Balcania et Slavia

Studi linguistici | Studies in Linguistics

Vol. 1 – Num. 1
June 2021

Edizioni
Ca'Foscari

Balcania et Slavia

Studi linguistici | Studies in Linguistics

Edizioni Ca' Foscari - Digital Publishing
Fondazione Università Ca' Foscari
Dorsoduro 3246, 30123 Venezia
URL https://edizionicafoscari.unive.it/en/edizioni4/riviste/balcania-et-slavia/

Balcania et Slavia
Studi linguistici | Studies in Linguistics

Semestral journal

Head office Università Ca' Foscari Venezia | Dipartimento di Studi Linguistici e Culturali Comparati | Ca' Bernardo | Dorsoduro 3199, 30123 Venezia, Italia | balcaniaetslavia@unive.it

Publisher Edizioni Ca' Foscari - Digital Publishing | Fondazione Università Ca' Foscari | Dorsoduro 3246, 30123 Venezia, Italia | ecf@unive.it

Certificazione scientifica delle Opere pubblicate da Edizioni Ca' Foscari - Digital Publishing: tutti i saggi pubblicati hanno ottenuto il parere favorevole da parte di valutatori esperti della materia, attraverso un processo di revisione anonima sotto la responsabilità del Comitato scientifico della rivista. La valutazione è stata condotta in aderenza ai criteri scientifici ed editoriali di Edizioni Ca' Foscari.
Scientific certification of the works published by Edizioni Ca' Foscari - Digital Publishing: all essays published in this issue have received a favourable opinion by subject-matter experts, through an anonymous peer review process under the responsibility of the Advisory Board of the journal. The evaluations were conducted in adherence to the scientific and editorial criteria established by Edizioni Ca' Foscari.

Balcania et Slavia
Vol. 1 – Num. 1 – June 2021

Table of Contents

Introduction
Iliyana Krapova, Svetlana Nistratova,
Luisa Ruvoletto, Giuseppina Turano 5

No Gender in 'Gender Agreement':
On Declension Classes and Gender in Serbo-Croatian
Boban Arsenijević 11

It's All Under Control!
On Perfective Present Forms in BCS Main Clauses
Marco Biasio 47

On the Diachrony of the Clitic Cluster in Bulgarian
Tsvetana Dimitrova 67

Postposed Articles
and DP Structures in Torlak
Jelena Živojinović, Beatrice Azzolina, Veronica Girolami 97

Language Ideologies and Multilingual Practices
of Post-Soviet Migrants in Western Europe
from a Translanguaging Perspective
Erzhen Khilkhanova 117

Balcania et Slavia
Vol. 1 – Num. 1 – June 2021

Introduction

Iliyana Krapova
Svetlana Nistratova
Luisa Ruvoletto
Giuseppina Turano
Università Ca' Foscari Venezia, Italia

Balcania et Slavia. Studies in linguistics is published by the Department of Linguistic and Comparative Cultural Studies at Ca' Foscari University of Venice. The primary goal of the Journal is to bring together original publications in the field of theoretical, areal-typological, contrastive, and diachronic linguistics, which can provide a better understanding of the structure and the history of the Slavic and Balkan languages. In the academic landscape of Italy, the Journal promises to become a unique virtual domain of interdisciplinary research in which the Slavic and the Balkan linguistic areas are put to scrutiny so that they can feed off each other, in both empirical and theoretical terms. As editors, our aim is to publish contributions that bring to light different typological, areal or structural properties of these languages especially if such properties have remained unnoticed in current traditions or even if, however well-known, they call for innovative approaches in order to be better understood and analysed.

Issues of variation and micro-variation will be particularly relevant for the overall subject matter of the Journal because we believe that it is through a comparative description of the variation patterns in the various domains of language that we can gain a deeper view into the comparative grammar of the languages under study. This goes without saying for the well-known multilinguistic and multiethnic Balkan area that, due to the tormented Balkan history made of migrations, population mix and failed integrational policies, presents, as many scholars have pointed out, a unique dialectal continuum with patterns of isogloss variation that have not yet received the attention they merit.

The languages of the Balkan Linguistic Union, which comprises Balkan Slavic, Balkan Romance, Greek and Albanian, have come to develop, over the past six or seven centuries, a number of shared grammatical (morphosyntactic) and lexical features currently known as 'Balkan linguistic types'. But, as acknowledged by many scholars, 'Balkan linguistic types' are much more than a random collection of commonly acquired features (loan constructions, calques, syntactic borrowings, etc.) as it is generally the case with contact-induced innovations among two or more neighbouring dialects or languages. Balkanisms are deeply integrated into the structure of some or all Balkan languages, and the specific *Sprachbund* effects they have produced go beyond a simple areal explanation and raise a number of questions that are still open (for discussion, see the collection of articles in Krapova, Joseph 2018): the specific role of contact in bringing about the convergencies, the nature of the convergence itself (grammatical replication, some sort of copying or borrowing of surface structures), the mechanisms of transfer, etc. Another aspect worth exploring is the relation between the Balkan *Sprachbund* and other well-known linguistic areas including Standard Average European (SAE) conceived as a linguistic unity from a broad European perspective (Haspelmath 2001). Old questions relevant for the description of specific Balkan convergencies are also still open for discussion: the degree of Balkanisation of each language spoken in the Balkans; the potential source(s) of each common phenomenon; the number and fate of each Balkanism, in historically reconstructed, as well as in ongoing contact situations. At the current stage of Balkan linguistics, we need to know whether these processes are still going on, in which fields or subfields of grammar, and whether there are significant outcomes of these processes in recent times (see the collection of papers in Sobolev 2021, among many others not cited here for lack of space).

The problem of shared or common linguistic types is particularly relevant also for the Slavic area. A traditional concern in Slavic linguistics has been the issue of divergence in diachronic terms, namely in what ways Slavic languages have evolved in space and time as individual linguistic entities, with some attention also paid to parallel (and non-contact-induced) developments motivated by genetic origin. Recently, however, new perspectives have emerged and have enriched the traditional historic interest, namely those of convergence due to language contact, both intra-Slavic as well as between Slavic and the rest of Europe. There is a renewed interest toward a more profound comparative-typological profile of the Slavic languages in terms of their potential SAE membership, as part of the core or the periphery (see the recent collection of papers in Danylenko, Nomachi 2019). Our Journal seeks to promote original studies aiming at discussing those and similar vitally important issues that will undoubtedly help delineate the Slavic area and define its essential

characteristics in reference to the 'other' Europe. We intend to publish innovative research on all topics regarding the diffusion of certain areal properties, grammaticalization in space and time, conservative and innovative tendencies all of which have contributed in one way or another to shape the current Slavic landscape.

Balcania et Slavia is an online peer-reviewed journal, published annually in two issues and available in open access. For our inaugural volume, we placed a call for papers and we also invited some scholars to submit original contributions that present different methodological approaches to different fields of Slavic and Balkan linguistics: phonology, morphology, syntax, pragmatics, and sociolinguistics.

Here is a brief synopsis of the papers included in the first issue and appearing in alphabetical order.

Boban Arsenijević's paper "No Gender in Gender Agreement: On Declension Classes and Gender in Serbo-croatian" looks into gender agreement in Serbo-Croatian (SC) as a theoretical problem for this language, but also for theories of how agreement mechanisms work in general. The problem posed by SC gender agreement data is that in certain cases two different gender values are triggered by the controlling noun. In the face of these potentially problematic data, the author seeks to disentangle the agreement feature into primitives like [animate] and [human] and proposes a model of how the features compose with each other and interact both with system features (like declension class) as well as with semantic ones. The most important and challenging property of agreement according to the author and his generative-driven approach to the complexities of gender morphology is that gender does not play any role in syntax but gets instantiated only at the level of logical form (LF).

The paper by Marco Biasio "It's All Under Control! On Perfective Present Forms in BCS Main Clauses" aims at providing a unified analysis of the syntax-pragmatics interface of the (allegedly) anomalous licensing of the perfective present (henceforth $Pres^{PF}$) in Bosnian-Croatian-Serbian present-tensed main clauses. In Old- and Neo-Štokavian dialects, $Pres^{PF}$ forms cannot refer to eventualities that are anchored to the utterance time. However, there seem to be three apparent exceptions to this structural constraint. They comprise: 1) *abusive metonymic performatives*, i.e. a closed class of disparaging phrases that figuratively perform on the insulted hearer the action referred to in the utterance; 2) running commentaries; 3) contexts featuring the epistemic operator *mòžda* 'maybe'. Building on some theoretical tenets advocated by neoperformative hypotheses, it is claimed in the paper that, for $Pres^{PF}$ forms to be licensed in BCS main clauses, the uninterpretable feature of control needs to be realised as a functional projection in the SpeechActP layer at the level of the so-called 'Seat of Knowledge'.

Erzhen Khilkhanova's paper "Language Ideologies and Multilingual Practices of Post-Soviet Migrants from a Translanguaging Perspective" is dedicated to translingualism, which is a popular topic that attracts the attention of linguists in many countries for a number of reasons. Within the Slavic area, the increase of interest is motivated by migration phenomena and issues of resettlement of former Soviet citizens in the West. The paper describes, discusses and analyses multilingual practices of non-Russian migrants from the former Soviet Union whose native language is Russian. The translanguaging perspective uncovers language ideologies underpinning these practices and serving as the main motivation for their replacement with linguistic assimilation despite the rich repertoire of multilingual resources available to these first-generation migrants.

The study also reveals a change of language ideologies of post-Soviet migrants in the more democratic Western European context towards awareness of language equality and increasing symbolic value of languages for ethnic identification.

In their paper, "Postposed Articles and DP Structures in Torlak", Jelena Zivojinović, Beatrice Azzolina and Veronica Girolami investigate the enclitic article in the DP structure of Torlak, a non-standard Balkan Slavic variety spoken in the Southeastern area of Serbia and in the bordering areas of Bulgaria and Macedonia.

Torlak displays both Balkan and non-Balkan features: like Bulgarian and Macedonian, it presents a postposed enclitic article-like element, but, at the same time, it does not exhibit multiple determination, which in Bulgarian and Macedonian is characterised by the presence of a demonstrative and one or more definite article suffixes within the extended projection of the noun.

The authors' proposal is that, while in Macedonian and Bulgarian the articles seem to have undergone a full grammaticalization into purely functional elements and for this reason they are able to occur with a demonstrative in a multiple determination construction, in Torlak instead, the grammaticalization of the determiner is only partially resulting in an inflectional affix that maintains the demonstrative semantics.

Tsvetana Dimitrova's paper "On the Diachrony of the Clitic Cluster in Bulgarian" traces back the formation of the clitic cluster in Bulgarian starting from the Old Church Slavonic through Middle Bulgarian up to the Early Modern Bulgarian and beyond. The author proposes that the clitic cluster is split in two layers – the main layer consists of a (pronominal) core and a (verbal) periphery, while the secondary layer hosts elements that are not strictly clitic but are clitic-like or semi-clitics. The author reviews a number of data in various diachronic corpora and reaches the conclusion that there has been practically no change in the positions of the elements within the core, while the changes that have occurred have interested the

periphery of the cluster and are due not to structural changes but to changes in the set, i.e. some elements of the pronominal and the auxiliary system have been reanalysed while original clitics of the discourse type have been lost.

We hope that this Journal will become an open forum for interesting discussions where researchers and experts in the field share their work, stimulate novel research paths and advance the study on the various aspects of the Slavic and the Balkan languages.

We are grateful to the members of our Advisory Board for the international prestige they confer to the Journal and for their precious suggestions.

Many thanks go to the Editorial Board members who actively participated in the entire process.

We must also highlight the fundamental work done by our reviewers who are among the most reputable scholars worldwide. They ensured a double-blind peer-review process – with two scholars for each paper – and we would like to thank them all for their efforts and time.

Last but not least, we would like to thank our publisher Edizioni Ca' Foscari, directed by Massimiliano Vianello, for the constant support. Our special thanks go to Mariateresa Sala for her great efficiency and invaluable editorial help.

Bibliography

Haspelmath, M. (1998). "How Young in Standard Average European?". *Language Science*, 20(3), 271-87.

Haspelmath, M. (2001). "The European Linguistic Area: Standard Average European". Haspelmath, M.; König, E.; Oesterreichter, W.; Raible, W. (eds), *Language Typology and Language Universals: An International Handbook*, vol 2, 1492-510. New York: Mouton de Gruyter. Handbuch zur Sprach- und Kominkationswissenschaft 20.2.

Danylenko, A.; Nomachi, M. (2019). *Slavic on the Language Map of Europe*. Mouton de Gruyter. Trends in Linguistics. Studies and Monographs [TiLSM] 333. https://doi.org/10.1515/9783110639223.

Krapova, I.; Joseph, B. (2019). *Balkan Syntax and (Universal) Principles of Grammar.* Mouton de Gruyter. Trends in Linguistics. Studies and Monographs [TiLSM] 285. https://doi.org/10.1515/9783110375930.

Sobolev, A. (2021). *Between Separation and Symbiosis. Southeastern European Languages and Cultures in Contact*. Mouton de Gruyter. Language Contact and Bilingualism. https://doi.org/10.1515/9781501509254–202.

Balcania et Slavia
Vol. 1 – Num. 1 – June 2021

No Gender in 'Gender Agreement': On Declension Classes and Gender in Serbo-Croatian

Boban Arsenijević
Institute of Slavic Studies, University of Graz, Austria

Abstract The present paper argues for a view of gender agreement without either grammatical or natural gender being represented as syntactic features. Rather than deriving declension classes in terms of realisation, I postulate them as the only relevant feature that is lexically specified on the noun. Agreement copies the declension class and triggers presuppositions. When these presuppositions clash with those already active in the discourse, default agreement is realised. The paper moreover provides a quantitative analysis of semantic correlates of declension classes and a novel analysis of SC declension classes.

Keywords Declension class. Gender. Agreement. Properties of quantity. Serbo-Croatian.

Summary 1 Introduction. – 2 Declension class in terms of realisation. – 3 Basics for a better model of declension classes. – 4 Declension Classes, Gender and Properties of Quantity: A Quantitative Corpus-based Analysis. – 5 The proposal: nouns are specified for declension classes, interpretive effects emerge in pragmatic competition. – 6 The morphology of declension classes. – 7 Conclusion.

Edizioni Ca'Foscari

Peer review

Submitted	2021-07-20
Accepted	2021-10-15
Published	2020-12-20

Open access

Citation Arsenijević, B. (2021). "No Gender in 'Gender Agreement': On Declension Classes and Gender in Serbo-Croatian". *Balcania et Slavia*, 1(1), 11-46.

1 Introduction

Serbo-Croatian (SC) nouns are traditionally divided in four declension classes and three genders (e.g. Stevanović 1989). Declension classes can be designated by a pairing of the nominative and genitive singular ending, for example: class <Ø, a>, class <o/e, a>, class <a, e>, class <Ø, i>. Traditionally, they are referred to by numbers in the given order, as class I, class II, class III and class IV declension, which is also the way I refer to them in this article.[1] This is illustrated in **table 1** for the nouns *mrav* 'ant', *žig* 'stamp' (to illustrate the effect of animacy), *more* 'sea', *selo* 'village' (*-e* comes after soft consonants, *-o* otherwise), *riba* 'fish' and *noć* 'night'.

Table 1 Declension class and gender in Serbo-Croatian

	Class I (<Ø, a>)		**Class II (<o/e, a>)**		**Class III (<a, e>)**		**Class IV (<Ø, i>)**	
	Sg	**Pl**	**Sg**	**Pl**	**Sg**	**Pl**	**Sg**	**Pl**
Nom	mrav-Ø	mrav-i	mor-e(sel-o)	mor-a	rib-a	rib-e	noć-Ø	noć-i
Gen	mrav-a	mrav-a	mor-a	mor-a	rib-e	rib-a	noć-i	noć-i
Dat	mrav-u	mrav-ima	mor-u	mor-ima	rib-i	rib-ama	noć-i	noć-ima
Acc	mrav-(a)(žig-Ø)	mrav-e	mor-e(sel-o)	mor-a	rib-u	rib-e	noć-Ø	noć-i
Inst	mrav-em	mrav-ima	mor-em(sel-om)	mor-ima	rib-om	rib-ama	noć-ju	noć-ima
Loc	mrav-u	mrav-ima	mor-u	mor-ima	rib-i	rib-ama	noć-i	noć-ima

The relation between SC declension classes and gender has been the subject of numerous investigations.[2] Grammatical gender stands for the type of agreement a noun triggers on declinable modifiers and predicates, and relies on the assumption that the agreement markers on these items correlate with the declension class of the noun, as well as with the natural gender of the referent, as illustrated in (1). Neither correlation is absolute: a noun can show an agreement pattern that fails to match the natural gender of the referent or the declension class of the controlling noun, but not both (see (1d) for a mismatch with the natural gender).

1 Due to a high degree of syncretism, the first and the second class, i.e. classes <Ø, a> and <o, a>, are sometimes taken to be one class, yielding a classification with three classes, one of which has two subclasses.

2 Stevanović 1989; Corbett 1982; Corbett 1991; Wechsler, Zlatić 2003; Despić 2017; Alsina, Arsenijević 2012a; 2012b; Arsenijević, Gračanin Yuksek 2016; Puškar 2017 among others.

(1)

a.	bel-a	krava	b.	bel-i	bik	c.	bel-o	tele	d. mlad-a gospoda
	white-F	cow		white-MASC	bull		white-N	calf	young-F gentlemen
	'white cow'			'white cow'			'white calf'		'young gentlemen'

Inanimate nouns are orthogonal to gender, and their agreement is fully predictable from their declension class: class I nouns trigger masculine agreement, class II nouns neuter, class III and IV nouns feminine. Furthermore, considering animate nouns, class I includes only masculine members (i.e. only nouns that denote male or gender-unspecified individuals), class II only neuter (i.e. those referring to offspring, which is arguably conceptualised as genderless), class IV takes no animates at all.[3] This all suggests that gender could be eliminated from the system, by modelling agreement purely in terms of declension class: masculine gender agreement is class I agreement, neuter is class II, feminine class III and IV. The ground for relying on gender comes from the so-called hybrid gender agreement nouns, such as class III nouns triggering masculine agreement in singular as in (2), where the agreement in the singular is not the one indicated by the declension class, but arguably associated with the gender of the referents (Corbett 1991; Aronoff 1994).

(2)

moj-Ø	ded-a		prv-i	komšij-a		mlad-Ø	delij-a				
my-MASC	grandpa-III.NOM.SG		first-MASC	neighbour-III.NOM.SG		young-MASC	dude-III.NOM.SG				
'my grandpa'			'neighbour next door'			'young dude'					
moj-e	ded-e		prv-e	komšij-e		mlad-e	delij-e				
my-FEM.PL	grandpa-III.NOM.PL		first-MASC.PL	neighbour-III.NOM.PL		young-MASC.PL	dude-III.NOM.PLG				
'my grandpas'			'neighbours next door'			'young dudes'					

This step, however, introduces the question of the division of work between the two sources of information. Is agreement determined by the declension class, unless it clashes with gender, or is it determined by gender, unless a noun lacks it – in which case declension class jumps in? Or do actually all nouns, animate or not, have a gender feature, which remains uninterpreted in inanimates?

Before going deeper with the discussion, let me address some apparent exceptions to the generalisations above about the gender – de-

3 Prescriptive grammars mention nouns *kći* 'daughter' and *mati* 'mother' as class IV animates but, in most present day spoken varieties of SC, these nouns have migrated to class III, with a more or less changed shape.

clension class mapping, pointed by an anonymous reviewer. One pattern is that exhibited by two types of nouns which show endings that formally may be attributed to class II, but trigger the traditional masculine agreement. The first group are animate nouns, including proper names and derived nouns (mostly with the suffix *-ko* for expressive nouns denoting property bearers), as illustrated in (3).

(3)

a. Jan-k-o	b. Per-o	c. prljav-k-o	d. rumen-k-o
Jan-SUFF-?.NOM.SG	Per-?.NOM.SG	dirty-SUFF-?.NOM.SG	ruddy-SUFF-?.NOM.SG
'Janko'	'Pero'	'dirty little one'	'ruddy little one'
Jan-k-a	Per-a	prljav-k-a	rumen-k-a
Jan-SUFF-?.ACC.SG	Per-?.ACC.SG	dirty-SUFF-?.ACC.SG	ruddy-SUFF-?.ACC.SG
'Janko'	'Pero'	'dirty little one'	'ruddy little one'

In spite of sharing some endings with class II, these nouns rather belong to class I, since their accusative case in singular takes the ending as in class I rather than being syncretic with the nominative – as characteristic of class II nouns (see table 1). The ending *-o* in the nominative singular can be explained on a phonological ground, in the interaction of prosody and stem-final consonant clusters. In favour of the view that these nouns have an exceptional nominative singular, but otherwise belong to class I, speaks also the fact that for a subset of these nouns (those like the name *Pero* or the common noun *medo* 'little bear') in some varieties, people share the nominative singular in *-o*, but otherwise decline them in class III.

Two properties of the nouns in (3), however, point in the direction of class II. One is that they cannot have plural forms, which is a property they share with uncontroversial animate class II nouns (Arsenijević 2017). The other is that they have a hypocoristic, hence expressive meaning and, as discussed in § 5, expressive meanings are a type of markedness that triggers the assignment of declension classes of a higher degree of markedness (especially class II and III). The difference between the nominative and the accusative (not attested in uncontroversial class II animates) would then be an argument in favour of a realisational nature of declension classes. Still, this would not be in clash with the main thesis of this article: that grammatical gender is a derivative of declension classes and not vice versa. If declension classes too are derivatives of lexical semantic properties of the bases (i.e. of roots or complex structures), in particular properties of quantity, natural gender and expressiveness, this is a further step in theoretical reduction.

The second group are borrowed nouns ending in *-o* or *-e*, some of which are inanimate and therefore cannot be tested for the accusative singular, as in (4a-b) (in inanimate class I nouns, this form is syncretic with the nominative, just like in class II nouns). Still, looking at the animate among these nouns, as in (4c-d), they do show the class I pattern in the accusative (that they belong to the same class is confirmed by sharing the same pattern with accented final syllables, as in (4b, d); moreover, accented final syllables are not attested in class II). Furthermore, for most speakers, the nominative plural ending for all these nouns is the one of class I. While there are some speakers that prefer the class II ending, importantly, most of them also prefer the neuter agreement for these nouns, indicating that in their grammar they belong to class II.

(4)

a. radi-o	b. kabar-è	c. impresari-o	d. krupij-è
radio-?.NOM.SG	cabaret-?.NOM.SG	impresario-?.NOM.SG	croupier-?.NOM.SG
'radio'	'cabaret'	'impresario'	'croupier'
radi-a	kabar-è	impresari-a	krupijè -a*
radio-?.ACC.SG	cabaret-?.ACC.SG	impresario-?.ACC.SG	croupier-?.ACC.SG
'radio'	'cabaret'	'impresario'	'croupier'
radi-i	kabarè -i	impresari-i	krupijè -i
radio-?.NOM.PL	cabaret-?.NOM.PL	impresario-?.NOM.PL	croupier-?.NOM.PL
'radio'	'cabaret'	'impresario'	'croupier'

* Nouns of this type with an accented final syllable preserve this syllable as part of the stem in declension.

For speakers who agree them in masculine gender, all these nouns are hence class I nouns with exceptional endings in the nominative singular – conditioned by phonology and for some of them also the shape in the donor language, rather than being members of class II.

The other seemingly exceptional pattern pointed out by the reviewer are class II and IV nouns used for animate individuals, as illustrated in (5). Arguably, however, these nouns do not have animate denotation. They denote abstract notions, likely tokens of properties. Their animate use is either metaphorical, i.e. metonymic, or idiomatic – in both cases irrelevant for the productive properties of the declension classes (Moltmann 2004; Villalba 2009; Arsenijević 2012), and the use for animates, or humans (if at all they ever denote animate individuals, since for instance none of these nouns can be used referentially, e.g. as an answer to the question *Who told you that?* – see Arsenijević et al 2021 for a discussion and experimental support regarding this kind of non-referential nouns in class III).

(5)

a. lik-n-ost-Ø	b. velik-an-stv-o	c. čud-ov-išt-e
person-ADJ-N-IV.NOM.SG	great-ADJ-N-II.NOM.SG	marvel-ADJ-AUGM.N-II.NOM.SG
'personality'	'majesty'	'monster'

Back to the issue of gender agreement, different analyses have been proposed for it – whether specifically for SC, or for other Slavic languages (such as variation in patterns of agreement available regarding the mapping between gender and declension classes is limited, analyses for one Slavic language with declension are quite straightforwardly adapted for another). Corbett (1991), Aronoff (1994), Wechsler and Zlatić (2003) among others postulate two independent features for the two sources of information: each noun is specified for both gender and declension class (where the former may be even further split), and grammatical operations and hierarchies decide about agreement. Others, like Müller (2004), Puškar (2017) or Caha (2021) propose to fully discard the declension classes as syntactic features or diacritics, and derive them from other features. While Müller (2004) does not identify the particular features using formal labels instead (F_1, F_2...), Puškar (2017) indicates them to centrally involve gender, and Caha (2021) explicates a set of four relevant features: class, feminine, individuated, and a fully unmarked case feature. A significant advantage of the latter type of accounts is that they can explain regularities about intra- and interdeclension syncretism.

Alternative approaches, such as Hachem (2015), Fassi Fehri (2018), Arsenijević (2017), argue that declension classes as well as agreement reflect properties of quantity (which include animacy and gender as markers of a high degree of individuation, see e.g. Matasović 2004), rather than gender alone. Particularly relevant for the present paper is Corbett's (1982) point that grammatical gender can be derived from the declension class.

The central goal of this paper is to develop an explicit reductionist analysis of the relation between gender and properties of quantity on the one hand, and declension classes on the other, where it is the declension class that is represented by a feature, with gender effects deriving from it, rather than the other way around as previously argued. Quantitative as well as theoretical arguments are provided in favour of the proposed view.

The paper is organised as follows. § 2 outlines the main tendencies of the previous accounts for declension classes in SC and other Slavic languages. § 3 discusses the desiderata for an analysis and where previous analyses failed to achieve them. § 4 lays out the quantitative data obtained from an annotated database of the 4,718 most frequent nouns in SC. § 5 presents a novel analysis on which declension class is lexically assigned and occurs in syntax as a feature, and ef-

fects on gender and properties of quantity associated with declension classes emerge via pragmatic competition. § 6 proposes a novel morphological analysis of the SC declension class system on which class I lacks endings all together, and is entirely realised in terms of syncretism with other classes. The other three classes are ordered according to the balance between the role played by the theme vowel and case endings, and their tendency to be realised overtly.

2 Declension Class in Terms of Realisation

Realisational approaches to morphology, such as Distributed Morphology (Halle, Marantz 1993) and Nanosyntax (Starke 2010), take syntax to operate on features (and roots) only, and resolves the problems of morphology in an interaction between syntactic structures, the lexicon and the phonology, at their interface. Hence, instead of postulating both gender and declension class as two separate features which enter grammatical interactions, these approaches offer an alternative where declension class is a matter of realisation sensitive to gender. Indeed, Puškar (2017) indicates that declension class is best modelled in terms of gender-sensitive realisation – unfortunately without developing a concrete proposal. Privizentseva (2020) offers a more concrete analysis for Russian – but she simply stipulates an opaque formal feature which together with the feature of gender determines the realisation of case. The analysis hence merely formalises the view that declension class is a function of gender, but does not pinpoint the exact additional component that completes the specification of this mapping.[4] The only attempt to provide a transparent (in the sense that concrete, independently attestable features are used) and explicit realisational account of declension classes in Slavic languages is Caha (2021) with a model of Russian declension classes. In what follows I will give only a brief assessment of his analysis, trying to minimise the unavoidable injustice to the presented work enforced by spatial limitations.

Caha outlines the entire path of reasoning from having the relevant functional features realised by roots (or more broadly nominal stems), via a view where the roots realise a relatively small functional sequence, while the features relevant for declension classes are realised by the very endings, to a hybrid, where in some case forms the root realises a long span, while in others it shrinks and the end-

4 Similar in spirit is Müller's analysis (2004), where two opaque features are introduced to capture the syncretism between Russian declension classes. However, this analysis has no ambition to derive declension class from gender via realisation, but rather considers the two features that determine it independently present in the structure.

ing realises the relevant projections.

For illustration, the singular case forms of the Russian noun *zavod* 'factory' receive realisations as in (6) (Caha's (101)). In the nominative and the accusative form, the root realises the lexical content, the features REF(ERENCE), CLASS, IND(IVIDUATED), and the formal features combining to determine the particular cases. In the remaining case forms, the root shrinks, and the ending realises the features CLASS, IND, next to the relevant case features.

(6) Re-elaboration of Caha's analysis of *zavod* 'factory'

	xNP	**REF**	**CLASS**	**IND**	**F1**	**F2**	**F3**	**F4**	**F5**	**F6**
NOM	zavod									
ACC	zavod									
GEN	zavod	a								
LOC	zavod	e								
DAT	zavod	u								
INS	zavod	om								

The costs the analysis pays for its great reductionist capacity include treating all the members of one of the two bigger declension classes (the class in *-a* in the nominative singular) as idioms where the root is contextually specified to only occur in structures which include the feature FEM(ININE). Moreover, all these nouns involve mismatch between the LF and the PF, in the sense that their meaning matches the entire idiom, and their phonological realisation only the structure below FEM. Furthermore, as Caha explicitly admits, the feature FEM in his account has nothing to do with the feminine natural gender of the referent (otherwise it would not only make all the inanimate nouns in the class interpreted as female, but also those with animate male referents such as *d'ad'a* 'uncle' or the name *Ilya*). The natural gender is an independent factor, which is not syntactically represented, and competes for triggering agreement with the syntactic feature FEM. Recall that animate nouns with mismatches between the declension class and gender are the reason for postulating gender in grammar, in addition to declension classes. Once the feature is dissociated from the semantics of gender, the analysis is not transparent any more (i.e. the feature is no different from Privizintseva's 2020 additional opaque feature).

Caha considers different alternative analytic strategies, pointing out various problems that they face, and indicating that a realisational analysis of Russian declension classes superior to the one he proposes is not possible considering the facts. Yet it is questionable whether the sacrifices that his analysis makes are worth the goal which is achieved.

3 Basics for a Better Model of Declension Classes

The desiderata stated in § 1 include reduction and transparency. Reduction is achieved if one of the two phenomena, declension class or gender, is derived from the other. Transparency is achieved if the phenomenon that serves as the basis for the derivation of the other one is independently attested (i.e. in other domains of syntax, or in semantics or phonology). Transparency thus further contributes to reduction.

Gender is independently attestable in semantics, but only for animate nouns. Declension classes are attested in morphology, in the endings which different nouns take for the same case form (i.e. in the same syntactic contexts). A logical step for realisational approaches is to have the former represented as a syntactic feature (because it is LF interpretable), and derive the latter from it via some realisation rules (as realisational approaches generally do with morphological realisations).

For transparency, this relation needs to be total and systematic. There should be a set of features such that each declension class is characterised by a unique subset of those features. All the members of a declension class then should share this exact specification – they should not lack any of the features from the combination, nor bear any additional ones from the set.

In this paragraph, I consider the likely candidates for semantic differences and, based on a corpus investigation, argue that while tendencies are clearly observable, no sufficiently strong contrast can be identified. This leads to the conclusion that the declension class must be a syntactic primitive, i.e. in feature-based approaches: a feature or a combination of formal features whose direct interpretation is the specification of the set of endings realising case forms of the noun. To achieve reduction, then, the only option is to derive gender from declension classes, which is the path I take in § 2.

Another important observation, which previous analyses overlook, is that what is in the literature referred to as (grammatical) gender (as discussed above regarding Caha's view) is effectively also a choice between declension classes. As already outlined in § 1, grammatical gender is identified by the agreement patterns the noun triggers. More concretely, it is identified by the set of endings the agreeing item takes. There are altogether three such sets of endings, one that is described as masculine, one as feminine and the third as neuter. But not only, as pointed out in the discussion of Caha, these endings establish a correlation with gender in animate nouns – they also fit the definition of declension classes. The only difference is that agreeing items are lexical categories that are not restricted to one declension class – they can combine with all three of them. In other words, what is referred to as grammatical gender is a declen-

sion class too: masculine is the adjectival declension class I, neuter is class II and feminine class III (this is also reflected in the shared endings between nominal and adjectival declension classes or morphological components thereof). Approaches which derive declension classes from gender thus effectively derive the declension class of the controller from the declension class of the agreeing item. There is something paradoxical about it.

One big problem for predicting the declension class from the gender feature alone is quite obvious: declension classes III and IV are both described as bearing feminine gender (i.e. they trigger exactly the same agreement patterns). Only based on gender, it is hence impossible to predict which noun will realise its case by class III and which by class IV endings.

Obvious candidates for the features that conspire with declension class to determine agreement are those that Caha (2021) employs in his model. Caha's final account however does not make use of all the four features it includes – rather, features CLASS, FEM and IND always remain together, and may for the intents and purposes of his analysis be replaced by only one feature – or it could be only one of them that plays the relevant role.

