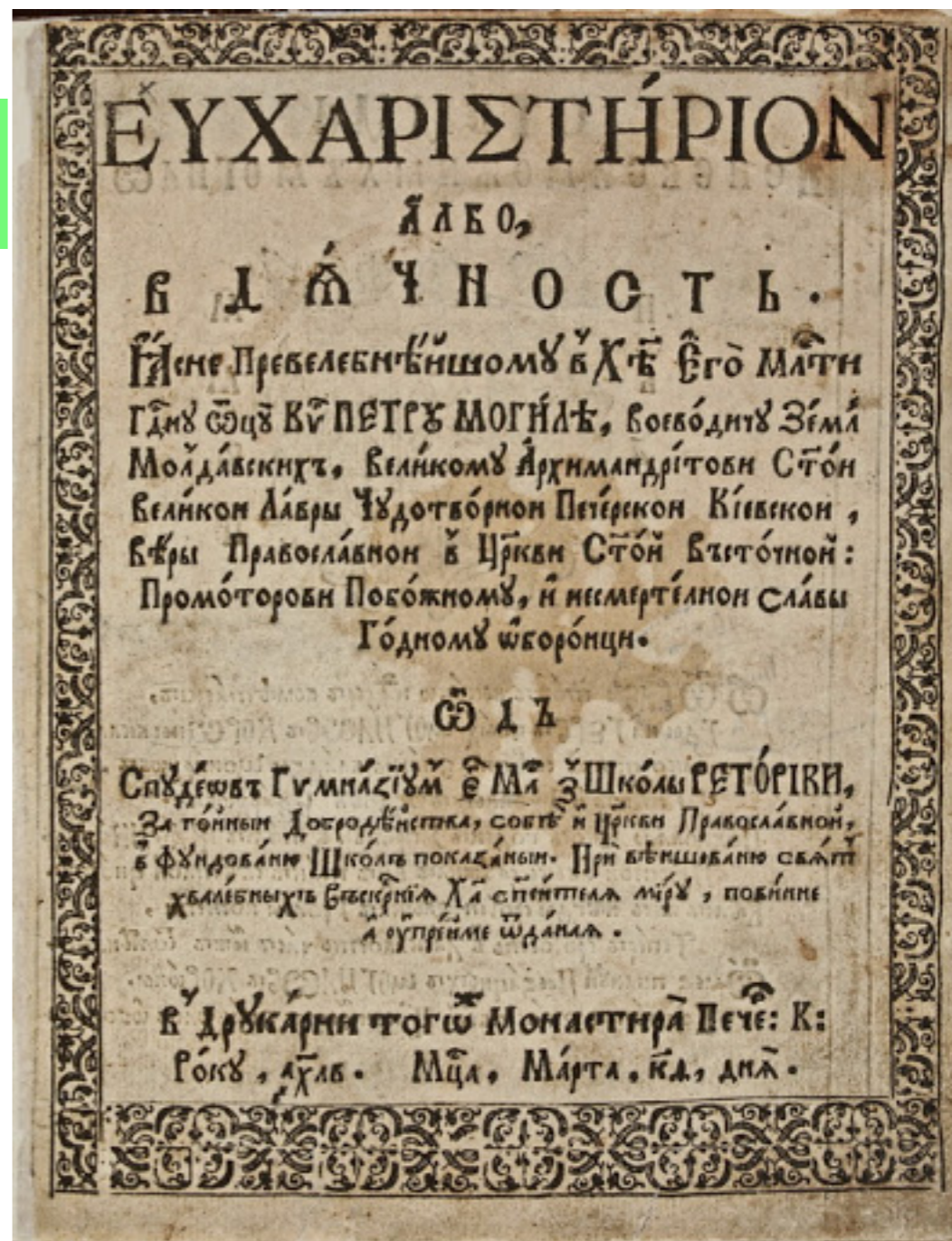


two letters in Greek,
two variants in OCS

Cyrillic 2: Kievan Printing, 17th c.

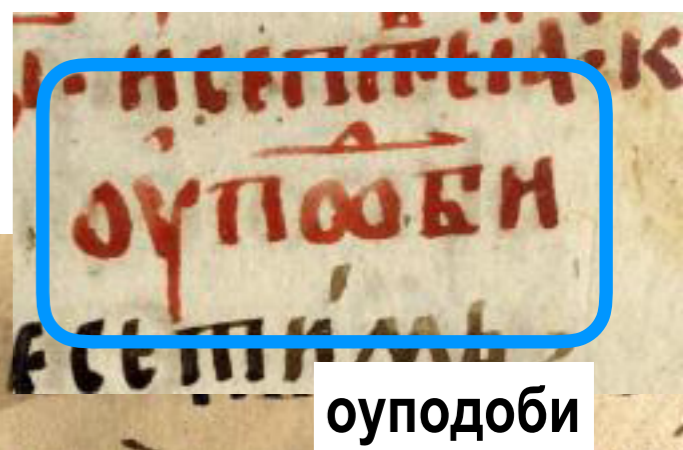
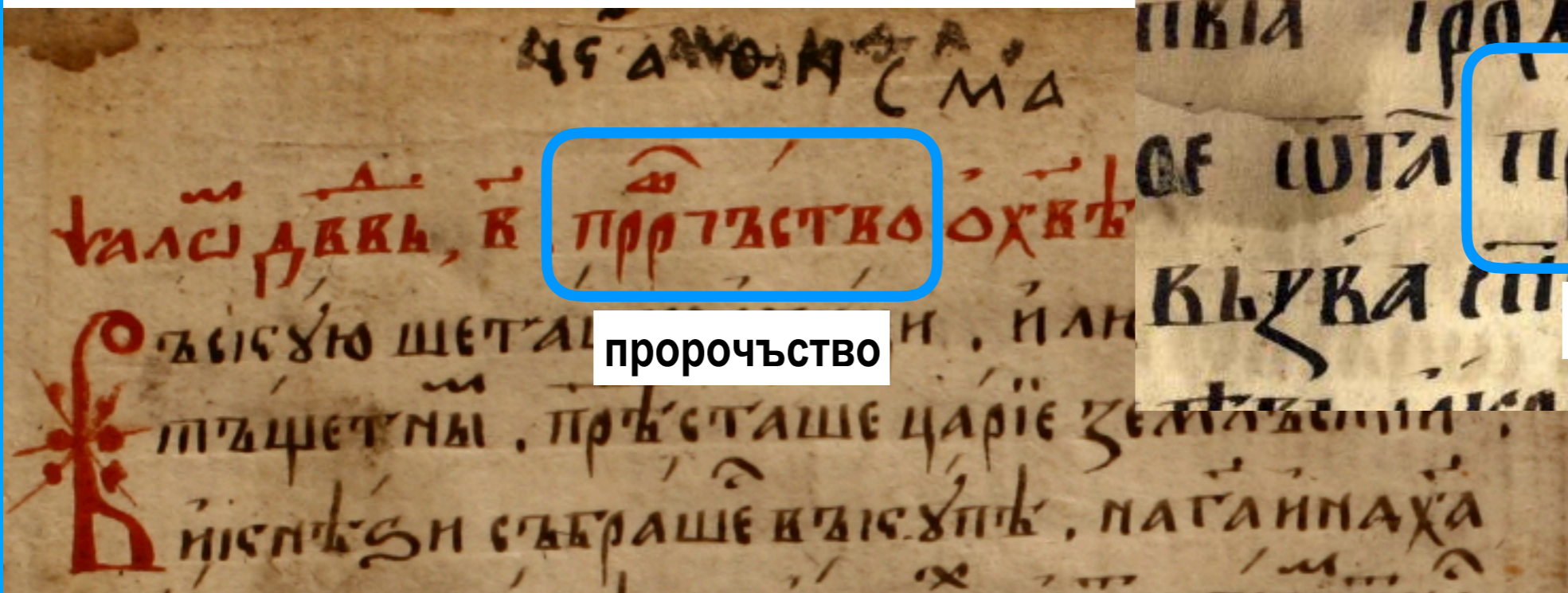
uses Greek Sigma, but sometimes...

it's real Greek
even in a Slavonic text

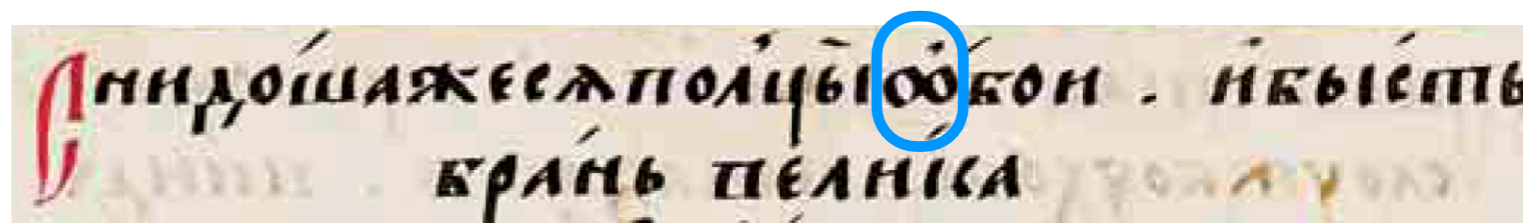


Cyrillic 3: Additional characters

Superscript ∞ ?



- two letters, non-sequential: proč"stvo
- two letters combined under one bow
- U+A699 not available as a superscript



A680	A681	A682	A683	A684	A685	A686	A687	A688	A689	A68A	A68B	A68C	A68D	A68E	A68F
Д	д	Ѕ	ѕ	ЗК	зк	Ѵ	ѵ	Дз	дз	Ѧ	ѧ	Ѩ	ѩ	Ѫ	ѫ
A690	A691	A692	A693	A694	A695	A696	A697	A698	A699	A69A	A69B	A69C	A69D	A69E	A69F
Ѭ	ѭ	Ѯ	ѯ	Ѱ	ѱ	Ѳ	ѳ	∞	∞	Ѵ	ѵ	Ѷ	ѷ	Ѹ	ѹ

Cyrillic 3: Additional characters

Vuk's third letter: **Д**.

Сад ја мислим: кад се њ са т саспави-
пи могло, оно се може (барем у печаша-
њу) и са д, л, н, саспавипи, ево ја ве-
лим овако: **Д, Љ, Њ.****) ето гошов посао.

— 9 —

Сад наша Азбуна, или, како прости Сер-
бљи по Србији и по Херцеговини говоре,
Буквица, има 29 писмена, међу којима
ни једно ни је сувнше, ни једно звукопре-
миенљиво, ниши које недоспаје, него и је
шаман онолико, колико својство Србског
језика изискује:

а, б, в, г, д, **Д**, е, ж, з, и, ј, к, л,
Љ, м, н, **Њ**, о, п, р, с, ш, ћ, у, ф, х,
ц, ч, ш. *)

— 10 —

може изговорити без помоћи успана, јези-
ка, зуба и оспали орудија говоренња, и ша-
кови писмена у нас има 24. која јесу: б,
в, г, д, **Д**, ж, з, ј, к, л, **Љ**, м, н, **Њ**,
и, р, с, ш, ћ, ф, х, ц, ч, ш.

Једно самогласно писме само, или са-
спављено са једним, или са више согласни
писмене, зовесе слог.

Cyrillic 3: Additional characters

Vuk's third letter: Д.

Vuk übernimmt Mrkaljs Alphabet vollkommen und bringt noch einige Verbesserungen an. Mrkaljs und Vuks Reform stehen demnach in folgendem Verhältnis:

Mrkalj: а, б, в, д, дѣ, е, ж, з, и, љ, н, лѣ, м, н, нѣ, о, п,

Karadžić: а, б, в, д, Д, е, ж, з, и, љ, н, лѣ, м, н, нѣ, о, п,

Mrkalj: р, с, т, тѣ, у, ф, х, ц, ч, ш.

Karadžić: р, с, т, ѣ, у, ф, х, ц, ч, ш. ⁸²

Die beiden Alphabete stimmen überein:

- 1) in der Zahl der Buchstaben;
- 2) auch bei Vuk fehlt der Buchstabe ѱ;
- 3) den Konsonanten j notiert Vuk genauso wie Mrkalj.

