Publishing Croatian Church Slavonic Manuscripts: Talk Outline

Andrew R. Corin

AATSEEL 1991

I.   Introduction

A.   Background

1. This talk will be more programmatic than a report on research results, though I will present some of the latter, including a text-based interpretation of the enigmatic 3 Sg. Aorist form *bisi*.

2. Its motivation: encourage greater recognition of the various forms of information in Church Slavonic manuscripts that can be easily encoded into computerized transcripts.

a. Computer technology is increasing the *feasibility of capturing and exploiting more of the information contained in manuscripts* than before, losing less in the process of transcription and publication.

b. The experience of entering data with personal computers leads to the realization that there is *more information worth being captured* than is generally recognized. While we have always been aware of this information, the effort required to record and systematically exploit much of it may have detracted us from seriously considering its potential. The emerging ease of transcribing more kinds of information presents an opportunity to reexamine assumptions concerning the value of certain types of information.

c. Lastly, the feasibility and desirability of capturing more detailed information creates the need for dissemination in forms which exploit the new capabilities.

4. So, these will be the two motifs of my talk:

- the benefits of *encoding more of the information contained in manuscripts* into our transcriptions
- the need for *forms of publication that exploit the new possibilities*: electronic publication and "comparative corpora"

B. Organization of the Talk

1. First, I will survey types of information typically ignored or lost during transcription, and the consequences of their omission. I address this question within the context of Croatian Church Slavonic (CCS), partially because that is my current field of work, partially due to the sense that American Slavists could benefit by becoming more familiar with CCS textual traditions. Analogous problems surely present themselves with other Church Slavonic redactions.

2. Then, I will review the way in which manuscript text is typically published, focusing on two areas:

a. reconciling conflicting needs regarding the level of detail in published text.

b. strategies for overcoming the striking paucity of published texts, as well as the limitations imposed by standard formats and media for text publication. I will present both a critique and some suggestions.

C. The goals of the presentation

- to stimulate discussion of the topic
- to influence the demands of consumers of published Church Slavonic text

II. Types of Information We Can Capture (and Why We Might Want to)

A. Word division

1. Writing most proclitics and enclitics together with the accentogenic word is a standardized characteristic of CCS orthography (Handout, example 1).

2. Deviation from this practice can be indicative of scribal signature.

3. Deviation might be indicative of (change in) the linguistic/orthographic status of particular unstressed words (e.g., *ot*).

B. Line endings

These are sometimes (not always) indicated in critical editions, but should always be indicated.

1. Hypercorrect štapić ("ι") or apostrophe for *a* as a space-saver can reflect date or scribal signature (Handout, example 2; Corin, *The New York Missal*, 1991, Table 7).

2. "Jor" ("ъ") is used in a few texts apparently as a space-filler, possibly reflecting

- scribal signature
- archaism

3. Any tendency to end lines in a vowel??? (*not* a regular CCS feature)

- noted in some OCS and Serbian texts
- may reflect scribal signature

C. Distinct Forms of Abbreviation

1. Ligatures

See the Handout, example 7, summarizing ligature usage in OT and NT lections from the Gospel reading for the fifth Tuesday through the Gospel reading for the fifth Friday of Lent. The only lacuna during this segment is probably the first three lines in the Novi Missal.

A. The inventory of ligatures employed in a codex shows considerable variation.

B. The number (tokens) of ligatures used or regularity of employment, overall or for any letter combination, also shows considerable variation.

C. Patterns of ligature use may correlate with

- date (cf. Ill4 vs. other mss.)
- scribal signature (cf. variation among NYM's scribes, especially Hand G)

- something "unusual" about a text (cf. the understandably small number of letter combinations in the 1483 *editio princeps*; the extraordinary number of letter combinations overall in NYM, likely reflecting at least in part Hand G's eccentricity; see my recent volume on the NYM)

2. Superscription

See the Handout, example 4.

