



Universität  
Zürich<sup>UZH</sup>

Slavisches Seminar

---

# Orthography as a window to diachrony

Barbara Sonnenhauser, Zurich

barbara.sonnenhauser@uzh.ch



## Introduction

### Orthography and language

- Orthography
  - Spelling
  - Punctuation
  - Spacing
- Orthography as a bridge between two systems of language: spoken and written (Vachek 1939: 113)



## Introduction

### Orthography and language

„We may hypothesize that medieval punctuation habits, at least in some languages, reflected the pragmatics of a text more closely than do modern punctuation rules.“ (Gvozdanović 2000: 332)

„The use of modern punctuation in philological editions fails to do justice to the language of the manuscripts.“ (Gvozdanović 1995: 177)

→ orthography and older texts?



## Approaching older texts

### Some pitfalls

- Bashing the scribe (see Fleischman 2000)
  - „negligence, defective linguistic knowledge, or mechanical copying methods of scribes“ (Fleischman 2000: 47)
  - variation as errors, language change as decay
  - e.g. ‚loss of‘: constructions, cases, ....



## Approaching older texts

КЪ ОУТРЪНЫ ЖЕ СЪНОМЪ СОБЪЮТЬ БЫВЪ, МНѢШЕ СЕ ЦАРИЦОУ  
НѢКОЮЮ НА ПРѢСВѢТЛѢМЪ СѢДЕШТОУ ПРѢСТОЛѢ ЗРѢТИ, И  
МНОЖСТВО МНОГО СВѢТЛЫХЪ ВОСННЪ СОКРЪСТЬ ТОЕ СТОЮШТІИХЪ  
(Novaković, p. 56)

И НА СЪМНВАНЕ ЗАСПѦ. И ВИДѢ ЕДНА ЖЕНА КАТО ЕДНА  
ЦАРИЦА, И СѢДИ НА СТОЛЪ ЗЛАТЬ СВѢТЛІКАВЪ. И ОКОЛО НЕА  
СТОАТЬ МНОГО ВОИНЕ ХЪБАВИ СВѢТЛІКАВЫ. (Tixonravovski, 56v)



## Approaching older texts

КЪ ОУТРЪНЫ ЖЕ СЪНОМЪ СОБЪЮТЬ БЫВЪ, МНѢШЕ СЕ ЦАРИЦОУ  
НѢКОЮ НА ПРѢСВѢТЛѢМЪ СѢДЕШТОУ ПРѢСТОЛѢ ЗРѢТИ, И  
МНОЖСТВО МНОГО СВѢТЛЫХЪ ВОСННЬ СОКРЪСТЬ ТОЕ СТОЮШТІИХЪ  
(Novaković, p. 56)

И НА СЪМНВАНЕ ЗАСПÀ. И ВИДѢ ЁДНА ЖЕНÀ КАТО ЁДНА  
ЦАРИЦА, И СѢДИ НА СТОЛ' ЗЛАТЬ СВЕТАИКАВЪ. И ОКОЛО НЕА  
СТОАТЬ МНОГО ВОИНЕ ХЪБАВИ СВѢТАИКАВЫ. (Tixonravovski, 56v)



## Approaching older texts

КЪ ОУТРЪНЫ ЖЕ СЪНОДЪ СОБЪЮТЬ БЫВЪ, МНѢШЕ СЕ ЦАРИЦОУ  
НѢКОУЮ НА ПРѢСВѢТЛѢДЬ СѢДЕШТОУ ПРѢСТОЛѢ ЗРѢТИ, И  
МНОЖСТВО МНОГО СВѢТЛЫХЪ ВОСННЬ СОКРЪСТЬ ТОЕ СТОЮШТІИХЪ  
(Novaković, p. 56)

И НА СЪМНВАНЕ ЗАСПÀ. И ВИДѢ ѐДНА ЖЕНÀ КАТО ѐДНА  
ЦÀРИЦА, И СѢДИ НА СТОЛ' ЗЛАТЬ СВѢТЛÏКАВЪ. И ОКОЛО НÏА  
СТОАТЬ МНОГО ВОИНЕ ХЪБАВИ СВѢТЛÏКАВЫ. (Tixonravovski, 56v)