Feature REF is supposed to be common for all nouns, or at least to have the same status in all of them (Baker 2003). Feature FEM basically stipulates agreement and effectively splits the set of declension classes in two (the masculine and the neuter declension class on the one hand, and the two feminine declension classes on the other). Features CLASS and IND, which are closely related to properties of quantity (such as count, mass, atomic), should then be responsible for distinguishing at least between the former and the latter two. As two features have the potential of specifying four different classes (absence of both, only CLAS, only IND, both CLASS and IND), the question emerges whether the remaining two features are needed at all.

Works like Hachem 2015, Fahri Fessi 2018 and Arsenijević 2017 follow exactly this path: they argue that the unit of counting (i.e. possibly its absence) and the type and degree of individuation are the properties behind not just declension classes, but actually also what is traditionally seen as gender. In this family of views, it is the finer division along these dimensions that plays a role: rather than reflecting mass vs. count denotations or a mere degree of individuation, the relevant features reflect particular kinds of mass, or different degrees of individuation. This view takes differences between non-atomised masses (*space*, *time*), atomised masses (*sand*, *powder*), collective denotations (*crew*, *clan*), vague count denotations (*ocean*, *field*, *plane*), together with the opposition between abstract and concrete denotations, animate, inanimate and human, or events, properties and individuals, as the base for the emergence of declension classes. Natural gender then presents only a fraction of the set of properties which

conspire to yield the traditional gender agreement and the division of nouns into declension classes cross-linguistically.

Arsenijević (2017) supports this view by data from SC showing that neuter nouns (which all also belong to declension class II) manifest a range of deficiencies regarding individuation. Examples as in (7) manifest that neuter agreement is not available to highly individuated plural uses such as in demonstratives used with pointing, animates generally or conjunction.

(7)

a. Ona *(deca) su neumorna. (pointing at children)
those.NEUT.PL child.NEUT.PL AUX.PL tireless.NEUT.PL
'Those children are tireless' / *'Those are tireless'.

b. dugm-e dugm-eta tel-e *tel-eta
button-NEUT.SG button-NEUT.PL calf-NEUT.SG calf-NEUT.PL
'button' 'buttons' 'calf'

c. Sel-o i polj-e su potopljen-i / *potopljen-a
village-NEUT.SG and field-NEUT.SG are.PL flooded-MASC.PL / flooded-NEUT.PL
'The village and the field are flooded'.

On this basis, Arsenijević argues that neuter corresponds to the absence of the classifier feature, i.e the absence of Landman's (2011) neat generating capacity (see also Arsenijević 2017), which blocks the formation of a proper plural. He claims that the form traditionally analysed as N.PL is rather a collective plural form (a weak division over a messy quantity structure in Landman's sense), unable to express highly individuated plural meanings.

In modelling the relation between gender, properties of quantity and declension classes, a quantitative insight into the structure of the nominal category in the language use is highly informative. Seeing what proportions of the nouns in each class are accounted by each of the approaches, and how many remain unaccounted for can give us a measure of plausibility of these accounts. This can help us both determine the minimal set of features needed to account for the facts, and the optimal direction of derivation, i.e. whether declension class should be derived from gender and properties of quantity, or it should be lexically specified, and display effects on the interpretations in the domain of gender and properties of quantity.

4 Declension Classes, Gender and Properties of Quantity: A Quantitative Corpus-Based Analysis

To empirically test the modelling options in the domain of declension classes, gender and properties of quantity, I performed a corpus-based research. The material was assembled by first excerpting the 5,000 most frequent SC nouns from the srWaC corpus (Ljubešić, Klubička 2016), and removing doubles and non-words, which reduced the material to a database of 4,718 nouns. The nouns were then annotated for the properties in **table 2** (the database is annotated for a significantly larger set of properties, but only the ones relevant for the discussion are given in the table).[5]

Table 2 Annotated properties of SC nouns in the database

Freq	Frequency (tokens per million words)
Class	Declension class (I-IV)
Gram. gender	What agreement the noun triggers (MASC, FEM, NEUT, HYB[rid])
Nat. gender	Lexical semantic restriction of the sex of the referent (NMASC, NFEM, NNEUT, Ø)
Anim	Animate (1, 0)
Count	Does the noun inflect for plural (1, 0)
Mass	Does the noun combine with mass quantifiers like *sve* 'all' (1, 0)
Coll	Does the Sg noun allow both Sg and Pl agreement (1, 0)
Group	Does the noun in the singular denote a group of entities (1, 0)
Name	Is it a proper name for any class of individuals (1, 0)
Suff	Does the noun have at least one suffix (1, 0)
Suff1	What is the first external (linearly last) suffix of the noun if it has any

Additionally, the mass nouns in the database have been annotated for the properties in **table 3**.

Table 3 Additional properties annotated only for mass nouns

Abstract	Is the noun abstract (1, 0)
Atomised	Does the denotation of the noun involve a level of minimal units (1, 0)
Atom-type	If atomised, what is the type of atoms (concrete, abstract, event, vague)
Mass-type	The ontological class (spatial, temporal, structure, domain, dimension, complex, assembly, substance, emotion, condition, capacity, property, event, vague count)

5 The database was annotated by two annotators, the author and Adisa Nanić, a master student and native speaker of SC. Inter-annotator disagreements were resolved by consulting other native speakers.

The database indicates significant differences between the sizes of declension classes. Class I and class III are the largest, with 2,259 and 1,563 member nouns, respectively. Class II is somewhat smaller, with 720 nouns, and class IV is very small with only 178 members. Moreover, it turned out that 71% of class II nouns are derived from the (predominantly deverbal) suffix *-VV.je* (Arsenijević 2010; Simonović, Arsenijević 2014), and 72% of class IV nouns from the deadjectival suffix *-ost*. There are hence only 51 nouns of class IV which are not derived from the suffix *-ost* in our sample. The proportion is probably even more drastic in the aggregate lexicon, as the suffix is highly productive, thus implying a large number of hapaxes.

Quantitative analyses have been performed for the relation of declension class with the properties Gender, Anim, Mass, as well as the within-mass properties Abstract, Atomised and Atom-type. The remaining annotated properties either did not show a significant effect, or their effect could be subsumed under that of one or a combination of the discussed properties.

Grammatical gender establishes a very tight match with declension classes. Class I systematically agrees with masculine, class II with neuter, class IV with feminine gender. The only locus of incomplete match lies in class III, which has a number of human members showing hybrid gender agreement: masculine in singular and feminine in plural (see (2)). In the sample of 4,718 nouns, with 1,563 class III nouns, 46 such nouns are identified (0.97% of the database and 2.94% of class III nouns). Grammatical gender is hence a very good predictor of declension class, and assuming that each gender value is represented as a different feature (combination), a realisational account based on it covers a very large portion of the data. In order to cover the small subset of class III nouns that trigger hybrid agreement, one first needs to see whether there is a feature that all of them share. Of all the properties annotated in the database, two are shared by all these nouns: they are all animate, and none of them is specified for feminine natural gender (i.e. 14 are male proper names, 8 are nouns denoting male roles such as *čika* 'uncle' or *papa* 'pope', and the remaining 24 are gender-general, like *sudija* 'judge' or *arhitekta* 'architect'). A realisational account is hence possible where +GMASC (i.e. grammatical masculine gender) triggers class I realisation, except in the context +GMASC +HUM-NFEM +SG (grammatical masculine, human, non-natural-feminine, singular), which triggers class III realisation. Inversely, grammatical gender can be derived from the combination of the declension class and the remaining features above (+HUM-NFEM +SG).

An additional problem for a realisational account of the declension class is that not just a majority of class III forms trigger feminine agreement, but also all the nouns in class IV. Hence, a noun specified for +G _ FEM is underspecified between class III and class

IV realisations. Yet, as the discussion of other properties in a quantitative perspective below shows, class III is open for all the values of all the considered properties, which means that there is no property which could be represented as a distinctive feature or a distinctive set of contexts for the two declension classes. Note that, if declension class is specified and gender agreement is realisational, no ambiguity arises as feminine agreement can be specified to be realised in both types of contexts: class IV and class III with the exception of III +HUM-NFEM+SG, where it is the masculine agreement that gets realised. Recalling further that the actual substance of gender agreement is declension class assigned to the agreeing item, it gets even more plausible to consider declension class of the noun as the primitive, rather than gender agreement.

Natural gender posed a problem for annotation, in there being different strengths of gender presuppositions. Among those nouns that qualify for it (all the animate nouns), some have a relatively weak presupposition, others moderate and yet others a strong one. Any test that was tried faced this issue, including the one eventually implemented, illustrated in (8).

(8)

a. ženski jaguar
female jaguar.I
'female jaguar'

b. ženski sudija
female judge.III
'female judge'

c. ?ženski papa
female pope.III
'female pope'

d. #ženski deda
female grandpa.III
'female grandpa'

e. ženka-jaguar
female-jaguar.I
'female-jaguar'

f. žena-sudija
woman-judge.III
'woman-judge'

g. ?žena-papa
woman-pope.III
'woman-pope'

h. #žena-deda
woman-grandpa.III
'woman-grandpa'

The problem for annotation actually only reflects the fact that the natural gender of a noun is not a discrete value, as it is strongly influenced by cultural constructs. Whether an animal will be prototypically represented as a male or a female, how immanent a property of a pope it is that it is a male, and how immanent it is for a grandpa – are not questions that render clear-cut categories. Generally, the degradation is stronger for humans than for animals, and for nouns with an established minimal pair in gender than for those without it. It is also typically stronger in singular than in plural. In the annotation, therefore, only human nouns were annotated for natural gender,[6] and it was specified for any noun that causes degradation

6 Very few animal-denoting nouns caused degradation in combination with the adjective *muški* 'male' (*svinja* 'pig', *krava* 'cow', *ovca* 'sheep', *koza* 'goat', *mačka* 'cat'), i.e.

when combined with the adjective *muški* 'male', i.e. in a (semi-)compound with *muškarac* 'man' (as +NFEM) or *ženski* 'female', i.e. *žena* 'woman' (as +NMASC), irrespective of the strength of degradation.

Even though in the annotation the mapping of this property onto discrete values was forced, looking only at the singular forms, the insights above suggest that natural gender cannot be a grammatical feature, as grammatical features are stable discrete formal objects, rather than continuous scalar ones, which moreover change depending on other features (number in this case).

There are two more reasons why natural gender is unlikely to be a feature. One is that it often emerges as a consequence of the existence of a minimal pair of lexical items specialised for the opposite genders (see also Merchant 2014; Sudo, Spathas 2019). This indicates that lexical competition in pragmatics is a possible source of natural gender, even if the lexical semantics does not necessarily include it (this is supported by the effect of number, as in plural, gender-general interpretation figures more prominently in the pragmatic competition, see also Puškar 2018; Mitić, Arsenijević 2019).

The other reason is the quantity of nouns carrying natural gender. In the entire database, 192 nouns are annotated as restricted regarding natural gender (139 to male and 53 to female referents). This makes 4.07% of the database. It is not just unusual, but also uneconomic that a system employs a feature marked for either of its two values on such a small fragment of the targeted lexical category. **Animate** nouns are most frequent within class I, with 32.64%. Class III is second with 10.71% of animates. In classes II and IV, animate nouns have a marginal presence with 5 (0.69%) and 4 animate members (2.25%), respectively. Even among these nine members, one in class II and two in class IV are collective nouns that are not traditionally considered proper animates. Moreover, as shown in Arsenijević (2017), even animate diminutives, which would add a couple of dozens of lower frequency animates to class II if the sample were expanded, all show a defective behaviour, manifested for example in failing to inflect for plural. Assuming that the seven exceptions are idiomatically stored, declension classes can be divided in two carrying or realising a -ANIM feature (class II and IV, assuming the animates are lexically stored as idioms) and two without it (class I and III).

All together there are 542 count animate common nouns and 377 personal names in the database – in sum 919 count animate nouns, i.e. 19.48% of the sample. Only these nouns are candidates to carry natural gender features, as they can only be interpreted on animates. In the remaining over 80% of nouns, gender features may only

ženski 'female' (*konj* 'horse', *pas* 'dog'). All of them have specialised nouns for the gender denoted by the adjective used for testing.

occur as purely formal. Nothing predicts their presence or absence, and they basically stipulate the agreement patterns, i.e. the declension classes of the agreeing items.

Finally, there is also a number of animate nouns controlling feminine agreement, which do not trigger female, and even tend to trigger male presuppositions, such as *pijanica* 'drunkard', *budala* 'fool', *spavalica* 'sleeper'. An approach stipulating the feature FEM on these nouns would correctly predict their agreement, but would incorporate a problematic property, namely that an animate noun carrying feature FEM triggers male presuppositions. Even worse, hybrid nouns like *komšija* 'neighbour', *gazda* 'master', *vladika* 'bishop' would need to be specified for not bearing the feature FEM in the singular, but receiving it in the plural, while triggering male presuppositions in both sets of forms.[7]

(9)

a.	teš-k-a	pijanic-a	b.	naš-Ø	komšij-a	c.	naš-e	komšij-e
	heavy-FEM	drunkard-III		our-MASC	neighbour-III		our-FEM.PL	neighbour-III.PL
	'heavy alcoholic'			'our neighbour'			'our neighbours'	

Mass nouns in the broad sense determined by our test make up 81,63% of class IV and 72,67% of class II nouns. Among class I and class III nouns, they are represented with only 7,16% and 26,21%, respectively.

In class II, this is mostly due to the productivity of the suffix *-VV.je* and the abstract mass individual-to-property denominal suffix *-stvo* (9,32%), as these suffixes derive almost only mass nouns.

Due to the productivity of the suffix *-ost*, the percentage for class IV is expected to significantly increase when lower frequency nouns are included.

However, the optimism from the sharp contrasts is relativised by the fact that still a quarter of class III and 162 nouns from class I have mass denotations and would need to be idiomatically stored in case mass semantics were used as a feature that takes part in deriving the declension class.

Collective nouns, which I take in the narrow SC-specific sense of nouns that denote groups of entities and allow both singular and plural number agreement, are generally rare – no class reaches 1% of collective members. Still, differences indicate a clear split in the

7 One may, of course, always introduce an additional gender feature, such that one is responsible for agreement, the other for the interpretation (see e.g. Puškar 2017), but this is a costly methodological choice – especially considering that an additional account is still needed for how these two features determine the realisation of the inflection (i.e. the declension class).

tendencies. It seems that class I is closed for collectives: even though it is the largest class, it does not have a single collective member. Class III has 0.19% collective members, class IV 0,56% and class II 0,83%. Part of the reason is that class II and class IV have morphological operations deriving collective nouns (suffix *-VV.je* with nominal bases and suffix *-ad* with class II nouns as bases, respectively).

Together, mass and collective nouns indicate 1) that classes II and IV are much more likely to be homogeneous (i.e. cumulative and divisive) in mereological terms (e.g. Krifka 1989), i.e. to have non-atomised and/or messy semantics in terms of Landman (2011), 2) that class I has a strong tendency for quantised, i.e. neat atomised semantics, and 3) that class III is open for all types of meaning. Similar can be concluded from animacy, considering that animacy ranks high on the individuation hierarchy, as opposed to mass and collective meanings. In all these properties, class III ranks in the middle, indicating that it is permissive for all the observed properties.

Let us now skip to the properties annotated for mass nouns only. Note that the test used to identify mass nouns is whether the noun combines with the mass quantifier *sav* 'all'. This quantifier combines not only with proper mass nouns, but also with all the nouns that have a vague quantity interpretation. The annotation aimed to identify as mass only those nouns that combine with the quantifier in their dominant interpretation, i.e. without accommodating the vagueness. Still, the nouns identified as mass ended up including a number of count vaguely bounded nouns, as illustrated below – where it was hard to put the boundary between dominant and non-dominant interpretations. The annotation of the type of atomisation identifies this subclass, thus splitting the broad class of mass nouns into a big part which is quite reliable, and a smaller one with compromised reliability.

The property **atomised** was annotated with three values: as atomised, when the noun denotes a homogeneous assembly of atomic units (*pasulj* 'beans', *pesak* 'sand', *društvo* 'society') or an assembly of units varying in the relevant properties (*divljač* 'game [wild animals]',[8] *smeće* 'trash', *komšiluk* 'neighbourhood'); as count vague, when the noun in the singular denotes a singular entity without internal atomisation, but with vague boundaries (*dah* 'breath', *svemir* 'space', *oblast* 'area'); or as non-atomised, when the noun denotes a proper mass: homogeneous, unbounded, without internal atomisation (*tečnost* 'liquid', *komfor* 'comfort', *ambicija* 'ambition').

Event-denoting nouns presented an additional challenge for annotation, as it is hard to properly map them onto the classification above. The following strategy was opted for nouns denoting states, and pure processes without any specification of a culmination like

8 The latter two types were separated in the property type of atom.

čekanje 'waiting', *egzistencija* 'existence', *rad* 'work' were annotated as narrow (i.e. non-atomised) mass. Those denoting single instances of a rather bounded eventuality like *isplata* 'payment', *podsticaj* 'encouragement', *bekstvo* 'escape', were annotated as vague, as they indeed have plurals and are characterised by vague boundaries like vague count nouns above. Finally, nouns preserving the secondary imperfective semantics and therefore having an iterative interpretation (*ispaljivanje* 'firing', *izvrtanje* 'turning inside out') were annotated as atomised – assuming that single iterations of the eventuality constituting potentially unbounded assemblies are equivalent to atoms constituting a homogenous unbounded mass. The aggregate distribution is given in **table 4**.

Table 4 The distribution of types of atomisation among mass nouns (%)[9]

	Class I	Class II	Class III	Class IV
atomised	12,55	38,86	21,64	5,91
vague	31,39	13,90	11,56	18,34
mass	56,05	47,23	66,79	75,73

Relative to other classes, class I has a higher proportion of vague count nouns, nouns that only marginally display mass semantics and essentially are weakly quantised. Regarding the remaining two options, class I has a distribution matching the overall average ratio for these two atomisation values. This additionally supports the treatment of this class as highly individuated. One may wonder why there are non-fully-individuated nouns in this class in the first place – but recall that we are anyway dealing with tendencies, and that the percentage of such nouns in the class was very low. Class II has a significantly higher proportion of atomised mass nouns than other classes, and class IV the opposite: a larger number of non-atomised, i.e. narrow mass nouns. This now provides a contrast, even if only tendential, between class II and IV. While they both have a tendency for mass interpretations, class IV tends to have a proper mass interpretation, and class II rather tends to be atomised. Class III is close to the average overall distribution again, except that it has a somewhat lower participation of vague nouns. This confirms again that this class does not take part in the competition in the domain of properties of quantity.

To make sure that the choices made regarding the annotation of event-denoting nouns did not distort the distribution, I consider also the distribution when event nouns are disregarded, as given in **table 5**.

9 $Chi^2 = 113{,}9331$, $p < 0{,}00001$.

Table 5 The distribution of types of atomisation among mass nouns without event nouns (%)

	Class I	Class II	Class III	Class IV
atomised	15,90	22,22	26,33	4,87
vague	32,38	35,85	10,97	17,68
mass	51,70	41,91	62,69	77,43

Only class II shows a notable difference, in particular in the increase in vague count nouns at the expense of the other two types. The percentage of vague count nouns has actually tripled, and makes class II the richest of all classes in this type of nouns. Considering that in class II pretty much all event-denoting members are derived by one suffix, the distribution without it is possibly more telling about the class.[10] Crucially, however, the change only makes stronger the contrast between class II and class IV regarding narrow mass meanings on the one hand, and the atomised interpretations on the other. Finally, the property **abstract** had the distribution in **table 6**.

Table 6 The distribution of abstractness among mass nouns (in %)[11]

	I	II	III	IV
abstract	64,12556	85,09213	76,49254	89,34911
concrete	35,87444	14,90787	23,50746	10,65089

Once again, the grouping is into class II and IV as more abstract, in class I as more concrete and in class III as having average distribution. Considering that abstractness is lower in the individuation hierarchy than concreteness, this too supports the view on which classes II and IV tend towards lower values of individuation, class I towards the higher end, and class III is open for various degrees of individuation. The strong contribution of the suffix *-VV.je* to class II has a significant influence regarding this property too. As the suffix almost only derives event nouns and events are abstract entities, without these nouns, this class ranks very low in abstractness (but other per-

10 Since in derived nouns it is the suffix that determines the declension class, the entire set of nouns derived by one suffix might need to be counted as one member of the declension class (this is even more obvious on certain views of derivational morphology, e.g. Lowenstamm 2014; Simonović 2020 for South Slavic). This principle of quantification makes the system even more sharply divided in two small and two large classes (class IV would have 41 members, class II 102, class I 1,559 and class III 699). Since, however, the same suffix does not necessarily assign the same properties of quantity, I chose counting each noun separately as the general approach, leaving an investigation along the lines above for the future, specifically for topics where issues that more strictly correlate with suffixes will be in focus.

11 $Chi^2 = 56.7812$, $p < 0.00001$.

centages would drop too, because event-denoting nouns are a significant group in all declension classes).

In summary, grammatical gender has a high capacity in predicting declension class, but due to the same agreement pattern of classes III and IV, the capacity is higher when the relation is inverted, i.e. when declension class predicts gender. Natural gender helps accounting for just over 4% of the data.

Properties of quantity have a broader coverage. Class I has a very strong tendency for individuation: it tends to include animate nouns, is an unlikely class of mass nouns in the broad sense – and when it includes them, it goes for the vague count and atomised, as well as for concrete, rather than non-atomised and abstract nouns. Still, this is a tendency and not a rule – class I includes 176 mass members, and 94 others that are ambiguous between a strict count and a mass interpretation in the broader sense.[12] The significance of these quantities is strengthened by the fact that class I is syncretic with class IV in the most frequent forms (NomSg, AccSg, NomPl) and with class II in almost all the remaining ones, and that nouns not only switch between class I and class IV, i.e. class II, but there are some that at least for some speakers may be declined in two classes (*bol* 'pain', *ološ* 'scam', *polen* 'pollen', i.e. *auto* 'car', *radio* 'radio', *dinamo* 'dynamo'). This latter fact might actually be offering a partial answer too: class IV is almost closed (see Simonović 2020) and its members are gradually switching to class I, probably exactly due to a reanalysis based on the frequent syncretic forms. Class I is possibly taking mass members due to a gradual disappearing of class IV and the syncretism between them, which drives the reanalysis.

Class I could hence in principle be modelled in terms of feature specification as +ATOM +NEAT (Landman 2011). For this, its 270 members that (also) have mass interpretations need to be treated as idioms, i.e. as class I members stored in the lexicon with a mass or vague interpretation in spite of the strict count feature specification that they bear. Already these numbers, received on a sample of 4,718 nouns, present quite a large quantity of nouns to be stored, and the actual numbers for an average speaker are likely five to ten times higher.

Classes II and IV are inanimate (the 7 overall individuated animate nouns they include are plausibly stored), with a strong tendency for

12 The reason why the regularities are only tendential is that they probably emerge as surface generalisations motivated by heterogeneous factors including language use. Their base is the interaction between the lexicon, grammar and pragmatics in language use, and they are restricted by various properties of the grammar and the lexicon (such as the gradual loss of class IV, the general preference to assign less marked declension classes, phonological properties of the stem favouring one declension class or another, specification of the derivational suffix for derived nouns or the shape of the noun in the donor language for the borrowed ones).

mass interpretations in the broad sense. When it comes to properties of quantity, class IV is tendentially specialised for non-atomised meanings, especially for property-denoting abstract nouns, and class II for atomised mass, especially for event-denoting abstract nouns. Class II could hence be modelled as bearing the -NEAT feature, and class IV as bearing the -ATOM feature. Again, this requires that proper count nouns in class II and atomised nouns in class IV be represented as idioms interpreted differently than indicated by their feature specification. Even though the percentage of these exceptions is higher, due to the small size of the classes, the absolute numbers are similar to class I. In the sample of 4,718 nouns, this yields 197 class II and 50 class IV nouns. Due to a high productivity of the suffixes *-VV.je* and *-ost*, the percentage of exceptions is likely to fall with the expansion of the sample to a speakers's entire lexicon, and the absolute numbers should not grow as much as in class I, but they are still several times higher than those attested.

The fatal hit for a realisational approach based on the properties of quantity comes from class III, which is open for all types of interpretations. If modelled in terms of features, this class can be characterised as lacking any of the relevant features, i.e. as having them all, and being realised based on the superset, i.e. subset principle. The problem is that there is no noun for which this class does not compete with at least one other class which is more specific and hence must win as the realisation of the declension class of the noun. The only way for this class to ever win is that its feature specification is irrelevant (hence probably absent), and all its members are stored as idioms (similar to Caha's 2021 proposal). Effectively, this means that, by extremely conservative means, at least 40% of the nouns in the lexicon are effectively assigned a declension class in whatever storage the framework postulates for idiomatic nouns, and at most 60% receive their declension class via realisation based on their feature specification. It is hard to argue that a model employing both realisation and an equivalent of lexical specification is better than one which entirely relies on the lexical assignment of the declension class as a feature. In the next paragraph, I propose an alternative model, in which declension class is a feature assigned in the lexicon, weakly depending on a range of semantic properties, some of which are independently represented in terms of other features (such as ATOM or NEAT), others being a matter of LF interpretation (such as gender). Tendential regularities discussed in this paragraph then must represent surface generalisations constituting the operation of declension class assignment, which has independently been argued to be sensitive to various often clashing factors (segmental phonology, prosody, properties of quantity, natural gender, borrowed suffix), see e.g. Nesset 1994. A tentative analysis would have all these interactions captured simply in terms of markedness: more marked values

of one feature (declension class) tend to match with more or with less marked values of other features.

5 The Proposal: Nouns are Specified for Declension Classes, Interpretive Effects Emerge in Pragmatic Competition

I have so far provided 4 major arguments against a realisational treatment of declension classes:

1. gender agreement is effectively the declension class displayed by the agreeing item, and hence the grammatical gender is effectively also a declension class specification on an item which can be specified for any of the three agreeing declension classes available;
2. the mapping from grammatical gender to declension classes is not fully defined, since feminine gender may map onto class III as well as class IV; the mapping from declension classes to gender (i.e. from the declension class on the noun to the declension class on the agreeing item) does not suffer from this problem;
3. grammatical gender needs the information about the natural gender to predict the declension class of hybrid gender agreement nouns, but natural gender is not a likely syntactic feature (it is not discrete, and the corresponding meaning may emerge via pragmatic competition);
4. properties of quantity also have a power to predict a significant portion of declension classes, but also with a non-marginal set of exceptions and with a particular problem to capture class III and its diachronic stability.

Instead, I argue, declension classes are better represented as features in the nominal projection forming a markedness hierarchy as follows: IV < I < II < III.[13] Agreement then unfolds in the following way, which is similar in spirit to the analysis proposed by Kučerova (2018), but with declension classes as the basic feature and with a more conventional rescue strategy (she postulates look-ahead outside of the language faculty, into the conceptual-intentional system). Syntax first determines the syntactic domain for the controller of agreement. Then, it inserts copies of a declension class value from this domain onto the agreeing item (as no copy can be inserted for class IV,

13 The formality of representation may be further increased by a Nanosyntax style of representation, where, for instance, class I is represented as just the class feature in the respective projection, class II as the absence of this projection (see Arsenijević 2017), class IV as the class feature and a mass feature adjacent to each other, and class III as two consecutive class projections.

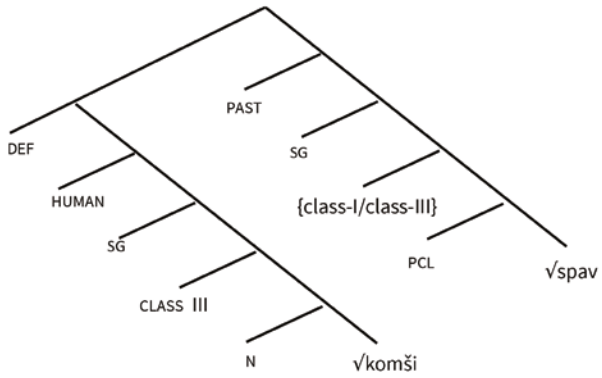
this class resorts to class III as the most marked). Ideally, it copies the features of the projecting noun, but there are alternative options too which I do not discuss in this paper (but see for instance Willer Gold et al. 2016 and Arsenijević et al. 2021). The structure is eventually spelled out and interpreted at PF and LF. In contexts in which the declension class is interpretable, it triggers presuppositions at LF. When these presuppositions clash with stronger presuppositions that are active in the discourse, the derivation crashes. Upon a crash, syntax may replace the declension class with the default and spell it out again. If no clash emerges, the derivation converges. Let me illustrate this on a pair of examples.

In both sentences in (10), the subject is a human-denoting nominal expression from class III. Therefore, this class is copied onto the agreeing item, and it is spelled out. Feature HUMAN causes the declension class feature to trigger gender presuppositions at LF. The gender presupposition triggered by class III is female. Here, the presupposition is not the content of the declension class, nor a property assigned to it – but rather a pragmatic mechanism relying on the procedure of declension class assignment in the lexicon, which includes natural gender as a parameter. When a declension class shows a strong tendency regarding some natural gender, pragmatics takes the occurrence of this declension class in agreement as a source of weak presupposition of the respective natural gender (similar holds for the properties of quantity). In (10a), this presupposition clashes with the presupposition independently available about the referent (here, a strong male presupposition is triggered by the fact that the noun *komšija* stands as the unmarked member in the natural gender opposition with the feminative *komšinica* 'female neighbour'; for various degrees of strength of the gender presuppositions depending on the source, and for additional, experimental evidence for this view, see Arsenijević et al. 2021).[14] The derivation crashes. As an attempt of repair, syntax now spells the structure out with the default declension class I on the agreeing item. This declension class triggers a male or gender-unspecified presupposition and the derivation converges. In (10b), all is the same until the structure is spelled out with the class III specified on the agreeing item. Here, however, the noun *sestra* 'sister' triggers a strong female presupposition due to being the marked member in the natural gender opposition with the noun *brat* 'brother'. Presupposition triggered by agreement matches the independently available strong presupposition, and the derivation converges.

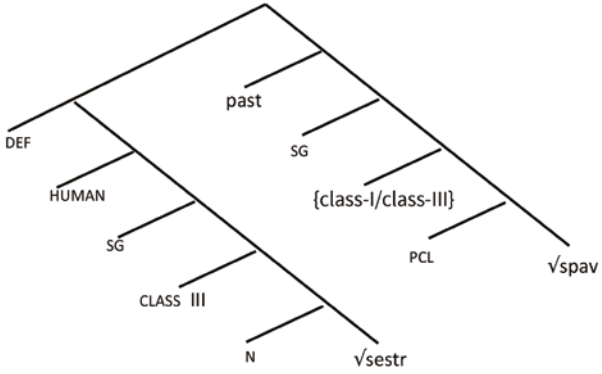
14 In the perspective of Distributed Morphology (Halle, Marantz 1993), the relevant information is whether there is a lexical realisation available for a more marked counterpart that may saliently realise the female gender. This information does not need access to the vocabulary as it is arguably available in the encyclopedia.

(10)

a. Komši-a je spava-o-Ø
neighbour-III.NOM.SG AUX.SG sleep-PCL-I.SG
'The neighbour was sleeping'.



b. Sestr-a je spava-l-a
sister-III.NOM.SG AUX.SG sleep-PCL-III.SG
'The sister was sleeping'.



Something similar happens when class II nouns are used in contexts presupposing neat quantity structures, as in conjoined singulars triggering the sum-interpretation illustrated in (11).

(11)

Sel-o	i	polj-e	su	potopljen-i	/	*potopljen-a
village-NEUT.SG	and	field-NEUT.SG	are.PL	flooded-MASC.PL	/	flooded-NEUT.PL

'The village and the field are flooded'.

Here, again, syntax first copies class II onto the participle. This triggers the presupposition of a messy quantity structure, which clashes with the strong presupposition of a neat quantity structure triggered by the conjunction of two count singulars. The derivation crashes, and syntax repairs it by inserting the default class I feature on the agreeing item. The derivation then converges.

The correlations that declension classes display with grammatical gender (i.e. with declension classes on agreeing items), natural gender and properties of quantity are best captured in terms of pragmatic competition between the degrees of markedness of declension classes sensitive to the degrees of markedness of natural gender, i.e. of properties of quantity. There are two different domains in which this competition shows effects. One is the lexical assignment of declension classes to nouns, where the competition contributes to the selection of declension class assigned to the noun. The other is the assignment of declension class to the agreeing items, traditionally referred to as gender agreement, where the competition determines the presuppositions triggered in the domain of gender and/or properties of quantity. Let me briefly assess each of the relevant cases.

When it comes to properties of quantity, I assume the four types of meanings already considered in § 4 to be ordered regarding markedness in the following way: non-atomised mass < atomised mass < vague count < strict count. The extreme values on the scale map onto the first two declension classes: non-atomised mass to class IV, strict count to class I. The remaining two values: atomised mass and vague count, which are also quantitatively more marked (i.e. they form significantly smaller classes in the lexicon, see § 4), target the next declension class in the hierarchy: class II. Hence, class IV tends to take mass non-atomised nouns, class I strict count nouns, and class II tends to take mass atomised and vague count nouns. Class I is more likely to take a vague count or an atomised mass noun than a non-atomised mass noun, and class IV is more likely to take an atomised mass noun or a vague count noun than a strict count noun. For reasons yet to be investigated, class III remains outside this mapping and receives members based on other properties, such as expressiveness, segmental phonology, prosody. Note that it is independently clear that other properties play a role, from the fact that even classes I, II and IV do not fully map to properties of quantity, but rather show a tendency. Two of those additional properties, natural gender and expressiveness, are discussed below.

