Daniel Bunčić

Selected Papers in Slavic Linguistics

Nine papers originally published in languages other than English, translated by the author

Daniel Bunčić
Selected papers in Slavic linguistics

Daniel Bunčić

Selected Papers in Slavic Linguistics

Nine papers originally published in languages other than English, translated by the author

USB MONOGRAPHS
Cologne 2026

The green book cover pays homage to the 1980 edition of the Latin-German school dictionary *Der kleine Stowasser*, whose nineteen-page "Sprachgeschichtliche Einführung", a compressed historical grammar of Latin—the only truly linguistic text I was ever given during thirteen years of school—got me fascinated with linguistics when I was thirteen. More recent editions of the dictionary have more beautiful covers (some of them designed by Friedensreich Hundertwasser), but unfortunately they do not contain this Indo-Europeanist introduction anymore.

Published in Cologne by USB Monographs, 2026

University of Cologne · University and City Library of Cologne
USB Monographs
Universitätsstraße 33, 50931 Köln, Germany

The electronic version of this publication is freely available online at USB Monographs: https://monographs.ub.uni-koeln.de/ (open access).
DOI: 10.18716/omp.64

ISBN (hardback): 978-3-565382-20-0
ISBN (e-book): 978-3-946275-26-8

Bibliographic information published by the *Deutsche Nationalbibliothek*.
The German National Library lists this publication in the *Deutsche Nationalbibliografie*; detailed bibliographic data are available on the internet at https://portal.dnb.de/.

Text, translation, layout, and book cover by Daniel Bunčić

For Barbara

Preface

The reason for this volume is that English has become the language of science, even within Slavic Studies, and I have noticed that especially my papers published in German are not read half as much as my papers published in English. Sometimes they are cited in ways that make it obvious that the colleague has not read more than the title because they do not understand the language of the paper.

However, I do think that some—certainly not all!—of the papers I have written in German or Russian during the last quarter of a century might still benefit readers and prevent them from doing the same work again. These are the papers I collected in this volume.

All these papers are witnesses of the time they were written in, and I did not try to update or modernize them, because that would basically have meant to rewrite them completely. Only in a few cases did I find some adjustment or comment unavoidable; these places are marked with "D. B. 2026". Misspellings and similar obvious mistakes were corrected tacitly, and the citation format was unified according to the Generic Style Rules for Linguistics. In general, however, you have to take these papers as texts of their time, not of 2026. I marked the original page breaks, so feel free to refer to the original publications (in addition to or instead of the translations).

The subjects of the papers are very diverse, reflecting my varying interests over seventeen years: false friends, Serbo-Croatian standard varieties, histories of concepts, biaspectual verbs, 'definiteness' in Russian, diglossia in 18th-century Russian, the sociolinguistics of writing. Since this makes a thematic grouping impossible, the papers are here presented in chronological order (with one exception).

The good thing about doing this now, rather than ten years ago, is that translating has become so much easier. For this volume, I heavily relied on DeepL (the company behind it, incidentally, being based just 3½ km from my university office in Cologne). Even though the machine provides excellent results, I checked every sentence thoroughly and changed about every other one, aligning the texts with my own voice and adjusting them to an English-speaking audience. Two of the papers collected here were actually originally written in English but then turned out to be published in Russian and German, respectively, so that I could use my original English wording wherever the published version did not deviate from it.

I am indebted to Katja Halassy and Raphael Thiele for accepting my book for USB Monographs, to Lilia Troanska and Bryan Ebel for diligently proofreading the whole book, to my academic teachers, Helmut Keipert and Tilman Berger, for teaching me the methods of our discipline, and to my wife, for everything.

Table of contents

Criteria for determining the degree of 'danger' associated with false friends

False friends (also known as 'deceptive cognates', 'treacherous twins', etc.) are a source of lexico-semantic interference. They can be defined as lexemes in different languages with formal congruency (i.e. similarity or regular correspondence) that do not provide the best semantic equivalent for each other, e.g. R. *stol* 'table' vs. Bg. *stol* 'chair'. While some false friends cause a lot of trouble for almost all bilinguals, other pairs are hardly ever confused. What determines how 'dangerous' a pair of false friends is? The degree of 'danger' can be analyzed on the basis of two factors: the probability of confusion and the seriousness of the potential consequences. This paper shows that the former factor is much more important than the latter. Among the various items discussed, the risk of confusion is increased especially by an overlap of the aggregate denotations of the words in question and by a semantic connection (proximity or antonymity) between their individual definitions. With criteria like these at hand, the 'danger' of any pair of false friends can be estimated, and certain particularly nasty pairs can be singled out in order to compose effective 'false friends' exercises for learners or to mark them in bilingual dictionaries.

Translated from "Kriterii dlja opredelenija stepeni 'opasnosti' psevdo-analogonimov ('ložnyx druzej')". Submitted in 2000 for publication in a Bulgarian volume that never saw the light of day and self-published online in 2001 (DOI: 10.18716/bun/krit).

DOI: 10.18716/omp.64.1

Hayward & Moulin (1984: 190) define the concept of "false friends of the translator" using Saussure's terminology: they are two words in different languages whose *signifiers* (forms) are "more or less similar" ("mehr oder weniger ähnlich", Gauger 1989: 581), but whose *signifieds* (meanings) differ. Examples of this phenomenon include Russian *stol* 'table' vs. Bulgarian *stol* 'chair', Russian *gora* 'mountain' vs. Bulgarian *gora* 'forest', and Russian *strana* 'country' vs. Bulgarian *strana* 'country; side; cheek'.[1] Unfortunately, the term *false friends* is used by different authors with different meanings, This article uses a clear synchronic definition (see Bunčič 2000: 17) that refers to lexemes of the same part of speech and avoids defining semantic differences.[2]

[1] More examples from Slavic languages can be found in the wikibook *False Friends of the Slavist* (https://en.wikibooks.org/wiki/False_Friends_of_the_Slavist). [The original article referred to my private website before it was turned into a wikibook.—D. B. 2026.]

[2] The original German article uses the term *pseudo-analogonyms* introduced by Hengst (1977: 252). Since this term never caught on, I here translate it as *false friends*. (D. B. 2026)

However, even with such a clear definition, which excludes much of what is often interpreted as “false friends”,[3] this turns out to be a very heterogeneous phenomenon. That applies, first of all, to the probability of errors associated with a given pair of words.

(1) Russian ***bulka*** = Bulgarian *xlebče, bjal xljab, simid*
Bulgarian ***bulka*** = Russian *nevesta; devuška*

(2) Russian ***čitat'*** = Serbian ***čitati*** ‘to read in a book, newspaper, etc.’
Russian ***čitat'*** = Serbian *recitovati* ‘to read aloud’

It is clear that among the examples given here, Russian *bulka* and Bulgarian *bulka* are fairly harmless false friends, while Russian *čitat'* and Serbian 2 *čitati* can easily be confused. Even a professional philologist made a mistake here when translating a conversation with Aleksandr Blok from Russian into Serbian (original: Aljanskij 1972: 109; translation: Aljanski 1971: 64). Although Marojević (1989: 23) sharply criticizes this mistake,[4] this type of false friends is incomparably more complex than the one in (1).

What is the difference? What factors influence the varying degrees of ‘danger’ of false friends? It seems that formal similarity only provides *an opportunity* for error, but a greater or lesser degree of similarity probably does not affect the likelihood of confusion between two similar forms (after all, even the forms of ‘true friends’ can be more or less different, cf. Russian *vysyxat'* vs. Bulgarian *izsăxvam*, Russian *ot* vs. Ukrainian *vid*). Therefore, this article is limited to the *semantic* side of this issue.

Most words (both false friends and others) have several meanings. This is because *polysemy* is a linguistic universal, necessary for expressing an infinite number of possible situations with a limited number of signs (lexemes) that human mem-

---

[3] The following phenomena are excluded, among others:

- formal interference, which Svobodová-Chemlová (1982) calls “presqu’amis” (“almost-friends”), e.g. Russian *svëkla* (feminine) vs. Bulgarian *cvekló* (neuter) ‘beetroot’, Russian *plóščad'* (feminine) vs. Bulgarian *ploštád* (masculine) ‘square, plaza’;
- “barbarisms”, also called “sans-amis” (“without-friends”, Svobodová-Chemlová 1982) or “mots-fantômes” (“phantom words”, Murav'ev 1985: 5f.), i.e. words that do not exist in a given language, e.g. Russian *teatr* vs. Serbian **teatar* ‘theater’ (which is *pozorište* in Serbian);
- word forms in different “grammatical positions” (Karpov 1998: 57), e.g. Russian *greška* (gen.sg. of *grešok* ‘trivial offense’) vs. Bulgarian *greška* (nom.sg.) ‘mistake’;
- words of different parts of speech, e.g. Russian *xotel* (verb, past tense) ‘wanted to’ vs. Bulgarian *xotel* (noun, nominative case) ‘hotel’.

[4] “Није јасно како се преводилац није досетио да се не може читати «опирући се обема рукама о наслон столице» него само рецитовати, говорити.” (“It is unclear why it did not occur to the translator that one cannot read while ‘leaning both hands on the back of a chair’, but only recite from memory, speak.”)

ory can retain. Consequently, a language without polysemy would be a "linguistic hell" ("лингвистический ад", Budagov 1963: 242).[5]

Therefore, when it comes to the meaning of a word, a distinction should be made between *individual meanings* (at the level of *parole*), of which a single lexeme may have several depending on the context in which it is used, and *the aggregate meaning* (at the level of *langue*) as the sum of the individual meanings. For example, the Russian word *volja* has the aggregate meaning 'desire; power; freedom', but in a specific utterance, *volja* can only appear in one of the three separate meanings, viz. 'desire', 'power', or 'freedom'.[6]

Against the backdrop of the **aggregate meaning of a word-form** as the sum of all its individual meanings, homonymy turns out to be only a special case of polysemy, so that within this article it is quite possible to consider the relationship between the meanings of two false friends without focusing on the difference between homonymy and polysemy.

D3 At the level of the aggregate meaning, we can speak of *complete false friends* when the meanings of two lexemes are in *a relationship of exclusion*:

(3) Russian ***život*** = Bulgarian *korem, tărbux* 'belly'
Bulgarian ***život*** = Russian *žizn'* 'life'

(4) Russian ***majka*** = Bulgarian *potnik, riza, košulja* 'undershirt, T-shirt, jersey'
Bulgarian ***majka*** = Russian *mat'* 'mother'; *matka* 'womb', *gajka* 'nut (for a bolt)'

Such false friends are easiest for language learners to remember, since the logical proposition that learners need to master is relatively simple: A ≠ B, e.g. Bulgarian *život* ≠ Russian *život*.

In the case of *partial false friends*, the language learner must additionally remember the conditions under which A = B and the conditions under which A ≠ B. The group of partial false friends can be further divided into subgroups of *inclusion* and *intersection*. In the case of *inclusion,* one of the lexemes covers all the meanings of the other, but the latter does not have all the meanings of the former:

(5) Bulgarian ***kraj*** = Russian ***kraj*** 'boundary line; country, region'
Bulgarian ***kraj*** = Russian *konec* 'end'

(6) Russian ***brak*** = Bulgarian ***brak*** 'marriage; scrap, waste'
Russian ***brak*** = Bulgarian *nedostatăk, greška* 'flaw, defect'

[5] Budagov attributes this expression to Otto Jespersen, but unfortunately without specific bibliographical data.

[6] Of course, it is possible to distinguish even more individual meanings; for example, Ožegov & Švedova (1999: 96) give two lexemes, *volja*[1] and *volja*[2], with five and two meanings, respectively.

Here, difficulties arise only in one direction, namely when translating from a language in which the word has more meanings. For example, when translating the Bulgarian word *kraj* with the meaning 'end', there is a danger of mistranslating it with the Russian word *kraj*. In the opposite case, the Russian word *kraj* is always correctly translated with the Bulgarian word *kraj*, since it covers all the meanings of the Russian lexeme.

Even more complex is the *intersection* of meanings. In this group of false friends, *both* lexemes have meanings that are not found in the false equivalent:

(7) Russian ***isxod*** = Bulgarian ***izxod*** 'result; way out of a bad situation'
Russian ***isxod*** = Bulgarian *kraj* 'end'
Bulgarian ***izxod*** = Russian *vyxod* 'exit'

Thus, this type of false friends has disadvantages in terms of both exclusion and inclusion: because of the meanings common to both lexemes, they are difficult to distinguish, and mistakes can be made in translation in both directions.

Milan (1989: 399), who conducts a similar semantic analysis of false friends, points out that Ivir (1968: 152–154) also mentions the relationship of equivalence, in which the semantics of the lexemes are the same. Such "false friends" differ, for example, in frequency of use. However, different frequencies of lexemes are always associated with stylistic, pragmatic, or minor semantic differences (unless the words refer to realities that occur with different frequencies in the compared language communities), so 🗋4 it seems insufficient to call this phenomenon (semantic) equivalence. Instead, the existing differences should be explored in more detail.

The specific meaning that a word has in *parole* is closely related to the situation in which the word is used. This means that in a specific situation, the danger of confusing the meaning of one word with the meaning of another exists only if both lexemes can semantically be used in the same situation. Therefore, many pairs of words that formally are false friends are unlikely to ever lead to errors due to inappropriate context. This applies primarily to those false friends whose members have accidentally converged (or become almost identical) in the course of the history of the language:

(8) Bulgarian ***krik*** = Russian *domkrat* 'car jack'
Russian ***krik*** = Bulgarian *vik* 'shout, cry'

(9) Bulgarian ***razboj*** = Russian *tkackij stanok* 'loom'
Russian ***razboj*** = Bulgarian *obir* 'robbery'

However, there are also randomly merged false friends with a high probability of error, and etymologically related words can diverge semantically so far that misunderstanding is almost impossible. The latter occurred with the following lexemes, which go back to the verbal root **dum-* 'to think':

(10) Bulgarian ***duma*** = Russian *slovo* 'word'
Russian ***duma*** = Bulgarian *(Narodno) sǎbranie* '(national) parliament'; *misǎl* 'thought'; *iztočnoslavjanska narodna pesen* 'East Slavic folk song'

It should be noted that when analyzing the risk of errors, we can only talk about probabilities. Since the number of possible situations is infinite, we can never rule out a situation in which two words are in a paradigmatic relationship that could cause confusion. Thus, "a 'false friend' does not exist objectively between two languages but always affects an individual speaker" ("besteht ein 'faux ami' nicht objektiv zwischen zwei Sprachen, sondern es ist immer ein individueller Sprecher davon betroffen", Haschka 1989: 149), for whom some false friends pose no problem, while others that seem obvious lead them astray. Since false friends are a concept of comparative lexicology, not only didactics, it covers all pairs of words that formally meet its criteria. However, some particularly difficult cases can be identified among them.

D5 The risk of confusion exists, first of all, when false friends belong to the same lexicosemantic field, e.g. the field of entertainment activities:

(11) Russian ***igrat'*** = Bulgarian ***igraja*** 'to play, have fun; to play a game; to play a role'
Russian ***igrat'*** = Bulgarian *svirja* 'to play music'
Bulgarian ***igraja*** = Russian *tancevat', pljasat'* 'to dance'

Often, though not always, it is possible to find a specific generic concept for false friends, so that the words being compared are cohyponyms:

(12) Russian ***dynja*** = Bulgarian *pǎpeš* 'cantaloupe, muskmelon, *Cucumis melo*'
Bulgarian ***dinja*** = Russian *arbuz* 'watermelon, *Citrullus lanatus*'

(13) Russian ***godina*** (poetic) = Bulgarian *vreme, moment* 'time, period, moment'
Bulgarian ***godina*** = Russian *god* 'year'

The generic concepts in (12) and (13) are, respectively, 'soft-fleshed sweet-flavored fruit of the gourd family' and 'time'.

It is worth noting that words with opposite meanings also belong to the same lexicosemantic field; for example, both of the following false friends can denote human qualities:

(14) Russian ***vrednyj*** = Bulgarian ***vreden*** 'harmful, dangerous'
Russian ***vrednyj*** = Bulgarian *neprijaznen* 'hostile'
Bulgarian ***vreden*** = Russian *lovkij, umelyj* 'neat, skillful, capable'

Furthermore, one of the false friends can be a hypernym of the other.

(15) Russian ***jagoda*** = Bulgarian *sočen zǎrnest plod ot xrastovi rastenija* 'berry'
Bulgarian ***jagoda*** = Russian *klubnika* 'strawberry'

(16) Russian ***palec*** = Bulgarian *prǎst* 'finger; toe'
Bulgarian ***palec*** = Russian *bol'šoj palec* 'thumb; big toe'

Čongarova (1992: 77) refers to this as *inclusion* of one lexeme by another. However, this term should not be confused with the aforementioned *relationship of inclusion*, which refers to the range of meanings in the sense of polysemy, whereas this *inclusion* is related to the "breadth of meaning" ("Bedeutungsbreite", Hengst 1977: 257; "širokoznačnost'", Čongarova 1992: 78). Čongarova also cites the following false friends as examples of *inclusion*:

(17) Russian ***nedelja*** = Bulgarian *sedmica* 'week'
Bulgarian ***nedelja*** = Russian *voskresen'e* 'Sunday'

(18) Russian ***bor*** = Bulgarian *borova gora* 'pine grove, pine forest'
Bulgarian ***bor*** = Russian *sosna* 'pine tree'

Indeed, the phenomenon here is similar, but not the same. After all, in (15)–(16), we are dealing with generic and specific concepts (i.e. *hyponymy*), so it is always logically true that 'a strawberry is a berry', 'a thumb is a finger', etc. In (17)–(18), however, such identification (*'Sunday is a week', *'a pinetree is a pine grove') is impossible because they are in a part-whole relationship (i.e. 🗋6 *paronymy*). Therefore, it would be more accurate to distinguish between *hyponymic inclusion* and *paronymic inclusion*.

Sometimes the semantic difference between false friends consists only in a small nuance, for example, in the degree of coldness in (19).

(19) Russian ***xolod*** = Serbian *hladnoća*, *studen*, *zima* 'cold'
Serbian ***hlad*** = Russian *proxlada* 'coolness'; *ten'* 'shade'

Here, it can be clearly determined that these words, which are similar in meaning, are nevertheless not equivalent.

So far, we have only discussed *semantic* differences. However, words that are completely identical in meaning can also be false friends, differing from each other, for example, stylistically. Thus, the Russian words *dlan'* 'hand, palm' and *usta* 'mouth' are considered obsolete in Russian and therefore have an archaic or poetic connotation when they appear in modern texts. By contrast, in Bulgarian, *dlan* 'hand, palm' and *usta* 'mouth' are neutral words corresponding to the Russian *ladon'* and *rot*, respectively.

While the *probability* of error is influenced by the relationship between the meanings of two words, it is largely **the meaning itself** that determines the consequences of the error. For example, false friends containing a precise definition of time or place can result in people not arriving on time for a meeting or not finding the right road.

(20) Russian: ***napravo*** = Bulgarian: *na djasno* 'to the right'
Bulgarian ***na pravo*** = Russian *prjamo* 'straight on'

(21) Russian: ***s utra*** = Serbian: *ujutro* 'in the morning'
Serbian: ***sutra*** = Russian: *zavtra* 'tomorrow'

The danger of awkward situations arises when the word used turns out to denote a concept from a specific sphere of usage:

(22) Bulgarian ***kal*** = Russian *grjaz'* 'dirt'
Russian ***kal*** = Bulgarian *ekskrementi* 'excrements'

The Russian verb *pisat'* is a special case: with the stress on the second syllable, it has the same harmless meaning as its Bulgarian counterpart *piša*. However, in the latter, the stress is placed on the first syllable, and with this accentuation in Russian, the word takes on a different meaning and a colloquial tone:

(23) Rus. ***pisát'*** = Bulg. ***píša*** 'to write'
Russian ***písat'*** (colloquial) = Bulgarian *pikája* 'to pee'

However, the factor of the meaning of the two words is modified by the relationship between their meanings, as seen in (24).

(24) Russian ***grudastyj*** = Serbian ***grudast*** 'grainy, granular'
Russian ***grudastyj*** = Serbian *sa velikim grudima* 'big-breasted'

Here, due to the relationship of inclusion between these words, a mistake can only be made when translating from Russian into Serbian. This means that the only 7 possible mistake is that a Russian speaker who wants to describe a woman with large breasts in Serbian uses a technical term with the meaning 'grainy'. False friends that are swearwords in both languages are also relatively safe.

(25) Russian ***prostak*** = Serbian *glupak* 'simpleton'
Serbian ***prostač*** = Russian *grubijan, nevež̌a* 'rude person'

Thus, to determine the 'danger' of false friends, we can say that the meanings of the two words being compared in themselves, which may influence the consequences of a potential error, are of secondary importance. More important is the type of relationship between the meanings, which affects the probability of error. Pairs of false friends pose a particular danger when the meanings of polysemous words overlap or when there is a semantic connection (proximity or antonymy) between their individual meanings.

References

Aljanski, Samuil M. 1971. *Kako je umirao Blok.* (Transl. Stanojlo Bogdanović.) Beograd: Slovo Ljubve.

Aljanskij, Samuil M. 1972. *Vstreči s Aleksandrom Blokom.* Moskva: Detskaja literatura.

Budagov, Ruben A. 1963. Sravnitel'no-semasiologičeskie issledovanija (romanskie jazyki). Moskva.

Bunčič, Daniel′. 2000. Psevdo-analogonimija: "Ložnye druz′ja perevodčika" kak edinica sopostavitel′noj leksikologii. In Ivaško, Ljudmila I. & Lutovinova, Irina S. (eds.), *Materialy XXIX mežvuzovskoj naučno-metodičeskoj konferencii prepodavatelej i aspirantov*, vyp. 11: Sekcija leksikologii i leksikografii (russko-slavjanskij cikl), 14–18. Sankt-Peterburg: Filologičeskij fakul′tet Sankt-Peterburgskogo gosudarstvennogo universiteta.

Čongarova, Irina. 1992. Mež″jazykovye leksičeskie sootvetstvija omonimičeskogo xaraktera (na materiale russkogo i bolgarskogo jazykov). Diss. kand. filol. nauk. Moskva: Institut Russkogo jazyka RAN.

Gauger, Hans-Martin. 1989. "Falsche Freunde" zwischen dem Deutschen und dem Spanischen: Ein lexikalisches Problem für Lernende (und Lehrende). In Cartagena, Nelson & Gauger, Hans-Martin, *Vergleichende Grammatik Spanisch-Deutsch*, vol. 2, 581–615. Mannheim, Wien, Zürich: Dudenverlag.

Haschka, Christine. 1989. Zur Entwicklungsgeschichte der "faux amis"-Forschung. *Lebende Sprachen* 1989(4). 148–152.

Hayward, Timothy & Moulin, André. 1984. False friends invigorated. In Hartmann, Reinhard R. K. (ed.), *Lexeter '83 proceedings: Papers from the International Conference on Lexicography at Exeter, 9–12 September 1983*, 190–198. Tübingen: Niemeyer.

Hengst, Karlheinz. 1977. Lehnwort – Internationalismus – Analogonym: Zur Semantik fachsprachlicher interlingualer Analogonyme im Russischen und Deutschen: H. H. Bielfeldt zum 70. Geburtstag. *Zeitschrift für Slawistik* 1977. 250–259.

Ivir, Vladimir. 1968. Serbo-Croat–English false pair types. *Studia Romanica et Anglica Zagrebiensia* 25/26. 149–159.

Karpov, Vladimir A. 1998. Vnutrijazykovaja i mež″jazykovaja omonimija bolgarskogo jazyka. In Safronov, German I. & Lilič, Galina A. (eds.), *Slavističeskij sbornik: V čest′ 70-letija professora P. A. Dmitrieva*, 56–63. Sankt-Peterburg: Izdatel′stvo Sankt-Peterburgskogo universiteta.

Marojević, Radmilo. 1989. *Lingvistika i poetika prevođenja (međuslovenski prevod).* Beograd: Naučna knjiga.

Milan, Carlo. 1989. Falsche Freunde: Ein besonderes Problem der kontrastiven Lexikologie (deutsch-italienisch). *Sprachwissenschaft* 1989. 384–404.

Murav′ev, Vjačeslav L. 1985. *"Ložnye druz′ja perevodčika": Posobie dlja učitelja francuzskogo jazyka.* 2nd edn. Moskva: Prosveščenie. [1st edn. Moskva 1969.]

Ožegov, Sergej I. & Švedova Natalija Ju. 1999. *Tolkovyj slovar′ russkogo jazyka.* Moskva: Az.

Svobodová-Chemlová, Jitka. 1982. *Problèmes de la traduction.* Vol. 2 (Interférences lexicales: Internationalismes, faux internationalismes et les problèmes liés à leur traduction). Praha: Státní pedagogické nakladatelství.

Intralingual and crosslinguistic enantiosemy as a communicative problem

There is an amazing phenomenon in languages: a combination of antonymy and homonymy, i.e. the expression of opposite meanings by one word. One of the best-known examples is Latin *altus*, meaning both 'high' (e.g. *altus mons* 'high mountain') and 'deep' (e.g. *altus lacus* 'deep lake'). However, remarkable as this fact may seem, it does not impede communication.

A phenomenon similar to this can also be observed between different languages, e.g. Polish *czerstwy chleb* 'stale bread' vs. Czech *čerstvý chléb* 'fresh bread', or English *absolutely* (as a reply) 'yes, absolutely' vs. Polish *absolutnie* 'no, absolutely not'. Cases like these, however, can indeed produce severe misunderstandings.

The difference between these two phenomena clearly lies in the different communicative situations. So how does "normal" communication among native speakers work, and what is communication in a second language lacking? An important role seems to be played by context and cotext, cultural background knowledge, redundancy, and strategies employed by the sender to prevent the receiver from misunderstanding the message.

Translated from "Ėnantiosemija vnutrijazykovaja i mež"jazykovaja kak problema kommunikacii". In Grabska, Marcelina (ed.), *"Słowa, słowa, słowa"… w komunikacji językowej II*, 207–212. Gdańsk: Fundacja Rozwoju Uniwersytetu Gdańskiego 2004. DOI: 10.18716/bun/enan.

DOI: 10.18716/omp.64.2

There is a remarkable phenomenon in languages—a combination of homonymy[1] and antonymy, i.e. the expression of opposite meanings by the same word. Some of the most famous examples include Latin *altus* 'high; deep', Russian *ob"exat'* 'to visit everywhere; to pass by', German *abdecken* 'to cover; to uncover',[2] and Polish *pożyczać* 'to lend; to borrow'.[3]

This phenomenon has been known for a long time. Cicero declared *tollendum esse Octavium*, meaning either 'Octavius should be elevated' or 'Octavius should be

[1] I would now argue that enantiosemy is more commonly a case of polysemy than of homonymy (D. B. 2026).

[2] The first meaning refers to translations such as 'remove (a roof), clear (the table), skin (an animal)', while the second refers to 'block (a light source), cover (an opposing team player)'.

[3] Unless otherwise indicated, the examples of enantiosemes in this article are taken either from my own collection or from Markovski (1986: 63ff.), Müller (1963: 51), Novikov (1973: 181–192), and Šercl' (1973 [1883]).

removed' (cf. Novikov 1973: 182; Hegel [1831] 1964: 120ff.). Georg Wilhelm Friedrich Hegel used the double meaning of the word *aufheben* in his philosophy:

> "Aufheben hat in der Sprache den gedoppelten Sinn, daß es so viel als aufbewahren, erhalten bedeutet, und zugleich so viel als aufhören lassen, ein Ende machen. Das Aufbewahren selbst schließt schon das Negative in sich, daß etwas seiner Unmittelbarkeit und damit einem den äußerlichen Einwirkungen offenen Daseyn entnommen wird, um es zu erhalten. – So ist das Aufgehobene ein zugleich Aufbewahrtes, das nur seine Unmittelbarkeit verloren hat, aber darum nicht vernichtet ist." (Hegel [1831] 1964: 120)

> "*Aufheben* has a double meaning in language, signifying as much as 'to preserve, to maintain' and at the same time as much as 'to cause to cease, to make an end'. The act of preserving itself already includes the negative aspect that something, in order to be preserved, is removed from its immediacy and thus from an existence open to external influences.—Thus, what is removed *(aufgehoben)* is at the same time preserved, having only lost its immediacy, but not destroyed."

208 This "intra-word antonymy" ("vnutrislovnaja antonimija", Novikov 1973: 182) has received the name *enantiosemy* [and also been called *contronymy*, *autoantonymy*, *antilogy*, *antagonymy* as well as *Janus words*—D. B. 2026], but it is still considered to be some kind of surprising flaw in language that should die out in modern literary languages, or, as Šercl′ (1973 [1884]: 259) wrote: "the more ancient the language and the more primitive the people, the more often this phenomenon occurs" ("чем язык древнее и чем народ примитивнее, тем чаще встречается это явление"). Against the backdrop of a more modern worldview, this idea sounds like this: "This phenomenon is unproductive in modern language and serves mainly as a kind of relic of the semantics of ancient roots" ("Это явление непродуктивно в современном языке и служит в основном своеобразным реликтом семантики древних корней", Novikov 1973: 192). This opinion is based, on the one hand, on the fact that there are very few examples of enantiosemy. For example, Andrzej Markowski (1986: 63) did not find a single example of enantiosemy among Polish adjectives.[4] On the other hand, it is usually assumed that enantiosemy hinders communication even more than ordinary homonymy, so that the language organism tries to suppress these phenomena.

Although there is indeed a tendency to eliminate ambiguity in language, the elimination of homonymy is much less common than is usually suggested (cf.

[4] Some works on more "exotic" (or "primitive") languages provide numerous examples of enantiosemy, many of which, however, are highly controversial: For example, Beznosikova (1984: 9) gives an example from the Komi language: *mešökys′ kis′tny* 'to pour out of a bag' and *mešökö kis′tny* 'to pour into a bag', where the verb *kis′tny* obviously means simply 'to pour', and the case endings of the noun 'bag' in this agglutinative language perform the functions of prepositions in inflected languages.

Budagov 1963: 243). In fact, enantiosemy in European standard languages is not dying out at all; on the contrary, new examples of this phenomenon are constantly being formed, cf. *abzac* in Russian youth slang 'something that causes sharp disapproval; something very good that causes approval' (Nikitina 1998: 9); from German youth slang, we can cite a whole series of adjectives with the same meanings: *ätzend, fett, fetzig, geil, höllisch, gut, irre, krass*. Evaluative enantiosemy is also present in the Russian slang verb *prikolot'sja*, cf. the following examples (taken from Nikitina 1998: 353ff.):

(1) *Vot svoloč', malo togo, čto dva stavit, ešče i* ***prikoletsja*** *nad toboj.*
'Here's a bastard, as if giving you an F weren't enough, he also **pokes fun** at you.'

(2) *Ja k ėtoj gerlé* ***prikololsja.***
'I have **fallen in love** with this girl.'

Of course, one could argue against such examples from slang, saying that they actually confirm the theory of the "primitiveness" of enantiosemy. 209 However, similar cases can be found in more scientific areas of vocabulary: for example, an important sphere of modern society, the field of economy, includes the Polish word *kupiec* 'trader, merchant, vendor; buyer, client' and the Polish example *pożyczać* 'to lend; to borrow' given above, as well as its German and Russian equivalents *leihen* and *odolžit'* (although the use of the latter in the sense of 'to borrow' is considered incorrect; cf. Novikov 1973: 182).[5]

All this indicates that enantiosemes (like other homonyms) do not pose a serious problem for communication, "since they are almost always neutralized either by the speech situation or by the context" ("так как почти всегда нейтрализуются или речевой ситуацией, или контекстом", Šanskij 1964: 43). One might even suggest that enantiosemy, like homonymy in general, is "a useful phenomenon" ("явление полезное", Malaxovskij 1990: 24). Hegel (1964 [1831]: 21ff.), discussing the ideal properties of language for science, even noted:

> "[D]ie deutsche Sprache hat […] viele Vorzüge vor den anderen modernen Sprachen; sogar sind manche ihrer Wörter von der weiteren Eigenheit, verschiedene Bedeutungen nicht nur, sondern entgegengesetzte zu haben […]; es kann dem Denken eine Freude gewähren, auf solche Wörter zu stoßen […]."
>
> "The German language has […] many advantages over the other modern languages; some of its words even have the further peculiarity of having not only different meanings, but opposite ones […]; it can be a joy for the mind to encounter such words […]."

[5] In non-rhotic varieties of English (i.e., variants with vocalized /r/), the prefixes *hyper-* 'above' and *hypo-* 'below', which are often used in scientific terminology (including for the formation of new terms), are pronounced identically. Thus, there are potentially many such homophones as *hypertension* 'high blood pressure' and *hypotension* 'low blood pressure'.

If we accept the definition of so-called "false friends" (or "pseudo-analogonyms", cf. Hengst 1977: 252 and Bunčič 2000: 15) as "crosslinguistic homonyms", then we can also speak of crosslinguistic enantiosemy. This phenomenon can be observed in many "false friends", cf. Russian *zapomnit'* 'to memorize' and Polish *zapomnieć* 'to forget', Russian *von'* 'stench' and Polish *woń* 'smell, aroma', Russian *kriminalist* 'criminalist' and Polish *kryminalista* 'criminal', Polish *no* 'well, yes' and Italian *no* 'no' or Polish *absolutnie* 'no, under no circumstances' and English *absolutely* 'yes, in any case'.[6] There are many more enantiosemes at this crosslinguistic level than within a single language. Firstly, this is due to the fact that there are no mechanisms that eliminate homonyms as there are within a language. 🗋210 Secondly, the reasons for the emergence of crosslinguistic enantiosemy are more diverse. For example, one of the most famous pairs of enantiosemes, Russian *urodlivyj* 'very ugly' and Polish *urodziwy* 'beautiful', arose purely by chance, namely, by the merger of two prefixes with the meaning 'not-' and 'up, rise-' in **u-*.

It is obvious that enantiosemy between two languages, compared to that within one language, can interfere with mutual understanding much more. To establish the reasons for the different status of this phenomenon in the two communicative situations, let us try to identify the differences between intralingual and interlingual communication.

As noted above, the meaning of intralingual enantiosemes is determined by context. This is done in two ways: First, different meanings of a word are used only in certain situations, so that the listener (or reader) can exclude those meanings that are not applicable in a given situation. Second, due to redundancy in language, almost everything is expressed several times, so that the addressee can usually reconstruct even completely indistinguishable words.

In contrast to this, a person speaking a foreign language cannot (or does not have time to) perceive all the details of the text, so redundancy for them is reduced to zero.[7] Of course, they do not know in which situations which meanings of a word are used. Often they do not know the word at all and are therefore forced to guess its meaning. The context does not help, because due to a common element in their meanings, antonyms can be used in the same contexts, e.g. Polish *czerstwy*

[6] Additional examples are given by Plotnikov (1979: 33ff.). Among gestures, we can cite nodding the head as a sign of agreement in most European countries, and in Bulgaria as a sign of negation.

[7] Therefore (as probably everyone who has ever studied a foreign language has experienced), a foreign language needs to be heard louder and under better circumstances than one's native language. This observation fits very well with Lewis's (2000: 21–24) approach to the problem of "false friends": He describes them as a special kind of noise in the interlingual channel of information transmission. It can be added that this type of noise is special in that it does not interfere with the message but, on the contrary, removes part of the information.

'stale' and Czech *čerstvý* 'fresh' in relation to bread: *czerstwy chleb* vs. *čerstvý chléb*. In addition to this, foreigners perceive the whole situation based on their native cultural and linguistic worldview. For all these reasons, feeling uncertain, they base their guesses on formal similarities with words in their native language (especially when the languages are closely related). This is the trickery of "false friends".

211 Moreover, when the sender and receiver both know the code they are using, then the speaker (and even more so the writer) almost instinctively avoids utterances that might be misunderstood. In interlingual communication, of course, the native speaker does not know which words might cause a misunderstanding by the non-native interlocutor.

Enantiosemy, this extreme case of semantics, has merely prompted these thoughts about the differences between communication in a non-native language and "normal" communication. The processes that ensure mutual understanding in communication in a language common to both interlocutors, but inaccessible to a non-native speaker, are always present, and the problems associated with this should be kept in mind when speaking to a non-native speaker.

References

Beznosikova, Ljucija M. 1984. *Ėnantiosemija v komi jazyke: Doklad na zasedanii Prezidiuma Komi filiala AN SSSR, 29 nojabrja 1984 g. (k VI Meždunarodnomu kongressu finno-ugrovedov, № 6).* (Serija preprintov "Naučnye doklady", 110.) Syktyvkar: AN SSSR.

Budagov, Ruben A. 1963. *Sravnitel'no-semasiologičeskie issledovanija (romanskie jazyki).* Moskva: Izdatel'stvo Moskovskogo gosudarstvennogo universiteta.

Bunčič, Daniel'. 2000. Psevdo-analogonimija: "Ložnye druz'ja perevodčika" kak edinica sopostavitel'noj leksikologii. In Ivaško, Ljudmila I. & Lutovinova, Irina S. (eds.), *Materialy XXIX mežvuzovskoj naučno-metodičeskoj konferencii prepodavatelej i aspirantov*, vyp. 11: Sekcija leksikologii i leksikografii (russko-slavjanskij cikl), 14–18. Sankt-Peterburg: Filologičeskij fakul'tet Sankt-Peterburgskogo gosudarstvennogo universiteta.

Hegel, Georg Wilhelm Friedrich. 1964 [1831]. *Wissenschaft der Logik: 1. Teil: Die objektive Logik.* (Sämtliche Werke: Jubiläumsausgabe in zwanzig Bänden, Hg. Hermann Glockner, Bd. 4.) 4th edn. Stuttgart: Frommann.

Hengst, Karlheinz. 1977. Lehnwort – Internationalismus – Analogonym: Zur Semantik fachsprachlicher interlingualer Analogonyme im Russischen und Deutschen: H. H. Bielfeldt zum 70. Geburtstag. *Zeitschrift für Slawistik* 1977. 250–259.

Lewis, Kristian. 2000. *Problem buke u priopćajnom kanalu te njezino oprimjerenje u slavenskoj homonimiji.* (Unpublished diploma thesis.) Zagreb.

Malaxovskij, Lev Vladimirovič. 1990. *Teorija leksičeskoj i grammatičeskoj omonimii.* Leningrad: Nauka.

Markowski, Andrzej. 1986. *Antonimy przymiotnikowe we współczesnej polszczyźnie na tle innych typów przeciwstawień leksykalnych.* Wrocław: Ossolineum.

Müller, Wolfgang. 1963. Über den Gegensatz in der deutschen Sprache. *Zeitschrift für deutsche Wortforschung* 19. 39–53.

Nikitina, Tat'jana G. 1998. *Tak govorit molodež': Slovar' slenga; Po materialam 70-x–90-x godov.* 2nd edn. Sankt-Peterburg: Folio-Press.

Novikov, Lev A. 1973. *Antonimija v russkom jazyke (Semantičeskij analiz protivopoložnosti v leksike).* Moskva: Izdatel'stvo Moskovskogo gosudarstvennogo universiteta.

Plotnikov, Bronislav A. 1979. Vidove leksikalni otnošenija meždu ednokorenni dumi v slavjanskite ezici. (Transl. Ilinka Rajkova.) *Săpostavitelno ezikoznanie* 4(2). 31–35.

Šanskij, Nikolaj M. 1964. *Leksikologija sovremennogo russkogo jazyka.* Moskva: Prosveščenie.

Šercl', Vikentij I. 1973. O slovax s protivopoložnymi značenijami (ili o tak nazyvaemoj ėnantiosemii). In Filin, Fedot P. (ed.), *Xrestomatija po istorii russkogo jazykoznanija*, 259–264. Moskva: Vyssaja škola. [Considerably shortened version of the article originally published in *Filosofskie zapiski* (Voronež), issue V–VI of 1883, 1–39, and issue I of 1884, 41–84.]

The integration of foreign words from European languages into the Cyrillic and Latin alphabets

There are several ways of integrating foreign words (including proper names) graphically. The aim of practical transcription is to represent the foreign pronunciation with native orthographic rules, whereas transplantation, transliteration, and exact phonetic transcription conserve the orthography and/or pronunciation of the original. A special feature of Cyrillic writing systems as opposed to Latin ones seems to be that they completely rely on native pronunciation rules and do not allow the integration of foreign orthographic conventions. That this is indeed a matter of alphabet rather than of language type becomes quite obvious when comparing Serbian with Croatian: The two languages essentially have the same grammar and the same orthography, but Croatian is always written in the Latin alphabet and Serbian uses Latin and Cyrillic letters equally, i.e. it has one orthography but two scripts (with biunique correspondences between the graphemes). Consequently, foreign words are always transcribed according to their pronunciation in Serbian (e.g. *Cirih* or *Цирих*), whereas into Croatian especially proper names are usually transplanted in their original orthography (e.g. *Zürich*). However, due to globalization and the use of computer-based text processors, transplantation of Latin-alphabet (especially English) words into Cyrillic texts is becoming more frequent, particularly in Russian.

Originally written in English as "Foreign words and Cyrillic" for the thematic session "Integration of foreign words" at the 24th annual conference of the Linguistic Society of Germany (DGfS) in Mannheim and presented on 28 February 2002, but published in Russian as "Integracija inostrannyx slov iz evropejskix jazykov v kirillice i latinice" in Kotjurova, Marija P. (ed.), *Filologičeskie zametki: Mežvuzovskij zbornik naučnyx trudov*, vypusk 2, vol. 1, 122–150. Perm': Permskij gosudarstvennyj universitet 2003. DOI: 10.18716/bun/iskl. — The original publication contains endnotes rather than footnotes.

DOI: 10.18716/omp.64.3

1. Introduction

This paper examines the peculiarities of graphical integration of foreign words into languages written in a Cyrillic alphabet, as opposed to languages that use the Latin alphabet. The features discussed here are primarily found in foreign words that are in the early stages of primary integration (according to Filipović 1990: 28, 48), rather than in adapted borrowings that are already in the process of secondary integration. Therefore, special attention is paid to foreign proper names, because they usually retain a connection with the source language for a very long time.

Different traditions of integrating foreign words into languages with different writing systems, in this case Cyrillic and Latin, are shown in this paper. Languages

that use the Latin alphabet prefer to preserve the original spelling of loanwords and readily introduce foreign grapheme-phoneme correspondences, whereas languages that use the Cyrillic alphabet prefer to convey the pronunciation of foreign words using their own spelling and never go beyond its boundaries.

To demonstrate this hypothesis and as a benchmark, three languages were selected that have different relationships to the Cyrillic alphabet: Russian (which uses exclusively Cyrillic), Serbian (which uses both Cyrillic and Latin script in parallel and equally), and Croatian (which is very close to Serbian but uses only the Latin alphabet). The languages of the former 🗋123 Soviet Union, which adopted the Cyrillic script in the 20th century, were not taken into account.[1] Foreign words from non-European languages, which are few in number in the languages of interest to us, were also left out of consideration.[2]

2. General possibilities for the graphic integration of foreign words

Before we turn to the peculiarities of Cyrillic, let us first have a general look at how foreign words (and names) can be integrated graphically. There are four different ways to do this: 1. transplantation, 2. transliteration, 3. practical transcription, and 4. scientific transcription.

2.1. Transplantation

Transplantation is a term borrowed from Superanskaja (1978: 26), defined as "direct inclusion of a foreign name in the text while preserving its spelling" ("непосредственное включение в текст иностранного имени с сохранением его графики", Giljarevskij & Starostin 1985: 11), as in the following examples:

(1) Сперва madame за ним ходила,
Потом monsieur ее сменил.
(Puškin, *Evgenij Onegin*, I, 3)

⟨Sperva *madame* za nim xodila,
Potom *monsieur* ee smenil.⟩
'At first *madame* looked after him,
Then *monsieur* took her turn.'

(2) After the concert, the crowd headed *en masse* for the parking lot.
(Webster 1993: 1355)

(3) "The cooking here is *wunderbar*," he said. (Webster 1993: 1355)

[1] 🗋141 For more details on the features of these languages, see note 29.

[2] Thus, we are primarily dealing with foreign words from the languages that Giljarevskij & Starostin (1985) write about integrating: English, Bulgarian, Hungarian, Dutch, Danish, Spanish, Italian, German, Norwegian, Polish, Portuguese, Romanian, Serbo-Croatian, Slovak, Finnish, French, Czech, and Swedish. For other languages, see Serdjučenko (1967).

Sometimes the graphic image of a foreign element changes slightly due to typographical issues:[3]

(4) German *K**ö**nigstra**ß**e* 'King Street' > *K**oe**nigstra**ss**e* (or, incorrectly: *K**o**nigstra**B**e*)
Croatian ***đ**uveč* 'rice dish with vegetables' > ***dj**uvec*

A problem arises when you have to add inflection or other native morphemes to transplanted words because then two different sets of grapheme-phoneme correspondences appear within one word. For example, in Croatian when a native morpheme is added to a foreign word stem the last letter of the stem is adapted to suit Croatian pronunciation:

(5) *Casablan**c**a*, gen. sg. *Casablan**k**e*
*Kenned**y***, gen. sg. *Kenned**yj**a*
*Haa**g***, adj. *haaš**k**i*

🗋 124 Slovene goes even a step further in that it treats foreign names like native nouns, i.e. if possible they are integrated into a native noun paradigm, for which the end of the foreign stem has to be reinterpreted morphologically:

(6) *Max Web**er***, gen. sg. *Maxa Web**r**a*
*Harold Pint**er***, gen. sg. *Harolda Pint**er**ja*
*Ol**ie** in Stanle**y*** 'Olie and Stanley', gen. sg. *Ol**ie**ja in Stanle**yj**a*

In Polish, the graphical word stem is left untouched, but where necessary the boundary between foreign and native pronunciation rules is indicated by an apostrophe (for more details, see Bartmińska & Bartmiński 1978: 40–50):

(7) *Harry*, gen. sg. *Harry'ego*, pronounced without the ⟨y⟩: [xa'rɛgɔ];
but loc./instr. sg. *Harrym* pronounced with the ⟨y⟩: ['xarɪm]

Such problems become even more complex when words are transplanted from one alphabet into another one that uses identical letter shapes with different denotations; e.g. the Serbian dative ending *-у* ⟨-u⟩ would be read as ⟨-y⟩ in a Latin context. Consequently, in Serbian orthography there is the following rule:

"Само изузетно се и у ћириличком тексту понеко име може навести и у изворној латиници (нпр. кад не можемо сазнати како се чита и преозвучава). Ако се таком имену додаје наш наставак, он се одваја цртицом, нпр.