## Approaching older texts

КЪ ОУТРЪНЫ ЖЕ СЪНОМЪ СОБЪЮТЬ БЫВЪ, МНѢШЕ СЕ ЦАРИЦОУ  
НѢКОЮЮ НА ПРѢСВѢТЛѢМЪ СѢДЕШТОУ ПРѢСТОЛѢ ЗРѢТИ, И  
МНОЖСТВО МНОГО СВѢТЛЫХЪ ВОИНЫ СОКРЪСТЬ ТОЕ СТОЮШТІИХЪ  
(Novaković, p. 56)

И НА СЪМНѢВАНЕ ЗАСПѢ. И ВИДѢ ЕДНА ЖЕНА КАТО ЕДНА  
ЦАРИЦА, И СѢДИ НА СТОЛЪ ЗЛАТЬ СВѢТЛІКАВЪ. И ОКОЛО НЕА  
СТОАТЬ МНОГО ВОИНЫ ХЪБАВИ СВѢТЛІКАВЫ. (Tixonravovski, 56v)





## Approaching older texts

### Some pitfalls

- Conceptual inertia (see Fleischman 2000)
  - application of concepts/categories of modern languages
    - e.g. ‚renarrative‘ as (alleged) grammatical category
      - inconsequent usage of renarrative / evidential forms
      - sometimes the author gets so much involved in the story that he forgets that he reports about events that are not witnessed by him personally (Cojnska 1979: 213-214)
      - “diffuse and rambling”, “unsettled language”, “chaotic language” (Moser 1972: 44)



## Approaching older texts

### Some pitfalls

- Conceptual inertia (see Fleischman 2000)
  - application of concepts/categories of modern languages
    - e.g. ‚renarrative‘ as (alleged) grammatical category
    - e.g. ‚sentence‘, ‚word‘ etc. as linguistic entities
      - spacing, punctuation



## Approaching older texts

### Variation

- within and across
  - texts and scribes
  - space and time
  - norms



## Approaching older texts

### Variation

- No native speakers available
  - Text as ‚native speaker‘ (Fleischman 2000: 46)
  - No negative evidence
- (Partial) control by textual parameters (Herring et al. 2000)



## Approaching older texts

### Variation

- linguistic reality
- prerequisite for language change

→ indispensable linguistic source



## Approaching older texts

### Variation

- But: what kind of variation?

„The language of a medieval [...] text is already a hybrid, reflecting not the idiolect of an individual language user, but the language of an author/composer filtered through one or more textual choices and subject to greater or lesser modification in the process.“ (Fleischman 2000: 46)



## Approaching older texts

### Corpora

- How to deal with variation in establishing diachronic corpora?
  - Faithfulness to the data
    - Leave interpretation to users
  - Accessibility of the data
    - Normalisation without interpretation?



## Bulgarian Church Slavonic → Balkan Slavic

Intended project: basic idea

- Diachronic corpus
  - Linguistic diachrony and textual history
  - Linguistic and philological perspective

→ Increase in reliability of data





## Bulgarian Church Slavonic → Balkan Slavic

Intended project: problems

- Language change
  - Structure
    - synthetic > analytic
  - Categories
    - new morphological categories (e.g. definiteness)
    - new conceptual categories (e.g. renarration)
- text evolution
  - textual structure (e.g. narrative chunks)
  - narrative structure (e.g. descriptive, mimetic, narrative)



## Bulgarian Church Slavonic → Balkan Slavic

Intended project: problems

- Comparability across time
    - Genre, textual and linguistic structure
    - Textual history (texts and witnesses)
  - Comparability across space
    - Translations (South and East Slavic)
    - Traditions (different text versions)
- Restriction on one text: the Life of Petka Tärnovska
- Diachronic corpus as parallel corpus



## Bulgarian Church Slavonic → Balkan Slavic

Intended project: problems

- technical make-up
  - transcription of manuscripts (OCR?)
  - annotation and alignment (linguistic and philological)
- trade-off between description and generalisation

→ example: orthography



## Orthography

### Standardized period

- Basic spelling principles
  - phonetic: *Piši kao govoriš!* (Vuk Karadžić)
  - etymological: *Govori za uši, a piši za oči!* (Ljudevit Gaj)



## Orthography

### Standardized period

- Basic punctuation principles
  - Rhetorical
    - reflection of orality (intonation, pauses, etc.)
    - not necessarily reflected in syntax
  - the more ‚rhetorical‘, the more insight into language use
- Syntactic
  - reflection of syntactic structures
  - not necessarily reflected in oral speech
- the more ‚syntactic‘, the more subject to and result of linguistic theory



## Orthography

Pre-standardized period: 16th-19th century Balkan Slavic

- transitional period in linguistic and literary development
- ‚intuitive‘ literacy

→ Orthography

- bridge between written and spoken language
- linguistic insight

→ Normalisation?