To illustrate this, consider the nouns in (12). The base *konj* 'horse' denotes an animate individual, hence with properties atomised and neat, and sitting relatively high in the individuation hierarchy. This supports the assignment of class I. The base *sen* 'hay' denotes an atomised mass, which supports the assignment of class II, and the suffix *-ost* introduces the nominal category for a base that denotes a property, supporting the assignment of class IV. Finally, it remains unclear on the proposed account why the suffix *-b* is specified for selecting class III, as both semantically and phonologically (i.e. to secure a vowel after the consonant cluster which it typically forms), it could as well take class II.[15]

(12)

konj	sen-o	mlad-ost	ber-b-a
horse.I	hay-II	young-N.IV	pick-N-III
'horse'	'hay'	'youth'	'harvest'

In agreement, as stated above, the declension class copied from the controller is the first option, but the default is an available option too. Hence, the competition is more intensive than at declension class assignment, as the choice of the declension class is made anew in every context (a noun is assigned class once and keeps it in all its uses). In salient contexts, the selection of a marked declension class triggers the presupposition of the holding of a marked property along the respective dimension. This is the reason for the agreement quirks of class II nouns reported in Arsenijević (2017), including those reported in (7), repeated as (13). Class II is the most marked class when it comes to properties of quantity – a class latently mapped with vague count and atomised mass nouns. Hence, it triggers the presupposition of the corresponding meaning. The pointing context, animate plurals, as well as conjunction, imply assemblies of highly individuated entities, yielding a clash. The pointing is resolved by using the class I form of the demonstrative (*oni* 'they.II') and the agreeing form of the predicate, the plural by using the collective plural form from class IV (*telad* 'calves.COLL'), and the conjunction again by resorting to class I, this time on the agreeing item only.

15 An anonymous reviewer suggests that arbitrary choices and mismatches of the type discussed could be a trace of diachronic processes. If it is correct that declension class assignment proceeds in the interaction of grammar, pragmatics and the lexicon in language use, it is indeed plausible that current assignments are traces of some earlier stages in which they were the optimal choices – under the condition that the later stages of development did not render them (strongly) suboptimal.

(13)

a. Ona *(deca) su neumorna. (pointing at children)
those.II.PL child.II.PL AUX.PL tireless.II.PL
'Those children are tireless' / *'Those are tireless'.

b. dugm-e dugm-eta tel-e *tel-eta tel-ad
button-II.SG button-II.PL calf-II.SG calf-II.PL calf-SUFF.IV.SG
'button' 'buttons' 'calf'

c. Sel-o i polj-e su potopljen-i / *potopljen-a
village-II.SG and field-II.SG are.PL flooded-I.PL / flooded-II.PL
'The village and the field are flooded'.

Natural gender is a property of animate nouns. Moreover, it only shows on the strictly count ones among them – as natural gender is a property of an individual, not of a group. Since strictly count animates are very high on the individuation hierarchy, and since, due to the mapping to properties of quantity, classes II and IV do not take nouns so high on this hierarchy, the opposition in natural gender only establishes competition between classes I and III. Regarding markedness, masculine natural gender is unmarked and feminine is marked. The mapping is such that the value of gender characterised by higher markedness maps onto the declension class that is more marked. As a result, all the nouns denoting females are assigned class III, and a majority of nouns denoting males are assigned class I (of the 139 natural masculine gender nouns in the database, 12 belong to class III, together with all the 52 natural feminine nouns). Again, class III turns out to be open for all the values, while class I is exceptionless in not taking one of the values – the feminine natural gender.

To illustrate this, in the respective culture, the prototypical ant is represented as male, and the prototypical bee as female. As the only two classes that can take a (non-diminutive) animate noun are class I and III, class I wins for the *ant* because it is does not clash with the male presupposition and is the default class, and class III for the *bee*, as it does not clash with the female presupposition. The agentive suffix *-ar* derives the human denoting noun *rudar* 'miner' and, as the prototypical human in the culture is conceptualised as a male, the noun receives class I. The phonologically empty suffix *-k* is added to this noun to derive a marked minimal pair. This markedness can be interpreted along several lines: as a diminutive, as the opposite gender, as an expressive variant (a similar observation is made in Caha 2021; see Potts 2007 for a discussion of expressive semantics). In the particular case, the functional load determines the marked gender, i.e. female interpretation as the relevant dimension. Since the corresponding presupposition matches class III, this is the class assigned to the noun.

(14)

mrav	pčel-a	rud-ar	rud-ar-k-a
ant	bee-III	mine-er.I	mine-er-*k*-III
'ant'	'bee'	'miner'	'she-miner'

In agreement too, the competition only emerges with humans and entities higher than that in the hierarchy. Hence, the competition is between classes I and III only. The more marked class, class III, triggers the presupposition of the more marked natural gender value, feminine (since the opposition is binary, class I favours the masculine gender by antipresupposition). When the declension class is copied in agreement, since agreement is reinterpreted for presupposition in each context as discussed above, a clash may emerge, which gets resolved by the spell out of the default declension class.

Caha (2021, 37) generalises affectiveness as a property of male-denoting members of the Russian counterpart of class III. Indeed, expressiveness occurs to be a third dimension of markedness that interacts with declension classes that arguably enters a mapping of the above type. Assuming that nouns with expressive meanings are marked compared to those without them, a mapping can be established where the marked value of expressivity maps higher rather than lower on the declension-class-markedness scale. Unfortunately, the database does not yet have a complete annotation of the nouns for expressivity, but of the 26 hybrid gender class III nouns in the database, 15 have a clear expressive meaning, ranging from the positive (*deda*, *deka* 'grandpa', *čika*, *teča* 'uncle', *beba* 'baby'), via the ambivalent (*vođa* 'leader', *gazda* 'master', *pristalica* 'supporter'), to the negatively biased ones (*ubica* 'murderer', *budala* 'fool', *ustaša* 'member of the Croatian Nazi movement').[16]

Consider the examples in (15). The noun *čovek* 'man' and the proper name *Petar* refer to men. Therefore, they are assigned declension class I. However, their hypocoristic forms *čova* and *Pera* (a typical nickname for the name used by close persons), which also denote men, are assigned class III because of their expressive component.

(15)

čovek	čov-a	Petar	Per-a
man.I	man-III	Petar.I	Per-III
'man'	'man (hypocoristic)'	'Petar'	'Petar (hypocoristic)'

16 The noun *ustaša* in fact has also been bearing a positive expressive value for certain speakers, but it has never been neutral.

Similarly, as already presented in (3) above, an expressive component may trigger a noun to display mixed class I/II behaviour, as illustrated in **table 7** below (note that in the nominative and accusative of class II, the endings *e* and *o* are in phonologically conditioned alternation), where the expressive hypocoristic noun *smrdljivko* 'stinky-SUFF-II', unlike its neutral counterpart *smrdljiv-ac-Ø* 'stinky-SUFF-I', shares with class II its typical nominative singular ending and the lack of plural forms for animate nouns, while having the accusative ending characteristic for class II and showing class II agreement behaviour.

Table 7 Expressivity-induced hybrid between classes I and II

	Class I (<Ø, a>)		**Expressive noun**		**Class II (<o/e, a>)**	
	Sg	Pl	Sg	Pl	Sg	Pl
NOM	smrdljiv-c-Ø	smrdljiv-c-i	smrdljiv-k-o	/	pil-e	/
GEN	smrdljiv-c -a	smrdljiv-c-a	smrdljiv-k-a	/	pil-et-a	/
DAT	smrdljiv-c -u	smrdljiv-c-ima	smrdljiv-k-u	/	pil-et-u	/
ACC	smrdljiv-c-a	smrdljiv-c-e	smrdljiv-k-a	/	pil-e	/
INST	smrdljiv-c-em	smrdljiv-c-ima	smrdljiv-k-om	/	pil-et-om	/
LOC	smrdljiv-c-u	smrdljiv-c-ima	smrdljiv-k-u	/	pil-et-u	/

Declension class assignment is sensitive to other factors too, such as phonology (bases with final consonant clusters tend to take declension classes with overt endings in the nominative rather than those with null morphology) or the shape of the noun in the donor language for loan nouns (nouns ending in *-a* in the donor language take class III declension, considering that there are no obstacles for such assignment, since class III is semantically flexible), but I do not discuss these aspects here beyond observing that an account realising declension class from gender is not likely to capture them in a transparent way.

The presented view is parsimonious compared to the alternatives, as it only specifies nouns for one relevant feature (class) instead of at least two (grammatical and natural gender, let alone the role of atomisation and neat vs. messy structure). It derives grammatical gender, i.e. the declension class on the agreeing item, from that on the controller of agreement rather than the other way around and does not face the problem of ambiguity of feminine agreement between class III and IV realisation. It is compatible with the findings regarding the pragmatic competition between minimal pairs (Merchant 2014; Sudo, Spathas 2019). Finally, it does not have to store a significant amount of lexical material or forms as idioms interpreted contrary to their feature specification or otherwise face a large number of exceptions. The approach is similar in spirit to Kučerova's (2018), in invoking a semantic evaluation of agreement, but differs both in the scope and in the technical implementation.

6 The Morphology of Declension Classes

In this paragraph, I propose a novel morphological analysis of declension classes in SC, which supports the markedness ordering assumed in the discussion and models provided in the preceding paragraphs: IV < I < II < III. The analysis capitalises on the fact that class I has not a single ending that is not syncretic with at least one other class, counting only same case and number counterparts, as obvious from table 1, repeated as **table 8**. In the singular, class I nominative and accusative are syncretic with class IV, class I genitive, dative/locative and instrumental with class II. In the plural, nominative is syncretic with class IV, accusative with class III, genitive with both classes II and III, and dative/locative and instrumental with classes II and IV. I take this as an indication that class I has no endings of its own: it is always realised by the endings from other declension classes.

Table 8 Declension classes in Serbo-Croatian

	Class I (<Ø, a>)		**Class II (<o/e, a>)**		**Class III (<a, e>)**		**Class IV (<Ø, i>)**	
	Sg	**Pl**	**Sg**	**Pl**	**Sg**	**Pl**	**Sg**	**Pl**
NOM	mrav-Ø	mrav-i	mor-e(sel-o)	mor-a	rib-a	rib-e	noć-Ø	noć-i
GEN	mrav-a	mrav-a	mor-a	mor-a	rib-e	rib-a	noć-i	noć-i
DAT	mrav-u	mrav-ima	mor-u	mor-ima	rib-i	rib-ama	noć-i	noć-ima
ACC	mrav-(a) (žig -Ø)	mrav-e	mor-e(sel-o)	mor-a	rib-u	rib-e	noć-Ø	noć-i
INST	mrav-em	mrav-ima	mor-em(sel-om)	mor-ima	rib-om	rib-ama	noć-ju	noć-ima
LOC	mrav-u	mrav-ima	mor-u	mor-ima	rib-i	rib-ama	noć-i	noć-ima

Moreover, I use the same feature, DIV, to represent both the genitive case and the plural (see Arsenijević 2005 for an elaboration). Genitive plural has the feature twice, once in the head realising grammatical number and once in the head realising the case.

(16) GEN SG: $[_{KP}$ DIV $[_{NumP}$ Ø$]]$ NOM PL: $[_{KP}$ Ø $[_{NumP}$ DIV$]]$ GEN PL: $[_{KP}$ DIV $[_{NumP}$ DIV$]]$

This accounts for the fact that the genitive singular forms of classes II, III and IV are identical to their nominative plural forms. Class I takes the plural form of class II for reasons to be briefly assessed. It also explains why genitive plural forms have the form of the nominative plural with a doubled ending (class I and III take the form from class II), again for reasons to be briefly discussed.

Finally, accusative is syncretic with the nominative, except in class III singular, where it has its own ending. I leave aside the differen-

tial object marking in class I, where animates have an accusative ending that is not copied from the nominative cell, but rather from the more marked genitive. Similarly, in plural, class I takes an accusative plural that is not identical to the nominative plural that it takes from class IV, plausibly due to a tendency to pick a more distinctive marking.

Assume further that class II has a mid vowel (V_{mid}) as its theme vowel, class III simply a vowel (V), and class IV *i*. By default, V_{mid} is realised as *o* (as confirmed by its realisation after a vowel),[17] but after palatal consonants it gets realised as *e*. By default, V is realised as *a*. When the case ending starts with a vowel, in those forms that attach with a theme vowel, the theme vowel remains in the form of a zero vowel, and blocks the phonological interaction of the stem with the suffix. In some forms this is subject to variation, as illustrated in (17), where speakers of one grammar do and speakers of the other do not (drop the zero theme vowel and) palatalize final velars in class II in front of endings with an initial front vowel.[18]

(17) pazuh-ima → %pazusima / %pazuhima (i.e. pazusima / pazuh*V*ima)
klupk-ima → %klupcima / %klupkima (i.e. klupcima / klupk*V*ima)

If all observable instances of syncretism are coded, we arrive at **table 9**, where all instances of syncretism between declension classes are marked by reference to the declension class and by the background colour in the respective cell, and endings are otherwise labelled according to their primary feature content (theme vowels are indexed for the class). The table is purely descriptive, for the purpose of establishing the amount of intra- and interclass syncretism in each class. No theoretical significance is intended (e.g. how the instances of syncretism emerge is not implied), assuming without further evidence that all various approaches to morphology available can derive it one way or another.

17 I am grateful to Marko Simonović (personal communication) for this piece of evidence, as well as that in (17).

18 See Arsenijević et al. (2021) for the same effect of a theme vowel without a full vocalic realisation preventing the preceding consonant from undergoing palatalization in the verbal domain, as illustrated in (i).
(i)

za-jah-a-ti	→	za-jah-V-iv-a-ti	→	zajahivati / *zajašivati
hind-ride-TH-INF		hind-ride-TH-IPF-TH-INF		
'to mount.PRF'		'to mount.IPF'		

Table 9 Sebo-Croatian declension class system re-analysed

	Cl3Sg	Cl2Sg	Cl1Sg	Cl4Sg	Cl3Pl	Cl2Pl	Cl1Pl	Cl4Pl
NOM	Th_3	Th_2	Class4	Ø	Pl	Pl	Class4	Th_4
GEN	Th_3-Acc	Nom	Class4	Nom	Nom	Nom	Class3	Nom
DAT	Th_3-Pl	Th_2-Pl	Class2	Th_4	Class2/4	Plx2	Class2	Plx2
ACC	Dat	Th_2-Dat	Class2	Th_4	Th_3-Inst	Th_2-Class4	Class4	Th_4-Inst
INST	Th_3-Inst	Th_2-Inst	Class2	Th_4-Inst	Th_3-Inst	Th_2-Class4	Class4	Th_4-Inst

As the table shows, class III only has one syncretic slot, genitive plural, where it takes a class II or class IV ending.[19] Class II ending is the default, and class IV ending is conditioned on particular phonological environments (see e.g. Simonović, Baroni 2014). The selection of the class IV ending does not violate the hierarchy. The default ending is taken from the closest declension class. When a phonological context requires a different ending, the alternative is taken from the next closest class with a different ending. As class I is syncretic with class II in this form, the only remaining choice, i.e. the closest declension class with a different ending for genitive plural, is class IV. Class III has the lowest amount of syncretism within itself and with other classes, and the highest number of different own endings. Class IV only participates in syncretism relations as a donor.

Other interesting generalisations include that nominative has no ending in any singular form, but while in classes II and III it involves the theme vowel of the class in a strong realisation (i.e. as a full vowel), in class IV it even lacks this component. The istrumental case universally has the ending *-m* in the singular and *-ma* in the plural, except for the singular form of class IV, where it reduces to *-u*.[20] Otherwise, class IV has very few endings generally, but also never (observably) reduces the theme vowel, i.e. it generally relies more on the theme vowel than on endings to realise case forms.

There is only one instance of syncretism under this analysis which skips one declension class in the hierarchy: the genitive plural form in class I which is taken from class III. However, since the accusative on the proposed analysis tends to be a function of the nominative rath-

19 I thank an anonymous reviewer for reminding me of the fact that a number of nouns of this class with stem-final consonant clusters take the ending of class IV.

20 This alternation is also observed in 1st person singular of the present tense, where the ending is *-m*, but several verbs realise the ending as *-u*, as well as between the adjectival and the nominal declension where in classes I and II the former takes *-m* and the latter *-u* in the dative/locative singular (*bel-o-m zid-u* 'white-TH-DAT wall-DAT'). It is obviously far-fetched to make any claims, but it is thinkable that in certain contexts *-m* occurs as the default when *-u* realises the marked alternation.

er than an independent ending, this instance of syncretism is plausibly a property of the function rather than a syncretic ending skipping the hierarchy. For instance, it could be that the accusative poses a requirement for the plural number feature to be visible, which is not satisfied by the two neighbouring candidates, as class IV only has a theme vowel in this slot, which also occurs in the singular, and class II has a deficient plural (see Arsenijević 2017), but is satisfied by class III. The analysis therefore may well feed models with further decomposition and reduction, such as for instance Caha's (2009).

The reason why class I avoid class II nominative in the plural is probably also the reason why both class I and class III take the genitive plural ending from class II. Genitive plural involves a double division, once in the number projection and one more time by the genitive case, as in (16) above. Classes I and III include highly individuated members (e.g. animate and human nouns), and their plural is hence interpreted as a neat division over the restriction. For different possible reasons which may involve avoidance of effective higher order divisions or the fact that classes I and III also have members with messy semantics, a weaker plural is preferred in the context of its double implementation. This is exactly what class II plural does, as it is accommodated to the messy semantics of this class (recall that a majority of its members have messy semantics, in particular as atomised masses or as vague count denotations). I leave a deeper elaboration of this question for future research.

The analysis in table 7 confirms the hierarchical ordering of declension classes as introduced in § 5: IV < I < II < III. Not only is this supported by the patterns of syncretism, where the given order involves no skipping (under a derivational realisation of the accusative), but also the numbers of different endings monotonically decrease, and the occurrences of the theme vowel monotonically increase from left to right.

7 Conclusion

I argued, based on the data from SC, against both views that independently represent both declension classes and grammatical gender as features of the noun, and those that try to derive the declension class from the combination of grammatical gender and other features. The arguments that I put forth include the fact that gender is predictable from the declension class of the noun in almost 96% of the nouns (on a sample of the 4,718 most frequent nouns) while gender needs additional information to distinguish between the classes III and IV; that gender boils down to the declension class of the agreeing item and hence cannot be a more basic property of a noun than its own declension class; and that at least four families of properties corre-

late with declension classes (agreement, natural gender, properties of quantity, expressivity) but none correlates absolutely. This last argument is supported by a detailed quantitative study of the relevant properties on a database of annotated SC nouns.

An alternative model is proposed, where the declension class is assigned to nouns in the lexicon and get copied in agreement from the noun onto the agreeing item. Exceptions emerge when the copied value triggers false presuppositions regarding the natural gender, properties of quantity or expressive content, in which case the default declension class I is inserted instead.

Finally, I offered a novel analysis of the SC declension classes, in which the optimal arrangement to capture the patterns of syncretism and the roles of the theme vowels and case endings yields exactly the hierarchy of markedness postulated in the analysis of the mapping of declension classes to natural gender and properties of quantity: IV < I < II < III.

Bibliography

Alsina, A.; Arsenijević, B. (2012a). "Hierarchies and Competing Generalizations in Serbo-Croatian Hybrid Agreement". Butt, M.; King, T.H. (eds), *17th International Lexical Functional Grammar Conference (LFG 2012)*. Bali: Udayana University, 6-22.

Alsina, A.; Arsenijević, B. (2012b). "The Two Faces of Agreement". *Language*, 88, 369-79.

Arregi, K.; Nevins, A. (2012). *Morphotactics: Basque Auxiliaries and the Structure of Spellout*. Dordrecht: Springer.

Arsenijević, B. (2007). "Značenje srpskog imenskog sufiksa *-je*" (The Meaning of the Serbian Nominal Suffix -je). *Jezik i društvena kretanja* (Language and social transitions). Kragujevac: FILUM, 71-176.

Arsenijević, B. (2017). "Gender, Like Classifiers, Specifies the Type of Partition: Evidence from Serbo-croatian". *CLS*, 52, 21-37.

Arsenijević, B. (forthcoming). "Adjectives as a Lexical Category – a Story of Striving for Extension". Mitrović, M.; Panagiotidis P. (eds), *A°: The Lexical Status of Adjectives*. Amsterdam: Jon Benjamins.

Arsenijević B.; Mitić, I.; Puškar-Gallien, Z. (2021). "Serbo-Croatian Hybrid Gender Agreement: Availability of the Restrictive Use Predicts Agreement" [unpublished manuscript] University of Graz.

Arsenijević, B.; Gračanin-Yuksek, M. (2016). "Agreement and the Structure of Relative Clauses". *Glossa: A Journal of General Linguistics*, 1(1), 17. http://doi.org/10.5334/gjgl.12.

Arsenijević, B.; Mitić, I. (2016a). "Effect of Animacy and Agentivity on the Processing of Agreement in Serbo-Croatian". Halupka-Rešetar, S.; Martínez-Ferreiro, S. (eds), *Studies in Languages and Mind, Selected papers from third Novi Sad Workshop on Psycholinguistics, Neurolinguistic and Clinical Linguistic Research*. University of Novi Sad: Faculty of Philosophy, 41-77.

Arsenijević, B.; Mitić, I. (2016b). "On the (In)dependence of Gender with Respect to Number in Agreement with Coordinated Subjects". *Journal of Slavic Linguistics*, 24(1), 41-60.

Arsenijević, B. (2012). "Semantic Ontology of Deadjectival Nominalizations". *Recherches Linguistiques de Vincennes*, 40, 53-72.

Baker, M. (2003). "Nouns as Bearers of a Referential Index". *Lexical Categories: Verbs, Nouns and Adjectives*. Cambridge: Cambridge University Press, 95-189. http://doi.org/10.1017/CBO9780511615047.003. Cambridge Studies in Linguistics.

Caha, P. (2021). "Modeling Declensions Without Declension Features". https://doi.org/10.1556/2062.2021.00433.

Caha, P. (2009). *The Nanosyntax of Case* [PhD dissertation]. Tromsø: University of Tromsø.

Corbett, G. (1991). *Gender*. Cambridge: Cambridge University Press.

Despić, M. (2017). "Investigations in Mixed Agreement: Polite Plurals, Hybrid Nouns and Coordinate Structures". *Morphology*, 27(3), 253-310. https://doi.org/10.1007/s11525-017-9301-3.

Fassi Fehri, A. (2018). *Constructing the Feminine to Mean. Gender, Number, Numeral, and Quantifier Extensions in Arabic*. New York: Lexington Books.

Fortson, B.W. IV (2010). *Indo-European Language and Culture: An Introduction*. 2nd ed. Oxford: Wiley-Blackwell.

Hachem, M. (2015). *Multifunctionality. The Internal and External Syntax of D- and W-Items in German and Dutch*. Utrecht: LOT-Publication.

Halle, M.; Marantz, A. (1993). "Distributed Morphology and the Pieces of Inflection". Hale, K.; Keyser, S.J. (eds), *The View from Building 20*. Cambridge (MA); MIT Press, 111-76.

Krifka, M. (1989). "Nominal Reference, Temporal Constitution and Quantification in Event Semantics". Bartsch, R.; Van Benthem, A.K.; Van Emde Boas, P. (eds), *Semantics and Contextual Expressions*. Dordrecht: Foris: 75-115.

Kučerová, I. (2018). "φ-Features at the Syntax-Semantics Interface: Evidence from Nominal Inflection". *Linguistic Inquiry*, 49(4), 813-45. https://doi.org/10.1162/ling_a_00290.

Landman, F. (2011). "Count Nouns – Mass Nouns – Neat Nouns – Mess Nouns". Glanzberg, M.; Partee, B.H.; Šķilters, J. (eds), *Baltic International Yearbook of Cognition, Logic and Communication*, 6. https://doi.org/10.4148/biyclc.v6i0.1579.

Ljubešić, N.; Klubička, F. (2016). "Serbian Web Corpus srWaC 1.1, Slovenian Language Resource Repository CLARIN.SI". http://hdl.handle.net/11356/1063.

Lowenstamm, J. (2014). "Derivational Affixes as Roots: Phasal Spell-out Meets English Stress Shift". Alexiadou, A.; Borer, H.; Schafer, F. (eds), *The Syntax of Roots and the Roots of Syntax*. Oxford: Oxford University Press, 230-59. https://doi.org/10.1093/acprof:oso/9780199665266.003.0011.

Moltmann, F. (2004). "Properties and Kinds of Tropes: New Linguistic Facts and Old Philosophical Insights". *Mind*, 113, 1-41. https://doi.org/10.1093/mind/113.449.1.

Matasović, R. (2004). *Gender in Indo-European*. Heidelberg: Winter.

Merchant, J. (2014). "Gender Mismatches Under Nominal Ellipsis". *Lingua*, 151, 9-32. https://doi.org/10.1016/j.lingua.2014.01.008.

Mitić, I.; Arsenijević, B. (2019). "Plural Conjuncts and Syncretism Facilitate Gender Agreement in Serbo-Croatian: Experimental Evidence". *Frontiers in Psychology*, 10, 65-79. https://doi.org/10.3389/fpsyg.2019.00942.

Müller, G. (2004). "On Decomposing Inflection Class Features: Syncretism in Russian Noun Inflection". Müller, G.; Gunkel, L.; Zifonun, G. (eds), *Explorations in Nominal Inflection*. Berlin: Walter de Gruyter, 189-228. https://doi.org/10.1515/9783110197501.189.

Nesset, T. (1994). "A Feature-Based Approach to Russian Noun Inflection". *Journal of Slavic Linguistics*, 2, 214-37.

Potts, C. (2007). "The Expressive Dimension". *Theoretical Linguistics*, 33, 165-97.

Privizentseva, M. (2020). "Declension Tracks Gender: Insights from Mixed Agreement in Russian". Talk given at NELS 51. https://bit.ly/3rhvGFc.

Puškar, Z. (2017). *Hybrid Agreement: Modelling Variation, Hierarchy Effects and Phi-feature Mismatches* [PhD dissertation]. Leipzig: University of Leipzig.

Puškar, Z. (2018). "Interactions of Gender and Number Agreement: Evidence from Bosnian/Croatian/Serbian". *Syntax*, 21(3), 275-318. https://doi.org/10.1111/synt.12154.

Simonović, M. (2020). "Categories, Root Complexes and Default Stress: Slovenian Nominalizations Revisited". *Linguistica*, 60(1), 103-17. https://doi.org/10.4312/linguistica.60.1.103–117.

Simonović, M. (2015). *Lexicon Immigration Service - Prolegomena to a Theory of Loanword Integration*. LOT Publication 393.

Simonović, M.; Arsenijević, B. (2014). "Regular and Honorary Membership: On Two Kinds of Deverbal Nouns in Serbo-Croatian". *Lingue e linguaggio*, 13(2), 185-210.

Simonović, M.; Baroni, A. (2014). "Lexicon, Markedness and Grammar in the Serbo-Croatian Wobbly a". Stanković, B.; Janić, A. (eds), *SinFoniJA 6 Proceedings*, 217-42. https://izdanja.filfak.ni.ac.rs/zbornici/2019/download/1878_4a637f0f1783cc70d10b818a1ebe29f6.

Starke, M. (2010). "Nanosyntax: A Short Primer to a New Approach to Language". *Nordlyd*, 36(1). https://doi.org/10.7557/12.213.

Stevanović, M. (1989). *Savremeni srpskohrvatski jezik* (Contemporary Serbo-Croatian Language). Belgrade: IRO Naučna knjiga.

Sudo, Y.; Spathas, G. (2019). "Natural Gender and Interpretation in Greek: Comments on Merchant (2014)". *Glossa: A Journal of General Linguistics*. https://doi.org/10.5334/gjgl.1173.

Villalba, X. (2009). "Definite Adjective Nominalizations in Spanish". Espinal, M.T.; Leonetti, M.; McNally, L. (eds), *Arbeitspapier*, vol. 124. Konstanz: Fachbereich Sprachwissenschaft, Universität Konstanz, 139-53.

Wechsler, S.; Zlatić, L. (2003). *The Many Faces of Agreement*. Stanford: Center for the Study of Language and Information.

Wechsler, S.; Zlatić, L. (2012). "The Wrong Two Faces". *Language*, 88, 380-7.

Willer-Gold, J.; Arsenijević, B.; Batinić, M.; Čordalija, N.; Kresić, M.; Leko, N.; Marušič, F.; Milićev, T.; Milićević, N.; Mitić, I.; Nevins, A.; Peti-Stantić, A.; Stanković, B.; Šuligoj, T.; Tušek, J. (2016). "Morphosyntactic Production of Coordination Agreement in South Slavic – A Comparative Study". *Journal of Slavic Linguistics*, 24(1), 187-224.

Balcania et Slavia
Vol. 1 – Num. 1 – June 2021

It's All Under Control! On Perfective Present Forms in BCS Main Clauses

Marco Biasio
Università degli Studi di Padova, Italia

Abstract This study aims to provide a unified analysis of the syntax-pragmatics interface of the (allegedly) anomalous licensing of the perfective present (Pres^{PF}) in BCS present-tensed main clauses. Although Pres^{PF} forms cannot usually refer to eventualities that are anchored to the utterance time (UT), there seem to be three apparent exceptions to this structural constraint. They are as follows: 1) abusive metonymic performatives; 2) live demonstrations; and 3) nonveridical contexts introduced by the epistemic operator *možda* 'maybe'. It is claimed that for Pres^{PF} forms to be licensed in BCS main clauses, control needs to be specified as a variable at the level of the so-called 'Seat of Knowledge' in the SpeechActP layer.

Keywords Verbal aspect. Perfective present. Control. Štokavian. Speech act phrase. Syntax-pragmatics interface.

Summary 1 Introduction. – 2 Breaking the Law? Anomalous Pres^{PF} Forms in Main Clauses. – 2.1 Abusive Metonymic Performatives. – 2.2 Live Demonstrations. – 3 Control as a Variable in SpeechActP. – 4 Conclusions.

Edizioni Ca'Foscari

Peer review

Submitted	2021-07-20
Accepted	2021-10-15
Published	2020-12-20

Open access

Citation Biasio, M. (2021). "It's All Under Control! On Perfective Present Forms in BCS Main Clauses". *Balcania et Slavia*, 1(1), 47-66.

1 Introduction

It is widely known that both in the Old- and the Neo-Štokavian dialectal groups of the BCS[1] continuum, there is a pervasive grammatical constraint on the licensing of independent perfective present forms (henceforth Pres^{PF}) that refer to different kinds of eventualities in certain syntactic environments, such as present-tensed main clauses.[2] The unavailability of Pres^{PF} forms in BCS main clauses is exemplified in (1) below, where the replacement of the inflected Pres^{IPF} form of the ambitransitive predicate *pisati* 'to write' with the corrispective Pres^{PF} form of *napisati* yields absolute ungrammaticality:[3]

(1) *Pišem*$_{\text{WRITE.1.sg.Pres}^{IPF}}$ (**Napišem*$_{\text{WRITE.1.sg.Pres}^{PF}}$) sve ovo kao upozorenje novopostavljenim direktorima na moguća neprijatna iznenađenja. [SrpKor, viva0104_n.txt, Viva (April 2001)]
'I'm writing all this to warn all the newly appointed directors of possible unpleasant surprises'.

The Author wishes to express his gratitude to the editors of *Balcania et Slavia*, three anonymous reviewers of FASL 30 for their thought-provoking remarks which helped me greatly in reworking and fine-tuning the core of the original proposal, two anonymous reviewers for their meticulous comments on an earlier draft of the paper, and all the Štokavian native speakers, members of the Facebook groups South Slavic Linguistics (`https://www.facebook.com/groups/697148470712789`) and Kako biste VI rekli? (`https://www.facebook.com/groups/1649886005251662`), who kindly agreed to fill out an online questionnaire. Special thanks go to James Hartstein for checking the correctness and the internal consistency of the text. The usual disclaimer applies.

1 Throughout this article I will adopt the umbrella label BCS (shortly for Bosnian-Croatian-Serbian) in order to refer collectively to the different varieties of the South Slavic dialectal continuum which are nowadays spoken in most of the national republics that emerged after the collapse of Jugoslavia. BCS is to be taken here as a synonym of 'Serbo-Croatian', which was the standard working term in linguistics at least up to the first major outbreak of Jugoslav Wars in 1991. The term however has been sidelined in this paper to prevent the emergence of unwarranted political associations. The Author of the present study acknowledges the scientific validity of the views stated in the Declaration on the Common Language (*Deklaracija o zajedničkom jeziku*) and supports the general claim that Bosnian, Croatian, Serbian, and Montenegrin should be viewed as national variants (each with equal rights) of the same standard South Slavic polycentric language (for more technical evidence in support of this claim see also Bailyn 2010). The text of the Declaration is available at the following link: `https://jezicinacionalizmi.com/deklaracija`. Where not specified otherwise, all translations are by the Author.