"Only in exceptional cases can a name sometimes also be cited in original Latin script in a Cyrillic text (e.g. if we cannot find out how it is read and pronounced). If such a name is given our ending, it is separated by a hyphen, e.g. *Huet*-a, za

[3] This method can also be used to represent Cyrillic letters using Latin letters and numbers, e.g. *COBETCKUÜ CO1O3* < *COBETCKИЙ CO1O3*. However, this approach is usually only used when the shapes of the letters are identical, as in the title of Paul McCartney's 1987 album *CH0BA B CCCP* on non-Cyrillic web pages: https://www.mcbeatle.de/macca/a/choba.html (last checked 22 Mar 2026; note the file name!).

Huet-а, за Huet-овог противника […]."[4] *Huet*-ovog protivnika […]."[4]
(Pešikan et al. 1995: 103 f.)

Russian in such cases also uses a separating hyphen, but usually an apostrophe is used before case endings:

(8) IBM-совместимость ⟨*IBM*-sovmestimost'⟩ 'IBM compatibility'
Laptop, gen. sg. Laptop'a

2.2. Transliteration

🗋 125 If the source has to be rewritten in another alphabet, the best way to enable the educated reader to restore the original orthography is through transliteration (cf. Superanskaja 1978: 16–25), i.e. an expression written in a foreign alphabet is transferred into the alphabet of the borrowing language letter by letter. In an ideal case of transliteration, there is a one-to-one correlation between foreign graphemes and transliteration graphemes, as is customary in scientific transliteration systems, e.g. in the system for the scientific transcription of Cyrillic into Latin letters of the International Organization for Standardization (ISO) and its national variants, see (9), and in Serbo-Croatian transliteration, see (10):

(9) Russian Горбачёв > *Gorbačëv*
Russian: Хрущёв > ISO *Hrušëv*, German *Chruščëv*, English *Xruščëv*

(10) Croatian (and Serbian) *U početku bješe riječ, i riječ bješe u Boga, i Bog bješe riječ.*
↔ Serbian *У почетку бјеше ријеч, и ријеч бјеше у Бога, и Бог бјеше ријеч.* (Vuk 1847: John 1:1)

Such systems were first created for Prussian library catalogs (cf. Instruktionen 1899: 50–55; Superanskaja 1978: 17), which collected books written in a wide variety of writing systems that had to be compiled into a single alphabetical catalog. Libraries in countries using the Cyrillic alphabet did not need such transliteration systems, as books in Latin script were entered into a special catalog based on the Latin alphabet.

However, such transliterations, while conveying the spelling of the source, do not reflect its pronunciation. For the correct pronunciation of the surname ⟨Gorbačëv⟩, one has to know that ⟨č⟩ is pronounced as [t͡ʃ], ⟨ë⟩ as [ɔ], etc., just as with ⟨en masse⟩ transplanted into English it is necessary to know the French pronunciation rules.

[4] This refers to the French scholar and bishop Pierre-Daniel Huet [ɥɛ] (1630–1721).

2.3. Practical transcription

Practical transcription is a method for representing pronunciation rather than spelling.[5] In this approach, the spelling of the source language is only needed for 🗋126 "the determination of the phonemic composition of the foreign word and the translation of foreign phonemes into phonemes of the borrowing language" ("определение фонемного состава иноязычного слова и перевод иноязычных фонем в фонемы заимствующего языка", Superanskaja 1978: 103). This adapted pronunciation must be conveyed exclusively by "the means of normal orthography" ("средствами нормальной орфографии") of the recipient language, i.e. its own grapheme-phoneme correspondences (Reformatskij 1955: 296):

(11) Russian *Уолфиш-бей* ⟨Uolfiš-bej⟩ < English *Walfish Bay*
Serbian *Хјуман рајтс воч* ⟨Hjuman rajts voč⟩ < English *Human Rights Watch*
German *Büro* 'office' < French *bureau*
Slovene *pižama* 'pajamas' < French *pyjama*

Of course, such grapheme-phoneme correspondences exist only for the native sound system, so that in many cases the pronunciation of the original cannot be conveyed completely:

(12) Russian: *Оруэлл* [ɔ.ru.ˈɛɬ] < *Orwell* [ˈɔː(ɹ).wəɬ]
Slovene *bife* [biˈfeː] 'buffet' < French *buffet* [byfɛ]
German *Samowar* [zamoˈvaːɐ̯] (often [ˈzamovaˑ]) < *Russian самовар* [səmʌˈvar]

In cases where this does not cause pronunciation difficulties, additional information about the original spelling is included in the practical transcription using the transliteration method, e.g. when reduced vowels or double consonants in English are reflected orthographically in Cyrillic:

(13) Russian *Лондон* ⟨London⟩ < English *London*,
Ллойд Уэббер ⟨Llojd Uėbber⟩ < English *Lloyd Webber*,
Хоуард ⟨Xouard⟩ < English *Howard*,
Гюнтер Грасс ⟨Gjunter Grass⟩ < German *Günter Grass*

In her book on practical transcription, Superanskaja (1978) presents this method as a kind of compromise between transliteration, which conveys only the spelling while neglecting the pronunciation, and phonetic transcription, which delves too deeply into the subtleties of foreign phonetics, making the text incomprehensible to the average reader. Therefore, these extremes, i.e. "blind adherence to the letter" ("слепое следование букве") and "excessive phoneticism" ("излишний фонетизм"), should be avoided (Superanskaja 1978: 131–134). However, due to the

[5] For details on practical transcription, see Superanskaja (1978). The term has been in use for a long time; it already appears in the article "Transkripcïja" in the Russian Brockhaus (Arsen'ev & Petruševskij 1901).

limitations of the grapheme-phoneme correspondences of the recipient language, it is impossible to consistently convey the phonetics, phonology, *or* spelling of the integrated words. The only advantage 🗋 127 of practical transcription is the organic integration of transcribed words into the borrowing language. Such words behave orthographically and phonetically like words of the recipient language. Therefore, practical transcription in all languages is an important element of secondary adaptation (cf. Filipović 1990: 28, 48) in the further integration of borrowed words, cf. the German example *Büro* in (11), which a few decades ago was usually written as ⟨Bureau⟩. However, this article is not about this second step of borrowing but about the graphic integration of foreign words that have not yet been adapted, as in the Russian and Serbian examples in (11).

2.4. Scientific transcription

An appropriate representation of the original pronunciation can only be achieved by scientific transcription. This term is used here to refer to transcription that is not oriented towards preserving the orthographic norms of the borrowing language. Outside scientific contexts (in which you might use IPA or the like), this method is employed especially for languages with a non-alphabetical writing system, such as the Pinyin[6] system for Chinese:

(14) Pinyin *Dèng Xiǎopíng* [tɤŋ\ɕi̯aʊ̯ˇpʰiŋˊ]
Pinyin *kūnqǔ* [kʰu̯ən˥ t͡ɕʰyˇ] ‘form of Chinese opera’

In these examples, ⟨x⟩ is used for a voiceless alveolo-palatal fricative [ɕ], and ⟨q⟩ for an aspirated voiceless alveolo-palatal affricate [t͡ɕʰ]. Though this transcription makes use of Latin letters, their pronunciation differs from the general pronunciation rule of ⟨x⟩ as [ks] and ⟨qu⟩ as [kw] in a target language like English.[7] Thus, both scientific transcription and transliteration (and, of course, transplantation) introduce unfamiliar grapheme-phoneme correspondences; one might call this **graphical code-switching**. Only practical transcription makes do with the pronunciation rules (the code) already present in the native orthography.

[6] This system is used in many languages and also within China. In Russian, the most common transcription system is that of Father Palladij (legal name: Pëtr Ivanovič Kafarov; cf. Serdjučenko 1967: 236), which, like the Pinyin system, conveys almost all the distinctive oppositions of Chinese phonology (except for suprasegmental ones), but does not go beyond the means of Russian orthography. For more details on this system, cf. Serdjučenko (1967: 231–257).

[7] In the Russian Palladij system, the pronunciation of the corresponding letters *c* ⟨s⟩ and *ц* ⟨c⟩, [sʲ] and [t͡sʲ], is much closer to the Chinese sounds [ɕ] and [t͡ɕʰ].[8] The endings, of course, changed for morphological rather than graphic reasons.

3. Differences between Cyrillic and Latin scripts

🗋 128 A special feature of Cyrillic writing systems as opposed to Latin ones seems to be that they completely rely on native pronunciation rules and do not allow the integration of foreign orthographic conventions (whether from a Latin, Cyrillic, or other source). Yet, this has not always been the case. Pešikan et al. (1995: 103) correctly observe that the medieval Cyrillic alphabet included all the letters of the Greek alphabet (many of which were not needed for Slavic words) and could therefore represent words of Greek origin in their original orthography. This applies especially to the following letters:

(15)	Greek letters:	⟨ι⟩ ⟨η⟩ ⟨υ⟩	⟨φ⟩ ⟨θ⟩	⟨ο⟩⟨ω⟩	⟨ξ⟩	⟨ψ⟩
		| | |	| |	| |	|	|
	Old Cyrillic letters:	⟨і⟩ ⟨и⟩ ⟨ѵ⟩	⟨ф⟩⟨ѳ⟩	⟨о⟩⟨ѡ⟩	⟨ѯ⟩	⟨ѱ⟩
	Slavic pronunciation:	[i] [i] [i]	[f] [f]	[o] [o]	[ks]	[ps]
		\ | /	| /	| /	¦	¦
	modern Russian letters:	⟨и⟩	⟨ф⟩	⟨о⟩	⟨к⟩ + ⟨с⟩	⟨п⟩ + ⟨с⟩

Consequently, e.g. in the words for 'incense' and 'Europe', (16a), which in Church Slavonic were written strictly according to their etymology, (16b), some letters in Russian were replaced with other Cyrillic letters, (16c), which also derive from Greek letters, but different ones, (16d):

(16)	a.	Greek:	θυμίαμα	Εὐρώπη
	b.	Church Cyrillic:	ѳѵмїамъ[8]	Єѵρωпа[8]
	c.	modern Russian:	фимиам	Европа
	d.	corresponding Greek letters:	*φημήαμ(α)	*Ἐβρόπα

Over time this Cyrillic alphabet, which had been designed for Old Church Slavonic, was adapted to various (Slavic and non-Slavic) national languages, and in these new orthographies the representation of foreign words immediately shifted to (practical) transcription:

> "Али се са стварањем националних ћирилица угасила могућност систематског изворног писања имена, не само из грчког него и из других словенских ћириличких језика."
> (Pešikan et al. 1995: 103)

> "But with the creation of national Cyrillic alphabets it became impossible 🗋 129 to write names, not only from Greek but also from other Cyrillic Slavonic languages, systematically according to the original."

This conclusion, however, is not at all self-evident. It is not impossible in the German or English 'national' alphabet to write, for example:

(17) German *Nishnij Nowgorod* (more frequently than *Nischni Nowgorod*)
< Russian *Нижний Новгород* ⟨Nižnij Novgorod⟩,

[8] The endings, of course, changed for morphological rather than graphic reasons.

German *Walesa* (not **Wauengßa*) < Polish *Wałęsa*,
German *Washington* (not **Woschingten*) < English *Washington*,
English *Birobidzhan* (not **Beerabeejahn*) < Russian *Биробиджан* ⟨Birobidžan⟩,
English *Weiss* (not **Vice*) < German *Weiß*

Note how natural it seems to use unfamiliar grapheme-phoneme correspondences for foreign names and how unthinkable most of the alternative 'phonetic' representations are. However, just as unthinkable as this use of native grapheme-phoneme correspondences for foreign proper names in English or German is the use of foreign grapheme-phoneme correspondences in Russian or Serbian:[9]

(18) Russian **Схакеспеаре* ⟨Sxakespeare⟩ (pronounced [ʃɛˈkspir]) < English *Shakespeare*
Serbian **буффетӣ* ⟨buffet⟩ (pronounced [ˈbifɛː]) < French *buffet*

Instead, all the Cyrillic-written languages use practical transcription to convey foreign words (including proper names):

(19) Russian *Шекспир* ⟨Šekspir⟩ 'Shakespeare'
Serbian *бифе* ⟨bife⟩ 'buffet'

Practical transcription is even used for words borrowed from a source language that uses the Cyrillic alphabet, although at first glance it would seem that nothing needs to be rewritten here:

(20) Russian *Джинджич* ⟨Džindžič⟩ < Serbian *Ђинђић* ⟨Đinđić⟩

(21) Serbian *Љењин* ⟨Ljenjin⟩ < Russian *Ленин* ⟨Lenin⟩
Serbian *Кијев* ⟨Kijev⟩ < Russian[10] *Киев* ⟨Kiev⟩

Note that the Russian words in (21) do not contain any letters not present in the Serbian alphabet.

In Andrej Kurkov's book *Dobryj angel smerti* (*The Kind Angel of Death*, 2000), 130 even the direct speech of characters speaking Ukrainian (and the standard language at that) is subjected to practical transcription, i.e. a large amount of text is not integrated into the Russian context in any way:[11]

[9] This statement holds for all languages written in a Cyrillic alphabet. The name of a former Macedonian band, *Блооб ӣхе буббле* ⟨Bloob the bubble⟩ (cf. *Makedonija Europe*, 16–29 Nov 2001, 12), is obviously 142 an artistic deviation from the rule, making fun of exactly this restriction. With the usual Macedonian grapheme-phoneme correspondences this name would have to be read as [ˈblɔ.ɔp txɛ ˈbub.blɛ] rather than like English *Bloob the bubble* [ˈblʊb ðə ˈbʌ.bl̩].

[10] In Ukrainian the name of the capital is *Київ* ⟨Kyïv⟩ (correctly pronounced [ˈkɪjɪu̯]), but most languages have borrowed it either from Russian or from an older stage of East Slavic, where the pronunciation must have been *[ˈkijɛw].

[11] According to the author, this peculiar method of conveying the Ukrainian language is a reaction to the same treatment of Russian words in Ukrainian by Ukrainian purists

(22) — Завтра мы з Галэю йидэмо в Кыйив, — наконец произнес он. — Вы залышытэсь тут, я з батькамы вжэ поговорыв. Якщо щось узнаю — подзвоню!… — Хорошо, — кивнул я. (Kurkov 2000: 264)

In Ukrainian orthography:
— Завтра ми з Галею їдемо в Київ […]. Ви залишитесь тут, я з батьками вже поговорив. Якщо щось узнаю — подзвоню!… […]

'"Tomorrow, Halja and I are going to Kyiv," he finally said. "You stay here, I've already talked to my parents. If I find out anything, I'll call you!"
"Okay," I nodded.'

3.1. Serbian

Serbian has a rather peculiar orthographic situation: It uses not only the Cyrillic alphabet (as seen above) but *latinica* and *ćirilica* equally. These alphabets are mutually transliteratable on the basis of biunique correspondences between the graphemes.[12] Transliteration tables can be found at the beginning of every grammar, orthography, dictionary, or textbook, and the two alphabets have absolutely equal rights. In this respect Serbian seems to be unique in the world.[13] However,

(personal communication, Cologne, 3 October 2003). To emphasize their incorrectness, they are often transcribed, e.g. "Канєшно, знают, но, понімаєш, они на то, навєрно, вніманія не обращают…" (Hnatkevyč 2000: 6), where the Russian words *конечно*, *понимаешь*, *наверно*, *внимания*, etc. are used instead of the Ukrainian "Звичайно, знають, але, розумієш, вони на це, напевно, не звертають уваги" ("Of course they know, but, you understand, they probably don't pay attention to that").

[12] The only exception is that the Cyrillic letters ⟨љ⟩, ⟨њ⟩, and ⟨џ⟩ correspond to the Latin digraphs ⟨lj⟩, ⟨nj⟩, and ⟨dž⟩, respectively. However, the latter are also valued as single letters (even in post-1990 Croatian orthographies) and appear as separate entries in dictionaries and other alphabetical lists (because they represent the distinct sounds [ʎ], [ɲ], and [d͡ʒ]): *a b c č ć d dž đ e f g h i j k l lj m n nj…*

[13] Similar situations can only be observed in history when there was a slow shift from one writing system to another (e.g. from Glagolitic to Cyrillic in Old Church Slavonic). Nowadays the literacy of large parts of the population and a public school system demand for quick changes sanctioned by official laws. The alphabetic dualism in Serbian is a vestige of Illyrism and Yugoslavism, which wanted to merge Serbian and Croatian traditions.
Today the situation is not resolved because this would be a decision between 'eastern' Serbian national traditions (Cyrillic) and a clearly 'western' European orientation (Latin). Nikčević's new "Montenegrin" orthography (1997; 2001: 19) has inherited this problem (although Nikčević himself apparently only uses the Latin alphabet in his books). Officially the same holds true for Bosnian, which in practice, however, is written only in the Latin alphabet.
Other contemporary languages with two parallel writing systems are not comparable to Serbian: In China, Korea, and Japan, the Latin alphabet is not at equal rights with the

the preference of one or the other of the two scripts can be a political choice rather than one of taste: until 1990 the state encouraged the use of the Latin alphabet in the 'eastern variant of Serbo-Croatian' to stress the unity of Yugoslavia, whereas with the rise of Milošević's nationalism the Cyrillic alphabet has become more fashionable.

Therefore there is only *one* Serbian orthography, which can be realized in *two* alphabets, with the Latin realization being at the same time the official transliteration of the Cyrillic realization, and vice versa. The most influential newspaper of Serbian diaspora, *Vesti* (or *Вести*), for example, appears every day in two editions with completely identical text, one in Latin and one in Cyrillic script.

131 Consequently, in both Cyrillic and Latin Serbian texts, all foreign words (including names) are transcribed according to their pronunciation in Serbian:[14]

(23) *Cirih* or *Цирих* < German *Zürich*
Brisel or *Брисел* < French *Bruxelles* or Dutch *Brussel*

In 1950 the Serbian linguist Aleksandar Belić (1950: 110) stated this as a rule in his *Serbo-Croatian orthography*:

> "Према основном принципу нашег правописа – да се речи пишу онако како се изговарају – треба и речи страних језика писати онако како се изговарају."
>
> "According to the general principle of our orthography that words are written as they are pronounced, words from foreign languages must be written as they are pronounced as well."

After the 1954 Novi Sad Agreement, which officially declared Serbian and Croatian one language, the unified orthography of 1960 (cf. Pravopis 1960a, 1960b), as well as subsequent orthographies, accepted both transcription and transplantation for foreign names in a Latin text; this is true even of the newest Serbian orthographies, e.g. Pešikan et al. (1995: 103):

> "Туђа властита имена пишу се у српском стандардном језику на два начи-
>
> "Foreign proper names are written in two ways in Standard Serbian: adapted spell-

respective East Asian writing system [the original unpublished English version of 2002 added: "which is implied by the term *digraphia*"—D. B. 2026]. The equal use of roman type and blackletter (or 'Gothic') in German and other languages until World War II or of civil and church Cyrillic in Russian and other languages were a question of typeface and not of writing system: Obviously ⟨abcdefghi…⟩ and ⟨abcdefghi…⟩ or ⟨аввгдежзи…⟩ and ⟨абвгдежзи…⟩ are no more different alphabets than, say, 143 ⟨abcdefghi…⟩ and ⟨*abcdefghi*…⟩ are. [This seem to have been my very first published thoughts about the topic later treated extensively in Bunčić 2016.—D. B. 2026]

[14] This rule is sometimes followed in Latin script even when the need for clarification is so obvious that the original spelling is given in brackets: "Enterprajz Ajrland (Enterprise Ireland)" (Radomirović 2001: 7).

на: прилагођено писање (применљиво у ћирилици и у латиници, нпр. – *Минхен*, *Шекспир* – *Minhen*, *Šekspir*) и изворно писање (применљиво у латиници, нпр. *München*, *Shakespeare*)."

ing (applicable to the Cyrillic and Latin script, e.g. Минхен, Шекспир—*Minhen*, *Šekspir*) and original spelling (applicable to the Latin script, e.g. *München*, *Shakespeare*)."

However, in practice foreign proper names in Serbian texts have always been transcribed, cf. ibid.:

"У нашој пак двоазбучној пракси прилагођено писање широко се примењује и у латиници […]. Пишући прилагођено страна имена, аутор може припремати свој рукопис не условљавајући да ли ће се штампати ћирилицом или латиницом."

132 "Yet in our bialphabetic practice the adapted orthography is widely used also in the Latin script. […] Writing foreign names in an adapted way, an author can prepare a manuscript with no implication of whether it will be printed in the Cyrillic or Latin alphabet."

Only in commercial advertisements can some exceptions be spotted, e.g. the words *pizza* and *pizzeria*, which are often transplanted, even on posters with otherwise Cyrillic inscriptions. In coherent Serbian texts, however, one finds almost exclusively *pica* and *picerija* if the text is Latin[15] and *пица* and *пицерија* if it is Cyrillic. Foreign abbreviations are usually transliterated according to the scheme for native Serbian words:

(24) *ЦХФ* ⟨CHF⟩ (for 'Swiss Franc')
УНХЦР ⟨UNHCR⟩ (for the United Nations High Commissioner for Refugees)

Note that for being transliteratable such abbreviations must not contain the letters *Q*, *W*, *X*, and *Y*, since they are not used in Serbian, so that there is no Cyrillic equivalent for them. If an abbreviation does contain these letters, the long form is translated into Serbian and then this translation is abbreviated (as is also customary in Russian):

(25) *СЗО = SZO < Светска здравствена организација < WHO < World Health Organization*

3.2. Croatian (for comparison)

It remains to be proven that this difference between Cyrillic and Latin-based languages is indeed a matter of alphabet rather than of language type. This becomes obvious when comparing Serbian with Croatian: The two languages have essentially the same grammar and the same orthography—and were indeed viewed as varieties of one Serbo-Croatian language until very recently. Moreover, many

[15] See footnote 17.

linguists would still agree with 🗋133 Miloš Okuka's (1998: 130) statement about Serbian, Croatian, and Bosnian:

> „Drei unterschiedliche Bezeichnungen stehen vom linguistischen Standpunkt her betrachtet für ein und dieselbe Sprache, suggerieren aber gleichzeitig vor allem das Unterschiedliche, obwohl ‚die Verschiedenheiten so gering sind, daß man sie wirklich suchen muß' (J. Raecke)[16]."
>
> "From the point of view of linguistics, three different names denote one and the same language, while at the same time suggesting first of all difference, although 'the differences are so slight that you really have to look for them' (J. Raecke)[16]."

This is why Croatian linguist Snježana Kordić (2001: 237), unlike many of her compatriots, quite correctly insists on the possibility of using the adjective *Serbo-Croatian* today as a linguistic term that unites the Serbian, Croatian, and Bosnian "political languages" as national varieties into a single "linguistic language."

The most obvious difference between Serbian and Croatian is that Croatian is *never* written in Cyrillic letters. This is why in contrast to the Serbian examples (11), (18), (19), and (23), in Croatian some foreign words and especially proper names are spelled as they are spelled in the source language:

(26) *Human Rights Watch*, *buffet*[17], *Zürich*, *München*, *Shakespeare*, *pizza*[17], *pizzeria*[17]

---

[16] This quotation, though without further references, is obviously taken from Raecke (1996: 21), who concludes (ibid. 22): "[Daraus] ergeben sich […] Probleme, die ein Slawist nur beschreiben, aber nicht lösen kann." ("This causes problems that a Slavist can only describe but not solve.")

[17] Although Anić (1991, 1994) and Barić et al. (1999) only give *bife*, Ćirilov (1994: 18) correctly describes the orthographic reality: "U Hrvata je pravilno *buffet*." ("Among the Croatians, *buffet* is correct.") The same applies to the words 'pizza' and 'pizzeria': Barić et al. (1999) present ⟨picerija⟩ as a better option than ⟨pizzeria⟩, and give only ⟨pica⟩, which is explained as "'pizza'". Anić (1991, 1994) only gives ⟨picerija⟩, but the definition of this word contains the word *pizza* (which in the second edition of 1994 is given in the form "pica (pizza)", whereas in the first edition of 1991 it does not appear as a headword at all). An internet search using the *Google* search engine (conducted on 17 July 2003) finds the following number of pages whose language the search engine has identified as "Croatian" and "Serbian":

	Croatian		Serbian	
buffet	1180	**83%**	40	6%
bife	236	17%	585	**94%**
pizzeria	1980	**89%**	198	29%
picerija	239	11%	476	**71%**

(Searching for the word 'pizza' is pointless, as *Google* does not distinguish ⟨c⟩ from ⟨ć⟩, i.e. *pica* 'pizza' from *pića* 'drink(s)'. In addition, the word *pica* [pronounced ['pi:ʌʧsa], in contrast

All Croatian typewriters contain keys for the letters *q*, *w*, *x*, and *y*, although they are not needed for native words, and Croatian computer keyboards even provide key combinations for such symbols as *ß*, *ü*, *é*, etc. Of course, they are used only for foreign words that are not yet integrated and for foreign proper names. When loanwords are integrated, they adopt a more native spelling:

(27) *paviljon* < French *pavillon*
koledž < English *college*

3.3. Russian

🗋 134 In contrast to Serbian, Russian is written exclusively in Cyrillic letters. However, in the years after the revolution, the Latin alphabet seemed more modern, so that some seventy languages of the Soviet Union were given a Latin-based orthography, and in 1929 a Subcommittee on the Latinization of the Russian Language was established under the Main Committee on Science of the People's Commissariat of Education, headed by Nikolaj Jakovlev (cf. Alpatov 2001, Ašnin & Alpatov 2001). However, by the time the subcommittee had finished their work, the enthusiastic belief in a quick world revolution was over and Stalin's policy then included a Russification of the whole country. Therefore, in 1937–1941 all the newly latinized languages of the USSR were cyrillicized.[18] Thus, the only changes in Russian orthography implemented by the communists were those drawn up by the Academy of Sciences on 11 May 1917 and confirmed by the Provisional Government's Ministry of People's Enlightenment on May 17. This reform retained the identical representation of morphemes (regardless of regular morphonological alternations) but removed the differentiation of homonyms, thus abolishing some

to ['pi\t͡sa] 'pizza'; cf. also Šipka 2008—D. B. 2026] is also a Serbian and Croatian slang word meaning 'cunt'; cf. Andrić 1976, Sabljak 2001). Of course, such a search cannot be considered a scientifically sound corpus analysis, but these numbers still give some idea of the linguistic reality. [In 2026, the absolute numbers are much higher, e.g. 475,000 vs. 87,600 for Croatian *buffet* vs. *bife*, but the relations are roughly the same, at least for Croatian. For Serbian, *pizzeria* has risen to 47% and *buffet* even to 65% in the Google hits (even with an additional restriction to websites from Serbia), but this seems mainly due to multilingual websites that were identified as Serbian and to article names like the brand *Buffet Crampon* for wind instruments.—D. B. 2026]

[18] For example, Azerbaijani, which had had a long literary tradition in an Arabic writing system, was written in a 'Socialist' Latin alphabet from 1925/29 to 1939 and in a Cyrillic alphabet from 1940 onward. In 1991/93 the language was re-latinized, but this time on the model of the Turkish orthography (rather than the Azerbaijani orthography used prior to 1939). The Moldovan language underwent even more 🗋 144 alphabet changes: 1924–1928 Latin, 1928–1932 Cyrillic, 1933–1937 Latin, 1937–1988 Cyrillic, since 1989 Latin again (cf. Vilku-Pustovaïa 2001: 395).

alternative graphical representations for one and the same sound (cf. (15)). As a result, the base of words like the following is now spelled identically:

(28) *миръ* ⟨mir″⟩ ‘peace’, gen. sg. *мира* ⟨mira⟩ → *мир* ⟨mir⟩, *мира* ⟨mira⟩
міръ ⟨mïr″⟩ ‘world’, gen. sg. *міра* ⟨mïra⟩ → *мир* ⟨mir⟩, *мира* ⟨mira⟩
мyро ⟨mÿro⟩ ‘chrism’, gen. sg. *мyра* ⟨mÿra⟩ → *миро* ⟨miro⟩, *мира* ⟨mira⟩

Another difference between Serbian and Russian is that there is no unified transliteration for Russian, neither from Cyrillic to Latin nor vice versa. In contrast to Japan and China, in Russia there does not seem to be any need for such a system. Instead, a transliteration is used for international telegrams written in Russian, an English transcription for international passports, and a French (!) transcription for naturalization documents in Germany, whereas western libraries and scholars use various 🗋 135 transliteration systems (all of them inspired by Czech orthography), e-mail transcriptions are chaotic, and the official ISO standard is almost never used.[19]

Despite all these differences, however, in Russian, just as in Serbian, foreign proper names are always transcribed according to native orthographic rules (cf. (11), (12), (19), (20); for details on the rules of adaptation, cf. Giljarevskij & Starostin 1985). But the phonological basis of practical transcription is often mingled with elements of transliteration, e.g. in the case of double consonants and English vowels. Double consonants are pronounced either as long or short consonants, and English vowels are often pronounced as they are written in English (rather than as they are pronounced, as suggested in the transcription):

(29) *Уолл-Стрит Джорнал* ⟨Uo**ll**-Strit Dž**o**rn**a**l⟩ for *Wa**ll** Street J**ou**rn**a**l*
Уол-Стрит Джернл* *⟨Uol**-Strit Dž**e**rn**l**⟩ < [ˈwɔːlˌstɹiːtˈd͡ʒə:(ɹ)nl̩]

This combination of transcription and transliteration and the change of transcription traditions through time has produced many doublets:

(30) *Юккер, Уэккер, Уеккер* ⟨Jukker, Uėkker, Uekker⟩ < *(Günther) Uecker*
Хоуард, Хауард, Говард ⟨Xouard, Xauard, Govard⟩ < *(Sir Ebenezer) Howard*

[19] Several transliteration and transcription systems were compiled in 1957 by A. A. Reformatskij (Superanskaja 1978: 22ff.), and in 2001 by Lydia Winschel and Andreas Prilop, on the website https://web.archive.org/web/20030605095123/http://www.unics.uni-hannover.de/ntr/russisch/umschrifttabelle.html (last checked 22 Mar 2026). [Nowadays, the Wikipedia article “Romanization of Russian” (https://w.wiki/Giar) gives a very good overview. — D. B. 2026.]
Although the ISO system is indeed somewhat clumsy in practice, Superanskaja’s criticism (1978: 20) is completely unjustified: She claims that the Russian surnames Ёлкин, Елькин, and Элькин are not distinguished in the ISO system and are transliterated identically as ⟨Elkin⟩, although Reformatskij’s table (ibid. 22ff.) clearly shows that according to the ISO system, these surnames are written as ⟨Ëlkin⟩, ⟨El′kin⟩, and ⟨Èl′kin⟩, respectively.

Уолпол, Вальпол ⟨Uolpol, Val'pol⟩ < *(Robert) Walpole*
Хадсон, Гудзон[20] < English *(Henry) Hudson* (and the river, bay, and strait named after him)

Occasionally, however, foreign words are transplanted into Russian texts in their Latin orthography: On the one hand, there are foreign quotations (especially in scholarly texts, where even Greek is usually latinized!). On the other hand, transplantations especially from English have recently become more and more frequent in popular texts. Consequently, the *Dictionary of the Russian Language of the End of the 20th Century* (Skljarevskaja 1998: 697–700) contains an appendix of computer terms frequently used in Latin orthography.[21] Outside of 136 advertising, electronics and science, however, Latin letters nowadays seem to be even rarer than they were in the 19th century.

4. Reasons for differences between Cyrillic and Latin scripts

All in all, it seems to be a fact that Cyrillic alphabets tend to resist the integration of foreign grapheme-phoneme correspondences, whereas Latin alphabets much more readily accept them. What is the reason for this difference? One might connect it to the observation that, taken as a whole, in Cyrillic orthographies the phonological principle is more important than it is in Latin-based orthographies. In Cyrillic there is a range from the moderately 'morphological' Russian to the radically 'phonetic' Serbian orthography, whereas aside from the rather perfect 'phonetic' representation of Finnish, Latin orthographies include almost 'logographic' examples like English or French and a lot of intermediate cases like German, Danish, or Polish.

Examining this question with regard to the rather phonological Spanish and the etymological French orthography, Meisenburg (1989: 251) tries to show "that the different structures of their writing systems are mainly caused by the specific phonetic evolution that took place in each language after it had first been recorded in writing" ("daß diese unterschiedliche Strukturierung vor allem durch die jeweilige lautliche Weiterentwicklung nach der ersten schriftlichen Fixierung der Sprache bewirkt worden ist"), i.e. a language which has had a long history since it

[20] The last example, ⟨Гудзон⟩, is cited by Vinogradov (1977: 149) as an "incorrect" transcription because it does not sufficiently reflect the original pronunciation. Despite his condemnation, it is this spelling that is still traditionally considered the main one, e.g. by Proxorov (1997).

[21] Before, entries in the Latin script had only existed in dictionaries of foreign words (if at all); there is even a two-volume dictionary that contains exclusively foreign expressions in their original spelling (Babkin & Šendecov 1966). The last universal Russian dictionary with a Latin appendix was Ušakov (1935–40: vol. 4, 1473–1484). It contained words and expressions like *conditio sine qua non* or *bon mot* (with a reference to *бонмо* ⟨bonmo⟩), some of which are illustrated by quotations from Puškin, Tolstoj, and Lenin.

began to be written tends to have a more etymological orthography than a more recently alphabetized language.

As for Cyrillic compared to the Latin alphabet, Meisenburg's approach might indeed explain the way Serbian behaves, although it would come rather short of Vuk Karadžić's deliberate choice of the radical phonological principle ("Пиши као што говориш!"—"Write as you speak!"). But Russian orthography is an unbroken continuation of the Old Church Slavonic one, with several slight corrective reforms. For example, the Russian letters ѣ and ѣ are used to this day, although the sounds they 🗋137 used to represent vanished in the 14th century. So Meisenburg's argumentation cannot explain the fundamentally different integration of foreign words (especially names) in Cyrillic and Latin orthographies.

However, when looking at 'intra-Cyrillic' transcriptions such as (20) and (21), one can come to the conclusion that our question might be explained by the typographic structure of the Cyrillic alphabets: As Kempgen (1993: 29) observes, the "diacriticity" ("Diakritizität") of Latin orthographies is higher than that of Cyrillic ones, i.e. Latin orthographies make much wider use of diacritical marks to represent sounds not covered by the canonical 26-letter Latin alphabet. They also use many more digraphs and trigraphs (so their 'digraphicity' is higher, too), whereas Cyrillic orthographies usually invent completely new letters:[22]

[22] The only exceptions I know (apart from languages very recently alphabetized on the basis of the International Phonetic Alphabet) are on the one hand the Latin letters ⟨þ⟩ (the 'thorn' rune) and ⟨ð⟩ (a slashed Insular ⟨d⟩) in Icelandic and ⟨ß⟩ (originally a ligature of ⟨ſ⟩ + ⟨ʒ⟩ or ⟨ſ⟩ + ⟨ſ́⟩, i.e. ⟨sz⟩ or ⟨ss⟩) in German. On the other hand, there are a few diacritics in Cyrillic: the breve above Russian ⟨й⟩ (and later also Belarusian ⟨ў⟩) has its origin in Church Slavonic accent notation (cf. Keipert 1999: 253–255), which in turn is an imitation of Greek suprasegmentals. Russian ⟨ё⟩ was invented by Karamzin (a definite 'westerner') in 1797 but is still mostly replaced by ⟨е⟩ without the diaeresis. [We now know that ⟨ё⟩ was invented in 1783 by Princess Ekaterina Daškova, up to now the only ever female president of the Russian Academy of Sciences (cf. Pčelov & Čumakov 2000: 13–16) and that the two dots are not a diaeresis but an umlaut (cf. Bunčić 2017: 25f.).—D. B. 2026.] Ukrainian ⟨ї⟩ was introduced only in 1873 (cf. Bilodid 1958: 314); previously, the number of dots above the ⟨і⟩ had played no role, so that there had been three allographs: ⟨ı, ї, і⟩. Macedonian ⟨ѓ, ќ⟩ and 'Montenegrin' ⟨ǵ, ć⟩ 🗋145 (cf. Nikčević 1997, 2001: 19) were obviously designed on a Latin (namely, Polish) model in the 20th century.
Kempgen (1992: 128) also mentions "diacritical elements" *within* the letter structure, e.g. the 'tails' in Russian ⟨щ⟩ vs. ⟨ш⟩ and ⟨ъ⟩ vs. ⟨ь⟩ or in Serbian ⟨ц⟩ vs. ⟨ц⟩ and ⟨ђ⟩ vs. ⟨ћ⟩. However, such elements are part of the letter itself and therefore differ significantly from diacritical marks in the strict sense.

(31)

orthography	based on	diacritics	digraphs	new letters
Slovene	Latin	č š ž	—	—
Polish	Latin	ą ć ę ł ń ó ś ż ź	cz ch dz dź sz[23]	—
Hungarian	Latin	á é í ó ö ő ú ü ű	cs gy ly ny sz ty zs	—
English	Latin	—	ch ck ee ea gh ng oo oa ou sh th wh …	—
French	Latin	à â è ê é ë î ô û	au ch eau gn ou oi …	—
German	Latin	ä ö ü	ch ck ie ng sch …	ß
Old Church Slavonic	Greek	—	ѕι	ѣ ѥ ж ц ч ш щ ъ ь ю ꙗ ѧ ѩ ѫ ѭ[24]
Russian	Cyrillic	ё, й	—	э
Serbian	Cyrillic	—	—	ђ ј љ њ џ (ћ)
Kazakh	Russian	ғ қ ң	—	ә ө ұ ү h

It is not hard to see that transplantation is easier if the only foreign elements you have to cope with are digraphs and trigraphs (which do not cause any problems) and diacritics (which can be combined, added by hand, or just left out, cf. (4)). But how should the Serbian name *Јовановић* ⟨Jovanović⟩ be transplanted into a Russian text if the letter case, typewriter, or computer keyboard[25] does not contain the letters *J* ⟨J⟩ and *ћ* ⟨ć⟩? Furthermore, even in a consistent transliteration you will have to use diacritics or at least digraphs if you want the transliteration to be complete. Cyrillic has never used digraphs for the primary Slavic languages,[26]

138

[23] Before a vowel, the acute above the palatal consonants *ć, dź, ń, ś, ź* is replaced by a postponed diacritical *i*, e.g. ***dzi****eń* 'day', not **dź**eń***, so that if this letter were part of graphic combinations denoting soft sounds, there would be even more digraphs in Polish orthography *(bi, ci, fi, gi, ki, mi, ni, pi, si, wi, zi)* and even one trigraph *(dzi)*, which would make the graphemic model very uneconomic.

[24] The letter ⟨ѱ⟩ is probably a variant of the Greek letter *psi* ⟨ψ⟩ with a new phonetic value. The traditional explanation of ⟨ѱ⟩ as a ligature of ⟨т⟩ under (!) ⟨ш⟩ to denote the sound [ʃtʲ] is implausible, especially because of the shape of this letter in Glagolitic (cf. Trunte 1997: 49, 55f.).

[25] With the introduction of *Windows 95*, it has become much easier to use different languages on a computer: the Russian system now includes not only an English keyboard but also options for many other languages, including Serbian. However, most users do not know (and are not sufficiently interested in) how to use them.

[26] The only exception are the digraphs ⰻⰾ ⟨ǵl⟩ and ⰻⰿ ⟨ǵn⟩ for [ʎ] and [ɲ] written by Franciscan friars in a Cyrillic script traditionally called *Bosančica*—without doubt on the

so there is no tradition to do so for foreign languages either. In the Latin alphabet, digraphs have already been used by the Romans to represent the Greek letters ⟨φ⟩, ⟨θ⟩, ⟨χ⟩, and ⟨ψ⟩ as ⟨ph⟩, ⟨th⟩, ⟨ch⟩, and ⟨ps⟩.[27]

We are dealing here with traditions deeply rooted in culture, as Jušmanov (1933, cited in Superanskaja 1978: 26) already pointed out: "whereas for the West it is important how to write, for the East it is important how to read" ("если для Запада важно, как написать, то для Востока — как прочесть"). The cultural concepts that Jušmanov referred to as the West and the East can be correlated in the Slavic languages with the concepts of *Slavia Romana* and *Slavia Orthodoxa* (cf. Picchio[28] 1984, 1991), which are similar to the border between the Latin and Cyrillic alphabets, so that Mikołajczak (2000) proposes replacing these terms with *Latinitas*, a term with a long tradition in the cultures of Central and Western Europe, and *Cyrillianitas*, a term which, according to him, "also focuses on the role of the Slavic alphabets" ("koncentruje także uwagę na roli słowiańskich alfabetów").

---

model of Italian *gl, gn* (cf. Trunte 2001: 57). — The 'diacritic letters' ѣ ⟨″⟩ and ѣ ⟨′⟩ in Russian originally denoted distinct vowel sounds. Consequently, although synchronically *mь* ⟨t′⟩ might be analyzed as a digraph for [tʲ], it has never been invented as such. Historically, ъı ⟨y⟩ is clearly a digraph, but this might be due to its diphthongal pronunciation at least in some South Slavic dialects that were the basis for the Glagolitic and/or Cyrillic alphabet (cf. Mareš 1965: 70–75; furthermore, the sound has developed from a Proto-Indo-European **ū* to /i/ in most Slavic languages, and the constituent letters ⟨ъ⟩ and ⟨ı⟩ stand for **ŭ* and **i*, respectively). In contrast to the digraphs for unitary phonemes in (31), digraphic representation of diphthongs and affricates is a matter of course. — In the non-Slavic languages of the former Soviet Union, the situation is somewhat different: The Cyrillic orthographies for these languages were developed 🗋 146 quite recently (since 1937) and mostly from scratch for languages that had been written in a Latin alphabet for a decade or two. Some of these orthographies do use combinations with diacritic letters (e.g. Abkhaz ⟨дә⟩ for [dʷ]). In 2000, a Cyrillic practical transcription for Kazakh called *Kazanovica* was developed in order to be able to write Turkic languages with a Russian character set, replacing e.g. Kazakh ⟨ə⟩, ⟨қ⟩, ⟨h⟩ with ⟨аь⟩, ⟨къ⟩, ⟨хъ⟩ (cf. Sergeev 2001). For a complete overview of the additional letters in sixty Cyrillic-written languages cf. Píška (1996: 96–97); Giljarevskij & Grivnin (1970: 275) enumerate 62 languages with a "Russian" alphabet.

[27] Diacritic marks are a bit younger: Their systematic use for the representation of the sound system of a language dates back to Jan Hus' *De Orthographia Bohemica* (ca. 1412, i.e. more than a century before the first introduction of French accents and cedilla by Geoffroy Tory in 1529). Previously, diacritical marks had already been used to convey intonational features, e.g. in Ancient Greek, or the results of sound changes that occurred after the introduction of writing, e.g. umlauts in German (initially ⟨å, ǒ, ů⟩, now ⟨ä, ö, ü⟩).

[28] According to the rules of Giljarevskij & Starostin (1985: 157ff.) for the practical transcription of Italian proper names, the surname *Picchio* should be transcribed as *Пиккьо* ⟨Pikk′o⟩, but both Russian-speaking Slavists and the author himself spell it *Пиккио* ⟨Pikkio⟩ (cf. Pikkio 2002).

Let us give another interpretation of this difference. In cultures that use the Cyrillic alphabet, the writers have great authority: they alone decide how a foreign word should be pronounced (and if there are several variants, e.g. British and American, they choose one) and how readers should adapt the original pronunciation to their own language. In Latin-script cultures, however, readers are more autonomous: they are free to choose the appropriate 🗋139 pronunciation, for which writers must preserve all the properties of the word that are inherent in the source language.

5. Analytical conclusions

For all these reasons transliteration, which would have to use digraphs or diacritics, will probably never be used in Cyrillic writing systems consistently. Yet there seems to be a tendency towards transplantation of certain words especially from English (computer terminology, advertising, etc.). Whether this is just a temporary phenomenon is hard to say. It might be that as soon as everybody knows a specific term, the necessity of transplantation will disappear and it will be adapted to native orthography. Many such words already have parallel forms today, e.g. *internet* and *интернет*, *notebook* and *ноутбук*, *PC* and *ПК* in Russian, or *e-mail* and *mejl/мејл* in Serbian. However, more and more speakers of languages that use the Cyrillic alphabet nowadays have a good knowledge of English, which will probably provide for an incessant influx of English words. Due to the strictly etymological orthography of English, which contains a large number of homophones that are only distinguished by orthography, such words are sometimes hardly recognizable when transcribed according to their adapted pronunciation. Furthermore, modern computer systems make it easy to switch between different alphabets and the respective (virtual) keyboards. This tendency is likely to increase even in Serbian, where this would further reduce the optical distance between Serbian and Croatian.

So far we have looked at the differences between Cyrillic and Latin orthographies from the point of view of Latin, treating any difference as a peculiarity of Cyrillic. But maybe it is the other way round? Maybe the inclination to use digraphs and diacritic marks is a characteristic of the Latin script that distinguishes it from all other alphabets? There are, of course, many reasons for the spread of the Latin alphabet throughout the world. One of them are sociological factors, such as the fact that the alphabet of the Romans has been used for three world languages: Latin, French, and English. 🗋140 But perhaps another reason is the adaptability inherent in the Latin alphabet, which makes it easy to apply it to any language using digraphs and diacritical marks and which makes graphic code-switching possible.[29]

[29] Similar trends can also be observed in the languages of the former Soviet Union that

To finally resolve this issue, it is, of course, necessary to conduct a thorough study of other writing systems, which is impossible to do within the scope of this paper. However, at first glance, in relatively isolated writing systems (Greek, Hebrew, Armenian), graphic code-switching (transplantation, transliteration) and diacritical marks are either not used at all or are used to a much lesser extent than in the Latin alphabet. It is even possible to establish a certain proportionality between the use of graphic code-switching in a single writing system and the area of its distribution. For example, the Arabic script is used for many languages other than Arabic: Persian (Farsi), Urdu, Azerbaijani (e.g. in Iran), Kurdish (in Iraq), Uyghur (in China), Sindhi (in Pakistan), Pashto, etc. The additional letters needed to represent the phonology of these languages are formed using a variety of diacritical marks, e.g. Persian letters such as ⟨پ⟩ [p], ⟨چ⟩ [t͡ʃ] and ⟨ژ⟩ [ʒ], from Arabic ⟨ب⟩ [b], ⟨ج⟩ [d͡ʒ] and ⟨ز⟩ [z], respectively (cf. Giljarevskij & Grivnin 1970: 170, Ager 1998). Borrowings from Arabic into these languages are usually transplanted in their original spelling. Similarly, the Chinese script has spread throughout much of East Asia and is used, for example, for Korean (in South Korea) and Japanese. Here, very free code-switching between Chinese characters and Korean or Japanese syllabograms is permitted; consequently, foreign words, especially from English, are often transplanted, retaining the Latin alphabet.

