Писменото наследство и информационните технологии

El'Manuscript–2014

Материали от V международна научна конференция
Варна, 15–20 септември 2014 г.

София · Ижевск
2014

Отговорни редактори: проф. дфн В. А. Баранов
                        доц. д-р В. Желязкова
                        д-р А. М. Лаврентьев

Редактори: Нели Ганчева, Веселка Желязкова (български текст)
             О. В. Зуга, В. А. Баранов (руски текст)
             Кевин Хокинс (Kevin Hawkins) (английски текст)

Сборникът съдържа материали от конференция, посветена на разработването и създаването на съвременни средства за съхраняване, описване, обработка, анализ и публикуване на ръкописни и старопечатни книжовни паметници и исторически извори, а също и на въпросите за подготвянето на електронни ресурси в областта на хуманитаристиката и тяхното използване в научните изследвания и преподаването.

Corpus-based Analysis of Changing Norms:
Tracing the Life of Paraskeva of Tărnovo
from Middle Bulgarian Church Slavonic to Balkan Slavic

Barbara Sonnenhauser, Jürgen Fuchsbauer

Diachronic corpus, Middle and Early Modern Bulgarian, normalization, Euthymius of Tărnovo, damaskini, Life of Petka

Back in 1976, Evgenija Demina formulated an 'interesting and important task' in sketching the linguistic development of Eastern South Slavic as evinced in written documents from the times of Evtimij to the beginnings of its standardization, i.e. the damaskini and the tradition originating from them:

> Интересной и важной представляется задача проследить, как <...> книжникам удалось решить сложнейшую проблему столкновения изысканного, нарочито архаизованного и усложненнаго языка и стиля Евфимия с демократическими чертами народной, близкой по своему духу к фольклору письменности дамаскинов <...> [Demina 1980, 187].

The linguistic development from OCS to the modern Eastern South Slavic standard languages as manifested in written documents reflects not only a change in linguistic features and language structure but more fundamentally a change in norms. Whereas the norms reflected in contemporary written documents are predominantly based on intentionally established authoritative prescriptions, the norms reflected in older stages are essentially convention-based. The writers did not consciously adhere to a pre-established norm, nor did they intend to establish one by their writings [e.g., Cojnska 1979: 23–24]. Still, specific convergences can be observed [e.g., Sonnenhauser 2014], which suggests the existence of implicitly underlying conventions. These conventions are related to the functions of the texts and form the basis of two main types of norms: internally, peer-group-oriented norms with the main aim of adhering to an authoritative tradition at the expense of comprehensibility ('exclusive' norms), and externally, people-oriented norms with the main aim of achieving comprehensibility at the expense of tradition ('inclusive' norms). This correlates with the text types at hand, translation and (re-)narration being the most important ones. As a consequence, the investigation of changing norms presupposes an integrated approach that takes into consideration not only the temporal dimension but also further features related to language, function and text type. These components are summarized in figure 1.

Our research concentrates on changes in convention-based norms from Middle Bulgarian Church Slavonic to Balkan Slavic. We rely on the assumption that 'norm' as reflected in written documents dating to this period is a relative notion; norms emerge from the interplay of linguistic means, on the one hand, and function and text type, on the

other. Variation preceding the consolidation of norms is encountered on all possible levels: graphic, linguistic, textual and content-related.

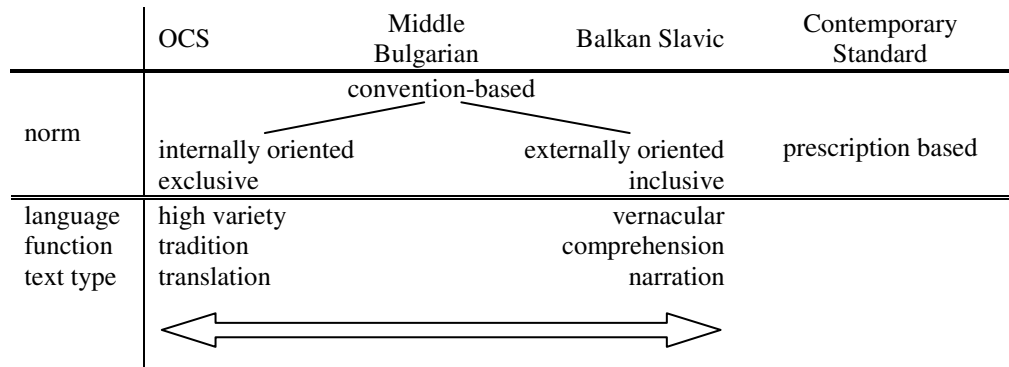|	OCS	Middle Bulgarian	Balkan Slavic	Contemporary Standard
norm	convention-based			prescription based
	internally oriented exclusive		externally oriented inclusive	
language function text type	high variety tradition translation		vernacular comprehension narration	