Die Kyrilliza Vuks ist also in ihrer Grundlage die von Mrkalj. Vuk hat aber noch ein übriges getan; er verbesserte das Alphabet weiter, und zwar genau nach den Anweisungen Mrkaljs. Wie bereits gezeigt wurde, sprach Mrkalj den Zeichen лѣ, нѣ, тѣ, дѣ einen Übergangsscharakter zu, schlug die Schaffung eines einfachen Buchstabens für die einzelnen Laute vor. Der weiche Halbvokal hätte dann entfernt werden können und sein Prinzip ein Laut - ein Zeichen wäre realisiert gewesen. Vuk tat dann nichts anderes, als daß er den weichen Halbvokal mit dem Zeichen des zu palatalisierenden Lautes verschmelzen ließ: ѣ, ѣ, Д. Er führte für diese Laute keine

Cyrillic 3: Additional characters

Vuk's third letter: **Ђ**.

also used in print

Вук – ПИСМЕНИЦА из 1814. године - 29 СЛОВА СРПСКОГ ПИСМА

а	б	в	г	д
Ђ	е	ж	з	и
ї	к	л	љ	м
н	њ	о	п	р
с	ш	ћ	у	ф
х	џ	џ	Слово џ Вук не ставља у ову србицу	ш

Cyrillic 4: Transliteration of Glagolitic

- a) Glagolitic is usually transliterated in Cyrillic
- b) Glagolitic and Cyrillic have superscripts
- c) Can every Glagolitic superscript be reproduced in Cyrillic?

1E000	1E001	1E002	1E003	1E004	1E005	1E006	1E007	1E008	1E009	1E00A	1E00B	1E00C	1E00D	1E00E	1E00F
Ɱ	Ɱ̇	Ɱ̈	Ɱ̉	Ɱ̊	Ɱ̋	Ɱ̌	Ɱ̍	Ɱ̎	Ɱ̏	Ɱ̐	Ɱ̑	Ɱ̒	Ɱ̓	Ɱ̔	Ɱ̕
1E010	1E011	1E012	1E013	1E014	1E015	1E016	1E017	1E018	1E019	1E01A	1E01B	1E01C	1E01D	1E01E	1E01F
Ɱ̖	Ɱ̗	Ɱ̘	Ɱ̙	Ɱ̚	Ɱ̛	Ɱ̜	Ɱ̝	Ɱ̞	Ɱ̟	Ɱ̠	Ɱ̡	Ɱ̢	Ɱ̣	Ɱ̤	Ɱ̥
1E020	1E021	1E022	1E023	1E024	1E025	1E026	1E027	1E028	1E029	1E02A	1E02B	1E02C	1E02D	1E02E	1E02F
Ɱ̦	Ɱ̧	Ɱ̨	Ɱ̩	Ɱ̪	Ɱ̫	Ɱ̬	Ɱ̭	Ɱ̮	Ɱ̯	Ɱ̰	Ɱ̱	Ɱ̲	Ɱ̳	Ɱ̴	

LARGE characters: no superscripts

Cyrillic 4: Transliteration of Glagolitic

Glagolitic superscripts (non-spacing)

1E000	1E001	1E002	1E003	1E004	1E005	1E006	1E007	1E008	1E009	1E00A	1E00B	1E00C	1E00D	1E00E	1E00F
ⴐ	ⴑ	ⴒ	ⴓ	ⴔ	ⴕ	ⴖ		ⴘ	ⴙ	ⴚ	ⴛ	ⴜ	ⴝ	ⴞ	ⴟ
1E010	1E011	1E012	1E013	1E014	1E015	1E016	1E017	1E018	1E019	1E01A	1E01B	1E01C	1E01D	1E01E	1E01F
ⴠ	ⴡ	ⴢ	ⴣ	ⴤ	ⴥ	⴦	ⴧ	⴨			⴬	ⴭ	⴮	⴯	ⴰ
1E020	1E021	1E022	1E023	1E024	1E025	1E026	1E027	1E028	1E029	1E02A	1E02B	1E02C	1E02D	1E02E	1E02F
ⴲ	ⴳ		⴩	⴪		⴮	⴯	ⴰ	ⴱ	ⴲ					

Cyrillic superscripts (non-spacing)

2DE0	2DE1	2DE2	2DE3	2DE4	2DE5	2DE6	2DE7	2DE8	2DE9	2DEA	2DEB	2DEC	2DED	2DEE	2DEF
Б	В	Г	Д	Ж	З	К	Л	М	Н	О	П	Р	С	Т	Х
2DF0	2DF1	2DF2	2DF3	2DF4	2DF5	2DF6	2DF7	2DF8	2DF9	2DFA	2DFB	2DFC	2DFD	2DFE	2DFE
Ц	Ч	Ш	Щ	Ѡ	ѡ	Ѣ	ѣ	Ѥ	ѥ	Ѧ	ѧ	Ѩ	ѩ	Ѫ	ѫ

A674	A675	A676	A677	A678	A679	A67A	A67B
Ѭ	ѭ	Ѯ	ѯ	Ѱ	ѱ	Ѳ	ѳ

A69E	A69F
Ѵ	ѵ

Cyrillic 4: Transliteration of Glagolitic

Glagolitic

1E000	+
1E010	Ɔ
1E020	Ɑ

Ā ǃ ħ Ɔ Ǫ Ǫ
HĚ Ĭ Ě Ǫ MǪ ΓΛǪ ρ̂̂ Ě ∴

ħ Ɔ Ĥ
ĚMǪ Ǫ HĚ

Evangeliarum Assemani. Codex Assemani 3. slavicus glagoliticus. Ediderunt Josef Vajs & Josef Kurz. Tomus II. Edidit Josef Kurz. Praga 1955, 3, 89

Asemanovo ev.
Kurz, ed., p. 2

I. Schrift und Aussprache. 9

§ 6. Umschreibung des glagolitischen Alphabets durch das kyrillische. Bei Herausgabe glagolitischer Denkmäler in kyrillischer Umschrift werden wiedergegeben: Ɑ durch ѡ (ein neu erfundenes Zeichen), daher auch ⱭⱭ durch ѡѡ; Ɔ durch ѣ, daher auch ⱭƆ durch ѡѣ; Ǫ durch ѣ, daher auch ⱭǪ durch ѡѣ; Ǫ durch ѣ, daher auch ⱭǪ durch ѡѣ; Ǫ durch ѣ, ein aus der altserbisch-kyrillischen Schrift entlehntes, im altbulg. Alphabet nicht gebrauchtes Zeichen; Ǫ, in den beiden Bedeutungen (= kyrillisch *ě* und = kyrill. *ja á*), durch ѣ; im übrigen durch die Buchstaben, wie sie in den Alphabeten § 1

A. Leskien, Handbuch der altbulgarischen Sprache, Heidelberg 1969, 9.













Cyrillic su




2DE0	Б
2DF0	Ц
A674	Є

00F	Ѓ
01F	Ѓ
02F	
DEF	Ѧ
DEF	Ѧ
DEF	Ѧ

Cyrillic 4: Translit. of Glagolitic

Missing:   > cyr. І і

	→		×		→		×
	→		×	---	→		×
	→		×		→	??	×
	→		×	(phonetically, something like 'i-', or [ö] according to Trubetzkoy)			

* for certain Cyrillic orthographies, where  = [e] (instead of ) and  = [je]. Has no counterpart in Glagolitic.








actual use:

Transliteration of Glagolitic nasal vowels using Cyrillic letters

6. Кромѣ исчисленныхъ согласныхъ былъ въ ц. слав. еще одинъ если не чисто и согласный, то полусогласный звукъ, для

96

§ 8. СИСТЕМА СОГЛАСНЫХЪ.

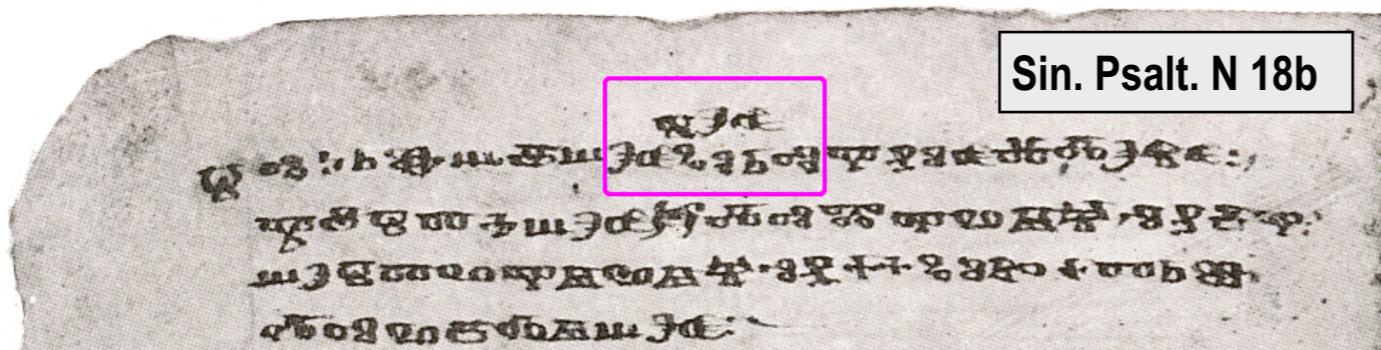
котораго ни кириллица, ни глаголица не представляютъ самостоятельнаго начертанія: это заднеязычный ^{дыхательный} . Св. Кириллъ обозначалъ его обыкновенно знакомъ , но лишь въ соединеніи іота съ послѣдующими гласными А, Е, Ъ, У, напр. , , . Предъ І іотъ не обозначался, если не считать его знакомъ первую палочку  осмеричнаго, разлагая его на +І. Выше упомянуто впрочемъ, что знакъ іотаціи имѣлъ предъ гласными и другое назначеніе—показывать мягкость предыдущаго согласнаго. Первое значеніе имѣеть іотація въ началѣ словъ и слоговъ послѣ зіянія, второе же—въ средозвучномъ и позвучномъ расположеніи іотованныхъ буквъ.

Budilovič, Načertanie cerkovnoslav. gramm., Warszawa 1883, 5-6

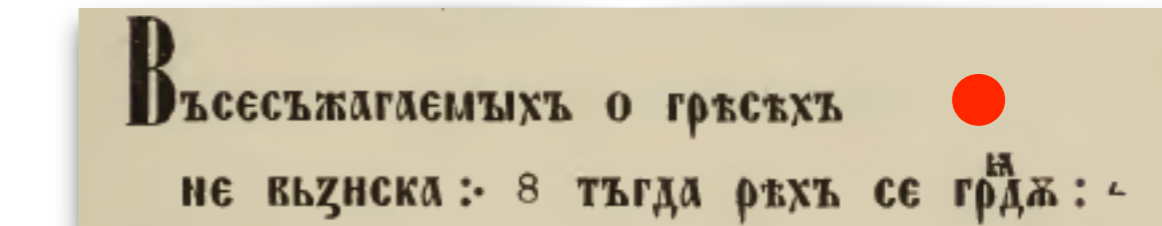
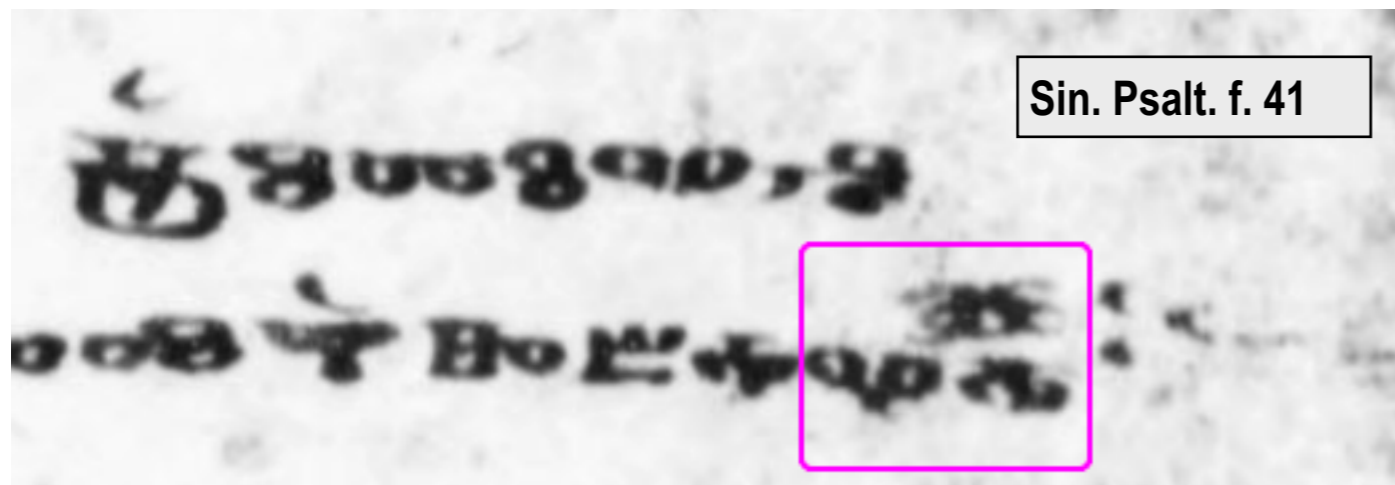
Cyrillic 4: Translit. of Glagolitic