From these observations, it can be concluded that graphic code-switching by means of transplantation or transliteration is an important factor in the efficiency of a writing system, not 🗋 141 only for the imperialistic spread of a script to other countries, but also for international exchange, especially in the era of globalization.

References

Ager, Simon. 1998. Persian (Fārsī/فارسی). In *Omniglot: A guide to writing systems.* https://www.omniglot.com/writing/persian.htm (last checked 22 Mar 2026).

Alpatov, Vladimir. 2001. Un projet peu connu de latinisation de l'alphabet russe. *Slavica Occitania* 12. 13–28.

Andrić, Dragoslav. 1976. *Dvosmerni rečnik srpskog žargona i žargonu srodnih reči i izraza.* Beograd: Beogradski izdavačko-grafički zavod.

Anić, Vladimir. 1991. *Rječnik hrvatskoga jezika.* Zagreb: Novi Liber.

Anić, Vladimir. 1994. *Rječnik hrvatskoga jezika.* 2nd edn. Zagreb: Novi Liber.

Arsen'ev, Konstantin K. & Petruševskij, Fëdor F. (eds.). 1901. *Ėnciklopedičeskïj slovar'.* Vol. 33[a]. S. Peterburg: Brokgauz & Efron.

adopted the Cyrillic alphabet in the 20th century. Their orthographies sometimes use diacritical marks and digraphs, and borrowings from Russian (or via Russian as an intermediary language) often retain Russian spelling (cf. Superanskaja 1978: 94ff.). This is because Cyrillic is used here outside the cultural context of Slavia Orthodoxa as the alphabet of a colonial language (cf. Stern 2003).

Ašnin, Fedor D. [† 2000] & Alpatov, Vladimir M. 2001. Putin za realjnyje celi: My mogli perejti na latinskoe pis′mo. *Nezavisimaja gazeta* 31 Mar 2001. 8.

Babkin, Aleksandr M. & Šendecov, Valentin V. 1966. *Slovar′ inojazyčnyx vyraženij i slov, upotrebljajuščixsja v russkom jazyke bez perevoda.* 2 vols. Moskva, Leningrad: Nauka.

Barić, Eugenija et al. 1999. *Hrvatski jezični savjetnik.* Zagreb: Institut za hrvatski jezik i jezikoslovlje.

Bartmińska, Izabela & Bartmiński, Jerzy. 1978. *Nazwiska obce w języku polskim: Problemy poprawnościowe; Słownik wymowy i odmiany.* Warszawa: PWN.

Belić, Aleksandar. 1950. *Pravopis srpskohrvatskog književnog jezika.* Beograd: Prosveta.

Bilodid, Ivan K. (ed.). 1958. *Kurs istoriï ukraïns′koï literaturnoï movy.* Vol. 1. Kyïv: Vydavnyc′tvo Akademiï Nauk Ukraïns′koï RSR.

[Bunčić, Daniel. 2016. *Biscriptality: A sociolinguistic typology*. Edited by Daniel Bunčić & Sandra L. Lippert & Achim Rabus. With contributions by Anastasia Antipova, Carmen Brandt, Ekaterina Kislova, Henning Klöter, Alexandra von Lieven, Sandra L. Lippert, Helma Pasch, Achim Rabus, Jürgen Spitzmüller, and Constanze Weth. Heidelberg: Winter.—D. B. 2026]

[Bunčić, Daniel. 2017. Factors influencing the success and failure of writing reforms. *Studi Slavistici* 14. 21–46.—D. B. 2026.]

Ćirilov, Jovan. 1994. *Srpsko-hrvatski rečnik varijanata – Hrvatsko-srpski rječnik inačica.* 2nd edn. Beograd: Bata.

Filipović, Rudolf. 1990. *Anglicizmi u hrvatskom ili srpskom jeziku: Porijeklo – razvoj – značenje.* Zagreb: Školska knjiga.

Giljarevskij, Rudžero S. & Grivnin, Vladimir S. 1970. *Languages identification guide.* (Transl. Lev Navrozov.) Moscow: Nauka.

Giljarevskij, Rudžero S. & Starostin, Boris A. 1985. *Inostrannye imena i nazvanija v russkom tekste: Spravočnik.* 3rd edn. Moskva: Vyssaja škola.

Hnatkevyč, Jurij V. 2000. *Unykajmo rusyzmiv v ukraïns′kij movi! Korotkyj slovnyk-antysuržyk dlja deputativ Verxovnoï Rady ta vsix, xto xoče, ščob joho ukraïns′ka mova ne bula sxožoju na movu Vjerky Serdjučky.* Kyïv: Prosvita.

Instruktionen für die alphabetischen Kataloge der preußischen Bibliotheken und für den preußischen Gesamtkatalog: Vom 10. Mai 1899. Berlin: A. Asher & Co.

Jušmanov, Nikolaj V. 1933. Problema oformlenija evropejskix sobstvennyx imen i naučno-texničeskix terminov v latinizovannyx jazykax SSSR. *Revoljucija i pis′mennost′* 11–12. 54–66.

Keipert, Helmut. 1999. *I contractum redivivum*: Zur Wiedereinführung des *i kratkoe* im russischen Buchdruck. *Zeitschrift für Slawistik* 44(3). 251–267.

Kempgen, Sebastian. 1993. Spezifika slawischer Schriften. In idem (ed.), *Slavistische Linguistik 1992: Referate des XVIII. Konstanzer Slavistischen Arbeitstreffens, Bamberg 14.–18. 9. 1992*, 111–143. München: Sagner.

Kordić, Snježana. 2001. Naziv jezika iz znanosti gledan. *Republika* 57(1/2). 236–243. [The discussion continues in the following issues of the journal.]

Kurkov, Andrej. 2000. *Dobryj angel smerti.* Xar′kov: Folio.

Mareš, František Václav. 1965. *Die Entstehung des slavischen phonologischen Systems und seine Entwicklung bis zum Ende der Periode der slavischen Spracheinheit.* (Transl. Peter Rehder & C. Schacht.) München: Sagner. [Original 1956: Vznik slovanského fonologického systému a jeho vývoj do konce období slovanské jazykové jednoty. *Slavia* 25. 443–495.]

Meisenburg 1989: Trudel Meisenburg. Romanische Schriftsysteme im Vergleich: Französisch und Spanisch. In Eisenberg, Peter & Günther, Hartmut (eds.), *Schriftsystem und Orthographie*, 251–265. Tübingen: Niemeyer.

Mikołajczak, Aleksander Wojciech. 2000. Między Latinitas i Cyrillianitas – rozdroże europejskiej tożsamości. In Kostyrko, Teresa & Zgółka, Tadeusz (eds.), *Kultura wobec kręgów tożsamości: Materiały konferencji przedkongresowej, Poznań, 19–21 października 2000 (Kongres Kultury Polskiej 2000)*. Poznań, Wrocław: Nakład Wielkopolskiego Towarzystwa Kulturalnego. https://web.archive.org/web/20030907110838/http://www.orw.poznan.pl/orw/archiwum/20001019/mikolajczak.html (last checked 22 Mar 2026).

Nikčević, Vojislav. 1997. *Pravopis crnogorskog jezika.* Cetinje: Crnogorski PEN Centar.

Nikčević, Vojislav. 2001. *Gramatika crnogorskog jezika*. Podgorica: Dukljanska Akademija Nauka i Umjetnosti.

Okuka, Miloš. 1998. *Eine Sprache – viele Erben: Sprachpolitik als Nationalisierungsinstrument in Ex-Jugoslawien.* Klagenfurt: Wieser.

[Pčelov, Evgenij V. & Čumakov, Viktor T. 2000. *Dva veka russkoj bukvy Ë: Istorija i slovar'.* Moskva: Narodnoe obrazovanie.—D. B. 2026]

Pešikan, Mitar & Jerković, Jovan & Pižurica, Mato. 1995. *Pravopis srpskoga jezika: Školsko izdanje.* Novi Sad: Matica Srpska.

Picchio, Riccardo. 1984. Guidelines for a comparative study of the language question among the Slavs. In Picchio, Riccardo & Goldblatt, Harvey (eds.), *Aspects of the Slavic language question*, vol. 1, 1–42. New Haven (CT): Yale Concilium on International and Area Studies.

Picchio, Riccardo. 1991. Slavia ortodossa e Slavia romana. In idem, *Letteratura della Slavia ortodossa (IX—XVIII sec.)*, 7–83. Bari: Dedalo.

Pikkio, Rikkardo. 2002. *Drevnerusskaja literatura.* Moskva: Jazyki slavjanskoj kul'tury.

Píška, Karel. 1996. Cyrillic alphabets. *TUGboat: The communications of the TeX Users Group* 17.2 (Proceedings of the 1996 Annual Meeting), 92–98. https://www.tug.org/tugboat/tb17-2/tb51pisk.pdf (last checked 22 Mar 2026).

Pravopis hrvatskosrpskoga književnog jezika. 1960a. Zagreb: Matica Hrvatska.

Pravopis srpskohrvatskoga književnog jezika. 1960b. Novi Sad: Matica Srpska.

Proxorov, Aleksej M. 1997. *Bol'šoj ėnciklopedičeskij slovar'.* 2nd edn. Moskva, Sankt-Peterburg: Bol'šaja Rossijskaja Ėnciklopedija, Norint.

Radomirović, Vladimir. 2001. Ekonomija na prvom mestu. *Reporter* 7 Nov 2001. 7.

Raecke, Jochen. 1996. Zum Problem einer „bosnischen Sprache". *Bulletin der Deutschen Slavistik* 2. 19–22.

Reformatskij, Aleksandr A. 1955. Transkripcija. In idem, *Vvedenie v jazykoznanie*, 295–297. Moskva: Učpedgiz.

Sabljak, Tomislav. 2001. *Rječnik hrvatskoga žargona.* Zagreb: V. B. Z.

Serdjučenko, Georgij P. 1967. *Russkaja transkripcija dlja jazykov zarubežnogo vostoka.* Moskva: Nauka.

Sergeev, Andrej V. 2001. Praktičeskaja transkripcija Kazanovica kak proekt reformy tjurkskix alfavitov na osnove kirillicy. (Conference presentation at *Meždunarodnaja naučno-praktičeskaja konferencija "XXI vek: Jazyk, vremja i prostranstvo"*, 19–20 Apr 2001.)

[Šipka, Milan. 2008. Muke oko pice. In idem, *Standardnojezička preispitivanja*, vol. 1, 61–66. Novi Sad: Prometej.—D. B. 2026.]

Skljarevskaja, Galina N. (ed.). 1998. *Tolkovyj slovar' russkogo jazyka konca XX v.: Jazykovye izmenenija.* Sankt-Peterburg: Folio-press.

Stern, Dieter. 2003. Kolonialsprache Russisch. In Gladrow, Wolfgang (ed.), *Die slawischen Sprachen im aktuellen Funktionieren und historischen Kontakt: Beiträge zum XIII. internationalen Slawistenkongress vom 15. bis 21. August 2003 in Ljubljana*, 69–97. Frankfurt: Peter Lang.

Superanskaja, Aleksandra Vasil'evna. 1978. *Teoretičeskie osnovy praktičeskoj transkripcii.* Moskva: Nauka.

Trunte, Nikolaos. 1997. Ex Armenia Lux: Konstantin-Kyrill und Mesrop-Mašt'oc'. *Palaeoslavica* 5. 31–57.

Trunte, Nikolaos. 2001. *Slavenskij jazyk: Ein praktisches Lehrbuch des Kirchenslavischen in 30 Lektionen, zugleich eine Einführung in die slavische Philologie.* Vol. 2. 2nd edn. München: Sagner. [1st edn. 1998.]

Ušakov, Dmitrij N. 1935–40. *Tolkovyj slovar' russkogo jazyka.* 4 vols. Moskva: Sovetskaja Ėnciklopedija.

Vilku-Pustovaïa, Irina. 2001. Le roumain/moldave: Histoire d'une guerre des alphabets. *Slavia Occitania* 12. 391–416.

Vinogradov, Viktor A. 1977. Transkripcija. In Proxorov, Aleksandr M. (ed.), *Bol'šaja sovetskaja ėnciklopedija*, 3rd edn., vol. 26, 148–149. Moskva: Sovetskaja Ėnciklopedija.

Vuk Stefanović Karadžić (transl.). 1847. *Novi Zavjet Gospoda našega Isusa Hrista.* Beč: Štamparija jermenskoga manastira.

[Webster.] 1993: *Webster's new encyclopedic dictionary.* Cologne: Könemann.

The (re)nationalization of the Serbo-Croatian standards

The question how many languages are spoken on the language territory of the Neo-Štokavian standards has been a subject of heated debate for a long time. Although many linguists have become reluctant to engage in this topic, the problem is not solved. The two sides still accuse each other of ideologism. But is everybody who deals with this subject actually an ideologist? Or might it be possible that Bosnian, Croatian, Montenegrin and Serbian are really no "normal" standard languages, just as Serbo-Croatian is not a "normal" pluricentric language? To be able to meet a decision on this, the characteristics of the speech forms in question have to be compared with those of other languages. Therefore, neither the model of the "literary variants" nor the one of the "autonomous languages" can be used, since these models are valid only for those linguistic situations for which they were developed. In this context Ulrich Ammon's model is often quoted, the internationally most widely accepted model permitting a decision on what is an individual language and what is a national variety of a language; however, the details of this model in its entirety have not been cited so far by any of the sides. Therefore, in this paper, I present some less commonly known parallels and differences between our speech forms on the one hand and such languages as German, English or Portuguese on the other. This shows that the Serbo-Croatian speech forms really have much in common with the standard varieties of pluricentric languages but that they also exhibit specificities which are not envisaged in Ammon's model. Above all it can be shown that something really happened around 1991, which on the basis of Ammon's model can be termed "nationalization" (or "renationalization"). This nationalization explains a great deal of the theoretical problems and demonstrates that the linguistic situation in former Yugoslavia is not as unique a case as is frequently assumed.

Translated from "Die (Re-)Nationalisierung der serbokroatischen Standards" in Kempgen, Sebastian & Gutschmidt, Karl & Jekutsch, Ulrike & Udolph, Ludger (eds.), *Deutsche Beiträge zum 14. Internationalen Slavistenkongress, Ohrid 2008*, 89–102. München: Sagner 2008. DOI: 10.18716/bun/bkms.

DOI: 10.18716/omp.64.4

1. Introduction

To date, there is still no agreement in the sometimes heated debate that has been going on for many years about the now proverbial question "How many languages?" (which, for example, Thomas 1994, Bugarski 2000, Alexander 2002/03, and many others already mention in the titles of their articles). This can hardly be attributed to a lack of arguments or proposed solutions. Therefore, this article will not attempt to answer this question directly, but will instead analyze, at a metalevel, why the linguistic situation in Bosnia, Croatia, Montenegro, and Serbia is so difficult to grasp linguistically.

It is undisputed that serious linguistic changes took place in (former) Yugoslavia at the beginning of the 1990s, but there is uncertainty about what exactly happened. While some assume that Serbo-Croatian split into several languages (Radovanović 2003, Rehder 1998), others argue that the changes can only be seen in metalinguistic discourse, which either means that there is still only one "linguistic" language (Kordić 2004a, Pranjković 2006), albeit several "political" ones (Thomas 1994, Bugarski 2000, Alexander 2002/03), or that it has now finally been officially recognized that a Serbo-Croatian language never existed (Katičić 1997, Auburger 1999).

A major problem in assessing the situation is that it has so far been described almost exclusively using models developed specifically for this situation, making it difficult to compare these descriptions with those of other languages. For this reason, the situation will be presented here strictly according to a model that has been tested on dozens of languages, namely the model of pluricentric languages developed by Ulrich Ammon.[1]

2. Presentation of the model of national varieties

2.1 Practical application: using Austrian German as an example

When Austria joined the European Union in 1995, it insisted that Protocol No. 10 be appended to the accession treaty, which contains a list of the following words in the annex:

> 90 *Beiried* 'striploin',[2] *Eierschwammerl* 'chanterelles', *Erdäpfel* 'potatoes', *Faschiertes* 'minced meat', *Fisolen* 'green beans', *Grammeln* 'cracklings', *Hüferl* 'beef hip',[2] *Karfiol* 'cauliflower', *Kohlsprossen* 'Brussels sprouts', *Kren* 'horseradish', *Lungenbraten* 'tenderloin',[2] *Marillen* 'apricots', *Melanzani* 'eggplant', *Nuss* 'sirloin tip',[2] *Obers* 'cream', *Paradeiser* 'tomato', *Powidl* 'plum butter', *Ribisel* 'currant', *Rostbraten* 'forerib',[2] *Schlögel* 'hind leg',[2] *Topfen* 'curd cheese', *Vogerlsalat* 'lamb's lettuce', *Weichseln* 'sour cherries' (European Union 1995: 330)

These 23 Austriacisms therefore have "the same status and may be used with the same legal effect as the corresponding expressions used in Germany" ("den glei-

[1] This article addresses the fact that the written languages of Bosniaks, Croats, Montenegrins, and Serbs are clearly distinguishable from one another by using the terms *Bosnian*, *Croatian*, *Montenegrin*, and *Serbian*, while the fact that the members of the four peoples can communicate better with each other than, for example, with Slovenes or Macedonians, is referred to by the traditional term *Serbo-Croatian*. This is done in a completely neutral manner, simply because none of the proposed alternatives (*Central South Slavic*, *Standard Neo-Štokavian*, *BCS*, *SerBoCroatian*, *naški*, etc.) is entirely convincing.

[2] Apparently, Austrian, German, American, and British butchers all cut up cows, pigs, lambs, and chickens differently, so that the pieces they sell (and name) only partially overlap but never coincide. This is a terminological nightmare for a translator (not only for a vegetarian one).—D. B. 2026.

chen Status und dürfen mit der gleichen Rechtswirkung verwendet werden wie die in Deutschland verwendeten entsprechenden Ausdrücke", i.e. *Roastbeef* 'striploin', *Pfifferlinge* 'chanterelles', *Kartoffeln* 'potatoes', *Hackfleisch* 'minced meat', *grüne Bohnen* 'green beans', etc.; ibid.; for the background and impact of this protocol, see Markhardt 2006).

In 2003, however, a product labeled as *Marillenmarmelade* 'apricot jam' was nevertheless rejected by the food inspection authority, as *Marmelade* must, according to an EU directive, contain citrus fruits, whereas jam must be labeled as *Konfitüre*. While this regulation, which is based on the English model and contradicts common German usage, had been tacitly accepted in Germany, it led to the "Apricot Jam War" ("Marillenmarmeladekrieg") in Austria (Markhardt 2006: 22)—especially since it was mistakenly considered to be linguistic paternalism from Germany (ibid. 23)—and ultimately resulted in the adoption of an exemption by the European Council.

As early as 1949, in order to avoid any hint of linguistic affinity with Germany, the name of the school subject *Deutsch* 'German' was abolished in Austrian schools and replaced by the term *Unterrichtssprache* 'language of instruction'; however, the glottonym was reintroduced in 1952 (cf. Ammon 1995: 126f.). The *Österreichisches Wörterbuch* 'Austrian Dictionary' (Back & Fussy 2006), published in Vienna since 1951, does not contain the word *Deutsch* even in a subtitle. And the manifesto "Österreichisch als eigene Sprache!" ("Austrian as a language in its own right!"), signed in 2004, called on the government to amend the article on language of the federal constitution to read "Die Staatssprache ist Österreichisch in einem europäischen Kontext" ("The official language is Austrian in a European context"), "Die Staatssprachen sind Deutsch und Österreichisch" ("The official languages are German and Austrian"), or something similar, and to advocate for the recognition of Austrian as a separate EU language (cf. Manifest 2004).

These brief snapshots should demonstrate how closely national identity can be linked to a speech form that, despite its high symbolic and identificatory value, is classified by all serious linguists as 'only' a *national variety* and not as a separate language.

2.2 The theory: Ammon (1995)

On the question of how to objectively decide whether two linguistic expressions belong to different languages or to the same language, Kloss (1952) already provided some answers, which Ammon (1995) further refined and, particularly for pluricentric languages, developed into a comprehensive model with a clearly defined algorithm. This will be briefly presented here.

According to Ammon, two criteria are decisive for answering this question: first, roofing, which correlates with standardization (or Kloss's *ausbau*), because "standard varieties cannot be roofed, and non-standard varieties cannot roof other

varieties" ("Standardvarietäten können nicht überdacht werden, und Nonstandardvarietäten können nicht überdachen", Ammon 1995: 2); and second, the similarity of the varieties being compared, for which Ammon (1995: 6) suggests measuring "the number of matching words in texts that are identical in meaning and translated as literally as possible" ("in sinngleichen, möglichst wörtlich übersetzten Texten […] die Zahl der übereinstimmenden Wörter") as an approximation. The interaction of the two criteria is illustrated in the following table: 🗋91 Starting from a given standard variety, a variety compared with it belongs to the same language or not, depending on its standardization and its similarity to the initial variety, as shown in the table.

<table>
<tr><th></th><th>high similarity</th><th>medium similarity</th><th>low similarity</th></tr>
<tr><td>identical words</td><td rowspan="2">> 50 %</td><td>< 50 %</td><td rowspan="2">< 50 %</td></tr>
<tr><td>recognizable cognates</td><td>> 50 %</td></tr>
<tr><td>standard variety (roofing)</td><td>same language
(pluricentric)</td><td colspan="2">other language
(ausbau language) | (abstand language)[3]</td></tr>
<tr><td>nonstandard variety (roofed or roofless)</td><td colspan="2">same language
(dialect, sociolect, etc.)</td><td>other language
(foreign dialect etc.)</td></tr>
<tr><td>Ammon's example:
German
Standard German</td><td>Austrian Standard German, Swiss Standard German</td><td>Standard Luxembourgish, Standard Dutch, Standard Afrikaans, Standard Yiddish
Swabian dialect</td><td>Standard French</td></tr>
</table>

If, according to these criteria, several standard varieties belong to one language, that language is *pluricentric* (e.g. Arabic, Armenian, Chinese [Mandarin!], German, English, French, Hindi-Urdu, and many other languages; cf. Clyne 1992). In this case, according to Ammon (1995: 96), a further distinction can be made as to whether a center in which one of the standard varieties applies has its own internal codification (as e.g. for German in Germany, Austria, and Switzerland) and is thus a *full center*—or whether it has specific standard variants but has not codified them independently in normative dictionaries, making it a *semi-center* (such as, for German, Liechtenstein, Luxembourg, South Tyrol, and East Belgium).

[3] Ammon (1995: 1–11) does not use these two internationally accepted terms coined by Kloss (1952: 16–20) here, but this is precisely what is meant; Kloss (1952: 19f.) already speaks of a "minimum distance" ("Mindestabstand") that must be given in order to speak of different (ausbau) languages, and of a "maximum distance" ("Höchstabstand") that requires the existence of different (abstand) languages even without ausbau.

Furthermore, Ammon (1995: 84f., 92, 95) distinguishes between "national", "state", and "regional" standard varieties (the latter two of which can be further specified as "subnational" or "substate"). Abstracting somewhat from Ammon's explicit wording, one could say that *state* varieties are distinguished by an area of validity that is clearly defined by state borders, while *regional* standard varieties usually have an area of validity with fuzzy borders, with a center and a periphery; one example would be the south(east)ern German regional standard with its center in Munich, whose individual variants (*Semmel* 'bread roll', *Brotzeit* 'lunch snack', apical [r], *ich bin gesessen* vs. *ich habe gesessen* 'I have sat', etc., cf. König 2005: 232–245) are spread across areas of varying sizes. *National* varieties, on the other hand, can also be state varieties in the case of a nation state, but they are primarily linked to the members of the nation, regardless of how many countries the nation lives in. (At the time of the division of Germany, for example, the German nation lived in two states, the Federal Republic of Germany and the German Democratic Republic, while the Kurdish nation is currently spread across several 'foreign' countries.) From this, it can be concluded that although a national variety is generally (cf. Ammon 1995: 77) dependent on state institutions for the codification and enforcement 92 of norms, its scope encompasses "all individuals of the nation or the language community within the nation" ("alle Individuen der Nation bzw. der Sprachgemeinschaft in der Nation", ibid.), i.e. it is ultimately defined by individuals and not primarily by territory.

3. Attempt to apply Ammon's model

3.1 Similarity tests

In the following, we will now apply this model to Bosnian, Croatian, Montenegrin, and Serbian, while naturally retaining Ammon's terminology, even though it differs in part from the terms commonly used in former Yugoslavia. In particular, it should be noted that a) a linguistic system[4] is referred to here exclusively as a *variety*, while *variant* always refers to a linguistic unit, e.g. a word, a phoneme, a syntactic structure, etc.; that b) *standard variety* is synonymous with *književni jezik* (i.e. it is also used when a language has only one standard variety);[5] and that

[4] Since Brozović (1970: 38) and Silić (1996) claim that a single variety is only a question of Saussure's *parole* (or Coseriu's *habla*), while the *langue* (or *systema*) encompasses all standard and non-standard varieties of the (Serbo-)Croatian language as a whole, it should be clearly stated here that every dialect, sociolect, idiolect, and of course every standard variety has its own *langue* and *parole* (and also Coseriu's *norma*), while a language as a whole is merely a 'system of systems' (cf. Ammon 1995: 1).

[5] In order to avoid the misunderstanding that the term *standard varieties* must *a priori* refer to several varieties of *one and the same* language, I have chosen the term *standards* for the title of this article.

c) Ammon simply calls *language* what Brozović (1970: 66) introduced the term *jezik-dijasistem* 'language-diasystem' for.

Therefore, the similarity test described above must first be applied to the standard varieties (which serve as official languages) in the four (five if Kosovo is included) Serbo-Croatian-speaking countries. All differences are underlined and then counted:[6]

> *Bosnian:* Već u slavenskoj prapostojbini dolazi do uočljivih razlika između pojedinih jezičkih grupa. Tako se formiraju buduće etničko-jezičke zajednice: južnoslavenska, zapadnoslavenska i istočnoslavenska. Prije doseljenja Južnih Slavena na Balkan postojala je i unutrašnja raslojenost ove grane na njen istočni i zapadni dio, te se u nauci govori o istočnojužnoslavenskome i zapadnojužnoslavenskome prajeziku. Iz zapadnojužnoslavenskoga prajezika na zapadnijem dijelu Balkanskoga poluotoka razvijaju se tri dijalekatske cjeline: štokavska, kajkavska i čakavska. (Halilović 1998: 13)

> D 93 *Croatian of Croatia:* Već u slavenskoj prapostojbini dolazi do uočljivih razlika između pojedinih jezičnih grupa. Tako se formiraju buduće etničko-jezične zajednice: južnoslavenska, zapadnoslavenska i istočnoslavenska. Prije doseljenja Južnih Slavena na Balkan postojala je i unutrašnja raslojenost ove grane na njezin istočni i zapadni dio, te se u znanosti govori o istočnojužnoslavenskome i zapadnojužnoslavenskome prajeziku. Iz zapadnojužnoslavenskoga prajezika na zapadnijem dijelu Balkanskoga poluotoka razvijaju se tri dijalekatske cjeline: štokavska, kajkavska i čakavska. (translated by Amir Kapetanović)

> *Serbian of Serbia:* Već u slovenskoj prapostojbini dolazi do uočljivih razlika između pojedinih jezičkih grupa. Tako se formiraju buduće etničko-jezičke zajednice: južnoslovenska, zapadnoslovenska i istočnoslovenska. Pre doseljenja Južnih Slovena na Balkan postojala je i unutrašnja raslojenost ove grane na njen istočni i zapadni deo, te se u nauci govori o istočnojužnoslovenskom_ i zapadnojužnoslovenskom_ prajeziku. Iz zapadnojužnoslovenskog_ prajezika na zapadnijem delu Balkanskoga poluostrva razvijaju se tri dijalekatske celine: štokavska, kajkavska i čakavska. (translated by Biljana Golubović)

Of these 68 words from the Bosnian original, which are not intended as 'evidence' but merely as an illustration of the method, 94% are identical in the Croatian translation and 81% in the Serbian translation;[7] the Croatian and Serbian transla-

[6] One problem not specifically addressed by Ammon (1995) is that of writing. He himself also takes into account purely orthographic differences such as Swiss ⟨ss⟩ versus German/Austrian ⟨ß⟩. However, the Serbian example makes it clear that purely graphic differences can only be included to a limited extent, because otherwise they would lead to the obviously false statement that the inhabitants of Belgrade *speak* a common standard variety, but *write* two different languages that are completely foreign to each other, since the Cyrillic and Latin versions of a text differ from each other by almost 100%. Therefore, I have only counted those differences that are also audible when the text is read aloud.

[7] Bosnian texts that are less similar to Croatian are closer to Serbian. Lehfeldt (1999: 88) notes the rarity of specific Bosniakisms even in the newspaper *Ljiljan*.

tions are 75% identical with each other.[8] Bosnian Serbian or Montenegrin—Ijekavian—translations, which cannot be analyzed in this paper, would be even closer to Croatian and Bosnian than the Ekavian Serbian of Serbia, while Bosnian Croatian texts generally differ little from the Croatian of Croatia.[9]

This clearly fulfills Ammon's condition for treating the standard varieties mentioned as a pluricentric language, namely an identity of over 50% of the words. (Tests with slightly larger and fictional text excerpts always yield slightly different numerical values, but overall the same result.) Burgenland Croatian, however, is a special case, as its standardization took place outside Yugoslavia and it has integrated Čakavian and Ikavian elements into the standard. Therefore, this will be dealt with separately here:

> *Burgenland Croatian:* Širom svita haraju i dandanas nemiri i krize, na peldu u Iraku, Afganistanu, Pakistanu, Bliskom istoku itd. itd., kot da se čovičanstvo i u tisućljetnoj prošlosti u tom pogledu ništ_ nije naučilo. A i uzroki nemirov uvijek su isti: siromašni 🗋94 protiv bogatih, borba za vlast, za moć, vjerska nepodnošljivost ali i uskraćivanje ljudskih prav_. Vrlo rijetko kroz ljeto čuju se riči državnikov, ki pozivaju narode i pojedince na pravično podiljene zemaljskih dobar_, na ispunjenje zakonov, na toleranciju – ali božićno vrime zaistinu je pravo vrime za takovo opomenjivanje i razmišljavanje. (Sučić 2007: 1)

> *Croatian of Croatia:* Širom svijeta haraju i dandanas nemiri i krize, na primjer[10] u Iraku, Afganistanu, Pakistanu, na Bliskom istoku itd. itd., kao da čovječanstvo i u tisućljetnoj povijesti u tom pogledu ništa nije naučilo. A i uzroci nemira uvijek su isti: siromašni protiv bogatih, borba za vlast, za moć, vjerska nesnošljivost, ali i uskraćivanje ljudskih prava. Vrlo rijetko tijekom godine čuju se riječi državnika, koji pozivaju narode i pojedince na pravično _dijeljenje zemaljskih dobara, na ispunjenje zakona, na toleranciju – ali božićno vrijeme doista je pravo vrijeme za takovu opomenu i razmišljanje. (translated by Amir Kapetanović)

Of the 91 word forms in this text excerpt, 74% are retained in the Croatian translation. This means that Burgenland Croatian differs from Croatian at least as much

[8] Compare the (admittedly constructed) texts of 160 words given by Ammon (1995: 9–11), in which the German German and Swiss German versions are 85% identical (not counting orthographic differences). In the versions in Yiddish, Luxembourgish, and Dutch, the languages most closely related to German, respectively only 16%, 6%, and 5% of the words are identical with the German versions.

[9] Okuka (2000: 72) considers Bosnian Croatian to be "identical to Croatian in Croatia, in some cases even more 'Croatianized' than in Croatia itself" ("идентичан хрватском језику у Хрватској, у неким случајевима чак и више 'похрваћен' него у самој Хрватској"), with the last clause, however, suggesting a certain independence of the Bosnian Croatian variety.

[10] Although Amir Kapetanović translated this correctly, the German paper published in 2008 contained a mistake here, for which I alone am responsible. (D. B. 2026.)

as Ekavian Serbian.[11] In Ammon's model, however, these subtleties are not relevant, as all texts coincide by well over 50%, meaning that their similarity is *high* and they must therefore be classified as varieties of a pluricentric language.[12]

It is often claimed that it was the 'Croatian Vukovites' ('hrvatski vukovci') who forcibly 'united' the Croatian and Serbian languages, which are actually foreign to each other, into one language. Unfortunately, no native speakers from the time of the National Renaissance period are alive who could be asked 95 to produce 'equivalent texts translated as literally as possible' in order to verify this. But even a glance at actual translations is revealing. For this purpose, we will use the beginning of the translations of Herder's (1791) chapter on the Slavs, which is available in an "Illyrian" version from Ljudevit Gaj's *Danica* (1835: 231) and two Serbian versions, of which the one by Stamatović (1832: 27) is quoted here, which, as Keipert (2000) has shown, served as a basis for the *Danica* translation alongside the German original. Where the older Serbian version by Davidović (1816: 248) is closer to the *Danica* text, it is added in parentheses:[13]

Serbian: Славенски народи обузымаю на земльи, већій просторъ, нежели (негол') _у повѣстницы, између прочи причина и збогъ тога, што су подалеко отъ Римляна живили. Мы и[x] познаємо найпре на Дону, а доцніе (потомъ) на Дунаву,

[11] The fact that many Croatian Croats find Burgenland Croatian much more difficult to understand than Serbian has nothing to do with linguistic similarity but rather with their limited contact with Burgenland, as a result of which Burgenlandisms such as *pelda* 'example' (= *primjer*) are largely unknown in Croatia, while Serbianisms such as *ostrvo* 'island' (= *otok*) are well known. However, for a variety of reasons, comprehensibility is not a criterion for determining linguistic status (cf. Ammon 1995: 5f.).

[12] Another variety that might be considered missing here is Molise Slavic ("Molise Croatian"). In the Standard Croatian and Molise Slavic parallel translations from the *Decameron* that are printed in Duličenko (2003: 155f.) and, although not strictly prepared according to Ammon's criteria, can serve as a guide, a total of 10% of the words match. However, the clear majority of Molise Slavic words can be recognized as cognates to their Croatian parallels, so that the similarity of the texts can be classified as *medium*. (These figures would probably not change significantly even with a translation "as literal as possible", as there are too many phonological, morphological, syntactic, and lexico-semantic differences, cf. Breu 1998.) Thus, what is decisive for the classification of this variety is standardization. Breu (1998: 275) notes: "Das Moliseslav. ist keine Schriftsprache. Gelegentlich wird es von Intellektuellen dennoch schriftlich gebraucht […]" ("Molise Slavic is not a standard language. Occasionally, however, intellectuals use it it in writing […]"). This would make it a dialect (a roofless external dialect) of Serbo-Croatian. For Duličenko (2003: 32), the enthusiast of microlanguages, however, a grammar book printed in Canada in 1968 is reason enough to describe Molise Slavic as a "functionally weak literary microlanguage" ("функционально слабый литературный микроязык"), i.e. as an ausbau language. However, Molise Slavic is by no means a standard variety of Serbo-Croatian, as this would require it to be *highly* similar to the other standard varieties.

[13] My sincere thanks go to Helmut Keipert, who was kind enough to give me copies of the poorly accessible Serbian translations.

ондѣ мѣђь Гоѳима (Готы), а овдѣ мѣђь Хуннима (Хунны) и Булгарима (Бугары), съ коима су често Римско Царство яко узнемиривали, по већой части само као сопутни, помагаюћи или служећи народи […].

'Illyrian': Slavenski puki obstiraju veći prostor na zemlîi, nego vu dogodovšćini, a to med ostalim takajše zato, jer su od Rimlîanov daleko živeli. Mi ih poznamo najpervo na Donu, zatim na Dunaju, onde med *Goti*, a ovde med *Huni* i *Bulgari*, s koimi su rimsko carstvo jako uznemiravali, ali vekšinum kakti skupputujući, pomagajući ili služeći puki. [Kursiv im Original]

Of the 55 words in this section, 55% are identical, so that the similarity of the texts according to Ammon can still be described as *high*.[14] Thus, according to Ammon's model and contrary to Brozović (1970: 124), the standard varieties of Serbian and Croatian can already be classified as a common pluricentric language in 1835 due to their high similarity, even though their standardization at that time was still almost completely independent of each other.

Between 1835 and today, there have been phases of convergence and divergence between the standard varieties discussed here, but it can be assumed that the similarity never became less than it was in 1835 and that the differences were never completely eliminated. Thus, at least since 1835, we have been dealing with standard varieties of high similarity—the basic prerequisite for the application of Ammon's model of national varieties is therefore met.

3.2 Development of the variety structure over time

Unfortunately, the development of the variety structure prior to 1945 must be excluded here, but it is clear that political terms such as *Illyrian*, *Yugoslav* (cf. 🗋96 Okuka 1998: 19f.), *Serbo-Croato-Slovene* (*srpsko-hrvatsko-slovenački*, cf. Okuka 2000: 66–67), the *Bosnian language* of the Austro-Hungarian administration (cf. Okuka 1998: 54–59) or even Vuk's (1849) *Serbian* for everything Štokavian did not refer to uniform standard varieties.

Yugoslavia's language policy only received a linguistic foundation in socialist Yugoslavia. The 1954 Novi Sad Agreement established that there were two *varijante* (i.e. standard varieties), a western one centered in Zagreb and an eastern one centered in Belgrade. While the agreement only addresses the differences in the representation of **ě* and the two alphabets (cf. its text in Greenberg 2004: 172–174),

[14] By comparing other passages, the texts could be brought closer to the requirement of a translation "as literal as possible", which would certainly increase their similarity even more. For example, the 'Illyrian' translation in other places also uses *narod* instead of *puk* (Danica 1835: 231, 232), *jer* is also found in the Serbian text (Stamatović 1832: 29), etc. Incidentally, it should be noted that some of the differences can be attributed to Kajkavian elements in the 'Illyrian' text, which can hardly be described as Croatian-Serbian differences (e.g. 'Croatian' *živeli* vs. 'Serbian' *živili*), since these isoglosses run differently from the boundary between Serbian and Croatian.

other linguistic literature also and above all deals with lexical differences (cf. Ivić in Brozović & Ivić 1988: 47 [1990: 65]). The term *književni jezik* later came into use as a synonym for *varijanta*, especially in Croatia (cf. Brozović in Brozović & Ivić 1988: 103 [1990: 88], Brozović 1992: 359). A "polarization" was assumed, according to which the two varieties "overlap" (Brozović 1970: 36) or "neutralize" each other (which is criticized by Janković 1982) in the areas between Belgrade and Zagreb (especially in Bosnia and Montenegro).[15] This description is typical of *regional standard varieties* that have a center and a periphery. It can be assumed that this bipolar structure was inherited from the period before 1945 and that the lack of independence of the varieties of Bosnia and Montenegro can be explained by the centralism of the First Yugoslavia.

Since the 1960s, however, voices were raised calling for the recognition of the language used in the Bosnian and Montenegrin republics as the third and fourth *varijanta* (Hraste 1965, Isaković [1970] 1998), but most Serbian and Croatian linguists adhered to the bipolarity of Serbo-Croatian. Instead, they spoke of *bosansko-hercegovački* or *crnogorski književnojezični izraz* ('literary-language expression', Brozović in Brozović & Ivić 1988: 102f. [1990: 88], also *standardnojezička upotreba* 'standard-language usage', Janković 1982: 849), sometimes also *subvarijanta*, etc. (cf. Isaković 1998 [1970]: 241). This difference in terminology meant a certain degree of unequal treatment, which was reflected, among other things, in the fact that no separate internal codes (dictionaries, grammars) were developed for Bosnia and Montenegro. In this respect, *varijanta* can be translated into Ammon's terminology as 'standard variety of a full center' and *književnojezični izraz* or *subvarijanta* as 'standard variety of a semi-center'.

With the recognition of specific standard varieties for each of the four Yugoslav republics, they were each given areas of validity with clearly defined borders, thus changing from *regional* to *(sub)state* varieties "on a 'republican level'" ("na 'republičkom nivou'", Isaković 1998 [1970]: 245). (For the varieties of such federal 'states within a state' with great cultural autonomy, a new term would have to be added to 97 Ammon's model.) In this case, too, it can be assumed that this change had already been taking place gradually and did not only occur at the moment when science and politics recognized it.

The development of the variety structure between 1945 and 1995 proceeded so rapidly that the next stage of development was initiated almost simultaneously with the one just described. The much-praised "Declaration on the name and

[15] This overlap and the emphasis on the equality of both *varijante* sometimes led to surprising results. For example, in 1972, a Sarajevo edition of the *Matica* spelling rules of 1960 was published under the title *Pravopisni priručnik srpskohrvatskoga/hrvatskosrpskoga jezika* 'Spelling guide of the Serbo-Croatian/Croato-Serbian language', in which the chapters (which had been specially reduced from 17 to the even number 16 for this purpose) were printed alternately in Cyrillic and Latin script (cf. Brborić 2000 [1972]: 45).

status of the Croatian literary language" ("Deklaracija o nazivu i položaju hrvatskog književnog jezika") of 13 March 1967 is often misunderstood in this regard. It did not claim that Croatian was an independent language (as the incorrectly abbreviated title of the 1991 new edition, *Deklaracija o hrvatskome jeziku*, suggests); rather, the term *književni jezik* is used throughout the text, which, as mentioned, can be translated as 'standard variety'. It is merely a question of the naming and use of this variety: instead of *zapadna varijanta* 'western variant' or even *hrvatskosrpska varijanta hrvatskosrpskoga/srpskohrvatskoga jezika* 'Croato-Serbian variant of the Croato-Serbian/Serbo-Croatian language', the signatories now wanted to call their standard variety *hrvatski književni jezik* 'Croatian literary language' (scil. *hrvatskosrpskoga jezika* 'of the Croato-Serbian language'), and they demanded that texts intended for Croatia be written in this variety.[16]

While these demands should actually be self-evident, the significance of the *Deklaracija* only becomes apparent in the reaction of Serbian writers in the *Predlog za razmišljanje* of March 19. They now demanded that "also all Croats living in the territory of the SR Serbia and all Serbs living in the territory of the SR Croatia" ("и свим Хрватима који живе на територији СР Србије и свим Србима који живе на територији СР Хрватске", Grupa članova UKS 2000 [1967]: 17) be granted the right to media and school instruction "in their language and script" ("на свом језику и писму", ibid.), which had not been the case previously (cf. Katičić 1984: 295).[17] This meant that the standard varieties were individualized for the first time, i.e. they were no longer tied to the borders of the republic in which a speaker lived, but to their nationality. To the extent that this demand was implemented, the (sub)state varieties of Bosnia, Croatia, Montenegro, and Serbia became the *national* varieties of the Bosniaks, Croats, Montenegrins, and Serbs.

This process (which has already been outlined by Šipka 2003: 277) could be described as the *nationalization* of the Serbo-Croatian standard varieties, and its significance can hardly be overestimated. (Compare that during the Cold War, East German and West German each had their own internal codifications, which means that they were full centers, but only *state* varieties, whereas Austrian and Swiss German are *national* varieties.) Here, varieties previously determined by republic borders and state institutions were placed in the hands of individuals, and people

[16] The final trigger for this demand was the publication of the conclusions of the 5th Congress of the League of Composers (sic—not Communists!) of Yugoslavia in the Serbian variety, Slovene, and Macedonian only (Deklaracija 1991 [1967]: 9).

[17] Even the renaming of the official language in the Croatian constitution of 1974 apparently did not change this, as the language article clearly stated that Croatian Serbs used the same variety as Croats: "U SR Hrvatskoj u javnoj je upotrebi hrvatski književni jezik – standardni oblik narodnog jezika Hrvata i Srba u Hrvatskoj, koji se naziva hrvatski ili srpski" ("In the SR Croatia, the Croatian literary language is in public use—the standard form of the national language of Croats and Serbs in Croatia, which is called Croatian or Serbian"; cf. also Okuka 2000: 68f.).

felt this qualitative difference very clearly and put it into words 🗋98 by saying that their language had become 'independent' and had 'split off' from Serbo-Croatian. However, since this process—despite all purist neologisms *(tuđmanice)* and all "language censorship" ("Sprachzensur", Kordić 2004b)—was not characterized by any significant change in the substance of the language, these descriptions are inaccurate from a linguistic point of view. What began in 1967 and ended in Croatia with the 1990 constitution, which made the Croatian standard variety the official language but allowed for other "languages" and the Cyrillic script at the local level, was the elevation of this (sub)state variety to a *national* variety. In Bosnia, the authors of the 1993 constitution still assumed a uniform state variety, which, according to a law of 29 August 1993, could be called *Bosnian*, *Croatian*, or *Serbian* (Gröschel 2003: 162), and it was only the Dayton Constitution of 1995 (written in English in the original), which contains no provision on the state language but was practically translated into three different varieties, that completed the individualization and nationalization of the standard varieties. In Montenegro, it is still unclear whether nationalization has really been completed. Regardless of the renaming of the national language to *Montenegrin* in the Constitution of 19 October 2007 (which, like the *Deklaracija* of 1967, is 'only' a question of naming), I am not yet sure whether the 32% Serbs in Montenegro use a different standard variety from the 43% Montenegrins. Nikčević's orthography (1997), which actually prescribes a significant difference from the Serbian standard varieties of Serbia and Bosnia and would therefore probably not be accepted by Montenegrin Serbs, has apparently not been able to gain acceptance so far, and, conversely, Montenegrin Serbs seem to continue to use the Ijekavian standard of Montenegro, which they share with their ethnic Montenegrin neighbors. In this respect, Montenegrin should perhaps still be classified as the 'state variety of the Montenegrin nation state' for the time being and not yet as a *national* variety.