2 *Eventuality* is here used (in the original spirit of Bach 1986) as a cover term for both states and non-states (processes, events), while the definition of a *main clause* (i.e. a clause which can stand on its own, in opposition both to matrix and subordinate clauses) is given after Haspelmath (2020, 603).

3 In linguistic theory the concept of *grammaticality* has proved to be notoriously hard to disentangle from the (apparently overlapping) notion of *acceptability*. For the sake of clarity we stick to the results of the discussion recently elaborated in Leivada, Westergaard 2020 and Leivada, Murphy 2021, 4-5, where it is claimed that (un)grammaticality, unlike (un)acceptability, is more of a twofold concept (on the axis 'absolute'-'relative') than a scalar notion.

This constraint has drawn considerable attention over recent decades. In the cognitive literature the functional markedness of grammatical forms with present time reference – resulting from the combination of the aspectual category 'perfective' and the temporal category 'present' – has been labelled as the *present perfective paradox* (cf. among others De Wit 2017). It has been proposed that such markedness holds cross-linguistically, and is mainly triggered by the cognitive difficulty of processing and representing as a single whole an eventuality token still unfolding at the UT.[4] In other words, either the eventuality is predominantly represented as holding at the UT (thus focusing on its outer temporal anchoring), or is seen as a single whole (thus zooming in on its inner temporal properties). In more formal terms, adopting a Kratzerian definition of PF according to which the event time interval needs to be included within the reference time interval (i.e. for present-tensed eventualities, the near-instantaneous UT), Todorović (2015, 87-8) argues that the impossibility of satisfying the inclusion requirement of PF within the time interval selected by the UT in SpecT leads to the impossibility for the event to be temporally located, and consequently triggers the absolute ungrammaticality of main clause Pres^{PF} forms.[5]

De Wit (2017, 38) claims that languages can solve the present perfective paradox resorting to three main repair strategies, i.e. assigning a futural (the *prospective strategy*), past (the *retrospective strategy*), or non-actual reading (the *structural strategy*) to eventualities realised with Pres^{PF} forms. Unlike East Slavic languages, where main clause Pres^{PF} forms have been prototypically reanalysed as aspectual futures (e.g. Russian *pročitaju*$_{\text{READ.1.sg.Pres}^{PF}}$ 'I will read'), or even West Slavic languages, where main clause Pres^{PF} forms can be alternatively used either as aspectual futures or modally nuanced non-actual presents (e.g. Czech *koupím*$_{\text{BUY.1.sg.Pres}^{PF}}$ 'I will buy', 'I buy'), in almost all the South Slavic languages spoken in the Balkans[6] Pres^{PF} forms have retained only their non-actual function, and thus are unable to refer to eventuality tokens which are anchored to the UT.

Old- and Neo-Štokavian dialects of BCS seemingly adopt a reinforced structural strategy[7] as the licensing of non-actual Pres^{PF} forms

4 The Utterance Time (UT) is here identified with the so-called *reference time interval* which, according to the proposal of Demirdache and Uribe-Extebarria (2004), occupies a structural position in SpecT.

5 Note that this formalisation, despite managing to nicely capture the BCS pattern, is not unproblematic *per se*, for it requires several successive adjustments in order to hold true from a cross-linguistic perspective.

6 Exceptions are made for some Kajkavian and Čakavian dialects of the Croatian variety of BCS, which will not be taken into account in this survey.

7 In this respect BCS imposes even more restrictions on the licensing of Pres^{PF} forms than other South Slavic languages (e.g. Bulgarian) and behaves more similarly to oth-

is restricted to embedded clauses (cf. (2), whereby *pročitam*$_{\text{READ.1.sg.PresPF}}$ 'I read' is licensed inside a clause of purpose introduced by the complementizer *da*),[8] interrogative-negative contexts[9] (cf. *ne dođete*$_{\text{COME.2.pl.PresPF}}$ 'you don't come' and *ne probate*$_{\text{TRY.2.pl.PresPF}}$ 'you don't try' featured in the whimperative structure in (3)) or even chains of

er languages of the *Balkansprachbund*. In Modern Greek, for instance, present tense forms built from the aoristic stem (*synoptikoi* 'momentaneous') cannot occur in main clause contexts and can be licensed either in embedded environments introduced by the subjunctive marker *na* (e.g. *Mporó na alláxō*$_{\text{CHANGE.1.sg.PresPF}}$ *tis glósses óses forés thélō* 'I can change languages as many times as I want') or, as emphasised by an anonymous reviewer, as complements of the future marker *tha* (e.g. *Tha to peis*$_{\text{SAY.2.sg.PresPF}}$ 'You will say it'; cf. Giannakidou 2009). This last use bears some resemblance to the distribution of Pres^{PF} forms in the Serbian variety of BCS, which can replace the bare infinitive and be licensed by the complementizer *da* after an inflected proclitic form of the (volitional) future marker (e.g. *On će* [*da dođe*]$_{\text{COME.3.sg.PresPF}}$ 'He will come' instead of *On će doći*$_{\text{COME.InfPF}}$). The possible reasons for the difference between BCS and Modern Greek on one hand, the other South Slavic languages on the other are left for future research.

8 It must be added, however, that the licensing of Pres^{PF} forms in BCS subordinate clauses is subject to further restrictions. Todorović (2017, 88-90) points out that, on the one hand, Pres^{PF} forms cannot occur in propositional complements of intensional verbs such as *v(j)erovati* 'to believe', at least when their temporal argument is not ordered according to a relation of anteriority (<) or posteriority (>) with respect to the UT, hence the ungrammaticality of sentences such as **V(j)eruje da Jovan prevede*$_{\text{TRANSLATE.3.sg.PresPF}}$ *p(j)esmu* (intended: '(S)He believes that Jovan has translated a poem (just now)'). In addition, as a result of a more general constraint on the distribution of PF (which possibly holds cross-linguistically), in the Serbian variety of BCS Pres^{PF} forms cannot be licensed as complements of phasal verbs such as *počinjati* 'to begin' (**Počinjem da prevedem*$_{\text{TRANSLATE.1.sg.PresPF}}$ *pesmu*, intended: 'I'm beginning to translate the entire poem (just now)'). Generally speaking, as pointed out by an anonymous reviewer, predicates that select either subjunctive or infinitive forms in Romance languages can license Pres^{PF} forms in BCS. This is, for instance, the case of verbs such as *pokušavati* 'to try', whose tenseless complements do not yield a future interpretation whatsoever (**Pokušava da sutra prevede*$_{\text{TRANSLATE.3.sg.PresPF}}$ *pesmu*, intended: '(S)He is trying to translate the entire poem tomorrow') and can always enter an aspectual opposition with the corresponding Pres^{IPF} forms. On the other hand, predicates selecting indicative forms in Romance languages block the licensing of Pres^{PF} forms in BCS. Among them are assertive predicates such as *tvrditi* 'claim', e.g. **Tvrdim da Ivan dođe*$_{\text{COME.1.sg.PresPF}}$ (intended: 'I claim/am claiming that Ivan is coming').

9 Pres^{PF} forms can appear in other nonveridical or anti-veridical contexts as well. With a small group of telic eventualities (e.g. *dati* 'to give', *pustiti* 'to let', *dozvoliti* 'to allow', *doći* 'to come'), negative Pres^{PF} forms are regularly licensed. These forms, which fall outside the scope of this paper, seemingly behave like light verbs and are prototypically marked with 1st (singular and plural) or 2nd (singular and plural) bound person-number indexes, as in (i):

> (i) E, vidite, ja *ne dam*$_{\text{GIVE.1.sg.PresPF}}$ taj trenutak, tu epohu, tu kristalnu kocku vedrine, "svečanost tragične vedrine", *ne dam*$_{\text{GIVE.1.sg.PresPF}}$ da to nestane. [HNK, vj20021110ku03]
> 'So, you see, I won't give away that moment, that age, that "crystal cube of serenity" [Kristalna kocka vedrine – a house of culture in Sisak], the "celebration of a tragic serenity" [a theatre play by Tomislav Durbešić], I won't let it disappear'.

Vojvodić's (2015, 55-6) hypothesis that BCS elliptical constructions such as *ne dam* are derived via transformation of future-oriented negative volitional sentences of the type *Neću* [*da dam*] (lit. 'I won't give (it)') requires a more thorough examination.

habitual eventualities which however need to be explicitly marked as such[10] (in (4) the non-actuality of the eventualities realised with the Pres^{PF} forms *popijem*$_{\text{DRINK.1.sg.Pres}PF}$ 'I drink' and *iscijedim*$_{\text{SQUEEZE.1.sg.Pres}PF}$ 'I squeeze' is signalled by the complex temporal adjunct *svaki dan <nakon jela>* 'every day <after meals>' and by the conjoined adjunct *ujutro i navečer* 'in the morning and in the evening'):[11]

10 Main clause (Štokavian) Pres^{PF} forms can also be licensed in isolation, as long as the temporal anchoring of the clause remains unordered with reference to the UT. Unlike West and East Slavic languages, however, in contemporary BCS modal (dynamic) readings of Pres^{PF} have been – to the best of my knowledge – replaced entirely either by overt modal constructions governed by *moći* 'can' or by (stylistically unmarked) Pres^{IPF} forms, cf. (ii):

(ii)? On u jednoj ruci *digne*$_{\text{LIFT.3.sg.Pres}PF}$ (može [da *digne*]$_{\text{LIFT.3.sg.Pres}PF}$ / može *dići*$_{\text{LIFT.Inf}PF}$ / *diže*$_{\text{LIFT.3.sg.Pres}IPF}$) 25 kg. (Galton 1976, 92)
'He can lift 25 kg with one arm'.

Other similar contexts include popular sayings (cf. (iii)) and semi-lexicalised expressions (cf. (iv)). Interestingly, in both cases 3rd p. sg. Pres^{PF} forms can freely alternate with Aor^{PF} forms, which in some cases are homographs and can thus be distinguished only prosodically (cf. *dôđe*$_{\text{COME.3.sg.Pres}PF}$ 'he/she comes' vs. *dóđe*$_{\text{COME.3.sg.Aor}PF}$ 'he/she came'). Note also that Aor^{PF} forms in the given contexts do not trigger any anterior reading whatsoever:

(iii) Ko *se* dima *ne nadimi*$_{\text{SMOKE UP.3.sg.Pres}PF/\text{3.sg.Aor}PF}$, taj *se* vatre *ne nagrije*$_{\text{HEAT UP.3.sg.Pres}PF}$ / nagreja$_{\text{3.sg.Aor}PF}$.
'Results demand sacrifice (lit. 'He who does not catch the smoke catches no heat from the fire')'.

(iv) On mi *dođe*$_{\text{COME.3.sg.Pres}PF}$ brat. (Dickey 2000, 199)
'He is my brother (lit. 'He comes to me as a brother')'.

Additionally, in several Kajkavian and Čakavian dialects Pres^{PF} forms can also be reinterpreted as aspectual futures, most frequently alongside a temporal operator such as *sutra* 'tomorrow': compare *sutra dođem*$_{\text{COME.1.sg.Pres}PF}$ 'Tomorrow I'll come', *sutra ti javim*$_{\text{LET KNOW.1.sg.Pres}PF}$ 'I'll let you know tomorrow', *sutra naprintam*$_{\text{PRINT.1.sg.Pres}PF}$ 'I'll print tomorrow' and the like (I am thankful to Mladen Uhlik and an anonymous reviewer for providing me with these as well as similar other examples). As for Štokavian areal variation, Polovina (1985, 98-9) illustrates an interesting spoken exchange between two Belgrade-based youngsters around mid-eighties where an apparently independent Pres^{PF} form seems to be freely licensed in a future-oriented context (*Dobro, izađeš*$_{\text{COME OUT.2.sg.Pres}PF}$ *u 7, kažeš imam autobus u 7.30…* 'Fine, you come out at 7 o'clock, you say I've got my bus at 7.30…'). However, a closer look at the communicative situation reveals that the speaker is revisiting a potential future schedule for their interlocutor, in a way similar to the so-called narrative future (Dickey 2000, 149-54), which in turn is usually considered a peculiar instantiation of the historical present template. Since perfective aspectual futures are virtually ungrammatical in Štokavian dialects, then, a comprehensive analysis of these contexts falls outside the scope of the present paper.

11 It should be kept in mind that in both interrogative-negative and explicitly habitual contexts, Pres^{IPF} forms (*dolaziš*, *pijem*, *cijedim*) can always occur. Unlike Pres^{IPF} forms however, the use of Pres^{PF} forms in examples like (3) is linked to the emergence of particular pragmatic inferences which have to do with the contextual actualisation either of circumstantial or dynamic modal meanings (among others Tanasić 1996, 169-70), while in examples like (4) the use of Pres^{PF} forms is generally preferred with telic eventualities which are measured out by their (referential) internal argument (among others Dickey 2000, 68-71; Vojvodić 2019, 17).

(2) Kada sam stigao kući, seo sam da *pročitam*$_{\text{READ.1.sg.Pres}^{PF}}$ nekoliko radova koje su naši mali domaćini napisali na času veronauke povodom Vaskrsa. [SrpKor, poli100402.txt, Politika (02.04.2010.)]
'Once I came back home, I sat and read some works which were written by our little hosts during their religion hour for Easter'.

(3) Mnogo prirodnih ledenih zidova svih stepena težine Vas očekuje u i oko Livinja. Zašto *ne dođete*$_{\text{COME.2.pl.Pres}^{PF}}$ i *probate*$_{\text{TRY.2.pl.Pres}^{PF}}$? [srWaC.4276.7, #4275]
'Lots of ice walls with all degrees of difficulty await you in and around Livigno. Why don't you come and try yourself?' (⊃ 'Come and try! You are invited!').

(4) "Svaki dan nakon jela *popijem*$_{\text{DRINK.1.sg.Pres}^{PF}}$ jedno Tomislav pivo, a ujutro i navečer čašicu žestokog pelinkovca. U kavu obavezno *iscijedim*$_{\text{SQUEEZE.1.sg.Pres}^{PF}}$ pola limuna, i doktor, vjerujte, meni ne treba […]" [HNK, gs20040327sb29520]
'Every day after meals I drink a bottle of Tomislav beer, while in the morning and in the evening [I drink] a shot of Pelinkovac. In my coffee I squeeze half a lemon by default, and believe me, I don't need any doctor'.

However there seem to be three apparent exceptions to this higher-level constraint applied to present-tensed main clauses. The first one involves the so-called abusive metonymic performatives (Dickey 2015), as shown in (5) below:

(5) Da prevedem na srpsko-hrvatski-bošnjački: *poserem se*$_{\text{SHIT.1.sg.Pres}^{PF}}$ na svoj život u državi u kojoj ste vi ministar. (Dickey 2015, 254)
'Let me put it in Serbo-Croato-Bosniak: I shit on my life in a state in which you are a minister'.

The second one involves the aspectual marking of a special subtype of running commentaries, live demonstrations, such as in the recipe instructions reported in (6):

(6) *Zapalim*$_{\text{SET ON FIRE.1.sg.Pres}^{PF}}$ joj kožu na vratu, pa *preklopim*$_{\text{FOLD.1.sg.Pres}^{PF}}$ na leđa i *učvrstim*$_{\text{FASTEN.1.sg.Pres}^{PF}}$ čačkalicama. Potom grabilicu *napunim*$_{\text{FILL.1.sg.Pres}^{PF}}$ rakijom lozovačom, *zagrijem*$_{\text{WARM UP.1.sg.Pres}^{PF}}$ i *zapalim*$_{\text{SET ON FIRE.1.sg.Pres}^{PF}}$, te *polijem*$_{\text{POUR.1.sg.Pres}^{PF}}$ puricu. [HNK, vj20091226pis010]
'I sear the skin [of the turkey] on the neck, then I fold it on the back and I fasten it with toothpicks. Then I fill the ladle with lozovača [home-made grape rakija], warm it up and set it on fire, then I pour it on the turkey'.

The third one involves the possible licensing of Pres^{PF} forms in a particular type of nonveridical context featuring the epistemic operator *možda* 'maybe', as in (7):

(7) Nakon teniske karijere, posvetit ću se humanitarnom radu, a možda *nađem*$_{\text{FIND.1.sg.Pres}^{PF}}$ i ženu s kojom ću zasnovati obitelj, kazao je 31-godišnji Chang. [HNK, vj20030202sp08]

'After my tennis career is over I'll devote myself to voluntary work, and maybe find a woman I can start a family with, said the 31-year-old Chang'.

In this paper it is claimed that these three contexts of Pres^{PF} use do not constitute an exception to the pattern we regularly find in Štokavian dialects. On the contrary, they can be subsumed under a unified analysis which, building on some theoretical tenets advocated by the so-called neoperformative hypotheses (among others Speas, Tenny 2003), aims at mapping the feature of (cognitive) control in the higher portion of the clausal spine, above CP. The article is structured as follows: in § 2 I will discuss the aforementioned 'exceptions' in more detail with regard to the suppression of independent Pres^{PF} forms from Štokavian main clauses; in § 3 I will reframe the issue at the syntax-pragmatics interface, providing some evidence for a proper syntacticization of the parameter of (cognitive) control inside SpeechActP; in §4 I will draw some preliminary conclusions.

2 Breaking the Law? Anomalous Pres^{PF} Forms in Main Clauses

In this section I will address each of the three case studies of this paper: abusive metonymic performatives (§ 2.1), live demonstrations (§ 2.2), and nonveridical contexts featuring *možda* (§ 2.3).

2.1 Abusive Metonymic Performatives

The label *abusive metonymic performatives* (henceforth AMP; the term was first coined in Dickey 2015, 254-9; 268-70) applies to a closed class of obscene and disparaging phrases which, although lacking most of the characteristics ascribable to explicit performative utterances,[12] share some interesting illocutionary similarities

12 By *explicit performative utterance* we define a token of the utterance set U(x), actualising a proposition P(x) with an explicit performative verb in its abstract logical form, which amounts to the action expressed in P(x) when uttered under felicitous circumstances by an agent who is endowed with (or granted) the necessary authority. An explicit performative utterance has to be self-referential, pragmatically felicitous and grammatically well-formed; moreover, it must not satisfy any inherently antiperformative condition. Finally, it has to be categorically defined by the speaker in the speech act event and recognised as such by the hearer (Eckardt 2012, 47-8).

Unlike AMPs, the prototypical tempo-aspectual form licensed in explicit performative utterances (cf. the commissive illocutionary act exemplified by *obećavati*^{IPF} 'to promise' in (v)) is Pres^{IPF}, while Pres^{PF} forms are regularly ruled out. It should be mentioned, however, that Pres^{IPF} forms can be opposed by Fut^{PF}, although only to a limited extent and as long as certain preconditions at the syntax-pragmatic interface (which cannot be explored in detail here) are met:

with them as well. In other words, AMPs, which refer to purely resultative eventualities and are typically a main clause phenomenon, do not constitute a simple depiction of the eventuality token actualised in the sentence; they also figuratively perform on the insulted hearer the action that the utterance refers to (hence the term 'metonymic'). Relevant examples of Štokavian AMPs are reported in (5) above and (8) below:

(8) *Nabijem*$_{\text{STICK.1.sg.PresPF}}$ te na kurac! (Dickey 2000, 200)
'I thrust you onto my cock!'

Although AMPs are coded with Pres^{PF} forms in other Slavic languages, most notably those of the Western branch (which is in line with the Slavic East-West aspect division laid out in Dickey 2000), such marking stands out as peculiar in BCS. Comparing this usage with some exceptional cases of Pres^{PF} forms used performatively – both in Old Church Slavonic and Old Slovene – Dickey (2015, 268-70) claims that perfective AMPs should be considered a grammatical archaism inherited from Common Slavic, and credits 'absolute control of the speaker' as a key factor for their aspectual marking.

Even though in Dickey (2015) the working definition of 'absolute control of the speaker' is not formalised, I have argued elsewhere (Biasio, forthcoming) that the author's claim is sound, and that there is both historical, formal (i.e. morphosyntactic), as well as even evolutionary evidence to back it up.[13] It is also worth mentioning that AMPs can neither be coded as Pres^{IPF} in their derogatory, performative-like function (#*Nabijam*$_{\text{STICK.1.sg.PresIPF}}$ *te na kurac* could only report an ongoing event, if any), nor can they be marked with person-number bound indexes different from 1st sg. (**Nabiješ*$_{\text{STICK.2.sg.PresPF}}$ *ga na kurac!*) or negated (**Ne nabijem*$_{\text{STICK.1.sg.PresPF}}$ *te na kurac!*). Additionally, other outrageous expressions that do not denote purely resultative eventualities resist marking with Pres^{PF} forms: cf. *Jebem*$_{\text{FUCK.1.sg.PresIPF}}$ (**Odjebem*$_{\text{FUCK OFF.1.sg.PresPF}}$ / **Zajebem*$_{\text{SCREW UP.1.sg.PresPF}}$) *ti mater* 'Fuck you', lit. 'I fuck (*fuck off / *screw up) your mother'. The data seemingly suggests that the parameter of control may play a pivotal role in the aspectual coding of resultative eventualities in egophoric utterances.

(v) Navijačima *obećavam*$_{\text{PROMISE.1.sg.PresIPF}}$ (**obećam*$_{\text{PROMISE.1.sg.PresPF}}$ / ?*ću obećati*$_{\text{PROMISE.1.sg.FutPF}}$) maksimalno zalaganje, borbu za napad i svaku odbranu i to će, valjda, doneti željeni rezultat. [SrpKor, poli010303.txt, Politika (03 March 2001)]
'I promise our supporters absolute commitment and a relentless fight for attack and defence and this, I hope, will bring about the desired outcome'.

13 Although I believe it is safer to assume that AMPs and other 'perfective' performatives in Old Church Slavonic and Old Slovene actually belonged to a general aspect (Biasio, forthcoming), in contemporary BCS such preverbed predicates are unambiguously interpreted as PF.

2.2 Live Demonstrations

There is a large body of literature concerning the aspectual marking of running commentaries in Slavic languages (cf. among others Galton 1976, 95-7; Dickey 2000, 155-74), which are usually analysed as a peculiar instantiation of the historical present – irrespective of the significant variation they display both in their temporal anchoring and their extradiegetic properties. Instead, for the sake of clarity, in this subsection only a particular subset of running commentaries is taken into account, i.e. live demonstrations (experiments, recipes etc.). As Dickey (2000, 161) appropriately points out, live demonstrations are interesting because of their complex temporal nature. Phrased otherwise, they are at the same time both (non-actual) generalisations over a definite set of propositions (possibly built upon a deficient anankastic-oriented model of the type 'if you want to X, you have to Y') and the report of their ongoing actualisation in the utterance (i.e. the eventualities are actually unfolding at the UT).

As shown in (6) above and again in (9) below, Pres^{PF} forms can be freely licensed in the given context. They can be contextually replaced by PF imperative or infinitive forms (which however constitute different types of speech acts) or even alternate with Pres^{IPF} forms, especially with unbounded eventualities. Person-number bound verbal indexes may vary, mostly (but not exclusively) along the conjugation axis 1st p. sg./pl.:[14]

(9) Pecnicu *zagrijemo*$_{\text{WARM UP.1.pl.Pres}^{PF}}$ na 250 stepeni i *pecemo*$_{\text{BAKE.1.pl.Pres}^{IPF}}$ lepinje na drugoj od vrha mrezi […] Pecene lepinje *pokriti*$_{\text{COVER.Inf}^{PF}}$ krpom i *ostaviti*$_{\text{LEAVE.Inf}^{PF}}$ desetak minuta da se ohlade prije rezanja. [sic!] (https://www.coolinarika.com/recept/bosanske-lepinje-f9c5e5f6-63db-11eb-8e23-0242ac120027) 'We preheat the oven to 250° and bake our buns on the rack (not the one on the top) […] Cover the baked buns with a cloth and let them cool down for some ten minutes before cutting'.

[14] Variation in number (e.g. from singular to plural) is likely to be triggered by a different structural configuration of discourse roles, which can be brought about in turn by a variation in the intensity of the illocutionary force assigned to the utterance. For instance, in her analysis of Russian 'mental performatives' (i.e. a subclass of expositives which are frequently used to mold the rhetorical structure of written texts and public speeches in order to make them logically consistent, e.g. *opredelit'*^{PF} 'to identify', *predpoložit'*^{PF} 'to suppose', *podčerknut'*^{PF} 'to emphasise' and the like), Rjabceva (1992, 24) argues that the use of the 1st p. pl. of the perfective non-past (*opredelim…*, *predpoložim…*, *podčerknem*) is instrumental in underlining the impositive illocutionary character of such predicates, in that the speaker, while unfolding their reasoning in a step-by-step fashion, expects their addressee(s) to follow the same logical steps and infer the same conclusion(s).

Dickey (2000, 168) further notes that in BCS this kind of PF-marked live demonstration cannot be overtly anchored to the UT by means of temporal operators such as *sad*(*a*) 'now' (here we gloss over some technicalities on the multilayered semantic nature of such deictic elements). For Pres^{PF} forms to be licensed, reference needs to be made instead to the steps of the whole procedure as it is performed by the speaker. Moreover, if the eventualities are not presented as bounded, Pres^{IPF} forms can replace Pres^{PF}, cf. (10):

(10) Najpre *uzmem*$_{\text{TAKE.1.sg.Pres}^{PF}}$ (*uzimam*$_{\text{TAKE.1.sg.Pres}^{IPF}}$) bocu sodium nitrata i *sipam*$_{\text{POUR.1.sg.Pres}^{IPF}}$ sadržinu u retortu; onda *upalim*$_{\text{LIGHT.1.sg.Pres}^{PF}}$ (*palim*$_{\text{LIGHT.1.sg.Pres}^{IPF}}$) Bansenovu lampu i *zagrejem*$_{\text{WARM UP.1.sg.Pres}^{PF}}$ (*zagrejavam*$_{\text{WARM UP.1.sg.Pres}^{IPF}}$) do tačke ključanja… (Dickey 2000, 168)
'First I take the flask of sodium nitrate and pour the contents into this beaker; then I light the Bunsen burner and heat it to a boil…'

Dickey (2000, 173) goes on to suggest that the ban on the contextual presence of *sad*(*a*) alongside PF-marked live demonstrations is linked to the reportive character of the speaker's speech, who would be in control of the flow of the events as they present them to their addressee(s) who are already ordered in a causal chain. Again, then, the aspectual coding of externally and/or internally bounded eventualities in semi-actual contexts seems to be dependent on the role of the point of view (POV) of a sentient individual in egophoric utterances.

2.3 Contexts with *možda* 'Maybe'

Pres^{PF} forms are regularly found in present-tensed main clause environments featuring the epistemic operator *možda* 'maybe', as in (7) above or (11) below. Variation can occur with respect to aspectual marking (unbounded eventualities, or presented as such, can be realized with Pres^{IPF} forms) and person-number bound verbal indexes other than 1st p. (sg./pl.), as shown in (12):

(11) Kad je mogao Dule, zašto ne bismo mi koristili taj recpet [*sic!*]. Možda *pozovem*$_{\text{INVITE.1.sg.Pres}^{PF}}$ Vujoševića da provede dva meseca s nama. [srWaC.883681.21, #883680]
'If Dule could, why wouldn't we use this recipe? Maybe I can invite Vujošević to spend two months with us'.

(12) Prvo je i osnovno pravilo da sebe učiniš srećnom osobom, da budeš srećan čovek. Onda imaš šanse i nekog drugog da usrećiš, možda *napraviš*$_{\text{MAKE.2.sg.Pres}^{PF}}$ i *odgajiš*$_{\text{RAISE.2.sg.Pres}^{IPF}}$ neku srećnu decu. [srWaC.1189461.10, #1189460]
'The first and most basic rule is to make yourself a happy person, to be happy. Then you can have chances to make someone else happy, maybe you can make and raise some happy children'.

This type of Pres^{PF} use, which is sometimes mentioned in passing even in prescriptive grammars (cf. Barić et al. 1997, 409), is worthy of attention as, unlike AMPs and running commentaries, it involves purely nonveridical contexts, seemingly aligning itself with interrogative-negative structures (cf. (3)) and negative presents (cf. (i)). Should this pattern be consistent, we would expect Pres^{PF} forms to be licensed even alongside other epistemic operators semantically akin to *možda*, such as *verovatno*$_{SR}$/ *vjerojatno*$_{HR}$ 'probably' or *moguće* 'possibly', but this is not the case as (13) clearly shows:

(13) – Za sledeći put, i neki vikend... – Da, možda (*verovatno / *moguće) *dođem*$_{COME.1.sg.Pres^{PF}}$ na vikend ovde, da. [srWaC.1128436.4, #1128435]
'– Next time, maybe some weekend… – Yes, maybe I'll come here for a weekend, sure'.

On the basis of (surface) syntactic and prosodic evidence, one could be inclined to look at *mòžda* as the result of the univerbation of the epistemic modal *mȍže*$_{CAN.3.sg.Pres^{IPF}}$ (from *mȍći* 'can') and a 'hidden' complementizer *da*, which, following Veselinović (2019, 46-58), occupies a position in SpecC and takes a clausal complement (much in the spirit of the 'declarative' *da* proposed in Vrzić 1996).[15] While this may very well account for the diachronic derivation of the operator, it fails to provide a satisfactory structural explanation as it does not provide reasons for the general acceptability of (14a) and (15),[16] nor does it clarify why the same pattern does not hold for *vàljda* 'probably', 'I guess' (cf. (16)-(18)).[17]

15 Note, however, that there is no general agreement on the structure of the clausal architecture in BCS and that alternative hypotheses involving different makeups of the BCS clausal spine have been put forth as well (cf., for instance, Todorović 2012, where it is proposed that the two *da* in BCS are not complementizers *senso strictu*, but rather operators introducing either veridical or nonveridical complements). Note also that Kovačević (2008, 197-202) points out that *možda* alongside Pres^{PF} can license Fut^{PF} and Fut^{IPF} forms as well (*Možda kupim*$_{BUY.1.sg.Pre^{PF}}$ */ ću kupiti*$_{BUY.1.sg.Fut^{PF}}$ */ budem kupio*$_{BUY.1.sg.FutI-IPF.masc.}$*/-la*$_{1.sg.FutII^{PF}.fem.}$ *kuću na moru*). A further analysis of the range of variations within the pattern, particularly in relation to the licensing properties of *valjda*, is left for future research.

16 According to my consultants, both (14a) and (15) are quite marginal in literary Štokavian and occur with a relative frequency in oral speech (according to an anonymous reviewer, (14a) occurs most naturally as an elicited answer from a prior context). The majority of the Štokavian native speakers I consulted claimed that (14a) can be better parsed as an elicited answer to a previous context, although there is plenty of evidence for such contexts appearing in isolation as well, cf. (vi) below:

(vi) Možda da jumbo-plakate *dignemo*$_{RAISE.1.pl.Pres^{PF}}$ i kat više, pa uz onu narodnu "daleko od očiju, daleko od srca", uživamo u nasmiješenim modelima s reklama. [HNK, gs20030701os21879]
'Maybe we should raise the billboards one floor up and then, in accordance with the old saying "out of sight, out of mind", we enjoy these smiling models from the advertisements'

17 According to my consultants, (16) can be uttered either if *kupim*$^{Pres^{PF}}$ is replaced by *kupujem*$^{Pres^{IPF}}$ or if a relative temporal adverb such as *dotad* 'by then' is added to the

Indeed, *vàljda* displays the same prosodical contour as *mòžda* and can be diachronically derived in a similar manner (i.e. via univerbation of *vàlja* $_{3.sg.Pres.IPF}$ from *vàljati* 'to be worth' and *da*):[18]

(14) a. Možda da *kupim*$^{Pres.PF}$ kuću na moru. (Intended: 'Perhaps I should buy a beach house')
b. *Može da da *kupim*$^{Pres.PF}$ kuću na moru.

(15) Možda može da *kupim*$^{Pres.PF}$ kuću na moru.

(16) ?/*Valjda *kupim*$^{Pres.PF}$ kuću na moru. (Intended: 'I guess I should buy a beach house')

(17) ?/*Valjda da *kupim*$^{Pres.PF}$ kuću na moru.

(18) ?/*Valjda valja da *kupim*$^{Pres.PF}$ kuću na moru.

Interestingly, Mirič (2004, 218) points out that in the Serbian variant of the BCS continuum both *možda* and *verovatno* can be used if the realization of the eventuality actualized in the utterance rests entirely on the speaker, while *valjda* implies that such realization is not (completely) under the speaker's control:

(19) – Hoćeš li ti doći večeras? – Možda. / – Verovatno. / – *Valjda.[19]
'– Are you coming this evening? – Maybe. / – Probably.'

context (the issue why a Pres^{PF} form here could be interpreted as a future perfect cannot be addressed here). (17) and (18) are acceptable only as interrogatives.

18 Although both *možda* and *valjda* can be used either as full modal (epistemic-evidential) operators or parenthetical elements (Zvekić-Dušanović 2019), the two (classes of) items are still assigned different intonational patterns. However, the difference in the epistemic commitment associated with the two operators (with *valjda* being epistemically weaker and, possibly, doxastically stronger than *možda*) does not seem to affect their syntactic properties.