Missing: Ѧ > cyr. ІА

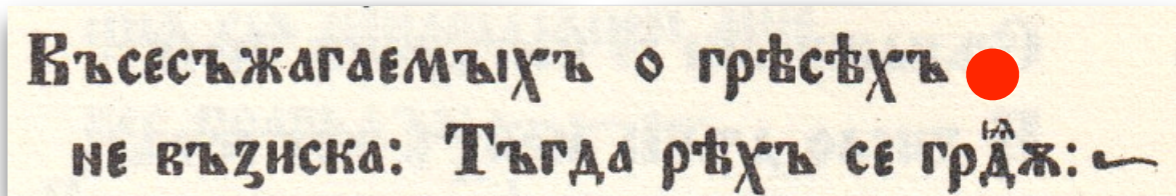
actual uses:



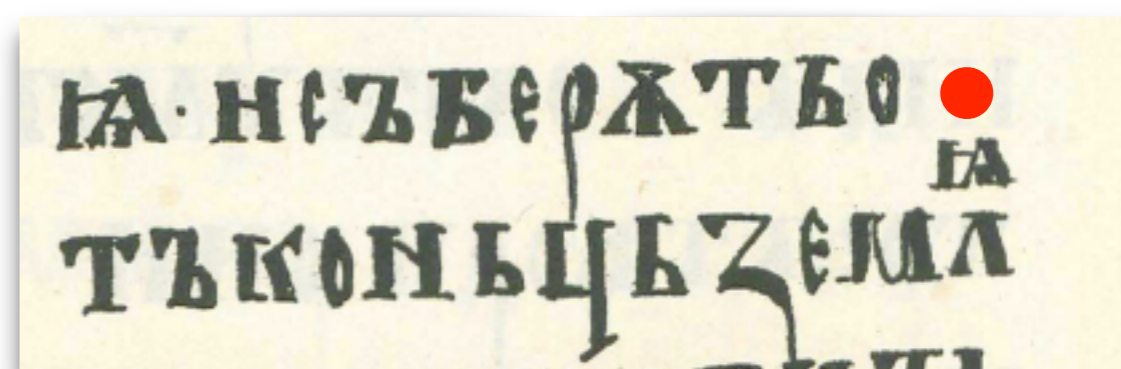
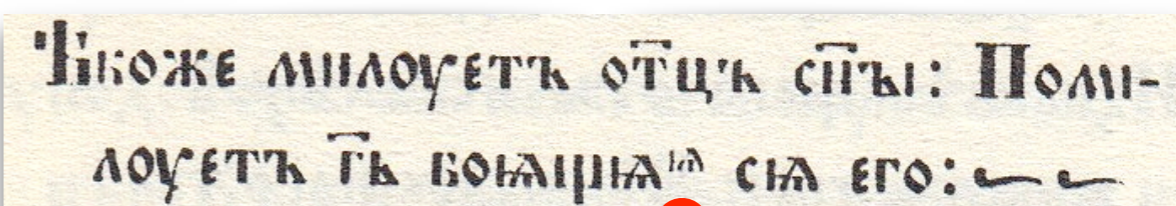
Glagolitic originals:



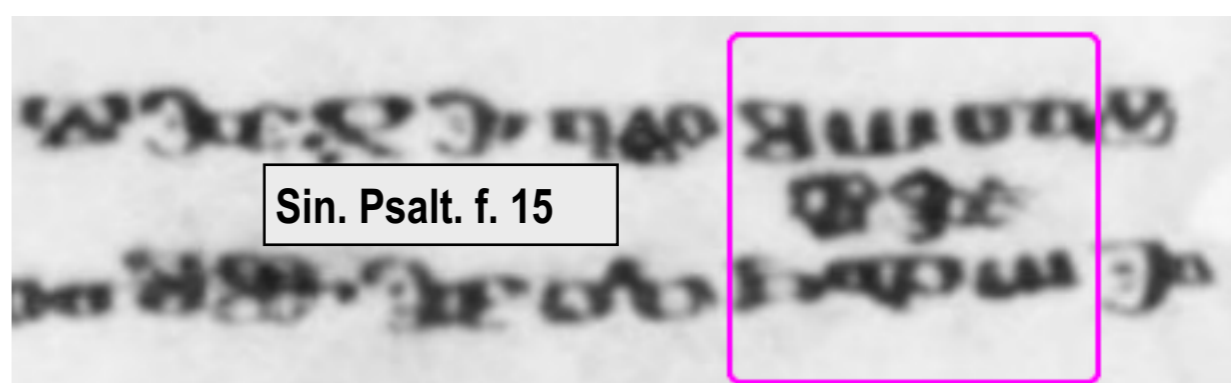
Psalterium Sinaiticum, ed. Geitler, Zagreb 1883, p. 85



Psalterium Sinaiticum, ed. Sever'janov, M. 1922, p. 52 / 132



Ostromir f. 241d-16 (noted by V. Baranov)



Cyrillic 4: Translit. of Glagolitic

Missing: ⱮⱮ, ⱮⱮ > cyr. ЪІ

ⱮⱮ	ⱮⱮ	—	ЪІ	—	у
----	----	---	----	---	---

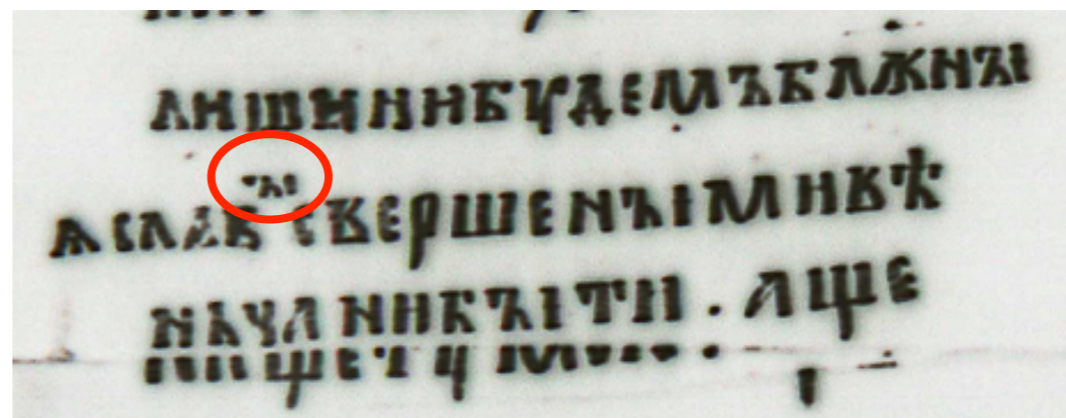
Horace Lunt, Old Church Slavonic Grammar 1959, 17

ЪІ	U+A650; CYRILLIC CAPITAL LETTER YERU WITH BACK YER
ъі	U+A651; CYRILLIC SMALL LETTER YERU WITH BACK YER
ы	U+A679; COMBINING CYRILLIC LETTER YERU

Only < ы > available as Cyrillic superscript; combination with back yer is missing, although the inline version is available

actual use:

spotted by Viktor Baranov

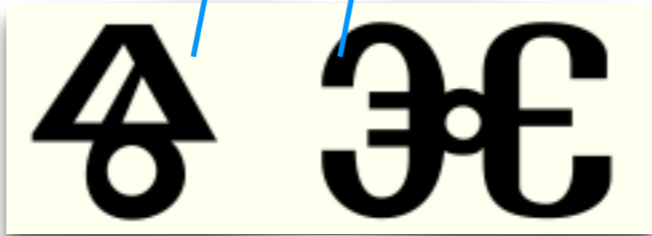


Cyrillic 4: Transliteration of Glagolitic

Glagolitic transliterated



1E000	1E001	1E002	1E003	1E004	1E005	1E006	1E007	1E008	1E009	1E00A	1E00B	1E00C	1E00D	1E00E	1E00F
Ǻ	Ǫ	ǫ	Ǧ	Ǻ	Ǽ	ǰ	ǰ̇	Ǫ	Ǫ̇	Ǫ̇*	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇
Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇
1E010	1E011	1E012	1E013	1E014	1E015	1E016	1E017	1E018	1E019	1E01A	1E01B	1E01C	1E01D	1E01E	1E01F
Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇
Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇
1E020	1E021	1E022	1E023	1E024	1E025	1E026	1E027	1E028	1E029	1E02A	1E02B	1E02C	1E02D	1E02E	1E02F
Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇	Ǫ̇



red dot: superscript not available in Cyrillic Extended-A (superscripts)

> 5 additional Cyrillic superscripts are needed: I ı ħ t EA

* without jota, transliteration needs to be U+A676, ⁱ

Glagolitic: Status Quo


two blocks: inline + superscripts


2C00	2C01	2C02	2C03	2C04	2C05	2C06	2C07	2C08	2C09	2C0A	2C0B	2C0C	2C0D	2C0E	2C0F
ⴐ	ⴑ	ⴒ	ⴓ	ⴔ	ⴕ	ⴖ	ⴗ	ⴘ	ⴙ	ⴚ	ⴛ	ⴜ	ⴝ	ⴞ	ⴟ
2C10	2C11	2C12	2C13	2C14	2C15	2C16	2C17	2C18	2C19	2C1A	2C1B	2C1C	2C1D	2C1E	2C1F
ⴠ	ⴡ	ⴢ	ⴣ	ⴤ	ⴥ	⴦	ⴧ	⴨	⴩	⴪	⴫	⴬	ⴭ	⴮	⴯
2C20	2C21	2C22	2C23	2C24	2C25	2C26	2C27	2C28	2C29	2C2A	2C2B	2C2C	2C2D	2C2E	2C2F
ⴰ	ⴱ	ⴲ	ⴳ	ⴴ	ⴵ	ⴶ	ⴷ	ⴸ	ⴹ	ⴺ	ⴻ	ⴼ	ⴽ	ⴾ	ⴿ
2C30	2C31	2C32	2C33	2C34	2C35	2C36	2C37	2C38	2C39	2C3A	2C3B	2C3C	2C3D	2C3E	2C3F
ⴐ	ⴑ	ⴒ	ⴓ	ⴔ	ⴕ	ⴖ	ⴗ	ⴘ	ⴙ	ⴚ	ⴛ	ⴜ	ⴝ	ⴞ	ⴟ
2C40	2C41	2C42	2C43	2C44	2C45	2C46	2C47	2C48	2C49	2C4A	2C4B	2C4C	2C4D	2C4E	2C4F
ⴠ	ⴡ	ⴢ	ⴣ	ⴤ	ⴥ	⴦	ⴧ	⴨	⴩	⴪	⴫	⴬	ⴭ	⴮	⴯
2C50	2C51	2C52	2C53	2C54	2C55	2C56	2C57	2C58	2C59	2C5A	2C5B	2C5C	2C5D	2C5E	2C5F
ⴰ	ⴱ	ⴲ	ⴳ	ⴴ	ⴵ	ⴶ	ⴷ	ⴸ	ⴹ	ⴺ	ⴻ	ⴼ	ⴽ	ⴾ	ⴿ