Since the settlement areas of the four nations have different borders than the four now independent republics, nationalization has given rise to complex variety structures: In addition to the full center of Croatia, the Croatian national variety includes Burgenland Croatian, Bosnian Croatian, and possibly Vojvodina Croatian as further state varieties; the Serbian national variety comprises the state varieties of Serbia ("srbijanski", Radovanović 2003: 238) and Bosnia, as well as possibly Montenegro (see above), Croatia (Krajina and Slavonia), and soon perhaps also Kosovo; and even the Bosnian national variety has state varieties outside Bosnia and Herzegovina in the Serbian and Montenegrin Sandžak (as well as in Croatia, cf. Halilović 1998: 19). In some of the cases mentioned, it would of course be necessary to examine more closely whether these are really independent varieties, but for the time being we can assume that they are, since according to Ammon (1995: 45; cf. also 64f.), "even the smallest linguistic differences" ("schon kleinste linguistische Unterschiede") are sufficient for this.

At the same time, Bosnia and Montenegro have risen from semi-centers to full centers in some respects, with the publication of prescriptive grammars and thus the beginnings of internal codifications. However, there are still no normative monolingual dictionaries, Bosnian grammars sometimes contradict each other considerably (cf. Lehfeldt 2003), and the standard set by Nikčević's (2001) Montenegrin grammar is followed by almost no one, so 🗋99 that Ammon's classification would have to be further differentiated here in order to more accurately represent the status of Bosnia and Montenegro as 'three-quarter centers'.

Whether the described nationalization of the Serbo-Croatian standard varieties was a renationalization depends on whether the language forms used in the 19th century, for example by Ljudevit Gaj and Vuk Karadžić, were *national* varieties (which would then have been regionalized later) or perhaps rather regional or (sub)state varieties. In my opinion, there is some evidence to suggest that this was the case, i.e. that the chosen form of linguistic expression in the 19th century did not only depend on the area in which one lived, but that, for example, within Slavonia, Bosnia, or Krajina, Catholics and Orthodox Christians, even if they spoke the same dialect, used different standard varieties that, due to their religious affiliations, differed not only in their alphabets but also lexically and structurally. But then again, in the case of three of the four peoples it is at least questionable whether they can be described as fully developed nations in the 19th century. However, this question would require a detailed investigation that cannot be carried out within this framework.

4. Conclusion

The most important conclusion to be drawn is that, although no actual 'language division' took place around 1990, what happened was more than just a change in metalinguistic discourse or the creation of mere 'political languages' (since the national varieties undoubtedly also exist outside politics): the standard varieties of the four peoples were *nationalized*.

As stated in the introduction, it is by no means the aim of this article to use Ammon's terminology to demand that only the pluricentric totality of the varieties discussed should be referred to as *language*. In everyday life, there is no reason not to use the word *language* in all shades of its rich polysemy and therefore, just as one speaks of the *language of music*, the *language of bees*, *colloquial language*, *programming languages*, etc., one can also refer to Bosnian as a language or to Serbo-Croatian, depending on the context. In this sense, the *bon mot* attributed to Mark Twain (about England and America) and Karl Kraus (about Germany and Austria), among others, can of course also be applied (cf. Hägi 2006: 17): Bosniaks, Croats, Montenegrins, and Serbs are four nations divided by their common language.

In linguistics, however, the polysemous word *language* must either be converted into a clearly defined term, or it must be avoided in the same way as the word

word, which is also replaced by terms such as *lexeme*, *word form*, *word position*, *orthographic word*, *phonetic word*, etc. What is important is that the same terms are used for the same things, and different terms for different things. Therefore, for example, the Croatian national standard variety and the abstract English language unit that overarches American, British, and Australian English cannot be referred to with the same term; and if a term that deviates from international usage is introduced, like Brozović's (1970: 66) *diasystem* (cf. Mønnesland 2003), this term must also be applied consistently to German, English, French, etc.

Politics is not affected by the terminological requirements of linguistics, and so each country is free to name its official language 100 as it wishes, and of course each country can also demand that laws, treaties, EU regulations, etc. be drafted in its respective national variety, as Austria has demonstrated. However, linguistics could help explain to politicians what this means and that a national variety is much, much more than a list of words.

References

Alexander, Ronelle. 2002/03. Bosnien, Croatian, Serbian: One language or three? *International Journal of Slavic Linguistics and Poetics* 44/45. 1–35.

Ammon, Ulrich. 1995. *Die deutsche Sprache in Deutschland, Österreich und der Schweiz: Das Problem der nationalen Varietäten.* Berlin, New York: De Gruyter.

Auburger, Leopold. 1999. *Die kroatische Sprache und der Serbokroatismus.* Ulm (Donau): Hess.

Back, Otto & Fussy, Herberg (eds.). 2006. *Österreichisches Wörterbuch.* 40th edn. Wien: öbv, hpt. [1st edn. 1951.]

Brborić, Branislav. 2000. Šta novo donosi bosanski pravopis? In idem, *O jezičkom raskolu (Sociolingvistički ogledi I)*, 45–47. Beograd, Novi Sad: CPL, Prometej. [First published in *Front* 23(49) (748) of 8 Dec 1972, 35.]

Breu, Walter. 1998. Das Moliseslavische. In Rehder, Peter (ed.), *Einführung in die slavischen Sprachen (mit einer Einführung in die Balkanphilologie)*, 3rd edn., 274–278. Darmstadt: Wissenschaftliche Buchgesellschaft.

Brozović, Dalibor. 1970. *Standardni jezik: Teorija, usporedbe, geneza, povijest, suvremena zbilja.* Zagreb: Matica Hrvatska.

Brozović, Dalibor. 1992. Serbo-Croatian as a pluricentric language. In Clyne, Michael (ed.), *Pluricentric languages: Differing norms in different nations*, 347–380. Berlin, New York: De Gruyter.

Brozović, Dalibor & Pavle Ivić. 1988 [1990]. *Jezik, srpskohrvatski/hrvatskosrpski, hrvatski ili srpski.* Zagreb: Jugoslavenski Leksikografski Zavod Miroslav Krleža. [Main publication 1990 in *Enciklopedija Jugoslavije*, 2nd edn., vol. 6, 48–94. Zagreb: Jugoslavenski Leksikografski Zavod Miroslav Krleža.]

Bugarski, Ranko. 2000. Serbo-Croatian: How many languages? In Kunzmann-Müller, Barbara (ed.), *Die Sprachen Südosteuropas heute: Umbrüche und Aufbruch*, 192–199. Frankfurt am Main: Peter Lang.

Clyne, Michael (ed.). 1992. *Pluricentric languages: Differing norms in different nations.* Berlin, New York: De Gruyter.

Danica 1835 = Slavenski puki (iz Herdera, Bečkoga izdanja). *Danica Horvatska, Slavonska i Dalmatinska* 1(33) (22 Aug 1835). 231–232. [Reprint in *Danica ilirska* (ed. by Ivo Frangeš and Mladen Kuzmanović), vol. 1, Zagreb: Liber 1970.]

[Davidović, Dimitrije.] 1816. Istoria Slavjana (G. Herderom). *Zabavnik* 1816. 248–254.

Deklaracija 1991 = *Deklaracija o hrvatskome jeziku: s prilozima i Deset teza.* (Ed. by Stjepan Babić, Dalibor Brozović and Josip Pavičić.) Zagreb: Matica Hrvatska. [First published in *Telegram*, 17 Mar 1967. 1.]

Duličenko, Aleksandr D. 2003. *Slavjanskie literaturnye mikrojazyki: Obrazcy tekstov.* Vol. 1. Tartu: Izdatel'stvo Tartuskogo universiteta.

European Union. 1995. *Documents concerning the accession of the Republic of Austria, the Republic of Finland and the Kingdom of Sweden to the European Union.* Vol. 2. Luxembourg: Office for Official Publications of the European Communities.

Greenberg, Robert D. 2004. *Language and identity in the Balkans: Serbo-Croatian and its disintegration.* Oxford: Oxford University Press.

Gröschel, Bernhard. 2003. Postjugoslavische Amtssprachenregelungen – Soziolinguistische Argumente gegen die Einheitlichkeit des Serbokroatischen? *Srpski jezik* 8(1/2). 135–196.

Grupa članova UKS. 2000. Predlog za razmišljanje [povodom hrvatske jezičke deklaracije]. In Brborić, Branislav, *O jezičkom raskolu (Sociolingvistički ogledi I)*, 16–17. Beograd, Novi Sad: CPL, Prometej. [First published in the Zagreb edition of *Borba* 32(90) of 2 Apr 1967.]

Hägi, Sara. 2006. *Nationale Varietäten im Unterricht Deutsch als Fremdsprache.* Frankfurt am Main: Peter Lang.

Halilović, Senahid. 1998. *Bosanski jezik*. 2nd edn. Sarajevo: Baština.

Herder, Johann Gottfried. 1791. Slavische Völker. In idem, *Ideen zur Philosophie der Geschichte der Menschheit.* Vol. 4, chapter 16.4. Riga, Leipzig: Hartknoch. [Editions: *Sämmtliche Werke*, vol. 14, Berlin: Weidmann 1909, 277–280; *Herders Werke in fünf Bänden*, vol. 4, Berlin, Weimar: Aufbau 1964, 392–394.]

Hraste, Mate. 1965. O trećoj varijanti hrvatskosrpskog književnog jezika. *Jezik* 13(4). 106–113.

Isaković, Alija. 1998. Varijante na popravnom ispitu. In Halilović, Senahid, *Bosanski jezik*, 2nd edn., 241–258. Sarajevo: Baština. [Originally published in *Život* 19(11–12), 54–71.]

Janković, Srđan. 1982. Distinktivni pokazatelji standardnojezičke varijantnosti. *Naše teme* 26(5). 841–851.

Katičić, Radoslav. 1984. The making of Standard Serbo-Croat. In Picchio, Riccardo & Goldblatt, Harvey (eds.), *Aspects of the Slavonic language question*, 261–295. New Haven: Yale Concilium on International and Area Studies.

Katičić, Radoslav. 1997. Undoing a "unified language": Bosnian, Croatian, Serbian. In Clyne, Michael (ed.), *Undoing and redoing corpus planning*, 165–191. Berlin, New York: Mouton De Gruyter.

Keipert, Helmut. 2000. Herders „Slaven-Kapitel" in der „Danica Horvatska, Slavonska i Dalmatinska" (1835). In Jekutsch, Ulrike & Kroll, Walter (eds.), *Slavische Literaturen im Dialog: Festschrift für Reinhard Lauer zum 65. Geburtstag*, 129–140. Wiesbaden: Harrassowitz.

Kloss, Heinz. 1952. *Die Entwicklung neuer germanischer Kultursprachen von 1800 bis 1950.* München.

König, Werner. 2005. *dtv-Atlas Deutsche Sprache.* 15th edn. München: dtv. [1st edn. 1978.]

Kordić, Snježana. 2004a. Pro und kontra: „Serbokroatisch" heute. In Krause, Marion & Sappok, Christian (eds.), *Slavistische Linguistik 2002: Referate des XXVIII. Konstanzer Slavistischen Arbeitstreffens, Bochum, 10. 9.–12. 9. 2002*, 97–148. München: Sagner.

Kordić, Snježana. 2004b. Die aktuelle Sprachzensur in Kroatien. In: Symanzik, Bernhard & Birkfellner, Gerhard & Sproede, Alfred (eds.), *Sprache · Literatur · Politik: Ost- und Südosteuropa im Wandel*, 259–272. Hamburg: Dr. Kovač.

Lehfeldt, Werner. 2003. O stanju kodifikacije bosanskoga jezika. *Wiener slawistischer Almanach,* Sonderband 57. 147–151.

Lehfeldt, Werner. 1999. Zur gegenwärtigen Situation des Bosnischen. *Wiener slavistisches Jahrbuch* 45. 83–90.

Manifest 2004 = [Henisch, Peter & Hintze, Christian Ide & Muhr, Rudolf & Schindel, Robert & Streeruwitz, Marlene.] „Österreichisch" als eigene Sprache! Keine „deutsche" Rechtschreibreform mehr! Besondere Beachtung des europäischen Kontextes! *Der Standard* 14./15.08.2004, 31. https://archiv.sfd.at/sprache/ (last checked 22 Mar 2026).

Markhardt, Heidemarie. 2006. 10 Jahre „Austriazismenprotokoll" in der EU: Wirkung und Nichtwirkung – Chancen und Herausforderungen. In Muhr, Rudolf & Sellner, Manfred B. (eds.), *Zehn Jahre Forschung zum Österreichischen Deutsch: 1995–2005; Eine Bilanz*, 11–38. Wien, Frankfurt am Main: Peter Lang.

Mønnesland, Svein. 2003. O pojmu "dijasistem". *Wiener slawistischer Almanach*, Sonderband 57. 153–160.

Nikčević, Vojislav. 1997. *Pravopis crnogorskog jezika.* Cetinje: Crnogorski PEN Centar.

Nikčević, Vojislav. 2001. *Gramatika crnogorskog jezika.* Podgorica: Dukljanska Akademija Nauka i Umjetnosti.

Okuka, Miloš. 1998. *Eine Sprache – viele Erben: Sprachpolitik als Nationalisierungsinstrument in Ex-Jugoslavien.* Klagenfurt: Wieser.

Okuka, Miloš. 2000. Jezik u ustavima triju Jugoslavija, Hrvatske i Bosne i Hercegovine. *Riječ* 6(1/2). 66–76.

Pranjković, Ivo. 2006. Hrvatski i srpski su jedan jezik [razgovarala Vesna Laušić]. *Slobodna Dalmacija* 7 Feb 2006. https://web.archive.org/web/20090318214353/http://arhiv.slobodnadalmacija.hr/20060207/kultura01.asp (last checked 22 Mar 2026).

Radovanović, Milorad. 2003. Srpski jezik na početku milenijuma: Inventar eksternih i internih pitanja. *Wiener slawistischer Almanach*, Sonderband 57. 233–243.

Rehder, Peter. 1998. Das Serbokroatische. In idem (ed.), *Einführung in die slavischen Sprachen (mit einer Einführung in die Balkanphilologie)*, 3rd edn., 300–309. Darmstadt: Wissenschaftliche Buchgesellschaft.

Silić, Josip. 1996. Hrvatski jezik kao sustav i kao standard. In Turk, Marija (ed.), *Riječki filološki dani: Zbornik radova s Međunarodnoga znanstvenog skupa*, vol. 1, 187–194. Rijeka: Pedagoški fakultet.

Šipka, Milan. 2003. Jezik Bošnjaka, Hrvata, Srba i Crnogoraca – problem klasifikacije i nominacije idioma. *Wiener slawistischer Almanach*, Sonderband 57. 255–282.

Stamatović, Pavle. 1832. Slavenski narodi. *Serbskïj lětopis, nastavlěn Ġeodorom Pavlovičem* 8(3) (30). 27–31.

Sučić, Robert. 2005. "… mir ljudem … ki su dobre volje!". *Glasilo Hrvatskoga kulturnoga društva u Gradišću* 2007(6). 1.

Thomas, Paul-Louis. 1994. Serbo-croate, serbe, croate…, bosniaque, monténégrin: Une, deux…, trois, quatre langues? *Revue des études slaves* 66(1). 237–259.

Vuk Stefanović Karadžić. 1849. Srbi svi i svuda. In idem, *Kovčežić za istoriju, jezik i običaje Srba sva tri zakona*, 1–27. Beč: Štamparija Jermenskoga manastira. [Reprinted in *Sabrana dela*, vol. 17, 31–48. Beograd: Prosveta 1972.]

On the usefulness of Google Books & co. for histories of concepts (not only in the Slavic Studies): *glasnost'*, *standard language*, and *digraphia/bigraphism*

The article shows how new insights for the history of concepts can be gained using Google Books and other large collections of full texts. Various problems that are encountered are discussed. By demonstrating the method, new light is shed on the history of the concepts of *glasnost'*, *standard language* and *digraphia/bigraphism*.

Translated from "Über den Nutzen von Google Books & Co. für (nicht nur slavistische) Begriffsgeschichten" in Podtergera, Irina (ed.), *Schnittpunkt Slavistik: Ost und West im wissenschaftlichen Dialog; Festgabe für Helmut Keipert zum 70. Geburtstag*, vol. 1, 393–407. Göttingen: V&R unipress – Bonn University Press 2012. DOI: 10.18716/bun/gglb.

DOI: 10.18716/omp.64.5

One of Helmut Keipert's many contributions to the history of Slavic words and terms,[1] the essay on *glasnost'* published shortly before his retirement, begins as follows:

> „Die folgenden Ausführungen […] sollen vor allem daran erinnern, daß man bei begriffsgeschichtlichen Studien in der Russistik von der Lexikographie nicht diejenige Unterstützung erhält, die dafür erforderlich wäre" (Keipert 2006a: 1).
>
> "The following remarks […] are primarily intended to remind readers that for studies on the history of concepts, Russian lexicography does not provide the support that would be necessary for them."

However, help may now be available from a completely different source—namely, databases of digitized full texts. In this short paper, I would like to explore how this works, how much support they can provide, and what they cannot accomplish, using a few examples.

Among the full-text databases currently available, the Google Books project run by the American company Google Incorporated is by far the most ambitious and advanced. Among other things, the plan is to scan the entire library holdings

[1] The systematic list of writings published in the festschrift for his 65th birthday (Bunčić 2006) contains as many as 15 papers devoted to the history of a single word or term (positions 27, 58, 115–127). At least two more recent papers should now be added to this list, namely another on *cerkovnoslavjanskij* (Keipert 2006b) and one on *narodnost'* (Keipert 2008).

of the University of Michigan. In addition to many American libraries, several European libraries (and one in Japan) have now also decided to collaborate with Google, which means that the historical collections of the Austrian National Library, which are no longer protected by copyright and are so important for Slavic Studies, as well as the Bavarian State Library, which held the Special Subject Collection for Slavic Studies in Germany (funded by the German Research Foundation) until 1997, are now also to be digitized. Unfortunately, 394 no library in a Slavic country has joined the project yet (as of April 2011; https://web.archive.org/web/20110806014418/http://www.google.de/googlebooks/partners.html). Nevertheless, a considerable number of Slavic books—from the collections of the participating American and European libraries—are already available via the book search.

In the entire collection of books scanned to date, you can search for words, phrases, and exact word forms (in quotation marks) in the usual way using the Google internet search engine (at https://books.google.com/). Unlike an Internet search, however, the advanced search form allows you to limit your search to a specific period and display the results in chronological order (unfortunately only in reverse chronological order, as the technicians working on this project apparently did not think of the possibility of searching for the oldest reference instead of the most recent hit). If the copyright for the book found has expired, the page on which the hit is located is displayed as a facsimile, with the words searched for usually highlighted in color, and often you can even download the entire book as a PDF file if you wish. In the case of newer books that are still under copyright, sometimes only individual pages are displayed (and the hit may or may not be on one of the pages displayed), sometimes you only get a small "snippet" with the page number, and sometimes even just the general information that a particular book contains the word. However, even such sparse information can be helpful in finding the desired reference in a library in the traditional way.

Since only keywords can be entered in the book search engine, this tool is only suitable for those types of the history of concepts that take a "middle path between word history, factual history, and problem history" ("vermittelnder Weg zwischen Wortgeschichte, Sachgeschichte und Problemgeschichte", Keipert 2006a: 4) and thus have a semasiological element. An onomasiological search for expressions with a specific meaning is, of course, not possible. However, in the case of histories of concepts that trace the development of different expressions for a term, full-text databases can also help to illuminate the history of each of these expressions if the expressions used for the term are known from other sources.

As our first search term to illustrate the method, let us take the keyword ***glasnost'***, which became known internationally through Gorbačëv and which Keipert (2006a) uses as an example to criticize the inadequate representation of the history of the term in Russian dictionaries. Even in the 19th century, the 395

keyword, which some consider a neologism of the 1980s, played a political role when it came to the *publicity* of court proceedings in particular *(glasnost' sudoproizvodstva)*. The word itself is, of course, somewhat older; the first known evidence of its use can be found in Trediakovskij (1752). This is also the oldest hit that Google Books brings up (https://books.google.de/books?id=s0oEAAAAYAAJ&pg=PA9, digitized on 3 Jul 2007[2]). There, however, *glasnost'* is used only in the sense of "pure pronunciation, clear articulation" ("reine Lautung, klare Aussprache", Keipert 2006a: 7, 12). The central problem, which even Keipert (2006a: 13f.) was unable to solve with the means available at the time, is precisely that "we cannot yet trace the path of metaphorization of *glasnyj* from 'loud' to 'openly apparent, generally known'" ("wir den Weg der Metaphorisierung von 'laut' zu 'offen zutage liegend, allgemein bekannt' bei *glasnyj* bisher nicht verfolgen können").

Here, Google Books really seems to be able to help. Three hits from a search for *glasnyj* are particularly interesting: the first is from a dictionary that Keipert (2006a: 13) also quotes—in a later edition—, the *Polnyj nĕmecko-rossijskoj leksikon, iz bol'šago gramatikal'no-kritičeskago Slovarja gospodina Adelunga sostavlennyj* (vol. 2, Sanktpeterburg 1798). However, Keipert searched for *glasnyj* as a counterpart to *öffentlich* 'public' and did not find it in that article. The digital full-text search has the great advantage that one does not have to know in advance where one might find something. One would not have readily looked up *ruchtbar* 'notorious', but this very adjective is translated as "извѣстный, вѣдомый, гласный, явный" and illustrated, among other things, by the example "Eine ruchtbare That, гласное дѣло" ("a notorious act"). The corresponding noun can also be found here: "die Ruchtbarkeit, (plural: неуп:) извѣстность, явность, гласность" (…?id=Dc0GAAAAQAAJ&pg=PA320, 2 Mar 2007).

The precise usage of *glasnyj* at the end of the 18th century is illuminated by a letter from Catherine the Great to her emissary at the Sublime Porte, Jakov Ivanovič Bulgakov, dated 30 December 1782, which is reprinted in Dubrovin (1889: 974–976; and this collection of sources from 1781 and 1782 is sensibly filed under 1782 and not 1889 on Google Books). At the end of lengthy instructions on diplomatic maneuvers, the tsarina writes:

> «Покуда Порта о семъ не станетъ вызываться, вы сохраните все сіе въ непроницаемой *тайнѣ*, да и нѣтъ нужды сообщать сіе предварительно 396 интернунцію вѣнскаго двора; но когда

> "Until the Porte raises this issue, keep all this in strict *secrecy*, and there is no need to inform the internuncio of the Viennese court in advance; but when it already *becomes public knowledge*, then you may

[2] In the following, the addresses of the documents on Google Books are shortened by the constant "https://books.google.com/books", so that only the identification number and the reference to the page with the hit are given, supplemented by the date of digitization. The retrieval date is generally omitted—all information relates to April 2011 [and has been updated to March 2026 for the English translation—D. B. 2026].

уже оно *учинится гласнымъ*, тогда вы можете сказать интернунцію о вопросахъ, отъ Порты вамъ сдѣланныхъ, и о вашемъ отвѣтѣ […]» (ibid. 975f.; …?id=GSYbAAAAYAAJ&pg=PA976, 23 May 2008; my emphasis).

inform the internuncio of the questions put to you by the Porte and of your reply […]."

In a similar context, Levašov (1790: 155f.) uses the word in a report about his imprisonment in the Ottoman Empire:

«[…] Апреля 1го числа явился у насъ […] вѣстникъ съ объявленїемъ, что мы будемъ отпущены въ отечество […], и чтобы мы сїе содержали до времени въ *тайнѣ*, дабы народъ несталъ на Правительство роптать […]; но осторожность сїя не была сохранена въ Царь-градѣ съ надлежащею точностїю, гдѣ немѣдленно разпространился слухъ, что насъ выпустятъ на волю […], и мы увѣдомясь чрезъ прїятелей своихъ, что освобожденїе наше вездѣ почти *содѣлалось* уже *гласнымъ*, опасались, чтобъ сїе самое неоставило насъ и надолго въ Демотикѣ […].» (…?id=07o9AAAAYAAJ&pg=PA155, 8 Jan 2009; my emphasis).

"[…] On April 1, a […] messenger arrived with the announcement that we would be released to return to our homeland […], and that we should keep this information *secret* for the time being so that the people would not begin to complain to the government […]; but this caution was not maintained with due precision in Constantinople, where rumors immediately spread that we would be released […], and when we learned from our friends that our release had already *become public knowledge* almost everywhere, we feared that this very fact would keep us in Demotika for a long time […]."

In both cases, *glasnyj* is used to refer to a secret that has *become public knowledge*—something that should have been *kept secret* is now obvious. This could be the missing link between the original and the transferred meaning: when a secret is no longer a secret, it has usually been *spoken about*, *gossiped* or even *announced*; in any case, the silence has been broken by a voice *(glas)*. This development in meaning would be parallel to the German *ruchbar (werden)*, because this goes back to Middle Low German *ruchte* 'call, cry, reputation' (and, like *Gerücht* 'rumor', *berüchtigt* 'infamous' and also *anrüchig* 'objectionable', not to *riechen* 'to smell', cf. Pfeifer 2003: s.v.). From there, it is only a small step to the meaning 'public', because when *glasnost'* became a buzzword in the 19th century, it was, as Keipert (2006a: 9) rightly notes, a "future concept" ("Zukunftsbegriff"), i.e. the court proceedings were still secret at that time. What was being negotiated in secret was therefore yet to *become public knowledge*. Only when *glasnost'* is applied to an already existing practice does its meaning approach that of *publičnost'*, *javnost'* 'publicity'. Significantly, however, Gorbačëv's use of this keyword was also about the *introduction* of freedom of information and expression, so here, too, something that had previously been secret was now to be spoken about aloud.

The fact that not only in Russian or, as Keipert (1998: 122) has stated elsewhere, "in most Slavic languages the lexicographical prerequisites for more extensive etymological studies are lacking" ("für weiter ausgreifende wortgeschichtliche Studien in den meisten slavischen Sprachen die lexikographischen Voraussetzungen fehlen") but that German and English dictionaries, for example, 🗋397 also leave room for improvement, particularly with regard to technical terms, can be seen in the history of the term ***standard language***. Gröschel (2009: 92–95) has traced this history on the basis of specialized and general dictionaries: While Joseph (1987: 5) finds the first attestation of *standard language* in English in 1858, dictionary research shows that *standardni jezik* has been documented in Serbo-Croatian since 1964, *literaturnyj (standartnyj) jazyk* in Russian since 1966, and *Standardsprache* in German since 1968. Gröschel (2009: 93) supplements the evidence from Russian dictionaries with a reference in the title of a much earlier paper by Polivanov (1927). Gröschel (2009: 93) has "more than just slight doubts" ("mehr als nur gelinde Zweifel") about this chronology, especially since he cannot imagine that "such a large time gap should yawn" ("eine derart große zeitliche Lücke klaffen sollte") in German between *Standard*, which was already documented in 1804, and *Standardsprache*, which was first documented in 1968 (ibid. 94).

As Google Books shows, these doubts are entirely justified. For example, *Standardsprache* can already be found in several journal articles from the end of the 19th century, first in two reviews by Siebs (1897a: 220, …?id=18cqAAAAMAAJ &pg=PA219, 12 Feb 2007; 1897b: 555, …?id=A-IevztN2O8C&dq=Standardsprache&pg=PA555, 4 Jun 2007; Siebs' orthoepic dictionary, on the other hand, still exclusively writes *Hochsprache* for 'standard language' in its 16th edition of 1957). Jespersen (1904: 39) also refers to "Einheitssprache, Gemeinsprache, Standardsprache oder wie man sie sonst nennen will" ("uniform language, common language, standard language, or whatever else you want to call it"). Up to 1967, Google Books yields a total of 146 books containing the term *Standardsprache*, which clearly shows that the term was used more than sporadically before it first appeared in a dictionary. In Serbo-Croatian, references in journals can be found from 1950 onwards (first in Tabak 1950: 320; …?id=FwVBAQAAIAAJ&q=standardnim). For Russian, Polivanov (1927) has not yet been digitized, but bibliographic references to this essay can be found, as well as a number of attestations of *standartnyj jazyk* between 1927 and 1966.

The first recorded use of the English term *standard language* in its modern sense[3] can also be traced back to 1858, at least 81 years earlier than previously

[3] In an obviously different sense, probably as a nonce word, *standard language* appears in Fell (1729: 302), in the report on the events following the capture of the city of Hippo by the Vandals in 431, one year after the Church Father Augustine had died there: "[…]but the Barbarians ſhew'd Reſpect to the Saint's Body and Writings, which made it evident that the Almighty reſtrain'd them from inſulting the Remains of His Servant, and reſerved

thought. In Google Books, 🗋398 a review of Clarke (1777) can be found in *The Critical Review: or, Annals of Literature* 45 from 1778 (pp. 248–252; …?id=7cI-PAAAAQAAJ&pg=PA248, 13 Mar 2008) in which the book is quoted. Although the book itself has not been digitized by Google Books, it is available online in all German university libraries through a national license. It deals with the education of young men, and in this context the author explains:

> "If ſuch attention is alſo given to the dialect and pronunciation, as to form their language to the *national ſtandard*, they may one time or other be benefited thereby; eſpecially if they have occaſion to be in places diſtant from that of their nativity. I have ſeen men who knew not *p* from *q*, by being habituated in youth to hear and ſpeak the *ſtandard language*, have more addreſs and ſentiment, than others who had been taught reading, writing, and the uſe of numbers […]" (Clarke 1777: 18; my emphasis)

A document written five years later makes it clear that the word is also used in philology itself. In the dispute over the authenticity of the 15th-century poems by Thomas Rowley, which were forged by Thomas Chatterton in the 18th century (there was also an "invented past" ("erfundene Vergangenheit", cf. Keipert 2001) in Western Europe), Tyrwhitt (1782: 3) opposes two arguments put forward by those who believe in the poems' authenticity:

> "1. That the Poems are written in a provincial dialect, and therefore are not reducible to the rules of the *ſtandard-language*. 2. That there was no *ſtandard-language* in the XV century, by which they can be tried." (…?id=JkZWAAAAYAAJ&pg=PA3, 3 Mar 2011; my emphasis)

Here, too, Google Books is more helpful than traditional lexicography.

Highly specialized scientific terms, such as those describing the phenomenon of **biscriptality**—i.e. the fact that several scripts are used for one and the same language (which is the subject of my postdoctoral research project)—are hardly ever included in dictionaries. Of particular interest in terms of the history of the concept is the term *digraphia*, which is based on diglossia according to Ferguson (1959) and which was coined independently several times. It was popularized by DeFrancis (1984), who notes in a footnote that after completing his manuscript, he learned that this term had already been used in a similar sense by Dale (1980). However, Grivelet (2001: 1f.) cites Zima (1974) as "[t]he first discussion of the notion of digraphia", and Unseth (2008: 3) adds Jaquith (1976) as another "inventor" of this term. A Google Books search 🗋399 for "digraphie" finds Gebhardt

---

his Works as the Standard Language of the Church on the important and difficult Articles of Grace and Predeſtination" (…?id=CRgQ6lqsuOwC&pg=PA302). There, *language* is still to be understood in a sense that is no longer common, which the *Oxford English Dictionary* (s.v. *language* 5a) describes as "That which is said; talk, report, rumour", here perhaps more freely translated as *doctrine* or *teaching*.

(1974: 175; …?id=CpIrAAAAMAAJ&q=digraphie), among others. Unfortunately, this book can only be viewed in "snippet view", so one must first borrow it from a library to see that Gebhardt here refers to Lafont (1971). Lafont's article is freely accessible on the journal portal *Persée* (doi:10.3406/lfr.1971.5576[4]), and so arguably the very first evidence of the linguistic term *digraphie* for 'biscriptality' has been found, which Lafont mentions only in passing and which was certainly not taken up by the other "inventors". This passage discusses the fact that Occitan not only competes with Standard French, but also has two competing orthographies:

> « La situation de diglossie occitane n'est donc pas semblable absolument à celles qu'on peut trouver en d'autres lieux de contacts linguistiques : elle se complète par une situation de digraphie » (Lafont 1971: 95).
>
> "The situation of Occitan diglossia is therefore not at all similar to those found in other areas of language contact: it is complicated by a situation of digraphia."

As an aside, let me remark that the *Bibliographie linguistique*, whose online edition, after a free trial and promotional period, is now unaffordable for many libraries, can be viewed in a limited preview on Google Books, so that a keyword search for *digrafia* there leads to references to the second and third parts of Consani's (1988, 1989, 1990) trilogy of articles under nos. 7030 and 7031 in the *Bibliographie linguistique* for the year 1990 (…?id=os9TMERY9OsC&pg=PA357). This allows us to trace a fifth and (so far) last independent "inventor" of this expression.

Similarly interesting and complex is the prehistory of sociolinguistic concepts of biscriptal writing, namely the emergence of expressions for 'biscriptal', e.g. for documents that contain the same text in two scripts (but in the same language). This has not yet been researched at all. Google Books finds the oldest relevant mention of *digraphic* in Pierides (1876; …?id=IBsXAQAAIAAJ&q=digraphic), but only small excerpts are displayed. If you then obtain the article from the library, you will find that Pierides (1876: 38) actually acknowledges having invented the term himself:

> "In the summer of 1873 I became possessed of an inscription in Greek and Cypriote, then discovered in Larnaca, the ancient Citium. […] As the language is the 400 same in both parts, and only the writing differs, I prefer calling this inscription *digraphic*, instead of *bilingual*, until a better definition is proposed."

Like so many temporary solutions, this one also proved to be surprisingly durable, as the term is still in use today, especially in Greek philology—Consani (1988, 1990) undoubtedly continues this usage with his coinage of *digrafia*. At almost the

[4] Unlike the *Uniform Resource Locator (URL)*, the *Digital Object Identifier (DOI)* is permanent. If the browser does not automatically resolve such addresses, "doi:" has to be replaced with "https://doi.org/".

same time, von Sallet (1875: 132) independently introduced the German term *zweischriftig* in a completely different discipline, numismatics, for coins that contain inscriptions in two different scripts—here, too, the two scripts are the Cypriot syllabary and the Greek alphabet:

> „[…] einige dieser Münzen, welche als zweischriftig – sit venia verbo – besonders interessant sind, geben neben der cyprischen auch die griechische Legende […]."

> "[…] some of these coins, which, being *zweischriftig*—if you permit the expression—, are particularly interesting, also bear the Greek inscription in addition to the Cypriot one […]."

This hit can also only be viewed as a "snippet" on Google Books (…?id=CU1VIO 6izyEC&q=zweischriftig, 17 May 2005) and had to be obtained in paper form for the complete text. Another synonym for *digraphic* and *zweischriftig* is *bigraphic*, which is documented much earlier than the other two, namely by Pierquin de Gembloux (1840). However, the content of this book is more than dubious, as it describes, as a preliminary study for a *History of the Fatherland before the Roman Conquest* (*Histoire de la Patrie avant la conquête romaine*, ibid. X), the migrations of the Celts across half the world, including America. For example, on a large rock in the Mississippi there allegedly is "a bigraphic Celtic inscription, i.e. half hieroglyphic, half alphabetic" (« une inscription celtique bigraphique, c'est-à-dire moitié hiéroglyphique, moitié alphabétique », ibid. 248; …?id=jh1GAAAAcAAJ&pg= PA248, 25 Jan 2007). Here again, the advantage of the method becomes apparent, because no one would have thought to look for the first evidence of this linguistic term in such an abstruse work. Of course, it is also unlikely that this work influenced the later use of this term in linguistics, so that we are probably dealing here with a nonce coinage. Later users of the same term, including Siméon (1889: IX), who comments on Mexican pictographic manuscripts with Latin glosses in Nahuatl, saying that they "could more exactly be called *bigraphic*" (« pourraient être plus exactement appelés *bigraphiques* »), are certainly independent of Pierquin. For the latter book, however, Google Books only provides a reference that it contains the word sought (…?id=3V8SAAAAYAAJ&pg=PR9, 5 Dec 2007), after which one must borrow it in paper form and search for the relevant passage. However, the address referred to in the hit list contains a code for the page number, even if it is not displayed on the page ("PR9" in this case is p. IX, with 'R' for Roman numerals, as opposed to "A" for Arabic numerals).

401 The extent to which the "forgetting of knowledge" is not only commonplace "in Slavic philology" ("Vergessen von Wissensbeständen […] in der Slavischen Philologie", Keipert 2006c) but also in other fields, is demonstrated, apart from the sixfold 'invention' of *digraphia*, also by the fact that Blake (1995: 463) still speaks of a "'bigraphic' scribe (to invent a suitable term)" ("escriba 'bígrafo' (para inventar un término adecuado)", …?id=sq7jAAAAMAAJ&q=bígrafo, "snip-

pet view"), even though, as seen above, this term has not needed to be invented since 1889 (or 1840).

One of the earliest examples of a noun describing 'biscriptality' (of texts), incidentally, comes from Russia, specifically from Ol'denburg (1899: 208), who writes about an Indian manuscript:

> «[Ч]резвычайно любопытную особенность этого отрывка составляетъ то, что въ немъ мы имѣемъ образчикъ *биграфизма*, а именно на листѣ 27b. мы встрѣчаемъ *одновременно* и письмо характера *индійскаго* gupta и *кашгарскаго* […].» (original emphasis)
>
> "An extremely curious feature of this passage is the fact that it provides us with an example of *bigraphism*; namely, on folio 27b, we encounter *simultaneously* writing in both the *Indian* gupta script and the *Kashgar* script […]."

However, this text has not yet been included in Google Books. Nevertheless, when searching for *Bigraphismus*, one comes across a "snippet" from Barthold (1899: 140; …?id=byfWAAAAMAAJ&q=bigraphismus&pg=PA140, 29 May 2009), which, as one can discover by examining a paper copy of this journal, reports on Ol'denburg's essay in German. In this way, despite Google Books's focus on Western libraries, so far neglected Slavic works sometimes leave their mark.[5]

These brief excerpts from the history of the concept of biscriptality may suffice to demonstrate the possibilities but also some of the **obstacles** of working with Google Books. Here, the impossibility of an onomasiological search proves to be a serious problem, because the multitude of terms that have been used for the phenomenon of biscriptality (in addition to *digraphia*, *bigraphism*, and *Zweischriftigkeit*, there are also *bialphabetism*, *biscriptalism*, *orthographic diglossia*, and others) must be entered individually into the search engine—and in all languages in which relevant evidence is hoped to be found. In doing so, one can never be entirely sure that one has not overlooked an important discussion that could actually be found with the available means, because another term was used there that one happens not to know.

402 Almost as large a problem as the synonyms used for a concept to be investigated is posed by polysemy and homonymy. For example, the relevant hits for a search for *digraphie* have to be fished out of a sea of irrelevant hits, since *digraphie* a) is also used as an alternative form of *digraph* (referring e.g. to Polish *cz*, *ch*, *rz*, etc.), b) refers to a method in radiology used to display the lungs in both

[5] A much more detailed treatment of the history of this term will be published as part of my habilitation thesis. [It was published as chapter 2 of Bunčić (2016: 27–50), which elaborates on the history of the notion itself but does not say much about the methods used to investigate this history. — D. B. 2026.]

the inhaled and exhaled state on the same X-ray image, and c) is the usual word for double-entry bookkeeping in French.

Moreover, typing and reading errors can lead to incorrect hits. When searching for *bigraphisch* 'bigraphic', for example, the vast majority of hits are misspellings or incorrect readings by the text recognition program for *biographisch* 'biographical' (because if the *o* could not be recognized correctly, something like *bi•graphisch* is returned as a hit for *bigraphisch*). Poor image quality can lead to absurd errors—for example, Google Books reads the following passage in Schneidewein (1740: 466)

as "Ut quod in curia гласности non brincar" (…?id=mcnXB3uy4CcC&pg=PA466, 13 Jan 2010; with my human eyes, I read "Ut quod in curia mercatorum non obtineat"). Another problem is word separation at the end of a line, which means that a search for *glasnyj* also returns a number of hits where *glasnyj* appears at the beginning of a line, but the previous line ends, for example, with *so-*.[6]

When searching for the use of a word in the past that is very common in the present, it is often annoying that the metadata, including the publication year, always refer only to entire volumes. Thus, one is surprised to find a word such as *Standardsprache* 'standard language' in a text from 1484 (…?id=AMwJAQAAIAAJ &q=standardsprache, 11 Feb 2009)—but of course this is not the oldest attested use of the term in German but a hit in the accompanying text of a scholarly edition (in this case of Stephan von Landskron's *Hymelstrasz* of 1484, published by G. J. Jaspers in 1979 "mit einer Einleitung und vergleichenden Betrachtungen zum Sprachgebrauch in den Frühdrucken", i.e. "with an introduction and comparative observations on language use in early prints"). In addition, there are a surprising number of typing errors in the metadata (or are they also obtained through automatic text recognition?). One example is what at first glance appears to be the earliest evidence of *standard language* in English in a text dated to 1708 (where, incidentally, the term is used in the sense of standardized English-language biological terminology). However, this turns out to be a letter to the editor from the *Monthly Magazine* of February 17**9**8 (…?id=arXaiokRtygC&pg=PA110, 21 Sep 2005).

403 Google Books is controversial primarily because of its handling of **copyright**, and for good reasons. The opt-out procedure sought by Google, whereby the author (or copyright holder) of a printed work must be aware that Google is scanning

[6] As of 2026, the first two of these problems have been solved, probably by improved text recognition; the passage from Schneidewein is now contained as "Utquòd in curia mercatorum non obtineat" in Google Books. False positives due to hyphenation are also much less common now. (D. B. 2026.)

this work in order to object to it, is certainly not ideal from the perspective of commercial interests. Admittedly, though, the most comprehensive possible collection of printed literature, as intended by Google and desirable for scientific purposes, will never be achieved if all authors who have not been dead for more than 70 years must give their express consent (*opt-in* procedure). It seems to me that a slightly expanded "snippet view" (which, especially in the case of journals and anthologies, would have to contain the bibliographic information of the individual article) would be a sensible compromise for all works that are not in the public domain or have been released by their copyright holders. This would not harm the authors but would provide researchers with enough information to decide whether a text contains relevant information and, if necessary, should be obtained by other means. However, this issue seems to be of little relevance to the cases discussed here, as the hits all come from works whose copyright has either expired or which have been made available by the publisher on the basis of an agreement with Google (in these cases, the date of digitization is usually missing).

A much bigger problem than the copyright dispute is Google's monopolistic position. It is highly questionable whether a single private company, which is not subject to any government or intergovernmental control, should be allowed to make a large part of the world's knowledge universally accessible (even if this company has the unofficial motto "Don't be evil"). Therefore, **alternative projects** are very welcome. The most comprehensive such institution besides Google Books (which, according to Armstrong, has scanned 15 million books to date) is *HathiTrust* (https://www.hathitrust.org/), a very promising network of American research libraries that offers not only a union catalog but also a unified search of all digital full texts available at these libraries—including those digitized as part of the Google Books program. Unfortunately, there is no advanced full-text search function yet, which means that searches in the digital documents cannot be limited in time, filtered, or sorted, making this tool unsuitable for the purposes of conceptual history at present.

404 *Europeana*, an association of European libraries, museums, and archives, according to its own information, contains "more than 15 million objects" (including images, audio recordings, and videos, https://europeana.eu/about-us). Unfortunately, there is no functional full-text search feature yet. However, as with HathiTrust, we can expect improvements in the near future, which will hopefully make these databases genuine alternatives and additions to Google Books.

Gallica, the digital library of the French National Library (https://gallica.bnf.fr/), has proven helpful in several searches. It currently has 1.5 million digitized documents, but these are almost exclusively French texts and texts related to France. A Polish equivalent is the *Federacja bibliotek cyfrowych*, which currently has over 600,000 digitized publications (https://fbc.pionier.net.pl/), but so far it is only

possible to search the metadata. The same applies to the *Ėlektronnaja biblioteka* of the Russian State Library in Moscow, which has digitized 630,000 dissertations and over 8,300 old prints, among other things (http://elibrary.rsl.ru/).

A project with a different focus is JSTOR, which specializes in digitizing older issues of scientific journals (https://www.jstor.org/). Although the organization itself is non-profit, its services are so expensive that most libraries only purchase individual sub-collections. Nevertheless, it is possible to perform a full-text search of the entire collection; for articles that are not subscribed to by your own library and that contain the search term, only their bibliographical information is displayed. (Some articles can also be purchased directly online—according to my random samples, at prices ranging from $9 to $38.) However, searches conducted in JSTOR in parallel with the Google Books searches described above revealed that most of the relevant information could not be found there. In addition to the deliberate restriction to journals, this is probably also due to the strong focus on English-language sources.

Overall, I hope to have shown that Google Books can be useful for researching the history of concepts (and other full-text databases will certainly be useful in the future as well). Even with this tool, reading the original paper texts will not become unnecessary in the foreseeable future, and this method also requires a certain overview of the use of the term being investigated. In this respect, book searches cannot be compared to using Google's internet search engine as a corpus, which Robb (2003) describes as "quick 'n dirty". Since you often only get text snippets or bibliographic information and 🗋405 then, in order to read the entire text, may have to resort to traditional interlibrary loans, searching via Google Books is not necessarily fast. On the other hand, the results obtained in this way are by no means unreliable (as long as one is aware that there may be other uses of the search term that are not covered by Google Books). Therefore, given the often inadequate representation of the history of terms in dictionaries (not only Slavic ones), the method presented here can be a valuable aid.

References

Armstrong, Kim. 2011. One million books scanned and returned to CIC University libraries. *Inside Google Books*. https://web.archive.org/web/20110204173021/http://booksearch.blogspot.com/ (last checked 22 Mar 2026).

Barthold, W. [= Bartol'd, Vasilij V.]. 1899. Russische Arbeiten über Ostasien: Jahresbericht für 1898. *Mittheilungen des Seminars für Orientalische Sprachen an der Königlichen Friedrich-Wilhelms-Universität* 2. 140–151.