19 It goes without saying that in (19) *valjda* is not ungrammatical *per se*, however its distribution can only be generalised to contexts where the feature of the speaker's control is not prominent. In order to better exemplify the contrast between operators I resort to the following example (vii) (I thank an anonymous reviewer for suggesting it to me):

> (vii) Pavlović se nada da nije kraj: Valjda ću odigrati još neki meč za Partizan. (`http://beta.admin.mozzartsport.com/fudbal/vesti/pavlovic-se-nada-da-nije-kraj-valjda-cu-odigrati-jos-neki-mec-za-partizan/354132`)
> 'Pavlović hopes this is not the end: I will probably play some more games for Partizan' (⊃ 'I hope / suppose I will play some more games for Partizan, but it does not depend on me').

On a side note, *možda* can also stack together with *valja*, as in the following example (viii):

> (viii) Da, možda valja uključiti i moldavski i ukrajinski. (`https://twitter.com/srdjan-drago/status/1290913720856698880`)
> 'Yes, probably it is worth including both Moldavian and Ukrainian'.

This would suggest that the licenseability of $Pres^{PF}$ forms is not (exclusively) triggered by either the compositionally derived syntactic properties of *možda* (with reference to, for instance, $vjerojatno_{HR}$/ $verovatno_{SR}$ and *moguće*) or by the nonveridical flavour of the context (cf. the contrast with *valjda*). Another aspect seems to bring *možda*-contexts together with the other case studies under analysis, i.e. the prominent salience of the speaker's control-related POV. In the next section I will elaborate on this idea and argue that control, beside its semantic contribution, is a syntactically relevant element as well.

3 Control as a Variable in SpeechActP

The idea that linguistic components traditionally assigned to discourse-related modules (e.g. the role of discourse participants in the speech act event, other logophoric roles, and the notion of POV) should be encoded in the syntactic structure dates back to the early seventies, and would be brought to light some decades later on the basis of consistent crosslinguistic evidence. Speas and Tenny (2003, 320) were among the first to argue for the existence of a separate functional layer above CP, the so-called 'SpeechActP', where discourse roles (traditionally, the Speaker and the Hearer) are mapped and defined in terms of their structural positions.

In their joint paper, Speas and Tenny note that the restrictions on the inventory of phenomena linguistically dependent on the POV of a sentient individual tend to be crosslinguistically stable, which may hint at the hybrid (i.e. not entirely pragmatic) nature of the properties of such phenomena. They thus propose a further configurational structure in the scope of SpeechActP, a projection called 'Evaluation Phrase', which takes two arguments: the Seat of Knowledge (a "sentient 'mind,' who can evaluate, or process, or comment on the truth of a proposition", p. 332, henceforth SOK), mapped to the specifier position, and the Evidential Phrase, which is linked to the evidential inference(s) on the truth of a given proposition and mapped to the internal argument position (Speas, Tenny 2003, 327-31).

Since Speas and Tenny (2003), there has been much speculation on the exact composition of the inner layer of SpeechActP, including the number of functional projections included within the maximal projected structure and their respective ordering. In this paper I will adopt Zu's (2018, 73 ff., 101-2) slightly revised idea that the SOK appears in the specifier position of a Sentience Phrase and is bound by the base-generated Speaker in declaratives, rhetorical questions and quiz questions. Within this framework, I define cognitive control as the hierarchical and contingent (i.e. not given *a priori*) rela-

tion between the Speaker and the Hearer in the speech act event,[20] which can be modelled as an abstract variable specified at the level of the SOK, checked against by the Speaker (which may or may not be coindexed with the syntactic subject) and spelled out lower in the structure in AspP as PF.

The idea that the cognitive feature of the speaker's control may additionally have a syntactic salience stems from the following two observations. The first one echoes the treatment of speech-style particles and markers of politeness in Korean discussed in detail by Portner, Pak and Zanuttini (2019). Abstracting away from the (quite dense) details of their proposal, by replacing SpeechActP with another functional projection called 'cP' ('c' stands for 'context'), the authors argue that different allocutive pronominal forms, syntactically bound by the Hearer ('Interlocutor' in their terms), are spelled out via an operator-variable relationship mediated by a λ-abstractor type that would reflect the alternation of abstract features on the functional head c (Portner, Pak, Zanuttini 2019, 24-6). Judging from our data sample, control is always related to the Speaker notwithstanding the presence of an actual coindexing of the discourse participant role with the syntactic subject (cf. above (12)). In other words, control always seems to be tightly linked with the expression of the Speaker's POV in egophoric utterances. Following Zu's (2018) account, we can thus assume that in declarative contexts the SOK is always bound by the Speaker and that control, which can be overtly expressed whenever a hierarchical relationship between (at least) two participants in the same speech act event is established, seemingly functions as a variable mediated by a λ-operator at the level of the SOK, as tentatively sketched out in (20):

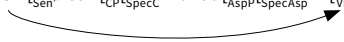
(20) $[_{SpeechActP}[Speaker[_{SenP}[_{SpecSen}SOK[_{Sen}\lambda con]]]]]$

The second observation boils down to the correlation between control and perfective morphology. Dickey's (2000, 25-7) classical proposal of a Slavic East-West aspect division on the basis of the prototypical meaning assigned to PF places BCS in a transitional zone, with Štokavian varieties sharing more commonalities with the Eastern group. The crucial point here is that control can be seen as a contextually determined effect stemming from the aspectual semantic core of PF in Eastern Slavic languages, namely temporal definiteness, i.e. the assignment of an eventuality to a uniquely determined point on the time axis, over which the speaker would then yield full control (Dickey 2018). If we assume that in Štokavian the variable

20 A comprehensive discussion of the pragmatic variables of the participant structure lies beyond the scope of the present paper.

for control is specified at the level of the SOK and thus bound by the Speaker, it is possible to claim that the SOK enters an agreement relation with the aspectual projection at a distance (much like probe-goal agreement), consequently triggering the spell-out of the morphological element(s) associated with PF.[21] If this hypothesis is indeed on the right track, an example like (21a) would then have the structural representation in (21b) (note that only the relevant projections of the clausal spine are featured below):

(21) a. Možda $kupim_{1.sg.Pres_{PF}}$ kuću na moru.
b. $[_{SpeechActP}$[Speaker$[_{SenP}[_{SpecSen}$SOK$[_{Sen}$,λ*con*$[_{CP}[_{SpecC}$možda$[_{AspP}[_{SpecAsp}$PF$[_{VP}[_{V}$kupiti$[_{DP}$ kuću na moru]]]]]]]]]]]



4 Conclusions

In this paper, I presented a unified analysis for several contexts featuring the (apparently) anomalous licensing of $Pres_{PF}$ forms in BCS present-tensed main clauses, i.e. abusive metonymic performatives (§ 2.1), live demonstrations (§ 2.2), and nonveridical contexts with *možda* (§ 2.3) *inter alia*. I claimed that for all these cases the feature of (cognitive) control, which can be defined in terms of the hierarchical relation between the Speaker and the Hearer – and thus encoded in the syntactic structure as an abstract variable at the level of the SOK in SpeechActP – bears special relevance. The association between control and perfective morphology can be modelled in terms of a probe-goal agreement between the SOK and the aspectual projection and can be explained as a semantic extension of the prototypical function assigned to PF in the Štokavian varieties of the BCS continuum, i.e. temporal definiteness.

Regrettably, a number of pertinent questions could not be answered in this survey. For instance, it remains unclear why $Pres_{PF}$ forms cannot be licensed in similar semi-actual contexts, whereby control is also shown to play a significant role, e.g. performative utterances (cf. (v)).[22] Furthermore, the role of verbal tenses (more spe-

21 It is important to mention that in BCS an eventuality can be realised as PF only if it is externally and/or internally bounded. Internal boundedness here corresponds to what is commonly labelled '(inherent) telicity' (resultativity). Externally bounded eventualities do not need to be telic.

22 It may be the case that, differently from AMPs, the temporal anchoring of performative utterances is assigned a different ordering with reference to the UT. It may also be the case that in regular performative utterances such as (v) the hierarchical relation between the speaker and the hearer is modelled in a slightly different fashion with reference to AMPs, with the speaker still yielding control over the eventuality in their scope (domain) yet being unable to intrude in the hearer's domain and take

cifically, the aorist and the two future tenses) and their interaction with the higher layer of the syntactic structure should be taken into greater consideration. Finally, dialectal internal variation should be reviewed more thoroughly so as to reinforce the core principles of the main argument. All these issues and their interplay are thus worthy topics for future research.

List of abbreviations

Aor^{PF}	perfective aorist
fem	feminine
Fut^{PF}	perfective future
Fut^{IPF}	perfective future perfect
HNK	Hrvatski nacionalni korpus
HR	Croatian
Inf	infinitive
IPF	imperfective
masc	masculine
p	person
PF	perfective
pl	plural
POV	point of view
$Pres^{PF}$	perfective present
$Pres^{IPF}$	imperfective present
sg	singular
SOK	Seat of Knowledge
SR	Serbian
SrpKor	Korpus savremenoga srpskog jezika
srWaC	Serbian Web Corpus
UT	Utterance Time

it over completely (Dickey, personal communication). I am grateful to Stephen Dickey for the fruitful discussion on these data.

Bibliography

Bach, E. (1986). «The Algebra of Events». *Linguistics and Philosophy*, 9(1), 5-16. https://doi.org/10.1007/BF00627432.

Bailyn, J.F. (2010). «To What Degree Are Croatian and Serbian the Same Language? Evidence from a Translation Study». *Journal of Slavic Linguistics*, 18(2), 181-219. https://doi.org/10.1353/jsl.2010.0000.

Barić, E. et al. (1997). *Hrvatska gramatika* (Croatian Grammar). 2nd ed. Zagreb: Školjska knjiga.

Biasio, M. (forthcoming). «Slavic Performatives Were Originally Marked for General Aspect». Gherbezza, E.; Laskova, V.; Perissutti, A.M. (a cura di), *Atti del VIII Incontro di Linguistica Slava*. Roma: Aracne.

Demirdache, H.; Uribe-Extebarria, M. (2004). «The Syntax of Time Adverbs». Guéron, J.; Lecarme, J. (eds), *The Syntax of Time*. Cambridge; London: The MIT Press, 143-79.

De Wit, A. (2017). *The Present Perfective Paradox Across Languages*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198759539.001.0001.

Dickey, S.M. (2000). *Parameters of Slavic Aspect: A Cognitive Approach*. Stanford: CSLI.

Dickey, S.M. (2015). «The Aspectual Development of Performatives in Slavic». *Zeitschrift für Slavische Philologie*, 71(2), 249-304.

Dickey, S.M. (2018). «Thoughts on the 'Typology of Slavic Aspect'». *Russian Linguistics*, 42(1), 69-103. https://doi.org/10.1007/s11185-017-9189-x.

Eckardt, R. (2012). «*Hereby* Explained: An Event-Based Account of Performative Utterances». *Linguistics and Philosophy*, 35(1), 21-55. https://doi.org/10.1007/s10988-012-9110-4.

Galton, H. (1976). *The Main Functions of the Slavic Verbal Aspect*. Skopje: Macedonian Academy of Sciences and Arts.

Giannakidou, A. (2009). «The Dependency of the Subjunctive Revisited: Temporal Semantics and Polarity». *Lingua*, 119(12), 1883-908. https://doi.org/10.1016/j.lingua.2008.11.007.

Haspelmath, M. (2020). «On the Difference between a Clause and a Sentence». Kibrik, A.A. et al. (eds), Кибрик, А.А. и др. (под ред.), *VAProsy jazykoznanija: Megasbornik nanostatej. Sbornik statej k jubileju V.A. Plungjana* ВАПросы языкознания: Мегасборник наностатей. Сборник статей к юбилею В (VAProsy jazykoznanija: A Mega-anthology of Nano-articles. Collected Papers for V.A. Plungjan's Birthday). А. Плунгяна. Moskva: Buki Vedi, 597-605.

Kovačević, M. Ковачевић, М. (2008). «Konkurentnost dvaju futura i perfektivnog prezenta u savremenom srpskom jeziku» Конкурентност двају футура и перфективног презента у савременом српском језику (Competition of the Two Future Tenses and the Perfective Present in Contemporary Serbian). *Zbornik Matice srpske za slavistiku* Зборник Матице српске за славистику (Matica Srpska Journal of Slavic Studies), 73, 195-211.

Leivada, E.; Murphy, E. (2021). «Mind the (Terminological) Gap: 10 Misused, Ambiguous, or Polysemous Terms in Linguistics». *Ampersand*, 8, 1-9. https://doi.org/10.1016/j.amper.2021.100073.

Leivada, E.; Westergaard, M. (2020). «Acceptable Ungrammatical Sentences, Unacceptable Grammatical Sentences, and the Role of the Cognitive Parser». *Frontiers in Psychology*, 11(364), 1-9. https://doi.org/10.3389/fpsyg.2020.00364.

Mirič, D. Мирич, Д. (2004). «Značenje častic*y valjda* i ee russkie ekvivalenty» Значение частицы ваљда и ее русские еквиваленты (The Meaning of the Particle valjda and Its Russian Equivalents). Anan'eva, N.E.; Karceva, Z.I. (eds) Ананьева, Н.Е.; Карцева, З.И. (под ред.), *Slavjanskij vestnik (K 70-letiju V.P. Gudkova)* Славянский вестник (К 70-летию В.П. Гудкова) [Slavonic Bulletin (For V.P. Gudkov's 70th Birthday)], 2. Moskva: MAKS Press, 215-24.

Polovina, V. Половина, В. (1985). «O upotrebi glagolskih vremena u savremenom srpskohrvatskom razgovornom jeziku» О употреби глаголских времена у савременом српскохрватском разговорном језику (On the Use of Verbal Tenses in Contemporary Spoken Serbo-Croatian). *Naučni sastanak slavista u Vukove dane* Научни састанак слависта у Вукове дане [Scientific Meeting of Slavists in (o 'During') Vuk's Days], 14(2), 97-103.

Portner, P.; Pak, M.; Zanuttini, R. (2019). «The Speaker-Addressee Relation at the Syntax-Semantics Interface». *Language*, 95(1), 1-36. https://doi.org/10.1353/lan.2019.0008.

Rjabceva, N.K. Рябцева, Н.К. (1992). «Mental'nye performativy v naučnom diskurse» Ментальные перформативы в научном дискурсе (Mental Performatives in Scientific Discourse). *Voprosy jazykoznanija* Вопросы языкознания (Topics in the Study of Language), 4, 12-28.

Speas, P.; Tenny, C. (2003). «Configurational Properties of Point of View Roles». Di Sciullo, A.M. (ed.), *Asymmetry in Grammar*. Vol. 1, *Syntax and Semantics*. Amsterdam; New York: John Benjamins Publishing Company, 315-44. https://doi.org/10.1075/la.57.15spe.

Tanasić, S. Танасић, С. (1996). *Prezent u savremenom srpskom jeziku* Презент у савременом српском језику (The Present Tense in Contemporary Serbian). Beograd: Institut za srpski jezik SANU.

Todorović, N. (Nataša) (2012). *The Indicative and Subjunctive 'da'-Complements in Serbian. A Syntactic-Semantic Approach* [PhD dissertation]. Chicago: University of Illinois.

Todorović, N. (Neda) (2015). «Tense and Aspect (In)Compatibility in Serbian Matrix and Subordinate Clauses». *Lingua*, 167, 82-111. https://doi.org/10.1016/j.lingua.2015.04.007.

Todorović, N. (Neda) (2017). «If You Can Show the Future, I Know What Your Are Made of: Aspect + Modal-Temporal Domain in Serbian». *Linguistica Brunensia*, 65(2), 87-102.

Veselinović, D. (2019). *The Syntax and Acquisition of Modal Verb Flavors* [PhD dissertation]. New York: New York University.

Vojvodić, D.P. Войводич, Д.П. (2015). «Soveršennyj vid i otricatel'nyj kontekst: vyraženie otkaza vypolnit' dejstvie» Совершенный вид и отрицательный контекст: выражение отказа выполнить действие (Perfective Aspect and Negative Context: Expressing Refusal to Perform an Action). Mišanin, J.A. (pod red.) Мишанин Ю.А. (отв. ред.), *Russkij jazyk v kontekste nacional'noj kul'tury (Materialy III Meždunarodnoj naučnoj konferencii, posvjaščennoj 200-letiju so dnja roždenija M. Ju. Lermontova, Saransk, 21-23 maja 2014 g.; Tom 1)* Русский язык в контексте национальной культуры (Материалы III Международной научной конференции, посвященной 200-летию со дня рождения М.Ю. Лермонтова, Саранск, 21-23 мая 2014 г.; Том 1) [The Russian Language within the Context of National Culture (Materials of the III International Scientific Conference Dedicated to the Two-hundredth Anniversary of M. Ju. Lermontov's Birth: Saransk, May 21-23 2014, Volume 1)]. Saransk: Izdatel'stvo Mordovskogo universiteta, 55-8.

Vojvodić, D.P. Војводић, Д.П. (2019). «O vidsko-vremenskoj sinonimiji i metafori u srpskom, ruskom i drugim slovenskim jezicima» О видско-временској синонимији и метафори у српском, руском и другим словенским језицима (On the Aspectual-Temporal Synonymy and Metaphor in Serbian, Russian and Other Slavic Languages). *Slavistika* Славистика (Slavistics), 23(1), 11-44. https://doi.org/10.18485/slavistika.2019.23.1.1.

Vrzić, Z. (1996). «Categorial Status of the Serbo-Croatian 'Modal' *da*». Toman J. (ed.), *Annual Workshop on Formal Approaches to Slavic Linguistics: The College Park Meeting 1994*. Ann Arbor: Michigan Slavic Publications, 291-312.

Zu, V. (2018). *Discourse Participants and the Structural Representation of the Context* [PhD dissertation]. New York: New York University.

Zvekić-Dušanović, D. Звекић-Душановић, Д. (2019). «O modalnoj rečci *valjda* u kontekstu» О модалној речци ваљда у контексту (On the Modal Particle valjda in Context). *Naučni sastanak slavista u Vukove dane* Научни састанак слависта у Вукове дане [Scientific Meeting of Slavists in (o 'During') Vuk's Days], 48(1), 169-78. https://doi.org/10.18485/msc.2019.48.1.ch14.

Balcania et Slavia
Vol. 1 – Num. 1 – June 2021

On the Diachrony of the Clitic Cluster in Bulgarian

Tsvetana Dimitrova
Institute for Bulgarian Language, Sofia

Abstract The article traces back the formation of the clitic cluster in Bulgarian starting from the Old Church Slavonic through Middle Bulgarian up to the Early Modern Bulgarian and beyond. It offers a hypothetical two-layer structure of the cluster – with the main layer consisting of a (pronominal) core and a (verbal) periphery, and a secondary layer hosting ('quasi-clitical') elements that exhibit, both diachronically and synchronically, a behaviour that is not strictly consistent with that of the clitical elements. The language material from three corpora shows that there was no change in the positions of the elements in the core, while the observed changes in the periphery are mainly due to changes in the set of the elements (as a result of the restructuring of the pronoun system and changes in the auxiliary system, as well as the loss of some early clitics, such as the discourse markers).

Keywords Bulgarian language. Clitics. Clitic cluster. Diachrony. Corpora.

Summary 1 Introduction. – 2 Clitics in Bulgarian. – 3 Clustering: The History. – 3.1 Discourse Clitics. – 3.2 Pronominal Clitics. – 3.3 Verbal Clitics and Clitic-Like Elements.

Edizioni Ca'Foscari

Peer review

Submitted	2021-07-21
Accepted	2021-10-15
Published	2020-12-20

Open access

Citation Dimitrova, T. (2021). "On the Diachrony of the Clitic Cluster in Bulgarian". *Balcania et Slavia*, 1(1), 67-96.

1 Introduction

Slavic clitics have been the focus of attention in a number of works in the recent decades (Franks, King 2000; Franks 2017; Zimmerling 2013; Migdalski 2016; among many others), with accounts also on the diachrony of cliticization in (one or more) Slavic languages.[1]

The diachrony of the Bulgarian clitics has drawn particular interest because Bulgarian is part of the Balkan *Sprachbund* and as such it exhibits verb-adjacent cliticization pattern, but also because it displays various patterns in its history.[2] Although it has been claimed that the position of clitics in Old Church Slavonic (OCS)[3] texts mostly echoes their placement in the Greek originals, there are numerous exceptions reflecting patterns in the language of the scribes (especially in constructions without parallels in Greek; see Sławski 1946, among others). In her seminal study on the clitics in the history of Bulgarian, Pancheva (2005) argues that verb-adjacent clitics found in Old Bulgarian (and OCS) temporarily switched to second-position (up to and beyond the Middle Bulgarian period), only to become verb-adjacent again (in modern Bulgarian). Migdalski (2016) claims that in the majority of cases pronominal clitics in OCS were verb-adjacent while the second position (2P) was obligatory only for the so-called 'operator' clitics (discourse markers and 'ethical datives'); he relates the verb-adjacency to the presence of tense morphology in a language.

The discussion in the present article is based on data excerpted from manuscripts that presumably reflect the diachrony of the Bulgarian language and are closer to the vernacular than to the literary language. The main aim of the article is to sketch the evolvement of the (sentential) clitic cluster in the history of Bulgarian and to check

Part of the work presented in the article was performed within the Russian-Bulgarian bilateral project *Evolution of the Grammatical Structure in Russian and Bulgarian Languages in a Contrastive Context: Corpus Analysis and Formal Grammars*. The Bulgarian part has been financially supported by the Fund of Scientific Research at the Ministry of Education and Science with a contract N02/2 from 14.06.2018. A version of the text was presented at the 25th annual conference *Derzhavin Readings* in March 2020, within the section "Contemporary and Historical Problems of Bulgarian and Slavic Studies" (with a short paper included in the proceedings). I would like to thank the two anonymous reviewers of the present article for the constructive comments, as well as Andrej Bojadžiev for the discussion of some points and Iliyana Krapova for the extensive comments and suggestions. Needless to say, the Author is solely responsible for all the remaining errors.

1 Sławski 1946; Radanović-Kocić 1988; Pancheva 2005; Kosek 2011; Zimmerling 2013; Migdalski 2013; 2016.

2 Sławski 1946; Pancheva 2005; Zimmerling 2013; Migdalski 2013; 2016.

3 Here I will use the term Old Church Slavonic (OCS), which reflects the status of the language as used by the Slavic orthodox community (cf. Bujukliev et al. 1993). However, the linguistic features of the texts in the monuments are assumed to reflect those of Old Bulgarian.

whether there has been a significant change in the clustering pattern. The observations on this specific set of data confirm the old hypothesis that the clitics in Bulgarian are (and were) verb-adjacent (and also target 2P). The necessary ramification here is that any conclusions based on heterogeneous diachronic data highly depend on the sources and their analysis. The notion 'second position', for example, depends on syntactic constituency within the language and on the interpretation of the data that are taken into account. With diachronic texts, the segmentation is also a matter of interpretation: for the OCS data here, the texts were already segmented within the corpus used for reference; with the other data (Middle Bulgarian and Early Middle Bulgarian), however, segmentation was done aside. The difficulties with the specifics of the data remain: with the OCS data I have tried to consider variant readings if accessible, with interesting results;[4] this should be done over all the data if possible (for example, one should consider variant readings within the texts of the damaskin collections). In order for such a task to be accomplished, however, one needs a really big and smartly annotated corpus, which is unfortunately missing at this time.

2 Clitics in Bulgarian

To define 'what a clitic is' is not a trivial task because, as Franks (2017, 146) puts it, "the idiosyncrasies of clitics all go in the same direction, i.e., a *negative* one". It is easier to say what clitics are not rather than what they are, because they are – prosodically, morphologically, lexically, and syntactically – 'deficient'. Firstly, clitics are prosodically deficient as they do not have an independent stress, but form a prosodic unit with another, prosodically 'strong', word. However, in many languages clitics may host the stress when preceding or following other elements that remain prosodically 'weak' (or may receive secondary stress). In present-day standard Bulgarian, clitics may bear stress after negation as in (1) (there is also a hypothesis for a secondary stress here; Krăsteva (2020, 119), on the basis of experimental data, claims that in interrogative sentences with negation, even when the clitics receive stress, they do not have "an independent intonational peak").

4 A more detailed review cannot be supplied here but ca. 43% of the occurrences of the datives ми 'I.DAT' and ти 'you.DAT' in Codex Marianus (CM) have variant readings in other New Testament (NT) texts as witnessed in Codex Zographensis (CZ), Codex Assemani (CA) and Codex Sabbae (CSb), and/or in the respective Greek editions.

(1)

Не_**му**	даде	книгата,	нали?
NEG*_**he.DAT**	gave	book.the	right

* The following abbreviations are used in the glosses: NEG, negative particle; EMPH, emphatic particle, e.g. же; NOM, nominative; GEN, genitive; DAT, dative; ACC, accusative; ReflCl, ся; ReflPossCl, си; DEF, word form with a definite article; COP, copula; AUX, auxiliary; COND, conditional; INF, infinitive; IND, indicative; AOR, aorist; IMPF, imperfect; RENARR, renarrative; SUBJ, subjunctive. The elements at focus are given in bold.

Secondly, clitics are a 'bag' of grammatical features (Sadock 1991) and do not have lexico-conceptual features, i.e. they are also semantically deficient (Franks 2017, 154). Historically, the short dative forms of the first- and second-person singular pronouns ми 'I.DAT' and ти 'you.DAT' were already clitics in Old Bulgarian/OCS (as were the parallel pronouns in Greek, see Wackernagel 2009; Večerka 1989; 1993); other short pronominal forms, however, became clitics as a result of early attrition: 3pSg masculine accusative е-го (ѥ-го)[5] 'he.ACC' > го 'he.ACC', with further reconsideration of г- as a third-person marker resulting in a generalised third-person plural with the ending -и (ги 'they.ACC'); еѫ/е > ѫ > я 'she.ACC' (3pSg feminine accusative); емоу > моу 'he.DAT' (3pSg masculine and neuter dative), еи > ей > й 'she.DAT' (3pSg feminine dative), etc.; see Byjukliev et al. 1991, 234-6; Mirčev 1963, 165-6).

Thirdly, clitics are syntactically dependent although different elements may exhibit different behaviour. Clitics can be categorised into 'simple' (non-paradigmatic) and 'special' (paradigmatic) (Zwicky 1977; 1985): in present-day Bulgarian, the latter are present-tense auxiliaries/copulas, short forms of personal and possessive pronouns, which have corresponding 'full' (prosodically and syntactically 'strong') forms.

Clitics are known for their ability to clusterise – in some languages the clitic cluster has a fixed position in the clause or in the phrase (in standard Bulgarian these are the 'special' clitics and the 'simple' interrogative ли). The so-called Wackernagel clitics in Indo-European languages were all second-position (2P) clitics. At issue here would be the definition of the second position: whether the clitic is positioned after the first word in the sentence, as originally formulated by Wackernagel (2009), or after the first syntactic constituent (including a phrasal one) (Halpern 1995), or after the first constituent that immediately follows the so-called rhythmic-syntactic barrier of Zaliznjak (2008). Zaliznjak introduced this barrier as an ad-

5 Orthographic variants throughout the text are generalised if possible (e.g. ero = ero; ми = мі etc.), except for certain examples. The same is true for jotated variants (i.e. ero is given instead of ѥго etc.).

ditional (and obligatory) condition to make the Wackernagel effect work for the Old Russian clitics. Additionally, the restrictions known as the Tobler-Mussafia's law (Tobler 1875; Mussafia 1888), originally formulated for the Romance languages, do not allow 'special' clitics to be first in the sentence.

All these constraints are partially valid for present-day standard Bulgarian: clitics are mostly verb-adjacent, and are found either after the verb in the first position in the clause, as in (2a), or before the verb if there is another constituent in first position, as in (2b). Although (2c) is ungrammatical with clitics following both the subject and the verb and coming third in the linear order of the clause, clitics can still be found further up in the clause following more than one constituent, as in (2d). Clitics may also land third after a preceding verb that, however, is analysed as first after the rhythmic-syntactic barrier (with other sentential element to its left), as in (2e).

(2) a.

Даде	**му**	**я**	Иван	книгата
gave	**he.DAT**	**she.ACC**	Ivan	book.DEF

b.

Иван	**му**	**я**	даде	книгата
Ivan	**he.DAT**	**she.ACC**	gave	book.DEF

c.

*Иван	даде	**му**	**я**	книгата
*Ivan	gave	**he.DAT**	**she.ACC**	book.DEF

d.

В	събота	изненадващо	Иван	**му**	**я**	даде	книгата
In	Saturday	surprisingly	Ivan	**he.DAT**	**she.ACC**	gave	book.DEF

e.

Да,	даде	**му**	**я**	книгата	Иван
Yes,	gave	**he.DAT**	**she.ACC**	book.DEF	Ivan

According to Zimmerling (2013, 89), there are different language systems regarding the position of the clitics, and Bulgarian belongs to the class of languages with so-called 'extended WP+ system', in which the clustering clitics are expected to be verb-adjacent or rather the verb in sentences with 2P clitics is expected to be clitic-adjacent, i.e.

the positional constraint here is both on the clitic(s) and on the verb.[6] In addition, the clustering clitics are in more or less fixed position relative to the clause boundary – generally second after an element in complementiser position (conjunctions, wh-words, other relativisers, etc.) or following the first constituent that comes immediately after the so-called rhythmic-syntactic barrier of Zaliznjak (2008).

The following considerations are applicable for the discussion that follows. In the article, I assume that second-position of the clitics is the position following: (a) a specific set of conjunctions (да 'to', че 'that', ѣко 'as', аште 'if', ами 'but'); (b) wh-words; (c) other relativisers (дето, де, щото 'that, which');[7] (d) a constituent in focus or topic; or (e) verbal element in first position. Sentential constituents located outside the boundary of the clause such as vocatives and some extra-sentential particles do not count as first position. Differently from other Slavic languages, clitics in Bulgarian cannot be found first in the sentence, and in spite of some exceptions in historical texts, Bulgarian clitics in general comply with the Tobler-Mussafia's law if and only if the latter applies to the clause and not to the level of the intonational phrase (i.e. relative to the rhythmic-syntactic barrier of Zaliznjak (2008). In what follows, we will see that these considerations are also valid for the history of Bulgarian in spite of some counter-examples.

3 Clustering: The History

The hypothesis probed in the article is that the order of the elements within the cluster is a follow-up on the diachronic development of elements targeting the second position, from left to right. Thus, the oldest clitics (2P) would be on the far left (or first) of the clustering elements, while the newest ones would land on the (far) right. Beside the special clitics and the discourse clitics, elements that are found around the cluster and are considered (semi-)part of it by some authors (see more in Zaliznjak 2008), are: the interrogative clitic ли, which is found on the far left but also, given that it does not impose any syntactic restrictions on the preceding element, it may show up after each of the other elements in the cluster; the old adverbial particles и 'and', тоу 'then'; and the adverbs пакъ, пакы 'again'.

6 As one of the reviewers rightly put it "a decisive factor could be the position of the verb, and in that scenario the second position requirement is irrelevant", although there are still some exceptions, though obscure or stylistically marked.

7 These form prosodic words (linked by _) with the clitics in some texts – in fact, in these positions, the clitics are part of the first prosodic complex: DLv (17th c.; Mladenova, Velčeva 2013): **д̑ето_мꙋ_се** да́де во́лѧ ѻ҆номꙋ́ѕи влькꙋ. **щото͡е** сегà а҆нтихрі̑сть. **и҆_да_сѐ** поклонѧ́ть вль́кꙋ, **че_мꙋ͡е** ра́ботата ка҄то на_вль́кь.

Not all clustering elements in present-day Bulgarian show similar behaviour. Monosyllabic auxiliaries, other than the present-tense ones, such as the conditional би 'would', the past-tense бе 'was/were', the future ще 'will/shall', may also appear first in the clause. Both in present-day standard Bulgarian and in the historical data observed in the article, these auxiliaries may be found either to the left or to the right of the pronominal elements (the short forms of dative and accusative personal pronouns, the short forms of possessive personal pronouns, which have the same form as the dative personal ones, and the short reflexive pronouns – accusative се and dative си). Within the cluster, the pronominal clitics appear in a fixed order with respect to one another: the dative always precedes the accusative (with reflexives behaving in a parallel way), while the present-tense auxiliaries show a split behaviour: all persons except third person singular show up to the left of the pronominals, while the third-person singular auxiliary (е'is.3pSg') shows up to its right.

In the present article, I propose that the cluster is composed of a primary layer and a secondary one. The primary layer is further split into a core (pronominal) and a periphery (verbal). This is also historically motivated as the auxiliary clitics have joined the clustering elements later (present-tense auxiliary clitics are derived via attrition to the left or to the right of the remaining elements: first-person singular е-смь > съм 'am.1pSg'; second-person singular е-си > си 'are.2pSg'; third-person singular е-стъ > е 'is.3pSg'; third-person plural сѫ-тъ > са 'are.3pPl'). The secondary layer hosts elements that are clitic-adjacent such as past-tense auxiliaries whose position also varies as in (3a) vs. (3b), as well as the conditional (see Nicolova 2008).