A.   *ěk̃* , other scribes *ě̃k* , for *ěko*

s̃h , other scribes *sh̃* , for *stihι*

B. Patterns of superscription suggest the solution to the origin of CCS *bisi* for *bistι …*

3. Compression

Compression is a common manner of abbreviation through superscription found in some Cyrillic textual traditions (Handout, example 4c):

<*rě̌č*> for *reče*, <*slᵛo*> for *slovo*, <*vᵈoju*> for *vodoju*, <*posrᵈě*>   for *posrědě*)

Compression is not used in mature CCS, but must have been known in early CCS, since:

bistι is abbreviated as <*bis̃*>, but also as <*bˢi*> (Cf. NYM 62a1).

It is therefore likely that:

<*bis̃*> (abb. of *bistι*) altered to <*bˢi*> by analogy to <*rě̌č*> (compression of *reče*)

<*bˢi*> was reinterpreted as compression of *bisi* by analogy to <*rě̌č*> as compression of *reče*

III. Strategies for Publication and Dissemination of Text

The expanding possibilities for electronic publication discussed above suggest that hard copy dissemination of Church Slavonic manuscript and early printed text is no longer a sufficient approach to meeting research needs. Two issues arise concerning the way forward: the level of detail to be recorded, and strategies for selecting and disseminating transcribed text.

A. Level/Forms of Detail Recorded in Transcribed Text

Attitudes will differ regarding the desirability of reproducing certain types of information that can now be easily encoded. Mostly, this concerns orthographic and/or paleographic details.

1. Users Whose Primary Concern is *Content*

Literary or liturgical historians, others interested primarily in textual *content*, even the occasional philologist who just wants to compare texts to those of some other recension, require an *easily readable* text. They'll likely want abbreviations resolved, word division modernized, and minimal diacritic distraction (for ligatures, line endings, tildes, accents [in Serbian texts], etc.). Such a readable version is

also useful for scholars knowledgeable in, say, Russian Church Slavonic, but who initially feel uncertain when confronted with CCS texts. These users may prefer something like Version 1 (Handout, Example 5). Readability is also paramount when making artefacts of great cultural significance available to the public in a realistic, yet approachable, manner. This purpose is evident, for example, in the luxurious edition of the Hrvoje Missal.

2. Users Requiring Linguistic, Orthographic, and Paleographic Detail

Linguists and philologists benefit when "extra-textual" attributes are preserved. They may prefer to see abbreviations as in the original (ligatures, superscription, contractions, suspensions, compression), original line endings, marginal notations, punctuation marks, tildes, variant forms of tildes and other letters (e.g., dual forms of superscript "t"), etc.

1. Researchers interested in language, orthography, and filiations may prefer something like Version 2 (Handout, Example 5), which represents the text with only the actual Glagolitic letter forms stripped away, but retaining quantifiable information.

2. Researchers interested in computerized analysis may prefer something like Version 3, which encodes, in an analyzable format:

- beginning and end of ligatures
- place of *titla* (tilde)
- word breaks at line end
- superscription indicated by an asterisk

We could additionally encode boundaries between accentogenic words and the clitics written together with them.

3. Paths toward Resolving the Conflict

1. Publish hard copy in version 1 (or 1 and 2), with versions 2 and 3 (or just 3) available on diskette.

2. The result: a book accompanied by a diskette.

4. Cost-Benefit

a. The risk: will hard copy publishing of manuscripts be undermined if consumers can produce their own hard copy from software?

b. One solution: sell hard copy and on-line versions together.

c. The markets for hard-copy publication, especially of high resolution color facsimiles and "research-optimized" transcripts may be sufficiently distinct that no danger exists. If so, a small electronic publishing firm to meet the Slavic linguistic/philological community's need for low-cost dissemination of machine-readable text could be of immense benefit. The firm could start small, with an editor and research assistant utilizing one Macintosh and one PC, operating with just Microsoft Word and Excel. The firm could develop gradually as the stock of machine-readable text and demand grow.