Blake, Robert J. 1995. El latín notarial de un escriba bilingüe o 'bígrafo' del XIII. In Pérez Gonzáles, Maurilio (coord.), *I Congreso Nacional de Latín Medieval: Actas (León, 1–4 de diciembre de 1993)*, 463–468. León: Universidad de León.

Bunčić, Daniel. 2006. Systematisiertes Schriftenverzeichnis von Helmut Keipert. In Bunčić, Daniel & Trunte, Nikolaos (eds.): *Iter philologicum: Festschrift für Helmut Keipert zum 65. Geburtstag*, 21–47. München: Sagner.

[Bunčić, Daniel. 2016. *Biscriptality: A sociolinguistic typology*. Edited by Daniel Bunčić & Sandra L. Lippert & Achim Rabus. With contributions by Anastasia Antipova, Carmen Brandt, Ekaterina Kislova, Henning Klöter, Alexandra von Lieven, Sandra L. Lippert, Helma Pasch, Achim Rabus, Jürgen Spitzmüller, and Constanze Weth. Heidelberg: Winter. — D. B. 2026.]

Clarke, Cuthbert. 1777. *The true theory and practice of husbandry*: *Deduced from philosophical researches, and experience*, Newcastle: T. Slack.

Consani, Carlo. 1988. Bilinguismo, diglossia e digrafia nella Grecia antica I: Considerazioni sulle iscrizioni bilingui di Cipro. In Campanile, Enrico & Cardona, Giorgio Raimondo & Lazzeroni, Romano (eds.), *Bilinguismo e biculturalismo nel mondo antico: Atti del Colloquio Interdisciplinare tenuto a Pisa il 28 e 29 settembre 1987*, 35–60. Pisa: Giardini.

Consani, Carlo. 1989. Bilinguismo, diglossia e digrafia nella Grecia antica II: Le lettere di Filippo v e i decreti di Larissa (Schwyzer, *DGEEP*, 590). *Annali del Dipartimento di Studi del Mondo Classico e del Mediterraneo Antico* 11. 137–159.

Consani, Carlo. 1990. Bilinguismo, diglossia e digrafia nella Grecia antica III: Le iscrizioni digrafe cipriote. *Orientamenti linguistici* 25. 63–79.

Dale, Ian R. H. 1980. Digraphia. *International Journal of the Sociology of Language* 26. 5–13.

DeFrancis, John. 1984. Digraphia. *Word* 35. 59–66.

Dubrovin, Nikolaj F. (ed.). 1889. *Prisoedinenie Kryma k Rossïi: Reskripty, pis'ma, reljacïi i donesenïja.* Vol. 4 (1781–1782). Sanktpeterburg: Glazunov.

Fell, Charles. 1729. *The lives of saints, collected from authentick records of church history*. Vol. 3. London: T. Osborne.

Ferguson, Charles A. 1959. Diglossia. *Word* 15. 325–340.

Gebhardt, Karl. 1974. *Das okzitanische Lehngut im Französischen*. Frankfurt am Main: Peter Lang.

Grivelet, Stéphane. 2001. Introduction. *International Journal of the Sociology of Language* 150. 1–10.

Gröschel, Bernhard. 2009. *Das Serbokroatische zwischen Linguistik und Politik: Mit einer Bibliographie zum postjugoslavischen Sprachenstreit*, München: Lincom.

Jaquith, James R. 1976. Digraphia in advertising: The public as guinea pig. *Visible Language* 10(4). 295–308.

Jaspers, Gerardus Johannes. 1979. Einleitung. In Stephan von Landskron, *Die Hymelstrasz: Mit einer Einleitung und vergleichenden Betrachtungen zum Sprachgebrauch in den Frühdrucken (Augsburg 1484, 1501 und 1510)*. Amsterdam: Rodopi.

Jespersen, Otto. 1904. *Phonetische Grundfragen*. Leipzig: Teubner.

Joseph, John E. 1987. *Eloquence and power: The rise of language standards and standard languages*. New York: Blackwell.

Keipert, Helmut. 1998. Überlegungen zur Terminologie der sog. nationalen Wiedergeburt der Slaven. In Bjørnflaten, Jan Ivar & Kjetsaa, Geir & Mathiassen, Terje (eds.), *A centenary of Slavic Studies in Norway: The Olaf Broch Symposium, Oslo 12–14 September 1996: Papers*, 121–141. Oslo: Norwegian Academy of Science and Letters.

Keipert, Helmut. 2001. Erfundene Vergangenheit? *Jahrbücher für Geschichte Osteuropa* 49(2). 264–267.

Keipert, Helmut. 2006a. *Glasnost′:* Zu den lexikographischen Voraussetzungen für begriffsgeschichtliche Untersuchungen im Russischen. In Thiergen, Peter (ed.), *Russische Begriffsgeschichte in der Neuzeit: Beiträge zu einem Forschungsdesiderat*, 1–21. Köln: Böhlau.

Keipert, Helmut. 2006b. *Cerkovnoslavjanskij* in verordnetem Vergessen. In Jesenšek, Marko & Zorko, Zinka (eds.), *Jezikovna predanost: Akademiku prof. dr. Jožetu Toporišiču ob 80-letnici*, 129–145. Maribor, Ljubljana: Slavistično društvo, SAZU.

Keipert, Helmut. 2006c. Über das Vergessen von Wissensbeständen in der Slavischen Philologie *oder* Jan Kollárs Abhandlung „Über die literarische Wechselseitigkeit" in der „Danica Ilirska" von 1836. *International Journal of Slavic Linguistics and Poetics* 44/45. 261–278

Keipert, Helmut. 2008. Ljudevit Gajs „Zauberwörtchen" *narodnost*. In Di Salvo, Maria & Moracci, Giovanna & Siedina, Giovanna (eds.), *Nel mondo degli Slavi: Incontri e dialoghi tra culture; Studi in onore di G. Brogi Bercoff*, vol. 1, 314–326. Firenze: Firenze University Press.

Lafont, Robert. 1971. Un problème de culpabilité sociologique: la diglossie franco-occitane. *Langue française* 9(1). 93–99.

Levašov, Pavel A. 1790. *Plěn i stradanïe Rossïjan u Turkov ili Obstojatel'noe opisanïe bědstvennyx priključenïj preterpěnnyx imi v Car-gradě po ob"javlenïi vojny i pri vojskě, za kotorym vlačili ix v svoix poxodax; s prioščeniem dnevnyx zapisok o voinskix ix dějstvïjax v prošedšuju vojnu, i mnogix strannyx, rědkix i ljubopytnyx proizšestvïj*. Sanktpeterburg: Bogdanovič.

Ol'denburg, Sergej Ḟ. 1899. Otryvki kašgarskix i sanskritskix rukopisej iz sobranïja N. Ḟ. Petrovskago, II. *Zapiski Vostočnago otdělenïja Imperatorskago russkago arxeologičeskago obščestva* 11 (1897–1898). 207–264.

Pfeifer, Wolfgang (ed.). 2003. *Etymologisches Wörterbuch des Deutschen*. München: dtv.

Pieridés, Demetrios. 1876. On a digraphic inscription found in Larnaca (read 6th April, 1875). *Transactions of the Society of Biblical Archaeology* 4(1). 38–43.

Pierquin de Gembloux, Claude-Charles. 1840. *Histoire monétaire et philologique du Berry.* Vol. 1. Bourges: Ménage.

Polivanov, Evgenij D. 1927. O literaturnom (standartnom) jazyke sovremennosti. *Rodnoj jazyk v škole* 1. 225–234.

Robb, Thomas. 2003. Google as a quick 'n dirty corpus tool. *TESL-EJ* [= Teaching English as a Second Language—Electronic Journal] 7(2). https://tesl-ej.org/ej26/int.html (last checked 22 Mar 2026).

Schneidewein, Johannes. 1740. *In quatuor institutionum imperialium D. Iustiniani libros commentarii*. Colonia Agrippina: Metternich. [1st edn. Argentoratum [Strasbourg]: Rihelius 1571.]

Siebs, Theodor. 1897a. Review: van Helten, W. L. Zur Lexikologie [sic; spelled "lexicologie" in the reviewed book] des Altwestfriesischen. […] Amsterdam […] 1896. *Literaturblatt für germanische und romanische Philologie* 18(7). 219–223.

Siebs, Theodor. 1897b. Review: Friesch woordenboek, bewerkt door Waling Dijkstra en dr. F. Buitenrust Hettena, […] Leeuwarden 1896. *Zeitschrift für deutsche Philologie* 29. 552–557.

Siebs, Theodor. 1957. *Deutsche Hochsprache: Bühnenaussprache*. 16th edn. (Ed. by Helmut de Boor and Paul Diels.) Berlin: De Gruyter. [1st edn. 1898.]

Siméon, Rémi. 1889. Introduction. In idem (ed.), *Annales de Domingo Francisco de San Anton-Muñon Chimalpahin Quauhtlehuanitzin: Sixième et septième relations (1258–1612)*, V–XLIV. Paris: Maisonneuve et Leclerc.

Tabak, Josip. 1950. Kako Vinaver prevodi Andersena. *Republika* 5. 316–321.

Trediakovskij, Vasilij K. 1752. *Sočinenïja i perevody kak stïxami tak i prozoju.* Vol. 1, S.-Peterburg.

Tyrwhitt, Thomas. 1782. *A vindication of the appendix to the poems, called Rowley's, in reply to the answers of the Dean of Exeter, Jacob Bryant, Esquire, and a third anonymous writer; with some further observations upon those poems, and an examination of the evidence which has been produced in support of their authenticity*, London: Payne.

Unseth, Pete. 2008. The sociolinguistics of script choice: An introduction. *International Journal of the Sociology of Language* 192. 1–4.

von Sallet, Alfred. 1875. Die Münzen der griechischen Könige von Salamis in Cypern und die denselben zugetheilten modernen Fälschungen. *Zeitschrift für Numismatik* 2. 130–137 + Tafel V.

Zima, Petr. 1974. Digraphia: The case of Hausa. *Linguistics* 124. 57–69.

Biaspectual verbs as polysemes
On homonymy, aspectual neutrality, and the conative reading

The paper starts from the question whether biaspectual verbs like Russian *ispol'zovat'* 'to use (IPFV/PFV)' or *motivirovat'* 'to motivate (IPFV/PFV)' are homonymous pairs of a perfective and an imperfective verb or just one aspectually neutral verb that receives its aspectual meaning from the context. If they are homonyms, they ought to have all the meanings of non-homonymous verbs, including the so-called conative reading of attempt and success, as in *On sdaval*$_{IPFV}$*, no ne sdal*$_{PFV}$ *ėkzamen* 'He took the test but did not pass'. If they are aspectually neutral, sentences given out of the blue without context ought to be read as perfective or as imperfective with equal probabilities. A small corpus analysis and an online survey with 1295 native speakers test these hypotheses. The results show that biaspectual verbs are neither homonyms (since they cannot have the conative reading) nor aspectually neutral (since they all show a clear tendency towards either a perfective or an imperfective interpretation). Consequently, they must be viewed as polysemous, with an α-meaning according to their lexical-actional function and a β-meaning achieved by recategorization where the context demands it.

Translated from "Biaspektuelle Verben als Polyseme: Über Homonymie, Aspektneutralität und die konative Lesart" in *Die Welt der Slaven* 58(1) (2013). 36–53. DOI: 10.18716/bun/2asp.

DOI: 10.18716/omp.64.6

0. What is this about?

This article deals with the so-called biaspectual verbs in Russian,[1] which are traditionally regarded as verbs with homonymous aspectual forms (Section 1). The interpretation in terms of homonymy is contrasted with an alternative analysis, according to which these verbs are aspectually neutral and derive their aspectual meaning only from the context (2). I will try to resolve this question by testing if they can have the so-called conative reading (3), using a corpus analysis (4) and a survey (5). Since the results do not fit either homonymy or aspect neutrality, biaspectual verbs are finally analyzed as polysemes with α- and β-meanings (6). On this basis, an explanation for biaspectuality is then proposed (7), which seems more adequate than previous purely semantic and purely formal explanations.

As is well known, some Russian verbs are an exception to the otherwise all-encompassing aspect opposition in that they do not show any formal distinction

[1] Since the phenomenon also exists in other Slavic languages, it is not impossible that the results of this study can also be transferred to other languages. However, I have not examined them specifically and do not want to assume an identity *a priori*.

for the two verbal aspects. These include, for example, verbs such as *ispol'zovat'*$_{\text{IPFV/PFV}}$ 'use' or *krestit'*$_{\text{IPFV/PFV}}$ 'baptize', but above all many foreign words such as *rekomendovat'*$_{\text{IPFV/PFV}}$ 'recommend', *motivirovat'*$_{\text{IPFV/PFV}}$ 'motivate', *bojkotirovat'*$_{\text{IPFV/PFV}}$ 'boycott', etc. Even verbs that do have a morphologically marked aspect partner 🗋37 can be biaspectual. Examples are *obeščat'*$_{\text{IPFV/PFV}}$ 'promise', which now has a perfective partner *poobeščat'*$_{\text{PFV}}$ but which can still be used in the perfective aspect, or *obrazovat'*$_{\text{IPFV/PFV}}$ 'form', which has an imperfective correlate *obrazovyvat'*$_{\text{IPFV}}$, which, however, has not (yet) completely displaced its derivational base in imperfective use. According to Gorobec (2008: 12), 19% of biaspectual verbs have such aspect partners.[2]

The phenomenon of biaspectual verbs is often dismissed as unimportant. Klein (1995: 671), for example, calls it "atypical […] and not of particular interest", and Zaliznjak & Šmelëv (1997: 65, 2000: 75) also state:

> «Двувидовые глаголы в современном русском языке представляют собою периферийное явление, так как языковая система стремится к тому, чтобы грамматическим различиям соответствовали формальные.»
>
> "In Modern Russian, biaspectual verbs are a peripheral phenomenon, as the language system strives to ensure that grammatical distinctions correspond to formal ones."

Dickey (2000: 11) also includes perfectiva tantum and imperfectiva tantum in his judgment: "unpaired verbs constitute various kinds of peripheral or marginal cases."

In contrast to this, Elena Gorobec (2008) compiled a list of 1,378 (!) biaspectual verbs in Modern Russian, which she checked individually to see whether they still function in both aspects today.[3] This number does not give the impression of a peripheral phenomenon. Admittedly, this total only provides information about dictionary entries (types) and consists partly of technical vocabulary such as *anodirovat'* 'anodize' or *azotirovat'* 'to nitride', so that it says little about the frequency in the text (tokens). However, corpus research has shown that, on average, about every 430th word in the text (or every 75th verb token) is a biaspectual verb.[4]

---

[2] Čertkova & Čang (1998: 17) count 41% verbs with aspect partners, but on the basis of only 412 biaspectual verbs.

[3] She herself gives the total number as 1,394, but 16 verbs (from *pauperizovat'* 'pauperize' to *perekul'tivirovat'* 'cultivate (the ground) again with a cultivator', cf. Gorobec 2008: 193f., 204, 95) are counted double by mistake. The total also contains reflexive formations with *-sja*, some of which are listed as separate lexemes.

[4] In the Russian National Corpus, you can search for verbs that the automatic tagger has tagged as both perfective and imperfective (by searching for the grammatical feature "pf" at a distance of "0" from "ipf" in the non-disambiguated subcorpus). Among the 203,258,919 words (including 35,458,500 verbs) comprising the subcorpus, 1,211,432 hits

Random samples in printed books have 🗋38 confirmed these values: On average, you read a biaspectual verb about every 2–3 pages, and in specialist literature sometimes almost on every page. The phenomenon is therefore not completely negligible.

Furthermore, it is often assumed that the number of biaspectual verbs is shrinking (Zaliznjak & Šmelëv 1997: 62; Breu 2009: 213; cf. Čertkova & Čang 1998: 15) because new aspect partners are constantly being formed and these displace the original verbs in one of the aspects, so that one day there will be no more biaspectual verbs in Russian. However, new biaspectual verbs are constantly being added through borrowing (Zaliznjak & Šmelëv 2000: 71f.; Čertkova & Čang 1998: 14, 31), e.g. *skanirovat'* 'scan', *šrederovat'* 'shred', *dajdžestirovat'* 'summarize into a digest', *dampirovat'* 'put into a dump', etc. (all from Gorobec 2008: 177–204), because "the mode of formation of such verbs, which can express both aspects, is unrestrictedly productive" ("[d]ie Bildungsweise solcher Verben, die beide Aspekte ausdrücken können, ist unbeschränkt produktiv", Isačenko 1962: 354, §204). At the moment, it seems impossible to say which tendency will gain the upper hand in the long term.

The disqualification of biaspectual verbs as a marginal phenomenon is probably also due to the desire to back Slavic verbal aspect as a grammatical category, while the very expression of aspect by devices of word-formation is already atypical of a grammatical category. The concession that aspect is not expressed obligatorily in all verbs without exception would further undermine the foundation of this category. (However, as will be shown below, a closer examination of biaspectual verbs can confirm rather than cast doubt on the all-encompassing validity of the aspect category).

---

were obtained. However, not all of these are biaspectual verbs, as the tagger also assigns both aspect labels to:

- homographs such as *priznaet* (for *priznáet* from *priznat'*$_{\text{PFV}}$ or *priznaët* from *priznavat'*$_{\text{IPFV}}$ 'recognize') or *otrezat'* (for *otrézat'*$_{\text{PFV}}$ or *otrezát'*$_{\text{IPFV}}$ 'cut off')
- lexical homonyms such as *naxodit'sja* ('be situated (IPFV)' or 'walk to exhaustion (PFV)') or *pokupat'* ('buy (IPFV)' or 'bathe for a while (PFV)')
- words not included in the tagger's vocabulary list that it classifies as verbs, such as the proper name *Džaljaj* or the slang verb *zaknokat'*$_{\text{PFV}}$ 'catch (?)'
- forms that are homonymous (or homographic) with biaspectual verbs, such as *kazni* (with the stress *kázni* GEN.SG of *kazn'* 'execution' instead of IMP.SG *kazní* from *kaznit'*$_{\text{IPFV/PFV}}$ 'execute') or *bežat'* ('run (IPFV)' instead of 'flee (IPFV/PFV)')

A check of the first 300 hits showed that approx. 39% of them are actually biaspectual verbs. Extrapolated to the 1.2 million hits of this search, the partial corpus used should therefore contain approx. 470,000 biaspectual verbs.

1. The traditional interpretation: homonymy

The treatment of biaspectual verbs as a case of homonymy is also based on the desire to assign two aspectual forms to all Russian verbs without exception (cf. Isačenko 1962: 355, §204; Maslov 1984: 69f.; Zaliznjak & Šmelëv 2000: 71). Most researchers 🗋39 regard aspect pairs as forms of one and the same lexeme (e.g. Isačenko 1962: 352, §203; Maslov 1984: 88; Lehmann 1999: 227f.; Dickey 2000: 45; Anstatt 2003: 58; Breu 2000: 23, 2009: 210). Biaspectual verbs are then regarded as lexemes with homonymous aspectual forms:

ispol'zovat' (IPFV: *ispol'zuju* etc.; PFV: *ispol'zuju* etc.) 'use'

If aspect pairs are interpreted as two different verb lexemes with the same lexical meaning but different aspects, biaspectual verbs are viewed as lexical homonyms (cf. Švedova 1980: §1388; very consistent also in notation: Gorobec 2008: 45f.):

ispol'zovat'$_1$ 'use (IPFV)'
ispol'zovat'$_2$ 'use (PFV)'

According to this interpretation, biaspectual verbs have two different forms for the perfective and the imperfective aspect, just like aspect pairs, but the two forms 'happen to be' identical.

2. The antithesis: aspectual neutrality

Koschmieder (1934: 11), however, described biaspectual verbs as "aspectless" ("bezaspektowe") or "communia". This approach suggests itself particularly if one considers that the few biaspectual verbs of Slavic origin preserve a state prior to the grammaticalization of the aspectual opposition and that the loanwords owe their biaspectuality to the lack of an aspectual opposition in the source language. Nevertheless, this explanation has been dropped in recent decades in favor of the homonymy thesis. My idea of taking it up in a new form and subjecting it to a tangible test was inspired by Carlota Smith (1997). In her aspect theory, there are three "viewpoint aspects": perfective, imperfective and, additionally, neutral aspect.[5] Among the examples she uses to justify the necessity of the latter is French, where, for example, one can say in the past tense:

(1) *Le cheval gagnait*$_{\text{IPFV}}$ *le course, mais il n'a pas gagné*$_{\text{PFV}}$.
'The horse was winning the race but it did not win.'

In the future tense, however, one cannot say:

[5] As part of her general linguistic study, Smith also deals with Russian and briefly mentions biaspectual verbs there, but without assigning them to the *neutral viewpoint*: "The […] few bi-aspectual verbs […] may serve to express either viewpoint" (Smith 1997: 258).

(2) **Le cheval gagnera*$_{\text{FUT}}$ *le course, mais il ne gagnera*$_{\text{FUT}}$ *pas.*
'The horse will be winning the race but it will not win.' (Smith 1997: 80)

🗋 40 Admittedly, the assumption of a third aspect seems unnecessarily complicated to me. In Smith's case, it is based on universal grammar: A "neutral viewpoint aspect" is also present in languages that have no grammatical aspect opposition at all. It seems more reasonable to me to assume that the French future tense does not participate in the aspect opposition.

However, it seems useful to assume that belonging to an aspect also conveys meanings that an aspectually neutral verb cannot convey (or that would then have to be expressed by other elements in the sentence). And, as can be seen here, this also applies to the imperfective aspect, which is often regarded as an unmarked member of the aspect opposition. Applied to Russian biaspectual verbs, the question would therefore be whether a sentence like (3) is possible with a biaspectual verb or not:

(3) ?*Bioximiki v laboratorii izolirovali*$_{\text{IPFV/PFV}}$ *virus, no ne izolirovali*$_{\text{IPFV/PFV}}$ *ego.*
'The biochemists isolated the virus in the laboratory, but they did not isolate it.'
≟ 'The biochemists tried to isolate the virus, but without success.'

According to the unanimous opinion of grammars and the aspectual literature, such a construction is possible with paired verbs, since the so-called conative reading of the imperfective aspect comes into play here, and Zel'dovič (2009: 287) also considers this to be unproblematic for most biaspectual verbs.

The fact that in (3) the same word form occurs twice in a sentence, making it potentially ambiguous, should not render it ungrammatical, because the same phenomenon can also be observed with lexical homonyms such as in (4) or with homonymous word forms such as the singular and plural of English *fish* in (5), and yet both sentences are undoubtedly grammatical:

(4) *Miru mir!* 'Peace (world) for the world (peace)!'

(5) *The big fish ate the small fish.*

3. The criterion: the conative reading

It is therefore worth taking a closer look at the conative reading. *Conative* (from Latin *cōnātus* 'attempt, effort') is the term for the use of the imperfective aspect of an accomplishment verb to denote an attempt (without saying anything about its success). Some examples of this use that are repeatedly cited are:

(6) *utešal*$_{\text{IPFV}}$*, no ne utešil*$_{\text{PFV}}$
'he comforted$_{\text{IPFV}}$ but did not comfort$_{\text{PFV}}$' (Zaliznjak & Šmelëv 1997: 20)

(7) *sdaval*$_{\text{IPFV}}$*, no ne sdal*$_{\text{PFV}}$ *ėkzamen*
'he took$_{\text{IPFV}}$ the exam but did not pass$_{\text{PFV}}$' (ibid.) 🗋 41

(8) *ubivali*$_{\text{IPFV}}$*, da ne ubili*$_{\text{PFV}}$
'they killed$_{\text{IPFV}}$ but did not kill$_{\text{PFV}}$ (him)' (ibid.; A. N. Tolstoj, Švedova 1980: §1441)

(9)

[Antipa:]	*A začem mundir snjal, službu brosil?*	'[Antipa:] Why did you take off your uniform, quit the service?
[Tarakanov:]	*Ob"jasnjal*$_{\text{IPFV}}$ *ja ėto.*	[Tarakanov:] I have already explained$_{\text{IPFV}}$ this.
[Antipa:]	*Ob"jasnjal*$_{\text{IPFV}}$*, da ne ob"jasnil*$_{\text{PFV}}$.	[Antipa:] Explained$_{\text{IPFV}}$ yes, but not explained$_{\text{PFV}}$.'

(Gor'kij, Švedova 1980: §1441)

If 'biaspectual' verbs really have two homonymous aspectual forms, they should be able to fulfill all the functions of the imperfective and perfective aspect, i.e. they should also allow the conative reading. If they do not allow this reading, at least their 'imperfective forms' are not full-fledged imperfectives. Then biaspectual verbs probably ought to be classified as aspectually neutral, although, as explained above, I would not assume a "neutral aspect" as Smith (1997) does but would suppose that the biaspectual verbs then possibly derive their aspectual meaning exclusively from the context.

4. Corpus study

My first approach to answering this question was to search the Russian National Corpus (*Nacional'nyj korpus russkogo jazyka*, https://ruscorpora.ru/), which currently contains over 200 million words. As a search pattern I used Vi_{IPFV} … *no/a/da* 'but' … *ne* 'not' … Vi_{PFV}, where Vi denotes the same verb twice, entered once in the imperfective and once in the perfective aspect, and "…" represents 0–4 words.

First, 14 classic examples of the conative reading from the literature on verbal aspect were searched for in this way. The only verb pair that occurs more frequently in this construction, namely 12 times in total, is *lovit'/pojmat'* 'hunt/catch' (e.g. *Tri dnja ego lovili*$_{\text{IPFV}}$ *v lesu u derevni, no tak i ne pojmali*$_{\text{PFV}}$*, potom sdalsja sam* 'For three days he was hunted$_{\text{IPFV}}$ in the forest near the village but not caught$_{\text{PFV}}$, then he turned himself in', Mixail Tarkovskij, 2002). The pair *ubivat'/ubit'* 'kill' occurs exactly three times, exclusively in works by A. N. Tolstoj from 1928, which is where the standard example cited in (8) comes from. Similarly, *utešat'/utešit'* 'comfort' (cf. (6)) can be found three times, with the most recent of the three occurrences coming from Gor'kij and dating from 1928. With the aspect pairs *sdavat'/sdat'* 'take/pass (a test)' and *ob"jasnjat'/ob"jasnit'* 'explain' (cf. (7) and (9), as well as with *pisat'/napisat'* 'write' and *ubeždat'/ubedit'* 'convince'), this construction does not occur at all in the Russian National Corpus. Together with a few other examples (2 × *vstrečat'/vstretit'* 'meet, pick up' in Ostrovskij, 1879, and 🗋42 A. Bovin, 1999; 2 × *rešat'/rešit'* 'solve' in *Nezavisimaja Gazeta*, 2003, and L. Zorin, 2008; 1 × *dogonjat'/dognat'* 'catch up' in Furmanov, 1923; 1 × *davat'/ dat'* 'give' in Leskov, 1864; 1 × *ugovarivat'/ugovorit'* 'persuade' in Dostoevskij, 1872; 1 × *postupat'/postupit'*

'enter' in P'ecux, 2001; 1 × *"dokazyvat'"/"dokazat'"* 'prove' (in quotation marks) in S. Mejen, 1987), the search yielded a total of 27 hits (12 of them after 1945).

The same pattern was then used to search for 15 common biaspectual verbs (where $V_{i\text{IPFV}}$ and $V_{i\text{PFV}}$ were of course identical in each case), whose semantics allow for an interpretation of attempt–success (cf. Maslov 1984: 49f.; Zaliznjak & Šmelëv 1997: 20, 2000: 22), namely for *assimilirovat'(sja)* 'assimilate (oneself)', *bežat'* 'flee', *identificirovat'(sja)* 'identify (oneself)', *ispol'zovat'* 'use', *izolirovat'(sja)* 'isolate (oneself)', *kaznit'* 'execute', *likvidirovat'* 'liquidate', *obeščat'* 'promise', *obrazovat'* 'educate', *organizovat'* 'organize', *razminirovat'* 'demine', *rodit'(sja)* 'give birth/be born', *rusificirovat'(sja)* 'russify (oneself)', *velet'* 'command' and *ženit'(sja)* 'marry'. This search yielded no hits.

However, the difference is of course not meaningful. If only about every 75th occurring verb is biaspectual anyway, with 27 occurrences for aspect pairs, only about ⅓ of a hit for biaspectual verbs could be expected statistically. The corpus analysis obviously reaches its limits here.

It has to be noted, however, that the standard examples of the conative reading that are repeatedly cited are surprisingly rare in the corpus. It is therefore astonishing that biaspectual verbs, which are found 470,000 times in the Russian National Corpus, are often dismissed as a peripheral phenomenon, while a sentence pattern that is found 27 times in the same 200-million-word corpus can apparently not be left out of any treatise on Russian verbal aspect (cf. Maslov 1984: 48; Comrie 1976: 47f.; Klein 1995: 692f.; Smith 1997: 238; Zaliznjak & Šmelëv 1997: 20, 2000: 22). The conative reading of the imperfective aspect thus does not seem to have the status in the language that it is accorded in the aspectological literature. At least the sentence pattern $V_{i\text{IPFV}}$ 'but not' $V_{i\text{PFV}}$ is obviously more an artistic exploitation of the aspect category in literature than a part of everyday usage. In the few cases in which the imperfective aspect can actually regularly denote an attempt in everyday life, one can probably assume a lexicalized or idiomatized meaning. This applies in particular to the verbs *lovit'*$_{\text{IPFV}}$ 'hunt' (vs. *pojmat'*$_{\text{PFV}}$ 'catch') and *rešat'*$_{\text{IPFV}}$ 'take care of' (vs. *rešit'*$_{\text{PFV}}$ 'solve'), *lečit'*$_{\text{IPFV}}$ 'treat' (vs. *vylečit'*$_{\text{PFV}}$ 'heal') 43 and the phrase *sdavat'*$_{\text{IPFV}}$ *ėkzamen* 'take an exam' (vs. *sdat'*$_{\text{PFV}}$ *ėkzamen* 'pass an exam').[6]

[6] In *iskat'*$_{\text{IPFV}}$ 'seek' (vs. *najti*$_{\text{PFV}}$ 'find'), the character of an attempt is also clearly recognizable in the lexical meaning, but the difference here is so great that the two verbs do not function as aspectual partners at all: *iskat'* has the perfective partner *poiskat'* 'seek for a while' (cf. Breu 2000: 41f.) and *najti* has the imperfective partner *naxodit'* (since one can find without seeking, and maybe repeatedly or in the historical present tense). The verb *ždat'*$_{\text{IPFV}}$ 'wait' (vs. *dožd'at'sja*$_{\text{PFV}}$ or *dožidat'sja*$_{\text{IPFV}}$ 'expect successfully') behaves similarly. The same situation actually applies to *lovit'*$_{\text{IPFV}}$ and *pojmat'*$_{\text{PFV}}$, but these form an aspect pair according to Maslov's criteria, i.e. *lovit'* can not only mean 'hunt' but also 'catch', e.g. when repeated or in the historical present (Zaliznjak & Šmelëv 1997: 43).

Incidentally, the rarity of this construction in the corpus is not caused by the verbs' semantics. The fact that the verbs tested can certainly be used to express an attempt can be seen from the frequency with which they occur immediately after the verbs *(po)pytat'sja* 'try', *(po)starat'sja* 'make an effort' and *(po)probovat'* 'try': The perfective partners of the 14 aspect pairs tested above all in all occur 3,591 times in these combinations (of which *ob"jasnit'* 'explain' alone occurs 874 times, and *vstretit'* 'meet, pick up' is the rarest with 15 hits). For the 15 biaspectual verbs, the result is even clearer in relation to each other, because after the verbs expressing an attempt, they come to a total of approx.[7] 630 occurrences (of which *ispol'zovat'* 'use' has the most hits with 305, but *obeščat'* 'promise', *rusificirovat'* 'russify' and *razminirovat'* 'demine' are also used once each).

5. Online survey

Since the typical conative construction with biaspectual verbs could not be found in the corpus, an online questionnaire was constructed to find out whether it is considered acceptable by Russian native speakers.

A total of 1,295 respondents took part in the survey, almost all of whom came from Russia.[8] Of these, 65% were male 44 and 30% female (71 people did not specify their gender). 68% had a university degree, 13% of the respondents—relatively few for linguistic surveys—were students and 19% did not have a degree. 18% of the respondents (19% of the graduates and 39% of the students) identified themselves as philologists.

5.1 First part: Conative sample sentences

The first part of the questionnaire consisted of sentences of the familiar type V_{iIPFV}, *no ne* V_{iPFV}, which I constructed with 6 biaspectual verbs. These verbs were selected in such a way that they semantically allowed the interpretation attempt–success and quasi-synonymous sentences with paired verbs could be formed for them. For

[7] In the case of the verb *bežat'*, it was necessary to extrapolate how often the biaspectual verb *bežat'* 'flee' and how often the imperfective verb *bežat'* 'run' was meant.

[8] This large number is due to the fortunate circumstance that a respondent who had learned about the survey via the *vkontakte.ru* network was also one of the more than 35,000 authors of the popular collective blog *dirty.ru* under the pseudonym *charlespetzold* and wrote a short post about the survey there on 28 March 2012 (https://d3.ru/343475/, last checked 22 Mar 2026). He begins his post with the deliberately contradictory filler sentence *Oni nadejalis', čto on peredumaet, no ne nadejalis'* 'They hoped he would change his mind, but they did not hope', which he takes as an allusion to 4 March 2012, the date when Putin was elected president for the third time, and promises that the questionnaire will "potentially […] generate a bunch of funny memes" ("потенциально […] породит кучу забавных мемов").

example, quasi-synonymous sentences with the aspect pair *očiščat'*$_{\text{IPFV}}$*/očistit'*$_{\text{PFV}}$ *ot min* 'clear of mines' could be formed for sentences with the biaspectual verb *razminirovat'* 'demine'. In order to present each subject with each biaspectual verb and each quasi-synonymous aspect pair without juxtaposing them in the same sentence context, the verbs were each placed in two different contexts. In addition, each sentence was available in three different types. The first type consisted of the bare repetition of the same verb:

(10) *Pjat' let nazad na ėtom pole byli švejcarcy, kotorye ego razminirovali*$_{\text{IPFV/PFV}}$*, no ne razminirovali*$_{\text{IPFV/PFV}}$*.*
'Five years ago there were Swiss on this field who demined$_{\text{IPFV/PFV}}$ it, but did not demine$_{\text{IPFV/PFV}}$ it.'

The first half of the sentence was the same for all sentence types. In the second type, however, a clarifying addition was added in the second half:

(11) *… no ne razminirovali*$_{\text{IPFV/PFV}}$ *polnost'ju.*
'… but did not completely demine$_{\text{IPFV/PFV}}$ it.'

Finally, in the third type, the failure of the attempt was paraphrased with a different verb:

(12) *… no ne našli*$_{\text{PFV}}$ *vsex min.*
'… but they did not find$_{\text{PFV}}$ all the mines'.

In this way, a total of 6 × 2 × 2 × 3 = 72 sentences were obtained, 12 of which were presented to each test subject, in a pseudorandomized way so that each biaspectual verb and each aspect pair occurred once each. In addition, there were 24 🗋45 fillers taken from the Russian National Corpus, all of which also contained the words *no … ne* 'but … not' and half of which were manipulated so that they were logically contradictory or grammatically incorrect.

The respondents were asked to rate the sentences with grades between 5 and 1 (which also correspond to Russian school grades and were additionally highlighted with traffic light colors).[9] For the evaluation, these scores were converted into a scale from +1 to –1, as shown in the following table. (Detailed results for the individual sentences can be viewed at http://anketa.buncic.de/.)

[9] There was also a field for optional comments under each sentence, which was completed in 2.7% of the cases. The 423 comments submitted in this way provided information on whether the reasons for the negative evaluation of a sentence were actually to be found in the conative reading or whether individual sentences were perceived as problematic for other reasons. For example, there was a typo in one sentence, in sentences of the type *Palač Džona Li kaznil…* it was criticized that it was not possible to tell whether *Džona Li* was accusative or genitive, etc. To be on the safe side, a total of 13 of the 72 sentences were therefore not included in the statistics.

n ≥ 1731	Biaspectual verbs			Aspect pairs		
	bare	addition	paraphrase	bare	addition	paraphrase
+1.0	3%	6%	**32%**	**30%**	**37%**	**51%**
+0.5	4%	12%	21%	22%	25%	22%
±0.0	8%	17%	17%	19%	16%	14%
–0.5	15%	21%	14%	16%	12%	8%
–1.0	**70%**	**45%**	17%	13%	11%	5%
∅	**–0.72**	**–0.43**	**+0.20**	**+0.20**	**+0.32**	**+0.53**

As can be seen in this summary of the overall results, those sentences in which a biaspectual verb appears in both imperfective and perfective function are, as expected, judged very bad, while the corresponding sentences with aspect pairs show positive mean values. There is also a difference between biaspectual verbs and aspect pairs in the sentences with a paraphase in the second part, albeit a much smaller one.

5.2 Second part: Contextless aspect determination

The second part of the questionnaire is much smaller. Here, the 6 biaspectual verbs from the first part were each embedded in a sentence that gave no indication of which time it referred to. Since the verbs were each in the 'non-preterite' form, their temporal meaning depended on the aspectual interpretation (cf. Čertkova & Čang 🗋46 1998: 13): If the verb is interpreted as imperfective, this form has present tense meaning; if it is interpreted as a perfective, it is future tense. The task was to select one of three paraphrases that disambiguated the original sentence:

(13) *Ėlektrik izoliruet*$_{\text{IPFV/PFV}}$ *kabel' izolentoj.*
'The electrician insulates/is insulating/will insulate the cable with electrical tape …'

a) *Ėlektrik sejčas zanimaetsja*$_{\text{IPFV}}$ *izoljaciej kabelja.*
'The electrician is now dealing$_{\text{IPFV}}$ with the insulation of the cable.'

b) *Izoljacija kabelej javljaetsja*$_{\text{IPFV}}$ *reguljarnoj zadačej ėlektrika.*
'The insulation of cables is$_{\text{IPFV}}$ a regular task of the electrician.'

c) *Ėlektrik skoro pridët*$_{\text{PFV}}$*, čtoby obespečit' izoljaciju kabelja.*
'The electrician will come$_{\text{PFV}}$ soon to secure the insulation of the cable.'

Each respondent was presented with two of these sentences and four fillers, two of which contained an imperfective verb in the present tense and two of which contained a perfective verb in the future tense. The results are shown in the following table (where the answers for the present and habitual present tense are combined into one sum).

n ≥ 391	Present tense (→ IPFV)	Future tense (→ PFV)
kaznit' 'execute'	12%	**88%**
razminirovat' 'demine'	17%	**83%**
identificirovat' 'identify'	**70%**	30%
likvidirovat' 'liquidate'	**80%**	20%
assimilirovat'sja 'assimilate'	**90%**	10%
izolirovat' 'isolate/insulate'	**94%**	6%

A clear dichotomy can be seen here: While for the first two verbs more than four-fifths of the test subjects spontaneously tended towards the future-tense and therefore perfective-aspect interpretation, for the other four verbs more than two-thirds tended towards the present-tense and therefore imperfective-aspect reading. In other words: biaspectual does not equal biaspectual; obviously, some 'biaspectual' verbs are rather more imperfective and others are rather more perfective.

The fact that not all biaspectual verbs are completely biaspectual in all parts of the paradigm is, of course, nothing new. For example, *organizovat'*$_{\text{IPFV/PFV}}$ 'organize' in the past tense can only be used as a perfective verb (Zaliznjak & Šmelëv 1997: 65), *ženit'sja*$_{\text{IPFV/PFV}}$ 'marry' is only imperfective in the plural (ibid. 62), *arestovat'*$_{\text{IPFV/PFV}}$ 'arrest' cannot form a *budu*-future, i.e. in the future tense it is exclusively perfective (Isačenko 1962: 354, §204.1), and *realizovat'*$_{\text{IPFV/PFV}}$ cannot be combined with phase verbs (Čertkova & Čang 1998: 22f.). However, such 'gaps' in biaspectuality have so 🗋47 far only been explained by the existence of aspect partners (here: *organizovyvat'*$_{\text{IPFV}}$, *poženit'sja*$_{\text{PFV}}$, *arestovyvat'*$_{\text{IPFV}}$, and *realizovyvat'*$_{\text{IPFV}}$) that have already displaced the corresponding biaspectual verb in certain uses (and subsequently also penetrate other contexts; in the case of *arestovat'*$_{\text{IPFV/PFV}}$, Gorobec (2008: 53)—three decades after Isačenko—considers the imperfective use already obsolete). Maslov and Zel'dovič go a bit further by assuming that prefixed verbs (Maslov 1984: 87) and verba dicendi (Zel'dovič 2009: 279, 286f.) generally tend towards the perfective aspect.

However, the preference for a particular aspect that can be recognized here is of a different quality, since it is independent of the existence of an aspectual partner or membership in a particular morphological or semantic class and is obviously of a systematic nature.[10] Even though the number of verbs tested is very small, this finding, together with what we have already seen, suggests that the tendency towards one of the two aspects is not a symptom of biaspectuality coming to an end, but is inherent in all biaspectual verbs from the very beginning. The "true biaspectuality" ("подлинная двувидовость") as such, which

[10] Of the six verbs above, only *izolirovat'*$_{\text{IPFV/PFV}}$ has an aspect partner in *zaizolirovat'*$_{\text{PFV}}$ (cf. Gorobec 2008: 228–234), only *razminirovat'* is prefixed, and none is a verbum dicendi (not even *kaznit'* 'execute', cf. Zel'dovič 2009: 289).

Maslov (1984: 87) assumes to be the center of a scale between (rather) perfective and (rather) imperfective verbs, does not seem to exist.

5.3 Comparison of some individual results

The results of the first and second parts will be put into context using two examples. The first involves the verb *kaznit'* 'execute'. One of the two contexts reads *Preljubodejku kaznili*$_{\text{IPFV/PFV}}$*, no ne kaznili*$_{\text{IPFV/PFV}}$ 'The adulteress was executed$_{\text{IPFV/PFV}}$, but not executed$_{\text{IPFV/PFV}}$' or *Preljubodejku zabivali*$_{\text{IPFV}}$ *kamnjami, no ne zabili*$_{\text{PFV}}$ 'The adulteress was stoned$_{\text{IPFV}}$, but not stoned$_{\text{PFV}}$'. As an addition, *do smerti* 'to death' was inserted, the paraphrase was *no vdrug pojavilsja pervosvjaščennik i ostanovil kazn'* 'but suddenly the high priest appeared and stopped the execution'. The following table shows the mean values of the ratings.

n ≥ 211	biaspectual verb	Aspect pair
bare	–0.85	+0.27
addition	–0.69	+0.41
paraphrase	–0.49	+0.59

48 Compared with the overall result given in Section 5.1, this individual result is downright exemplary: the values for the biaspectual verb are clearly negative throughout (even in the sentence with the paraphrase), the values for the aspect pair are clearly positive, the difference between the two values for all three sentence types is approx. 1.1—i.e. more than half of the overall scale. However, the verb tested here was also the one that 88% of the test subjects interpreted out of context as future tense, i.e. as perfective. And if it tends towards the perfective aspect, it is not surprising that it is less acceptable in a special function of the imperfective aspect.

The other example is the verb *izolirovat'* 'isolate/insulate', here in the context already mentioned in (3), *Bioximiki v laboratorii izolirovali*$_{\text{IPFV/PFV}}$ *virus, no ne izolirovali*$_{\text{IPFV/PFV}}$ *ego* 'The biochemists isolated$_{\text{IPFV/PFV}}$ the virus in the lab, but did not isolate$_{\text{IPFV/PFV}}$ it' and *Bioximiki v laboratorii opredeljali*$_{\text{IPFV}}$ *virus, no ne opredelili*$_{\text{PFV}}$ *ego* 'The biochemists determined$_{\text{IPFV}}$ the virus in the lab, but did not determine$_{\text{PFV}}$ it'. As an addition, *im ne udalos'*… 'they did not succeed…' was inserted, and *no bez rezul'tata, tak kak proba okazalas' zagrjaznënnoj* 'but without result, as the sample was contaminated' was used as a paraphrase. The following table shows the mean values again.

n ≥ 209	biaspectual verb	aspect pair
bare	–0.72	+0.36
addition	–0.57	+0.28
paraphrase	+0.17	+0.44

While the two sentences with the repetition of the biaspectual verb are, as always, clearly unacceptable, the rating of the sentence containing the biaspectual verb *izolirovat'* in the first part and the paraphrase in the second part is above zero. The difference to the paired equivalent here is only 0.27—i.e. only a bit more than an eighth of the scale. To determine how significant these differences are, an analysis of variance (ANOVA) was necessary. It showed that even this small difference is still highly significant ($p < 0.001$) but the effect is small ($\eta^2 = 0.03$) due to high variance.[11] It should be borne in mind that the word *izolirovat'* was spontaneously classified as imperfective by 94% of the test subjects. This may 🗋49 explain why the difference to the 'real' imperfective is so small in the sentences with the paraphrase. In contrast to *kaznit'*, the relatively bad rating of the sentences with a repeated biaspectual verb could possibly be due to the repetition of the verb in perfective function in the second part of the sentence rather than the conative function in the first part.

6. Synthesis: polysemy

Notwithstanding the rare use of the conative construction with 'but not', the questionnaire revealed that it is significantly less acceptable for all biaspectual verbs than for quasi-synonymous 'pure' imperfectives. This difference cannot be explained by the common theory according to which biaspectual verbs have homonymous aspectual forms, because there is no indication why the imperfective form of a biaspectual verb should function differently from the imperfective member of an aspectual pair. However, in contrast to Smith's (1997: 80) explanation of the French future tense, aspect neutrality is not the solution either, because Russian biaspectual verbs are, as the second part of the questionnaire has shown, either somewhat perfective or somewhat imperfective, i.e. by no means completely aspectually neutral.

