Standardization in Balkan Slavic Diachronic Research

Стандартизация в историческом исследовании балкано-славянских языков

Ivan Šimko

University of Zurich,
Switzerland

Иван Шимко

Цюрихский университет,
Швейцария

Abstract

The present paper studies the problem of standardization of Bulgarian within the context of the emergence of the Balkan Sprachbund. Traditionally, standardization is considered to be a part of the nation-building process, understood as the codification of orthographic and other linguistic norms in authoritative documents. As they are legally binding within the national collective, the traditional view distinguishes texts from the era before standardization containing more dialectal phenomena and the standardized literature, where dialectal features are usually suppressed.

This study presents the hypothesis that the codification of the Bulgarian language in the 19th century did not have such an impact on the later development of language norms. Rather, the codification merely led to changes in orthography. Other norms of the literary language gradually developed within the manuscript tradition of the so-called damaskini. This hypothesis is supported by a quantitative analysis of a sample of texts from various centuries and dialectal areas.

Резюме

Настоящая статья посвящена проблеме стандартизации болгарского языка в контексте конвергентных процессов, приведших к образованию балканского языкового союза. Традиционно стандартизация языка рассматривается как часть процесса становления нации и подразумевает кодификацию орфографических и других языковых норм в авторитетных лингвистических документах. Поскольку эти документы имеют правовое значение для формирования национального коллектива, в истории литературного языка каждой нации обычно принято строго отличать тексты, созданные до стандартизации и сохраняющие различные диалектные явления, от литературы на стандартизированном языке, почти не допускающем появления диалектных черт.

В статье высказывается предположение о том, что кодификация болгарского языка в XIX в. не оказала существенного влияния на последующее развитие языковых норм, изменив всего лишь орфографию. Прочие лингвистические особенности развивались постепенно в рамках рукописной традиции так называемых дамаскинов. Данное предположение основывается на квантитативном анализе избранных текстов, относящихся к разным векам и диалектам.

Ключевые слова

балкано-славянские языки, церковно-славянский язык, дамаскины, канберрское расстояние, орфография

1. Problem

When linguists, focusing on synchronic research, take literature as a source for older stages of a language, they tend to follow some presuppositions worth to think about. One is the idea that the scribe or the author can be localized in time and space, writing in the language used in the area under analysis, and thus should be classified as a representative of the variety (e. g. [Friedman 1986: 282; Sonnenhauser 2015: 49]). If not, as for example in the situation of diglossia, there would be at least traces or tendencies of the local vernacular left behind in their literary production (e. g. [Miklošić 1871: 6]).

There are many factors complicating such a simple classification. First, it is not only the vernacular which causes the author to deviate from the literary norm: genre, contents, inspirations, the language of the source or that of the selected audience might have an impact, too. Additional factors might include the author's level of education, preferences, aspirations or actual place in the social hierarchy, and interferences from other languages acquired throughout his life. Second, the language of the literature also may or may not be

represented by a stabilized norm, even if it seems to reflect an older stage of the language. Earlier literary norms often lack explicit formulation (prescription), and thus require reconstruction.

Another problem is the presupposed dichotomy between the pre-standardized and standard languages. The former, but also informal registers of the standardized language, "the vernacular speech of ordinary people", are considered natural [Milroy 1999: 37]. Their natural character contrasts with that of a standard language, a set of linguistic norms promoted by an authority (e. g. an official prescription or academic consensus). These authorities symbolically evaluate language shifts or individual structural features of spoken or written practices as correct or incorrect ("mistakes"), as signs of corruption ("patois" [Weber 1976: 67f.]), or also as indices of inclusion or exclusion within the political community ("shibboleths"). Whether the norms produced by such an authority differ from other motivations behind the language shifts or not, remains an open question.

The question is about the nature of standardization itself. The emergence of official language norms is usually considered to be an important part of the nation-building processes, spreading either from "above", by means of the centralized education and mass media within the borders of a state [Weber 1976: 303f.; Anderson 2006], or from "below", through a network of educators and artists, which could gradually develop into a national political movement [Handelman 1977: 196; Hroch 1985]. There is a less clear consensus regarding the question, whether the standardization follows rather internal or external needs of the language community: whether it answers requirements of its new administrative function in the modern society [Bourdieu 1991: 48], or the need to establish clear boundaries between communities, criteria for membership and inclusion [Barth 1969: 15].

The differences between the individual processes of standardization are likely to be as numerous as those between individual national movements. For that reason, it is also hard to establish a clear boundary between the "natural" pre-standardized variety of a language and its "artificial" standard. Proponents of constructivist views of modern nations describe the emergence of standard languages in constructivist terms, that is, focusing on their artificial features [Hobsbawm 1990: 54], while their opponents stress the importance of features preserved from the pre-standardized literature instead [Hastings 1997: 3].

The problem of standardization has some methodological implications for diachronic studies among the branch of South Slavic showing the features specific for the Balkan area, like the postpositional definiteness marking or the use of the mid vowel (e. g. [Leake 1814: 380; Schleicher 1983: 210; Haarmann 1976: 85; Tomić 2006]). These usually include two literary standards

(Bulgarian and Macedonian), as well as related dialects, like that of the Prizren-Timok area [Friedman 2017: 2]. The diachronic spread of these features is a controversial topic. This can already be seen on the classification of the Prizren-Timok dialects. Although a separate ethnonym (Torlak) is known to both a colloquial use [Skok 1971 III: 484] and linguistic discourse [Vuković 2020], it is not uncommon to classify them as Southeast Serbian [Belić 1905] or Northwest Bulgarian [Стойков 1993: 104]. As the classification of these dialects as Serbian or Bulgarian is not devoid of suspicions of promoting national interests, an umbrella term "Balkan Slavic" was proposed for the whole area (e. g. [Sobolev 1996: 63]), including the Bulgarian and Macedonian standards.

The term "Balkan Slavic" also has been employed in diachronic studies as a label for the transitionary literature, which diverges considerably from the Church Slavonic-based literary language in the time when these features were developed, but before the actual standardization of modern Bulgarian and Macedonian (cf. [Sonnenhauser 2015: 50]). The process of their standardization is often described as a set of arbitrary decisions of the state executive concerning the most controversial and symbolically laden features (e. g. [Irvine & Gal 2000: 60f.; Fielder 2019]); thus modern standards tend to be perceived primarily as tools of the language policy and, again, national interests. On the other hand, local scholars (e. g. [Керемедчиев 1943; Конески 1967: 40; Андрейчин 1977: 166]) focus on establishment of a standard by their own predecessors, the grammarians. These include the publication of prescriptive grammar books and dictionaries (e. g. [Рилски 1989; Пуљевски 1875]) and polemical treatises (e. g. [Дринов 1911; Мисирков 1903]). Both these descriptions agree on the role of authoritative prescription in the development of the language standard. However, they both show less interest in the consequent acceptance (or rejection) of the norm in the use of daily communication and literature.[1]

Thus, the usual picture of a standardization process is a timeline of prescriptions: discrete steps of intruductions of individual norms concerning grammar, lexical items and graphic features. The acceptance of the prescription by the writing (or speaking) public is taken for granted. In this text, we will present an inverse picture: how the said prescriptions reflect the preceding practice in literature. In other words: an evolutionary model of standardization.

---

[1] Prescriptive decisions lacking appeal may be rejected. One of the possible reasons are political connotations. For example, the use of the letter jat—<ѣ> (cf. below section 3.6.) became an issue in struggle between Bulgarian nationalists and the Agrarian Party in the 1920s [Андрейчин 1977: 168]. The current Bulgarian alphabet (without <ѣ>) was officially established only in February 1945. Bulgaria's new status as a Soviet satellite simply put an end to the debate.

2. Historical Background

Church Slavonic[2] must have seemed very archaic and foreign to a 16th–18th century Bulgarian reader. It was accessible only to a limited number of people from an educated audience. The actual number of its active users (i. e. writers, hardly any speakers) was very low in the 18th century even in the countries where it enjoyed a high status, like Wallachia, Moldova and Russia [Trunte 2018: 5]. In these countries it was also used in secular administration, alongside very different vernaculars. But as these countries were replacing it with literary languages based on spoken Romanian and Russian, Bulgarians also asked themselves whether it would not be more adequate to write as they speak. The first attempts to write in a language accessible to the broad audience in post-medieval Bulgaria can be dated to the 17th century. One of the first documents written in such a language can be seen in the Catholic prayer book called *Abagar* from 1651. The language of this booklet with "prayers used by the converted heretical Paulicians as amulets" [Stefanov 2008: 60] reflects the local dialect, but with considerable Croatian and Italian influences [Tsibranska-Kostova 2016: 14]. However, other documents representing the Catholic literature in Bulgarian are scarce.

Another variant of a vernacular-based literary language can be seen in the documents called *damaskini*, translations of the collection of homilies *Thēsauros* by Damaskēnos Stouditēs. Published in print first in the 1560s in Venice, the book became famous for its use of the language of the commoners. It was soon translated into Church Slavonic, but was less accessible to a less educated audience than the original. Early in the 17th century, a new translation into a language called *simple Bulgarian*[3] emerged. As its Greek original, it diverges from the usual literary language of its period both from the points of view of stylistics and of grammar. Compare the following sentence from the Church Slavonic hagiography (*Life*) of St. Petka [Vuković 1536: 196v; ex. 1] and its damaskini edition [*Tixon. d.* 56v–57r; ex. 2]:

---

[2] The term *Church Slavonic* is used in this article in a broad sense of the term, denoting a supraregional, polycentric literary language (cf. [Кайперт 2017: 23–29]). In a narrower sense of the word, it denotes the norms explicitly described by Constantine of Kostenets [Ягич 1895: 387–487] and Smotrickyj [Смотрицкий 1648]. Such use is not very common in Bulgaria, where the term usually denotes the later Russian redaction (e. g. [Демина 1985: 14]; maybe "Ruthenian" would be more suitable in this aspect). A proper equivalent would be *Middle Bulgarian*, which is, however, ambiguous from the perspective of the relation between spoken and written language.

[3] The term is based on the headings of the texts authored by Stoudites: *metaphrastheis eis tēn koinēn glōssan* 'translated into the common language' (e. g. [*Thēsauros* 1751: 5]). Church Slavonic editions translate the phrase *eis tēn koinēn glōssan* literally *obʼštуmь ęzykomь* (SG.INST '[by the] common language'), while their translations into early modern Bulgarian use adjectives *prostуmь* 'simple' or *bolgarskуmь* 'Bulgarian'.

(1) *stráxomъ*   *ábïe*   *óbъjetъ*   *bývъ*   *na+⁴* *zémlju*
     fear.SG.INST   suddenly   overtaken.SG.NOM   having been.SG.NOM   to   ground.SG.ACC

     sébe   *povrъže*
     REFL.GEN/ACC   throw.3SG.AOR

     'having been suddenly overtaken by fear, he threw himself on the ground'

(2) *i+*   *tói+*   *se*   *upláši*   *i+*   *pade*   *ná+*   *zemlja*
     and   M.3SG.NOM   REFL.ACC   scare.3SG.AOR   and   fall.3SG.AOR   to   ground.SG

     'and he got scared, and he fell to the ground'

Stylistically, Church Slavonic of example (1) prefers longer, complex sentences with participles in subordinate clauses. Example (2) shows a simplified stylistic structure with two separate sentences with finite verbs. It also reflects the morphosyntactic developments of the local dialects: marking spatial relation (*na zemlja* 'to the ground') by the preposition only, without the specific case marker as in Slavonic example (1). Because of their radical break with both grammatical and stylistic norms, the damaskini have been extensively studied by modern linguists since their discovery in the 19th century (e. g. [Jagić 1877; Дринов 1911: 315f.; Аргиров 1895; *Kopr. d.*; *Svišt. d.*; Петканова-Тотева 1965; *Trojan d.*; *Tixon. d.*; *Loveč d.*⁵]).