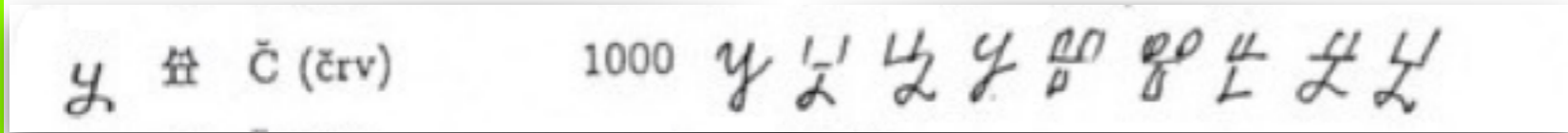
1E000	1E001	1E002	1E003	1E004	1E005	1E006	1E007	1E008	1E009	1E00A	1E00B	1E00C	1E00D	1E00E	1E00F
ⴐ	ⴑ	ⴒ	ⴓ	ⴔ	ⴕ	ⴖ		ⴘ	ⴙ	ⴚ	ⴛ	ⴜ	ⴝ	ⴞ	ⴟ
1E010	1E011	1E012	1E013	1E014	1E015	1E016	1E017	1E018	1E019	1E01A	1E01B	1E01C	1E01D	1E01E	1E01F
ⴠ	ⴡ	ⴢ	ⴣ	ⴤ	ⴥ	⴦	ⴧ	⴨			⴫	⴬	ⴭ	⴮	⴯
1E020	1E021	1E022	1E023	1E024	1E025	1E026	1E027	1E028	1E029	1E02A	1E02B	1E02C	1E02D	1E02E	1E02F
ⴰ	ⴱ		ⴳ	ⴴ		ⴶ	ⴷ	ⴸ	ⴹ	ⴺ					

Glagolitic 1: Update

Addition to Unicode v. 14 (2021)

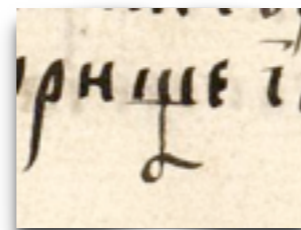
2C2F  **GLAGOLITIC CAPITAL LETTER CAUDATE CHRIVI**
This is used as an alternative to Glagolitic letter chrivi, combined with 0303 combining tilde or with 0483 combining Cyrillic titlo as an et cetera sign.

2C5F  **GLAGOLITIC SMALL LETTER CAUDATE CHRIVI**
(no superscript required)

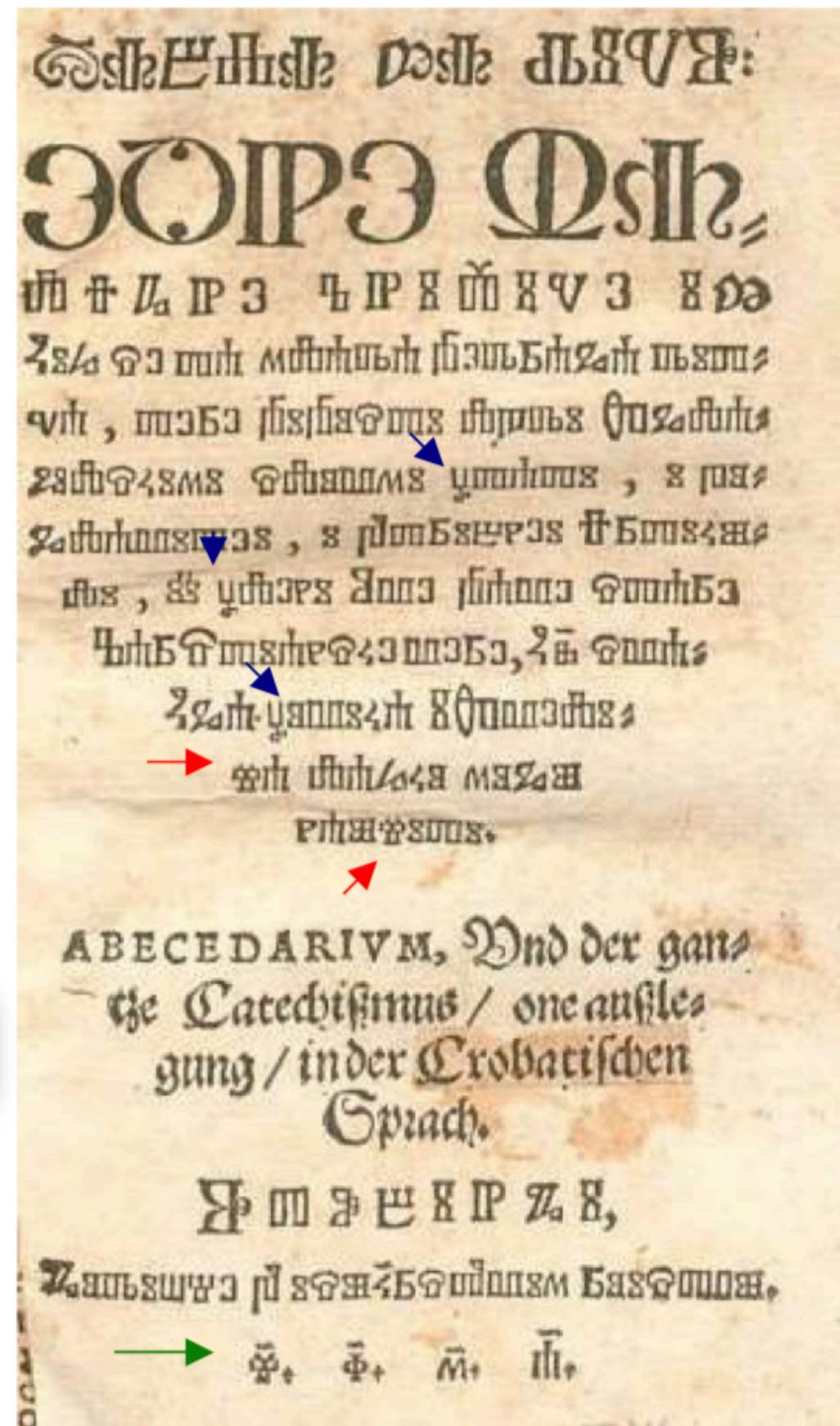


cursive glagoljica

Cyrillic




Zoran Vukojević 2019
Proposal L2/19-288




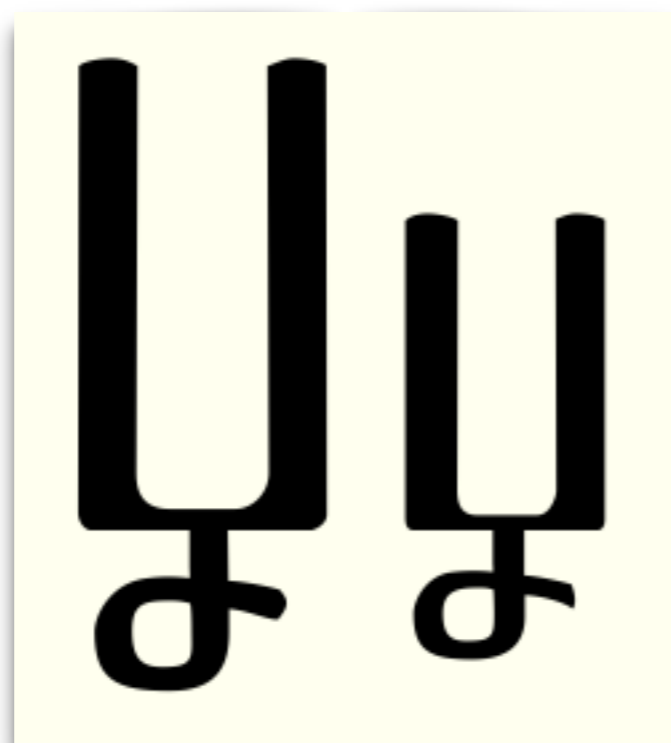
Trubar, Tabla za dicu, Tübingen 1561

Glagolitic 1: Update

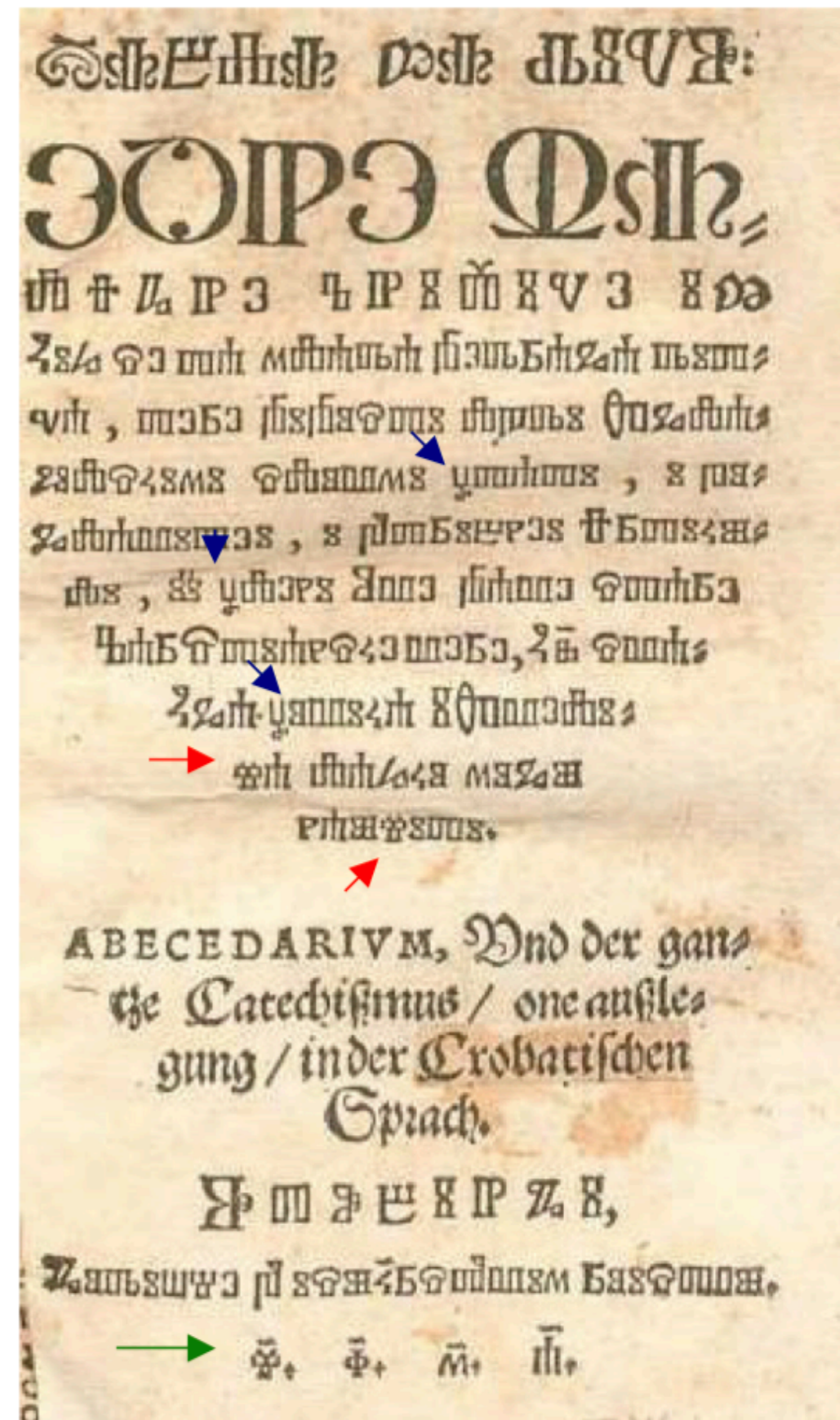
Addition to Unicode v. 14 (2021)

2C2F  GLAGOLITIC CAPITAL LETTER CAUDATE CHRIVI
This is used as an alternative to Glagolitic letter chrivi, combined with 0303 combining tilde or with 0483 combining Cyrillic titlo as an et cetera sign.