The difference between the more perfective and the more imperfective biaspectual verbs is obviously part of the lexicon entry and should also be taken seriously by linguistics, lexicography and language didactics. Instead of *kaznit' (PFV and IPFV)* and *izolirovat' (PFV and IPFV)*, dictionaries should therefore indicate such verbs as *kaznit' (PFV; also IPFV)* or *izolirovat' (IPFV; also PFV)* or similar.

The concept of α- and β-verbs, which goes back to Lehmann (1993), can therefore also be applied to biaspectual verbs:

[11] The high significance can be explained by the fact that the two mean values are based on 215 and 212 individual ratings, respectively, so that the result is easily replicable. The explained variance is denoted by $\eta^2 = 0.03$, i.e. 97% of the variance (the distribution across the five answer options) is due to random, individual, or in any case different reasons, and only 3% has anything to do with the verb selected. In the sentences with bare verb repetition, this value is still $\eta^2 = 0.45$ (a medium effect).

«Вид одного из партнеров всегда ближе к ЛАФ [лексической акциональной функции — D. B.], в то время как видовые функции другого (или других) основаны на рекатегоризации.»
(VARGOS 2011: 46)

"The aspect of one of the partners is always closer to the LAF [lexical actional function—D. B.], while the aspectual functions of the other (or others) are based on recategorization."

Similarly, one of the aspects of biaspectual verbs is also closer to the lexical actional function (cf. Lehmann 1999: 227–229; VARGOS 2011: 28–29), i.e. to Vendler's "situation type" that the lexical meaning of the verb expresses by default (Dickey 2000: 40–43). This aspect is its α-meaning, the other its β-meaning. Biaspectual verbs that denote processes or states are therefore α-imperfective; verbs that denote events are α-perfective. The β-meaning, i.e. the respective other aspect, is only possible in contexts that already make it sufficiently clear that the 50 α-meaning is not meant. This is because, while in other verbs recategorization is expressed morphologically, in the case of biaspectual verbs it must be triggered by other linguistic means.

Speaking of two different meanings already shows that biaspectual verbs are not a case of homonymy, but of polysemy. They behave similarly to lexical polysemes, such as the word *noga* 'leg', which in its basic meaning denotes a human limb, but in the context of talking about a chair it can also represent a component of that chair. The relationship between the α- and β-meaning of biaspectual verbs is that of metaphor: if an imperfective expression for 'execute' is needed, the perfective expression for 'execute' (the verb *kaznit'*) is used because of its similarity (which is already given because of the identical lexical meaning).

In this respect, biaspectual verbs differ from perfectiva tantum and imperfectiva tantum only in the fact that such metaphorical use is blocked in the latter: In contrast to *kaznit'*$_{\text{IPFV/PFV}}$, a verb such as *zaplakat'*$_{\text{PFV}}$ 'burst into tears' cannot simply fill this function in the rare cases where this meaning is needed in the imperfective aspect (e.g. in habitual use or in the historical present) but must be replaced by a paraphrase such as *načinat'*$_{\text{IPFV}}$ *plakat'*$_{\text{IPFV}}$ 'start to cry'.

The classification of biaspectual verbs as polysemes also explains why even those verbs that tend strongly towards the imperfective aspect allow the conative reading less than 'pure' imperfectives. The full functionality of aspect as a grammatical category only arises from the opposition between different aspectual forms. However, there is no opposition between the different meanings of one and the same polysemous sign (in Saussure's sense).

7. Why are biaspectual verbs biaspectual?

Two main reasons are generally given for biaspectuality: On the one hand, biaspectual verbs are seen as remnants of older language stages or a temporary phe-

nomenon in the borrowing process. The formation of aspectual partners and 'monoaspectualization' (cf. *odnovidovoj*, Maslov 1984: 87) is seen as an element of integration into the Modern Russian grammatical system. On the other hand, morphological obstacles are identified (cf. ibid. 69f.; Gorobec 2008: 52): The unstressed suffix *-ovat'* does not allow suffixation with *-yva-* (cf. *obrazovát'*$_{\text{IPFV/PFV}}$ → *obrazóvyvat'*$_{\text{IPFV}}$ 'form', but *zaímstvovat'*$_{\text{IPFV/PFV}}$ 'borrow' ↛ **zaímstvovyvat'*$_{\text{IPFV}}$), and some lexical prefixes prevent further aspectual prefixation (cf. *kodírovat'*$_{\text{IPFV/PFV}}$ → *zakodírovat'*$_{\text{PFV}}$ 'encode', 🗋51 but *perekodírovat'*$_{\text{IPFV/PFV}}$ 'recode' ↛ **zaperekodírovat'*$_{\text{PFV}}$).[12] However, this does not explain why many unprefixed biaspectual verbs ending in *-ovat'* do not form an aspectual partner by prefixation.

Gennadij Zel'dovič (2009) therefore attempts to find a semantic commonality. Based on the verba dicendi, which he regards as the "prototype of the biaspectual verb" ("прототипом двувидового глагола", ibid. 287), he claims that in all biaspectual verbs the actants must be obligatorily specified, i.e. they can only denote a very specific situation, which partially neutralizes the opposition between uniqueness and multiplicity, since such a specific situation does not usually repeat (ibid. 296). According to Zel'dovič (2009: 315f.), this greater specificity means that biaspectual verbs are 'more binding' than quasi-synonymous paired verbs. Therefore, the following sentences, among others, are better with a 'normal' perfective verb than with a biaspectual one:

(14) *Načal'nik prikazal*$_{\text{PFV}}$*/velel*$_{\text{IPFV/PFV}}$ *prinesti dokumentaciju, no nikto ne pošelox nulsja.*
'The boss commanded$_{\text{PFV}}$/ordered$_{\text{IPFV/PFV}}$ to bring the documentation, but nobody stirred.' (Zel'dovič 2009: 317)

(15) *V delo vložili*$_{\text{PFV}}$*/?investirovali*$_{\text{IPFV/PFV}}$ *million dollarov, no zakončit' stroitel'stvo ne smogli.*
'A million dollars was put$_{\text{PFV}}$/?invested$_{\text{IPFV/PFV}}$ into the project, but the construction work could not be completed.' (ibid.)

(16) *Čas nazad ob"javili*$_{\text{PFV}}$*/??anonsirovali*$_{\text{IPFV/PFV}}$ *novuju peredaču, da tut že i otmenili.*
'An hour ago a new program was advertised$_{\text{PFV}}$/??announced$_{\text{IPFV/PFV}}$ but immediately canceled again.' (ibid. 316)

However, the information on the acceptability of these sentences seems to be based purely on introspection. I therefore also asked about these sentences in the first part of my online survey (with each questionnaire containing two of these sentences and four more fillers). The mean values of the ratings are shown in the following table.

[12] In the online survey, however, some respondents disregarded these restrictions and suggested 'improvements' for the sentences presented, including **assimilirovyvat'*, **razminirovyvat'* and *?dorazminirovat'*.

n ≥ 424	perfective verb	biaspectual verb
(14)	+0.18	+0.27
(15)	+0.91	+0.88
(16)	+0.61	+0.56
∅	**+0.57**	**+0.57**

🗋52 According to the comments made, the low values for sentence (14) are clearly due to the somewhat unusual verb *pošeloxnut'sja* 'stir'. An analysis of variance shows that the differences between the perfective and the biaspectual verb are neither significant nor meaningful ($p > 0.1$; $\eta^2 < 0.01$), both in the individual sentences and overall—despite the large amount of data involved. Zel'dovič's theory therefore appears to have no empirical basis, at least in the sub-area examined.

As far as the biaspectual verba dicendi are concerned, one may indeed assume a semantic reason for the biaspectuality, which has something to do with the performative character of these words. However, other explanations are needed for the entire set of biaspectual verbs.

One approach to such an explanation could lie in the α- and β-meanings of biaspectual verbs. By determining which aspect represents the α- and which the β-meaning, one ought to be able to make predictions about the formation of aspect partners: it can be assumed that aspect partners are only formed for the β-meaning of biaspectual verbs.[13] However, the formation of such a verb may be morphologically blocked if an α-perfective verb ends in unstressed *-ovat'*, which does not allow the formation of an imperfective β-verb with *-yva-*, or if an α-imperfective verb already has a prefix that prevents the formation of a perfective β-verb with another prefix. This could explain why some biaspectual verbs form aspect partners very quickly and become 'monoaspectual' over time, while others steadily remain biaspectual.

This explanation of biaspectuality based on the analysis of biaspectual verbs as polysemes would combine a semantic factor (the lexical actional function of the verb) and a formal factor (a morphological obstacle against the formation of the corresponding β-partner) and would thus do better justice to the nature of biaspectuality than purely formal or purely semantic explanations.

[13] An example also found in VARGOS (see above) is the β-partner *obrazovyvat'*$_{\text{IPFV}}$ of the α-verb *obrazovat'*$_{\text{IPFV/PFV}}$ 'form'. As with aspectual pairs, there are of course also 'diffuse' α/β-verbs (cf. ibid. 38–39). This could explain why some biaspectual verbs at least occasionally form both an imperfective and a perfective aspect partner, such as *organizovat'*$_{\text{IPFV/PFV}}$, for which there is (normatively) the imperfective β-partner *organizovyvat'* as well as (today rather colloquially and not in all meanings) the perfective partner *sorganizovat'* (Čertkova & Čang 1998: 24). Of course, 'anti-iconic' cases in which the α-verb is morphologically marked and the β-partner is unmarked also occur here as well as in paired verbs (Lehmann 1993: 288–293), e.g. *poprivetstvovat'*$_{\text{PFV}}$ vs. *privetstvovat'*$_{\text{IPFV/PFV}}$ 'greet, welcome'.

References

Anstatt, Tanja. 2003. *Aspekt, Argumente und Verbklassen im Russischen.* (Habilitation thesis.) Tübingen. https://d-nb.info/1154308057/34 (last checked 22 Mar 2026).

Breu, Walter. 2000. Zur Position des Slavischen in einer Typologie des Verbalaspekts (Form, Funktion, Ebenenhierarchie und lexikalische Interaktion). In idem (ed.), *Probleme der Interaktion von Lexik und Aspekt (ILA)*, 21–54. Tübingen: Niemeyer.

Breu, Walter. 2009. Verbale Kategorien: Aspekt und Aktionsart. In Kempgen, Sebastian & Kosta, Peter & Berger, Tilman & Gutschmidt, Karl (eds.), *Die slavischen Sprachen: Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung*, vol. 1, 209–225. Berlin, New York: De Gruyter.

Čertkova, Marina Ju. & Čang, Pej-Či. 1998. Ėvoljucija dvuvidovyx glagolov v sovremennom russkom jazyke. *Russian Linguistics* 22. 13–34.

Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems.* Cambridge: Cambridge University Press.

Dickey, Stephen M. 2000. *Parameters of Slavic aspect: A cognitive approach.* Stanford: CSLI.

Gorobec, Elena A. 2008. *Dvuvidovye glagoly v sovremennom russkom jazyke: Problemy statusa i klassifikacii.* (PhD dissertation.) Kazan′.

Isačenko, Aleksandr V. 1962. *Die russische Sprache der Gegenwart. Teil I: Formenlehre.* Halle (Saale): Niemeyer

Klein, Wolfgang. 1995. A time-relational analysis of Russian aspect. *Language* 71(4). 669–695.

Koschmieder, Erwin. 1934. *Nauka o aspektach czasownika polskiego w zarysie: Próba syntezy.* Wilno: Nakład Towarzystwa Przyjaciół Nauk.

Lehmann, Volkmar. 1993. Die russischen Aspekte als gestufte Kategorien (Ein Beispiel für die Bedeutung der kognitiven Linguistik in der slavistischen Sprachwissenschaft). *Die Welt der Slaven* 38(2). 265–297.

Lehmann, Volkmar. 1999. Aspekt. In Jachnow, Helmut (ed.). *Handbuch der sprachwissenschaftlichen Russistik und ihrer Grenzdisziplinen*, 214–242. Wiesbaden: Harrassowitz.

Maslov, Jurij S. 1984. *Očerki po aspektologii.* Leningrad: Izdatel′stvo Leningradskogo Universiteta.

Smith, Carlota S. 1997. *The parameter of aspect.* 2nd edn. Dordrecht: Kluwer. [1st edn. 1991.]

Švedova, Natalija Ju. (ed.). 1980. *Russkaja grammatika.* Moskva: Nauka.

VARGOS = Mende, Julija & Born-Rauxeneker, Eva & Brjugeman, Natal′ja. 2011. *Vid i akcional′nost′ russkogo glagola: Opyt slovarja.* München: Sagner.

Zaliznjak, Anna A. & Šmelëv, Aleksej D. 1997. *Lekcii po russkoj aspektologii.* München: Sagner.

Zaliznjak, Anna A. & Šmelëv, Aleksej D. 2000. *Vvedenie v russkuju aspektologiju.* Moskva: Jazyki russkoj kul′tury. [New, revised edition of Zaliznjak & Šmelëv 1997.]

Zel′dovič, Gennadij. 2009. Počemu russkie glagoly byvajut dvuvidovymi? *Wiener Slawistischer Almanach* 64. 275–330.

Definiteness as a 'covert category' in Russian?

Various phenomena of Russian grammar have so far been explained on the basis of the hypothesis that Russian nouns have the 'covert' category of definiteness. Usually, individual examples are cited to prove that certain phenomena depend on this category. In this paper corpus analyses are used to refute the hypothesis and prove that those grammatical phenomena that are frequently associated with definiteness can be better explained on the basis of other categories. For example, the choice of accusative or genitive case for the direct object under negation depends on whether the object is presupposed or asserted and which elements of the meaning are in the scope of negation. Not a single one of the phenomena inspected in this paper is based on definiteness as it exists in languages with a definite article.

Translated from "Definitheit als 'verborgene Kategorie' im Russischen?" in Pitsch, Hagen (ed.), *Linguistische Beiträge zur Slavistik: XXI. JungslavistInnen-Treffen in Göttingen, 13.–15. September 2012*, 75–96. München: Sagner 2014. DOI: 10.18716/bun/DefR.

DOI: 10.18716/omp.64.7

1. Introduction: 'Covert categories'

With this essay I neither want to nor can I make a further contribution to the difficult question of what definiteness actually is. Therefore, uncharacteristically, it will feature neither formulas of predicate logic nor bald French kings. The much more modest aim of this paper is to critically question whether what is commonly described as 'definiteness' in descriptions of Russian is actually comparable to definiteness in languages with articles.

Since Russian, as is well known, does not have a definite article, there is no obvious formal expression for definiteness. It is therefore often assumed that definiteness in Russian is a 'covert category':

> „Will man den Artikel nicht, wie das im 19. Jh. vorgeschlagen wurde, als „Luxus der Sprache" (LA GRASSERIE 1896) betrachten und will man umgekehrt auch nicht behaupten, daß artikellosen Sprachen – wie etwa dem Russischen – etwas fehle, so hat man von der Konstanz einer grammatischen Funktion auszugehen. Diese Funktion wird in jeder Sprache realisiert, jedoch in unterschiedlichen ‚grammatischen Gewändern'. Kategoriale

> "If one does not want to regard the article as a 'luxury of language', as was suggested in the 19th century (LA GRASSERIE 1896), and conversely, if one does not want to claim either that articleless languages—such as Russian—lack something, then one has to assume the constancy of a grammatical function. This function is realized in every language, but in different 'grammatical guises'. Categorial functions that are realized in com-

> Funktionen, die in komplexen, noch unbekannten und damit noch nicht dokumentierten Mustern realisiert werden, bleiben dem Betrachter in der Regel zunächst verborgen. Sie sind so gut getarnt […], daß der Beobachter das Vorhandensein solcher Regeln nicht einmal vermutet." (Leiss 2000: 3)
>
> plex, as yet unknown and therefore so far undocumented patterns usually remain hidden from the observer at first. They are camouflaged so well […] that the observer does not even suspect the existence of such rules."

The term *covert category*, incidentally, comes from Benjamin Lee Whorf, who first used it publicly in 1936 to describe a category "which has no overt mark actualized along with the words of the class but which operates through an invisible 'central exchange' of linkage bonds in such a way as 🗋76 to determine certain other words which mark the class" (Whorf 1956 [1936]: 69).[1] Whorf calls this effect on other words, which replaces a direct linguistic expression, *reactance*. As a synonymous term for *covert category*, he also suggests *cryptotype*, "a name which calls attention to the rather hidden, cryptic nature of such word-groups" (Whorf 1956 [1937]: 92).[2] The following example given by Whorf (1956 [1936]: 71) demonstrates this cryptic peculiarity well:

> "Another English cryptotype is that of the transitive verbs of a covering, enclosing, and surface-attaching meaning, the reactance of which is that UN- may be prefixed to denote the opposite. Hence we say 'uncover, uncoil, undress, unfasten, unfold, unlock, unroll, untangle, untie, unwind', but not 'unbreak, undry, unhang, unheat, unlift, unmelt, unopen, unpress, unspill'."

Anyone who has the 1996 Diane Warren song "Unbreak my heart" sung by Toni Braxton in their ears will already be taken aback by Whorf's first counterexample. A search in the *Corpus of Contemporary American English (CoCA)* shows that most of the other supposedly impossible words are also used, e.g. "And here was Beatriz […], using a long-handled hook to *unhang* a slinky black dress" (Johnny Payne, "The ambassador's son", 1995) or "She desperately wanted to *unheat* this guy, put him at ease" (Scott Turow, *Personal Injuries*, 1999). The following sentence from a scientific text shows that this is not only done with artistic intent: "It is much easier to restore a newly found mural than to try to *unrestore* and rerestore an old

[1] Whorf here uses grammatical gender in English as an example of a covert category, in contrast to Latin gender, which is an open category. On closer inspection, both statements seem to be problematic.

[2] According to *Google Books*, the term *cryptotype* is found in a 26-page pencilled manuscript by Whorf, dated 25 January 1930 and kept in New Orleans, entitled "'Notes' (on works in Aztec in the Department of Middle American Research, Tulane University of Louisiana)". Unfortunately, the context in which the word is used here is not available online (on the advantages and limitations of *Google Books*, see Bunčić 2012 [English translation in this volume, pp. 55–69—D. B. 2026]).

one" (Mark Sullivan, *Archaeology*, 2003). The "cryptotype" found by Whorf is therefore not really a linguistic, i.e. grammatical or at least lexical, category but a purely semantic classification. As soon as the context indicates that the action denoted by a verb can be undone, the prefix *un-* can also be used.

77 Of course, it is true that a grammatical category does not require a concrete morpheme to express it directly (cf. also Fillmore 1968: 3). With purely semantic definitions in the style of Whorf, however, anything and everything can be classified as a 'category'. Hulanicki (1973), for example, analyzes the general factual meaning of the Russian imperfective aspect as a 'covert perfect'. If the concept of 'covert categories' is to be given any meaning at all, it must be restricted to regular *grammatical* effects of *linguistically* defined classes. An example of such a category could be countability, which has grammatical effects in many languages and which is clearly linguistically, i.e. lexically, and not semantically defined—as can be seen from the fact that Russian *gorox* and *kartofel'* (and several similar words) are uncountable, while their English equivalents *peas* and *potatoes* are countable. (To denote and count individual peas or potatoes, the derivatives *gorošina* and *kartofelina* or *zerno goroxa* 'a grain of peas' and *kluben' kartofelja* 'a bulb of potatoes' are needed in Russian).

2. Definiteness

Sometimes definiteness is understood as a logical and thus universal category which, as a generic term, encompasses the entire continuum from generic to specific to definite expressions in the narrower sense (cf. Daiber 2012: 214–217). Of course, this article does not deny that there is also a logical difference for speakers of Russian between generic and non-generic expressions, for example, and that this difference sometimes manifests itself linguistically. (For example, the use of a demonstrative pronoun such as *ėtot* 'this' indicates that the qualified nominal phrase is not to be interpreted generically.) Rather, the point is to show that definiteness *in the narrower sense*, as a grammatical category marked by the definite article in article languages, is not present in Russian (cf. Birkenmaier's 1979 title *Artikelfunktionen in einer artikellosen Sprache* 'Article functions in an articleless language').

Christopher Lyons (1999) still saw the definite article as expressing two functions at once: on the one hand, it denotes identifiability for the recipient (e.g. in *I bought the car this morning* or *They've just got in from New York. The plane was five hours late*; ibid. 2–3) and on the other hand it marks "inclusivity" (e.g. in *Beware of the dog* or *You are the first visitors to our house*, ibid. 9–10). However, Klaus von Heusinger (1996, 2006) has 'reunited' definiteness by tracing all its effects back to a basic function: a definite marking indicates that reference is made to the *most salient* representative 78 of the category denoted by the NP in a given situation. In many languages, definiteness is marked by a definite article. Indefiniteness is

simply expressed by the absence of this definite article; an 'indefinite article' is not necessary (and in many languages, such as Bulgarian and Macedonian, it does not exist). Where there is such an article, it obviously takes on other functions, e.g. indicating countability. Lyons (1999: 34) therefore calls this article a "cardinal article". Vater (2005: 106–107) is even more explicit, calling it a "quantifier" ("Quantor") and thus placing it in the same category as *some*, *all*, *every*, etc. This observation is important because it makes the assumption of a 'zero article' superfluous.[3]

What is important for our further analysis is the distinction between definiteness and specificity (cf. von Heusinger 2011). Abstracting from the many subtypes of specificity, it is traditionally summarized in such a way that a specific expression refers to a referent 'that the speaker knows and/or refers to'. However, von Heusinger (2011: 1046) points out that specificity can also be based on the perspective of another person who is salient in the given context. In any case, specific expressions refer to referents for which the information necessary for their identification—in contrast to definite expressions—is inaccessible to the hearer.

3. Definiteness in Russian?

At least in German-speaking Slavic Studies, definiteness in Russian has been discussed at least since 1979, when the two monographs *Artikelfunktionen in einer artikellosen Sprache* 'Article functions in an articleless language' by Willy Birkenmaier and *Die Determination des Substantivs im Russischen und Deutschen* 'The determination of the noun in Russian and German' by Wolfgang Gladrow appeared almost simultaneously (the former in West Germany, the latter in East Germany). In a very similar way, using a large number of individual examples from fiction, both authors attempt to show that various mechanisms in Russian fulfill the same functions as definite articles in translations of these text passages (or in the originals from which 🗋79 the Russian examples are translated). However, no comprehensive corpus analysis is carried out. Whether there is also counter-evidence that contradicts the claimed function remains unclear. Moreover, the argument is based exclusively on equivalence in translation, without any closer examination of the phenomena within the language.

Despite the methodological difficulties of these monographs, which, to be sure, were quite up to date 34 years ago, the view that definiteness exists in Russian is

[3] The difference between *Sie ist eine Lehrerin* and *Sie ist Lehrerin* 'She is a teacher' in German therefore has nothing to do with definiteness—both NPs are indefinite due to the absence of the definite article. The assumption of a zero article in languages with definite and 'indefinite' articles produces absurd results, such as the following statement from an Italian–Serbo-Croatian comparison (Stifanić 1980: 56): "Kontrastivna analiza pokazuje da je najveći broj imenica sa nultim članom preveden bez formalnog obeležja za član" ("The contrastive analysis shows that the majority of nouns with zero articles are translated without formal article marking").

still widely held today. For example, Vater (2005: 110) says about "'articleless languages' (such as Latin and the Slavic languages)" ("'Artikellose Sprachen' (wie das Lateinische und die slawischen Sprachen)"):

> „In solchen Sprachen wird offenbar Definitheit besonders im Bereich der Deixis angezeigt, während sie in den anderen Bereichen unmarkiert bleibt oder durch andere Mittel – z. B. durch Wortstellung und Aspektgebrauch – mit bezeichnet wird."
>
> "In such languages, definiteness is apparently indicated especially in the area of deixis, whereas it remains unmarked in the other areas or is co-indicated by other means—e.g. by word order and use of aspect."

As a Slavist, Kempgen (2009: 10) not only knows better that there are also Slavic languages with articles, but is also much more cautious with his wording in other respects:

> „Die Definitheit tritt nur in einem Teil der slavischen Sprachen auf […]. Als ‚Ersatz' dienen den slavischen Sprachen Möglichkeiten unterschiedlicher Wortstellung."
>
> "Definiteness occurs only in some of the Slavic languages […]. As a 'substitute', Slavic languages use the possibilities of different word orders."

Although Kempgen distances himself from the word *substitute* here by using gnomic quotation marks, it should be noted that current Slavic linguistics also assumes a certain functional equivalence between the word order in Russian and the article in German. The assumption of a universal grammar can also lead to the postulate that a category which is expressed by the article in some languages is also "definitely present in the deep structure" in other languages ("in der Tiefenstruktur auf jeden Fall vorhanden", Obst 1981: 19). An even stronger thesis is put forward by Leiss (2000: 14): "Aspect and article are realizations of one and the same grammatical function" ("Aspekt und Artikel sind Realisierungen ein und derselben grammatischen Funktion", see Section 3.2 below). However, the still 'current' Academy Grammar of Russian from 1980 (Švedova et al. 1980: §§2667, 2671, 2672) also refers to "definiteness" ("opredelennost'") to describe the object case of negation (see Section 3.5 below).

Overall, definiteness in the extant literature is claimed to be responsible for a variety of phenomena. These include word order (NP–VP = definite vs. VP–NP = indefinite) and the opposition of nominative (= definite) vs. genitive (= indefinite) in the subject of negated existential clauses. However, if both formal phenomena express the same grammatical category (whether covertly or overtly), they must always 🗋80 behave in the same way in one and the same sentence. Accordingly, only the following two combinations should be possible:

(1) *moroz* *ne* *čuvstvovalsja.*
frost.NOM not was.felt.M
'The frost was not felt.' (definite → NP–VP and NOM)

(2) *Ne čuvstvovalos′ moroza.*
not was.felt.N frost.GEN
'There was no frost to be felt.' (indefinite → VP–NP and GEN)

The following two sentences should accordingly be impossible, since the two expressions for definiteness contradict each other, but the frost can hardly be definite and indefinite at the same time:

(3) *Moroza ne čuvstvovalos′.*
frost.GEN not was.felt.N (NP–VP = definite, but GEN = indefinite)

(4) *Ne čuvstvovalsja moroz.*
not was.felt.M frost.NOM (VP–NP = indefinite, but NOM = definite)

However, as native speakers confirm and corpus research shows, (3) and (4) are completely grammatical and sentences of this type are documented many times. From this we can only conclude that the word order and case in these sentences must fulfill *different* functions. At least one of the grammatical phenomena *does not*, as claimed, express definiteness. Incidentally, Christa Hauenschild (1985: 238), who in the context of machine translation is interested in the coding of referential identity between a noun phrase and a preceding noun phrase, also assumes that different factors can contradict each other, which is why she works with different degrees of probability for reference identity.

In the following, we will therefore examine the functions of the most important grammatical phenomena in Russian that have so far been explained with reference to definiteness or have been suspected of expressing definiteness, with a view to find out for which of them this might be the case and for which it is not.[4]

[4] The selection of phenomena considered is not exhaustive, but to me this selection seems sufficient to invalidate the basic idea that definiteness must also have grammatical functions in an articleless language like Russian. Among other things, I have not dealt with the genitive-accusative alternation in affirmative sentences with verbs of giving, taking and wishing. Schlund (2011) has examined this in detail: Based on a very careful corpus analysis, she manages to severely challenge the traditional view that definiteness is decisive for the choice of case. In her opinion, definiteness only plays a subordinate role here as one of many factors influencing the statistical probabilities for case selection, and even that only as a result of metaphorization, in that the function of the genitive to denote partitivity in mass nouns is transferred to count nouns. However, even this last concession to definiteness can still be questioned. For example, Schlund (2011: 203f.) paraphrases a comment by Levontina (2007) on the sentence *Ona ždët avtobusa* 'She is waiting for the bus' (in which the bus is in the genitive) as follows: "the genitive implies that the person is not waiting for a concrete bus but that she is simply waiting (hoping) for some bus to pass by that would take her closer to her destination." As the wording *a concrete bus* (instead of e.g. *the bus already mentioned*) suggests, this is obviously more about a form of specificity than about definiteness in the sense of identifiability for the addressee. It remains to be inves-

81 Incidentally, Svetlana Friedrich (2009: esp. 36–58, 191–197) also argues against this explanation of many of these phenomena. However, her arguments are mostly based on a terminological misunderstanding (cf. Daiber 2012), a narrow focus on single sentences (which, however, she shares with many of the criticized authors) and on her incorrect counter-hypothesis that definiteness in Russian is expressed by the sentence accent (see Section 3.3). I will therefore not go into the details of her arguments.

3.1 Determiners

Determiners, i.e. demonstrative and indefinite pronouns, are the most obvious expressions that can have something to do with definiteness, because articles, where they exist, have generally developed from them. However, the Russian determiners can be dealt with quickly here, as *ėtot* and *tot* clearly have a deictic function. Although they can also be used anaphorically (just like demonstrative pronouns in languages with articles), they are—in contrast, for example, to their Polish and Czech counterparts—generally unsuspected of being articles (cf. e.g. Berger & Weiss 1987: 15), although of course "due to the absence of the article […], demonstrative pronouns fulfill some of the tasks that are typical of the article in other languages" ("auf Grund des Fehlens des Artikels […] erfüllen die DPr [= Demonstrativpronomina] einen Teil der Aufgaben, die typisch für den Artikel in anderen Sprachen sind", Kordić 2002: 90).

82 The indefinite pronouns could be suspected of being 'negative definite pronouns', but as mentioned above, definiteness is universally marked, not indefiniteness (cf. Lyons 1999: 34). More detailed studies of the determiners *koe-kakoj*, *kakoj-to*, *kakoj-nibud'*, *odin* and *nekij* show that they have other functions, including the marking of specificity (cf. Geist 2010; Gorishneva 2014).

3.2 Aspect

The idea that aspect and article correspond to each other and both express a form of definiteness is, as mentioned above, the central thesis of Leiss (2000).[5] Indeed, Dickey (2000: 19–27) also describes the basic function of the Russian perfective aspect as "temporal definiteness" (in contrast to more western Slavic languages, where the main function is "totality"). Examples of this correspondence can already be found in Birkenmaier (1979: 114):[6]

---

tigated whether the observed statistical effect of definiteness is not in fact only a side-effect of a regularity that is actually based on specificity.

[5] On the lack of a Germanistic foundation for her thesis, cf. the review article by Hans Ulrich Schmid (2003).

[6] Previously, on the basis of a few very specific examples, exactly the opposite relation-

(5) *Vy* *uže* *perevodili* *francuzskie* *stixotvorenija?*
you already translated.IPFV French poems
'Have you ever translated French poems?

(6) *Vy* *uže* *pereveli* *francuzskie* *stixotvorenija?*
you already translated.PFV French poems
'Have you translated the French poems yet?

Indeed, a certain correlation between the temporal definiteness and the definiteness of the object cannot be dismissed out of hand. In (6), the perfective aspect refers to a concrete situation that is clearly delimited in time, and in such a situation it is probable that the poems in question are also very likely concrete poems. However, one can ask oneself how one would express another—admittedly less frequent—situation in Russian, in which, for example, a lecturer has given the (temporally determined) task that all students in a seminar should choose any three (and therefore indefinite) poems from 🗋83 anthologies of French poetry and translate them.

But there is no need for theoretical speculation. A corpus search quickly yields real counterexamples (here, for example, one from Lev Tolstoj's novel *Voskresenie* 'Resurrection' and its English translation, found in the parallel subcorpus of the Russian National Corpus):

(7) *Nexljudov* *perevodil* *slova* *angličanina* […],
Nexljudov translated.IPF words of the Englishman […],
ne vnikaja v smysl ix […].
without paying attention to their meaning […].
"Nekhludoff translated the Englishman's […] words without paying any attention to their meaning […]."

Using the methods of aspectology, it is easy to explain why Tolstoj uses the imperfective aspect here: The sentence gives a more precise description of the way in which a longer-lasting action in process is being carried out. This is the processual meaning of the imperfective aspect. The assertion that the imperfective aspect expresses the indefiniteness of the object does not help here at all, because the object is clearly definite.

For anyone who is seriously interested in the functions of the Russian verbal aspect, this example should be enough to get back to the study of aspectology and away from the idea that temporal definiteness and nominal definiteness are basically one and the same thing.

---

ship had been assumed, namely that the perfective aspect correlates with an indefinite object and the imperfective with a definite object (cf. Obst 1981, 87–92).

3.3 Word order and sentence accent

As is well known, the word order of Russian is not limited by the need to mark case roles as in English, nor by something like the verb-second rule in German, nor by enclitic position rules as in Serbo-Croatian.[7] Instead, it is available almost exclusively for the expression of information structure. Junghanns & Zybatow (2009) have described even more precisely what the Prague linguists tried to summarize under the headings "functional sentence perspective" ("funkční větná perspektiva") or "aktuální větné členění" ("actual division of the sentence") with the terms *theme* and *rheme* by distinguishing between the topic-comment structure and the 🗋84 focus-background structure. Leiss (2000: 6), however, also sees this as a covert expression of definiteness:

> „Es dürfte kein Zufall sein, daß die Vertreter der Funktionalen Satzperspektive Sprecher einer artikellosen slavischen Sprache waren. Was sie mit der Thema-Rhema-Gliederung entdeckt haben, ist in erster Linie die unsichtbare Enkodierung von Definitheit/Indefinitheit in artikellosen Sprachen."
>
> "It is probably no coincidence that the proponents of the Functional Sentence Perspective were speakers of an articleless Slavic language. What they discovered as theme-rheme structure is first and foremost the invisible encoding of definiteness/indefiniteness in articleless languages."

There is no shortage of examples of this connection. Here is one of them (from Birkenmaier 1979: 55):

(8) *Iz* *perednej* *donosilsja* *šëpot.*
from antechamber came whisper
'A whisper came from the antechamber.'

(9) *Šëpot* *donosilsja* *iz* *perednej.*
whisper came from antechamber
'The whisper came from the antechamber.'

So the rule seems to be: NP–VP → definite NP; VP–NP → indefinite NP. It should be noted, however, that this only applies to neutral intonation, where the intonational center is located towards the end of the utterance (cf. also Birkenmaier 1979: 57). In colloquial speech in particular, however, the intonational center is often at the beginning of the utterance. (In a sense, following the stream of thought, the most important new information is conveyed first and then the remaining information necessary for the understanding and grammaticality of the utterance is 'handed in later'; cf. Zemskaja in Zemskaja & Kitajgorodskaja & Širjaev 1981: 33–34).

[7] This does not mean, however, that there are no rigid positional rules at all. For example, prepositions must always immediately precede the noun phrase they govern, *ne* 'not' must immediately precede the verb or the negated constituent, etc.

In this respect, Friedrich (2009: 200) establishes the following alternative rule after excluding several special cases:

> „Liegt keiner der ausgeschlossenen Fälle vor, so richtet sich die Möglichkeit von definiten und indefiniten Bedeutungen – generische sind immer möglich – nur noch nach der Akzentuiertheit des referentiellen Ausdrucks (ohne Satzakzent – definit, mit Satzakzent – indefinit)."
>
> "If we are dealing with none of the excluded cases, the possibility of definite and indefinite meanings—generic ones are always possible—depends only on the accentuation of the referential expression (without sentence accent—definite, with sentence accent—indefinite)."

However, Daiber (2012: 212, 220) notes in his review of this monograph that the author does not distinguish clearly enough between definiteness and information structure. This is related to the fact that the example sentences cited, on which her argument is based and which almost never include the context, can easily be transformed into evidence that contradicts the argument by 🗋85 manipulating the context accordingly. For example, Friedrich (2009: 168) illustrates her "preliminary hypothesis" ("Vorläufige Hypothese") with the following sentence, which she apparently considers to be completely unproblematic:

(10)	*Včera*	*my*	*smotreli*	*novyj*	*fil'm.*
	yesterday	we	saw	new	movie

If the sentence accent is on *fil'm* with neutral intonation and there is no contrasting accent, (10) is, according to Friedrich, to be interpreted as 'Yesterday we watched a new movie'. However, this is by no means mandatory. If there is a preceding cotext, according to which, for example, several possibilities for evening activities were discussed at length, including an exhibition opening, two interesting theater plays, an evening of games and also a new film in the cinemas (which has thus already been mentioned and possibly named by title and described in more detail), and that a decision was finally made, then sentence (10) (with an accent on *fil'm*) can easily be expressed to mean 'Yesterday we watched the new film' (with a definite direct object). This does not require a contrasting accent, nor does the definite meaning have to be expressed by a "definiteness qualifier" ("Definitheitsqualifikator", Friedrich 2009: 179–182) such as *ėtot (novyj fil'm)* 'this (new movie)'. This would merely be a context in which the referent of the noun phrase is already known (definite), but the information about what was watched is new (rhematic, focused). Similarly, many sentences on which Friedrich's argument is based can be dismantled, so that, for example, *pevica podarila cvety vosxiščënnomu zritelju* (ibid. 184f.) can also mean 'The singer gave the flowers to the enthusiastic spectator', etc.

In the literature to date, hardly any counterexamples against the connection between definiteness and word order or sentence accent are discussed. Friedrich (2009: 49–55) also does not provide any really conclusive counterexamples against word order as a definiteness marker, but merely refers to "expressive" sen-

tences in which the sentence accent is on the beginning. Obst (1981: 190) concludes on the basis of a detailed analysis of individual textual evidence that word order is not a reliable definiteness marker:

> „Ob ein Subjekt indefinit ist, läßt sich letzten Endes immer nur auf Grund eines negativen Kriteriums angeben, nämlich auf Grund des Kriteritums, daß […] der außersprachliche Referent der betreffenden Subjekts-NP im gesamten vorhergehenden Text noch nicht vorgekommen ist."
>
> "Whether a subject is indefinite can ultimately only ever be stated on the basis of a negative criterion, namely based on the criterion that […] the extralinguistic referent of the subject NP in question has not yet occurred in the entire preceding text."

86 However, Obst (1981: 74) has previously excluded intonation from his consideration, which makes his analysis vulnerable to Friedrich's (2009) argument.

A more systematic analysis would therefore be desirable here, and so I looked for evidence in corpora. Since it is not possible to search for definite or indefinite noun phrases in a corpus of Russian, I used parallel corpora for this purpose: *ParaSol (Parallel Corpora of Slavic and Other Languages)* and the parallel subcorpus of the Russian National Corpus *(Nacional'nyj korpus russkogo jazyka, NKRJa)*. In them, I searched in article languages e.g. for English *A* or German *Ein(e)* at the beginning of a sentence (i.e. after a period). Three of the examples found in this way (all of which have no expressive intonation and their focus at the end of the sentence) are given below.

The first comes from the novel *Piknik na obočine (Roadside Picnic)* by Arkadij and Boris Strugackij, which can be found in the NKRJa. The sentence we are interested in is preceded by the following context: «Сердце стучало бешено, он поставил портфель на асфальт, торопливо разорвал пачку сигарет, закурил. Он глубоко затянулся, отдыхая, как после драки […]» ("His heart was beating wildly. He put the briefcase on the pavement and impatiently tore open the pack of cigarettes. He lit one, inhaled deeply, and rested, as if after a fight […]"):

(11)	*Dežurnyj polismen*	*ostanovilsja*	*rjadom*	*i*	*sprosil:*
	A cop	stopped	near him	and	asked:

— Vam pomoč', mister?
"Need help, mister?"

This is an example of a sentence with *maximum focus* (cf. Junghanns & Zybatow 2009: 689f.): A new character appears (the policeman) and immediately performs a new, unexpected action (he stops and asks a question). Neither the NP nor the VP is therefore in the background. In the sense of Junghanns & Zybatow (2009: 689), "the entire sentence can thus be understood as an answer to questions of the type *What is going on? What happened?*" ("der gesamte Satz [kann] als Antwort auf Fragen der Art *Was ist los? Was ist passiert?* aufgefasst werden"—and not as an answer to the question *What did the policeman do?* or *Who stopped next to him?*).

The following passage from Bulgakov's *Master i Margarita* found in *ParaSol* contains two sentences in direct succession that contradict the postulated rule. Iešua says shortly after the sentence «Я советовал бы тебе, игемон, оставить на время дворец и погулять пешком» ("I'd advise you, Hegemon, to leave the palace for a while and go for a stroll"): 87

(12) *Groza* *načnëtsja* […] *pozže k večeru.*
A storm will come […] later on, towards evening.

(13) *Progulka* *prinesla by tebe bol'šuju pol'zu* […].
A stroll would do you much good […].

Sentence (12) can again be analyzed as maximum focus. The walk in (13), however, is not in focus, as it has already been mentioned. It forms the topic of the sentence, while the focus is on *bol'šuju pol'zu* 'great benefit'. However, although the walk is mentioned, it is not definite, since it is only suggested and therefore not yet realized and thus not identifiable for the listener (i.e. Pilate).

(14) *Poslyšalis'* *dva* *golosa.* *Bas* *skazal* *bezžalostno:* […]
were.to.hear two voices bass said mercilessly
„Zwei Stimmen wurden laut. Ein Baß sagte erbarmungslos: […]"
'Two voices were raised. A bass said mercilessly: […]'

In this example, again from *Master i Margarita* (this time with a German translation), *bas* 'bass' is topical, as it is a hyponym of the mentioned *golosa* 'voices'. However, this concrete bass voice is not yet known to the reader, not yet identifiable for him or her, so that *bas* is indefinite.

All in all, it can therefore be stated that the information structuring that determines Russian word order in the form of topic-comment structure and focus-background structure has a completely different function than definiteness. The former is primarily a category of text linguistics, the latter a semantic-pragmatic one. Although there are certain statistical correlations between these two categories (e.g. topics are frequently definite NPs), (12)–(14) have shown that they are fundamentally independent categories. In addition, the two categories are expressed quite differently: while the definite article allows binary marking of each NP as definite or indefinite, independent of the further co-text, so that there can be several definite or several indefinite NPs in a sentence, word order must have something on the left and something else on the right, so that it only allows a relative marking of 'newer' information compared to 'older' one.

3.4 Negated existential clauses

88 These last two subsections deal with the genitive of negation, which can occur with subjects and direct objects. In negated existential clauses, the subject can be

in the genitive case. For a start, this is illustrated by a standard example (Apresjan 1985: 292; Partee & Borschev 2007: 148; Späth 2008: 163):

(15) *Otvet* *ne* *prišël.*
answer.NOM not came.M
'The answer did not come.'

(16) *Otveta* *ne* *prišlo.*
Answer.GEN not came.N
'There was no answer.'

Another much-discussed example, which was already cited at the beginning of this article, is repeated here with the interpretations mentioned by Partee & Borschev (2007: 149, 160):

(1′) *Moroz* *ne* *čuvstvovalsja.*
frost.NOM not was.felt.M
'The frost was not felt.' (For example, we were dressed warmly.)

(3′) *Moroza* *ne* *čuvstvovalos′.*
frost.GEN not was.felt.N
'No frost was felt.' (There was no frost.)

Späth (2008: 168) uses precisely this sentence, among others, to demonstrate that the case here by no means directly expresses definiteness:

(17) *Na* *gore* *byl* *moroz.* *No* *moroza* *ne* *čuvstvovalos′.*
on mountain was frost.NOM but frost.GEN not was.felt.N
'There was frost on the mountain. But the frost was not felt.'

Here, the genitive is the natural choice in the second sentence, although the frost introduced immediately before is definite. For further details, refer to the comprehensive account by Späth (2008). He comes to the conclusion that the case change between nominative and genitive not only has nothing to do with definiteness 🗋89 but also cannot be captured by specificity: "Rather, an event is asserted as non-localizable in a given situation" ("Ein Ereignis ist vielmehr in einer gegebenen Situation als nicht lokalisierbar behauptet", ibid. 173).

Hauenschild (1985: 234) concludes with respect to the genitive of negation and the identity of reference between a nominal phrase (NP_2) and a preceding nominal phrase (NP_1) in the text: "If NP_2 as a subject or direct object is in the genitive, identity of reference with NP_1 is relatively unlikely" ("Wenn NP_2 als Subjekt oder direktes Objekt im Genitiv steht, ist Referenzidentität mit NP_1 relativ unwahrscheinlich"). This observation is quite compatible with localizability. If the genitive marked indefiniteness, one would expect that reference identity is not "relatively unlikely", but excluded.

3.5 The case of the direct object under negation

Transitive verbs exhibit a similar competition of cases to existential sentences, but this time it concerns the direct object. In affirmative sentences, the latter is in the accusative case, but under negation it traditionally has to be marked as genitive. In addition to the genitive, it has increasingly become possible to retain the accusative under negation. Since language change is ongoing, the genitive of negation is basically always possible according to the old rules. Therefore, the only relevant question is in which cases the accusative is (now) possible (or even preferable) and in which it is not. Traditionally, definiteness is used to explain this, including e.g. the Academy Grammar of 1980:

> «На выбор падежа оказывает влияние также определенность или неопределенность объекта.» (Švedova et al. 1980: §2667)
>
> "The choice of case is also influenced by the definiteness or indefiniteness of the object."

> "There is, however, a growing tendency in Russian for genitive case in negative contexts to be restricted to indefinites." (Lyons 1999: 201)

This connection can be illustrated by the following sentences (Birnbaum 1970: 23):

(18) *Ja ne vižu devuški.*
I not see girl.GEN.SG
'I do not see a/the girl.' 🗋90

(19) *Ja ne vižu devušku.*
I not see girl.ACC.SG
'I do not see the girl.'

The semantics of the verb obviously also plays a role here: while the retention of the accusative under negation is possible without problems with *verba afficiendi*, it seems questionable with *verba efficiendi*, in which the direct object thus only comes into existence through the activity expressed by the verb, as the following sentences (from Birkenmaier 1979: 139f.) show:

(20) *On ne zametil ošibki.*
he not noticed mistake.GEN.SG
'He did not notice any mistake.'

(21) *On ne sdelal ošibki.*
he not made mistake.GEN.SG
'He did not make any mistake.'

(22) *On ne zametil ošibku.*
he not noticed mistake.ACC.SG
'He did not notice the mistake.'

(23) ?*On ne sdelal ošibku.*
he not made mistake.ACC.SG

Sentence (23) can only be interpreted with a semantic repair by reading *ošibka* 'mistake' not as a concrete mistake, but as a generic term for a certain kind of mistake: 'He did not make this (i.e. this kind of) mistake.'