    The earliest transcripts are anonymous and practically mechanical, supplementing the lack of printing technology in the area. In the 18th century, many "authored" editions appear. Works of such writers as the monk Josif Bradati (ca. 1714–1757) and priest Stojko of Kotel (1739–1813) are often considered a fusion of Church Slavonic morphology with modern syntax [Вътов 2001: 7]. This is also the case of the *Slavenobulgarian Chronicle* by monk Paisius of Hilendar (1722–1773), which spread in transcripts from the 1760s. In the following example [Иванов 1914: 42], Paisius retains archaic PRS.3SG ending *-tъ* and a F.SG.GEN for the name, as in Church Slavonic, but not the expected locative (**žitïï*) after the preposition:

(3) *kako*   *pišetъ*   *vъ*   *žitiè*   *prepodobnie*   *Paraskevì*
     as   write.3SG.PRS   in   life.NOM/ACC   reverend.F.SG.GEN   parascheva.GEN

     '[...] as it is written in the *Life of St. Parascheva*'

Unlike the *simple Bulgarian* of the damaskini, the language of the *Nedělnik* and the *Chronicle* preserved some of the holy aura of the liturgic language, while being (likely) sufficiently comprehensible to less educated public. The language is sometimes called *Slavenobulgarian* [Керемедчиев 1943: vii],

---

⁴ The <+> marks tokens which are written together with the following word in the original. See Table 2 below for transcription rules.

⁵ For a more detailed overview of the early history of the damaskini studies, cf. [Демина 1968: 11f].

eponymous with the most famous work written in it, stressing the continuity with the Church Slavonic literature. The grammarians of the early Bulgarian national movement in the 19th century were still torn between the "tyrants", like Neophyte of Rila, proponents of using *Slavenobulgarian*, who would rather force the common folk to learn the archaisms, on the one hand, and the "demagogues" like Petăr Beron persuading their fellow literates to abandon their idealized inflection markers and write in the *simple* language of the commoners for the sake of contemporary trends, on the other [Ibid.: v].

In short, the situation of the Bulgarian literature of the 16–18th century is not simply one of diglossia. Literature was written in Church Slavonic, *simple Bulgarian* and *Slavenobulgarian*, following various orthographic and grammatical norms, and existing alongside dialects with varying sociolinguistic status. In the words of Marin Drinov [Дринов 1911: 274], the written language from the 16th to the 18th century is ruled by "endless chaos" (*bezkrajna bărkanica*). Thus we will look at possibilities of studying the interferences between these norms.

3. Comparison

Our model of *standardization* is based on the gradual adoption of various linguistic practices, which we perceive as features of a historical text. Certain practices are established as a *norm* binding the written (or even spoken[6]) production of the community. For some of these practices we can also find counterexamples of *destandardization*: linguistic practices originally adopted by an earlier standard, which gradually fall out of use due to incompatibility with standardized features, due to alienation by language change or due to its redundant character. A standardized feature does not have to be an *innovation*: a linguistically archaic feature can be adopted or simply withstand attempts for removal. Nor does a linguistic innovation need to be standardized; it may retain the status of a substandard or foreign feature, being systematically avoided in the texts and speech of higher status. Features which are avoided either systematically[7] or by promoting an incompatible alternative can be dubbed *non-standardized*.

We try to separate the concepts of *codification* and *standardization*—the publication of an authoritative document calling for an adoption of an explicitly formulated linguistic norm, and the adoption itself. But these documents

---

[6] Although standardization does effect the spoken practices too (cf. [Milroy 1999: 47–59]), but our study focuses on developments attested only in text sources.

[7] Fuchsbauer [2010: 177] describes one such case in the Church Slavonic translation of *Dioptra* by Philippos Monotropos, which avoids the postponed demonstratives with the *t*-root (only *on-* and *s-* are used), abundant in other Church Slavonic redactions and the damaskini (cf. below section 3.1.).

are still of great importance for our analysis. On their basis we identify the features, which can be used to illustrate the development of a standard. In the case of Standard Bulgarian, our body of authoritative literature includes early primers and grammar books [Берович 1824; Рилски 1989; Богоров 1844; Хрулев 1859; Момчилов 1868[8]], influential polemics [Дринов 1911], as well as decisions of the state executive [*Упътване* 1899; *Наредба* 1945]. It is harder to find such documents for older stages. Norms of older literature were not codified in the modern sense, lacking means of enforcement comparable to those of modern standard languages. For this reason, historical grammars [Ягич 1895, Смотрицкий 1648] are only of limited relevance. It is necessary to use secondary descriptions (e. g. [Велчева 1966; Гълъбов 1968; Христова 1991; Вълчев 2007] and modern Church Slavonic grammars issued by ecclesiastical authorities [Бончев 1952; Миронова 2010] for reference as well.

Let us assume that standardization includes both aspects: the grammar is taught together with the orthography. If, on the other hand, orthographic differences are bigger, then it is reasonable to expect more dialectal influence among the linguistic features of a source. For this reason, we will also discuss purely graphic features, like the script, accentuation and abbreviations. Each (grammatical or orthographic) feature can be represented as a variable, a property of an individual text source. These variables then can serve as a basis of comparison between the sources.

Table 1 lists the features represented as variables for our analysis. Standardized features reflect practices codified by Bulgarian grammarians of the 19th–21st century—present-day Standard Bulgarian. The second column lists the features, which are not only present to some degree in the literature of the pre-standardized period (16th–19th century) of all three (Church Slavonic, *simple Bulgarian* and *Slavenobulgarian*) literary traditions, but also mentioned

---

[8] The choice tries to focus on influential sources. Keremedčiev designated Neophyte of Rila as the "undoubted leader" [Керемедчиев 1943: xii] of the *Slavenobulgarian* school of grammar. However, he was not the first one publishing a systematic text on the matter. Another *Slavenobulgarian* grammar was published shortly before the Neophyte's by Emanuil Vaskidovič and Neophyte of Hilandar-Bozveli (cf. [Вълчев 2007: 81]), and there were also other influential texts with similar premises, appearing soon afterwards (e. g. [Павлович 1836, Венелин 1838]). In a similar vein, Bogorov's grammar from 1844 was described by Keremedčiev as one having a "strong influence on all teachers and grammarians of the period" [Керемедчиев 1943: xxi]. Even contemporary scholars like Vălčev consider it a "landmark" for modern Bulgarian philology and grammar [Вълчев 2007: 222]. The choice of an authoritative grammar for the period after Bogorov is harder due to the sheer number of publications in the period—the availability to the author in the time of writing the article being a major argument. While Xrulev's grammar more or less reiterates the principles set by Bogorov, it is interesting for us, as our corpus includes a text written by the same person (i. e. [*Nedělnik* 1856]). Momčilov's grammar is one of the most voluminous among the grammars of 1860s, and it also receives most attention by Vălčev [Вълчев 2007: 335–356].

by Church Slavonic grammarians. Non-standardized features can be observed in older literature (especially from the *simple* and *Slavenobulgarian* traditions), but are not adopted by today's standard, nor are they present in Church Slavonic.[9]

Overview of analyzed features

Standardized innovations	Slavonicisms/archaisms	Not standardized features
1.1. Postnominal article	2.1. CS nominal inflection	3.1. Inflected articles
1.2. Postadjectival article	2.1a. Non-NOM endings	3.2. Articled short form adjective
1.2a. M.SG adjectival ending -*ija*	2.1b. M.SG -*a*	
1.3. Extended demonstrative	2.1c. F.SG -*u*, -*ь* or -*ǫ*	3.3. "Future indefinite" tense
1.4. DAT possessive pronoun	2.1d. M.SG -*u*	3.4. Differential object marking
1.5. *šte* particle for FUT tense	2.2. Long-form adjective	3.4a. Object doubling
1.6. Analytical infinitive marking	2.2a. M.SG adjectival ending -*ij*	3.4b. 3SG.ACC for indirect objects
1.7. Unified orthography	2.3. GEN possessive pronoun	
1.7a. Non-final/non-palatal /ă/	2.4. Proximal deixis marking	3.5. Non-Cyrillic script
1.7b. /i/ in all positions	2.5. Synthetic infinitive marking	3.6. Specific letter for /dʑ/
1.7c. /ja/ and final /jă/	2.6. Old 2/3PL aorist forms	3.7. Simplified accentuation
1.8. Separation of unaccented words	2.7. Archaic letters	
1.9. No accent markers	2.7a. Use of <ѣ>	
1.10. Arabic numerals	2.7b. Use of <ы>	
	2.7c. Use of <ѧ> for /ja/	
	2.8. Loanword-specific letters	
	2.9. Word-final jers	
	2.10. CS accentuation	
	2.10a. Use of all four markers	
	2.10b. Breve on syllable-final vowel	
	2.10c. Writing of *spiritus lenis*	
	2.11. Lexicalized abbreviations	

4.1 Grammatical features

The first variables reflect the most visible difference between Standard Bulgarian and Church Slavonic: the amount of definiteness markers following nouns (1.1), adjectives (1.2), as well as the amount of nouns with non-nominative endings (2.1a). These were intensively debated in the 1830s grammars like

---

[9]   The choice is roughly based on the lists of features specific for Church Slavonic of the Resava redaction and the language of the damaskini employed by Velčeva [Велчева 1966: 117] for their comparison. Non-standardized features were not listed, but they are relevant from the point of view of discussed topics.

that of Neophyte of Rila, who used Smotrickyj's grammar as a model [Рилски 1989: xvii]. Trying to preserve at least traces of old nominal inflection,[10] he introduced dialectal variants of the article for the M.SG animate paradigm, fusing the demonstrative function of the article with the syntactic function of the case ending [Ibid.: 86]:

(4) N *stárec**o***
G *na-stárc**a*** or *na-stárec**a***
D *na stárec**atъ***
A *stárec**atъ***

According to Fielder, Neophyte thought that the case endings and articles would occupy the same morphological slot [Fielder 2019: 46]. He indeed uses terms *člénъ* 'article' and *padéžъ* 'case' interchangeably [Рилски 1989: 163]. His idea of employing dialectal differences in phonetics to mark the case was not accepted by the writers, but nominal inflection can in a limited extent be observed in the literature of the time. The article variant -*a* was homographic with the old animate M.SG.OBL (GEN/ACC) case ending, common at least in literature with proper names and *nomina sacra*. For this reason, the marker is ambiguous. In our study, Variable 2.1b (M.SG -*a*) thus reflects the presence of any -*a* ending in M.SG nouns. Variable 1.1 (postnominal article) reflects only the situation when a token contains the root of a demonstrative pronoun, positioned after the morphological case ending of a noun. It does not include modern Bulgarian suffixes -*t*, -*ta* etc. only, but also Church Slavonic short demonstratives (*sь*, *onъ* etc.) following the noun.

Neophyte did not address the use of articles inflected for case, which can be found in some peripheral (e. g. Rhodopean and Timok) dialects even nowadays, as well as in some lexicalized relics in Standard Bulgarian (e. g. *pettjax* 'about five' [Мирчев 1978: 201]). Such instances are reflected in the Variable 3.1. The variation between nominative and oblique endings is attested in older damaskini, as well as in *PPS* (1796: 11r[11]) for both MASC and FEM articles:

(5) *póče* *avrámъ* *da+* *ljúbi* *róbinju+* *tu*
begin.3SG.AOR abram.NOM to love.3SG.PRS servant.F.SG.ACC DEF.F.SG.ACC
'Abram fell in love with the servant'

---

[10] Neophyte's proposal is actually a compromise between the "tyrant" and "demagogic" positions on the matter of cases and articles. Venelin argued against the standardization of articles, because he found them absent in Macedonia [Венелин 1838: 46]. Pavlovič accepted some of the articles, wroting them as separate words (e. g. *prosty o ęzyk* 'the simple language'), but he argued for more inflection (e. g. in plural), because of many fossilized forms attested in dialects (e. g. *sъ bogomъ* [Павлович 1836: 8]).