2C5F  GLAGOLITIC SMALL LETTER CAUDATE CHRIVI



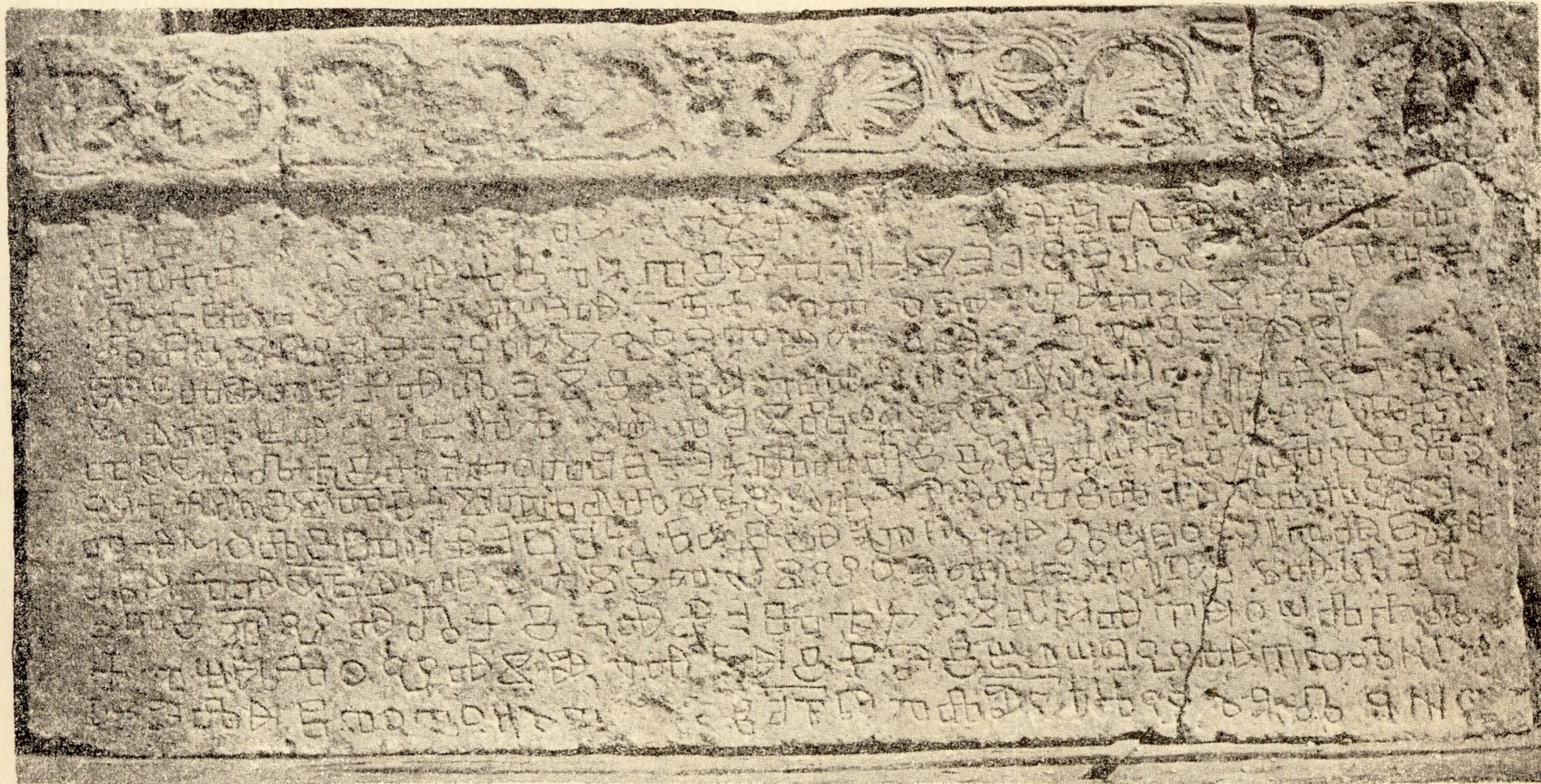
Zoran Vukojević 2019
Proposal L2/19-288



Trubar, Tabla za dicu, Tübingen 1561

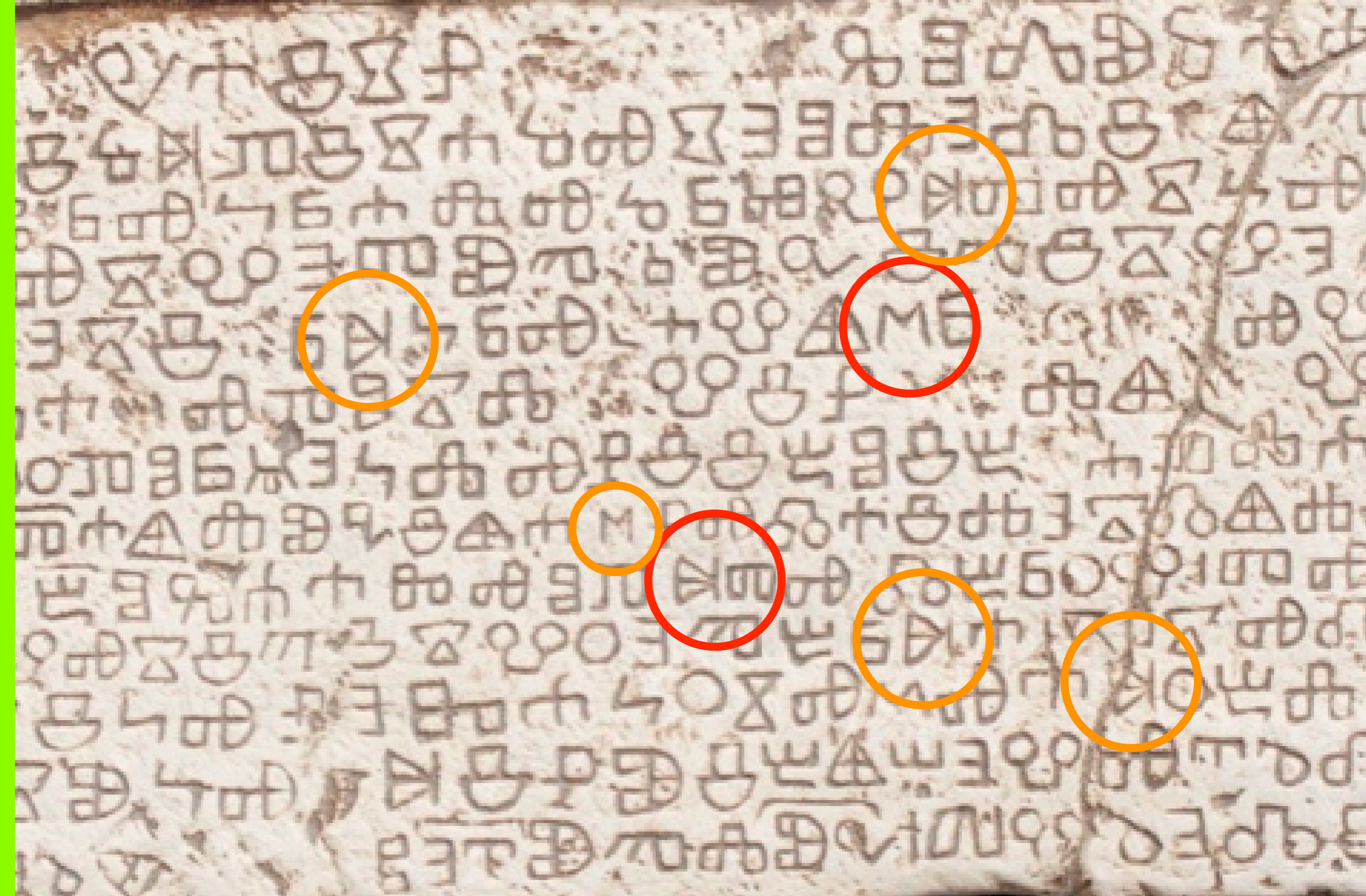
Glagolitic 2: More Additions?

Bašćanska ploča, Krk, ca. 1100



The original tablet

Encoded variants: M and A

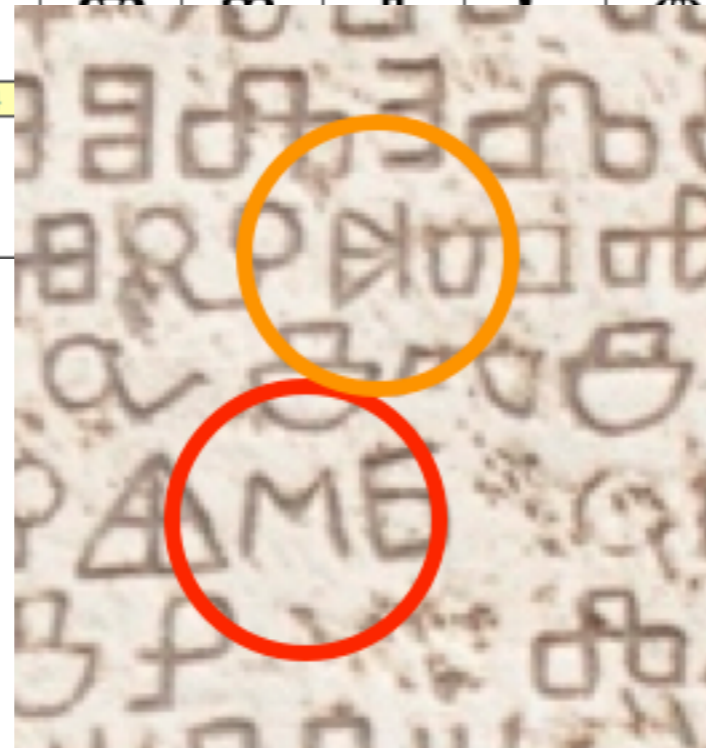


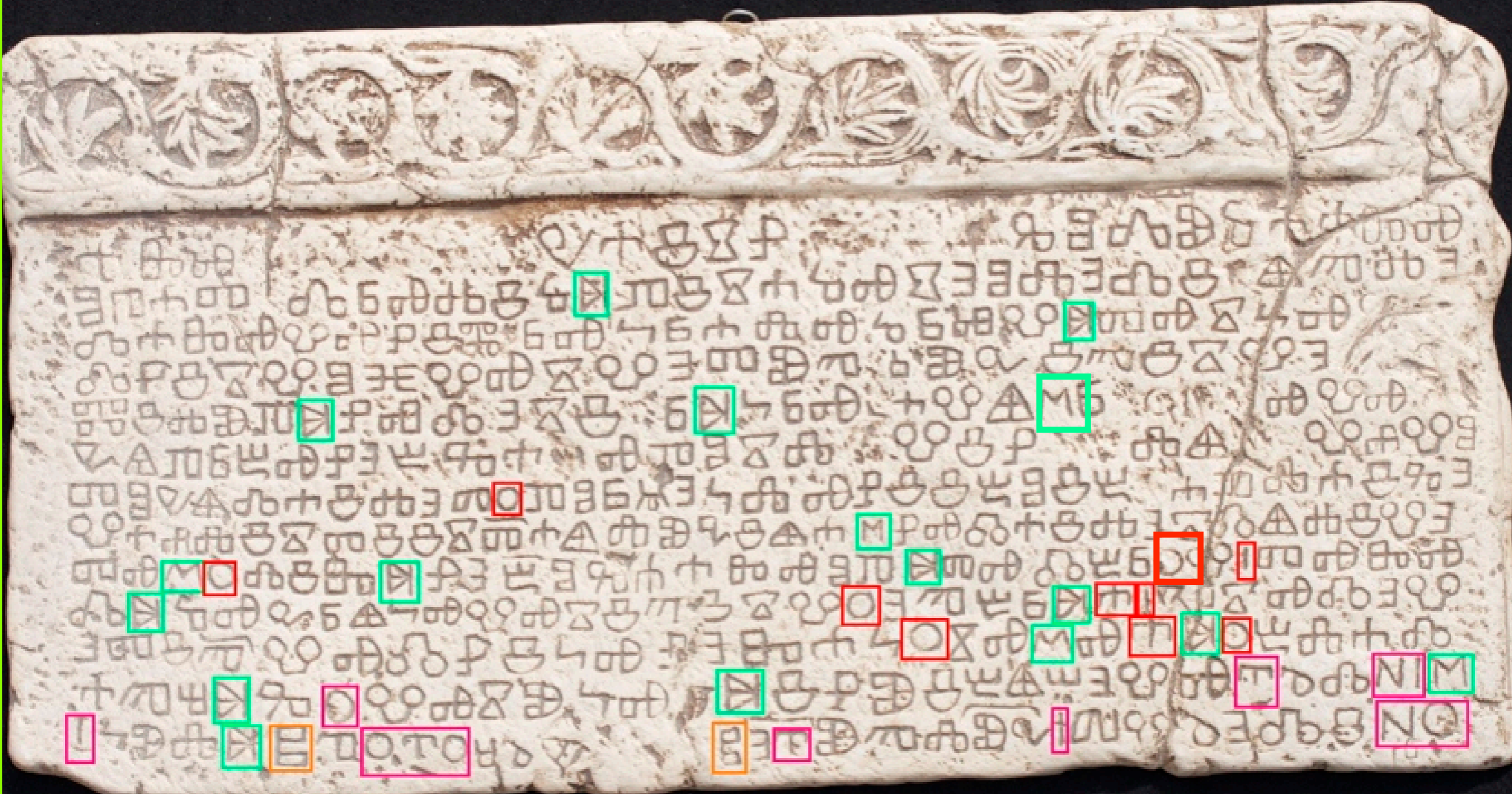
Encoded variants: M and A

Font - Bukyvede [/.../ BukyVede/ BukyVede-Regular.vfb]