As with the existential clauses (cf. 3.4 above), the question of the existence of the direct object seems to have something to do with the choice of case: With the *verba efficiendi*, where the genitive of negation is obviously still obligatory, the creation of the object is assertive, i.e. in the scope of negation. With the *verba afficiendi*, on the other hand, which allow the accusative under negation, the existence of the object is presupposed, i.e. not in the scope of negation.

91 However, the following two examples not only show that the accusative under negation is also possible with indefinite objects but also require a revision of the interpretation just presented:

(24) *Počemu vy ne poslali zapisku s sovoj?*
why you not sent message.ACC with owl
'Why didn't you send us a letter by owl?'

(25) *Spasibo skaži, čto ne pišu protokol* […].
Thank_you say.IMP that not I_write report.ACC
'You can thank me for not writing a report […].'

Both (24) (from *Harry Potter* by J. Rowling, found in ParaSol) and (25) (from *Pinočet* by B. Ekimov, found in NKRJa) contain *verba efficiendi*, because *poslat'* 'to send' is not meant here in the sense of *otoslat'* 'to send off', but also encompasses the writing of the letter itself. Consequently, in both sentences, the object does not exist at all due to the negation. Nevertheless, the accusative is used here. Why? The context shows that the speakers assume that the action and therefore the emergence of the letter or the report would have been the 'default case'. In other words, what is presupposed is not the *existence* of the object but a kind of 'defaultness' of this object. The scope of negation is thus perhaps restricted to something like the realization of an assumed plan. The object is already imagined as a very concrete, finished final result of the activity: The teacher in (24) does not intend her question as a vague 'write to me again some time' but knows very precisely what should have been in the letter Harry should have sent her. Similarly, the policeman in (25) knows exactly what the contents of the report should have been. Only the transfer of these objects from their very concrete existence in the imaginary world to the real world is negated.

A similar interpretation was already suggested by Tomson (1903: 218) in an earlier phase of the gradual extension of the accusative to negation:

«Отрицательныя повѣствовательныя предложенія с В[инительнымъ] п[аде-	"Negative declarative sentences with the object in the accusative case occur when

> жомъ] объекта являются тогда, когда отвергается дѣйствіе, съ увѣренностью предполагаемое, ожидаемое или желаемое говорящимъ или слушателемъ.»
>
> an action is negated that is assumed with certainty, expected or desired by the speaker or hearer."

92 However, the many examples Tomson (1903: 218–220) gives of this observation can for the most part also be explained by definiteness, topicality or existential presupposition. The only one of his proofs to which this does not apply, similarly to (24) and (25), is the following (ibid. 219):

(26) *Ona nam obed ne prigotovila.*
she us.DAT lunch.ACC.SG not cooked
'She did not cook lunch for us.'

Unfortunately, I cannot offer a formalization of these regularities. It seems to me that they could be understood as a kind of specificity (the concrete form of the object is accessible to the speaker without the object being identifiable for the addressee). Ueda (1992: 97) summarizes the many factors he analyzed, which seem to have varying degrees of influence on the choice of case, in a sense that could point in this direction:

> "The examples from this section confirm that G[enitive case] is correlated with reference to a general set which consists of members, without presenting any of them as distinct, while A[ccusative case] is correlated with reference to a distinct member of a set."

Perhaps, however, a solution similar to the one proposed by Späth (2008) for existential clauses makes sense. But for the argument made in this paper it is only relevant that the choice of the object case under negation is determined by categories other than definiteness in the sense defined in Section 2.[8]

4. Conclusion

Russian determiners mark deixis, specificity, and similar categories; verbal aspect expresses temporal definiteness, which is different from nominal 93 definiteness; word order expresses information structure, which, unlike definiteness, is a text-linguistic function; the opposition between nominative and genitive in negated

[8] Since the displacement of the genitive of negation is a linguistic change that is currently underway, there is of course still a lot in flux here, which makes it difficult to recognize an emerging system. Moreover, a case form may be avoided in some places because it leads to an ambiguity in number. In both the inanimate feminines and the inanimate neuters, the endings of the GEN.SG and ACC.PL have merged (the forms are therefore homonymous, or homographic with different accents), which in this case, where in principle both a genitive and an accusative are possible, can lead to ambiguity about the number. Sometimes the accusative singular or the genitive plural is chosen to avoid this ambiguity.

existential sentences is based on the localizability of the referent in a given situation; and the opposition between accusative and genitive for the direct object in negated sentences seems to have something to do with presupposition and the scope of negation. All five of the phenomena examined here thus turned out not to be based on definiteness, but to have a better explanation based on other categories. Consequently, the category of definiteness is not necessary to describe the Russian language system. Therefore, according to the principle of economy, it can be assumed that definiteness is not a category of Russian.

The existence of definiteness in article languages, where there is a morpheme that seems to serve only this one purpose, apparently obscures the wide variety of options available to languages to narrow down the set of possible referents that an NP can refer to: In addition to the definite vs. indefinite opposition, there are also specific vs. non-specific NPs (with various subtypes, e.g. scopally specific or epistemically specific), presupposed vs. asserted, background vs. focus, topic vs. comment, existence vs. location, and many others. The expression of each of these oppositions contributes equally to understanding by giving the recipient clues as to how to select the referent meant in the given situation from the set of possible referents of an NP. A language that does not express definiteness (but perhaps specificity or elements of information structure) therefore lacks nothing. *The Russian language is complete without definiteness as well.*

The communicative function that all these categories have in common, namely ensuring that the receiver understands which real-world referent the sender means, is so fundamental that it is difficult to speak of equivalence based on this common function. The individual pragmatic contents expressed—such as 'I have not yet mentioned the referent of this NP', 'this information is new to you', 'this NP refers to any representative of its extension', etc.—are so different that their intersections are relevant only as statistical quantities. Hastily equating these functions will not increase our knowledge about them.

Since definiteness is often regarded as a 'covert category' in Russian and it now turns out that this 'covert category' does not exist, the question arises whether the concept of 'covert categories' makes sense at all. Perhaps it would be more appropriate to speak only of grammatical categories in general. If a category is part of a language system, it should also be recognizable in 94 morphological or syntactic regularities. From a linguistic point of view, assuming a category that cannot be detected using scientific methods—which also include examining the entire linguistic material, including possible counterexamples and the search for alternative explanations—because it is supposedly so well 'hidden' makes no sense.

References

Apresjan, Jurij D. 1985. Sintaksičeskie priznaki leksem. *Russian Linguistics* 9. 289–317.

Berger, Tilman & Weiss, Daniel. 1987. Die Gebrauchsbedingungen des Anaphorikums *tot* in substantivischer Verwendung. In Freidhof, Gerd et al. (eds.), *Slavistische Linguistik 1986: Referate des 13. Konstanzer Slavistischen Arbeitstreffens (16.–19. 9. 1986)*, 9–93. München: Sagner.

Birkenmaier, Willy. 1979. *Artikelfunktionen in einer artikellosen Sprache: Studien zur nominalen Determination im Russischen.* München: Fink.

Birnbaum, Henrik. 1970. *Problems of typological and genetic linguistics viewed in a generative framework.* The Hague, Paris: Mouton.

Bunčić, Daniel. 2012. Über den Nutzen von Google Books & Co. für (nicht nur slavistische) Begriffsgeschichten. In Podtergera, Irina (ed.), *Schnittpunkt Slavistik: Ost und West im wissenschaftlichen Dialog, Festgabe für Helmut Keipert zum 70. Geburtstag*, vol. I, 393–407. Göttingen: V&R unipress. [English version in this volume, pp. 55–69.]

Daiber, Thomas. 2012. [Review of Friedrich 2009.] *Kritikon litterarum* 39(3/4). 212–221.

Dickey, Stephen M. 2000. *Parameters of Slavic aspect: A cognitive approach.* Stanford (CA): CSLI Publications.

Fillmore, Charles J. 1968. The case for case. In: Bach, Emmon et al. (eds.), *Universals in linguistic theory*, 1–88. London, New York: Holt, Rinehart & Winston.

Friedrich, Svetlana. 2009. *Definitheit im Russischen.* Frankfurt am Main: Peter Lang.

Geist, Ljudmila. 2010. Indefinitpronomina im Russischen und Spezifizität. *Zeitschrift für Slawistik* 55(2). 206–221.

Gladrow, Wolfgang. 1979. *Die Determination des Substantivs im Russischen und Deutschen.* Leipzig: Enzyklopädie.

Gorishneva, Elena. 2014. *Odin* in der restriktiven Funktion. In Pitsch, Hagen (ed.), *Linguistische Beiträge zur Slavistik: XXI. JungslavistInnen-Treffen in Göttingen, 13.–15. September 2012*, 97–118. München: Sagner.

Hauenschild, Christa. 1985. *Zur Interpretation russischer Nominalgruppen: Anaphorische Bezüge und thematische Strukturen im Satz und im Text.* München: Sagner.

Hulanicki, Leo. 1973. The actional perfect in Russian. *Slavic and East European Journal* 17(2). 174–183.

Junghanns, Uwe & Zybatow, Gerhild. 2009. Grammatik und Informationsstruktur. In Kempgen, Sebastian et al. (eds.), *Die slavischen Sprachen: Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung*, vol. 1, 684–707. Berlin, New York: de Gruyter.

Kempgen, Sebastian. 2009. Phonetik, Phonologie, Orthographie, Flexionsmorphologie. In Kempgen, Sebastian et al. (eds.), *Die slavischen Sprachen: Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung*, vol. 1, 1–14. Berlin, New York: de Gruyter.

Kordić, Snježana. 2002. Demonstrativpronomina in den slavischen Sprachen. In Symanzik, Bernhard & Birkfellner, Gerhard & Sproede, Alfred (eds.), *Die Übersetzung als Problem sprach- und literaturwissenschaftlicher Forschung in Slavistik und Baltistik: Beiträge zu einem Symposium in Münster, 10./11. Mai 2001*, 89–116. Hamburg: Dr. Kovač.

Leiss, Elisabeth. 2000. *Artikel und Aspekt: Die grammatischen Muster von Definitheit.* Berlin, New York: de Gruyter.

Levontina, Irina. 2007. "A on, mjatežnyj, prosit buri": Kaverznye padeži [Interview with Radio Svoboda by Lilja Pal'veleva, 8 Aug 2007]. https://www.svoboda.org/a/406097.html (last checked 22 Mar 2026).

Lyons, Christopher. 1999. *Definiteness.* Cambridge: Cambridge University Press.

Obst, Ulrich. [1981.] *Studien zur zweidimensionalen syntagmatischen Substitution in modernen Prosatexten des Russischen.* Münster: Aschendorff.

Partee, Barbara & Borschev, Vladimir. 2007. Existential sentences, *be*, and the genitive of negation in Russian. In Comorovski, Ileana et al. (eds.), *Existence: Semantics and syntax*, 147–190. Dordrecht: Springer.

Schlund, Katrin. 2011. Genitive vs. accusative as a direct object in affirmative sentences in Russian. In Karl, Katrin Bente & Krumbolz, Gertje & Lazar, Marija (eds.), *Beiträge der Europäischen Slavistischen Linguistik (POLYSLAV) 14*, 201–210. München: Sagner.

Schmid, Hans Ulrich. 2003. Linguistische Theorie, philologische Empirie und (grammatische) Kompetenz [Review of Leiss 2000]. *Zeitschrift für Dialektologie und Linguistik* 70(3). 298–313.

Späth, Andreas. 2008. Lokalisierung von Situationen unter Satznegation im Russischen. In Geist, Ludmila et al. (eds.), *Linguistische Beiträge zur Slavistik: XIV. JungslavistInnen-Treffen in Stuttgart, 15.–18. September 2005*, 163–175. München: Sagner.

Stifanić, Marija. 1980. Mogućnosti iskazivanja vrednosti italijanskog člana u srpskohrvatskom jeziku. In *Studije iz kontrastivne analize italijanskog i srpskohrvatskog jezika*, vol. 1, 37–61. Beograd: Filološki Fakultet Beogradskog Univerziteta.

Švedova, Natalija Ju. et al. (eds.). 1980. *Russkaja grammatika.* Vol. II (Sintaksis). Moskva: Nauka.

Tomson, Aleksandr I. 1903. Vinitel'nyj padež prjamogo dopolnenija v otricatel'nyx predloženijax v russkom jazyke. *Russkij filologičeskij vestnik* 49(1–2). 192–234.

Ueda, Masako. 1992. *The interaction between clause-level parameters and context in Russian morphosyntax: Genitive of negation and predicate adjectives.* München: Sagner.

Vater, Heinz. 2005. *Referenz-Linguistik.* München: Sagner.

von Heusinger, Klaus. 1996. Definite Kennzeichnungen, Anaphora und Salienz. *Linguistische Berichte* 163. 197–226.

von Heusinger, Klaus. 2006. Salience and anaphoric definite noun phrases. *Acta Linguistica Hafniensia* 38. 33–53.

von Heusinger, Klaus. 2011. Specificity. In von Heusinger, Klaus et al. (eds.), *Semantics: An international handbook of natural language meaning*, vol. 2, 1024–1057. Berlin, New York: de Gruyter.

Whorf, Benjamin Lee. 1956. *Language, thought and reality.* (Edited by J. B. Carroll.) New York: Wiley.

Zemskaja, Elena A. & Kitajgorodskaja, Margarita V. & Širjaev, Evgenij N. 1981. *Russkaja razgovornaja reč': Obščie voprosy, slovoobrazovanie, sintaksis.* Moskva: Nauka.

‘Diastratic diglossia’ in 18th-century Russia, *or:* When did Church Slavonic become a foreign language?

Boris Uspenskij assumes that the “Third South Slavic Influence”, which began in the 1640s, resulted in the speedy decline of diglossia in Russia and its replacement with bilingualism. This brought about a change of status of Church Slavonic: While in diglossia H and L serve as varieties of one language, in bilingualism Church Slavonic is a foreign language for Russian speakers. This paper, however, provides proof of the fact that Church Slavonic remained a variety of the Russian language with high prestige until the 1760s or 1780s. Among other things, this follows from Lomonosov’s inclusion of Church Slavonic into his Theory of the Three Styles. Consequently, the second half of the 17th and the first half of the 18th century have to be regarded as a transitional period, during which the essential features of diglossia were preserved, but the distribution of H and L gradually turned from a strictly diaphasic model into a predominantly diastratic one. We call this kind of situation ‘diastratic diglossia’.

Translated from “‘Diastratische Diglossie’ im Russland des 18. Jahrhunderts, *oder:* Wann wurde Kirchenslavisch zur Fremdsprache?” in Dieser, Elena (ed.), *Linguistische Beiträge zur Slavistik: XX. JungslavistInnen-Treffen in Würzburg, 22.-24. September 2011*, 29–45. München: BiblionMedia 2015. DOI: 10.18716/bun/digl.

DOI: 10.18716/omp.64.8

1. Introduction

This paper was originally written as part of my postdoctoral thesis on biscriptality. Since biscriptality is defined as the simultaneous use of two writing systems, script variants, or orthographies for the same language (or varieties of the same language), the linguistic situation in 18th-century Russia poses a problem. It is well known that Tsar Peter I introduced the so-called *graždanskij šrift* ‘civil type’ in 1708, without completely replacing the traditional Old Cyrillic script. The latter remained in use as a typeface for ecclesiastical texts—and even to this day for certain church books used in Orthodox worship (cf., e.g., the acolouthia for the martyrs of the Soviet era, published in 1992, excerpts of which are printed in Trunte 2014 [1998]: §30.7c). From today’s perspective, however, some will argue that this cannot be considered biscriptality in the sense of the above definition because the Old Cyrillic script has only been used for Church Slavonic since the alphabet reform, while texts in the Russian language have since been printed ex-

clusively in *graždanskij šrift*. Thus, the two script variants would be used for different languages, so that both Russian and Church Slavonic would be monoscriptal.

Although closer examination shows that the distribution of scripts among the language forms in the 18th century was by no means so clear-cut, the question arises: Were Russian and Church Slavonic actually already different languages at the beginning of the 18th century? Since when have they been different languages, and how can this be determined? This paper addresses this complex of questions.

2. Questions of distance

30 How can we determine whether two varieties belong to the same language or to different ones? The first approach is, of course, to determine the linguistic distance, i.e. the degree of similarity and difference between the varieties. It is intuitively clear that two varieties with only minimal differences (e.g. British and American English) can hardly be considered different languages, while completely different varieties (e.g. English and Hungarian) can under no circumstances be regarded as varieties of one language. There is extensive specialist literature on the question of how this distance can be measured (e.g. Ammon 1989, Casad 1992). The fundamental problem is to weigh up the facts of different language levels (e.g. phonetics, grammar, vocabulary) against each other and thereby obtain a measure that reflects as accurately as possible what is commonly described as "intelligibility". In contrast to distance, which, at least in theory, can actually be expressed in objective numbers, intelligibility is *a priori* a subjective variable that depends to a large extent on the individuals involved, their experience with texts in the respective language, their basic motivation to understand these texts, and, last but not least, the topics of the texts used. Furthermore, intelligibility is not symmetrical, as it is possible that speakers of one language find it easier to understand speakers of another language than vice versa (a well-known example is that Danes apparently understand Swedish more easily than Swedes tend to understand Danish). Therefore, if the answer to the question of the linguistic status of a variety is to reflect more than just the attitudes of the speakers concerned, there is no way around measuring the distance.

Ammon (1995: 6) proposes a simple test which, although it does not provide a concrete figure for the distance between two languages, allows a distinction to be made between high, medium, and low similarity—which is entirely sufficient for our purposes—and, because it is based on the comparison of texts, largely eliminates the need to weigh up different language levels.[1] For this test, "texts

[1] The question of how to weigh up graphic/orthographic and phonetic/phonological differences, however, is not entirely clear. Since written texts are usually compared with each other, one is sometimes confronted with orthographic differences that have no effect

with the same meaning, translated 🗋31 as literally as possible" ("sinngleiche, möglichst wörtlich übersetzte Texte", ibid.) are compared. A high degree of similarity (i.e. a small distance) between the varieties can be said to exist if more than half of the words are completely identical, and a low degree of similarity (i.e. a large distance) if more than half of the words are completely different.

> „Bei mittlerer Ähnlichkeit zwischen zwei Varietäten ist […] die Zahl der divergierenden Wörter zwar größer als die Zahl der übereinstimmenden Wörter; bei der Mehrzahl der paarweise verglichenen Wörter ist jedoch die Ähnlichkeit so groß, daß sie leicht als einander entsprechend erkannt werden können." (Ammon 1995: 6)

> "In the case of medium similarity between two varieties, […] the number of diverging words is greater than the number of identical words, but for the majority of the words, when compared in pairs, the similarity is so great that they can easily be recognized as corresponding to each other."

This test does not require the evaluation of massive parallel corpora; even with very short texts, it is usually quickly clear which category a pair of varieties belongs to. With respect to historical language states, however, we encounter the additional problem that it is no longer possible to ask a native speaker to produce a translation "as literally as possible". However, existing parallel texts can be used as a workaround and adapted slightly if necessary. To illustrate this, here is a small excerpt from an exercise book by Mixail Ivanov from 1726–1728, quoted by Uspenskij (2002: 509, §19.2):

> **Russian:** Я видѣлъ малчика, которои стоячи в' цр҃кви осудил чл҃ка, немного погодя и сам осудился от ныхъ [sic! — Б. У.] людей. Я ето смотрячи молвил ему: вот ежели бы ты не осудил то бы и сам не_осужденъ былъ.

> **Church Slavonic:** Азъ видѣхъ отрока, иже в' цр҃кви осуди чл҃ка. Послѣди и сам осужденъ бысть от иных чл҃къ. Аз сіе зрящи рекъ ему: аще бы ты не осудилъ еси не бы и сам осужденъ бысть [sic! — D. B.].

All deviations are underlined. For comparison, let us look at the Russian translation of the Gospels from 1818, which was not completed until the beginning of the 19th century but is characterized by a very literal and yet modern translation (here Matthew 7:24–25 is quoted from the third edition: Rossijskoe Biblejskoe Obščestvo 1820: 21):[2] 🗋32

on pronunciation, or with orthographically identical words that are pronounced very differently. However, these problems are unlikely to be of any significance for Church Slavonic and 18th-century Russian.

[2] The facsimile comes from *Google Books* (http://books.google.de/books?id=pTs7AAAAcAAJ&pg=PA21, accessed 24 April 2012).

к҃д. Всѧкъ ѹ̀бо, и́же слышитъ словеса̀ моѧ̑ сїѧ̑, и твори́тъ ѧ́, уподо́блю єго̀ мꙋ́жꙋ мꙋ́дрꙋ, и́же созда̀ хра́минꙋ свою̀ на ка́мени:	24. И такъ всякаго, кто слушаетъ сіи Мои слова, и исполняетъ ихъ, уподоблю мужу благоразумному, который построилъ свой домъ на камнѣ:
к҃є. И сни́де дождь, и прїидо́ша рѣ́ки, и возвѣ́ѧша вѣ́три, и нападо́ша на хра́минꙋ тꙋ́, и не паде́сѧ:	25. И пошелъ дождь, и разлились рѣки, и подули вѣтры, и устремились на домъ тотъ, и онъ не упалъ;

As can be seen, between about one-third (13 out of 38 words in the Russian Bible text) and just under half of the words (in the exercise book) match exactly, and for the vast majority of words the correspondence is obvious. The test therefore shows a *medium* degree of similarity between Church Slavonic and the Russian vernacular.

While a high degree of similarity between two varieties means that they automatically belong to the same language, and a low degree of similarity means that they are different (abstand) languages *per se*, the classification in cases of medium similarity depends on sociolinguistic criteria such as standardization: It is precisely this medium distance that allows a dialect to become a standard language through standardization (for a graphical representation of this relationship, see Bunčić 2008: 91 [English translation: p. 41 in this volume—D. B. 2026]). For our question about the status of Church Slavonic, this means that its medium distance from the Russian vernacular allows for both answers. Therefore, what is decisive is the sociolinguistic function of Church Slavonic and its relationship to Russian.

3. Diglossia vs. bilingualism

Uspenskij (1987, 2002) describes the sociolinguistic situation of the East Slavs in the Middle Ages (or, in Western terminology, also in the early modern period) as *diglossia* in the sense of Ferguson (1959; on the history of the transfer of this concept to the East Slavic language situation, see Rehder 1995: 362–367). According to this, the East Slavic—and later Russian—variety of Church Slavonic (traditionally called *recension*, *redaction* or, in Russian, *izvod*) functioned as the H variety and the East Slavic 'vernacular' as the L variety.[3] In Russia, according to Uspenskij

[3] Uspenskij (2002: 24–26, §2.2) calls H *knižnyj jazyk* 'bookish language' or *literaturnyj jazyk* 'literary language, standard language' and L *nekniżnyj jazyk* 'non-bookish language' or *razgovornyj jazyk* 'colloquial language', which is not always appropriate when one considers, for example, that books such as *Russkaja pravda* were also written in Old East Slavic and that Ferguson's concept is fundamentally open to different distributions of functions. I will therefore use the more neutral terms *H* and *L* here.

🗋33 (2002: 418, §16.3), the so-called Third South Slavic Influence,[4] which began in the 1640s, led to the dissolution of diglossia and its replacement with bilingualism. Gradually, Russian took over almost all the functions of the H variety, while at the same time Church Slavonic now also came to be used for secular purposes that had previously been reserved for the L variety, e.g. for everyday conversations or private letters (Uspenskij 2002: 477–485, §18.2; in the Ruthenian, i.e. modern-day Ukrainian and Belarusian, areas, diglossia had already been terminated before that, cf. ibid. 386–408, §15). This change led to competition between the varieties within both 'high' and 'low' text types, which is best illustrated by the emergence of parallel texts in both varieties (Uspenskij 2002: 474, §18.1). However, this does not mean that Church Slavonic lost its high prestige:

> «[П]рименение языка определяется при этом не отношением к предмету речи (как это имело бы место ранее), но исключительно уровнем образования пишущего — владение церковнославянским языком демонстрирует ученость, принадлежность к элитарной культуре.» (Uspenskij 2002: 482, §18.2)
>
> "In this, the choice of language is not determined with respect to the subject of the text (as would have been the case previously) but exclusively by the writer's level of education; mastery of the Church Slavonic language demonstrates erudition and belonging to the elite culture."

After 1640, Church Slavonic and the Russian vernacular did not lose their status as 'high' and 'low' varieties. It was not at all the case at this point in time that the two varieties "function more or less in parallel" ("функционируют более или менее параллельно"), as Uspenskij (2002: 24, §2.2) describes 🗋34 their relation in bilingualism. All that changed were the factors determining the choice of these varieties. In Gregory's (1967) terminology, the previous model can be described as purely *use-oriented* and the subsequent one as partially *user-oriented*, while in Coseriu's (1992: 280–292) terminology, the distribution of varieties had previously been *diaphasic*[5] and now became mainly *diastratic*. Ferguson (1959), however, does

[4] This widely used term emphasizes similarities between the orthographic symptoms of this development and those of the Second South Slavic Influence in the 14th century. However, the "Third South Slavic Influence" had nothing to do with the South Slavs but was in fact an influence of Ruthenian language practice on Russian language usage (cf. Uspenskij 2002: 411–412, §16.1).

[5] Unfortunately, the term *diaphasic* is used with different meanings. When Coseriu introduced it into linguistic usage, he used it to refer to different stylistic levels, which fits the basic meaning of Greek *φάσις* 'expression' (Coseriu 1992: 283). In later works, however, other authors also include factors relating to the overall communication situation (cf. Ammon 2000). Still others use the term with reference to the current meaning of the word *phase*, so that it either becomes very similar to *diachronic* (cf. ibid.) or refers to differences between different age cohorts (e.g. Goossens 1977: 9–10). In this paper, the term is used

not envisage *user-oriented diglossia*. Both his four model cases and his description refer exclusively to diaphasic, and sometimes also to medial and diamesic[6] differences, which can be summarized as *use-oriented*. In the ideal case of diglossia, H and L not only apply to an entire language community, but each individual also masters both varieties, so that they can use them depending on the communication situation.[7] Therefore, Uspenskij concludes from the change in the Russian language situation described above that diglossia had been abolished and calls the new situation *bilingualism*.

🗋35 In addition, Uspenskij repeatedly compares the new sociolinguistic situation in Russia with the language situation in Western Europe:

> «Соответственно, церковнославянский язык предстает как язык ученого сословия, т. е. приобретает функции, свойственные латыни на Западе, и становится вообще функциональным эквивалентом латыни.» (Uspenskij 2002: 479, §18.2)
>
> "Accordingly, Church Slavonic appears as the language of the educated class, i.e. it acquires functions characteristic of Latin in the West and, in general, becomes a functional equivalent of Latin."

It is precisely this situation in Western Europe, in which Latin, as the language of education and, in some cases, the official language, roofed various vernacular languages that were not closely related to it, that Auer (2005: 9–12) describes as *exoglossic diglossia*, because Latin functionally assumes exactly the role that the H variety plays in diglossia and the situation differs from Ferguson's diglossia only in that the H variety is not closely related (endoglossic) to the respective L variety.

Putting this information together, the following picture emerges: in the second half of the 17th century, the linguistic situation in Russia changed from endoglossic diglossia to a kind of bilingualism that is functionally comparable to exoglossic diglossia—but without abandoning the endoglossic H variety, Church Slavonic. Obviously, this does not add up.

---

exclusively in Coseriu's original meaning, i.e. in relation to language styles and the text types in which they are realized.

[6] This practical term, coined in Italian linguistics (cf. Kabatek 2000: 313f.), is used here to describe the difference between the *language of proximity* and the *language of distance* as defined by Koch & Oesterreicher (1985).

[7] Ferguson does not address the fact that the school education necessary for acquiring active competence in the H variety is not accessible to all sections of society and that, as a result, even in his four model cases, some speakers are/were unable to produce acceptable Standard German, Standard French, High Arabic, or Katharevousa. However, this is not specific to diglossia: Even in a 'normal' standard-dialect continuum (in a *diaglossia*, cf. Auer 2005: 22–23), there are always speakers who do not have a sufficient command of the standard language and need help, for example, when writing job applications, so as not to be rejected from the outset because of dialect features and poor spelling.

The central problem with this description is, once again, that Uspenskij (2002: 24, §2.2) assumes that the decisive criterion for distinguishing between diglossia and bilingualism is that "in the case of diglossia, the functions of the two coexisting languages are in complementary distribution" («В случае диглоссии функции двух сосуществующих языков находятся в дополнительном распределении»). However, as Rehder (1989: 368) rightly points out, a strict complementary distribution is not part of Ferguson's diglossia at all. Ferguson (1959: 328, 329) himself put it this way:

> "In one set of situations only H is appropriate and in another only L, with the two sets overlapping only very slightly. […] In all the defining languages some poetry is composed in L, and a small handful of poets compose in both, but the status of the two kinds of poetry is very different […]."

🗋36 Consequently, it is reasonable to assume that a slight overlap, a few texts in the 'wrong' variety, a handful of translations from the H variety into the L variety and vice versa, which Uspenskij (2002: 472–512, §§18–19) describes for the situation after the Third South Slavic Influence, would not have led Ferguson to assume a complete dissolution of diglossia. Moreover, Church Slavonic never lost its high prestige compared to the East Slavic vernacular, and language teaching in schools remained limited to Church Slavonic for a long time. The first Russian grammar book aimed at native speakers and therefore written in Russian, Adodurov's "detailed grammar" ("пространная грамматика"), was probably written between 1738 and 1741 (cf. Uspenskij 1975: 44, 49), but it was not printed and was probably only known to a small circle of people (ibid. 91; cf. also Keipert 2002: 119). The first Russian grammar for Russians that was also printed, Lomonosov's *Rossijskaja grammatika*, did not appear until 1755, more than a century after the beginning of the Third South Slavic Influence.

Overall, therefore, the sociolinguistic shift from diglossia to bilingualism does not seem to have taken place as quickly as Uspenskij assumes. We may assume a transitional phase that lasted until the mid-18th century, during which the main characteristics of diglossia continued to apply, but the distribution of the varieties gradually changed from *use-oriented* to *user-oriented*, or in other words: the formerly diaphasic diglossia increasingly took on diastratic elements. This phenomenon could be described as *diastratic diglossia*.

The actual emancipation of Russian from Church Slavonic only began gradually as a result of Peter I's reforms. By granting secular texts a higher status than before, by creating secular educational institutions such as the St. Petersburg Academy of Sciences, and by establishing printing houses that were not under ecclesiastical authority, the foundation was laid for Russian to compete with Church Slavonic as an H variety, and for the latter to eventually become unnecessary outside the religious sphere.

4. Native variety vs. foreign language

All of this is relevant to our initial question of when Church Slavonic came to be regarded as a foreign language, because there is a mutual dependence between the 37 sociolinguistic relationship of two varieties used and their linguistic status in relation to each other:

«Соответственно, в условиях двуязычия оба языка […] осознаются именно как два разных самостоятельных языка. Между тем, в условиях диглоссии сосуществующие языки не противопоставляются, а отождествляются.» (Uspenskij 2002: 26, §2.2.1)	"Accordingly, under the circumstances of bilingualism, both languages […] are perceived as two different independent languages. Meanwhile, under the condition of diglossia, the coexisting languages are not opposed to but identified with each other."

Elsewhere, I have already shown that traditionally Church Slavonic was not regarded as a foreign language by its users: Ivan Uževyč wrote a single *Grammatica Sclavonica* in 1643–1645, in which he only occasionally distinguished between *lingua sacra* (i.e. Church Slavonic) and *lingua popularis* (i.e. Ruthenian; cf. Bunčić 2006: 144). In a 17th-century manuscript preserved in Copenhagen, the Russian translator of Noël de Berlaimont's popular phrasebook (Sørensen 1962; cf. Keipert 1993) did the same, referring to Church Slavonic alternatives to the vernacular form in only 19 places, e.g. **ѧ҇ топе́рь пришо́лъ из̆ школы / пришо́лъ из ꙋчи́лища** 'I have just come from school' (fol. 8; cf. Bunčić 2006: 143). For decades, there has been debate as to whether the language of the Bible translations printed by Francysk Skaryna between 1517 and 1522 is Church Slavonic or 'Old Belarusian' (i.e. Ruthenian) or even a mixture of both (cf. Barkoŭski 1977: 4), and to this day there is no consensus on the matter (cf. Žuraŭski 2002: 71). Lunt (1990: 8) states unequivocally for the earliest period of East Slavic literature: "'Old Rusian' and OCS [Old Church Slavonic] are mere variants of one language."[8]

By contrast, Uspenskij (2002: 25, §2.2) remains rather vague about the actual linguistic status of the speech forms involved in a situation of diglossia: 38

«[…] в языковом сознании при диглоссии книжный и некнижный языки воспринимаются как один язык […]. Между тем, для внешнего наблюдателя (включая сюда и исследователя-	"[…] in a situation of diglossia, the literary and non-literary languages are perceived as one language in linguistic consciousness […]. Meanwhile, for an outside observer (including a linguistic re-

[8] One consequence of this interpretation, mentioned in Bunčić (2006: 146–147) and, to my knowledge, not yet fully explored elsewhere, is that the recensions of Church Slavonic must be regarded as belonging to two languages at the same time: For example, Russian Church Slavonic is then both a variety of Church Slavonic alongside Croatian Church Slavonic, Serbian Church Slavonic, etc., and a variety of Russian alongside the Russian dialects.

> лингвиста) естественно в этой ситуации видеть два разных языка. Таким образом, […] диглоссию можно определить как такую языковую ситуацию, когда два разных языка воспринимаются (в языковом коллективе) и функционируют как один язык.»
>
> searcher), it is natural in this situation to see two different languages. Thus, […] diglossia can be defined as a linguistic situation in which two different languages are perceived (by the linguistic community) and function as one language."

The question is: Who is right—the linguist Uspenskij or the Russian language community in the 18th century? Even an "outside observer" such as the early linguist Johann Werner Paus called the language of his grammar "Slavonic-Russian language" ("Slavonisch-Russische Sprache") in 1729 and spoke out in favor of the unity of the two varieties: "after all, two *jazyki* can become brothers and 2 languages, sisters" ("zwey ɥзыки können jawohl brüder u[nd] 2. Sprachen Schwester[n] werden", quoted from Živov 2004: 214). Ferguson (1959: 325) leaves no doubt about the status of the language forms in diglossia: he clearly defines diglossia as a relationship between "two or more varieties of the same language". His definition of diglossia expressly excludes "the analogous situation where two distinct (related or unrelated) languages are used side by side throughout a speech community, each with a clearly defined role". Conversely, this means that Church Slavonic and the Russian vernacular were not separate languages as long as they were in a situation of diglossia.

This view is also held by Demidov (2006: 42). Regarding the mid-18th century, he uses the expression *cerkovno-slavjanskij 'jazyk'* 'Church Slavonic "language"'—with *jazyk* 'language' in quotation marks!—and explains this as follows:

> «Термин "язык" здесь понимается как норма и, шире, функциональная разновидность русского литературного языка.»
>
> "The term 'language' is understood here as a norm and, more broadly, as a functional variety of the Russian literary language."

🗋 39 This stance is based on his observation that many Church Slavonic features introduced in the 1751 Elisabeth Bible "also have a certain usage as variants in the secular literary norm" ("имеют некоторое вариантное хождение и в светской литературной норме", Demidov 2006: 42).

5. Lomonosov's Theory of the Three Styles

Even Lomonosov, who is rightly regarded as one of the pioneers of modern Standard Russian, proves in a way that Church Slavonic was not yet a foreign language in his time: In his *Predislovie o pol'ze knig cerkovnyx v rossijskom jazyke (Preface on the utility of church books in the Russian language)* from 1758, he includes Church Slavonic in his influential version of the Theory of the Three Styles. The highest style of the Russian language, as he imagines it, consists exclusively of Church Slavonic

words, albeit preferably those that are "comprehensible for Russians and not too antiquated" ("россиянам вразумительных и не весьма обетшалых"). Summarizing his thoughts on the high style, Lomonosov remarks: "Through this *štil'*,[9] Russian surpasses many contemporary European languages, making use of the Slavonic language from the church books" («Сим штилем преимуществует российский язык перед многими нынешними европейскими, пользуясь языком славенским из книг церковных», Lomonosov 1952 [1758]: 589). In this context it is remarkable that Lomonosov's (1755: 199, §467) *Rossijskaja grammatika* even includes such specific Church Slavonic grammatical forms as the dative absolute and expressly recommends its use in the high style (cf. Vomperskij 1970: 164).

The title of the *Predislovie* should probably also be taken seriously: the author considers *church books in the Russian language* (as the syntactically ambiguous title can also be understood) to be part of the Russian language. Keipert (1994: 26–27) points out that Lomonosov's (1952 [1758]: 590) ideas on the alleged comprehensibility of the "Russian language since the reign of Vladimir" ("российский язык от владения Владимирова") primarily refer to the Church Slavonic translation of the Bible (and not to more 🗋40 vernacular texts such as the *Russkaja pravda* or the Primary Chronicle, as one might assume). Only in this way is the statement understandable, since the Church Slavonic text of the Gospels has indeed changed little over the centuries, and an 18th-century Russian would therefore have been able to read a text from, for example, the Ostromir Gospels with almost the same ease as—or rather, with hardly any greater difficulty than—the Elizabeth Bible. This comprehensibility clearly does not apply to secular texts in Old East Slavic, considering how extensively 20th-century scholars have debated how the individual provisions of the *Russkaja Pravda* ought to be understood.

Vomperskij (1970: 139) interprets Lomonosov's view to mean that Church Slavonic was historically a language distinct from Russian (which modern linguistics would of course not dispute), but that "the Russian and 'Slavonic' elements ceased to be separate languages and, in the course of a complex interaction, formed functional varieties of a unified literary language on a Common Russian [i.e. East Slavic—D. B.] vernacular basis, which he calls *štil'*" ("русская и «славянская» стихии перестали быть отдельными языками и в результате сложного взаимодействия создали функциональные разновидности единого литературного языка на общерусской [т. е. восточнославянской — Д. Б.] народной основе, которые он называет «штилями»").

It quickly became apparent that Lomonosov's Theory of the Three Styles was not, in fact, a suitable model for the new Russian standard language. However, the great advantage of this theory was the fact that the previous two varieties, Church

[9] Instead of the modern Russian form *stil'*, Lomonosov uses the now obsolete form *štil'*, which I convey here in its original because some of the research literature on Lomonosov distinguishes between *štil'* in the sense he describes it and *stil'* in the modern sense.

Slavonic and the Russian vernacular, were regrouped into three styles. As a result, the middle style had to be composed of both Church Slavonic and vernacular elements. Only after the high and low styles were abandoned, i.e. no longer considered parts of the literary language, did the Russian standard language become completely independent of Church Slavonic, which from that point on can really be considered a foreign language. This development took place mainly during the second half of the 18th century. Vomperskij (1970: 180) dates the dissolution of the three styles to the mid-1770s and the early 1780s:

> «Начиная с этого времени становятся наглядными процессы смешения и разрушения между языковыми стилями и выработки в недрах старой системы «штилей» функционально-речевых стилей, когда сближение норм разговорной и 41 письменной речи создает единую национально-языковую норму, которая нашла свое выражение в творчестве А. С. Пушкина.»
>
> "Starting from this time, the processes of mixing and destruction between linguistic styles and the development of functional styles of speech within the old system of *štili* become apparent, when the convergence of the norms of spoken and written speech creates a single national language norm, which found its expression in the works of A. S. Puškin."

6. 'Ecclesiastical' vs. 'civil' script variants

Interestingly, it is at the same period that the Old Cyrillic type was tied to the Church Slavonic language (cf. Bunčić & Kislova & Rabus 2016 for more details): until the 1760s, the 'civil' and 'ecclesiastical' script variants were used independently of the language form for secular and religious texts, respectively. The interpretation of a text as secular or religious depended not only on the subject matter within the text but also on the context in which the text was placed. For example, a sermon on a current political topic could be interpreted as religious in a church context, but as secular in a political context. One of many examples of this is Simon Todorskij's sermon *Božie osobennoe blagoslovenie (God's Special Blessing)*, which he gave in 1745 at the wedding of the heir to the throne, the future Tsar Peter III, to the future Tsarina Catherine the Great. The Synodal Printing House in Moscow printed the sermon in Old Cyrillic type (Zernova & Kameneva 1968: № 359; Guseva 2010: № 1356), whereas the Printing House of the Academy of Sciences in St. Petersburg published the same text, virtually unchanged in terms of language, as a political contribution in civil type (Kondakov 1963–1967: № 6476). The last such text printed in two scripts that I have been able to find so far is *Učreždenie Imperatorskago vospitatel'nago doma (The Establishment of the Imperial Orphanage)* by Ivan Beckoj (or Beckij), which was printed in 1763 in civil type (Kondakov 1963–1967: № 561) and in 1764 in Old Cyrillic type (Zernova & Kameneva 1968: № 659; Guseva 2010: № 101). Only after that did the Old Cyrillic script variant remain restricted to texts in Church Slavonic, while vernacular Rus-

sian texts were from then on printed exclusively in civil type. This strict separation of script variants according to language basically still applies today.

7. Conclusion

42 Drastically simplified and reduced to a few key protagonists, the development could be described as follows: Peter the Great's 'Enlightenment from above' initiated the separation of Russian language norms from Church Slavonic, Lomonosov and others (e.g. Karamzin) promoted it, and Puškin put the finishing touches to the now finally emancipated modern Russian standard language. The third South Slavic influence, in any case, was not the immediate reason for the transition to bilingualism—let alone the period of turmoil around 1600, as Uspenskij (2002: 472, § 18.1) suggests. The phenomena in the 17th and first half of the 18th century that he describes (translations, parallel texts, etc.) merely characterize a transitional situation, in which Church Slavonic and the Russian vernacular continued to function as H and L varieties of a unified 'Slavonic-Russian' language. However, their distribution was no longer exclusively use-oriented but increasingly user-oriented, which also facilitated the appearance of parallel texts. If one wanted to find a label for this sociolinguistic situation, *diastratic diglossia* would probably be the most appropriate description. It was not until the second half of the 18th century, namely during the 1760s–1780s, that Russian emancipated itself from Church Slavonic, so that from then on Russians can legitimately regard the latter as a foreign language.

References

Ammon, Ulrich. 1989. Towards a descriptive framework for the status/function (social position) of a language within a country. In idem (ed.), *Status and function of languages and language varieties*, 21–106. Berlin, New York: de Gruyter.

Ammon, Ulrich. 1995. *Die deutsche Sprache in Deutschland, Österreich und der Schweiz: Das Problem der nationalen Varietäten.* Berlin, New York: de Gruyter.

Ammon, Ulrich. 2000. Diaphasisch. In Glück, Helmut (ed.), *Metzler Lexikon Sprache*, 2nd edn., 157. Stuttgart, Weimar: Metzler.

Auer, Peter. 2005. Europe's sociolinguistic unity, or: A typology of European dialect/standard constellations. In Delbecque, Nicole (ed.), *Perspectives on variation: Sociolinguistic, historical, comparative*, 7–42. Berlin: Mouton de Gruyter.

Barkoŭski, Viktar I. 1977. Ad rėdaktara. In Aničenka, Uladzimir V. (ed.), *Sloŭnik movy Skaryny*, vol. 1, 3–6. Minsk: Vyšėjšaja škola.

Bunčić, Daniel. 2006. *Die ruthenische Schriftsprache bei Ivan Uževyč: unter besonderer Berücksichtigung seines Gesprächsbuchs* Rozmova/Beseda*, mit Wörterverzeichnis und Indizes zu seinem ruthenischen und kirchenslavischen Gesamtwerk.* München: Sagner.

Bunčić, Daniel. 2008. Die (Re-)Nationalisierung der serbokroatischen Standards. In Kempgen, Sebastian et al. (eds.), *Deutsche Beiträge zum Internationalen Slavistenkongress, Ohrid 2008*, 89–102. München: Sagner. [English version in this volume, pp. 38–54.]

Bunčić, Daniel & Kislova, Ekaterina & Rabus, Achim. 2016. In Bunčić, Daniel et al., *Biscriptality: A sociolinguistic typology*, 102–121. Heidelberg: Winter.

Casad, Eugene H. (ed.). 1992. *Windows on bilingualism.* Dallas: SIL.

Coseriu, Eugenio. 1992. *Einführung in die Allgemeine Sprachwissenschaft.* 2nd edn. Tübingen: Francke. [1st edn. 1988, based on a Spanish manuscript from 1981.]

Demidov, Dmitrij G. 2006. Rol′ knig cerkovnoj pečati v istorii russkogo literaturnogo jazyka. In Kornilov, Gennadij E. (ed.), *Čtenija, posvjaščennye dnjam slavjanskoj pis′mennosti i kul′tury: K 70-letiju so dnja roždenija doktora filologičeskix nauk, professora G. E. Kornilova: Sbornik statej meždunarodnoj konferencii*, 31–43. Čeboksary: Čuvašskij gosudarstvennyj universitet imeni I. N. Ul′janova.

Ferguson, Charles A. 1959. Diglossia. *Word* 15. 325–340.

Goossens, Jan. 1977. *Deutsche Dialektologie.* Berlin, New York: de Gruyter.

Gregory, Michael. 1967. Aspects of varieties differentiation. *Journal of Linguistics* 3(2). 177–197.

Guseva, Aleksandra A. 2010. *Svod russkix knig kirillovskoj pečati XVIII veka tipografij Moskvy i Sankt-Peterburga i universal′naja metodika ich identifikacii.* Moskva: Indrik.

Kabatek, Johannes. 2000. L'oral et l'écrit — quelques aspects théoriques d'un « nouveau » paradigme dans le canon de la linguistique romane. In Dahmen, Wolfgang et al. (eds.), *Kanonbildung in der Romanistik und in den Nachbardisziplinen: Romanistisches Kolloquium XIV*, 305–320. Tübingen: Narr.

Keipert, Helmut. 1993. Nochmals zur Kopenhagener Handschrift russischer Gespräche aus dem 17. Jahrhundert. *Zeitschrift für Slavische Philologie* 53(2). 285–303.

Keipert, Helmut. 1994. Die *knigi cerkovnye* in Lomonosovs "Predislovie o pol′ze knig cerkovnych v rossijskom jazyke". *Zeitschrift für Slavische Philologie* 54(1). 21–37.