[11] Some of our sources show multiple page/folio numberings. In such cases our citations refer to the original page numbering.

Punčo employs the oblique ending -*tu* along the otherwise generalized -*ta* (e. g. *na planináta* 'on the hill') for animate feminine nouns. Bogorov [Богоров 1844: 20] introduces a similar marking of F.SG.ACC nouns with -*ǫ* (Cyr. <ж>; e. g. F.SG. NOM/GEN/DAT *dušata* 'the soul', ACC *dušǫtǫ*) in his grammar. Although it is not clear whether this variation could be reflected in speech,[12] it was followed in literary practice until the late 1860s, when it was destandardized again by Momči-lov and Drinov (Var. 2.1c). These two also remove the marking of M.SG.DAT with -*u* from the standard. In earlier grammars, the dative ending could be at-tached to names, kinship terms and other nouns, which never carry an article (e. g. *Bogu* 'to God' [Рилски 1989: 91; Богоров 1844: 26]), in the grammars. Momčilov [Момчилов 1868: 28] destandardizes these forms as archaisms (Var. 2.1d).

One marker that survived Drinov's criticism was the m.sg ending -*a*. This ending works like the one defined by Neophyte: it fulfills both the syntactic function of an oblique case ending and the definiteness marking function of an article. Earlier literature still shows examples, where it is used as a general m.sg definiteness marker without the syntactic function [*Nedělnik* 1856: 257]:

(6) *diavola* *sę* *prestruvaše* *na* *razny* *zvěrove*
devil.DEF REFL.ACC change.3SG.IMPF to various.PL beast.PL
'the Devil changed himself to various beasts'

The current rules of its use were adopted into Ivančev's orthography (1899; cf. [Андрейчин 1977: 166]) and—despite recurring criticism [Fielder 2019]—has remained in written practice until today. Another homographic ending is used in the nominal count form (*brojna forma*), which is used in masculine nouns after numerals. This form is usually considered a fossilized dual nominative (e. g. [Мирчев 1978: 195; Маслов 1981: 149]), and it can be observed already in the damaskini [*NBKM 1064* 37v]:

(7) *utíduxa* *sítzki+ti* *pisjá* *du* *tzétiri* *pógleda*
go.3PL.AOR all.PL.DEF by foot to four shot.DL
'all went by foot four shots away'

In Standard Bulgarian, m.sg adjectives do not only express definiteness with the article, but also with an older root extension -*ij*-, as it is also seen in Neo-phyte's grammar (M.SG.NOM.DEF *svętýo* 'saint' vs. indefinite[13] variants *svętъ* or *sveti* [Рилски 1989: 102–103]. The extension is based on the old expression of definiteness by the suffixation of the pronoun **jь* at adjectives—also called compound or long-form [Lunt 2001: 142]. Bogorov's grammar [Богоров 1844:

---

[12] For the discussion of F.SG.ACC marking in damaskini, cf. [Велчева 1966: 117, Мирчев 1978: 168, Mladenova 2007: 306].

[13] Neophyte calls this form *so ousěčénïemъ* 'with shortening' (lit. 'cut'), which is similar to the terminology applied to Serbian adjectival short forms by Vuk [Karadžić 1974: 41].

35] preferred to write—as in Church Slavonic—the M.SG.NOM long-form without the article (e. g. *svetyĭ*), which was also preferred by Drinov [Дринов 1911: 283]. Although this practice may have been based on some Moesian dialects, where the unarticled m.sg long-form ending functionally fused with the article [Младенов 1963: 404f.], it was destandardized by Ivančev's reform in 1899 [Андрейчин 1977: 166]. Therefore, M.SG adjectives ending with variants of *-ii* are reflected by a separate variable (2.2a).

While the nominal M.SG ending *-a* is hard to distinguish from the short article or count form in a text, an articled adjective, based on the historical long-form (i. e. with ending *-ija* or *-ijat*), is unambiguous. Such forms, first attested in the 13th century [Мирчев 1978: 205], are avoided in later Church Slavonic redactions, where M.SG.GEN forms would be short *svęta*, long *svętago* (cf. [Миронова 2010: 101]). Adjectives with ending *-ija* are thus counted by Variable 1.2a as a standardized innovation. If an article follows an adjectival short form, as it is attested in many dialects across Bulgaria (e. g. *carskăt sín* 'royal son' [Mladenova 2007: 371]), the form is reflected by Variable 3.2.

Neophyte was indeed aware of the difference between the Church Slavonic long-form endings and the article [Рилски 1989: 170], as he removed long-forms in all positions of his paradigm. In the literature, adjectival long-forms appear in various genders and numbers too (e. g. F.SG.NOM *krasotà rá skaa* 'beauty of the Paradise', [*Ljub.d.* 97v]), although not very consequently.[14] Nonetheless, the fusion between definiteness marker and inflection, which in principle is not different from that of Church Slavonic and Vuk's Serbian [Karadžić 1974: 41f.], can be seen in other grammars. Bogorov [Богоров 1844: 35] codifies the M.SG oblique ending *-ago* (e. g. *svetago*) as an optional variant and Xrulev [Хрулев 1859: 28] even gives distinct short and long-forms for all three genders in SG (but only M.PL). All these endings were removed from the paradigm in Momčilov's [Момчилов 1868: 34] grammar, so we can consider them a destandardized feature (Var. 2.2).

As already mentioned above, Church Slavonic uses demonstrative pronouns, which can be placed both in front of their head noun (e. g. *v' tóĭ vési* 'in that village' [Rostovski 1689: 282v]) or following it (e. g. *putém' těm* 'by that road' [Ibid.]). The pronoun can be extended by a relative suffix *-žde* (e. g. *toę́žde nošti* 'in the same night' [Ibid.]). Modern Bulgarian uses a similar construction for the adnominal demonstrative, adding suffixes *-zi*, *-va* or *-ja* to the root (e. g. *tazi* F.SG 'that'). Such extended pronouns are already attested in the 12th century [Мирчев 1978: 182], but they are rare in Church Slavonic in the 17th–18th century. Neophyte [Рилски 1989: 116] codified the forms with suffixes *-ja* (M.SG.NOM *onyĭ*, F.SG *onáę*, PL *onýę*) and *-va* (M.SG.ACC *onogó-*

---

14 The same passage shows a long form *ráiskaa* in *Tixon. d.* (54v), but short *raiska* in *NBKM 709* (32 r).

va, N.SG *onovà*), considering forms with *-zi* phonetic variants redundant for the literary language [Рилски 1989: 178]. The *-zi* suffix was codified first by Xrulev [Хрулев 1859: 34]. We consider any variant of such extended pronoun as a standardized innovation (Var. 1.3).

Standard Bulgarian constructs the demonstratives from two roots: deictically unmarked *t-* (e. g. *tazi*) and *on-* (e. g. *onazi*) marked for distal deixis. Church Slavonic also uses a third root *s-*, marked for proximal deixis [Бончев 1952: 29].[15] The proximal root is also occasionally found in the damaskini, mostly in fixed phrases (e. g. *się staę* F.SG.NOM 'this saint' [*PPS* 66r]). It was not productive anymore in the language [Велчева 1964: 166] and the modern grammars do not even mention it. Proximal demonstratives thus can be considered a destandardized feature (Var. 2.4).

One of the typical balkanisms is the marking of syntactic objects with a second pronoun, although this construction serves different grammatic functions in particular languages (cf. [Tomić 2006: 239]). It is occasionally attested in Church Slavonic sources, but only rarely in *simple Bulgarian* damaskini. Early Bulgarian grammars do not mention this phenomenon at all. Mirčev [Мирчев 1978: 248] states that such marking is confined to Western dialects, and that it is rather avoided in literature. It is, indeed, a feature frequent in Macedonian dialects, and as such it has also been standardized there [Lunt 1952: 38]. According to Tomić [2006: 265, n. 69], Bulgarian grammarians rather tend to restrain the use of the feature. Among our sources, it is indeed most frequent in *NBKM 728* from South Macedonia, but it is also common in later damaskini from the East.[16] In *NBKM 1064*, the scribe systematically marks possessors with a second dative pronoun (*NBKM 1064* 33r):

(8) *fmirísa+   sa   paltá+   mu   na+ unugós   gimitzíe*
Stink.3SG.AOR REFL.ACC Flesh.SG.DEF M.3SG.DAT Of That.M.SG.OBL Sailor.SG.DEF

'the sailor's flesh started to stink'

Such construction is optional in the present-day standard Bulgarian, occuring in emphatic (cf. [Tomić 2006: 269, example 58b]) or emotional [Маслов 1981:303 §3v] environments. For the purposes of our study, we mark such instances with Variable 3.4a. The variable contains the number of such second pronouns.[17]

---

[15] Similar marking of proximal deixis has been standardized in Macedonian, but with another root (e. g. F.SG *ovaa* [Конески 1967: 342]), which is absent in the damaskini sources.

[16] In *NBKM 728* there are 8 instances or 1.67% of the total number of tokens in the text. The frequencies are smaller in *NBKM 1064* (0.51%) and *Berl. d.* (0.32%). A single instance is found in the *PPS* version of *Petka* as well as in [Vuković 1536].

[17] Marked as EXPL ("expletive dependent") in the Universal Dependencies annotation. The head of the dependency should be a noun (identified by a morphological tag), a syntactic object (UD tags NSUBJ or CSUBJ), oblique (OBL) or a nominal modifier (NMOD).

Although the use of short dative pronouns to mark possession is already attested in Old Church Slavonic [Lunt 2001: 149], it is not common in later redactions. It is mentioned by Bončev [Бончев 1952: 28], but redactions from the East Slavic area preferred either adjectival pronouns based on reflexives like *svo* [Смотрицкий 1648: 297], or genitive forms *ego/eę/ixъ* [Миронова 2010: 84]. The genitive forms are occasionally used in some damaskini sources, and systematically in the original *Nedělnik* (e. g. *živénie stýxъ egò* 'lives of His saints' [*Nedělnik* 1806: 184v], but they are not mentioned in such role in the grammars (e. g. [Богоров 1844: 47]). Thus, if a DAT pronoun is used to mark possession, it is reflected as a standardized feature (Var. 1.4); GEN pronouns in this role are considered destandardized (Var. 2.3).

Rarely, short accusative pronouns can also be used to mark indirect objects or possessors. As such forms are not discussed in available grammars, we consider them a non-standardized feature (Var. 3.4b). They appear in some damaskini, and most frequently in particular chapters of *PPS* (52v):

(9) | *maĭkja+* | *ju* | *pade* | *Xrtu* | *na* | *nozé+te* |
| mother.SG | F.3SG.ACC | fall.3SG.AOR | christ.DAT | to | legs.PL.DEF |

‘her mother fell to Christ‘s feet’

Verb morphology exhibits multiple characteristic changes in modern Bulgarian in comparison to earlier varieties. One is the expression of future tense. Church Slavonic shows two basic constructions: a simple form, formed by a present stem of a perfective verb, which is the only one codified by Smotrickyj (e. g. *pročtoù* 'I will read' [Смотрицкий 1648: 197r]); and a complex form, using an auxiliary verb *imati* 'have' and an infinitive form of the main verb (e. g. *ímatъ žíti* 'I will live' [Миронова 2010: 139]). This form originally expressed an obligation. It gradually replaced other complex forms, which used auxiliary verbs *xotěti* 'want' or *načęti* 'begin' [Мирчев 1978: 222; Lunt 2001: 154]). In contrast to Church Slavonic, the 'want'-auxiliary became predominant in the majority of Bulgarian dialects, replacing the simple future form as well. In the damaskini texts up to early 19th century, the future is usually built by a shortened 'want'-verb *šta* with the analytic infinitive, as in the following sentence from *Berl.d.* (185r):

(10) | *štéte+* | *da+ stánete* | *préd'* | *sъdóvište+to* | *xsvo* |
| want.2PL.PRS | to stand.2PL.PRS | in front | judgement seat.SG.DEF | christ's.N.SG |

‘you will stand in front of Christ‘s judgement seat’

The 3SG form of the auxiliary *šte* has been later fossilized. This stage was codified by Neophyte for *a*-stem verbs (e. g. 2PL *šte da dúmate* 'you will say' [Рилски 1989: 128]). For *e*- and *i*-stems he presents the present-day variant,

without the marker *da* (e. g. 2pl *šte píšete* 'you will write'; *šte nósite* 'you will carry' [Ibid.: 136–141][18]). Bogorov provides another variant, with an inflected auxiliary and without the *da* marker (e. g. 2PL *štete pišete* [Богоров 1844: 65]). First, Xrulev's grammar [Хрулев 1859: 44] uses the current variant for all verbs, although he still considers the *da* marker optional. The number of *šte* used as future markers is reflected by Variable 1.5 as a standardized feature.

