

## Synodal Church Slavonic Collation: Some Issues and Concerns

The traditional order of letters is given as

ѦѦ, ББ, ВВ, ГГ, ДД, ЕЕ, ЄЄ, ЖЖ, СС, ЗЗ, ИИ, Її, КК, ЛЛ, ММ, Оо, ОО, ПП, Рр, Гг,  
 ТТ, О҃҃, У҃҃, ХХ, ФФ, Ѡѡ, Ѣѣ, Цц, Чч, Шш, Щщ, Хх, Ыы, Ьь, Ъѣ, Юю, ІѦѦ,  
 ѦѦ, Ѣѣ, Ѡѡ, Ѣѣ, Ѡѡ, Ѣѣ, Ѡѡ, Ѣѣ

This listing contains all the basic letters which occur in the Synodal Church Slavonic language. In addition to the above letters, the diacritical marks shown in Table 3, Table 4, and Table 5. It can be noted that not all of the diacritical marks have the same meaning or use through time.

Table 1: Church Slavonic Diacritics

Diacritic Name	Proper Code-Point	Auxiliary Code-Point	Example
Acute Accent	U+ 0301		Ѧ́ Ѧ́
Grave Accent	U+0300		Ѧ̀ Ѧ̀
Titlo	U+0483	U+0303	Ѧ̑ Ѧ̑
Psili	U+0486	U+0313	Ѧ̑ Ѧ̑
Yerok	U+033E		Ѧ̑
Kamora	U+0311		Ѧ̑ Ѧ̑
Diaeresis	U+308		Ѧ̈ Ѧ̈
Breve	U+0306		Ѧ̆ Ѧ̆

Table 2: Composite Diacritics

Composite Diacritics	Proper Sequence	Auxiliary Sequence	Example
Apostrof	U+0486 U+0300	U+0313 U+0300	Ѧ̑ Ѧ̑
Iso	U+0486 U+0301	U+0313 U+0301	Ѧ̑ Ѧ̑

Table 3: Most Common Superscripted Letters. Forms in Brackets are Unattested in Synodal Church Slavonic.

Superscripted Letter	Code-Point	Superscripted Letter	Code-Point
( <sup>6</sup> А)	U+2DE0	( <sup>7</sup> А)	U+2DEE
<sup>н</sup> А	U+2DE1	( <sup>х</sup> А)	U+2DEF
<sup>с</sup> А	U+2DE2	( <sup>т</sup> А)	U+2DF0
<sup>а</sup> А	U+2DE3	<sup>у</sup> А	U+2DF1
( <sup>х</sup> А)	U+2DE4	( <sup>и</sup> А)	U+2DF2
( <sup>с</sup> А)	U+2DE5	( <sup>т</sup> А)	U+2DF3
( <sup>к</sup> А)	U+2DE6	( <sup>ѣ</sup> А)	U+2DF4
<sup>т</sup> А	U+2DE7	( <sup>а</sup> А)	U+2DF6
( <sup>и</sup> А)	U+2DE8	( <sup>е</sup> А)	U+2DF7
<sup>н</sup> А	U+2DE9	( <sup>ѣ</sup> А)	U+2DF9
<sup>о</sup> А	U+2DEA	( <sup>к</sup> А)	U+2DFA
( <sup>т</sup> А)	U+2DEB	( <sup>т</sup> А)	U+2DFB
<sup>п</sup> А	U+2DEC	( <sup>ѣ</sup> А)	U+2DFD
<sup>ѣ</sup> А	U+2DED		

### Proposed Collation Order

Since Unicode provides us with 4 levels of detail for specifying the order in which the different aspects of a proper collection sequence can be defined. Basically, I would suggest placing the basic characters to Level 1; the diacritical marks and superscript letters to Level 2; and capitalisation issues to Level 3. The following will provide greater detail and examples to illustrate the given suggestions.