Unicode 2C00 Glagolitic

2C00	2C01	2C02	2C03	2C04	2C05	2C06	2C07	2C08	2C09	2C0A	2C0B	2C0C	2C0D	2C0E	2C0F
Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ
2C10	2C11	2C12	2C13	2C14	2C15	2C16	2C17	2C18	2C19	2C1A	2C1B	2C1C	2C1D	2C1E	2C1F
Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ
2C20	2C21	2C22	2C23	2C24	2C25	2C26	2C27	2C28	2C29	2C2A	2C2B	2C2C	2C2D	2C2E	2C2F
Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ
2C30	2C31	2C32	2C33	2C34	2C35	2C36	2C37	2C38	2C39	2C3A	2C3B	2C3C	2C3D	2C3E	2C3F
Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ
2C40	2C41	2C42	2C43	2C44	2C45	2C46	2C47	2C48	2C49	2C4A	2C4B	2C4C	2C4D	2C4E	2C4F
Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ
2C50	2C51	2C52	2C53	2C54	2C55	2C56	2C57	2C58	2C59	2C5A	2C5B	2C5C	2C5D	2C5E	2C5F
Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ	Ɱ





green: in UC; red and orange: not in Unicode
Glagolica, Latinica, Kyrillica: I, N, O, T, B

Alphabet Table (D. Žubrinić)

Ɱ, Ɱ	A (az)	1	Ɱ, Ɱ	O (on)	80
Ɱ	B (buki)	2	Ɱ	P (pokoi)	90
Ɱ, Ɱ	V (vidi)	3	Ɱ	R (rci)	100
Ɱ	G (glagole)	4	Ɱ	S (slovo)	200
Ɱ	D (dobro)	5	Ɱ, Ɱ	T (trdo)	300
Ɱ	E ((j)est)	6	Ɱ	U (uk)	400
Ɱ	Ž (živite)	7	Ɱ	H (hir)	600
Ɱ	Z (zemla)	9	Ɱ	Št, Šć, Ć (šća)	800
Ɱ	Ā (iže)	10	Ɱ	C (ci)	900
Ɱ, Ɱ	I (i)	20	Ɱ, Ɱ	Č (črv)	1000
Ɱ	K (kako)	40	Ɱ	Š (ša)	
Ɱ	L (ljudi)	50	Ɱ, Ɱ	poluglasovi (jor, jer)	
Ɱ, Ɱ	M (mislite)	60	Ɱ	Ja, Je (jat)	
Ɱ, Ɱ	N (naš)	70	Ɱ	Ju (jus)	
			Ɱ	jest-je	



encoded



unencoded

Psalterium Sinaiticum

Handwritten text in Church Slavonic script, likely a title or introductory passage, written in a cursive style.

МѢСЯЦЪ

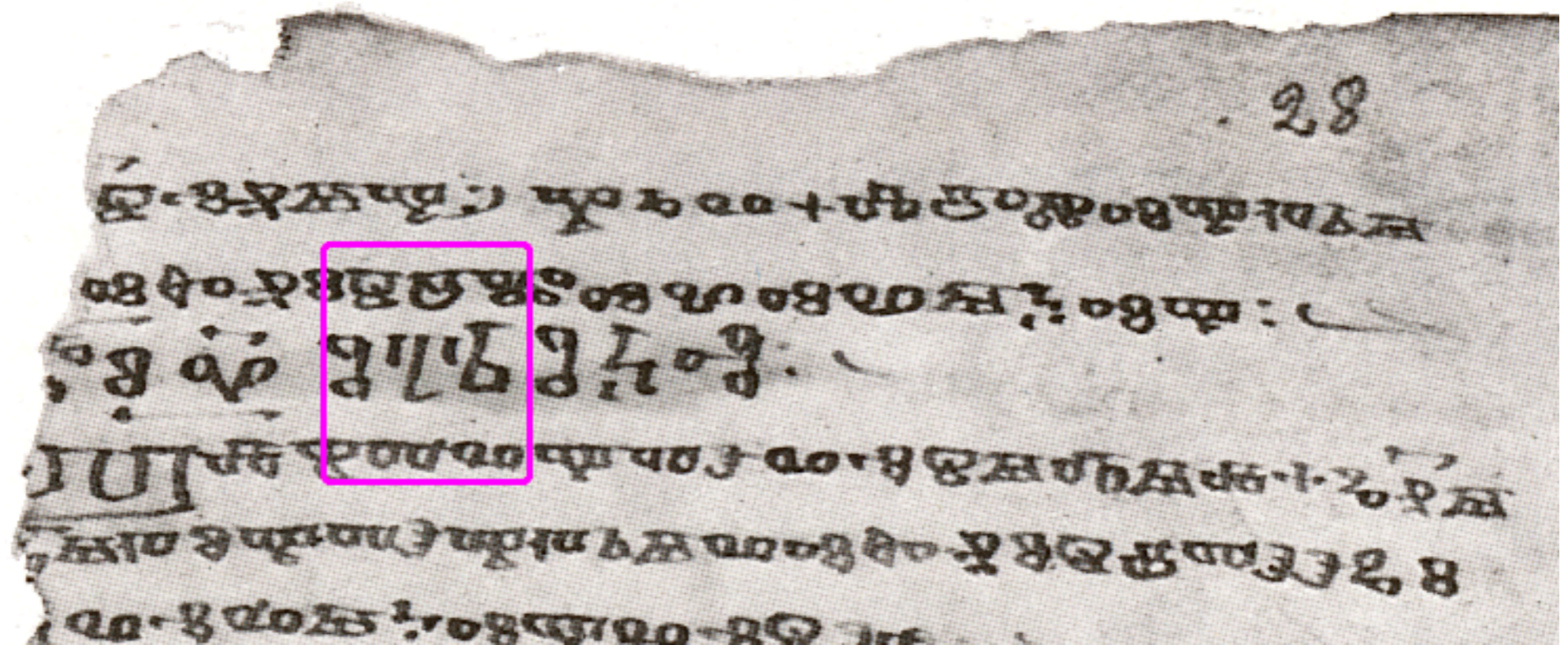
Handwritten text in Church Slavonic script, including a large decorative initial 'М' and several lines of text.

Handwritten text in Church Slavonic script, including a large decorative initial 'М' and several lines of text.

Geitler 1883:

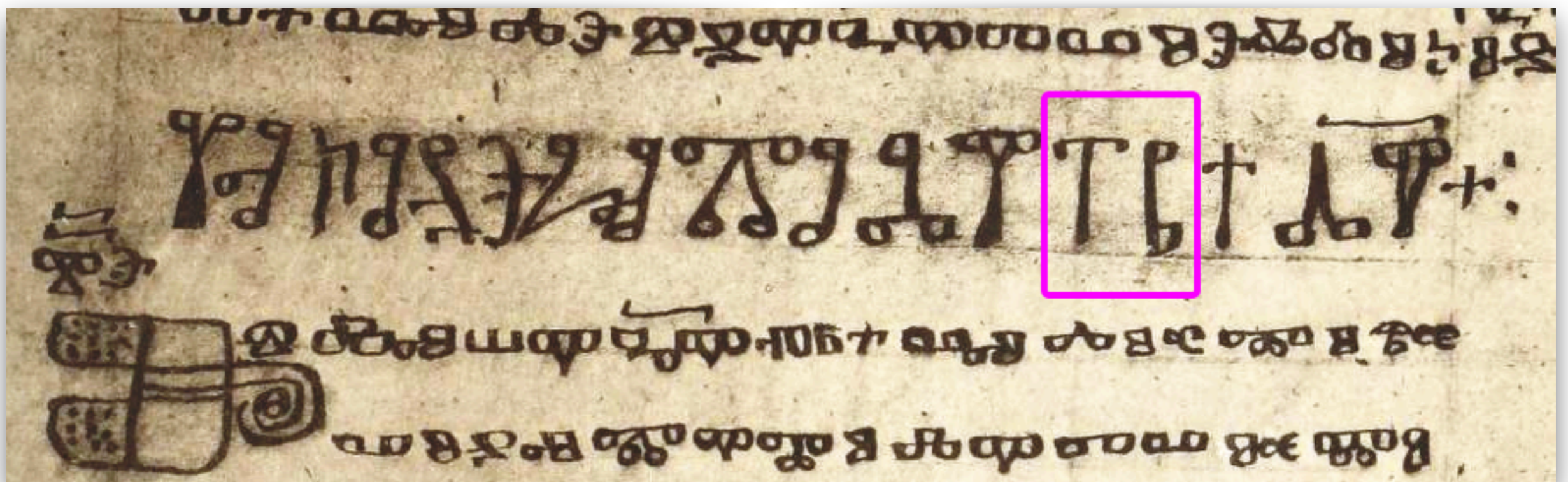
Psalterium sinaiticum. Das erste facsimile ist pag. 31 des originals: **сѣказалъ ми еси пакти**. Die zwei theile des **Ѣ** erscheinen noch getrennt geschrieben; in der aufschrift findet man die glagolitischen majuskeln **Т** und **В**. Das zweite, pag. 216: **твоего погубиѣхъ ежди рѣка твоѣ**. Das dritte, pag. 197: **подвижа ся і трепетѣна**. In der aufschrift die majuskel **М**.

Psalterium Sinaiticum, New Part



PsaltSin
N28a

Geitler 1883:



Geitler 1883, 184: Observations

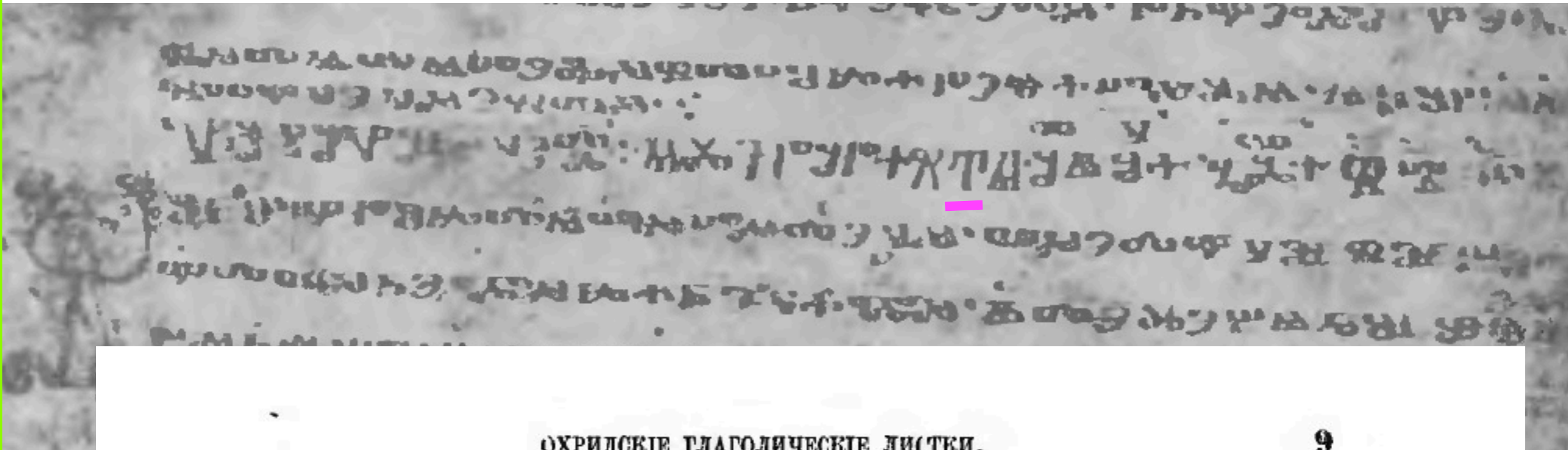
gleiche der ersten hand mit dieser zweiten, dass auch ziemlich bedeutende unterschiede des runden und eckigen typus zu gleicher zeit neben einander bestanden haben konnten. So ist die schriftart des psalters, insbesondere dessen zweite hand, ein uralter vorläufer der croatischen eckigen schrift. Auch in den speciellen gestalten seiner buchstaben sind keime verstreut, welche (allerdings erst in viel späterer zeit) weiter entwickelt, den charakter der eigentlichen croatischen schrift mitbestimmen. Obwohl wir dieselben im einzelnen schon behandelten, wollen wir, um ein übersichtliches bild zu gewinnen, hier an dieselben erinnern. Die grossgezeichneten aufschriften der psalmen zeigen die alten majuskeln der glagolica **M T B** (vgl. §. 157), höchst seltene buchstaben, welche sich auf der croatischen inschrift von Baška wiederfinden, **M** auf altbulgarischem boden nur im psalter, **B** nur auf jener inschrift und im psalter. Einigemale begegnen wir nun im psalter dem versuche, das grosse **M** in die kleine textschrift herabzuziehen, d. h. zur minuskel zu machen, so dass unser codex an einigen stellen zwei *m*-zeichen, **M** und **Ꝟ** besitzt, wie die ältere croatische glagolica. Ob diese schreibart in irgend einer bulgarischen schreiberschule zum vollen durchbruche gelangt ist, wissen wir nicht aus mangel an zeugnissen; aber schon in dem der bulgarischen glagolica so nahe stehenden und dennoch der serbo-croatischen literatur zuzuschreibenden fragment Mihanović herrscht dieses **M** über das sonst gewöhnliche **Ꝟ** (die eigentliche alte minuskel für *m*) und wird endlich alleiniges *m*-zeichen der Croaten. Wir hiel-

wird endlich alleiniges *m*-zeichen der Croaten. Wir hielten dies zeichen bis jetzt für croatisch, sein ursprung liegt im altbulgarischen psalter. Das eigenthümlich gestaltete croat. *k* fanden wir zuerst in den aufschriften des psalters, §. 140. Ein gewisses *b*, das sich auf der croatischen inschrift von Veglia findet, ist der zweiten hand des psalters eigen, §. 120. Das charakteristische einhörnige **Ꝟ** der prager fragmente, die der croatischen schrift so nahe stehen, muss auch auf croatischem boden bekannt gewesen sein (vgl. §. 113); im psalter ist es möglich, das werden dieses *ž*, das langsame zusammenfallen beider hörner in einen strich, bei verschiedenen händen zu verfolgen. Jene eigenthümliche abart des **Ꝟ**, die aus sechs quadraten besteht, §. 121, findet sich auf der inschrift von Baška, in spätcroatischen ligaturen und

Mnogobrojni nadpisi nad psalmi pisani su velikom stasitom glagolicom (majuskulom), s kakovom se upoznasmo već u izdanju glag. cloz., gdje Kopitar na snimku nekolicke redaka ovoga pisma navodi. U ovih nadpisih dolaze češće ona gore navedena **T M B**.

Što se tiče načina transkripcije, držao sam se istih već poprimljenih pravila, kao što u trebniku. Slova **T M B** imadu oblik u obće velikoga latinskoga i grčkoga t, m, v; da jih u transkripciji iztaknem, izrazio sam jih kursivnim latinskim *T, M, V*.

Evangelium Achridanum (XI. c.)



ОХРИДСКІЕ ГЛАГОЛИЧЕСКІЕ ЛИСТКИ.

9

ВЪ НОВѢ НѢА.: ЪЖЕ ПО ПАС

ТѢ О ИОА ГЛА Е І В :: С
Н

XX

10 19 **Θ**μηρὸν ποζδѣ въ тѣ денѣ. въ
единѣ свѣот[ѣ]
і двѣремѣ заренамѣ. ѿдеже бѣ
хж оучи^{иц} с[ѣ]
вѣрані. за страхѣ іудѣискѣ.
πολαε ἱς[ѣ]

19 Οὔσης οὖν ὀψίας τῆ ἡμέρα
ἐκείνη τῆ μιᾶ σαββάτων, καὶ
τῶν θυρῶν κεκλεισμένων ὅπου
ἦσαν οἱ μαθηταὶ [συνηγμένοι],
διὰ τὸν φόβον τῶν Ἰουδαίων,
ἦλθεν ὁ Ἰησοῦς καὶ ἔστη

Croatian Wikipedia

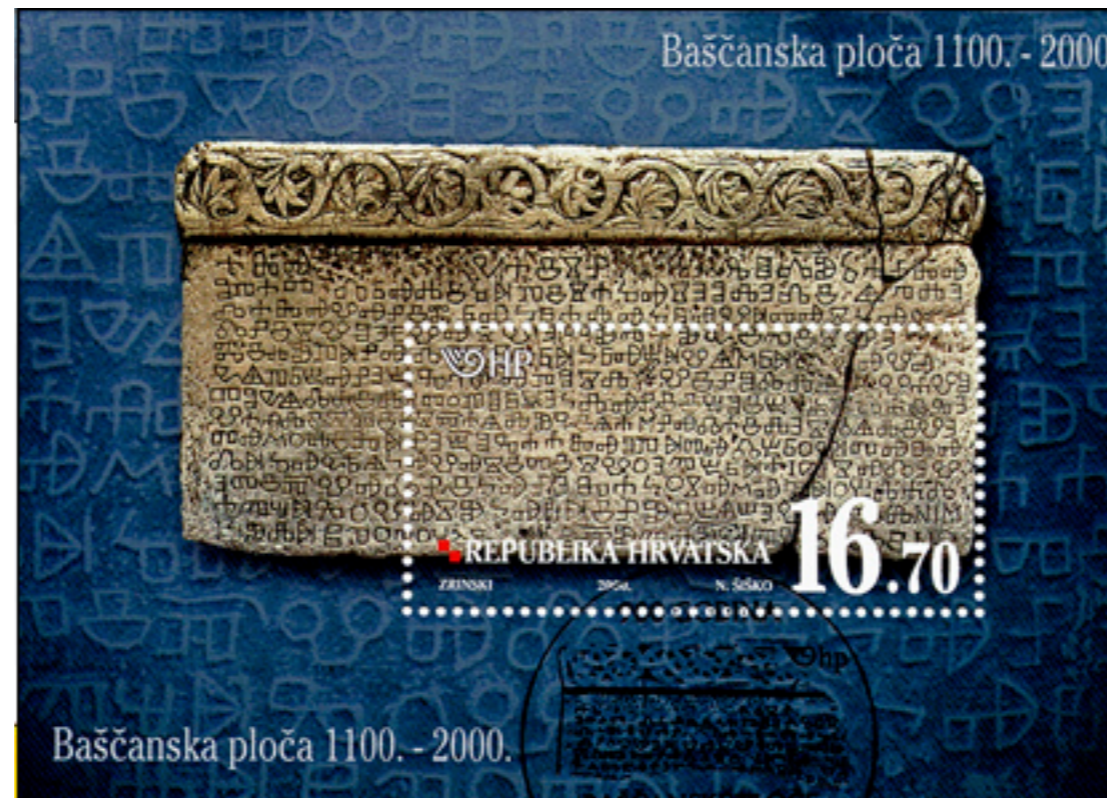
[https://hr.wikipedia.org/wiki/Baščanska_ploča](https://hr.wikipedia.org/wiki/Ba%C5%A1%C4%8Danska_plo%C4%87a)

"Usporedno s glagoljičkim slovima, javljaju se i neka **latinična** i **ćirilična** (*I, M, N, O, T, V*), a isti je slučaj i s drugim hrvatskim spomenicima pisanih glagoljicom iz **11. i 12. stoljeća.** "

Freq.:

M: 5

A: 12



Freq.:

O: 7

T: 5

I: 5

N: 2

B: 1

> five additional Glagolitic characters are needed: I, N, O, T, B

Glagolitic 3: "Modern Glagoljica"

used as a font variation of
Latin

VAŠE IME
na
GLAGOLJICI

h ɛ w ʏ ʉ ɔ ʋ
ɛ ɛ ɛ ɛ ɛ ɛ ɛ
ɛ ɛ ɛ ɛ ɛ ɛ ɛ
ɛ ɛ ɛ ɛ ɛ ɛ ɛ
ɛ ɛ ɛ ɛ ɛ ɛ ɛ



Glagolitic 3: "Modern Glagoljica"

a) one character split up into two

standard: variants

Ѣ, ѣ Ѣ, J (je) 30

Da bi se izbjegla dvoznačnost, stručno vijeće Društva prijatelja glagoljice (ѢѢѢ) odlučilo je (na prijedlog Vladimira Čepulića) da se za potrebe pisanja *suvremenim* hrvatskim jezikom rabe dva oblika đerva (ѣ, Ѣ):

a. glagoljsko slovo ѣ (đerv) za j,

b. đerv Ѣ za đ.

Na taj način riječ ѢѢѢѢ možemo napisati i kao ѢѢѢѢѢѢ. Ipak, običaj je da se Ѣ (jat) rabi kad god je to moguće (vidi detaljniji opis niže).

Glagolitic 3: "Modern Glagoljica"

b) new signs invented for X, Y, W, Q

Figure 7 New signs in Glagolitic script for X, Y, W, Q

Xylophon

Oxygen

Washington

Windows

Quality

Square

> further development needs to be observed

Non-Slavic Languages of the USSR

Ossetian

Andr. Joh. Sjögren, Ossetische Sprachlehre, SPb. 1844

ИРОН ÆБЗАГАХЪР

DAS IST

OSSETISCHE SPRACHLEHRE,

NEBST KURZEM

OSSETISCH-DEUTSCHEN UND DEUTSCH-OSSETISCHEN

WÖRTERBUCHEN,

VON

Dr. Andr. Joh. Sjögren,

Akademiker und Mitgliede mehrerer gelehrten Gesellschaften.

ST. PETERSBURG

GEDRUCKT BEI DER KAISERLICHEN AKADEMIE DER WISSENSCHAFTEN.

1844.

ZU HABEN BEI W. GRÄFF'S ERBEN U. LEIPZIG BEI L. VOSS.

Non-Slavic Languages of the USSR

Ossetian

Cyrillic!

Andr. Joh. Sjögren,
Ossetische Sprachlehre,
SPb. 1844

				Name					Name
А	а	А	а	a	Л	л	Л	л	el
Б	б	Б	б	be	М	м	М	м	em
В	в	В	в	we	Н	н	Н	н	en
Һ	һ	Һ	һ	ha	Ҟ	Ҟ	Ҟ	Ҟ	eng
Г	г	Г	г	ge	О	о	О	о	o
Ҝ	Ҝ	Ҝ	Ҝ	egh	Ӗ	Ӗ	Ӗ	Ӗ	ö
Ґ	ґ	Ґ	ґ	gje	П	п	П	п	pe
Д	д	Д	д	de	Пҕ	пҕ	Пҕ	пҕ	eph
Д̇	д̇	Д̇	д̇	dje	Р	р	Р	р	er
Æ	æ	Æ	æ	ä	С	с	С	с	ess
Е	e	Е	e	e	Ш	ш	Ш	ш	scha
É	é	É	é	é	Т	т	Т	т	te
З	з	З	з	æ	ҕ	ҕ	ҕ	ҕ	eth
Ж	ж	Ж	ж	she	Ҙ	Ҙ	Ҙ	Ҙ	tje
Д,	д,	Д,	д,	eds	Ω	ω	Ω	ω	ω
ДҞ	дҞ	ДҞ	дҞ	edsh	У	у	У	у	u
І	i	І	i	i	У	у	У	у	ü
Ј	j	Ј	j	je	Ф	ф	Ф	ф	ef
К	к	К	к	ka	Х	х	Х	х	cha
Ҝ	Ҝ	Ҝ	Ҝ	ekh	Ц	ц	Ц	ц	ze
К̇	к̇	К̇	к̇	kje	Ч	ч	Ч	ч	tsche
Ҝ̇	Ҝ̇	Ҝ̇	Ҝ̇	khje	Ц	ц	Ц	ц	etz
Q	q	Q	q	qa	Ч	ч	Ч	ч	ettsch

Non-Slavic Languages of the USSR

Ossetian

.ДЖ.