Use of an analytic construction for infinitive marking is another characteristic feature distinguishing Balkan Slavic from the rest of the Slavic family. Church Slavonic builds the infinitive from aorist stems by attaching a suffix *-ti*, while modern Bulgarian uses the construction similar to the one in Examples (5) and (10): a verb in the present tense following a *da* marker.[19] Such a construction actually does exist in Church Slavonic, too, but with the function of an optative, expressing exhortations and wishes. A classic example can be found in the Lord's prayer [*Mt* 6:9; Lunt 2001: 162; Миронова 2010: 171]:

(11) *da   svętitъ   sę   imę   tvoe*
      to   hallow.3SG.PRS   REFL.ACC name.SG.NOM   yours.N.SG.NOM
      'hallowed be Thy name'

The old infinitive form does not appear in modern Bulgarian grammars, although it is preserved in some isolated dialects (cf. [Мирчев 1978: 235]). Optative is not seen as a separate category of verbal morphology in present-day grammars, and it was described differently in earlier ones.[20] A synthetic infinitive construction does exist in Bulgarian, using an aorist stem without the suffix. Modern grammars agree that the form is only used after specific verbs (e. g. *stiga xodi* 'stop walking' [Мирчев 1978: 235]; *ne možeš go pozna* 'you cannot recognize him' [Маслов 1981: 284]). In the damaskini, the form is used to form the future tense as well, being placed in front of the auxiliary *šta*. Such sentences express a rather conditional meaning [*Tixon. d.* 97]:

(12) *i+   polovína   ot+   crstvoto+   sî.   dá+   štemъ.*
      even half.SG   of   kingdom.SG.DEF REFL.DAT   give.INF want.1SG.PRS
      'we would give a half of our kingdom'

This construction was codified as a specific type of a future tense by Neophyte [Рилски 1989: 129]. Momčilov distinguishes a "future definite" (*bǫdušte*

---

[18] Neophyte also defines a fourth conjugation for verbal *ja*-stems [Rilski 1835[1989]: 143–148], which is practically the same as that of *a*-stems, but without the *da* in future tense.

[19] There are different classifications of the *da* marker (cognate of the English conjunction *to*; cf. [Derksen 2008: 94]), e. g. a conjunctive [Маслов 1981: 286], a subordinating or modal particle [Friedman 2006: 661], or a subjunctive marker [Tomić 2006: 414].

[20] Optative is described in the earliest grammars: Neophyte provides a construction with marker *danò* and a verb in imperfect tense (e. g. *danò prodúmaxъ* 'may I have spoken' [Рилски 1989: 130]; cf. also [Богоров 1844: 67, Маслов 1981: 287, 334]).

oprĕdĕleno), denoting events happening in a given future moment, and a "future indefinite" (*neoprĕdĕleno*), when the moment is not given [Момчилов 1868: 52f.]. The former is expressed by the fossilized *šte* followed by the main verb in present tense, the latter by the construction using the synthetic infinitive as in Example (12). Although a similar distinction can be observed in Serbian (e. g. [Tomić 2006: 486]), the specific grammatical function of the indefinite future tense was called into question by Andrejčin [Андрейчин 1944: 252], who considered it an archaic variant of the "definite" future tense. More recent grammars (e. g. [Маслов 1981: 236; Radeva 2003: 74]) describe these forms (if at all) in a similar way.

Thus, Variable 1.6 reflects the presence of *da* markers dependent on auxiliary verbs, after which the use of synthetic infinitive is optional.[21] The number of old infinitives with the *-ti* suffix is counted by Variable 2.5. Instances of "future indefinite" tense constructions are counted by Variable 3.3 as a non-standardized feature.

Past tenses are morphologically similar in Church Slavonic and the modern standards of Bulgaria and Macedonia. However, generalizations and phonetic shifts levelled the difference between morphemes. Only 2/3SG forms are different between the aorist and imperfect; IMPF.1SG developed secondary forms and IMPF.PL forms were generalized for both tenses [Конески 1967: 420; Мирчев 1978: 212f.]. In the damaskini, verbs in plural already use the imperfect forms only. Specific aorist forms are attested (e. g. *pogrebóste* 'you buried' [*PPS* 67v], but their use is not systematic.[22] Neophyte has also codified only the innovative forms (e. g. AOR/IMPF.2PL *dúmaxte* 'you spoke' [Рилски 1989: 126]). Thus, old AOR.PL endings are handled as a destandardized feature (Var. 2.6) in our analysis.

4.2. Graphic Features

Alphabets in the damaskini sources slightly differ from the standards of the Church Slavonic and Greek literature. Table (2) show the characters common in this literary tradition, adapted for the Unicode standard. It does not reflect all the regularly employed allographs, like the initial vowel variants (<є>, <o>, <ꙍ>), the broad <m> and the space-saving <7> variants of <т>, nor ligatures and superscript letter variants:

---

[21] These include verbs with meanings 'want' (*šta*, *xoštu*, *xoču*; cf. Mirčev 1978:235), 'have' (*ima*, *nja+ma*) and 'begin' (*načena*, *počna*, *podbra*, *vzema*; cf. [Lunt 2001: 154]), as well as negative commands (*nedei*, *prestana*, *stiga*).

[22] For example, Punčo uses three forms for AOR/IMPF.3PL in the *Legend of Joseph, son of Rachel* (*PPS* 71r–83r): 83 times *-xu*, 10 times *-xa*, and only twice *-ša*. The two are likely copied from an East Slavic source: Resava orthography used *-še* ending in this position (OCS *-šę*).

Table 2

Damaskini alphabets and Latin transcription[23]

а	б	в	г	д	е	ж	ѕ	з	ı	ї	и	й	к	л	м	н	о	п
α		β	γ	δ	ε			ζ	ι		η		κ	λ	μ	ν	ο	π
a	*b*	*v*	*g*	*d*	*e*	*ž*	*ź*	*z*	*ı*	*ï*	*i*	*ĭ*	*k*	*l*	*m*	*n*	*o*	*p*

р	с	т	𐌏	оу	ф	х	ѡ	ѿ	ч	ц	ш	шт	щ	ъ	ы	ь	ˇ	ѣ
ρ	σ	τ	𐌏	ου	φ	χ	ω						ς					
r	*s*	*t*	*u*	*ou*	*f*	*x*	*w*	*wt*	*č*	*c*	*š*	*śt*	*št*	*ъ*	*y*	*ь*	*,*	*ě*

ѩ	ѥ	ю	ѫ	ѧ	ѯ̃	ѱ	ѳ	ѵ	ц̑
ѩ		ю		ξ	ψ	ϑ	υ		
ja	*je*	*ju*	*ǫ*	*ę*	*ˀ̣ʒ*	*ψ*	*θ*	*v*	*u̯*

Early damaskini show many rules of the above-mentioned Resava orthography. They use both jers: "orthographic" <ь> as the silent marker of word boundary and syllabic resonants, and "phonetic" <ъ> in prepositions, which are written together with the following word. As in this orthography, vowel letters may have an accent and/or spirit, while *pajerčik* (< ˇ > or < ˈ >) may appear above consonants instead of a following jer. Jat <ѣ> is written in its etymological place. Elsewhere, it shows influences of the vernacular: individual scribes sometimes employ their own modifications. The <ы> occurs not only on etymological places, but also as a variant of /i/. The old back nasal is regularily replaced by variants of <𐌏>, and also by both jers (preferably <ь> in the 17th century, <ъ> later) and <a>. The letter <ѫ>, also called big jus, scarcely appears, as well. For example, the main verb in the following sentence from *Tixon.d.* (95) is reflected in other editions in the following way:

(13)	*šte*	**búde**	*na+*	*krásnyi*	*i+*	*ne+veštestóvnyi*	*rái*
	FUT	be.3SG.PRS	at	beautiful.M.SG	and	immaterial.M.SG	paradise.SG
	'[your soul] shall be in the beautiful and immaterial paradise'						

Tixon.d.	*бꙋ́де*
Trojan d.	*бъ́де*
Ljub.d.	*бъ̈де*
NBKM 709	*ба́де*

The same form is written as *бѫде* and *баде* elsewhere in the version of the text in the damaskin of Koprivštica [*Kopr.d.* 11]. All these letters represent the

---

23 The character set is also used in the examples in this article, with omegas (*w*) replaced by *o* for reader's convenience. In this section, transcriptions will reflect the original script.

middle vowel /ă/. 19th century grammarians, trying to find an ideal representation of the vowel in the script, introduced various letters—<ă> and <ӑ> [Берович 1824], big jus <ѫ> [Рилски 1989: 123], using jers in positions, where they occur in Church Slavonic: e. g. 'first' is written as *пьрво* by Xrulev [Хрулев 1859: 12], but *пѫрв-* by Bogorov (e. g. in the very title of [Богоров 1844]). Finally, the reform of 1945 tried to unify its writing in Bulgarian under <ъ>, but the reform stumbled on the decision to discard orthographic jers at the end of words. Therefore, it is written as <a> in word-final positions, and as <я> after palatal consonants. To capture attempts of earlier literature to cope with the middle vowel problem, we reflect the use of a single letter for /ă/ in non-final, non-palatal positions as Variable 1.7a.[24] The writing of word-final orthographic jers is considered as a destandardized practice (Var. 2.9).

Another orthographic problem was the writing of the phonem /i/. Already Constantine-Cyrill adopted multiple variants rendering this phoneme from the Greek alphabet, which had been preserved as a part of orthographic tradition despite earlier phonetic shifts. A new letter (actually a digraph) has been established to reflect the Common Slavic *y*, which in later South Slavic merged with **i*. The damaskini literature took no less than four graphemes from the Resava orthography—<и>,[25] <ї>, <v> and <ы>—employing them according to the etymologic principle, phonotactic rules or free will. Of these four, the Cyrillic iota or <ї> was traditionally written for /i/ before other vowels and diphthongs [Ягич 1895: 415]. It was simplified to <i> by Bogorov and used up to Drinov's criticism [Дринов 1911: 285f.], after which it fell out of use. To analyze the practice in earlier literature, we count the writing of /i/ with a single letter as Variable 1.7b. As the writing of <ы> was supported by the East Slavic varieties (including the local redactions of Church Slavonic like that of Smotrickyj), we list it among the destandardized features (Var. 2.7b).