(3) a.

Иван	**бе/беше/бил**	**му**	**го**	дал.
Ivan	**be.AOR/IMPF/RENARR**	**he.DAT**	**he.ACC**	given

b.

Иван	**му**	**го**	**бе/беше/бил**	дал.
Ivan	**he.DAT**	**he.ACC**	**be.AOR/IMPF/RENARR**	given

In the pattern in (4), the primary layer is in bold. Some dialects have retained only present-tense AUX2, i.e. the verbal clitic is to the right (Antonova-Vasileva et al. 2016, map 144).

(4) (QU) (AUX0) **AUX1 DAT ACC AUX2** (AUX +)

The ordering in (4) illustrates the overall pattern in present-day standard Bulgarian and if we are to judge by the data discussed be-

low, it has undergone relatively few changes, which have rather affected the inventory of the elements that may appear in the respective slots rather than the pattern.

The earliest 2P clitics in Old Bulgarian/OCS are the discourse markers же and бо (the latter, however, was lost early). The pronominal clitics, which presumably joined the cluster at a later point and which build the core of the primary layer also in the history of Bulgarian, can be classified into several subclasses depending on their placement within the cluster: (1) the dative first-person singular ми 'I.DAT' and second-person singular ти 'you.DAT'; (2) the short forms of anaphoric pronouns (which would be later reanalysed as third-person pronouns): и 'he.ACC' (masculine singular; neuter and feminine dual), ѫ 'she.ACC' (feminine singular), е 'it.ACC' (neuter singular), ѧ 'they.ACC' (plural, masculine and feminine), ѣ/ѭ 'they.DU' (neuter plural; masculine dual); the accusative short forms of the first- and second-person singular personal pronouns мѧ 'I.ACC' and тѧ 'you.ACC' (and, possibly, plurals and duals), defined as semi-clitics (Vaillant 1948, 262; Večerka 1989, 42) as they may occur in 'strong' positions – after a preposition and (rarely) at the beginning of a sentence; and reflexives сѧ (accusative) and си (dative), also found after prepositions; (3) in later texts, the inventory expands with third-person singular datives моу 'he.DAT' (masculine and neuter) and й 'she.DAT' (feminine) derived from the earlier dative short forms емоу and еи that also tend to appear more or less adjacent to the position of the cluster in earlier texts.

Both historically and in present-day Bulgarian, verbal elements ('connectors'; Zimmerling 2013) among the clustering elements are located in the periphery of the primary layer and display variation in placement with respect to other clustering elements in that they can occupy either AUX1 or AUX2 in (5) below. These are: the monosyllabic auxiliaries of бъти (optative бѫ/би 'would'; aorist бы 'was/were. AOR'; imperfect бѣ 'was/were.IMPF', which, however, may be found also in first position); the future auxiliary ще/щѫ 'will/shall' (with variants ке, хте, че, etc.) appearing not earlier than Middle Bulgarian; present-tense auxiliaries and their successors: есмь (съм, смь) 'am.1pSg', еси (си) '2pSg', 3pSg естъ (е) 'is.3Sg', есме (сме) 'are.1pPl', есте (сте) 'are.2pPl', сѫтъ (сѫ, са) 'are.3pPl'.

The interrogative clitic ли (to the far left in both (4) and (5)) and adverbial particles are part of the secondary layer. These are: ти 'and', и 'and', пак(ь) 'again', тоу 'then', etc. The adverbs may show up on both sides of the primary layer, as indicated by the two positions of ADV in (5).

The pattern in (5) gives the relative order of the elements that tend to clusterise as witnessed in the earliest historical texts. With some changes, this order correlates with the one that has been observed by Zaliznjak (2008, 82) in the Novgorod charters. Note that not all elements that appear in bold in (5) are necessarily clitics; as will be seen below, some of them are semi-clitics, while others are clitic-like.

(5) DISC QU DISC ADV **AUX1 DAT (REFL)**[8]**ACC AUX2** ADV

According to the so-called 'historical principle', the order of the clitics in the cluster mirrors the chronology in which they have become 'weak', a pre-condition for their later reanalysis as clitics. This principle predicts that potential members of the clitic cluster would join the clustering elements only to their right. According to Zimmerling (2013), this principle cannot explain the order but can only account for the change in the status of the elements, which is a mechanism for expanding the cluster. Observations on the data that will be discussed, however, show that this is true for a limited set of elements with relatively fixed slots, either in the clause or in the phrase – the discourse markers and the pronominal clitics, which, as mentioned above, constitute the core of the clitic cluster (while verbal elements fluctuate).[9]

The discussion is based on data that has been excerpted from texts from three diachronic corpora. For Old Church Slavonic (OCS), these are the annotated texts of two monuments – Codex Marianus (CM) and Codex Suprasliensis (CSpr) within the TOROT corpus (Haug, Eckhoff 2011a; 2011b), which are used for the numbers; variant readings here are consulted according to three monuments – Codex Zographensis (CZ), Codex Assemani (CA), and Codex Sabbae (CSb) in the TITUS database. For Middle Bulgarian, the texts used are the Legend of Troy (LTr) (14th c.; Miklošič 1871) and the Vlacho-Bulgarian Charters (VB) (15th-16th c.; Bogdan 1902) that are available in the Diachronic Corpus of the Bulgarian Language (Totomanova 2015). For Early Modern Bulgarian (EMB), the texts in two damaskins (17th c.) are used: the Damaskin of Troyan (DTr; Ivanova 1967) and the Damaskin of Lovech (DLv; Mladenova, Velčeva 2013). The texts were chosen because of their accessibility and availability[10] for replica and the hypothetical closeness to the vernacular.[11]

8 The reflexive pronoun СА 'self.ACC' exhibits 'mixed' behaviour of a marker that reflexivises the verb and a semi-clitic but is also found among the clustering elements as we will see later. There are a couple of examples with other accusatives, therefore the reflexive is placed within the pattern here, albeit tentatively.

9 Expectably, some orders that are present in our data are not found in present-day standard Bulgarian: for example, pronominal clitics could occupy the first sentential position that is never the case in present-day standard Bulgarian (though this phenomenon is observed in some dialects, see Iliev 2018, Tiševa 2008, among other). Also expectably, clitics were found in the absolute second prosodic position as well; in some cases, they could split a noun phrase much like what happens in present-day Serbian. This is no longer the case in standard Bulgarian.

10 Within the TOROT corpus: `http://syntacticus.org/`; the TITUS database: `https://titus.fkidg1.uni-frankfurt.de/framee.htm?/index.htm`; and Diachronic Corpus of the Bulgarian Language: `https://histdict.uni-sofia.bg/textcorpus/list`

11 Texts in CSpr, VB, DTr and DLv have different linguistic properties, probably following different redactions (and translations) (in the case of CSpr, DTr, DLv; the collection of VB in the Bogdan's 1902 publication contains texts written over two centu-

3.1 Discourse Clitics

The discourse clitics же (a focus marker or an emphatic particle) and бо (a complementiser or a marker of subordination) are widely attested in OCS (with parallels in Greek; cf. Thayer 1889). These clitics usually appear in the second clausal and/or phrasal position: they are the earliest 2P clitics and both were already used as suffixes (as in иже 'who/which', никътоже 'nobody', неже 'than'; оубо 'therefore', ибо 'because', etc.). Migdalski (2016) calls them "operator clitcs" since they additionally 'operate' on the information structure of the sentence. If found together (rarely in the observed texts), the order is же > бо, as in (6a); pronominal clitics follow them (6b-c).

(6) a.

а҃ко	**же**	**бо**	вьлѣꙁе	къ	ȏученикомъ	CSpr, 251*r*, 19-20
if[i]	**EMPH**	**because**	entered	to	disciples	
Ὅτι	**μὲν**	**γὰρ**	εἰσῆλθε	πρὸς	τοὺς μαθητὰς	

i Due to space limitation, glosses are given if there is no appropriate translation and with elements that are in focus in the article, mainly clitics (e.g., is.AUX.3pSg: present-tense auxiliary, 3pSg; be.AUX.COND: conditional auxiliary; etc.).

b.

оунѣе	**бо**	**ти**	естъ		CM, Mt. 5,29
better	**because**	**you.DAT**	is.COP.3pSg		
συμφέρει	**γάρ**	**σοι**			
да	погъıблетъ	единъ	оудъ	твоихъ	
to	die	one	members.GEN	your.GEN	
ἵνα	ἀπόληται	ἓν	τῶν μελῶν	σου*	

* Three Greek NT editions were consulted. Wherever possible, I supply the corresponding Greek text with variants according to other editions if they attest for variant readings with respect to the clitics. If there is no reference to an edition, the text is given according to Nestle 1904; for a discussion on editions vs. manuscripts, see Toufexis 2010.

c.

въздастъ	**бо**	**ти**	**сѧ**	CM, Lk. 14,14
give	**because**	**you.DAT**	**ReflCL**	
ἀνταποδοθήσεται	**γάρ**	**σοι**		Nestle 1904
ἀνταποδοθήσεται	**δέ**	**σοι**		Tischendorf 8th ed.

ries ago, which reflect different language varieties and influences). The raw numbers given in the article show the overall trends without aiming at a thorough analysis as the data is quite heterogeneous.

во	вьскрѣшение	праведъныхъ
in	resurrection	righteous.GEN
ἐν	τῇ ἀναστάσει	τῶν δικαίων.
ἐν	τῇ ἀναστάσει	τῶν δικαίων.

The particle же is used for emphatic marking of the preceding constituent and has the semantics of a weak adversative conjunction, most often translating Greek δέ 'but, and' – as in (7a) – and more rarely μέν 'while', but also (in various contexts) τέ 'and, also', οὖν 'therefore', δή 'but, then', γάρ 'because', γέ 'in fact, only', καὶ 'and'. In some cases, there is no parallel element in Greek, as in (7b); see also the variation in (7c) where οὖν is translated as 'же оубо', only 'оубо' and only 'же' in the three manuscripts cited.

(7) a.

онъ	**же**	рече	имъ	врагъ	[...]	CM, CZ, Mt. 13,28
he	**EMPH**	told	they.DAT	enemy		
Онъ	**же**	[ре]че	имъ·	врагь		CA
ὁ	**δὲ**	ἔφη	αὐτοῖς,	ὁ Ἐχθρὸς		

раби **же**	рѣшѧ	емоу
slaves **EMPH**	told	he.DAT
Они **же**	рѣшѧ	[...]
Οἱ **δὲ** δοῦλοι	λέγουσιν	αὐτῷ

b.

[...]	приходитъ	**же**	неприѣзнь	[...]		CM, CZ, Mt. 13,19
	comes	**EMPH**	evil-one			
[...]	приходитъ		неприѣзнь	[...]		
[...]	ἔρχεται		ὁ πονηρὸς	[...]		

c.

егда	**же оубо**	придетъ	гн҃ь	винограда	[...]	CM, CA, Mt. 21,40
when	**therefore**	comes	owner	vineyard.GEN		
егда	**же**	придетъ	г҃нъ	винограда	[...]	CZ
ꙗ эгда	**оубо**	приде	г҃ь	виноградоу	[...]	CSb
ὅταν	**οὖν**	ἔλθῃ	ὁ κύριος	τοῦ ἀμπελῶνος	[...]	

Both же and бо may split a constituent – see (8) where же separates зѣло 'much' and зовѧштоу 'who is calling (him)' (επικράζοντος).

(8)

Зѣло **же** зовѧштоу	**ми**	ꙗ́го	много и́ глаголѧштоу	CSpr, 21*r*, 28-29
much **EMPH** calling	**I.DAT**	he.ACC	much and speaking	
Ἐπικράζοντος **δέ**	**μου**	αὐτῷ	καὶ λέγοντος πολλάκις·	

The complementiser бо (Sławski 1974, 285-6) is predominantly found after a single constituent, which can be preceded by a preposition, a conjunction or negation particle не (which can also precede бо). It translates the Greek emphatic conjunction γάρ 'because'. The use of оубо as an emphatic particle is a very early development, mostly in interrogative sentences – оубо is predominantly found after a wh-word (Cejtlin 1994, 721-2) and translates different elements in Greek: ἄρα μὲν 'then indeed', οὖν 'therefore', τοίνυν 'therefore' (each has specific placement: οὖν is predominantly in 2P, and μὲν can be found after the definite article, i.e. phrase-internally). There are adjacent placements of оубо and бо, with variant readings as in (9).

(9)

ѣко	**бо оубо**	събираѭтъ	плѣвелъı.		CM, Mt. 13,40
as	**therefore**	collect	weeds		
ѣ̄коже	**ꙋ̄бо**	плѣвелі	събіраѭ̄тъ	сѧ·	CA
ѣко	**оудобь**	събираѭтъ	плѣвелы		CZ
ὥσπερ	**οὖν**	συλλέγεται	τὰ ζιζάνια		

The constituent preceding же and бо can be a verbal element, a noun or a pronoun, an adjective, an adverbial, a wh-word such as което 'who', чѣто 'which/that', по чѣто 'because of which', etc., which can be preceded by a conjunction, negation particle, and/or followed by ми 'I.DAT', ти 'you.DAT', тѧ 'you.ACC', сѧ 'self.ACC', etc.

In Middle Bulgarian texts observed here, both discourse markers are rare. In LTr, же has only 9 occurrences, and бо is found 7 times. The raw numbers given by Dimitrova and Bojadziev (2014) show that many later non-canonical sources exhibit higher numbers for оубо but not for бо. In VB, бо is missing and же occurs 47 times, mainly in fixed expressions such as великъ же и малъ 'big but also small' (26 occurrences) and еще же (12 occurrences). же is also found after pronominals such as the dative ви 'you.2pPl' in: варе що **ви же** речет жꙋпан Кр(ъ)стѣ и жꙋпан Ханѧш 'something that župan Krastyo and župan Hanush told **you**'.

In the two 17th-century damaskins, же and бо are missing (except for in citations).[12]

3.2 Pronominal Clitics, Semi-Clitic and Clitic-Like Elements

The short forms of personal and anaphoric pronouns can be grouped into three subclasses according to their cliticisation status in OCS. The raw numbers in **table 1** support the assumption that they were all on the path to becoming clitics (in some of the accounts cited above, they are already analysed as such) but there are differences as well.

The first subclass includes the datives ми 'I.DAT' and ти 'you.DAT', which are the first pronominals to exhibit the behaviour of clitics: according to Wackernagel (2009), the Greek datives μοι 'I.DAT' and σοι 'you.DAT' were clitics, hence, their corresponding pronouns in OCS ми and ти might have been clitics, as well (Večerka 1989; Wackernagel 2009, 496, fn. 8 by D. Langslow). In the data observed here, they tend to appear in 2P (ca. 73%) preceded by a verb (ca. 67% of all 2Ps; and 72% of all occurrences) or after a nominal element, an adverb, a conjunction such as да 'to', аште 'if', ѣко 'as', and a wh-word (ca. 25% of all 2Ps; and ca. 21% of all occurrences). Although Migdalski (2016) claims that these were mainly ethical datives functioning as 'operator clitics', a study on these specific cases by Krapova and Dimitrova (2015) shows that many of these datives can be analysed either as affected participants, or as inalienable possessors, or as both (i.e. affected participants that are also inalienable possessors, see Minčeva 1964).

The second subclass of pronominal clitics comprises the so-called semi-clitics and can be further divided into two groups: a) the first- and second-person accusatives мѧ 'I.ACC' and тѧ 'you.ACC', which appear after a preposition and in first position; and b) the third-person accusatives (и 'he.ACC', ѫ 'she.ACC', ѧ 'they.ACC', е 'it.ACC', ѣ 'they.ACC'), which often translate non-clitic Greek pronouns. Elements of both groups show a strong tendency to appear in second position, as well as post-verbally. There is a slight precedence of the second group to appear post-verbally: ca. 78% (мѧ, тѧ) vs. ca. 93% (и, ѫ, ѧ, е, ѣ). When in second position, however, there is a clear difference between the two groups with respect to their pre-verbal placement: ca. 6% (мѧ, тѧ) vs. ca. 25% (и, ѫ, ѧ, е, ѣ).

The third subclass of pronominals – involving pronouns that will be defined as 'clitic-like' in the present discussion – includes the short dative bi-syllabic емоу 'he.DAT', еи 'she.DAT', имъ 'they.DAT', etc.,

12 In the so-called archaic damaskins, however, both are still present, in the earlier positions.

which have ambiguous behaviour with respect to discourse markers and the first two pronominal subclasses, and whose successors in later stages of Bulgarian (моу 'he.DAT', й 'she.DAT', etc.) are clitics. They are also found predominantly in second position (ca. 75%) and, (even) more often, post-verbally (ca. 92%). However, only 7% of the clitic-like pronouns are both 2P and pre-verbal – this percentage is comparable to the one seen in the preceding paragraph with respect to the first- and second-person accusatives мѧ 'I.ACC' and тѧ 'you.ACC' (the first subclass of the semi-clitics).

These results may indicate, albeit it is a tentative proposal, that 2P cliticisation was the dominant pattern, at least during the period witnessed by these OCS texts, and that this pattern was not in competition with pre-verbal placement but could co-exist with the latter in case the word order would allow for it. The 2P pattern and the post-verbal one were not in competition either and perhaps can be seen as alternative mechanisms motivated also by word order and/or information-structure principles and/or the syntax of the original Greek text.

Table 1 Pronominal clitics, semi-clitics and clitic-like pronouns in OCS (according to CM and CSpr)

CL[i]	2P, pre-V[ii]	2P, post-V	2P, non-V-adjacent[iii]	non-2P, pre-V	non-2P, post-V	non-2P, non-V-adjacent
ми, ти	75[iv]	204	26	12	97	4
емоу, еи, имъ	77	1053	17	23	362	7
мѧ, тѧ, нъı, въı	83	241	15	22	194	4
и, ѫ, ѧ, е, ѣ	49	788	11	12	278	3

i Clustering clitics are calculated as part of the cluster (as well as interrogative ли, monosyllabic adverbs и, ти, тоу), i.e. another clitic, which is part of the cluster, can precede or follow (e.g. приведѣте ми.DAT и.ACC сѣмо 'bring him to me here' will be calculated both for the dative ми and for the accusative и as 2P, post-V; рьци оубо намъ чъто ти.DAT сѧ.REFL мьнитъ 'threrefore, tell us what you think' will be calculated for the dative ти as 2P, pre-V).

ii The constituent in 1P can be preceded by a coordinating monosyllabic conjunction such as и 'and', а 'but', нъ 'but', etc., and the negation не. Elements after infinitives were calculated accordingly (повелѣ а҃нѳупатъ принести врѣтиште. и вьѣсадити и.ACC вь҃ не҃. 'order the proconsul to bring the back and to put it inside (in-it)' was 2P, post-V for the accusative и) but periphrastic verb constructions are not included.

iii One or more constituents (noun phrases, prepositional phrases, adverbial phrases, including non-monosyllabic adverbs such as пакъ, пакъı) can be placed between the clitic/clitic cluster and the verb.

iv Numbers do not involve periphrastic constructions, first-position elements or prepositional phrases.

If clusterised, the dative ми 'I.DAT' and ти 'you.DAT' tend to precede the accusative semi-clitics (and the reflexives), as in (10a-c). This is not true for the clitic-like datives, though, as in (10d).

(10) a.

приведѣте	**ми**	**и**	сѣмо.	CM, Mt. 17,17
bring	**I.DAT**	**he.ACC**	here	
φέρετέ	**μοι**	**αὐτὸν**	ὧδε.	

b.

принесѣте	**ми**	**ѭ**	сѣмо.	CM, Mt. 14,18
bring	**I.DAT**	**they.ACC**	here	
принесѣте		**а҃**	сѣмо	CA
принесъете			съемо	CSb
Φέρετέ	**μοι**	ὧδε	**αὐτούς**.	
φέρετέ	**μοι**	**αὐτοὺς**	ὧδε	Scrivener 1894

c.

рьци	оубо	намъ	чъто	**ти**	**сѧ**	мьнитъ	CM, Mt. 22,17
tell	then	we.DAT	what	**you.DAT**		think	
εἰπὲ	οὖν	ἡμῖν,	τί	**σοι**		δοκεῖ	

d.

наꙁнамена	**же**	**сѧ**	**ѥ́моу**	вѣньцъ	по	тѣлоу·	CSpr, 72*v*, 4
indicated	**EMPH**	**ReflCL**	**he.DAT**	wreath	on	body	
Ἐσημάνθη	**δὲ**		**αὐτοῦ**	ὁ στέφανος	καὶ διὰ	τοῦ σώματος.	

The placement of the pronominal elements often echoes the respective placement of the parallel elements in the Greek text but with variant readings (especially if there were no corresponding constructions in Greek – Sławski 1946 – such as the reflexive constructions and periphrastic verb constructions). The verbal elements are discussed further in the article; here I will only mention that clitics and semi-clitics demonstrate similar behaviour: both may precede the auxiliary and, by transitivity, also the participle (the active l-participle as well as the passive n/m/t-participles). Both may be found in 2P, in the order X > CL > AUX > Participle, where X can be a wh-word, as in (11a), a subjunction such as да 'to', as in (11b), or some other functional element in first position.

(11) a.

како	**и**	**бѫ҃** погоубили.	CM, Mk. 3,6
how	**he.ACC**	**are.AUX.3Sg** murdered	
ὅπως	**αὐτὸν**	ἀπολέσωσιν.	

b.

[…] съвѣшташѧ	да	**и бѫ**	оубили	CM, Mar. 3,6
[…] consulted	to	**he.ACC are.AUX.3Sg**	killed	
[…] ἐβουλεύσαντο	ἵνα	ἀποκτείνωσιν **αὐτόν**.		

Semi-clitic reflextive pronoun сѧ 'self.ACC' exhibits variation in its behaviour just like the other semi-clitic pronominals: it may appear after the (reflexive) verb (more often) but also among the clustering elements (more rarely) in a slot of another pronominal and, more specifically, in the slot of the accusative pronominals.

We have already distinguished between 'true' clitics (ми 'I.DAT' and ти 'you.DAT') and clitic-like pronouns (емоу 'he.DAT', etc.) above, but their different status with respect to the cliticisation behaviour also manifests with respect to сѧ 'self.ACC': ми and ти always precedes it, as in (10c), while the clitic-like pronouns typically follow it, as in (10d) above.

If the semi-clitic сѧ 'self.ACC', however, is found beside the elements in the core in (5), as indicated by (10c), it may also appear in 2P beside other elements as in (12) where it aspires for a slot within the cluster. In (21) below, we will see another example where сѧ is placed in front of the auxiliary but also after the main verb in compound tenses in a variant reading.

(12)

колико	**сѧ**	**бы**	троудилъ	CSpr, 49*r*, 11-12
how-much	**REFLCL**	**be.AUX.COND**	worked	
Πόσα		**ἂν**	ἔκαμες,	

As mentioned above, semi-clitics (reflexives included) are mainly post-verbal. Pronominal non-reflexive semi-clitics, though, may also appear either post-verbally, in most cases, or in 2P. If 2P, however, is occupied by a 'true' discourse clitic же or бо, the semi-clitic is often post-verbal, as in the examples in (13). This may signal that these semi-clitic pronouns do not behave as clitics yet.

(13) a.

потомь	**же**	обрѣте	**и**	и҃с.	въ	ц҃ркве.	CM, Jo. 5,14
later	**EMPH**	found	**he.ACC**	Jesus	in	temple	
потомь	**же**	обрѣте		и҃с	въ	цркъве·	CZ
Μετὰ	**ταῦτα**	εὑρίσκει	**αὐτὸν**	ὁ Ἰησοῦς	ἐν	τῷ ἱερῷ	

b.

седмь **бо** ихъ	имѣшѧ	**ѥѫ**	женѫ.	CM, Mk. 12,23
seven **because** they.GEN	had	**she.ACC**	wife.ACC	
седмь **бо**	ˊімѣшѧ	**ѥѫ**	женѫ·	CZ
οἱ **γὰρ** ἑπτὰ	ἔσχον	**αὐτὴν**	γυναῖκα.	

c.

они **же** оукоришѧ	**и**	[...]	CM, Jo. 9,28
they **EMPH** rebuked	**he.ACC**		
καὶ ἐλοιδόρησαν	**αὐτὸν**	[...]	
Ἐλοιδόρησαν **οὖν**	**αὐτόν**	[...]	Scrivener 1894

The inventory of clitics (and clitic-like elements) in OCS texts is rather limited. It is to be expanded in Middle Bulgarian but some forms are already attested in the early texts: for example, contracted forms of clitic-like datives, though very rarely and with variant readings, as in (14).

(14)

ˊI	въ	цръкъве	сѫштю	**моу·**	придошѧ	къ	Ѩемоу	CZ, Mk. 11,27
and	in	temple	being	**he.DAT**	came	to	he.AT	
ι	въ	цр̄кве	ходѧштю	**емоу**	·придѫ	къ	немоу	CM
καὶ	ἐν	τῷ ἱερῷ	περιπατοῦντος	**αὐτοῦ**	ἔρχονται	πρὸς	αὐτὸν	

The class of pronominal clitics expands in the Middle Bulgarian texts that I have considered. In LTr, the respective clitics occur in the same positions and order with the addition of third-person accusative го 'he.ACC' and dative моу 'he.DAT', which behave as 'true' clitics. The earlier forms, however (его and емоу), are still found in LTr but in different functions. The accusative clitic го 'he.ACC' is used for direct objects only, while the genitive (genitive-accusative) его predominantly expresses the possessive genitive. The dative clitic моу 'he.DAT' refers to external possessors and is found in 2P, while емоу refers to indirect object (and is found after a verb and after some prepositions such as противѫ 'against'). Expectably, его and емоу do not cluster.

The clustering elements exhibit the pattern in (5), as illustrated by the examples in (15), with the core (DAT > ACC) in 2P, and the verbal elements placed to its right (in AUX2).

(15) a.

[...]	оумыслити,	како	**го**	**и**	расыпати	LTr
[...]	think.INF	how	**he.ACC**	**also.EMPH**	scatter.INF	

b.

[...]	които	**моу**	**сѧ**	ѡ(те)цъ	бѣ	нарекль		LTr
[...]	who	**he.DAT**	**ReflCL**	father	was	called		

c.

[...]	какото	**ѫ**	**си**	**и**	пръвое	чюлъ.	LTr
[...]	how	**she.ACC**	**be.AUX.2Sg**	**also.EMPH**	first	heard	

d.

[...],	да	**мѵ**	**е**	имѧ	Прижїа	градъ.	LTr
[...],	to	**he.ACC**	**is.COP.3Sg**	name	Paris	town	

e.

[...]	помного	**моу**	**би**	лѣпотыі	прибыіло	LTr
[...]	much	**he.DAT**	**be.COND**	beauty	been	

f.

[...]	нѫ	**ми**	**еси**	тыі	оброкъ.	LTr
[...]	but	**I.DAT**	**is.COP.2Sg**	you.NOM.2Sg	oath	

In VB, pronominal clitics continue to cluster in the order DAT > ACC (including reflexives). As in (16a), the cluster is placed before the negation particle and can be separated by the verb (although both pronominal clitics and the cluster are already mainly verb-adjacent).

(16) a.

[...]	да	**мѵ**	**га**	не	липсат	ни	един	влас	VB
[...]	to	**he.DAT**	**it.ACC**	**NEG**	lack	no	one	hair	

b.

A	ако	**мѵ**	**се**	криво	видит	томѵ	[...]	VB
and	if	**he.DAT**	**ReflCL**	awry	see	this.DAT		

The numbers in **table 2** show clitics' overall preference to second position and pre-verbal placement in VB.

Table 2 Pronominal clitics in VB

	2P, pre-V	**2P, post-V**	**2P, non-V-adjacent**	**non-2P, pre-V**	**non-2P, post-V**	**non-2P, non-V-adjacent**
ми, ти, моу, й, им	560	43	21	14	15	2
га/го, ю, их, ме, те, ни, ви	237	18	12	8	3	2

In VB, a certain number of pronominal clitics are found after a mono-syllabic conjunction such as a 'but', и 'and', and after a pause, as well as in second prosodic position where it can split a noun phrase.

(17)

И	**ми**	доде	слɣга	моа,	да	**ми**	**е**	VB
and	I.DAT	came	servant	my	to	**I.DAT**	**is.AUX.3Sg**	

кɣпилъ	що	**мɣ**	**смь**	казал	ωт	все
bought	what	**he.DAT**	**am.AUX.1Sg**	said	from	all

In the two 17th-century damaskins that are took into consideration, pronominal clitics are also found in 2P but pre-verbal and post-verbal positions are almost equally distributed. Although rarely, the cluster may split a noun phrase, as in (18d). The example in (18e) illustrates the portion of the pattern in (5): DAT ACC AUX2.

(18) a.

ʐа_ѐди́нь	ча҇_**мɣ_се**	Ӣʐ҄корені̀	си́лата	Ӣ_Ю̄на́шьството	DLv
for_one	hour_**he.DAT_ReflCL**	eradicated	power.DEF	and_courage.DEF	

b.

и	Оста́ви	**си**	**го**	на	Очи́те	и	на	срц҄е́то,	DTr
and	leave	**ReflPossCL**	**it.ACC**	on	eyes.DEF	and	on	heart.DEF	

c.

а́	любовь	не	и́маᴹ	ни́що	**ми**	**се**	**не**	прида́ва.	DTr
and	love	NEG	have	nothing	**I.DAT**	**ReflCL**	**NEG**	lend	

d.

И	тогѝва	а́ггл҄	**мꙋ**	**се**	гнь	ꙗвѝ	DTr
and	then	angel	**he.DAT**	**ReflCL**	of-God	showed up	
ѿго́рѣ	на҆	нꙓ́го					
from-above	over	he.ACC					

e.

[...]	амѝ	сѝ	ꙋзмѝ	своѧ́та	кръ́па,	дето́	DTr
[...]	but	ReflPossCL	take	own	towel	that	
мї	**ѧ**	**сѝ**	да́ла	на	дѣ́ль		
I.DAT	**she.ACC**	**is.AUX.2Sg**	given	on	deed		

Table 3 Pronominal clitics in DTr and DLv

	2P, pre-V	2P, post-V	2P, non-V-adjacent	non-2P, pre-V	non-2P, post-V	non-2P, non-V-adjacent
ми, ти, моу, й, им...	538	418	52	26	49	3
ме, те, ны, вы...	147	58	19	4	5	1
го, я, гы, ѧ́...	586	425	44	17	33	2

To sum up, the inventory of the clitical pronouns that are found in second position and pre-verbally increases in the history of Bulgarian, mainly as a result of the changes in the inventory of pronominal elements. The accusatives го 'he.ACC', я 'she.ACC', гы 'they.ACC' and datives моу 'he.DAT', й 'she.DAT' have joined the cluster after the reanalysis of the earlier forms of anaphoric pronouns.

3.3 Verbal Clitics and Clitic-Like Elements

Both in the early and in later texts, the behaviour of most monosyllabic verbal elements such as be-auxiliaries/copulas, future auxiliaries щѫ/ще 'will/shall', is ambiguous. They may be found in first position, including immediately after negation, and preceding the discourse clitics – as in (19), but also in 2P, within a cluster and in second prosodic position, splitting a noun phrase, as in (19b) – in parallel to the orders in the Greek text.

(19) a.

бѣ	**же**	тоу	ма͠рѣ	магдалъынӣ.	[...]	CM, Mt. 27,61
was.COP	**EMPH**	there	Maria	Magdalene		
Ἦν	**δὲ**	ἐκεῖ	Μαρία	ἡ Μαγδαληνὴ	[...]	

b.

пѧть	**же**	**бѣ**	отъ	нихъ	боуи.	[...]	CM, Mt. 25,2
five	**EMPH**	**was.COP**	of	they.GEN	wise		
πέντε	**δὲ**	**ἦσαν**	ἐξ	αὐτῶν	φρόνιμοι		

Auxiliaries are found either to the left, or to the right of the pronominals among the clustering elements, as exemplified by the periphrastic verb constructions in (20).

(20) a.

[...]	ꙗ͠ко	да	**бы**	**ѭ͠**	и͠збавилъ	о͠тъ	н͠его·	CSpr, 280*r*, 26
[...]	if	to	**be.AUX.COND**	**she.ACC**	saved	from	he.GEN	

b.

[...]	**ι**	въ	водѫ	да	**и би** погоубилъ.	CM, Mk. 9,22
[...]	and	in	water	to	**he.ACC be.AUX.COND** murdered	
[...]	καὶ	εἰς	ὕδατα,	ἵνα	ἀπολέσῃ.SUBJ αὐτόν·.	

c.

да	**би**	**сѧ**	**ихъ**	коснѫлъ.	CM, Lk. 18,15
to	**be.AUX.COND**	**ReflCL**	**they.ACC**	touched	
ἵνα			**αὐτῶν**	ἅπτηται.SUBJ	

The positions of the reflexive and the auxiliary often vary, with the former more often found in 2P, while the latter is adjacent to negation, as in (21).