It should be noted that since the diacritical marks are often used as a form of disambiguation, the following rules are proposed for ordering “similar” examples (All examples drawn from the Elizabethan Bible in the Ponomar Project or appropriate Liturgical texts):

- 1) The nominative singular form should be sorted first ( $\chi\acute{\alpha}\tau\alpha < \chi\hat{\alpha}\tau\alpha$ ).
- 2) The singular forms should come before the plural/dual forms ( $\chi\alpha\tau\grave{\alpha} < \chi\alpha\tau\hat{\alpha}$ ;  $\kappa\acute{\omicron}\nu\eta < \kappa\acute{\omega}\nu\eta\alpha$ ).
- 3) Abbreviations should come before unabbreviated words with the same Level 1 forms ( $\Delta\acute{\kappa}\lambda < \Delta\kappa\lambda$ ).
- 4) Sorting of diacritical marks should be from left-to-right.

### Level 1

In order to implement the above suggestions, the following Level 1 sort order should be defined:

$\text{Αα} < \text{Ββ} < \text{ββ} < \text{Γγ} < \text{Δδ} < \text{Εε} < \text{εε} < \text{Ζζ} < \text{Σσ} < \text{Ξξ} < \text{Ηη} [\text{Ἠῆ} = \text{η} + \text{ῆ}] < \text{Ιι} < \text{Κκ} < \text{Λλ} < \text{Μμ} < \text{Οο} < \text{οο} < \text{Ωω} < \text{Ϟϙ} [\text{Ϟϙ} = \text{ω} + \text{ϙ}; \text{Ϟ} = \text{ω} + \text{ο} + \text{ϙ}] < \text{Ππ} < \text{Ρρ} < \text{Ϻζ} < \text{Ττ} < \text{Ϝϝ} = \text{Ϟϙ} = \text{γ} < \text{Χχ} < \text{Φφ} < \text{Ψψ} < \text{Ϡϡ} < \text{Ϣϣ} < \text{Ϥϥ} < \text{Ϧϧ} < \text{Ϩϩ} < \text{ϫκ} < \text{Ϭλ} < \text{ϭμ} < \text{Ϯν} < \text{ϯθ} < \text{ϰϱ} < \text{ϲρ} < \text{ϳσ} < \text{ϴτ} < \text{ϵε} < \text{϶ζ} < \text{Ϸξ} < \text{ϸχ} < \text{Ϲζ} < \text{Ϻζ} < \text{ϻγ} < \text{ϼλ} < \text{Ͻν} [\text{ῶῷ} = \text{ν} + \text{ῶ}] < \text{Ͽκ}$

Comments:

- 1) The base form of Ἠῆ is encoded, namely Ιι. This will simplify issues with diacritical marks.
- 2) The different forms of /o/ are arranged so rules 1 and 2 can be obeyed, namely  $\kappa\acute{\omicron}\nu\eta < \kappa\acute{\omega}\nu\eta\alpha$  and  $\omicron\eta\alpha < \acute{\omega}\eta\alpha$ .
- 3) Equating the different forms of /u/ will not give any issues as the first form is only used word initially while the second form is used elsewhere. The last form is only used in numerals. In older texts, I would suggest that if there is a need for ordering that the second form be given the highest priority and the digraph the lowest.
- 4) The ordering of the two forms of /ja/ ( $\text{Ἰἶα} < \text{Ἰἷα}$ ) is arbitrary, but traditional. There is a good set of words which are disambiguated by using these two forms:  $\text{Ἰἶζῆκκ} < \text{Ἰἷζῆκκ}$ .



- 4) The superscript letters should come last and they should be ordered in the same manner as the base letter.
- 5) Consider also, for example,  $\overset{\check{z}}{\text{z}}\text{eta}$  (lake) <  $\overset{\check{z}}{\text{z}}\text{eta}$  (pond/standing water<sup>1</sup>)?

### Level 3

Level 3 should consider the case of the letter with upper case coming before the equivalent lower case letter.

---

<sup>1</sup> All seven instances of this word in the Elizabethan Bible are translated by pond or standing water in corresponding verse in the King James Bible.