Figure 1

In order to empirically verify these assumptions and analyze the changing norms, the compilation of a well-designed corpus is required. Sketching the multilayered developments underlying the shifting norms cannot be achieved by analyzing single documents with respect to single features, dating to different, isolated points in times. It rather calls for a coherent set of texts that complies with the following requirements:

- reflects different diachronic stages
- reflects different dialectal bases
- reflects different writing traditions
- reflects different textual intentions

In order to obtain maximum comparability for our testing of the above presuppositions, we decided for the time being to focus on a single text of manageable scope, which nevertheless covers spatially and temporally the phenomena in question. Patriarch Euthymius' *Life of Petka* (Paraskeva of Epibatai) appears to be ideally suited. Together with its vernacular versions contained in damaskini it represents a case in point for the radical change in normalization occurring between the medieval and early modern period of Slavic written language in the Balkans. Composed presumably between 1376 and 1382 by Euthymius at the behest of Tsar Ioann Šišman [Demina 1980, 183], the Slavic *Life of Petka* is transmitted in several redactions differing, above all, in respect to their extent [Ivanova 2008, 246ff.]. A shortened version served as basis for two separate translations into the Bulgarian vernacular, one dating to the beginning of the 17th century, the other to its end [cf. Demina 1980, 184]. Demina, while deeming the older a free re-narration of the Church Slavonic text, characterizes the younger as a decidedly literal translation [op. cit. 187s.]. A thorough comparison of both, as well as a juxtaposition with their respective Church Slavonic models, can be expected to give detailed insight into different strategies of grammatical normalization, as well as into diverging approaches to the text as such. Hence, in order to provide the relevant information, a sample corpus of

the *Life of Petka* has to contain at least one instance of each textologically and linguistically relevant version of the text. Comparability will be established by applying means of textual criticism.

The resulting basic corpus is meant to be expanded over the course of time. The provisional restriction to only one comparatively short work will allow us to test existing technical means used in the creation of diachronic corpora in order to assess and, if possible, improve their applicability to premodern Bulgarian texts. Therefore, the 'Petka Tărnovska corpus' is valuable also from a technical point of view. In addition, the small size of this corpus allows it to be manually annotated. It can thus serve as training corpus for further automatic annotations and the development of an electronically searchable corpus of Middle and Early Modern Bulgarian. On the basis of this corpus, precise requirements for the digital compilation and editing of older texts can be formulated.

In our paper we shall outline predominantly from a philological point of view the specific challenges a combination of Middle and Early Modern Bulgarian texts poses for the creation of an electronically searchable diachronic corpus. Comparing the single versions we shall determine to what extent direct equivalence can be established and, if so, how far this may facilitate automated tagging. We shall also discuss whether the orthographic variation in premodern Bulgarian texts, which is greater than in other Slavic versions, can be expected to negatively affect the browsability of such corpora. By defining prerequisites for and obstacles to the processing of texts like the versions of the *Life of Petka* we hope to make a relevant contribution to the development of technical solutions for the difficulties resulting from considerable linguistic and textual variation.

References

Cojnska, Ralica. 1979. *Ezikăt na Joakim Kărčovski*. Sofija

Demina, Evgenija I. 1980. „Žitie Petki" Evfimija Tyrnovskogo v novobolgarskoj pis'mennosti. *Učenici i posledovateli na Evtimij Tărnovski. Vtori meždunaroden simpozium. Veliko Tărnovo, 20–23 maj 1976*. Sofija, 183–192.

Ivanova, Klimentina. 2008. Bibliotheca Hagiographica Balcano-Slavica. Sofija.

Sonnenhauser, Barbara. 2014. Erzählen im Bulgarischen zwischen Texttyp und Genre. Pressebericht und Evangelium, *Krimi i pravo* und *Tălkovanie. Die Welt der Slaven* 59/2, 355–381.