(21)

добрѣа͠	**би** бъ**ы**ло **емоу**	CM, Mt. 26,24
better	**be.AUX.COND** been **he.DAT**	
добрѣē	**ēмоу би** бъило·	CA

добрѣе	**емоу би** было·	CZ
добрꙗ̀ъее	было **бы емоу**·	CSb
καλὸν	**ἦν.IND αὐτῷ**	

аште **сѧ би** не родилъ	чл҃къ	тъ.
If **ReflCL be.AUX.COND** NEG born	man	this
āще **бі** не роділъ **сѧ**	чл҃кь	тъ·
ӑште **сѧ би** не родилъ	чл҃къ	тъ·
аще **сѧ бы** не родилъ	чл҃къ	тъ·
εἰ οὐκ ἐγεννήθη.IND	ὁ ἄνθρωπος	ἐκεῖνος.

In Middle Bulgarian, fluctuations in the positions of auxiliaries remain, as exemplified in (22) from LTr.

(22) a.

и	поиди	противѫ	Ектороу	кралю,	да	LTr
and	came	against	Hector	king	to	

би	**сѧ**	оставилъ	гръцкыѭ	воискы
be.AUX.COND	**ReflCL**	left	Greek	army

b.

[...]	и	метнѫ	прѣ҆д	фарижа	егова,	LTr
[...]	and	threw	in-front-of	Paris	his	

давно	**сѧ**	**би**	вьзвратилъ
long-ago	**ReflCL**	**be.AUX.COND**	come back

In VB, monosyllabic auxiliaries are mostly in AUX2, i.e. to the far right, where the latest clustering elements were joining, presumably, as in (23a-b), with 6 examples of the reverse order of which 5 are with смо ‘are.1pPl’ and one – with the conditional auxiliary би ‘would’ and negation particle – see (23c) (and 13 instances are with future auxiliary).

(23) a.

[...]	ере	не	**ми**	**е**	ꙋзел	вражмаш	господства	ми	VB
[...]	that	NEG	**I.DAT**	**is.AUX.3Sg**	taken	enemy	master	my	

b.

а	прьгари	заприхъ	ти	добитокь,	VB
and	burghers	shut	you.DAT	cattle	
та	**ме**	**сѹ**	молил(и)	мене.	
thus	**I.ACC**	**Are.AUX.3Pl**	asked	I.ACC	

c.

и	за	моѹ	слѹжбѹ	не	**би**	**мѹ**	мило	VB
and	for	my	service	NEG	**be.COP.COND**	**he.DAT**	dear	

vs.

ако	**мѹ**	не	**би**	платил	ѡт	волѧ.	VB
if	**he.DAT**	NEG	**be.AUX.COND**	paid	from	will	

The trend continues into the Early Modern Bulgarian with auxiliaries placed on the far right (including the present-tense ones), as in (24a-c), in contrast to the situation in present-day standard Bulgarian (compare also their positions in the dialects according to Antonova-Vasileva et al. 2016, map 144 where the majority of dialects instantiates the order AUX > DAT / ACC). There are again single instances of verbal clitics found before the pronominal ones, as in (24c). The order with respect to the reflexive, as in (24c), is the same as in present-day standard Bulgarian.

(24) a.

Защо	**го**	**си**	мъчи́ль,	и	защо́	**го**	**си**	испъ́диль	DTr
why	**he.ACC**	**is.AUX.2Sg**	tortured	and	why	**he.ACC**	**is.AUX.2Sg**	driven away	

b.

ак^{i}_**ĉь**	ѿстоюва́вале	пра́во	та_**e_ĉa**	ꙗ̀ле	DLv
if_**are.AUX.3pPl**	defended	rightly	then_**it.ACC_are.AUX.3pPl**	eaten	

c.

Затова́_**ca_e**	веќꙗ	и̑_покаа́нїе	[...]	DLv
therefore_**ReflCL_is.AUX.3pSg**	already	also_repentance	[...]	
ꙁатова́_**ca_ce**	веќꙗ	и̑_ра́искьіе	[...]	
therefore_**are.AUX.2pPl_ReflCL**	already	also_heavenly	[...]	

d.

а҅мѝ_**си_с҃а**	ѻ҆мрьси́ле	моми́н҅ството	прѣдѝ	сва́бⷣата	DLv
but_**ReflPossCL_are.AUX.3pPl**	sinned	virginity.DEF	before	wedding.DEF	

e.

Това́_са	д҃ето_**си**	нѣ_**с҃ь**	паʒ́иле	[...]	DLv
This_are.COP.3pPl	that_**ReflPossCL**	NEG_**are.AUX.3pPl**	kept		

Negation may separate clustering elements, with the reflexive to the left and the auxiliary to the right to form two prosodic words (whose constituents are linked via '_'), as in (25a) (see also (24e) with the possessive reflexive). Pronominal clitics are in second position while auxiliaries are verb-adjacent, as in (25b-c).

(25) a.

и҃_ѻще_**се**	нѣ_**ѐ**	добрѣ̀	[тѵ҆ка намѣстил]ь	[...]	DLv
and_yet_**ReflCL**	NEG_**is.AUX.3pSg**	well	here arranged	[...]	

b.

За́мь	догдѐ_**ви**	напрⷣѣ́	нѣ_**ѐ**	прѣва́рила	са́бїа	съмрьт̀нӓ	DLv
for	until_**you.DAT**	before	NEG_**is.AUX.3pSg**	overtook	sword	deadly	

c.

и҃_догдѐ_**са**	паʒа́рь	нѣ_**ѐ**	раʒва́лиль.	DLv
and_until_**ReflCL**	market	NEG_**is.AUX.3pSg**	broken	

Table 4 gives the order of clustering be- and future-auxiliaries (the latter in Middle Bulgarian and Early Modern Bulgarian) and pronominal clitics, semi-clitics, and clitic-like elements with the periphrastic verb constructions. The numbers attest for variation in the placement of auxiliaries among the clustering elements with preference to the right slot AUX2 of the pattern in (5). This is also true for the future auxiliary (the form щѫ was attested very early: Mirčev 1956, 202; Haralampiev 2001, 149). In the texts here, the position of щѫ is also ambivalent (in AUX1 and AUX2), but it is negation-adjacent and also found in second prosodic position.

Table 4 Pronominals and auxiliaries with periphrastic verb constructions

	Pron > Aux	Aux > Pron	Non-adjacent
OCS	53	18	18
MB	143	10	2
EMB	103	25	4

The clitic cluster in the noun phrase is left outside the scope of the present article but its phrasal position (2P) and order is very similar to that of the clausal cluster, with the earliest clitics to the left and the latest to the right (for comparison with Greek, see Eckhoff 2018). This cluster may not be considered viable in present-day Bulgarian if we assume that the definite article is an affix. The presumed pattern of the noun-phrase clitic cluster might have been DISC > DEM/ART > POSS/GEN/DAT > (AUX). There are very early examples of demonstrative pronouns that are in second position in the noun phrase and immediately follow a nominal element in first phrasal position: these are exactly the demonstratives that may be interpreted as clitic-like articles (Dimitrova-Vulchanova, Vulchanov 2012). However, there are also examples of discourse markers followed by demonstrative pronouns, clusterised in the same phrasal second position, which attests that the reanalysis has not been completed yet (as per Kurz 1963).

4 A Tentative Conclusion

The data that have been discussed in the article backs the assumption that, historically, there might have been no significant changes in the position of the elements in the core of the primary clitic cluster, i.e. the pronominal clitics, with variations happening (and still) among the elements in the verbal periphery. The earliest 2P clitics were lost, somewhat expectably, because they were non-paradigmatic unlike those that have remained. The set of elements filling the respective slots among the clustering elements was expanded as a result of the restructuring within the pronominal and verbal systems. Similar developments might have been at play for the (partial) loss of the noun-phrase cluster. A more thorough analysis of the variant readings across the monuments and constructions (matrix vs. subordinate clause, clause types, absolute constructions) is due in order to highlight the numerous intricacies of cliticisation and clitic patterns in the history of Bulgarian.

Bibliography

Antonova-Vasileva, L.; Vasileva, L.; Vitanova, M.; Vălčeva, E.; Keremidčieva, S.; Kostova, T.; Kočev, I.; Kjaeva, E.; Smedovska, K.; Tetovska-Troeva, M.; Urumova, P. (2016). *Bălgarski dialekten atlas. Obobštavašt tom IV. Morfologija* Български диалектен атлас. Обобщаващ том IV. Морфология (Bulgarian Dialect Atlas. Generalising. Vol. 4, Morphology). Sofia: Izdatelstvo na BAN.

Bogdan, I. (1902). *Documente şi regeste privitoare la relaţiile Ţării Rumineşti cu Braşovul şi Ungaria in secolul XV şi XVI* (Documents and Private Registers of the Romanian Counties of Brasov and Hungary in the 15th and 16th Centuries). Bucureşti.

Bujukliev, I.; Georgiev, I.V.; Davidov, A.; Dobrev, A.: Dogramadžieva, E.; Duridanov, I.; Zlatanova, R.; Ivančev, S.; Ilčev, P.; Minčeva, A.; Mirčeva, D. (1993). *Gramatika na starobălgarskija ezik* Граматика на старобългарския език (Grammar of the Old Bulgarian Language). Sofia: Izdatelstvo na BAN.

Cejtlin, R.; Večerka, R.; Blagova, E. (1994). *Staroslavjanskij slovar' (po rukopisjam X-XI vekov)* Старославянский словарь (по рукописям X- XI веков) [Old Slavic Dictionary (According to Manuscripts of the X-XI Centuries)]. Moskva: Russkij jazyk.

Dimitrova-Vulchanova, M.; Vulchanov, V. (2012). "An Article Evolving: The Case of Old Bulgarian". Jonas, D; Whitman, H, Garrett, A. (eds), *Grammatical Change: Origins, Nature, Outcomes*. Oxford: Oxford University Press, 160-78. https://doi.org/10.1093/acprof:oso/9780199582624.003.0008.

Dimitrova, T.; Bojadziev, A. (2014). "Historical Corpora of Bulgarian Language and Second Position Markers". *Proceedings of the First International Conference Computational Linguistics in Bulgaria*. Sofia: Institute for Bulgarian Language, 55-63.

Eckhoff, H.M. (2018). "Quantifying Syntactic Influence: Word Order, Possession and Definiteness in Old Church Slavonic and Greek". *Diachronic Slavonic Syntax*. De Gruyter Mouton, 29-62. https://doi.org/10.1515/9783110531435-003.

Franks, S. (1997). "South Slavic Clitic Placement is Still Syntactic". *Proceedings of the 21st Annual Penn Linguistics Colloquium. Penn Working Papers in Linguistics*, 4, 111-26.

Franks, S. (1998). *Clitics in Slavic*. Paper presented at Comparative Slavic Morphosyntax Workshop, Bloomington (IN).

Franks, S. (2017). *Syntax and Spell-Out in Slavic*. Bloomington (IN): Slavica.

Franks, S.; King, T.H. (2000). *A Handbook of Slavic Clitics*. New York: Oxford University Press.

Halpern, A.L. (1995). *On the Placement and Morphology of Clitics*. Dissertations in Linguistics. Stanford: CSLI Publications.

Haralampiev, I. (2001). *Istoričeska gramatika na bălgarskija ezik* Историческа граматика на българския език (Historical Grammar of the Bulgarian Language). Veliko Tărnovo.

Haug, D.; Eckhoff, H.M. (2011). "The PROIEL Corpus as a Source to Old Church Slavic: a Practical Introduction". Baranov, V. (ed.), *Pis'mennoe nasledie i sovremennye informacionnye technologii* Письменное наследие и современные информационные технологии (Written Heritage and Modern Information Technologies). Iževsk: Iževskij gosudarstvennyj technicheskij universitet, 37-55.

Haug, D.; Eckhoff, H.M. (2011). "Making the Most of the Data: Old Church Slavic and the PROIEL Corpus of Old Indo-European Bible Translations". *Łodź Studies in Language*, xxiv, 411-31.

Iliev, I. (2018). "Săpostavka meždu malomirovskija govor v Elhovsko i kajraklijskija i šikirlijskija govor v Besarabija i Priazovieto" Съпоставка между Маломировския говор в Елховско и кайракслийския в шикирлийския говор в Бесарабия и Приазовието (Comparison Between the Malomir Dialect in the Elhovo Region and the Kayrakli and Shikirli Dialects in Bessarabia and the Azov Region). *Dzjalo*, 6(12). https://www.abcdar.com/magazine/XII/Iliev_1314-9067_XII.pdf

Ivanova, A. (1967). *Trojanski damaskin* Троянски дамаскин (The Damaskin of Troyan). Sofia: Izdatelstvo na BAN.

Kosek, P. (2011). *Enklitika v češtině barokni doby* (Enclitics in the Czech Language of the Baroque Period). Brno: Host.

Krapova, I.; Dimitrova, T. (2015). "Genitive-Dative Syncretism in the History of the Bulgarian Language. Towards an Analysis". *Studi Slavistici*, 12, 181-208. https://doi.org/10.13128/Studi_Slavis-17976.

Krăsteva, H. (2020). "Sravnenie na prozodičnite osobenosti na klitičnija kompleks v săobštitelni i văprositelni izrečenija v bălgarskija ezik" Сравнение на прозодичните особености на клитичния комплекс съобщителни и въпросителни изречения в българския език (A Comparison of the Prosodic Features of the Clitic Cluster in Bulgarian Declarative and Interrogative Sentences). *Bălgarski ezik* Български език (Bulgarian Language), 77(3), 112-21. https://doi.org/10.47810/bl.67.03.09.

Kurz, J. (1963). "Problema člena v staroslavjanskom jazyke" Проблема члена в старославянском языке (The Problem of the Article in Old Church Slavonic). *Issledovanija po sintaksisu staroslavjanskogo jazyka* Исследования по синтаксису старославянского языка (Studies in the Old Church Slavonic syntax). Praha: Československa akademie věd, 121-82.

Migdalski, K. (2013). "Diachronic Source of Two Cliticization Patterns in Slavic". Salvesen, C.M.; Helland, H.-P. (eds), *Challenging Clitics*. Amsterdam; Philadelphia: John Benjamins, 135-58. https://doi.org/10.1075/la.206.06mig.

Migdalski, K. (2016). *Second Position Effects in the Syntax of Germanic and Slavic Languages*. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.

Miklošič, F. (1871). "Trojanska priča" (The Legend of Troya). *Starine* (Antiquity), 3, 147-88.

Minčeva, A. (1964). *Razvoj na datelnija pritežatelen v bălgarskija ezik* Развой на дателния притежателен падеж в българския език (The Development of Dativus Possessivus in the Bulgarian Language). Sofia: Izdatelstvo na BAN.

Mirčev, K. (1963). *Istoričeska gramatika na bălgarskija ezik* Историческа граматика на българския език (Historical Grammar of the Bulgarian Language). Sofia: Nauka i izkustvo.

Mladenova, O.; Velčeva, B. (2013). *Loveški damaskin. Novobălgarski pametnik ot XVII vek* Ловешки дамаскин. Новобългарски паметник от XVII век (The Damaskin of Lovech. Modern Bulgarian monument of the 17th century). Sofia: NBKM.

Mussafia, A. (1888). "Enclisi o proclisi del pronome personale atono quale oggeto". *Romania*, 27, 145-6.

Nestle, E. (1904). *Novum Testamentum Graece*. Stuttgart: Privilegierte Württembergische Bibelanstalt.

Nicolova R. (2008). *Bălgarska gramatika. Morfologija* Българска граматика. Морфология (Bulgarian Grammar. Morphology). Sofia: Universitesko izdatelstvo "Sv. Kliment Ohridski".

Pancheva, R. (2005). "The Rise and Fall of Second-Position Clitics". *Natural Language and Linguistic Theory*, 23, 103-67. https://doi.org/10.1007/s11049-003-2510-y.

Petersen, U. (2016). *Tischendorf's 8th Edition Greek New Testament with Morphological Tags*. Based on G. Clint Yale's Tischendorf text and on Dr. Maurice A. Robinson's Public Domain Westcott-Hort text. DiMedia, Department of Communication, Aalborg University.

Radanović-Kocić, V. (1988). *The Grammar of Serbo-Croatian Clitics: A Synchronic and Diachronic Perspective* [PhD dissertation]. University of Illinois, Urbana-Champaign.

Scrivener's Textus Receptus (1894). With Morphology. Prepared by M.A. Robinson. Bellingham 2002.

Sławski, F. (1946). *Miejsce enklityki odmiennej w dziejach języka bułgarskiego* (The Placement of Enclitics in the History of the Bulgarian Language). Krakow: Nakładem Polskiej Akademii Umiejętności.

Sławski, F. (1974). *Słownik prasłowiański* (Dictionary of the Proto-Slavic Language). Tom I, *A-B*. Wrocław-Warszawa-Krakow-Gdańsk: Wydawnictwo Polskiej Akademii nauk.

Thayer, J.H. (1889). *Greek-English Lexicon of the New Testament*. New York: American (Harper).

Tiševa, J. (2008). "Za bălgarskite ezikovi ostrovi" За българските езикови острови (On the Bulgarian Language Islands). *Plovdivski universitet "Paisij Hilendarski" – Naučni trudove. Filologija* Пловдивски университет "Паисий Хилендарски" - Научни трудове. Филология (Plovdiv University "Paisii Hilendardski" - Research Papers. Philology), 46(1), 133-45.

Tobler, A. (1875). *Vermischte Beitrage zur franzosischen Grammatik. Erste bis funfte Reihe*. Leipzig: Hirzel.

Totomanova, A.-M. (2015). "Proektăt 'Informatika, gramatika, leksikografija'i digitalnata obrabotka na srednovekovni slavjanski tekstove" Проектът "Информатика, граматика, лексикография" и дигиталната обработка на средновековни славянски текстове (The Project "Informatics, Grammar, Lexicography" and the Digital Processing of Mediaeval Slavic Texts)". *Sbornik dokladi i materiali ot zaključitelnata konferencija po proekta "Informatika, gramatika, leksikografija"* Сборник доклади и материали от заключителната конференция по проекта "Информатика, граматика, лексикография" (Proceedings of the Final Conference of the Project "Informatics, Grammar, Lexicography"). Sofia: Grafis, 5-16.

Toufexis, N. (2010). "One Era's Nonsense, Another's Norm: Diachronic Study of Greek and the Computer". Bodard, G.; Mahony, S. (eds), *Digital Research in the Study of Classical Antiquity*. Aldershot, 105-20.

Vaillant, A. (1948). *Manuel du vieux slave*. Paris: Institut d'etudes slaves.

Večerka, R. (1989). *Altkirchenslavische (Altbulgarische) Syntax*. Bd. I, *Die lineare Satzorganisation*. Freiburg.

Večerka, R. (1993). *Altkirchenslavische (Altbulgarische) Syntax*. Bd. II, *Die innere Satzstruktur*. Freiburg.

Wackernagel, J. (2009). *Lectures on Syntax. With Special Reference to Greek, Latin and German*. Ed. by D. Langslow. Oxford: Oxford University Press.

Zaimov, Y.; Capaldo, M. (1982-83). *Suprasălski ili Retkov sbornik* Супрасълски или Ретков сборник (Suprasălski or Retkov Collection), vols 1-2. Sofia: Izdatelstvo na BAN.

Zaliznjak, A.A. (2008). *Drevnerusskie enklitiki* Древнерусские энклитики (Old Russian Enclitics). Moskva: Jazyki slavjanskich kul'tur.

Zimmerling, A.V. (2013). *Sistemy porjadka slov slavjanskih jazykov v tipologičeskom aspekte* Системы порядка слов славянских языков в типологическом аспекте (Word-Order Systems in Slavic Languages in a Typological Perspective). Moskva: Jazyki slavjanskich kul'tur.

Zwicky, A.M. (1985). "Clitics and Particles". *Language*, 61(2), 283-305. https://doi.org/10.2307/414146.

Zwicky, A.M. (1977). *On Clitics*. Bloomington (IN). https://web.stanford.edu/~zwicky/on_clitics.pdf.

Balcania et Slavia
Vol. 1 – Num. 1 – June 2021

Postposed Articles and DP Structures in Torlak

Jelena Živojinović
Università degli Studi di Verona, Italia; UiT Norges arktiske universitet, Tromsø

Beatrice Azzolina
Università degli Studi di Verona, Italia

Veronica Girolami
Università degli Studi di Verona, Italia

Abstract This article sheds light on postposed articles and DP structures in Torlak, a non-standardised Balkan Slavic variety. Torlak and specifically Trgoviste-Torlak, unlike Bulgarian and Macedonian, does not exhibit MD. We argue that this scenery is due to a partial grammaticalization of the determiner, which is arguably an inflectional affix and maintains the demonstrative feature. In addition, we verify the nature of the Torlak DP and we make some considerations on the intermediate nature of this element with respect to the grammaticalization path, followed by the other Balkan Slavic varieties.

Keywords Torlak. Balkan Slavic. Multiple determination. Articles. DP structure.

Summary 1 Introduction. – 2 State of the Art. – 2.1 Early Proposals on the Balkan DP. – 2.2 Current Assumptions on the Structure of the Balkan DP. – 3 Testing the Nature of the Enclitic Article. – 4 Looking for the Torlak DP Layer. – 5 Discussion and Further Remarks. – 5.1 Multiple Determination. – 5.2 Grammaticalization Hypothesis. – 6 Conclusion.

Edizioni Ca'Foscari

Peer review

Submitted	2021-07-20
Accepted	2021-10-15
Published	2020-12-20

Open access

Citation Živojinović, J.; Azzolina, B.; Girolami, V. (2021). "Postposed Articles and DP Structures in Torlak". *Balcania et Slavia*, 1(1), 97-116.

1 Introduction

Torlak is a non-standard Old-Shtokavian Slavic variety spoken in the southern or southeastern area of Serbia and in the bordering areas of Bulgaria and Macedonia (see **fig. 1** below). It is also called the Prizren-Timok dialect in the attempt to delimit its distribution within the boundaries of Serbia. This variety has recently started getting more attention from the scientific community, mainly dealing with a variety of phenomena related to the central and north-eastern Timok area.[1] So far, the literature has focused on the postposition of the article and on the use of pronominal clitics, claiming that Torlak is a transitional area having both Balkan and non-Balkan features. For instance, it has second-position clitics as Serbo-Croatian, but makes consistent use of direct object clitic reduplication as Bulgarian and Macedonian.[2]

When it comes to the use of articles, the Trgovište-Torlak variety presents an overt postposed *t*-type particle deriving from the demonstrative pronoun *taj/ta/to* 'that', for instance:

(1) a. Bulgarian

dete-to
child.DET
'the child'

b. Albanian

ribar-at
fisherman.DET
'the fisherman'

c. Romanian

žena-ta
woman.DET
'the woman'

The particle can also enclitisize to the (first) adjective, for instance in (2), whereas the presence of a demonstrative pronoun incorporates the determiner as in (3).

1 Runić 2013a; 2013b; Vuković, Samardžić 2015; Makarova et al. 2020; Vuković 2021, among others.

2 Cf. Runić 2014, 11-94 for an overview of the clitic system; Živojinović 2021 for a comparative perspective on the use of clitics in Torlak.

(2)

ubavo-to	*malecko*	*dete*
beautiful.DET	little	child

'beautiful little child'

(3)

Ja	*gu*	*čuvam*	*kako*	*golupče,*	*u*
I	her	guard/raise.PRES.1SG	as	pigeon	in
onuj	*buljinu*	*što*	*kuka*	*cel*	
that	owl	that	complain/cry.PRES.3SG	entire	
a	*ona*	*izraste*			
but	she	grow.PRF.3SG			
noć					
night					

'I raise her like a pigeon (kindly), but she grew into that owl that cries all night'

The productivity and distribution of such particles seems to be considerably subjected to diatopic variation. Indeed, the literature on postposed articles and DP structures in the north-eastern and central areas shows instead some inconsistencies; for example, Runić (2014, 66) suggests that Torlak[3] is an article-less variety, whereas Vuković and Samardžić (2015) present the distribution of the postposed overt article. However, the former study is based on a data collection in urban areas, such as the central Niš and Leskovac area, whereas the latter relies on fieldwork data gathered in north-eastern isolated and urban areas of the Timok Valley. Indeed, Vuković and Samardžić (2015) show that the postposed article[4] is productive in isolated rural areas (e.g. *đubre-to* 'the garbage') that are less influenced by the Serbo-Croatian superstratum.

Our research presents a contribution in the study of the postposed articles and DP structures by providing novel data from the southern sub-variety of Trgovište, which borders Macedonia and Bulgaria (see the indication in **fig. 1**). The data was collected in the rural areas of Trgovište through the recordings of natural speech of 8 participants whose age range is 70-90.

3 PTS (Prizren-Timok Serbian) in Runić's terms (2014).

4 Specifically, Vuković and Samardžić (2015) argue that the Timok area maintains all three types of articles, namely the distal *t-* and *n-* types, deriving from demonstrative pronouns *taj, ta, to* and *onaj, ona, ono* 'that' as in *brat-at* 'the brother/that brother' and *vrata-na* 'the door/that door' respectively, and the proximal *v-*type, as in *baba-va* 'the grandmother/this grandmother'.

Figure 1 The overall distribution of Torlak

This investigation is an additional puzzle piece to Stanković (2013) who scrutinises three Torlak isoglosses[5] along with Macedonian in terms of Bošković's generalisations (2008), showing that DP/NP partition is not a sustainable description for these varieties. What

5 Stanković (2013) relies on the partition of Torlak varieties provided by Ivić (1994), who distinguishes the following isoglosses: Kosovsko-Resavski, Prizrensko-Južnomoravski, Svrljiško-Zaplanjski and Timočko-Lužnički. Trgovište-Torlak is located within Prizrensko-južnomoravski, an article-less variety (cf. Ivić 1994).

ALL and LWA indicate article-less languages and languages with articles respectively. Trgovište-Torlak is located within Prizrensko-južnomoravski (PJ).

Stanković (2013), who assumes that the southern Torlak isogloss is an article-less variety, proposes is rather a phonologically null vs. saturated DP structure on the basis of a set of structural and lexical properties which the isoglosses may or may not share.

In this article we will address the postposition of articles by testing the nature of the Torlak enclitics and we delineate its domain by framing Trgovište-Torlak within the Balkan Slavic context. Paragraph 2 of this article illustrates the existing proposals on the Balkan DP structure, which mainly focus on Bulgarian and Macedonian data. Paragraph 3 tests the nature of the Torlak article-like particles following Halpern (1992), whereas paragraph 4 provides further clues on the Torlak DP layer. Finally, paragraph 5 presents a brief comparative analysis with respect to the multiple determination phenomenon as opposed to Bulgarian and Macedonian and provides a note on the development of the Torlak article.

2 State of the Art

The issue surrounding the structure of postposed article has been a central topic in the generative literature.[6] Accordingly, several proposals were put forward in order to explain the structure of the DP in the Balkan languages.

In the following paragraphs, we will revise some of the most influential assumptions regarding the DP in Balkan Slavic essentially referring to Balkan Slavic languages, i.e. Bulgarian and Macedonian, to provide an appropriate starting point for the analysis of our Torlak data. In addition, some considerations on the nature of the postposed article will follow. As such, we will mostly build on Halpern (1992) and Franks (2001) to shed light on the affixal properties of the postposed article found in Trgovište-Torlak.

2.1 Early Proposals on the Balkan DP

The extensive bulk of literature on postposed articles has established a clear relationship between the presence of this affixal-like marker of definiteness in languages like Romanian, Albanian, Bulgarian and Macedonian to the Balkan *Sprachbund*, for instance:

6 Tomić 1996; Dimitrova-Vulchanova, Giusti 1995; 1998; Franks 2001; Rudin 2018a; 2018b for Balkan Slavic; Cornilescu 2016 for Romanian, among others.

(4) a. Bulgarian

kniga-ta
book.DET
'the book'

b. Albanian

krevat-i
bed.DET
'the bed'

c. Romanian

baiat-ul
boy.DET
'the boy'

Nonetheless, the assumption on whether the postposed article can be considered an actual inflectional marker has been widely debated. We will return to the affixal nature of the postposed article with specific reference to Torlak in paragraphs 3, 4 and 5. However, it seems crucial to point out two major lines of research that consider postposed articles either as (i) proper clitics (Scatton 1980; Tomić 1996) or (ii) inflectional affixes (Halpern 1992; Franks 2001; Rudin 2018a; 2018c; Embick, Noyer 2001 among others).

Tomić (1996), for instance, argues for a clitic-like nature of the article, which is to be found in a Wackernagel position within the DP. She claims that the postposed article is generated on the D°, triggering the movement of the highest, leftmost head to SpecDP. The movement of the noun, or whichever element follows the clitic, be it an adjective, numeral, or possessive, is considered a type of transformation applying from D-structure to S-structure in pure government and binding terms. In other words, the article as a nominal clitic is generated in D° and triggers the movement of N to SpecDP, which ends up in a spec-head relation with the article. However, even though Tomić (1996) assumes that nominal clitics, i.e. articles and possessive clitics, are Wackernagel clitics, she also admits that they seem to show some typical properties that are normally ascribed to affixes.[7]

For this reason, Dimotrova-Vulchanova and Giusti (1998) dispensed with N-to-D movement,[8] arguing that Bulgarian does not dis-

7 An anonymous reviewer also points out the fact that, if we were to follow the idea that the article is merged in D, coordinated Ns would display the article only on the first noun contrary to fact.

8 See also Bošković 2019 for a more detailed explanation to discard N-D raising in affixal article languages.

play any instance of movement of the noun to a higher position, as opposed to the optional movement of N, which is found in both Albanian and Romanian. This assumption allows the authors, on the one hand, to rule out the movement of N to D in the narrow syntax and, on the other hand, to rely on the movement of N to D at LF in order to check definiteness [Def] features. Franks' (2001) proposal goes along these lines, assuming movement of the highest head below D at LF including AP, which, following Abney's (1987) account, dominates NP.[9]

These two accounts, therefore, provide a solid basis to consider the article in Balkan Slavic (and non-Balkan Slavic languages) a proper affix, regardless of their assumptions related to movement operations taking place at LF.

2.2 Current Assumptions on the Structure of the Balkan DP

Abstracting away from the proposals that were analysed in § 2.1, we now review Rudin's (2018a; 2018b) assumptions as a starting point to better capture the structure of the Torlak DP. For now, we will consider the relationship between D° and X° as an agreement relation for definiteness, bearing in mind the existence of different proposals (Koev 2011; Petroj 2020 among others) which argue for phi-features and definiteness agreement.[10]

Considering the structure in (5), Rudin (2018a; 2018b) adopts an Abney-type of structure in which the AP dominates the NP.[11] She claims that D, being phonologically null, bears definiteness features that enter in an agreement relation with the articled word, namely the first head below the DP.

9 In Dimitrova-Vulchanova, Giusti 1998, the AP moves entirely to SpecDP and it checks [Def] features within the AP projection.

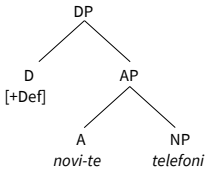
10 In a slightly different way, Giusti (2015) argues for the absence of a [Def] feature, introducing a scattered head bundled with other features proper to the nominal group, which could explain the presence of the inflectional article especially in the case of Romanian.

11 An anonymous reviewer points out that the head status of adjectives in structures like (2) is challenged by the fact that it fails to describe cases like the following:

> [$_{DP}$ [$_{AP}$ glavna-ta po značenie] pričina] Bg. (Cinque 2010, 47)
> Lit. 'the main in importance reason'.

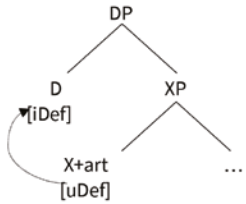
We acknowledge Cinque (2010) in asserting the phrasal status of adjectives and their adjunction to the NP. However, we believe that, for the purposes of this paper, both accounts, i.e. AP dominating NP or AP adjunction to NP, allow us to describe the behaviour of definiteness agreement with respect to post-posted articles in these Balkan Slavic varieties. Rudin (2018a) also points out that, even by adopting the perspective of adjunction, definiteness agreement still appears on the highest leftmost head, namely A.

(5)



As it is shown above, the agreement relation between the adjective and the [Def] feature allows us to explain the presence of the inflectional article on the highest head below D, in this case, AP. Following the idea put forward in Rudin (2018a; 2018b), we could try to describe definiteness agreement through the operation Agree (Chomsky 2000; 2001) by adopting a bidirectional approach (see for instance Baker 2008; Bjorkman, Zeijlstra 2019 among others).

(6)



As show in (6), the uninterpretable [Def] feature acts as a probe for the interpretable [Def] feature on D; it therefore looks up, as an instance of upward Agree, and checks its [uDef].[12]

According to these theoretical premises, in principle, it could be possible to apply the structure in (6) to the Torlak DP. However, before turning to the analysis of the DP of this Balkan Slavic variety, some considerations on postposed articles as inflectional markers for definiteness are needed.

12 Franks (2020) proposes to consider this type of agreement in terms of feature sharing, where D assigns definiteness features to XP that, in turn, shares them with the head X. As such, this account considers definiteness agreement as a more morphological process than the operation Agree and crucially as a more local relation.