The writing of the sequence /ja/ is another aspect, which distances not only Church Slavonic from Standard Bulgarian, but even more the single redactions of the former. Before the reform of 1945, two letters were used for /ja/: the <я> and the historical jat or <ѣ>. As the post-reform Standard Bulgarian, the Resava system had a single letter for it: the digraph <ꙗ>. Constantine of Kostenets, author of the standard, considered the jat an archaic letter pronounced /e/ or /je/ [Ягич 1895: 402]. It was likely pronounced dif-

---

[24] Only two among our sources fulfill this requirement: Xrulev's [*Nedělnik* 1856] and *NBKM 1064*, which uses the Greek alpha letter for /ă/. Still, Xrulev does not write the elsewhere preferred big jus in sequences with resonants (e. g. PRS.3SG *смѫрди* 'stinks', *дѫлбоко* 'deep'), as in his grammar. Since these resonants were likely considered syllabic in Church Slavonic literature (and schooling), such instances were disregarded.

[25] Graphemic status of <й>, the "short иже" (*i kratko*), is unclear in earlier texts. Among Bulgarian grammarians, the semivowel character of <й> is mentioned by Bogorov [Богоров 1844: 4], but it was not until Drinov [Дринов 1911: 285] that it was listed as a separate letter.

ferently in the dialect of the damaskini translator, who uses it occasionaly on the place of /ja/ or /jă/, too (e. g. PRS.3PL *чúнѣтъ* 'they cause' in *Tixon.d.* 95).[26] Furthermore, the damaskini use the mentioned small jus or <ѧ> for the same sequences.[27] Neophyte's grammar adopted the practice established by Smotrickyj [Смотрицкий 1648: 46r], using the letter <ꙗ> as the initial and <ѧ> as word-internal or final variant of /ja/. The jat was used instead of <ѧ> in etymological positions. With the adoption of the *graždanka* font (also seen in Momčilov's grammar), <ꙗ> and <ѧ> were replaced by the letter <я>. The 1945 reform replaced the jat <ѣ>, according to phonotactic rules, by <я> or <e>.

As according to the 1945 orthography the <я> in a word-final position can also denote the sequence /jă/, we consider the use of a single letter for both /ja/ and final /jă/ a standardized feature (Var. 1.7c). The use of both <ѣ> (Var. 2.7a) and <ѧ> (Var. 2.7c) are measured as two destandardized practices. The use of a single letter for the /o/ phoneme is thus considered a standardized feature (Var. 1.7d). The use of four special letters for Greek loanwords—<v>, <ꙁ>, <ψ> and <ѳ>—are taken as a destandardized practice (Var. 2.8). The use of <џ>, which has not been accepted by Church Slavonic grammarians, can be considered non-standardized (Var. 3.6).

Another graphic feature, distancing Standard Bulgarian from Church Slavonic, was the removal of accent markers (Var. 1.8), which can be first seen in Bogorov's grammar. Earlier literature, written before the Neophyte's grammar, prefers four different markers for accents (Var. 2.10a) and at least one spirit on word-initial vowels (Var. 2.10c). The writing of breves on syllable-final vowels other than <й> was a practice already abolished by Neophyte. The use of a simplified accentuation, e. g. with a single accent mark, is considered a non-standardized feature (Var. 3.7).

Earlier literature often writes monosyllabic words, like conjunctions and prepositions, together with longer words, characterized by a single accent per such orthographic "words" (e. g. *инамúсïа* or *i+na+mísïa* 'and in Moesia'; *Tixon.d.* 94). Most of them were separated, as can already be observed in Neophyte's grammar (Var. 1.8a). The reflexive pronouns remained to be written together with the preceding verb up to Momčilov's grammar (Var. 1.8b). As a standardized practice we also reflect the use of Arabic numerals (Var. 1.10).

---

[26] The use of jat reflects the struggle to create a supradialectal norm by the Bulgarian grammarians. The vowel marked by the jat in OCS was reflected as /e/ in the western and as /ja/ in the eastern dialects. The shifts were documented first in the 12th century [Мирчев 1978: 119]. The etymological rule employed by Resava redaction was practically reiterated by Drinov [Дринов 1911: 282].

[27] The Greek-script damaskin *NBKM 1064* reflects all three letters with epsilon (e.g. <ꙗ> in 3PL *κȣπαετ* 'they dig', *Ljub. d.*: *копаіатъ*; <ѣ> in *ποσίτζιιε σφέτ* 'all over the world', *Tixon.d.*: *по сúчкы свѣтъ*; <ѧ> F.3SG.ACC *ε*, *Tixon. d./Ljub. d.* F.3SG.ACC *ѧ*).

The use of lexicalized abbreviations (e. g. SG.OBL $\vec{xa}$ 'of Christ'; *Tixon. d.* 99) is considered a destandardized feature (Var. 2.11). The use of a non-Cyrillic (e. g. Greek or Latin) script is reflected in Variable 3.5.

5. Sources

We have analyzed the spread of the aforementioned features on a corpus of twelve texts dated from the 16th to the 19th century, representing two text traditions from the Balkan Slavic linguistic area—*Life of St. Petka* and *Legend of St. Thaïs*. Generally, the texts preserve the content and narrative structure, and thus linguistic differences can easily be compared between separate sources (print editions, manuscript collections) of the text. The sources used are a part of the digital corpus of pre-standardized Balkan Slavic.[28] Relations between the sources of the *Life of St. Petka* can be seen in Figure (1):
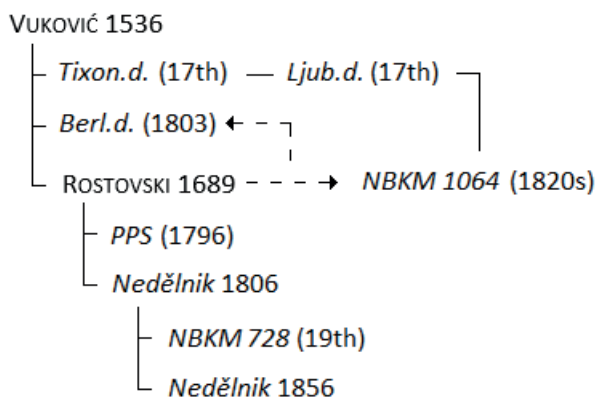

Figure 1. Relations between the sources used for *Life of St. Petka*[29]

The other text tradition is smaller, comprising only two versions of the *Legend of St. Thaïs*, translated from a Greek text by Josif Bradati in the 1740s. While the first text tradition covers a considerably broad area (including texts from Serbia and Kiev, various damaskini traditions and modern prints, 16th–19th century), the second one includes two sources closer to each other (Bradati writing in Samokov; Punčo in Mokreš near Danube, likely paraphrasing a transcript of Bradati's translation). The sources are listed with the approximate date of composition or publication, classification of the language (according to the categories defined above), typographic method, text and size in tokens in Table (3).

---

[28] See Šimko 2021 for a detailed description of the sources.

[29] Damaskini sources in the figure may represent hypothetical protographs of respective editions.

Table 3

Overview of sources

Source	Date	Language	Type	Text	Size[30]
Vuković 1536	1536	CS	printed	*Petka*	2222
Tixon. d.	early 17th	*simple* BG	manuscript	*Petka*	2472
Ljub. d.	late 17th	*simple* BG	manuscript	*Petka*	2503
Rostovski 1689	1689	CS	printed	*Petka*	1336
NBKM 328	1749	*Slaveno*-BG	manuscript	*Taisĭa*	891
PPS	1796	NW-BG dialect	manuscript	*Petka*	584
				Taisĭa	984
Berl. d.	1803	*simple* BG	manuscript	*Petka*	3120
Nedělnik 1806	1806	*Slaveno*-BG	printed	*Petka*	1905
NBKM 1064	1820s	east-BG dialect	manuscript	*Petka*	3340
NBKM 728	19th	a MK dialect	manuscript	*Petka*	686
Nedělnik 1856	1856	standard BG	printed	*Petka*	1249

Orthographic features were based on the analysis of originals or their fac-similes. Grammatical features were studied on annotated transcripts of the sources. As the sources use various scripts, they were transcribed into a diplo-matic set of Latin UTF-8-compatible characters. Each token is marked by tags reflecting its morphological structure and syntactic relations.[31] By comparing both grammar and orthographic features, we can quantify the differences be-tween individual sources. We have focused on two hypotheses:

(I.) First, we assumed the more orthographic rules are copied from an original, the more influence of the original can be expected in the grammar in spite of language change. Works orthographically similar should be gram-matically similar, too.

(II.) Second, we assume the modern standard developed from the lan-guage of the damaskini. If that is the case, *Slavenobulgarian* sources also should be placed somewhere between the sources representing the modern standard and Church Slavonic.

6. Analysis

In our analysis, the features listed in the Table 1 were represented either as fre-quencies (counted as absolute number of occurences divided by the size of the text in tokens) or as binary variables (considered "true" or "false" according to their presence or absence in the whole text). The method of measurement was

---

[30] As mentioned above, some lexical units like articles or negative prefixes were handled as separate tokens as well. Interpunction and similar markers were not considered.

[31] Also see [Šimko 2021] for a detailed description of the sources.

chosen according to the nature of these features (orthographic or linguistic) and their representation in the corpus. In general, morphologically and syntactically relevant features, which can be identified in the annotation of our source texts (even if their role could not be unambiguously determined), were counted as frequencies: the number of occurences divided by the length of the text (total amount of tokens). This allows us to compare the sources despite differences in size.

Table 4

Features counted for frequency

Standardized innovations	Slavonicisms/archaisms	Not standardized features
1.1. Postnominal article	2.1. CS nominal inflection	3.1. Inflected articles
1.2. Postadjectival article	2.1a. Non-NOM endings	3.2. Articled short form adjective
1.2a. M.SG adj. -*ija*	2.1b. M.SG -*a*	3.3. "Future indefinite" tense
1.3. Ext. demonstrative	2.1c. F.SG -*u*, -*ь* or -*ǫ*	3.4. Differential object marking
1.4. DAT possessive pronoun	2.1d. M.SG -*u*	3.4a. Object doubling
1.5. *šte* particle for FUT	2.2. Long-form adjective	3.4b. 3SG.ACC for indirect objects
1.6. Analytical infinitive marking	2.2a. M.SG adj. -*ij*	
	2.3. GEN possessive pronoun	
	2.4. Proximal deixis marking	
	2.5. Synthetic infinitive marking	
	2.6. Old 2/3PL aorist forms	

Orthographic features were measured on the basis of the whole text as binary variables. The presence of a single instance of specific letters (especially the archaic ones and <ц>) suffices for the variable to be "true":

Table 5

Features reflected as Boolean values

Standardized innovations	Slavonicisms/archaisms	Not standardized features
1.7. Unified orthography	2.7. Archaic letters	3.5. Non-Cyrillic script
1.7a. Non-final/non-palatal /ă/	2.7a. Use of <ѣ>	3.6. Specific letter for /dž/
1.7b. /i/ in all positions	2.7b. Use of <ы>	3.7. Simplified accentuation
1.7c. /ja/ and final /jă/	2.7c. Use of <ѧ> for /ja/	
1.8. Separation of unaccented words	2.8. Loanword-specific letters	
1.9. No accent markers	2.9. Word-final jers	
1.10. Arabic numerals	2.10. CS accentuation	
	2.10a. Use of all four markers	
	2.10b. Breve on syllable-final vowel	
	2.10c. Writing of *spiritus lenis*	
	2.11. Lexicalized abbreviations	

The classification of individual variables as "standardized" or "archaisms" does not play any role in the analysis itself. Works which we consider as protographs or older sources, as shown in Figure (1), do not necessarily score high among all archaic features (or low among the innovative or non-standardized ones). For example, both our Church Slavonic sources use short demonstratives after adjectives (e. g. *svĕtla ónà* 'the shiny [queen]' [Vuković 1536]; Var. 1.2), considered an innovation. The frequency of this feature in [Ibid.] and [Rostovski 1689] is close to earlier damaskini, while also some later sources show lower values.[32] Thus any source can be considered a reference point for comparison, in a similar way as a prototypical dialect in dialectology (cf. [Vuković 2020: 3]).