## Orthography

Pre-standardized period: 16th-19th century Balkan Slavic

- Aspects
  - Spacing → concept of ‚word‘
  - Punctuation → concept of ‚sentence‘
  - Spelling → concept of ‚language‘



## Orthography

Spacing: concept of ‚word‘?

– clitics

(1) bi nevr̃nuse (Punčo 363v)

(1′) i ne vr̃nu se (Angelov 1958: 106)

(2) ioñ se čudeše (Punčo 361v)

(2′) i oñ se čudeše (Angelov 1958: 106)





## Orthography

Spacing: concept of ‚word‘?

– negation

(3) *ne maľ děca* (Krčovski, 7)

(3′) *nemal deca* (Koneski 1974: 73)

– *ami – a mi*

(4) *a mi radujse* (Krčovski, 8)

(4′) *ami raduj se* (Koneski 1974: 74)



## Orthography

Spacing: concept of ‚word‘?

– *da*-constructions

(5) i gleda *danajde* togova človeka (Punčo 363)

(5′) i gleda *da* najde togova človeka (Angelov 1958: 107)

(6) sedogovorixu troicata *da* otrovatъ čřa (Punčo 361)

(6′) se dogovorixu tricata *da* otrovatъ čřa (Angelov 1958: 106)



## Orthography

Punctuation: concept of ‚sentence‘?

– Subordination

(7) nemu mu zapovědal sťyj apľъ Petrъ, ot tie dve ženi da ne se  
otdělue ot niXъ (Krčovski, 7)

(7′) nemu mu zapovedal svjatij apostol Petr ot tie dve ženi da ne se  
otdelue ot niX (Koneski 1974: 72)

(8) i *vide edna žena* kato edna carica (Demina 1971)

(8′) ? i *vide : edna žena* kato edna carica [to be checked with manuscript]



## Orthography

Punctuation: concept of ‚sentence‘?

- Subordination

(9) i mu kazala site s̃ii mĕsta po ime, *kako* gi vykatъ (Krčovski, 9)

(9′) I mu kazala site svjatii mesta po ime *kako* gi vikat. (Koneski 1974: 7)

(10) i e kazale sičko, *što* imъ učinilo na patъ (Krčovski, 9)

(10′) i e kazale sičko *što* im se učinilo na pat (Koneski 1974: 75)

→ syntactic structures?



## Orthography

Punctuation: concept of ‚sentence‘?

- Subordination

(11) togiva site člvěcy, čto imalo utoj gradъ, site se krestile (Krčovski, 9)

(11') togiva site čeloveci čto imalo u toj grad, site se krestile (Koneski 1974: 75)

→ different types of relative clause ?

[cf. contemporary Macedonian: only non-restrictive relative clauses are separated by a comma, Minova-Ćurkova 2007: 488]



## Orthography

Punctuation: concept of ‚sentence‘?

- ‚extra clausal‘ constituents

(12) *kěda umrešb, loša i gorka smertb* (Krčovski, 7)

(12') *k'e da umreš loša i gorka smert* (Koneski 1974: 73)

(13) *i edinb člvekb ot nixb imemu beše georgia vide sьnb* (Punčo 67v)





## Orthography

Punctuation: concept of ‚sentence‘?

- narrative rhythm

(14') Той црь Алексан'дрѣа, когн днгу войску да се вие сасъ грци  
н вн се за много време, н когн излезе от' дом' свой н отиде на  
войску н не връну се въ домъ свои четирдесеть лета. Све ходи н  
внеше се непрестано. Н каде удареше, сасъ божиємъ повелениемъ све е

(Angelov 1958: 106)





## Orthography

Spelling: concept of ‚language‘?