5. Bottom Line

Regardless of the choice, the boon to research now within reach is the ability to widely disseminate *analyzable (i.e., machine-readable) text* that consumers can manipulate with some level of automization. It is hardly optimal, given the technological possibilities, for an editorial team to prepare a text for publication on computer, and then publish text and apparatus solely in hard copy, with the result that researchers wishing to analyze data contained in the text must excerpt it anew. Surely it is preferable to publish in both hard copy and diskette, thus accommodating the needs of all consumers of published text. Electronic publication of even a single version can address the needs of multiple audiences to some extent. Certain information can be input as "hidden text" which can be made visible if desired. Of course, this can create some dilemmas, as examination of Example 5, Versions 1 and 3 in the Handout will show.

B. Overcoming the Dearth of Published CCS Text

Saying the obvious No. 1: the research community benefits from published texts.

- Not all researchers can access all original manuscripts all the time.
- Transcripts and transliterations are necessary for efficient work.

Saying the obvious No. 2: relatively few Church Slavonic codices or text corpora have been published, especially in digitized analyzable form, and relatively few are likely to be, at least in hard copy, in our lifetime. Some of the publication that does occur, moreover, is strategized to meet needs other than those of textual scholars (see above).

1. Two approaches to textual publication

The advent of personal computing changes the playing field for responding to this challenge. To understand this, we should distinguish two forms of text publication to address research needs.

a. The value of publishing *complete individual codices* remains unchanged.

b. The area in which portable personal computers can have a major impact is *cross-sectional publication*, which presents parallel text and/or individual variants from multiple witnesses and/or recensions. This enables readers to contextualize codices within their broader textual environment and to understand the development of that environment, including filiations, linguistic and orthographic features, approaches to copying and translation, etc., across entire manuscript traditions. Cross-sectional data is usually presented in an atomized fashion—i.e., a critical apparatus—that has limited usefulness to the broader research community in carrying out follow-up analysis or other cross-sectional research projects on the same textual corpus used to prepare the apparatus. The advent of personal computing enables us to accelerate and stimulate research by addressing this limitation.

2. Available Strategies

a. Critical editions

Critical editions are the traditional—and currently still the most common—approach to publishing CCS (or other) text in a form intended to highlight cross-sectional data. Critical editions present a single base text according to one of two principles:

- The base text is a *reconstructed* original, canonical, or authoritative text. Variants illustrate deviant development.
- The base text is a *witness* considered most archetypical, venerable, or representative. Variants contextualize that witness within the broader textual tradition. This is the general practice in Church Slavonic critical editions.

Alongside critical publication of full manuscripts, another common approach involves critical editions of individual texts (say, lectionary readings from a particular Biblical book), presenting one base text and variant readings from other witnesses, accompanied by textual, linguistic, and paleographic analysis.

i. Advantages

- Critical editions can provide variant readings from a broad range of witnesses.
- Enabling examination of a full codex *in extenso* is a prerequisite (though not sufficient) for recognizing discontinuities within the base text. HM and NYM provide telling examples; each shifts from Recension A to Recension B in at least at one location.

ii. Limitations

- Variants are limited to witnesses the editor(s) choose to excerpt, which may not include those significant for all research purposes.
- Readers depend on the editor(s) to include variants significant for their own research purposes.
- Atomized presentation of variants: readers cannot get a sense of how the connected text reads in other witnesses.
- If there are multiple recensions or canonical forms, each authoritative in a different time or place, this will not emerge as clearly as it would from connected text of multiple witnesses.
- If the base text, though culturally significant for esthetic or historical reasons, is atypical in some way, readers can gain a skewed impression of the broader context. Hrvoje's missal (currently the only missal manuscript with a modern critical edition) presents two examples:
    - HM is consciously vernacularized in a manner unlike any other missal of the 14th or 15th century (see the brief sample in the Handout, example 6). Impressions of the language of the CCS missals based on HM will be skewed. This would "leap out" at readers presented with connected synoptic (side-by-side parallel) text from multiple witnesses (Handout, example 6).
    - HM shifts at least once from recension A to recension B. Were texts from multiple locations in HM presented synoptically with parallel text from missals of both recensions, this crucial fact would "leap out."

b. Synoptic Single-Text Publication

An infrequent variant of the critical edition is the synoptic publication of base texts representative of multiple recensions of a single text (e.g., Hamm's presentation of Job (1953), with other variant readings in the critical apparatus. This approach provides additional valuable, albeit still filtered, cross-sectional information on the textual tradition.

i. Advantages

- This approach is relatively simple to apply.
- It highlights one major stemmatological aspect of the text's history.