Keipert, Helmut. 2002. Einleitung. In Keipert, Helmut & Huterer, Andrea (eds.), *Compendium Grammaticae Russicae (1731): Die erste Akademie-Grammatik der russischen Sprache*, 7–127. München: Verlag der Bayerischen Akademie der Wissenschaften.

Koch, Peter & Wulf Oesterreicher. 1985. Sprache der Nähe – Sprache der Distanz: Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch* 36. 15–43.

Kondakov, Ivan P. (ed.). 1963–1967. *Svodnyj katalog russkoj knigi graždanskoj pečati XVIII veka: 1725–1800.* 5 vols. Moskva: Gosudarstvennaja biblioteka SSSR imeni V. I. Lenina.

Lomonosov, Mixajlo. 1755. *Rossijskaja grammatika*. Sanktpeterburg: Pri Imperatorskoj Akademii Nauk. [Reprint: Leipzig 1972.]

Lomonosov, Mixail V. 1952 [1758]. Predislovie o pol′ze knig cerkovnyx v rossijskom jazyke. In idem, *Polnoe sobranie sočinenij* (ed. by V. V. Vinogradov, S. G. Barxudarov, and G. P. Blok), vol. 7, 585–592. Moskva, Leningrad: Izdatel′stvo Akademii Nauk SSSR.

Lunt, Horace G. 1990. History, nationalism, and the written language of Early Rus'. *Slavic and East European Journal* 34(1). 1–29.

Rehder, Peter. 1989. Diglossie in der Rus'? Anmerkungen zu B. A. Uspenskijs Diglossie-Konzeption. *Die Welt der Slaven* 34(2). 362–382.

Rossïjskoe Biblejskoe Obščestvo. 1820. *Gospoda našego Ïisusa Christa Svjatoe Evangelïe, ot Matfeja, Marka, Luki i Ïoanna, i Dějanïja Svjatyx Apostolov, na Slavjanskom i Russkom narečii.* Sanktpeterburg: Greč.

Schweier, Ulrich. 2001. Die "kirchenslavisch-ostslavische Diglossie" und ihre Auflösung als Kategorisierungswandel: Vom aristotelischen zum prototypischen Modell. *Die Welt der Slaven* 46(1). 27–42.

Sørensen, Hans Christian. 1962. *Ein russisches handschriftliches Gesprächsbuch aus dem 17. Jahrhundert.* København: Munksgaard.

Trunte, Nicolina. 2014. *Slavenskïj jazyk: Ein praktisches Lehrbuch des Kirchenslavischen in 30 Lektionen, zugleich eine Einführung in die slavische Philologie.* Vol. 2. 2nd edn. München: Sagner. [1st edn. 1998.]

Uspenskij, Boris A. 1975. *Pervaja russkaja grammatika na rodnom jazyke: Dolomonosovskij period otečestvennoj rusistiki.* Moskva: Nauka.

Uspenskij, Boris A. 1987. *Istorija russkogo literaturnogo jazyka (XI–XVII vv.).* München: Sagner. [3rd edn. Moskva: Aspekt Press 2002.]

Vomperskij, Valentin P. 1970. *Stilističeskoe učenie M. V. Lomonosova i teorija trex stilej.* Moskva: Izdatel'stvo Moskovskogo universiteta.

Zernova, Antonina S. & Kameneva, Tat'jana N. 1968. *Svodnyj katalog russkoj knigi kirillovskoj pečati XVIII veka.* Moskva: Gosudarstvennaja ordena Lenina Biblioteka SSSR imeni V. I. Lenina.

Živov, Viktor M. 2004. *Očerki istoričeskoj morfologii russkogo jazyka XVII–XVIII vekov.* Moskva: Jazyki slavjanskoj kul'tury.

Žuraŭski, Arkadzij. 2002. Die Sprache der Übersetzungen. In Rothe, Hans & Scholz, Friedrich (eds.), *Biblija ruska: vyložena doktorom Franciskom Skorinoju, Prag 1517–1519. Kommentare, Apostol* (Biblia Slavica, III.1), 69–103. Paderborn: Schöningh.

Right-to-left Cyrillic among the Bogomils?

In his famous book on scripts and religion, Mieses (1919: 325) reports that "The Manichaean-Bogomil Bulgarians left behind some inscriptions in a Cyrillic script written from right to left rather than from left to right". In my quest to find these inscriptions through a web of imprecise quotations without references, I found out that what he meant were actually members of the Bosnian Church. Indeed, in Bosnia one can find exactly two inscriptions that look inverted, on a *stećak* for Vlatko Vlađević and a tomb slab for Pavao Komlinović. After considering other possible explanations for the reversed inscriptions (secret writing, magic, direction of objects), the unspectacular solution seems to be that the stonemasons engraving the letters were illiterate and made technical mistakes. A brief inspection of other cases of "schismatic" changes in writing direction cited by Mieses reveals that none of them is actually caused by a change in religion, and his hypothesis can be rejected completely.

Originally written in English for my habilitation thesis as submitted to the faculty but then published separately in German as "Linksläufige kyrillische Schrift bei den Bogumilen?" in Meyer, Anna-Maria & Reinkowski, Ljiljana (eds.), *Im Rhythmus der Linguistik: Festschrift für Sebastian Kempgen zum 65. Geburtstag*, 115–134. Bamberg: University of Bamberg Press 2017. DOI: 10.20378/irbo-49095.

DOI: 10.18716/omp.64.9

> "[The Eiffel Tower] was the work of a German Jew, the Jewish response to the Sacré-Coeur, explained de Biez. He was perhaps the most combative anti-Semite in the group, who began his demonstration about Jewish inferiority from the fact that they write in the opposite way to normal people. 'The very form of this Babylonian construction,' he said, 'demonstrates that their brain is not made like ours. …'" (Umberto Eco, *The Prague Cemetery*, 437, ch. 23)

1. Introduction

In the course of my work on biscriptality, I came across a statement that I had to investigate and that ultimately offered me a deep insight, if not into the linguistics of writing itself (an interest that Sebastian Kempgen and I share), then at least into the history of the discipline:

> "Die manichäisch-bogumilischen Bulgaren hinterließen einige Inschriften in einer linksläufigen statt rechtsläufigen kyrillischen Schrift." (Mieses 1919: 325)

> "The Manichaean-Bogomil Bulgarians left behind some inscriptions in a Cyrillic script written from right to left rather than from left to right."

This statement appears in the book *Die Gesetze der Schriftgeschichte (The Laws of the History of Writing)* by Matthias Mieses (1919)—an important work, not only because it is one of the first scientific monographs on the linguistics of writing, but also because it refutes a notion of writing that was widespread at the time, namely that each 'race' had its own script. According to this view, the script specific to the Germanic peoples (or "Aryans") is blackletter (the "German script"), THE SCRIPT OF THE ROMANCE PEOPLES IS ANTIQUA, **THAT OF THE JEWS IS HEBREW, AND THE "REAL" SLAVS USE THE CYRILLIC ALPHABET**. Mieses refutes this racist theory of writing by showing that writing fundamentally follows religion and not race:

> „Von den Mitgliedern der uralaltaischen Rasse schreiben die Ungarn mit lateinischen Lettern, die Finnen und Esten mit Frakturbuchstaben und die Türken mit arabischen Schriftzeichen. Warum? Weil sie glaubensverschieden sind. Die Ungarn sind katholisch. Die Finnen und Esten verdanken ihre Glaubensform dem trotzigen Augustinermönch von Wittenberg. Die Türken beten mit dem Gesicht nach Mekka." (Mieses 1919: 10)
>
> "Of the members of the Uralo-Altaic race, the Hungarians write in Latin letters, the Finns and Estonians in blackletter, and the Turks in Arabic characters. Why? Because they have different beliefs. The Hungarians are Catholics. The Finns and Estonians owe their form of belief to the defiant Augustinian monk from Wittenberg. The Turks pray with their faces towards Mecca."

🗋116 Even though this certainly does not apply always and everywhere, and Mieses overshoots the mark in many places, his approach is definitely much closer to the truth than the racist theory of writing. In Slavic Studies, we are also familiar with Riccardo Picchio's (1958) distinction between *Slavia Romana* and *Slavia Orthodoxa*, which largely corresponds to the use of the Latin and Cyrillic alphabets, respectively.

2. In search of right-to-left inscriptions

Mieses (1919: 325–342) devotes a whole chapter to the question of writing direction.[1] In his opinion, it was primarily schisms within the same religion that led to the choice of a different writing direction. In this context, he also makes the statement that prompted this article: Bogomils who write Bulgarian in Cyrillic script, but in reverse? Does such a thing really exist? I wanted to check this out. Mieses

[1] The direction of writing is indeed deeply ingrained in the general consciousness. I still remember an experience from the time when one still had to place copy orders in the copy departments of Eastern European libraries. In L'viv, the photocopier dispensed the paper with the printed side facing up, thus reversing the order of the pages, which the library employee, who had probably been making copies for years, did not seem to have noticed. So when I asked her to please start copying the book from the back, she replied in astonishment: «Вы араб, что ли?» ("Are you an Arab or what?")

refers to Renner (1897), who, however, reports only one such inscription on a Bosnian *stećak*, one of those mostly cuboid or sarcophagus-shaped medieval gravestones typical of the region, which were included in the World Cultural Heritage on 15 July 2016 (cf. UNESCO 2016):[2]

> „Auf dem Wege nach Rogatica und um diese Stadt selbst finden sich zahlreiche römische und bogomilische Grabsteine. Mommsen hat schon einen in seiner Sammlung beschrieben, andere sind von Dr. Blau und Dr. Hoernes näher bestimmt worden. Unter den Bogomilensteinen sind einige bemerkenswerth wegen der sonst sehr selten vorkommenden Aufschriften. So lautet die eine, die auffallenderweise von rechts nach links zu lesen ist: ‚Va ime otca i sina i sv. duha. Ovdi leži Vlatko Vladjević […].'" (Renner 1897: 268 f.)

> "On the way to Rogatica and around that town itself, there are numerous Roman and Bogomil gravestones. Mommsen has already described one of them in his collection, while others have been specified in more detail by Dr. Blau and Dr. Hoernes. Among the Bogomil stones, some are noteworthy because of their inscriptions, which are otherwise very rare. One of them, which strikingly is to be read from right to left, says: 'Va ime otca i sina i sv. duha. Ovdi leži Vlatko Vladjević […].'"

🗋 117 Unfortunately, there are no more explicit references here other than the surnames and doctoral titles, so that it took some time to locate the sources on which Renner based his work. The cited "Dr. Hoernes" describes the cemetery of Lađevine (in the municipality of Rogatica, 70 km east of Sarajevo, today in Republika Srpska) in more detail:

> „Hier freuten sich nach den theils von links nach rechts, theils retrograd geschriebenen und nur theilweise lesbaren glagolitischen Epitaphien der riesigsten Grabsteinblöcke der Vojevode Miotoš mit seinem Sohne Stjepko und [der] gnädige Fürst Pavao und der Erzvojevode Miotoš und der […] edle Ban Vlatko Vlačević lange vor dem Einbruch der Türken ihres dunklen Daseins; hier wurden sie nach ihrem Tode […] in der tiefen, fruchtbaren Erde gebettet – ohne Sarg und ohne Gruft, wie unsere Nachgrabungen zeigten […]." (Hoernes 1888: 238)

> "Here, according to the Glagolitic epitaphs on the largest gravestones, only partially legible and some written from left to right, some backwards, the voivode Miotoš with his son Stjepko and [the] gracious Prince Pavao and Archvoivode Miotoš and the […] noble Ban Vlatko Vlačević enjoyed their dark existence long before the Turkish invasion; here they were laid to rest after their death […] in the deep, fertile earth—without coffins or tombs, as our excavations showed […]."

[2] The application was submitted jointly in 2009 by Bosnia and Herzegovina, Croatia, Montenegro, and Serbia.

All this is quite reminiscent of the children's game *broken telephone*: Hoernes, Renner and Mieses all report the striking fact that inscriptions are written from right to left, but they disagree on their number (one? many?), the ethnic affiliation of their authors (Bulgarians? Bosnians?) and the script in which they are written (Cyrillic? Glagolitic?). One wonders if any of them has really read the inscriptions with his own eyes.

118 A minor detail might illustrate the way these travelogues seem to have been written: Hoernes's (1888: 237) drawing of the graveyard (Fig. 1) seems to have been copied by Renner (1897: 265; Fig. 2). The latter's illustration shows the stones viewed from exactly the same angle and in exactly the same position, but while Hoernes (1888: 238) calls the scenery "extremely lovely" ("äußerst lieblich"), Renner shows the graveyard with a different, much rougher landscape around it. The different romantic impressions that the illustrators wanted to convey seem to have been more important than factual accuracy (to which, as the current photo of *stećci* in the immediate vicinity of Lađevine in Fig. 3[3] shows, Hoernes came closer than Renner).

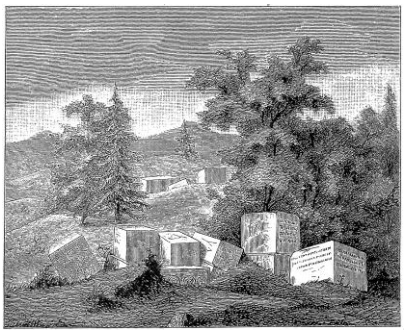

Fig. 1: Lađevine according to Hoernes

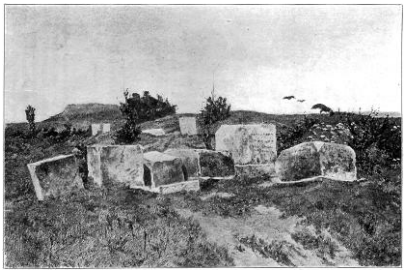

Fig. 2: Lađevine according to Renner

Fig. 3: Varošište today

3. The tombstone of Vlatko Vlađević

However, the name of Vlatko Vlađević mentioned by Hoernes and Renner is well-known. His gravestone from the beginning of the 15th century (between 1399 and 1415, cf. Vego 1981: 52) has been moved from Lađevine to the yard of the National Museum *(Zemaljski muzej)* in Sarajevo, and Marko Vego 119 (1981: 51–52), the

[3] Since the most interesting *stećci* from Lađevine are now in the National Museum in Sarajevo, the place no longer looks as it did at the end of the 19th century. Photos of Lađevine and other cemeteries in Rogatica can be found in Durmišević (2011), from which Fig. 3 is taken (http://oi47.tinypic.com/el4ihg.jpg, last accessed 28 August 2016 [and not accessible anymore as of 2026]). The villages of Lađevine and Varošište are 2 km apart.

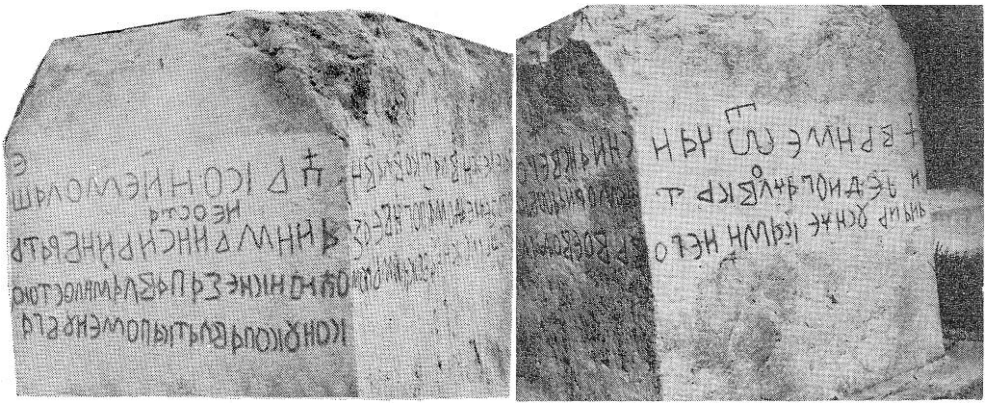

Fig. 4: Photos of Vlatko Vlađević's gravestone

museum's former director, has printed two photographs of the stone (Fig. 4).[4] These show that the inscription on the gravestone is in Cyrillic (not Glagolitic) and a variety of Serbo-Croatian (not Bulgarian)—and really runs from right to left, with all the letters being mirrored. Following Mazalić's (1941: 88) transcript, which was approved by Vego (1981: 51), the inscription reads as follows:

1:1	✝ вь имє wца и	† In the name of the Father and
2:1	сина и свѣго дх̑а	the Son and the Holy Spirit.
3:1	а сє лєжи влатко вла[ѣє]ви–	Here lies Vlatko Vla[đe]vić,
4:1	ѣь кои нє молаше	who did not bow
1:2	ни єдиога члов[и]кь[5] т–	to any man
2:2	акмо гн̑а а wвидє ми–	but his Lord and traveled
3:2	ого зємлє а дома погибє а з–	many countries and died at home and
4:2	а нимь нє оста ни сь ни брать	left no son or brother behind. 120
1:3	а на нь ꙋсиче камн иєго–	And the stone for him was carved
2:3	вь воєвода миw–	by his *vojevoda*[6] Miotoš

[4] The photos by Vego (1981: 51, 52) show all four inscribed sides of the stone: on the far right is side 1 and to its left part of side 2; the image on the left shows side 3 on the right and side 4 on the left. The transcription refers to this side numbering and also indicates line numbers.

[5] Mazalić (1941: 88) identifies the last letter of the word as ⟨ꙗ⟩—which would also be expected grammatically—although the photo shows that there is a clearly legible ⟨ь⟩, which here apparently represents the phoneme /a/, since ⟨ь⟩ can also represent /a/ in a 'strong' position elsewhere. (Given the consistency of the mirroring in the rest of the text, it is unlikely that this is simply an accidentally unmirrored form of ⟨ꙗ⟩.)

[6] The words *vojevoda* and *knez* are often translated as *duke* and *prince*, respectively. However, both words are so polysemous, especially in the Balkans, that I prefer to leave them unchanged in the translation. On Vlatko Vlađević's gravestone, on the one hand, *vojevoda* Miotoš is obviously subordinate to Vlatko, while *knez* Pavao (or Pavle) Radinović, who is

3:3	тошь с лѫжица божиѡмь пом–	of Lužice with God's help
4:3	оѣю и кнєза павла милостою	and with the grace of *knez*[6] Pavao,
4:4	кон ѫкопа влатка помєнѫ бга	who buried Vlatko in the name of God.

As we have seen, Renner (1897: 268f.) identifies the gravestone as Bogomil, which apparently made Mieses extrapolate that therefore it had to be Bulgarian. Truhelka (1908: 423) even argues that Vlatko Vlađević "was without doubt a high dignitary of the Bogomil Church" ("da je i Vlatko Vlagjević bio bez sumnje visoki dostojanstvenik bogumilske crkve"). However, as the inscription begins with the symbol of the cross, this interpretation has to be adjusted: What is meant here by "Bogomils" is the Bosnian Church, which was often slandered as Bogomil, especially since some Bogomils from Bulgaria had indeed found refuge in Bosnia. However, the Bosnian Church was clearly not identical to them. One of the differences is that "Bogomils abhorred the sign of the cross; the cross appears on the head of several Bosnian Church documents" (Malcolm 1994: 38).[7]

🗋121 Apart from that, Renner's attribution of the gravestone to the Bosnian Church is based on an outdated theory. As one can see 🗋122 in Figures 1–4, the *stećci* differ significantly in appearance from other Christian gravestones. Their rectangular shape is "more or less peculiar to the Bosnian region" (Malcolm 1994: 30), but it also occurs in neighboring regions of Croatia, Serbia, and Montenegro. It was previously assumed that these *stećci* were a peculiarity of the Bosnian Church. However, even during the heyday of the Bosnian Church large parts of the Bosnian population continued to be Catholic or Orthodox, and if all the *stećci* belonged to members of the Bosnian Church, then there would be no medieval gravestones left for the Catholics and Orthodox (Malcolm 1994: 30f.).

Nevertheless, it is at least possible that Vlatko Vlađević was a member of the Bosnian Church. However, none of the inscriptions and manuscripts that can be unambiguously attributed to the Bosnian Church (e.g. the three inscriptions referring to *krstjanin* Radohna, *gost* Mišljen and *gost* Milutin examined by Miletić 1957,

---

also known from historical sources to have ruled over a considerable area, is his lord (cf. Vego 1981: 52). On the other hand, in the inscription discussed below, *knez* Pavao Komlinović is clearly subordinate to *vojevoda* Sandalj Hranić Kosača, who ruled over an even larger area (cf. the map by Panonian 2010).

[7] Miletić (1957) has shown that much of what we know about the Bosnian Church, especially its titles (e.g., *krstjanin* 'Christian' for a monk, *gost* 'host' for an abbot), resembles the pre-Benedictine monastic organizations known from Celtic Christianity. She describes the medieval Bosnian Church as "a branch of the Basilians, more precisely of those ancient proto-Basilians whose monasticism preserved the oldest elements and was linked to the Egyptian tradition of the first hermits and anchorites" ("un ramo dei basiliani, e precisamente di quelli antichi, protobasiliani, il cui monachismo era il depositario dei più antichi elementi e si riallacciava alla tradizione egiziana dei primi eremiti ed anacoreti", Miletić 1957: 183; cf. the approving account in Malcolm 1994: 34–36).

or the manuscripts enumerated by Hercigonja 2006: 101, fn. 35) seem to be written from right to left.

4. Various theories for right-to-left writing

Vego (1981: 53) mentions only one other Cyrillic inscription with right-to-left text sections, namely the inscription on Pavao Komlinović's *stećak* in the cemetery of Bakri near Čitluk in Herzegovina. He explains the use of mirrored text in both inscriptions as "secret writing" ("tajno pismo", ibid.). Obviously, this conclusion is based on Mazalić (1941: 88, fn. 150), who mentions the writing direction of Vlatko Vlađević's gravestone only in a footnote:

> "Nadpis je ispisan tajnim načinom pisanja, naopako. Zašto ne može se reći. Svakako nije pod uticajem turskog pisma, našto se ranije pomišljalo."

> "The inscription is written in a secret form of writing, backwards. It is impossible to say why. In any case, it was not under the influence of the Turkish script, as was previously thought."

Indeed, before assuming that an obviously Christian gravestone inscribed before the Ottoman conquest mimics the Arabic writing direction, one would like to see a few dozen 123 Cyrillic documents by Muslim Slavs from the Ottoman era to be written from right to left. To my knowledge, there is not a single example of this.

However, the thesis that this is secret writing has the same problem: I do not know of any instance where this crude encryption method was used to encode a text that was really meant to be top secret. Those few cases of reversed writing that Sobolevskij (2007 [1902]: 53f.), Karskij (1928: 253), Čerepnin (1956: 395) and Ščepkin (1999: 162) mention seem not to have been used for serious purposes. Take, for example, the inscription "☩ ъшорг аꙁ ъшуܐг аⷨд умот ѣ҃тчорп еис отк а ☩" (quoted by all authors): 'Decoded', it reads "☩ а кто сие прочтѣ҃ тому даⷨ грушъ ꙁа грошъ ☩" ("† To whoever reads this, I will give pears for a penny †").

A better example, also from the East Slavic region, might be the inscription on a church bell of St. George's Cathedral in L′viv: After mentioning the year 1341, Saint George and the names of the current prince and abbot, the artisan doing all the work probably felt the desire to immortalize his own name as well but knew that this wish was not seemly. Therefore, he wrote his name in mirror-writing: "а писалъ скора ѧковъ" ("Jakov Skora wrote it", cf. Fig. 5; Peščak 1974: 22–23). Speranskij (1929: 43) cites a similar note in a Serbian manuscript from 1670 recording in reversed letters that "the hand of sinful, lazy Dimitrij" ("роука грѣшнаго Димитрıа лениваго") wrote the text.

Fig. 5: Inscription on a church bell (1341)

However, there does not seem to be any motivation for secret writing on Vlatko's *stećak*: Why should anyone want to keep the inscription on a gravestone secret?

Another reason why a text might be inverted, which does not seem to have been considered in the context of this inscription so far, is magic. An example of this is the fragmentary Novgorod *gramota* № 674 from the end of the 12th century, which was used as a talisman 🗋124 (*obereg*, cf. Zaliznjak 2004: 462–463). It bore the words "[ѿ] гласа вражиа и ѿ сотоужьниа грѣшьница [i.e. …стоужениа грѣшьничѧ]" ("… from the voice of the enemy and from the oppression of the wicked") from Psalm 54/55:4 (and originally perhaps something like "Keep me" before it, cf. Zaliznjak 2004: 463). The mirroring of such a text is a remnant of pagan magical beliefs and was supposed to enhance the power of the talisman: By writing bad things backwards, you turn them into something good. However, in the case of the birch bark, there is an obvious connection between the content of the text and the magic it was supposed to bring about: the Novgorod talisman was written from right to left to protect its bearer "from the voice of the enemy and from the oppression of the wicked". In contrast to this, the text of Vlatko's gravestone is completely referential and does not include any appellative elements. If this had been meant as a kind of magic spell, then the angels or demons who were supposed to work the magic would not have known what to do.

A final reason for reversed writing directions that should be briefly mentioned here is the idea that objects have a natural direction and that writing should follow this direction. For example, inscriptions on vehicles in China usually run from the front to the rear, which means that the writing on the left side of the vehicle runs from left to right, but on the right side it runs from right to left. For example, Wilde et al. (2005–2016) show a Chinese postal service vehicle with the inscription "中国邮政" on the left door and the English translation "CHINA POST" below it, but "政邮国中" and "TSOP ANIHC" on the right door (in each case without mirroring the characters). Similarly, one could imagine that in the case of the *stećci*, the deceased himself dictated the direction of the writing. Mirror writing could thus be justified by the fact that it should be legible from the inside—this could well explain the inscription on the stone block for Vlatko Vlađević. However, if this were the reason, one would expect to see more examples of it, similar to how most vehicles, market stalls, etc. in China are labeled from front to back, and similar to how anthropomorphic and zoomorphic characters in Egyptian burial chambers consistently face in the opposite direction to other texts (cf. fn. 12). However, since mirror writing occurs only twice on the more than 70,000 preserved *stećci*, this explanation is also rather unlikely.

5. The tombstone of Pavao Komlinović and the mystery's unspectacular solution

125 To solve the mystery of the reason for the right-to-left inscription on Vlatko Vlađević's *stećak*, it is necessary to look more closely at the other example of mirrored text from Bosnia, Pavao Komlinović's slab-shaped gravestone (which was inscribed between 1423 and 1434, cf. Vego 1962: 26; see Fig. 6[8]). The text reads:

outside top	✝ а се лежн кнеꙁь паваѡ	† Here lies *knez*[6] Pavao
outside right	комлнновнћь на свон	Komlinović on his
outside bottom	[пле]меннтон	[patri]mony
outside left	на проꙁрачцꙋ ꙋ днн воеводе	in Pozračac during the days of *vojevoda*[6]
inside top	сандаль кон га	Sandalj, whom he[9]
inside right	почтено н внрно слꙋжаше ꙋчрьто	honorably and faithfully served; engraved
inside bottom	ꙋчрьто [sic] на	engraved [sic] on the
inside left	племеннтен	patrimony

The inscription starts with the cross (in the top left corner of Fig. 6) and then goes clockwise around the slab, first winding along the outer edge, then spiraling inward. Vego (1962: 27) has marked two of the eight lines as "zdesna nalijevo" ("from right to left") and two as "slijeva 126 nadesno" ("from left to right"); the other four lines are unmarked. In any case he explicitly notes that "the letters of the inscription are carved sometimes in one and sometimes in the other direction" ("Slova natpisa su urezivana sad s jedne sad s druge strane", ibid.).

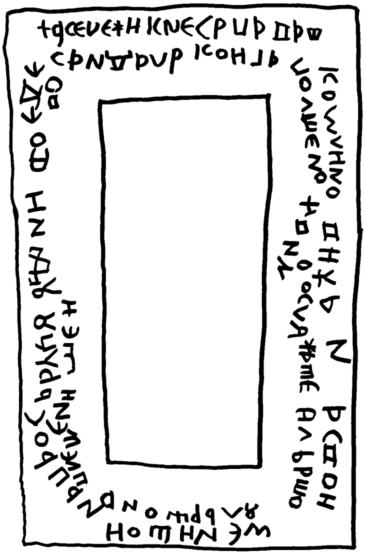

Fig. 6: Pavao Komlinović's gravestone

However, I see no reason to assume such a chaotic arrangement of the inscription. In my opinion there are two alternative but internally consistent ways to read the text, assuming that the tomb slab lay flat on the grave and that,

[8] As the photograph in Vego (1962: 26) is very hard to read because of the furrowed stone, I have made a tracing that shows only the lines of the letters according to Vego's (1962: 27) reading.

[9] This is the interpretation chosen by Vego (1962: 27) in his French translation. The word order would actually rather suggest "who … served him", but what we know from other sources about *vojevoda* Sandalj Hranić Kosača (cf. footnote 6), as well as the formulation *u dni* 'in the days of' indicate that *koji* 'who' has to refer to *knez* Pavao, and *ga* 'him' to Sandalj.

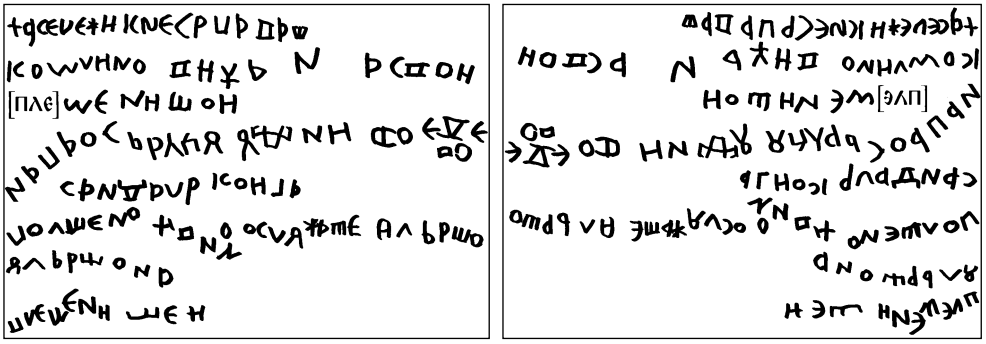

Fig. 7: Left-to-right and right-to-left reconstructions of the inscription

similar to the Humac Tablet (Nosić 1995: 33), one could walk around the slab and read the text in a spiral pattern. Either the inscription was meant be read from a viewpoint opposite each line, walking clockwise around the grave starting at the foot; in this case, the lines run from left to right. Or one should read from a viewpoint on the side of the slab nearest the current line, also walking around it clockwise but starting at the head end; in this case, all lines run from right to left. Figure 7 demonstrates these two possible interpretations by arranging the lines horizontally. It is obvious that in the left-to-right version many letters are upside down. This is the reason why one is tempted 🗋127 to read the lines on the side where one is standing, which makes them run from right to left. However, this reading is not flawless either: The letters ⟨ꙁ⟩ and ⟨ꙗ⟩ (with one exception) do not point in the direction of writing (i.e. they are not mirrored), and the ⟨ꙅ⟩ in line 5 is equally wrong in both versions.

The use of right-to-left or mirrored Cyrillic script here is therefore probably neither due to heresy or schism expressed in the direction of writing, as Mieses (1919: 325) assumes, nor to 'secret writing', as Vego (1981: 53) suggests, nor to magic. The text was probably meant to be read from the opposite side of the slab. However, the stonemason chiseled the 🗋128 letters into the slab standing (or kneeling) next to the slab on the nearest side. Since he was not very experienced (the placement of the letters does not speak in favor of his skills), he correctly mirrored the letters horizontally but forgot to mirror them also vertically, so that they turned out upside down.

Let us return to Vlatko Vlađević's *stećak*. It stands upright, so there is no confusion between top and bottom. Apart from that, the letters are almost consistently mirrored horizontally (except for ⟨ꙁ⟩, which is unmirrored in all three cases, see 3:2 and 4:3, and the ⟨ꙁ⟩ in the word ⟨ѯнєзд⟩ in 4:3). However, since this is probably the only inscription of this kind written from right to left in Bosnia and the surrounding area, and since it cannot be definitively attributed to any religious community, we can consider Mieses's claim to be refuted. There might be many reasons why a single text is written 'backwards'—even nowadays many children who have

only started to learn to write mirror individual letters or whole passages. Our scribe left out letters in several places, later adding "не оста" above the line (4:2), but never correcting the name of the deceased, which was misspelled as "Vlavić" (3:1). It is therefore probably best to explain the mirror writing in Vlatko Vlađević's case, as well as the upside-down letters in Pavao Komlinović's case, as the result of the scribes' incompetence. This brings us back to a conjecture already expressed by Truhelka (1889: 74), which, however, seems to have been forgotten, since it was not even mentioned by Mieses (1919), Mazalić (1941), or Vego (1981):

> "Ja sam sebi znao taj pojav samo tako protumačiti, da skulptor natpisa nije umio pisati, da je imao nacrt natpisa na papiru pred sobom, pa ga s obratne strane prenio na kamen te ga uklesao."
>
> "The only explanation I have been able to find for this phenomenon is that the sculptor of the inscription could not write, that he had a sketch of the inscription on paper before him, but that he placed it upside down on the stone and carved it in."

This assumption requires only one quite plausible assumption, namely that the stonemason lacked reading skills. According to Ockham's razor, this makes it preferable to hypotheses about a schismatic 🗋 129 change in writing direction, Arabic influence, secret writing, or magic.

6. Other cases of 'schismatic' changes in writing direction

In the case of the Bulgarian Bogomils, who turned out to be Bosnian Christians, Mieses was wrong. But what about the other examples he cites (Mieses 1919: 325–342)? Is a change in direction of writing for religious reasons otherwise a real phenomenon? If one takes a closer look at his examples, one finds that they are all distorted. I will briefly discuss only his first three examples here.

> „Die mohammedanische Bevölkerung der Malediven trat seit ihrer Islamisierung in Widerspruch mit allen ein indisches Alphabet gebrauchenden Völkern des asiatischen Südens und fing ihr bisheriges rechtsläufiges heimatliches Alphabet linksläufig zu schreiben an." (Mieses 1919: 325)
>
> "Since their Islamization, the Muslim population of the Maldives has come into conflict with all the peoples of the South of Asia who use an Indic alphabet and has begun to write their previous right-to-left native alphabet from left to right."

It is true that the Thaana script used for the Divehi language of the Maldives, in contrast to all other Indic scripts, is written from right to left. However, Thaana does not belong to the Indic script family, as it does not derive from the Brahmi script, but is based on the basic principles of the Arabic script (with vowel signs being mandatory, making it a syllabic alphabet; furthermore, it was probably orig-

inally invented as a secret script, as its consonant letters are based on the Arabic numerals 1 to 9).

> „Die Aethiopier setzten sich nach ihrer Christianisierung in einen Richtungsgegensatz mit der eigenen Vergangenheit und gaben ihrem bis damals linksläufigen semitischen Alphabet eine rechtsläufige Wendung." (Mieses 1919: 325)
>
> "After their Christianization, the Ethiopians took an opposite directional stance to their own past and gave their Semitic alphabet, which until then had been right-to-left, a left-to-right twist."

In this case, the Ethiopian script does indeed belong to the Semitic script family. However, even its direct predecessor, the ancient South Arabian script, was not necessarily left-to-right originally, but was initially written mainly in *boustrophedon* (alternating lines from left to right and right to left).

130 The different writing directions in Arabic and Ethiopian are therefore based solely on independent decisions to adopt a uniform direction after predecessors with varying writing directions.[10] Of course, it was certainly practical that the Ethiopian script had the same direction in which the Greek New Testament was written, just as it was probably convenient for Dhivehi speakers to write in the same direction as the Qur'an. But in neither case can there be any question of a conscious change in writing direction to express a religious difference.

131 The next case is a bit more complex:

> „Die buddhistisch gewordenen Mongolen vollzogen einen Bruch mit der bisherigen linksläufigen Richtung ihres ursprünglich den semitischen Syrern entlehnten Schriftsystems und führten eine rechtsläufige Front ein." (Mieses 1919: 325)
>
> "The Mongols, who had converted to Buddhism, broke with the previous right-to-left direction of their writing system, which they had originally borrowed from the Semitic Syrians, and introduced a left-to-right direction."

The Mongolian alphabet is written from top to bottom, column by column, from left to right. The change from the direction of writing commonly used in Semitic scripts (from right to left) already took place during the transition from the pre-predecessor of the Mongolian alphabet, the Sogdian script, to its predecessor, the Old Uyghur script (Coulmas 1996: 471–474, 512). In fact, they simply rotated the page by 90° counterclockwise[11] (Fig. 8) to conform to the top-to-bottom writing direction common in East Asia. However, the columns in Chinese, Japanese, and

[10] The same applies to the better-known case of the Greek alphabet, which is based on the right-to-left Phoenician script (Mieses 1919: 326–327): The oldest Greek texts were still written from left to right, but in the 6th century BCE, *boustrophedon* became the predominant form of writing, and finally, in the 5th century BCE, the current direction of writing was established.

[11] A clockwise rotation would have meant that one would have had to read from bottom to top.

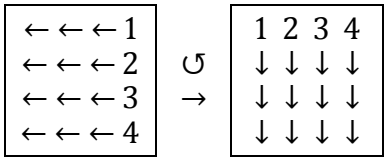

Fig. 8: Rotation from Sogdian to Old Uyghur

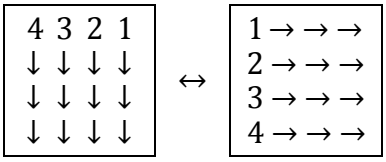

Fig. 9: East Asian writing directions

Korean texts are always arranged from right to left, so that a 90° counterclockwise rotation leads to 'European' left-to-right lines[11] (Fig. 9).

Consequently, the Uyghurs did not change the direction of their writing to left-to-right but rather rotated their pages so that they could read in columns from top to bottom (and more easily mix their writing with Chinese), and the arrangement of the columns reflects the original Semitic direction of writing, which they did not abandon.

7. Conclusion

Ultimately, therefore, nothing remains of Mieses's (1919: 325–342) theory about changes in writing direction due to religious schisms. Most of the world's scripts are written uniformly in one direction because this saves the reader the trouble of first having to determine the direction of writing.[12] Since there are four fundamentally equivalent options for determining a 🗋 132 uniform direction of writing (namely those shown in Figs. 8 and 9),[13] it is statistically very likely that different religious communities, if they use different scripts, will also have different directions of writing. And, of course, there are many cases in which a change of religion leads to a change in the writing system. However, there is no evidence anywhere of a religiously motivated change in writing direction alone.

Although this falsification does not add any new content to our knowledge of the linguistics of writing, it does at least provide a negative insight. In a typology

[12] This is usually less of a problem than it seems to us because we are not used to it. It is comparable to the problem of deciding in Serbia or Montenegro whether a sign should be read in Cyrillic or Latin (cf. Bunčić et al. 2016: 243–246): if it makes no sense, try the other script. In addition, in most cases the orientation of the characters helps. In Egyptian hieroglyphs, for example, all anthropoid and zoomorphic characters face against the direction of writing, i.e. toward the reader.

[13] Other writing directions would mean writing from bottom to top, which is virtually unheard of. The only exceptions I know of are the Irish Ogham alphabet and the as yet undeciphered Easter Island script Rongorongo (if it is a script in the full sense of the word). Both are closely linked to the writing material: Ogham was initially carved exclusively on memorial stones, along the vertical edges, starting at the bottom left and, if the edge was not long enough, continuing down the right edge; the classic Rongorongo documents are wooden tablets inscribed in a special type of *boustrophedon*, where the text begins at the bottom left and the tablet was rotated 180° after each line (Horley 2009: 251).

of biscriptality (Bunčić et al. 2016), therefore, in addition to the levels of script (Latin, Cyrillic, Chinese, etc.), glyphic script variants (e.g. Old Cyrillic, Bosančica, Graždanka, etc. within the Cyrillic script), and orthography, the direction of writing does not have to be taken into account as a further independent factor.

Perhaps even more important is the insight into how easily facts can be overinterpreted and that it is always worthwhile to re-examine a claim found in the scientific literature more closely—especially if it fits well into an otherwise sensible concept.

References

Bunčić, Daniel. 2016. *Biscriptality: A sociolinguistic typology*. Edited by Daniel Bunčić & Sandra L. Lippert & Achim Rabus. With contributions by Anastasia Antipova, Carmen Brandt, Ekaterina Kislova, Henning Klöter, Alexandra von Lieven, Sandra L. Lippert, Helma Pasch, Achim Rabus, Jürgen Spitzmüller, and Constanze Weth. Heidelberg: Winter.

Čerepnin, Lev V. 1956. *Russkaja paleografija.* Moskva: Gosudarstvennoe izdatel'stvo političeskoj literatury.

Coulmas, Florian. 1996. *The Blackwell Encyclopedia of Writing Systems*. Cambridge (MA): Blackwell.

Durmišević, Mirsad. 2011. Nekropole stećaka u Rogatici. *Rogatica – BiH*, 22 Jan 2011. https://rogatica-bih.blogspot.de/2011/01/nekropole-stecaka-u-rogatici.html (last checked 2 Apr 2026).

Eco, Umberto. 2012. *The Prague Cemetery*. Translated by Richard Dixon. London: Vintage Books.

Hercigonja, Eduard. 2006. *Tropismena i trojezična kultura hrvatskoga srednjovjekovlja.* 2nd edn. Zagreb: Matica Hrvatska. [1st edn. 1994.]

Hoernes, Moriz. 1888. *Dinarische Wanderungen: Cultur- und Landschaftsbilder aus Bosnien und der Hercegovina.* Wien: Carl Graeser.

Horley, Paul. 2009. Rongorongo script: Carving techniques and scribal corrections. *Journal de la Société des Océanistes* 129. 249–261.

Karskij, Evfimij F. 1928. *Slavjanskaja kirillovskaja paleografija.* Leningrad: Akademija Nauk SSSR. [Reprint: Leipzig 1972.]

Mazalić, Đoko. 1941. Borač, bosanski dvor srednjeg vieka. *Glasnik hrvatskih zemaljskih muzeja u Sarajevu* 53. 31–94.

Malcolm, Noel. 1994. *Bosnia: A short history.* London: Macmillan.

Mieses, Matthias. 1919. *Die Gesetze der Schriftgeschichte: Konfession und Schrift im Leben der Völker.* Wien, Leipzig: Braumüller.

Miletić, Maja. 1957. *I "Krstjani" di Bosnia alla luce dei loro monumenti di pietra.* Roma: Pontificium Institutum Orientalium Studiorum.

Nosić, Milan. 1995. Humačka ploča. *Riječ: Časopis za slavensku filologiju* 1(1–2). 33–41.

Panonian. 2010. Bosnia around 1412. In *Wikimedia Commons.* https://w.wiki/KgGp (last checked 2 Apr 2026).

Peščak, Marija M. (ed.). 1974. *Hramoty XIV st.* Kyïv: Naukova dumka.

Picchio, Riccardo. 1958. La "Istorija slavěnobolgarskaja" sullo sfondo linguistico-culturale della Slavia ortodossa. *Ricerche slavistiche* 6. 103–118.

Renner, Heinrich. 1897. *Durch Bosnien und die Hercegovina: kreuz und quer; Wanderungen.* 2nd edn. Berlin: Reimer.

Ščepkin, Vjačeslav N. 1999. *Russkaja paleografija.* 3rd edn. Moskva: Aspekt Press. [1st edn. 1918.]

Sobolevskij, Aleksej I. 2007. *Slavjano-russkaja paleografija: Kurs vtoroj.* 5th edn. Moskva: LKI. [1st edn. 1902.]

Speranskij, Mixail N. 1929. *Tajnopis' v jugo-slavjanskix i russkix pamjatnikax pis'ma.* Leningrad: Izdatel'stvo Akademii Nauk SSSR. [Reprint: Moskva 2011.]

Truhelka, Ćiro. 1889. Bosančica: Prinos bosanskoj paleografiji. *Glasnik Zemaljskog muzeja u Bosni i Hercegovini* 1(4). 65–74.

Truhelka, Ćiro. 1908. Crtice iz srednjeg vijeka (Na osnovu dubrovačkih arhivalija): I. Vlatko Vlagjević. *Glasnik Zemaljskog muzeja u Bosni i Hercegovini* 20. 419–423.

UNESCO. 2016. Five sites inscribed on World Heritage List. *UNESCO World Heritage Centre*, 15 Jul 2016. https://whc.unesco.org/en/news/1525 (last checked 2 Apr 2026).

Vego, Marko. 1962. *Zbornik srednjovjekovnih natpisa Bosne i Hercegovine.* Vol. 1. Sarajevo: Zemaljski muzej.

Vego, Marko. 1981. Novi i revidirani srednjovjekovni natpisi iz Bosne. *Naše starine: Godišnjak Zavoda za zaštitu spomenika kulture prirodnih znamenitosti i rijetkosti Bosne i Hercegovine* 14–15. 39–64.

Wilde, Danny et al. 2005–2016. Horizontal and vertical writing in East Asian scripts. In *Wikipedia: The free encyclopedia.* https://en.wikipedia.org/w/?oldid=731118466 (last checked 2 Apr 2026).

Zaliznjak, Andrej A. 2004. *Drevnenovgorodskij dialekt.* 2nd edn. Moskva: Jazyki slavjanskoj kul'tury. [1st edn. 1995.]

This volume contains English translations of seven papers previously published in German and two papers previously published in Russian on various subjects of Slavic Linguistics:

Criteria for determining the degree of 'danger' associated with **false friends** • Intralingual and crosslinguistic **enantiosemy** as a communicative problem • The integration of **foreign words** from European languages into the Cyrillic and Latin **alphabets** • The (re)nationalization of the **Serbo-Croatian standards** • On the usefulness of Google Books & co. for **histories of concepts** (not only in the Slavic Studies): *glasnost'*, *standard language*, and *digraphia/bigraphism* • **Biaspectual verbs** as polysemes: On homonymy, aspectual neutrality, and the conative reading • **Definiteness** as a 'covert category' in Russian? • **'Diastratic diglossia'** in 18th-century Russia, *or:* When did **Church Slavonic** become a foreign language? • **Right-to-left Cyrillic** among the Bogomils?