On the basis of our first hypothesis (the assumed relation between orthography and grammar) we would expect small differences between the analyses based on the two groups of variables. If the difference would be big, the standardization of orthography would be a process rather independent of the development of grammar or phonetics. The level of education would presumably be able to diminish the influence of the scribe's vernacular. On the basis of the second hypothesis (damaskini basis of the modern standard) we would expect results with two poles: Church Slavonic sources [Vuković 1536, Rostovski 1689] and presumably conservative works [*Nedĕlnik* 1806] on the one side, and works close to the modern standard (e. g. [*Nedĕlnik 1856*]) on the other.

Two statistical methods were used to measure the mutual distances between our sources. The distance represents the amount of variables with similar (if they are closer) or different (if more distant) values. For the selected features, the sources placed close to each other can be considered similar. They may also form clusters, which can then be interpreted as specific orthographic or linguistic varieties. First, we used the binary distance, using 20 Boolean variables based on graphic features of the texts. The other analysis concerned 22 float variables representing percentual frequencies of occurence of the selected linguistic features. Since the values of these variables are considerably small, we used the method of Canberra distance, which is based on the sums of series of fraction differences between the particular data sources (Kaur 2014). For our first hypothesis we expected similar results in both analyses: the distances between the sources should not vary much. For the second hypothesis we expected two or more clusters of sources, with works of a transitional or dialectal character in-between in both analyses.

The mutual distances can be represented using two-dimensional diagrams as an abstract map.[33]

---

[32] Frequencies given for the Var. 1.2 are: 0.31% in [Vuković 1536], 0.45% in both *Tixon. d.* and [Rostovski 1689], in *Ljub. d.* 0.64%. The values are comparably low in *PPS*: there is only one post-adjectival article in each of the texts from this source.

[33] The study was done in R v3.6.2 using the function DIST (URL: https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/dist; 5.5.2020). Diagrams
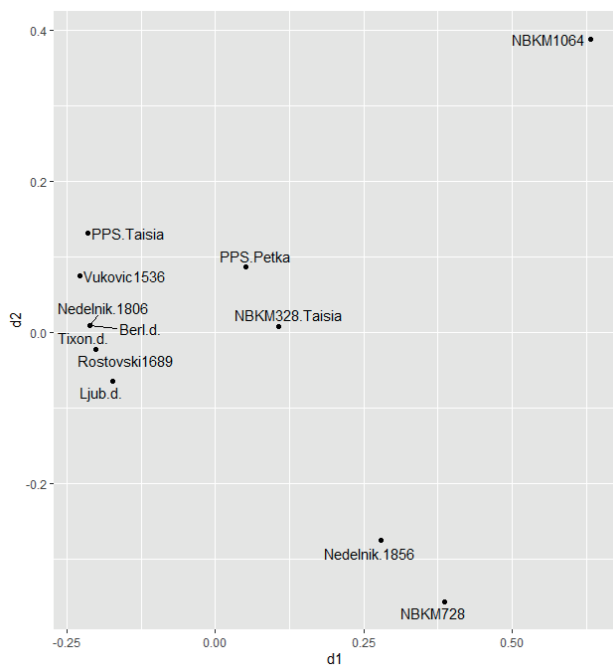
Figure 2.

Binary distances between the sources based on orthographic features

Figure 3.

Canberra distances between the sources based on grammatical features

were produced using the package GGPLOT2 v3.3.0 (URL: https://ggplot2.tidyverse. org/; 5.5.2020).

Figure (2) shows mutual distances based on similarities and differences in orthography. The damaskin *NBKM 1064* stands isolated from the rest due to its Greek script. The older damaskini (*Tixon. d.*, *Ljub. d.*, but also the later *Berl. d.*) are tightly clustered together with Church Slavonic sources ([Vuković 1536] and [Rostovski 1689]) and the [*Nedělnik* 1806]. Concerning graphic features, the older damaskini and the original *Nedělnik* do not represent a transitional stage between the Middle and Modern Bulgarian literature, but rather between the Resava and East Slavic redactions of Church Slavonic. Bradati's *NBKM 328* and *PPS* show some deviations from the arguably dominant Church Slavonic damaskini orthography, but at least in Punčo's case, these are not very systematic, given the distance between the two *PPS* texts. The orthography is similar between the 1856 edition of the *Nedělnik*, representing Bogorov's standard, and the late Macedonian damaskin *NBKM 728*.

Figure (3) shows the distances based on grammatical features. The sources form three clusters: (a) on the left, including the older damaskini, two later ones (*Berl.d.* and *NBKM 1064*) and the [*Nedělnik* 1856]; (b) one in the upper middle, including the sources from Western Bulgaria (*NBKM 328*, *PPS*) and the [*Nedělnik* 1806]; (c) and, finally, the Church Slavonic sources in the lower right corner. This shows a clear linguistic similarity between the *simple Bulgarian* of the damaskini and later dialects from both ends of the Balkan Slavic area (*NBKM 728* and *1064*), as well as the 1850's Bulgarian standard [*Nedělnik* 1856].

This distribution can be observed in spite of orthographic conservatism of the damaskini and textual relations. But let us compare the respective clusters from Figure (3) to observe mutual distances within them. When we exclude the Church Slavonic cluster (c) from the analysis, two sources become more isolated from respective clusters as shown in Figure (4): *NBKM 728*, our only source from Macedonia, and Bradati's *NBKM 328*. The "standardized" [*Nedělnik* 1856] remains close to most of the damaskini sources from the Eastern Bulgarian dialectal areas. *NBKM 728* also shows itself as linguistically different in Figure (5), where we exclude the cluster (b):

The (I.) first of our hypotheses is not supported by our test. The clusters in both analyses contain different sources. Orthographic similarity does not imply grammatical interferences by the original. The distances between *NBKM 728* and the cluster containing *simple Bulgarian* sources plus the [*Nedělnik* 1856] in Figures (4) and (5) implies stronger influence of dialectal differences. Concerning the (II.) second hypothesis, a striking similarity can be observed between the *simple Bulgarian* damaskini and the 1850's standard of [*Nedělnik* 1856] in the analysis of linguistic features. We can see there is a clear (likely dialectal) similarity between the standard of [*Nedělnik* 1856] and the linguistic norm of *simple Bulgarian*. The texts representing the *Slavenobulgarian* variety are distant from both the Church Slavonic and from the cluster including the other texts.
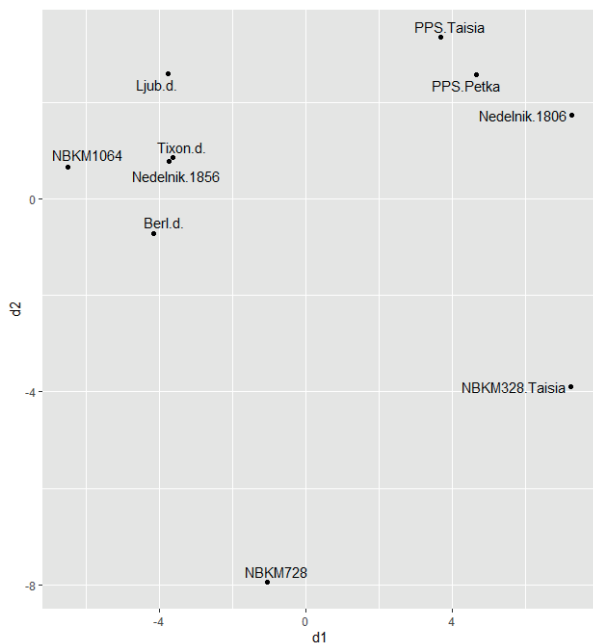
Figure 4.

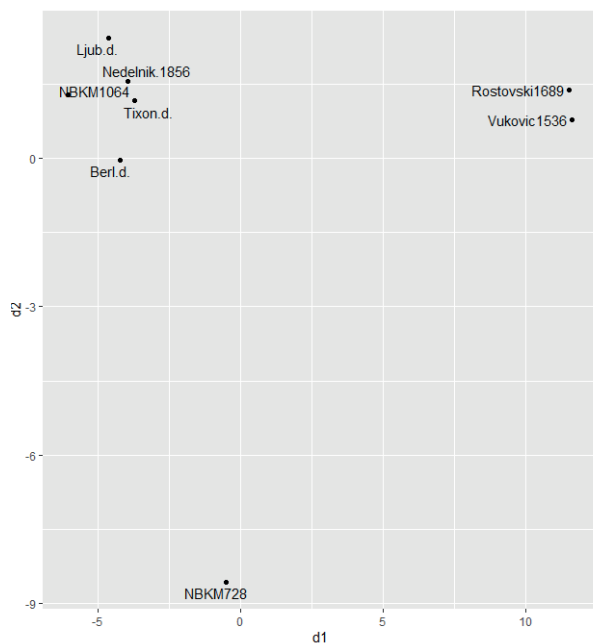Canberra distances based on grammatical features (excl. CS sources)

Figure 5.

Canberra distances based on grammatical features (excl. Slaveno-BG sources)

7. Discussion

Our study copes with the question, how prescriptive norms changes the language of the literature. Can we consider works, produced under the influence of an authoritative document (e. g. a grammar used in mass schooling) containing artificial or archaic rules, to be a reliable source for the development of the spoken language as well? Our task was to examine the usefulness of this dichotomy between the pre-standardized and standardized literature: the former reflecting a presumably "natural" language, the other "tainted" by arbitrary decisions of grammarians. This is an important question, especially within the Balkan Slavic area, where polyglossy and multiple literary norms existed along each other for centuries. But did the prescriptivists have the power to shape the grammar of the written language freely, or were they mere students, reiterating the common practice in terms of modern linguistics?

As discussed above in section 2., the Bulgarian national awakening of the 19th century began with a grammarian battle between the faction of "tyrants", proposing the *Slavenobulgarian* language of Paisius and Josif Bradati, like Neophyte of Rila, against the "demagogic" faction, promoting the *simple Bulgarian* of the damaskini tradition, like Beron and Bogorov. The analysis above aims to shed light on some of the elements mentioned in this conflict, or the possible strategy of the winner. The "demagogues" presented themselves as a movement reflecting the contemporary trends of making the language of literature closer to the vernacular. *Slavenobulgarian*, on the other hand, was considered an artificial variety [Керемедчиев 1943: v], aiming at the preservation of Church Slavonic (or, generally, non-Balkan Slavic) features. Concerning our analysis of the linguistic features, it is rather the faction of "demagogues", represented here by Xrulev's [*Nedělnik* 1856], which follows the older literary tradition, namely that of the damaskini. The relevance of this result is, of course, limited by the size of the sample, as well as the representative value of the selection. Future studies including more texts or text traditions will show relations between literary varieties more clearly. The result also does not say anything about "natural" or "artificial" character of the norm Xrulev followed. The analysis allows us to assume that the *simple Bulgarian* grammatical norm seems to have been quite stabilized even before the codification of the first "modern" grammar by Bogorov. This can be observed in spite of the orthographic variety of the sources following this norm.

What can be said then of this "artificial" *Slavenobulgarian*? Among our sources, this variety is represented by three sources (*NBKM 328*, *PPS* and [*Nedělnik* 1806]), all of which form the cluster (b) in Figure (3) above. Two of them were produced by men from Eastern Bulgarian dialectal areas—Josif Bradati from Elena and Sophronius from Kotel (both in the Eastern Subbalkan area). Yet the dialectal background is only one of the factors. Bradati

travelled extensively in the West, studying (and adapting his texts to) the language of the people around Rila, Vratsa and Eastern Macedonia. Sophronius was educated in Kotel, but most of his literary activity comes from his stays in Vidin and Vratsa in Northwest Bulgaria. Both these areas were parts of the former Peć Patriarchate, where Church Slavonic was likely still used. But it is also reasonable to assume that in their activities both Josif Bradati and Sophronius were influenced by the local dialects of their immediate audience. From the point of view of these dialects, their native Eastern speech likely sounded too foreign to the audience—with all the articles and the lack of inflection. The only one of these authors who wrote in his native dialect was Punčo [Шаур 1970: 62]. As the sources by Bradati and Sophronius are close to his texts from the lingustic point of view, we may assume the *Slavenobulgarian* may indeed have been based on dialects of the Northwest.