- Publication of numerous such studies can gradually fill out the developmental arc and cross-sectional context of the textual tradition.

ii. Limitations

- This approach still presents other aspects of variation in atomized form (critical apparatus), without cross-sectional text *in extenso* to enable others to confirm conclusions, expand upon them, or explore the textual evidence for other research purposes. There is still the need for *belief* that the editor has noted all variation, and from all witnesses, the reader might consider significant.
- This approach can convey the impression that witnesses chosen as base texts for each recension are more archetypical or representative than is the case, or of the discreteness of the recensional dichotomy when viewed across the full range of manuscripts. Circumstances may be more complex than a simple dichotomy. In my experience, editors typically do their best to mitigate this danger by examining the degree to which their chosen texts are indeed representative or archetypical.
- Readers still cannot gain a full appreciation of how the connected text of "non-base" witnesses reads in its text-internal context.
- Most important for us here, this approach is still adapted to the constraints of hard-copy publishing, and does not leverage the potential of computerized excerpting.

c. Bottom Line Regarding Critical Texts

As valuable and necessary as critical text publications are, they do not alone fulfill our current potential to meet CCS researchers' need for cross-sectional, contextualizing data.

d. Comparative Corpora

Comparative corpora (Handout, example 6) are a form of synoptic publication that has become feasible with the rise of portable personal computers. Comparative corpora present excerpts of parallel text (e.g., lections from one Biblical book) from as many witnesses as possible of the CCS missal, breviary, or other textual tradition. My doctoral dissertation and recent published volume on the NYM would have been impossible without prior compilation of a comparative corpus from manuscripts of the CCS *Missale Plenum*.

i. Advantages

- Connected text from each witness allows researchers to appreciate variation within its native documentary context.
- Comparative corpora provide full comparable parallel data from multiple witnesses, rather than an editor's selection.
- Recensional differences "leap out" in a way critical editions do not allow.
- Dissemination "force-multiplies" one researcher's effort in excerpting text, enabling others to analyze the same data, including relatively unstudied codices which will likely turn to dust before they are published in full.
- Dissemination allows distant researchers to gain, even through superficial examination, their own substantive, albeit incomplete, appreciation of the textual tradition, including the state of the text,

language, and orthography longitudinally within it. The expanded access to synoptic connected text also assists readers of published analyses of those texts.

- Multiple comparative corpora can be produced for the same effort as one full-codex publication.

ii. Limitations

- Printing requires many pages for relatively short passages. For on-line use, this is not a problem.
- Comparative corpora are useful primarily to linguists and philologists, less so to literary or cultural historians preferring "aggregation" of textual variation.
- Inconsistencies or discontinuities elsewhere within a witness will not emerge. This limitation also exists to some extent within traditional text publication strategies.
- Excerpting may be more time-consuming than excerpting individual variants. In my experience compiling one extensive comparative corpus from 14 CCS plenary missal manuscripts and the *editio princips* (lections—OT and NT—from the two final weeks of Lent) and a smaller one from 14 manuscripts plus one printed edition of the breviary (lections of 1 Samuel and Acts of the Apostles), this was not overly burdensome.
- Comparative corpora are less felicitous for encompassing extensive textual traditions such as the Vulgate Bible.
- More intellectual property issues can arise in disseminating a comparative corpus than a single manuscript.

IV. Conclusion: A Comprehensive Approach Should Include Comparative Corpora

All three approaches described here are needed to exploit the potential of electronic publication to meet research needs for whole-codex and cross-sectional studies of manuscript traditions. Traditionally, these needs have been met by individual researchers or research teams excerpting data and publishing results in hard copy, presenting cross-sectional data in atomized form. Moving forward, we can promote research by:

- publishing full-manuscript and cross-sectional text in electronic form
- encoding more analyzable information into our transcriptions
- encouraging the publishing of comparative corpora to complement other approaches