ДЖав D. n. pr. des Dorfes Dshaw im südlichen Ossetien.

ДЖивісхán D. ein Schmeichelwort von unbekanntem Ursprunge. S. Gram.

§. 138.

ДЖивіт D. n. gent. = Т. Дут.

ДЖивіттаг D. n. a. = D. Дуттаг.

ДЖіккó D. n. s. = Т. Діккú.

ДЖіңá D. n. s. = Т. Дvңá.

ДЖуаргін D. n. a. und s. = Т. Дуардін.

ДЖуаре D. n. s. Kreuz, Kirche. Vgl. Дуар.

"CYRILLIC CAPITAL LETTER DZZHE"

rather a blending of Д + Ж than Д + К?

available for Komi/Ossetian orthography:

052A	052B
ДЖ	ДЖ

Non-Slavic Languages of the USSR

Ossetian

h.

hа int. des Hinzeigens da (hast du)!

haj haj adv. ja ja! freilich! allerdings!

hајдá hајдá adv. u. int. weg! hinweg! fort! aus dem Türk. هاجدي (hajde od. hajda).

hæj int. des Verdrusses und des Rufens ach! ei!

hej int. des Rufens eí!

hej hej int. des Forttreibens fort! hinweg!

hó D. adv. ja! ja wol! nach dem Grusin. هوه (ho).

available for Old Abkhasian orthography:

A68A	A68B
Ტ	Ტ

Mayanist additions:

A726	A727
Ტ	Ტ

available for Azerbaijani orthography:

0526	0527
Һ	Һ

Һ Һ

Non-Slavic Languages of the USSR

Ossetian

ӕ.

ӕа int. des Hinzeigens da (hast du)!

ӕаӕ **ӕаӕ** adv. ja ja! freilich! allerdings!

ӕаӕдӕ ӕаӕдӕ adv. u. int. weg! hinweg! fort! а
od. hajda).

ӕӕ int. des Verdrusses und des Rufens ach!

ӕеј int. des Rufens eí!

ӕеј **ӕеј** int. des Forttreibens fort! hinweg!

ӕó D. adv. ja! ja wol! nach dem Grusin. ӕó

Алфавит Народов Севера Печатные формы букв

Aa	Bb	Cc	€€	Dd	Зз
Зз	Ee	Əə	Ææ	Ff	Gg
Hh	ӕӕ	Ii	Ьь	Jj	Kk
Ll	Łł	Mm	Nn	Ŋŋ	Oo
Өө	Pp	Qq	Rr	ƦƦ	Ss
Šš	Tt	Uu	Vv	Ww	Xx
Yy	Zz	ƷƷ	Latin!		

ӕӕ

A726	A727
ӕ	ӕ

"Unified Northern Alphabet", Latinisation in the Soviet Union; in use 1932 – 1937.

https://en.wikipedia.org/wiki/Unified_Northern_Alphabet

Russian Wikipedia:
Mayanist additions (Latin):

Non-Slavic Languages of the USSR

Ossetian

ӕ.

“There was a proposal in 2019 asking for some Komi Latin letters ... As a result of the discussion on this document, the UTC decided that for historical, transitional (oft failed) orthographies, it is ok to have mixing of scripts. This is to prevent ... a wholesale set of Cyrillic letters duplicated into Latin blocks, especially when the evidence is weak (and/or the letters were used for transitional orthographies, as noted). This is not to say that a case could be made to add certain letters based on evidence.”

Deborah Wanders, director, UTC, personal communication

Алфавит Народов Севера

Печатные формы букв

Аа	Вв	Сс	Єє	Dd	Ӗӗ
ӖӖ	Ee	Әә	Ææ	Ff	Gg
Hh	ӕӕ	Ii	Ьь	Jj	Kk
Ll	Łł	Mm	Nn	Ŋŋ	Oo
Өө	Pp	Qq	Rr	ƦƦ	Ss
ƦƦ	Tt	Uu	Vv	Ww	Xx
Yy	Zz	ƷƷ	Latin!		

"Unified Northern Alphabet", Latinisation in the Soviet Union; in use 1932 – 1937.

https://en.wikipedia.org/wiki/Unified_Northern_Alphabet

Russian Wikipedia uses:
Mayanist additions (Latin):

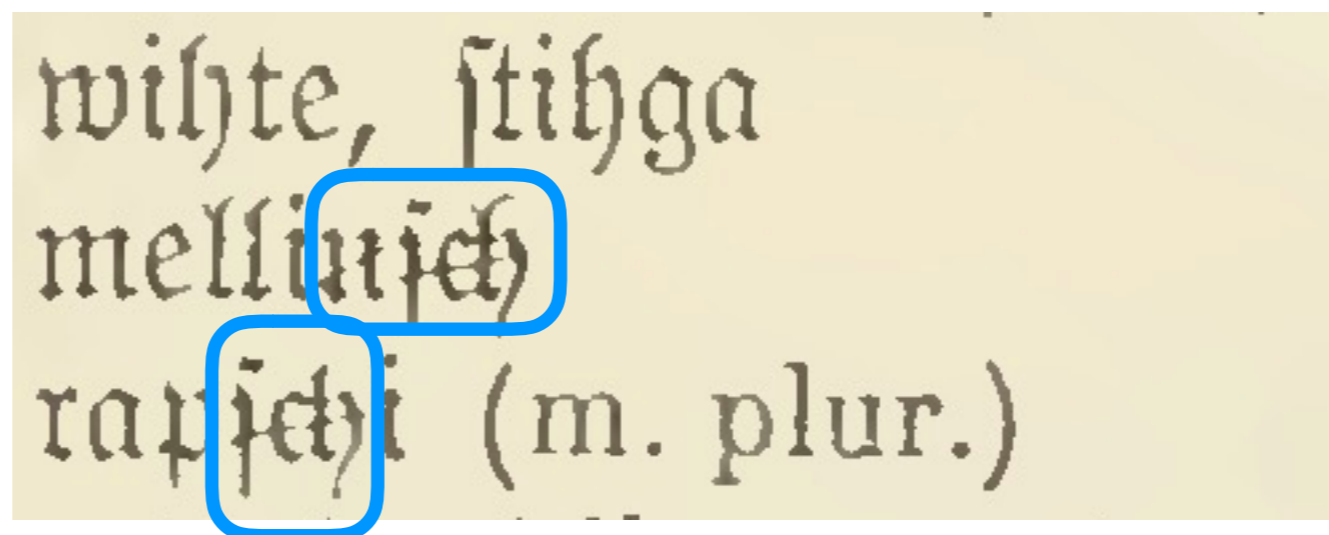
A726	A727
ӕ	ӕ

Varia

Latvian stroked characters

Ņ fch

red = missing



> two additional stroked characters are needed: c, h

A7A0	A7A1	A7A2	A7A3	A7A4	A7A5	A7A6	A7A7	A7A8	A7A9	A7AA	A7AB	A7AC	A7AD	A7AE	A7AF
Ġ	g	Ķ	ķ	Ņ	ņ	Ŗ	ŗ	Š	š	Ĥ	3	Ģ	Ł	Ṭ	Q
A7B0	A7B1	A7B2	A7B3	A7B4	A7B5	A7B6	A7B7	A7B8	A7B9	A7BA	A7BB	A7BC	A7BD	A7BE	A7BF
Ḳ	Ḷ	Ḷ	Ḳ	Ḷ	Ḷ	Ḷ	Ḷ	Ḷ	Ḷ						

Varia

Montenegrin Cyrillic additions (transitional 2009-2017)

The Montenegrin standard is still emerging. Its [orthography](#) was established on 10 July 2009 with the addition of two letters to the [alphabet](#). Their usage remained controversial and they achieved only limited public acceptance, along with some proposed alternative spellings.^[9] They had been used for official documents since 2009, but in February 2017, the [Assembly of Montenegro](#) removed them from any type of governmental documentation.^[10]

Linguistic considerations [\[edit \]](#)



Shtokavian subdialects in Montenegro.

Montenegrins speak [Shtokavian](#) subdialects of [Serbo-Croatian](#), some which are shared with the neighbouring Slavic nations:

- [Eastern Herzegovinian dialect](#) (in the west and northwest).
- [Zeta-Raška dialect](#) (spoken in the rest of the country).

Montenegrin alphabet [\[edit \]](#)

Main article: [Montenegrin alphabet](#)

The proponents of the separate Montenegrin language prefer using the [Latin alphabet](#) over the [Cyrillic alphabet](#). In both alphabets there are two additional letters (bold), which are easier to render in [digital typography](#) in the Latin alphabet due to their existence in [Polish](#), but which must be created *ad hoc* using [combining characters](#) when using Cyrillic.

Latin collation order

Latin	A	B	C	Č	Ć	D	Dž	Đ	E	F	G	H	I	J	K	L	Lj	M	N	Nj	O	P	R	S	Š	Ś	T	U	V	Z	Ž	Ǻ
Cyrillic	А	Б	Ц	Ч	Ћ	Д	Џ	Ђ	Е	Ф	Г	Х	И	Ј	К	Л	Љ	М	Н	Њ	О	П	Р	С	Ш	Ѓ	Ћ	У	В	З	Х	Ѕ

Cyrillic collation order

Cyrillic	А	Б	В	Г	Д	Ђ	Е	Ж	З	Ѕ	И	Ј	К	Л	Љ	М	Н	Њ	О	П	Р	С	Ѓ	Ћ	Т	Ћ	У	Ф	Х	Ц	Ч	Џ	Ш
Latin	A	B	V	G	D	Đ	E	Ž	Z	Ǻ	I	J	K	L	Lj	M	N	Nj	O	P	R	S	Ś	T	Ć	U	F	H	C	Č	Dž	Š	

Varia

y with ogonek
ž as modifier (sometimes even with acute above)
(Polish historical phonetics)

W części Śląska powstały nowe samogłoski nosowe z końcowych połączeń samogłoski ustnej ze spółgłoską nosową: zamiast *tam, śpiewám, tym babom, wiem, z dobrym chłopem, z nim* słyszy się *tą, śpyvu, ty babo, vų, z dobry xłope, ś nį.*

Na czeskim pograniczu Śląska, na odwrót, utrzymuje się *vým, z dobrym wyjasnie bez wzoru nauczyciela*.

Średniopolski rozwój $r^{\acute{z}}$ polega na powolnej redukcji elementu drżącego r na rzecz coraz pełniej rozwijającego się elementu szczelinowego $ž$, co ostatecznie doprowadzi w polszczyźnie XVIII w. do powszechnego artykulacyjnego utożsamienia dawnego $r^{\acute{z}}$ z $ž$. Wyjątkowo jest ono zaświadczone już w XV w. (1410); spotykamy je u Reja; udawane mazurzenie w *Peregrynacji Maćkowej* z r. 1612: *zec (= rzec), bzuch, dobrze* dowodzą go w początkach w. XVII. Przypuszczalny rozwój przedstawia się tedy następująco: $*r || *rj \Rightarrow r \Rightarrow r^{\acute{z}} \Rightarrow rž \Rightarrow ž$.

etc. etc. etc.

OK



Schmochtitz, Lusatia



This presentation was given during the conference
«EI'Manuscript 2021, Textual Heritage and Information Technologies»,
University of Freiburg, Germany, 12.-15.4.2021

© Prof. Dr. Sebastian Kempgen 2021

ORCID: 0000-002-2534-9423

D-96045 Bamberg, University of Bamberg, Germany

sebastian.kempgen@uni-bamberg.de

<https://www.uni-bamberg.de/slauling/personal/prof-em-dr-sebastian-kempgen/>