- Bulgarian, Macedonian, Serbian, Balkan Slavic?

(15) i momci *mlogo* stojutъ predъ neja (Punčo 67v)

(15′) ibie se *zamnogo* vreme (Punčo 360v)

(16) slepi prozirajut xromi *proxodutъ* (Punčo 68)

(16′) sveku godinu *daxodatъ* sasъ cřa na boi (Punčo 361)

→ Distribution across texts?



## To conclude

### Older texts

- Modern orthography
  - Imposes distinctions
  - Hides relevant information
    - variation
    - development
    - grammaticalisation



## To conclude

### Older texts

- Original orthography
    - textual alignment (if ‚sentence‘ and ‚word‘ are no valid units)
    - linguistic annotation
    - punctuation annotation
- two levels: normalisation & original



## References

Cojnska, R. 1979. *Ezikăt na Joakim Krčovski*. Sofija

Fleischmann, S. 2000. Methodologies and ideologies in historical linguistics: on working with older languages. Herring, S.C. et al. (eds.). *Textual parameters in older languages*. Amsterdam, 33-58

Gvozdanović, J. 2000. Parameters underlying punctuation in older Russian texts. Herring, S.C. et al. (eds.). *Textual parameters in older languages*. Amsterdam, 331-352

Gvozdanović, J. 1995. Parameters underlying the organization of mediaeval Russian texts. Andersen, Henning (ed.). *Historical linguistics 1993*. Amsterdam, Philadelphia, 177-190

Herring, S.C. et al. (eds.). *Textual parameters in older languages*. Amsterdam

Minova-Ġurkova, Liljana. 2000. *Sintaksa na makedonskiot standarden jazik*. Skopje

Moser, Ch. A. 1972. *A history of the Bulgarian literature 865-1944*. The Hague, Paris

Noel, P. & B. Sonnenhauser. 2013. Verschriftlichung zwischen System und Rede. Zur interpunktorischen Kodierung von funktionalen Performanzstrukturen. *Sprachwissenschaft* 38/3, 311-341

Sonnenhauser, B. 2015a. Sprachliche Strukturen, narrative Strategien. Zum funktionalen Sprachwandel im vor-standardisierten Balkanslavischen am Beispiel der Vita der Petka Tärnovska und des *Sbornik* von Pop Punčo. *Zeitschrift für slavische Philologie* (to appear)

Sonnenhauser, B. 2015b. Functionalising syntactic variance: declarative complementation with *kako* and *če* in 17th to 19th century Balkan Slavic. *Wiener Slavistisches Jahrbuch. Neue Folge*, 3 (to appear)

Sonnenhauser, B. & J. Fuchsbauer. 2014. Corpus-based analysis of changing norms: tracing the life of Petka Tärnovska from Middle Bulgarian Church Slavonic to Balkan Slavic. Baranov, V.A., V. Željazkova & A.M. Lavrent'ev (eds.). *Pismenoto nasledstvo i informacionite texnologii*. Sofija, 14-16

Vachek, J. 1939. Zum Problem der geschriebenen Sprache. *Travaux du Cercle linguistique de Prague* 8, 94-104



## Sources

Punčo: *Pop Punčov Sbornik*. 1796

Tixonravovski: Demina, E. I. 1971. *Tixonravovskij damaskin. Bolgarskij pamjatnik XVIIv. Issledovanie i tekst*. Sofija

Krčovski: Joakim Krčovski. 1817. *Čudesa presvetye bogorodicy*. Budimъ

Koneski: Koneski, B. (ed.). 1974. *Joakim Krčovski. Sobarni tekstovi*. Skopjek

Angelov: Angelov, B. 1958. *Iz starata bălgarska, ruska i srăbska literatura*. Sofija

Novaković: Novaković, S. 1877. *Život sv. Petke od patrijarha bugarskoga Jeftimija*. Starine 9, 48-59



## Acknowledgements

The research for this paper has been funded by the Austrian Science Fund FWF (project 'The emergence of narrativity in Early Neo-Balkan Slavic'; grant number M1536-G23). This support is gratefully acknowledged.

<http://pf.fwf.ac.at/de/wissenschaft-konkret/project-finder/29868>