This basis—both from the point of view of geography and the number of potential recipients—was likely much smaller than that of the standard proposed by the "demagogues". The cluster (a) in the Figure (3) contains sources which can be reliably attributed to the various locations in a wide area from Macedonia (if we include *NBKM 728*) to Sliven—the majority of the whole Balkan Slavic area. Furthermore, the changes the "tyrants" applied to the literary language were actually more innovative than the standard proposed by the "demagogues". They parted ways both with the damaskini and with the Church Slavonic literature. It was not so much a more conservative alternative, but rather a model based on different dialectal area and without an old literary tradition. *Slavenobulgarian* was also not normatively stabilized enough: the lingusitic differences between *NBKM 328* and other sources of this tradition, as seen in Figure (4), are not small. Finally, historical events in the Northwest in the second half of the 18th century (like the abolishment of the Peć Patriarchate in 1767 and the rebellion of Osman Pazvantoğlu in 1790s) did not provide good conditions for cultural and political integration, and hence the propagation of an overarching literary norm. Although many cultural centers in the Western area like Samokov, Vratsa or the Rila monastery produced followers for Bradati, in the long run they remained an isolated school. Bradati's students from the more eastern areas adapted his translations to a language closer to that of the damaskini.

8. Conclusion

Our study offers a method for measuring the spread of norms of a literary language among the writing community. The literary networks were likely not based on common schooling or source texts, but rather on mutual comprehensibility. Norms would be accepted so far as the texts produced under their influence were reproducible. When the text started to sound foreign, it

would provoke correctors—like Bradati's students in the East or Xrulev, when adapting *Nedělnik* to a language more similar to his Eastern Bulgarian dialect. The idea of "mistakes" which have to be corrected did indeed exist even before Neophyte's grammar. From an evolutionary point of view, *Slavenobulgarian* was a new form, which lacked appeal or compehensibility in the East—at least in comparison to the *simple Bulgarian* texts in the Northwest. While Josif Bradati and Sophronius were certainly very active writers, contributing greatly to the literature and learning of their time, they still remained too regionally inclined. In their time, the *simple Bulgarian* of the damaskini was already a literary norm affecting multiple dialectal areas, even in spite of the lack of schooling and orthographic experiments.

In short, *simple Bulgarian* of the damaskini was indeed a kind of a standardized language. The competing *Slavenobulgarian* did not stand a chance. The standardization in the 19th century was more or less a mere orthographic reform. Thus, newer literature can also be taken as a relevant source for the developments of Balkan Slavic dialects.

Bibliography

Primary sources

Abagar
Станиславов Ф., [*Абагар*], Свети Град, 1651.

Berl.d.
[Berlin damaskin], Cracow, Library of Jagellonian University, sign. Slav. fol. 36.

Kopr.d.
Милетич Л. (ед.), *Коприщенски дамаскин: новобългарски паметник от XVII век* (= Български старини, 2), София, 1908.

Loveč d.
Ловешки дамаскин, XVII в., Digital edition by National Library of Bulgaria. Available online: https://histdict.uni-sofia.bg/textcorpus/show/doc_192 (accessed 14.10.2021).

Ljub.d.
[Ljubljana damaskin], *Zbornik „Damaskin", bolgarski*, Ljubljana, National and University Library of Slovenia, sign. NUK Cod. Kop. 21.

NBKM 328
Дамаскин от 1750 г., Sofia, National Library of Bulgaria, sign. НБКМ 328.

NBKM 709
Дамаскин (сливненски) от XVII в., Sofia, National Library of Bulgaria, sign. НБКМ 709.

NBKM 728
Откъслек от простонароден дамаскин, век XIX, Sofia, National Library of Bulgaria, sign. НБКМ 728.

NBKM 1064
Дамаскин, от края на XVIII или началото на XIX в., Sofia, National Library of Bulgaria, sign. НБКМ 1064.

Nedělnik 1806
Софроний Врачански, еп., *Кириакодромион сиреч Неделник – Поучение*, Римник, 1806.

Nedělnik 1856
Софроний Врачански, еп., *Евангелие поучително*, Т. Т. Хрулев, пр., Нови Сад, 1856.

PPS
Поп-Пунчов сборник от 1796 год., Sofia, National Library of Bulgaria, sign. НБКМ 693.

Rostovski 1689
[Димитрїй Ростовскїй, еп.] [*Книга житий святых...на три м[е]с[я]цы первыя*], [Киев], 1689.

Svišt.d.
Милетич Л., ед., *Свищовски дамаскин: новобългарски паметник от XVII век* (= Български старини, 7), София, 1923.

Šimko 2021
Annotated Corpus of Pre-Standardized Balkan Slavic Literature, ver. 1.1. Slovenian language resource repository CLARIN.SI (http://hdl.handle.net/11356/1368).

Tixon.d.
Демина Е. И., *Тихонравовский дамаскин: болгарский памятник XVII в.: исследования и текст*, 2, София, 1972.

Trojan d.
Иванова А., *Троянски дамаскин: български паметник от XVII век*, София, 1967.

Thēsauros 1751
Δαμασκηνος Στουδιτης, *Θησαυρος Δαμασκηνου*, Ενετια, 1751.

Vuković 1536
[*Zbornik za putnike*, Venecia, 1536].

Аргиров 1895
Аргиров С., Люблянският български ръкопис от XVII век, *Сборник за народни умотворения, наука и книжнина*, 12, София, 1895, 463-560.

Иванов 1914
Иванов Й., стък., *История Славяноболгарская*, София, 1914.

Наредба 1945
Наредба-закон за правописа, *Държавен вестник*, 47, 27.2.1945.

Упътване 1899
Упътване за общо правописание, София, 1899.

Literature

Anderson 2006
Anderson B., *Imagined Communities: Reflections on the Origin and Spread of Nationalism*, London, New York, 2006.

Barth 1969
Barth F., *Ethnic Groups and Boundaries: The Social Organization of Culture Difference*, Boston, 1969.

Belić 1905
Belić A., *Dijalekti istočne i južne Srbije*, Beograd, 1905.

Bourdieu 1991
Bourdieu P., *Language and Symbolic Power*, tr. by G. Raymond and M. Adamson, Oxford, 1991.

Derksen 2008
Derksen R., *Etymological Dictionary of the Slavic Inherited Lexicon* (= Leiden Indo-European Etymological Dictionary Series, 4), Leiden, Boston, 2008.

Fielder 2019
Fielder G. E., The Semiotics of Ideology: The Definite Article Rule in Bulgarian, *Balkanistica*, 32/2, 2019, 45–70.

Friedman 1986
Friedman V. A., Macedonian Language and Nationalism During the Nineteenth and Early Twentieth Centuries, *Macedonian Review*, 16/3, 1986, 280–292.

——— 2006
Friedman V. A., Balkans as a Linguistic Area, K. Brown, ed., *Encyclopedia of Language and Linguistics*, 2nd ed., 1, Oxford, 2006, 657–672.

——— 2017
Friedman V. A., Language Ideology and Language Change in the Balkans: A View from the Early Twenty-First Century, *Die Welt der Slaven*, 62/1, 2017, 1–21.

Fuchsbauer 2010
Fuchsbauer J., *Die Übertragung der Dioptra ins Slavische* (Dissertation, Universität Wien, 2010).

Haarmann 1976
Haarmann H., *Aspekte der Arealtypologie: die Problematik der europäischen Sprachbünde*, Tübingen, 1976.

Handelman 1977
Handelman D., The Organization of Ethnicity, *Ethnic Groups*, 1(3), 1977, 187–200.

Hastings 1997
Hastings A., *The Construction of Nationhood: Ethnicity, Religion and Nationalism*. Cambridge, 1997.

Hobsbawm 1990
Hobsbawm E., *Nations and nationalism since 1780: programme, myth, reality*. Cambridge, 1990.

Hroch 1985
Hroch M., *Social Preconditions of National Revival in Europe: A Comparative Analysis of the Social Composition of Patriotic Groups among the Smaller European Nations*, B. Fowkes, trans., Cambridge, 1985.

Irvine, Gal 2000
Irvine J. T., Gal S., Language Ideology and Linguistic Differentiation, P. V. Kroskrity, ed., *Regimes of language: Ideologies, polities, and identities*, Santa Fe, 2000, 35–84.

Jagić 1877
Jagić V., Kako se pisalo bugarski prije dvjesti godina, *Starine*, 9, Zagreb, 1877, 137–171.

Karadžić 1974
Karadžić V. S., *Kleine Serbische Grammatik* (= Sagners Slavistische Sammlung, 1), übers. J. Grimm, München, 1974 [repr. 1824].

Kaur 2014
Kaur D., A Comparative Study of Various Distance Measures for Software fault prediction, *International Journal of Computer Trends and Technology*, 17/3, 2014, 117–120.

Leake 1814
Leake W. M., *Researches in Greece*, London, 1814.

Lunt 1952
Lunt H. G., *A Grammar of the Macedonian Literary Language*, Skopje, 1952.

——— 2001
Lunt H. G., *Old Church Slavonic Grammar,* 7th ed., Berlin, 2001.

Miklošić 1871
Miklošić F., *Trojanska priča*, bugarski i latinski, Zagreb, 1871.

Milroy 1999
    Milroy J., Milroy L., *Authority in Language: Investigating Standard English*, London, New York, 1999.

Mladenova 2007
    Mladenova O. M., *Definiteness in Bulgarian: Modelling the Process of Language Change* (= Trends in Linguistics: Studies and Monographs, 182), Berlin, New York, 2007.

Radeva 2003
    Radeva V., hg., *Bulgarische Grammatik: Morphologisch-syntaktische Grundzüge*, Hamburg, 2003.

Schleicher 1983
    Schleicher A., *Die Sprachen Europas in systematischer Übersicht: linguistische Untersuchungen* (= Amsterdam Classics in Linguistics, 4), Amsterdam, Philadelphia, 1983.

Skok 1971
    Skok P., *Etimologijski rječnik hrvatskoga ili srpskoga jezika*, Zagreb, 1971.

Sobolev 1996
    Sobolev A. N., Das Sprachgrenzenproblem im Balkanslavischen, H. Schaller., K. Haralampieff, W. Gesemann, *Sprache und Politik: Die Balkansprachen in Vergangenheit und Gegenwart* (= Südosteuropa-Jahrbuch, 27), München, 1996, 59–74.

Sonnenhauser 2015
    Sonnenhauser B., Functionalising syntactic variance declarative complementation with *kako* and *če* in 17th to 19th century Balkan Slavic, *Wiener Slavistisches Jahrbuch*, 3, 2015, 41–72.

Stefanov 2008
    Stefanov P., The Beginning of Bulgarian Printing (On the Occasion of its 500th Anniversary), *Occasional Papers on Religion in Eastern Europe*, 28/2, 2008, 59–67.

Tomić 2006
    Tomić O. M., *Balkan Sprachbund Morpho-syntactic Features* (= Studies in Natural Language and Linguistic Theory, 67), Dordrecht, 2006.

Trunte 2018
    Trunte N., *Kirchenslavisch in 14 Lektionen* (= Studienhilfen, 19), Wien, 2018.

Tsibranska-Kostova 2016
    Tsibranska-Kostova M., *Абагаръ* на Филип Станиславов: от артефакта към езиковия ресурс, *Zeszyty Cyrylo-Metodiańskie*, 5, 2016, 6–23.

Vuković 2020
    Vuković T., Degrees of non-standardness. Feature-based analysis of variation in a Torlak dialect corpus, *Journal of Corpus Linguistics*, 2020 (forthcoming).

Weber 1976
    Weber E., *Peasants into Frenchmen: The Modernization of Rural France 1870–1914*, Stanford, 1976.

Андрейчин 1944
    Андрейчин Л., *Основна българска граматика*, София, 1944.

——— 1977
    Андрейчин Л., *Из историята на нашето езиково строителство*, София, 1977.

Берович 1824
    Берович П. Х., *Буквар с различни поучения*, 1824.

Богоров 1844
    Богоров И., *Първичка българска граматика*, Букурещ, 1844.

Бончев 1952
    Атанасий (Бончев), арх., *Църковнославянска граматика*, София, 1952.

Велчева 1964

Велчева Б., Показателни местоимения и наречия в новобългарските паметници от XVII и XVIII в., *Известия на института за български език*, 10, София, 1964, 159–233.

——— 1966

Велчева Б., Норма и традиция в българския книжовен език от XVI-XVIII в., *Български език*, 2, 1966, 110–121.

Венелин 1838

Венелин Ю., *О зародыще новой болгарской литературы*, Москва, 1838.

Вълчев 2008

Вълчев Б., *Възрожденските граматики на българския език* (= Университетска библиотека, 477), София, 2008.

Вътов 2001

Вътов В., *Езикът на Софроний Врачански*, В. Търново, 2001.

Гълъбов 1968

Гълъбов И., Ранни школи за стария български книжовен език, *Български език*, 2–3, 1968, 141–148.

Демина 1968

Демина Е. И., *Тихонравовский дамаскин: болгарский памятник XVII в.: исследования и текст*, 1, София, 1968.

——— 1985

Демина Е. И., *Тихонравовский дамаскин: болгарский памятник XVII в.: исследования и текст*, 3, София, 1985.

Дринов 1911

Златарски В. Н., изд., *Трудове на М. С. Дринова*, 2, София, 1911.

Кайперт 2017

Кайперт Г., Церковнославянский язык: круг понятий, *Slověne*, 1, 2017, 8–75.

Керемедчиев 1943

Керемедчиев Г., *Борба за книжовен език и правопис*, София, 1943.

Конески 1967

Конески Б., *Граматика на македонскиот литературен јазик*, Скопје, 1967.

Маслов 1981

Маслов Ю. С., *Грамматика болгарского языка*, Москва, 1981.

Миронова 2010

Миронова Т. Л., *Церковнославянский язык*, Москва, 2010.

Мирчев 1978

Мирчев К., *Историческа граматика на българския език*, София, 1978.

Мисирков 1903

Мисирков К. П., *За македонцките работи*, София, 1903.

Младенов 1963

Младенов М. Сл., Членувани прилагателни форми на -*ий* в североизточните български говори, *Български език*, 4–5, 1963, 404–410.

Момчилов 1868

Момчилов И., *Грамматика за новобългарския език*, Русчук, 1868.

Павлович 1836

Павлович Х., *Грамматика славено-болгарска*, Будим, 1836.

Петканова-Тотева 1965

Петканова-Тотева Д., *Дамаскините в българската литература*, София, 1965.

Пуљевски 1875

Пуљевски Ђ. М., *Речник од три језика - С. Македонски, Арбански, и Турски*, Београд, 1875.

Рилски 1989

Неофит Рилски, *Болгарска Грамматика*, Крагуевац, 1835. [repr.: R. von Olesch, hrsg., nachdr., Neofit Rilski, *Bolgarska grammatika* (= Slavistische Forschungen, 41), Köln, Wien, 1989].

Смотрицкий 1648

Смотрицкий М. Г., *Грамматіка*, Москва, 1648.

Стойков 1993

Стойков Ст., *Българска диалектология*, София, 1993.

Христова 1991

Христова Б., За ресавските български ръкописи от XV–XVIII век, *Старобългаристика*, 4, 1991, 50–56.

Хрулев 1859

Хрулев Т. Т., *Българска грамматика: За ръководство на българските юноши*, Букурещ, 1859.

Шаур 1970

Шаур В., *Поп-Пунчов сборник как источник историко-диалектологических исследований*, Praha, 1970.

Ягич 1895

Ягич И. В., Книга Константина философа и грамматика о писменех, Idem, *Разсуждения южнославянской и русской старины о церковно-славянском языке* (= Исследования по русскому языку, 1), С.-Петербург, 1895, 366–581.

References

Anderson B., *Imagined Communities: Reflections on the Origin and Spread of Nationalism*, London, New York, 2006.

Andreychin L., *Iz istoriiata na nasheto ezikovo stroitelstvo*, Sofia, 1977.

Andreychin L., *Osnovna bulgarska gramatika*, Sofia, 1944.

Athanasius (Bonchev), arch., *Tsurkovnoslavianska gramatika*, Sofia, 1952.

Barth F., *Ethnic Groups and Boundaries: The Social Organization of Culture Difference*, Boston, 1969.

Bourdieu P., *Language and Symbolic Power*, tr. by G. Raymond and M. Adamson, Oxford, 1991.

Demina E. I., *Tikhonravovskii damaskin: bolgarskii pamiatnik XVII v.: issledovaniia i tekst*, 1, Sofia, 1968.

Demina E. I., *Tikhonravovskii damaskin: bolgarskii pamiatnik XVII v.: issledovaniia i tekst*, 3, Sofia, 1985.

Derksen R., *Etymological Dictionary of the Slavic Inherited Lexicon* (= Leiden Indo-European Etymological Dictionary Series, 4), Leiden, Boston, 2008.

Fielder G. E., The Semiotics of Ideology: The Definite Article Rule in Bulgarian, *Balkanistica*, 32/2, 2019, 45–70.

Friedman V. A., Balkans as a Linguistic Area, K. Brown, ed., *Encyclopedia of Language and Linguistics*, 2nd ed., 1, Oxford, 2006, 657–672.

Friedman V. A., Language Ideology and Language Change in the Balkans: A View from the Early Twenty-First Century, *Die Welt der Slaven*, 62/1, 2017, 1–21.

Friedman V. A., Macedonian Language and Nationalism During the Nineteenth and Early Twentieth Centuries, *Macedonian Review*, 16/3, 1986, 280–292.

Gulubov I., Ranni shkoli za stariia bulgarski knizhoven ezik, *Balgarski ezik*, 2–3, 1968, 141–148.

Haarmann H., *Aspekte der Arealtypologie: die Problematik der europäischen Sprachbünde*, Tübingen, 1976.

Handelman D., The Organization of Ethnicity, *Ethnic Groups*, 1(3), 1977, 187–200.

Hastings A., *The Construction of Nationhood: Ethnicity, Religion and Nationalism*. Cambridge, 1997.

Hobsbawm E., *Nations and nationalism since 1780: programme, myth, reality*. Cambridge, 1990.

Hristova B., Za resavskite bulgarski rukopisi ot XV–XVIII vek, *Palaeobulgarica*, 4, 1991, 50–56.

Hroch M., *Social Preconditions of National Revival in Europe: A Comparative Analysis of the Social Composition of Patriotic Groups among the Smaller European Nations*, B. Fowkes, trans., Cambridge, 1985.

Irvine J. T., Gal S., Language Ideology and Linguistic Differentiation, P. V. Kroskrity, ed., *Regimes of language: Ideologies, polities, and identities*, Santa Fe, 2000, 35–84.

Karadžić V. S., *Kleine Serbische Grammatik* (= Sagners Slavistische Sammlung, 1), übers. J. Grimm, München, 1974 [repr. 1824].

Kaur D., A Comparative Study of Various Distance Measures for Software fault prediction, *International Journal of Computer Trends and Technology*, 17/3, 2014, 117–120.

Keipert H., Conceptions of Church Slavonic, *Slověne*, 1, 2017, 8–75.

Keremedchiev G., *Borba za knizhoven ezik i pravopis*, Sofia, 1943.

Koneski B., *Gramatika na makedonskiot literaturen jazik*, Skopje, 1967.

Lunt H. G., *Old Church Slavonic Grammar,* 7th ed., Berlin, 2001.

Lunt H. G., *A Grammar of the Macedonian Literary Language*, Skopje, 1952.

Maslov Iu. S., *Grammatika bolgarskogo iazyka*, Moscow, 1981.

Milroy J., Milroy L., *Authority in Language: Investigating Standard English*, London, New York, 1999.

Mirchev K., *Istoricheska gramatika na bulgarskiia ezik*, Sofia, 1978.

Mironova T. L., *Tserkovnoslavianskii iazyk*, Moscow, 2010.

Mladenov M. Sl., Chlenuvani prilagatelni formi na -ii v severoiztochnite bulgarski govori, *Balgarski ezik*, 4–5, 1963, 404–410.

Mladenova O. M., *Definiteness in Bulgarian: Modelling the Process of Language Change* (= Trends in Linguistics: Studies and Monographs, 182), Berlin, New York, 2007.

Olesch R., von, hrsg., Neofit Rilski, *Bolgarska grammatika* (= Slavistische Forschungen, 41), Köln, Wien, 1989.

Petkanova-Toteva D., *Damaskinite v bulgarskata literatura*, Sofia, 1965.

Radeva V., hg., *Bulgarische Grammatik: Morphologisch-syntaktische Grundzüge*, Hamburg, 2003.

Saur V., *Pop Puncov sbornik kak istocnik istoriko-dialektologiceskich issledovanij*, Praha, 1970.

Schleicher A., *Die Sprachen Europas in systematischer Übersicht: linguistische Untersuchungen* (= Amsterdam Classics in Linguistics, 4), Amsterdam, Philadelphia, 1983.

Skok P., *Etimologijski rječnik hrvatskoga ili srpskoga jezika*, Zagreb, 1971.

Sobolev A. N., Das Sprachgrenzenproblem im Balkanslavischen, H. Schaller., K. Haralampieff, W. Gesemann, *Sprache und Politik: Die Balkansprachen in Vergangenheit und Gegenwart* (= Südosteuropa-Jahrbuch, 27), München, 1996, 59–74.

Sonnenhauser B., Functionalising syntactic variance declarative complementation with *kako* and *če* in 17th to 19th century Balkan Slavic, *Wiener Slavistisches Jahrbuch*, 3, 2015, 41–72.

Stefanov P., The Beginning of Bulgarian Printing (On the Occasion of its 500th Anniversary), *Occasional Papers on Religion in Eastern Europe*, 28/2, 2008, 59–67.

Stoykov St., *Bulgarska dialektologiia*, Sofia, 1993.

Tomić O. M., *Balkan Sprachbund Morpho-syntactic Features* (= Studies in Natural Language and Linguistic Theory, 67), Dordrecht, 2006.

Trunte N., *Kirchenslavisch in 14 Lektionen* (= Studienhilfen, 19), Wien, 2018.

Tsibranska-Kostova M., Abagarut na Filip Stanislavov: ot artefakta kum ezikoviia resurs, *Zeszyty Cyrylo-Metodiańskie*, 5, 2016, 6–23.

Velcheva B., Norma i traditsiya v balgarskiya knizhoven ezik ot XVI-XVIII v., *Balgarski ezik*, 2, 1966, 110–121.

Velcheva B., Pokazatelni mestoimeniia i narechiia v novobulgarskite pametnitsi ot XVII i XVIII v., *Papers of the Institute for Bulgarian Language*, 10, Sofia, 1964, 159–233.

Vŭlchev B., *Bulgarian language grammar from the period of the national revival* (= Universitetska biblioteka, 477), Sofia, 2008.

Vŭtov V., *Ezikut na Sofronii Vrachanski*, V. Tarnovo, 2001.

Weber E., *Peasants into Frenchmen: The Modernization of Rural France 1870–1914*, Stanford, 1976.

Ivan Šimko, Ph.D.
Universität Zürich
Slavisches Seminar
Plattenstrasse 43,
CH-8032, Zürich
Die Schweiz / Switzerland