3 Testing the Nature of the Enclitic Article

In the previous subsections, we mainly argued for the presence of an affixal-like postposed article specifically for Bulgarian and Macedonian. At this point of the analysis, it is necessary to understand where Trgovište Torlak stands concerning the nature of the enclitic particles and whether the assumptions that were put forward for Bulgarian and Macedonian still hold for this Balkan Slavic variety. The analysis covering our Torlak data is carried out below, after providing a quick review of the main phonological arguments to consider postposed articles as actual suffixes as outlined in Halpern (1992) and subsequently applied to Bulgarian in Franks (2001).

Halpern (1992) proposes four "tests" to validate the hypothesis that articles are actual suffixes, i.e. final-devoicing as shown in (7), consonant-schwa metathesis as shown in (8), changes in the placement of the stress as shown in (9) and the appearance of a special form of the stem in the articled word as shown in (10).

(7) a. Bulgarian

bratovčed [bratofčet]
'cousin' (Franks 2001)

b. Bulgarian

bratovčedât [bratofčedət]
cousin.DET
'the cousin' (Franks 2001)

(8) a. Bulgarian

grâk[*]
'Greek' (Franks 2001)
* The letter *â* represents a schwa.

b. Bulgarian

gârkât
Greek.DET
'the Greek' (Franks 2001)

(9) a. Macedonian

brAtučed
'cousin' (Tomić 1996, 531-2)

b. Macedonian

bratUčedot
cousin.DET
'the cousin' (Tomić 1996, 531-2)

(10) a. Bulgarian

interesen *grad*
interesting city
'the interesting city'

b. Bulgarian

interesnijat *grad*
interesting.DET city
'the interesting city'

These tests, carried out by Franks (2001), clearly show the affixal nature of the post-posted article in Bulgarian ruling out their clitic-like nature. Building on these assumptions, it is now crucial to understand whether the Torlak article displays any difference with respect to the neighbouring Slavic languages.

Firstly, final devoicing is a systematic phenomenon in Trgovište-Torlak, as in *brod-brot* 'ship', *grad-grat* 'town', *Glog-Glok*,[13] *leb-lep* 'bread'. Devoicing is blocked by the addition of the suffixal element that forms CVC syllables, for instance *grat* 'town' vs. *gradat* 'the town'.

The second test determined by Halpern (1992) is more problematic, as confirmed by Franks (2010, 111-13), who instead argues that the metathesis is the result of a schwa epenthesis (see (4) above). On this note, we do not identify relevant examples in our Trgovište-Torlak corpus either related to (ii) or (iii). Nonetheless, it is worth noting that 3+ syllable words containing a postposed *t-* particle bear the stress on the antepenultimate syllable, as in *trAktor-at* 'the tractor'.

The word length however plays a role, thereby allowing us to apply Halpren's fourth test, which broadly affirms that the masculine form of the article only attaches to long adjectival stems. This is clearly visible in adjectives such as *ubav* 'beautiful', or *nov* 'new', which do not allow forms such as **ubav-at muž* or **nov-at stol*, but require a restructuring of the syllable arrangement in the adjective, obtaining *ubavijat muž* 'the beautiful husband' or *novijat stol* 'the new table'.

The application of the tests related to the nature of the enclitics provides therefore a partial outcome. By claiming such results, we intend that the *t*-particle may not be a fully functional suffixal ele-

13 *Glog* is a village located in the area of Trgovište/Vranje.

ment, but it could be retaining some degree of lexical value, hence the absence of the particle in examples containing a demonstrative (cf. (3) illustrated above).

4 Looking for the Torlak DP Layer

Coming back to the analysis of the Balkan Slavic DP, it seems clear that the assumptions that we put forward in § 2.2 may be challenged by the behaviour of the Torlak article with respect to Halpern's tests. As we previously argued, these articles partially fit the tests, showing different outcomes compared to the Bulgarian and Macedonian ones. Considering these results, it may be useful to go deeper into the analysis of the Torlak DP, testing whether this understudied variety fits Bošković's (2008; 2012) diagnostics as an NP or DP language. Tasseva-Kurktchieva and Dubinsky's (2018) study already applied some of Bošković's (2012) 18 diagnostics on Bulgarian and their results showed that this Balkan Slavic language falls in a category that is neither the one of full DP nor a full NP language. Their proposal is that Bulgarian is, in fact, a weak DP language that projects the DP layer only in the presence of the definite article. Following from their results, we applied the same diagnostics, checking the behaviour of the structures listed in table 1, to Trgovište-Torlak in order to find out its status and the behaviour of the DP projection in this understudied variety.

Table 1 Syntactic contexts based on Bošković 2008

(i) Clitic doubling
(ii) Left-branch extraction
(iii) Negation raising
(iv) Scrambling
(v) Presence of majority superlative reading
(vi) Superiority and Multiple Wh-Fronting
(vii) Adnominal Genitives
(viii) Head-Internal Relatives and Locality

Specifically, clitic doubling and the related generalisations have been widely discussed in Živojinović (2021), claiming that Trgovište-Torlak stands in a compromise position showing overt postposed articles and allowing clitic doubling, but having Wackernagel-type clitics (11). On this matter, Bošković generalises that (i) second position clitic systems are only found in NP languages, (ii) only languages with articles may allow clitic doubling, (iii) there is no clitic doubling with second-position clitics.

(11)

Odamna	*ga*	*upozna*	*Milovana.*
long_time_ago	him.CL.ACC	met	Milovan

'I met Milovan a long time ago'.

When it comes to left-branch extraction, Trgovište-Torlak does not allow such structures (12), following Bošković (2008) who affirms that only languages without articles may allow LBE.

(12)

**Maleckoto*	*vido*	*dete.*
little.DET	saw	child

While Torlak does not allow scrambling, as predicted for languages with articles, the Neg-raising test requires some further explanation. Precisely, Bošković (2008) claims that the negation in a matrix clause negates the content of a subordinate clause and in such contexts it makes use of a licensed negative polarity item. However, Stanković (2013) observes that languages without articles, such as Serbo-Croatian, exhibit Neg-raising, but do not license an NPI with verbs such as *believe*. This observation can be extended to Trgovište-Torlak as well.

(13)

Ivan	*ne*	*veruje*	*da*	*Bog[k]*	*postoji.*
Ivan	not	believe	that	God	exists

'Ivan does not believe that God exists'. (Cf. Stanković 2013)

Trgovište-Torlak does not show the majority superlative reading unlike varieties such as Slovenian (for instance, *Največ ljudi pije pivo* 'Most people drink beer'). Similarly, it does not show strict superiority effects to multiple wh- fronting, an expected feature of varieties with articles (e.g. *Koj koga vidi? / Koga koj vidi?* lit. Who whom sees/ whom who sees). On the same line, our Torlak subvariety does not provide examples of two adnominal genitives and only allows head-external relatives.

Once again, the overall position of Trgovište-Torlak seems to be a compromise one, partially fitting Bošković's (2008) generalisations and presenting features belonging to article, but also article-less varieties.

5 Discussion and Further Remarks

The discussion on definite articles in Torlak and the analysis of the previous sections allows us to make some further remarks, specifically concerning the comparison among Bulgarian, Macedonian, and Torlak. An interesting phenomenon worth discussing, which interests mostly Bulgarian and Macedonian, falls under the name of Multiple Determination (MD). The presence (or absence) of MD provides us with the opportunity to shed some light on the properties of the suffixal articles found in the three languages in question. The framework that we adopt follows Rudin (2018c) with the addition of a novel proposal aiming at explaining the anomalous behaviour of Torlak compared to the other neighbouring varieties. Namely, we will call into question the process of partial or total grammaticalization of the article and its outcome in the different languages.

5.1 Multiple Determination

As anticipated, one of the phenomena that is worth analysing to better understand the characteristics and peculiarities of the DP in Torlak is Multiple Determination (MD). MD is defined as the presence, inside the DP layer, of a double or multiple realisation of the DP in certain environments (Alexiadou 2014). There are different hypotheses that try to explain the functioning of MD, two of the most relevant are (i) the split DP-Hypothesis and (ii) the 'distributed' DP-Hypothesis. According to the first one, the DP is divided into at least two layers that contribute differently to the meaning of the structure. This distinction is between a part where deixis is encoded and another part where determination is: [DP1 Deixis [DP2 Determination]]. According to the other hypothesis, instead, the Det can realise several other non-D related projections within the extended projection of the noun.

In the Balkan Slavic languages analysed, namely Bulgarian and Macedonian, MD is characterised by the presence of a Demonstrative and one or more definite article suffixes (Rudin 2018c), like the examples in (14) show:

(14)

a. Bulgarian

tija	*novi-te*	*telefoni*
these	new.DET	phones

'these new phones'

b. Macedonian

tie	*ubavi-te*	*fustani*
these	pretty.DET	dresses

'these pretty.DET dresses'

According to Rudin (2018c), in Bulgarian and Macedonian MD is used in colloquial speech and has a specific pragmatic reading that can be judged either positively or negatively by the speakers. However, as it will be demonstrated below, the situation concerning MD in Torlak is dramatically different. In particular, any instance of MD is judged ungrammatical by speakers of this language. Consider for instance (15).

(15) Torlak

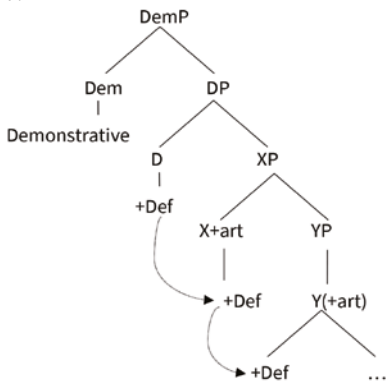
***ovoj**	*ubavo-**to***	*dete*
this	beautiful.DET	child

'this beautiful.DET child'

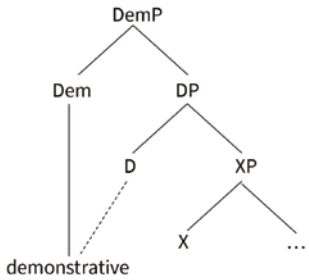
Going back to the structural analysis of the DP, we assume that in case of MD, the Dem is located higher up in the structure. Furthermore, we hypothesise that in Bulgarian and Macedonian there is a split in the DP and that the features that enter in an agreement relation are those present in the D head while those belonging to the Dem are only deictic. As such, since the featural content of the two heads does not interfere with one another, MD is allowed (a). In Torlak, instead, there is no split in the DP between the features of the Dem and of the D head and for this reason agreement is possible only from one head at a time (b). The two structures in (16) are taken from Rudin (2018c).

(16)

a.



b.



5.2 Grammaticalization Hypothesis

The peculiar characteristics of the DP in Torlak, which seems to behave differently with respect to other Balkan Slavic languages like Bulgarian and Macedonian, can be explained as caused by a process of partial grammaticalization. The grammaticalization path that is taken into consideration here is the one used to describe some Scandinavian varieties too, for instance Nynorsk Norwegian, namely *content item* > *grammatical word* > *clitic* > *inflectional affix* (Faarlund 2018, 618, the original theoretical framework from Hopper, Traugott 2003). In the case of Macedonian and Bulgarian, the articles seem to have undergone a full grammaticalization into purely functional elements and for this reason they are able to occur with a demonstrative in a MD construction. In Torlak, instead, the grammaticalization of the Det is only partial resulting in an inflectional affix that maintains the demonstrative semantics. We argue that the peculiarity of the situation in Torlak is that the Dem and the Det share the features that are semantically encoding part of their meaning, namely the demonstrative ones. This sharing is not allowed because the two heads are not differentiated enough due to the grammaticalization of the determiner, that is only partially undergone in Torlak. This claim is also supported by the results of the tests carried out in the previous sections, which confirm a compromise position of the article.[14] In the case of Bulgarian and Macedonian, instead, the complete process of grammaticalization renders the Det a purely functional element that does not enter into competition with the Dem for the semantic features encoding [+Dem] and hence the construction

14 An anonymous reviewer asks whether Torlak affixes encode deictic differences as the Macedonian article and the variety of Bulgarian spoken in the Rhodope mountains. As we previously argued, Trgovište-Torlak only displays a t-type of article; however, some other varieties of Torlak (cf. Vuković, Samardžić 2015 for the Timok area) may retain this distinction.

allows for MD. Being grammaticalization a gradual process, even in Bulgarian and Macedonian it is possible to find articles that have not completely undergone semantic bleaching.

6 Conclusion

In this paper we attempted to provide a contribution to the study of the Balkan Slavic postposed article by providing novel data from an understudied non-standardised variety. The study highlights yet again the transitional nature of Torlak that balances Balkan and non-Balkan features. Indeed, Torlak and specifically Trgovište-Torlak present a postposed suffixal article-like element, but they do not allow the use of multiple determination. We argued that this behaviour with respect to MD is explained by a lack of a split in the D features in Torlak, which is instead present in the neighbouring Bulgarian and Macedonian. To conclude the article, we noted that the peculiar behaviour of the DP in Torlak could be due to a partial grammaticalization.

Abbreviations

ACC	Accusative
CL	Clitic
CVC	Consonant-vowel-consonant
DET	Determiner
DP	Determiner phrase
LBE	Left branch extraction
NP	Noun phrase
NPI	Negative polarity item
PRES	Present
PRF	Perfect
SG	Singular
XP	X phrase

Bibliography

Abney, S. (1987). *The English Noun Phrase in Its Sentential Aspect* [BA thesis]. Cambridge (MA): Massachusetts Institute of Technology.

Baker, M. (2008). *The Syntax of Agreement and Concord*. New York: CUP.

Alexiadou, A. (2014). *Multiple Determiners and the Structure of DPs*. Amsterdam; Philadelphia: John Benjamins Publishing Company.

Bjorkman, B.M.; Zeijlstra, H. (2019). "Checking up on (-)Agree". *Linguistic Inquiry*, 50, 527-69. https://doi.org/10.1162/ling_a_00319.

Bošković, Ž. (2008). "What Will You Have, DP or NP?". Elfner, E.; Walkow, M. (eds), *Proceedings of the 37th Annual Meeting of the North East Linguistic Society*. Amherst: GLSA, University of Massachusetts, 101-14.

Bošković, Ž. (2012). "On NPs and Clauses". Grewendorf, G.; Zimmermann, T.E. (eds), *Discourse and Grammar: From Sentence Types to Lexical Categories*. De Gruyter Mouton, 179-246. https://doi.org/10.1515/9781614511601.179.

Bošković, Ž. (2019). "On Affixal Articles: An Argument from Bantu for Bulgarian, Romanian, and Icelandic". Franks, S.; Chidambaram, V.; Joseph, B.; Krapova, I. (eds), *Katerino Mome: Studies in Bulgarian Morphosyntax in Honor of Catherine Rudin*. Bloomington: Slavica Publishers, 37-52.

Chomsky, N. (2000). "Minimalist Inquiries: The Framework". Martin, R.; Michaels, D.; Uriagereka, J. (eds), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge (MA): MIT Press, 89-155.

Chomsky, N. (2001). "Derivation by Phase". Kenstowicz, M. (ed.), *Ken Hale: A Life in Language*. Cambridge (MA): MIT Press, 1-52.

Cinque, G. (2010). *The Syntax of Adjectives. A Comparative Study.* Cambridge (MA): MIT Press.

Cornilescu, A. (2016). "The Low Definite Article and the Evolution of the Romanian DP". *Generative Grammar in Geneva*, 9, 3-26.

Dimitrova-Vulchanova, M.; Giusti, G. (1995). "Quantified Noun Phrase Structure in Bulgarian". *Working Papers in Linguistics*, 5(1), 43-65.

Dimitrova-Vulchanova, M.; Giusti, G. (1998). "Fragments of Balkan Nominal Structure". Alexiadou, A.; Wilder, C. (eds), *Possessors, Predicates and Movement in the Determiner Phrase*. Amsterdam: John Benjamins, 333-60. https://doi.org/10.1075/la.22.12dim.

Embick, D.; Noyer, R. (2001). "Movement Operations after Syntax". *Linguistic Inquiry*, 32(4), 555-95.

Faarlund, J.T. (2009). "On the History of Definiteness Marking in Scandinavian". *Journal of Linguistics*, 45(3), 617-39. https://doi.org/10.1017/S0022226709990041.

Franks, S. (2001). "The Internal Structure of Slavic Noun Phrases, with Special Reference to Bulgarian". Przepiörkowski, A.; Banski, P. (eds), *Generative linguistics in Poland: Syntax and Morphosyntax*. Warsaw: Instytut Podstaw Informatyki PAN, 53-69.

Franks, S. [1998] (2010). "Clitics in Slavic". *Glossos*, 10, 1-157.

Franks, S. (2020). *Microvariation in the South Slavic Noun Phrase*. Bloomington: Slavica.

Giusti, G. (2015). *Nominal Syntax at The Interfaces*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Halpern, A. (1992). "The Balkan Definite Article and Pseudo-Second Position". *Proceedings of the Eighteenth Annual Meeting of the Berkeley Linguistics*

Society: General Session and Parasession on The Place of Morphology in a Grammar. Berkeley: Berkeley Linguistic Society, 338-49. https://doi.org/10.3765/bls.v18i1.1593.

Hopper, P.J.; Closs Traugott, E. (2003). *Grammaticalization*. Cambridge: Cambridge University Press.

Koev, T. (2011). "Definiteness as Agreement: Evidence from Bulgarian". Byram Washburn, M. et al. (eds), *Proceedings of the 28th West Coast Conference on Formal Linguistics*. Somerville (MA): Cascadilla Proceedings Project, 133-41.

Makarova, A.; Konior, D.V.; Vuković, T.; Sobolev, A.N.; Winistörfer, O. (2020). "Automatic Language Profiling of a Dialect Speaker: The Case of the Timok Variety Spoken in the Village of Berčinovac (Eastern Serbia)". *Acta Linguistica Petropolitana*, 16(2), 160-80. https://doi.org/10.30842/alp2306573716207.

Petroj, V. (2020). *Ignoring Language Barriers: Romanian-Serbian Code-Switching* [PhD dissertation]. https://opencommons.uconn.edu/dissertations/2473.

Rudin, C. (2018a). "On DP Structure in Balkan Slavic: Evidence from Multiple Determination". *Formal Approaches to Slavic Linguistics 27*. Stanford: Michigan Slavic Publications, 418-37.

Rudin, C. (2018b). "Multiple Determination in Bulgarian and Macedonian: An Exploration of Structure, Usage, and Meaning". Dickey, S.M.; Lauersdorf, M.R. (eds), *V zeleni drzeli zeleni breg: Studies in Honor of Marc L. Greenberg*. Bloomington: Slavica Publishers, 263-86.

Rudin, C. (2018c). "Demonstratives and Definiteness: Multiple Determination in Balkan Slavic". Blumel, A.; Gajic, J.; Geist, L.; Junghanns, U.; Pitsch, H. (eds), *Advances in Formal Slavic Linguistics 2018*. Berlin: Language Science Press, Ch. 12, 305-38.

Runić, J. (2013a). "On Some Doubling Constructions in Non-Standard Serbian and Slovenian Dialects". *Quaderni di lavoro ASIt*, 16, 21-44.

Runić, J. (2013b). "Cliticization Phenomena in Languages 'on the Border'". *University of Pennsylvania Working Papers in Linguistics*, 19(1), art. 21. https://repository.upenn.edu/pwpl/vol19/iss1/21.

Runić, J. (2014). *A New Look at Clitics, Clitic Doubling, and Argument Ellipsis: Evidence from Slavic* [PhD dissertation]. Storrs: University of Connecticut.

Scatton, E. (1980). "On the Shape of the Bulgarian Definite Article".Chvany, C.; Brecht, R. (eds), *Morphosyntax in Slavic*. Columbus: Slavica, 204-11.

Stanković, B. (2019). "DP or NP? The Case of Serbian Southeastern Dialects and Macedonian Language". Stanković, B.; Janić, A. (eds), *SinFonIJA 6 Proceedings (The 6th International Conference Syntax, Phonology and Language Analysis (SinFonIJA 6), held at the Faculty of Philosophy of University in Niš, Serbia* (September 26-28, 2013). Faculty of Philosophy, University of Niš, 83-118.

Tasseva-Kurktchieva, M.; Dubinsky, S. (2018). "On the NP/DP Frontier: Bulgarian as a Transitional Case". Franks, S.; Chidambaram, V.; Joseph, B.; Krapova, I. (eds), *Katerino Mome: Studies in Bulgarian Morphosyntax in Honor of Catherine Rudin*. Bloomington: Slavica, 287-312.

Tomić, O.M. (1996). "The Balkan Slavic Nominal Clitics". Halpern, A.L.;Zwicky, A., *Approaching Second Position Clitics and Related Phenomena*. Stanford: CSLI Publications, 511-36.

Vuković, T.; Samardžić, T. (2015). "Prostorna raspodela frekvencije postpozitivnog člana u timočkom govoru" (Areal Distribution of the Frequency of the

Post-positive Article in the Timok Vernacular of Southeast Serbia). Ćirković, S. (ed.), *Timok. Folkloristička i lingvistička terenska istraživanja 2015-2017* (Timok. Folklore and Linguistic Field Research 2015-2017). Knjaževac: Narodna biblioteka "Njegoš", 181-99.

Vuković, T. (2021). "Representing Variation in a Spoken Corpus of an Endangered Dialect: The Case of Torlak". *Lang Resources, Evaluation*. https://doi.org/10.1007/s10579-020-09522-4.

Živojinović, J. (forthcoming) "Torlak Clitic Doubling: A Cross-linguistic Comparison". Blümel, A.; Gajić, J.; Geist, L.; Junghanns, U.; Pitsch, H. (eds), *Advances in Formal Slavic Linguistics 2018*. Berlin: Language Science Press, 423-41.

Balcania et Slavia
Vol. 1 – Num. 1 – June 2021

Language Ideologies and Multilingual Practices of Post-Soviet Migrants in Western Europe from a Translanguaging Perspective

Erzhen Khilkhanova
Institute of Linguistics, Russian Academy of Sciences, Moscow, Russia

Abstract The main point of this paper is to describe, discuss and analyse multilingual practices of non-Russian migrants from the former Soviet Union from a translanguaging perspective uncovering language ideologies underpinning these practices. Using data collected through a 3-month linguistic ethnography supplemented by linguistic analysis of informal online communication, the Author found that fluid, translingual practices are generally not characteristic for the majority of well-educated post-Soviet migrants. Instead, we observe in the normative linguistic behaviour a lack of need or unwillingness to cross language boundaries and create hybrid linguistic forms.

Keywords Language ideologies. Multilingualism. Post-Soviet migrants. Non-Russian migrants. Translanguaging. Code switching. Code mixing. Social network groups. Western Europe.

Summary 1 Introduction. – 2 Theoretical Framework: Concepts of Translanguaging, Code Switching and Language Ideologies. – 3 Methodological Considerations. – 4 Translanguaging Practices of Post-Soviet Non-Russian Migrants. – 5 Language Ideologies at Work. – 6 Results and Discussion.

Edizioni Ca'Foscari

Peer review

Submitted	2021-07-20
Accepted	2021-10-25
Published	2020-12-20

Open access

Citation Khilkhanova, E. (2021). "Language Ideologies and Multilingual Practices of Post-Soviet Migrants in Western Europe from a Translanguaging Perspective". *Balcania et Slavia*, 1(1), 117-148.

1 Introduction

Although migration from both pre- and post-Soviet Russia is polyethnic, this aspect has been of interest to researchers mainly in relation to three groups: 'Russian Germans', 'Russian (Soviet) Jews' and ethnic Russians (see, e.g., Ben-Rafael et. al. 2006; Isurin 2011; Dietz 2000). Representatives of other ethnic groups were usually considered as an integral part of the undifferentiated mass of 'Russian' emigration, which in the ordinary sense means not only citizens of Russia, but also all immigrants from the former Soviet Union in general (hereinafter FSU).[1] Non-Russian ethnic groups from this region, especially in the post-Soviet period, mainly attracted scholarly interests of cultural anthropologists, ethnologists and sociologists. Although their research touched upon language issues as well (see, for example, Guchinova 2004), linguistic works concerning ethnic (minority)[2] languages of Russia in the migration context are scarce, compared to the countless literature on the Russian in migration (see e.g. Anstatt 2012; Besters-Dilger 2013; Glovinskaya 2004; Zemskaya 2001). These few linguistic and sociolinguistic works on non-Russian languages and migration address topics of language and identity (Yusupova, Nabiullina 2016; Bedretdin 2017; Khilkhanova 2017), structural and functional features of Tatar and other languages in foreign diasporas (Akhmetova 2004; Nabiullina, Yusupova 2015; Khilkhanova 2020; Bitkeeva 2018), mechanisms and consequences of language contacts with host country languages, e.g. Chinese and English (Nabiullina, Yusupova 2014; Yusupova et. al. 2013; Aliev 2017).

The Author thanks many people and organisations who contributed to this research. First, the project could not be done without the cooperation of migrants from Russia and Kazakhstan. Though I cannot name them here I nevertheless would like to thank all the anonymous informants who shared their life stories and sincerely answered my questions.

I would also like to express my gratitude to Dr. Günther Schlee and the staff and researchers of the Max Planck Institute for Social Anthropology (Halle/Saale).

Last but not least, I owe my deep thanks to the German Academic Exchange Service (DAAD) that provided financial support for this project. The reported study was also funded by the Russian Foundation for Basic Research (RFBR) and Deutsche Forschungsgemeinschaft (DFG), project number no. 21-512-12002 ННИО *Prognostic Methods and Future Scenarios in Language Policy – Multilingual Russia as an Example*.

1 The shortcoming of the definition of 'Russian' is also due to the fact that most Western European languages do not have the differentiation between the concepts of *rossijskij* and *russkij* appeared during the perestroika, where the first denotes *citizenship*, and the second means *ethnicity* (along with the term *ethnicity* in the Russian scientific tradition, the term *nationality* is often used).

2 In the Soviet and especially in the Russian context, all non-Russian languages fall under the definition of 'minority languages' related to the languages of those ethnic groups that represent a demographic minority in the country, and languages whose functions and scope of use are limited compared to the official language of this region (in this case, Russian).

I am interested in another linguistic aspect of post-Soviet migration, that is how language ideologies manifest themselves in the multilingual strategies and practices of the multiethnic body of migrants from the FSU, regulating their language attitudes and language practices. This problem is studied in this article in connection with the concept of translanguaging, which is becoming increasingly popular in modern studies of multilingualism, language contacts and migration. Translanguaging means "the ability of multilingual speakers to shuttle between languages, treating the diverse languages that form their repertoire as an integrated system" (Canagarajah 2011, 401).

This problem statement stems from my observations made during a field study conducted in 2016 within the framework of a project *Language and Ethnic Identity of Non-Russian Immigrants from the Former Soviet Union in Western Europe*. The aim of the project was to examine processes of negotiation, (re)construction and transformation of ethnic identity and its connection with language and culture among non-Russian migrants from the FSU. This multiethnic body of immigrants shared some common characteristics that were essential for my research aim and made them an interesting and little-studied research object: they were united by a common history, (Soviet) culture and (Russian) language, but at the same time represented other ethnicities and cultures. From the linguistic point of view, each ethnic group was multilingual and had at least three languages in their linguistic repertoire: the ethnic language (L1) (sometimes only at a symbolic level), the Russian language (L2) and the host country language/languages (L3-n).

The project aim did not include the study of translanguaging, but during the interview and participant observation I noticed a very limited use of the migrants' multilingual repertoire, lack of language creativity and switching between languages – what is usually defined as translanguaging. This fact was particularly striking against the background of quite frequent translanguaging strategies and practices of migrants in different regions of the world (Lee 2014; Li 2018; Pennycook 2008). I assume that this is due to language ideologies acquired in childhood that continue to guide (to a certain extent) the speech behaviour of migrants whose acculturation took place in the FSU. Therefore, the goal of this article is to examine multilingual practices of first-generation migrants from the FSU from a translanguaging perspective aiming to reveal and analyse translingual resources (or their absence) used by this multilingual and multicultural group and to uncover language ideologies underpinning these practices.

This research falls within the frame of critical ethnographic sociolinguistics. At the beginning of the article, the theoretical framework of this study is considered, namely concepts of translanguaging, code switching and language ideologies (paragraph 2). Then, I will

provide methodological considerations related to the characteristics of the social research object and research materials (paragraph 3). I will then examine translingual practices of post-Soviet migrants in their oral speech and informal written communication in social networks (paragraph 4). The next section is devoted to the description and analysis of several language ideologies in the historical context of the Soviet language policy and in the modern context of migrants' life in Western Europe. Finally, the summary of the results of this research is provided and conclusions are drawn.

2 Theoretical Framework: Concepts of Translanguaging, Code Switching and Language Ideologies

Language ideologies are the systems of ideas about social and linguistic relationships and how they relate to social values (Woolard 2003; Kroskrity 2000, 5). They have an affective dimension and are related to what Cavanaugh calls a "social aesthetics of language" (2009, 194-5). They are also defined as "the cultural conceptions of the nature, form, and purpose of language, and of communicative behavior as an enactment of the collective order" (Gal, Woolard 1995, 30).

Sociolinguists distinguish many ideologies related to the use or non-use of ethnic, or minority languages. These include folklorisation, hypertraditionalisation, and the association of the language only with the past and old times (Sallabank, Marquis 2018; Woolard 2020). Many ideologies revolve around the concepts of legitimacy, 'language ownership', and authenticity usually going hand in hand with language purism (Pischlöger 2016). In real life, these ideologies are intertwined with each other and their action is carried out in a complex way; often contradictory ideologies coexist in the minds of the same people.

In the context of my research, it is important to identify and show how language ideologies can act as an internal regulator of people's linguistic behaviour. Language ideologies are also closely related to translanguaging, since they are the source of linguistic variability and the driving force for the creation of new hybrid 'languages' by people who are on the border of two or more 'worlds'. This is why translingual strategies are so common in the migration environment. Migrants invent a new language to reflect their inner reality determined by their migration experience and destabilised identities (Hüppauf 2004).

The theory of translanguaging, which originates within the framework of pedagogy, has long gone beyond its framework and is actively used in modern studies of multilingualism and language contacts. Today, linguists are already saying that it is necessary to distinguish

between the pedagogical use of translanguaging and spontaneous translanguaging, the latter referring to the complex discursive practices of bilinguals (see García 2009). In addition to the well-known definition by Canagarajah given above, translanguaging is also defined as a smooth synergetic transition from one linguistic culture to another, as a result of which there is some merging of them, while there is no complete assimilation and the linguistic and cultural identity of language users is preserved, and a mixed discourse is created (Canagarajah 2002). Translanguaging implies the permeability of languages, their mutual influence, as a result of which a new quality of an enriched linguistic culture arises. The translingual area is characterised by blurriness, indistinctness, uncertainty, "linguistic fluidity" (Lee 2014). In language studies and in language pedagogy, the translingual approach focuses not on language as a system, but on language as a practice, i.e. activity (Lee 2014, 305).

It should be emphasised that translanguaging is the use of a speaker's entire communicative repertoire, as a result of which languages and cultures smoothly flow into each other, sometimes helping, and sometimes hindering (Pennycook 2008). Code switching and code mixing are considered one of the translingual strategies (Pennycook 2008, 30.4), or as an instance of translanguaging, alongside other bilingual phenomena such as translation, borrowing, and additional processes, in a range of modalities (MacSwan 2017, 191).

Thus, translanguaging is, firstly, a broader phenomenon including code switching (hereafter CS) and code mixing (hereafter CM).[3] Secondly, the more fundamental difference between them is the paradigmatic difference. The translanguaging concept belongs to the postcolonial paradigm (see Pennycook 2006), while the notions of CS and CM appeared in earlier studies of multilingualism and language contacts. This causes a difference in approaches: the terms CS and CM are based on the vision of languages as separate systems. Accordingly, speech, where there is a mixing and switching of codes, was often assessed rather as defective speech, since the reasons for these phenomena were seen in the speaker's insufficient linguistic competence in any of the languages involved, or even in both. This is the basis for the designations of such mixed speech as "macaronic" (Bert-

3 The term *code switching* is used here as a term with a broader meaning, which includes *codes mixing* (see also Myers-Scotton 1993, 1; Gumperz 1982, 59). The preference for the term *code switching* is also due to considerations of brevity and convenience. At the same time, I understand by code mixing the process of using two (or more) codes (languages) in speech, which has acquired a more or less regular form and has gone further than just code switching. But the most important differentiating criterion is grammatical: if the morphosyntactic rules of both languages are not violated, this is *code switching*. If the lexemes acquire morphological indicators of another language, this is *codes mixing*. In other words, as long as the morphosyntactic basis of a particular code is preserved, we can talk about CS. If not, about CM.

agaev 1969, 127), "mestizo language": in relation to the FSU, these characteristics were often used to describe speech, where Russian and other languages of the peoples of the USSR were mixed. A puristic, normative approach to language is clearly seen behind these epithets. The translanguaging approach is fundamentally different and positive. Translanguaging

> has a pronounced active and lingvocreative character, representing a new, creative, transformative and integrative use of all bilingual resources. Additional opportunities for the lingvocreative activity of bilinguals arise precisely in the 'indefinite zone' on the border of languages. (Proshina 2017, 161; see also MacSwan 2017, 190)

Since the main difference lies precisely in the approach, in the assessment of translanguaging and CS phenomena, I believe that a researcher can, however, work with the terminological apparatus traditionally used in CS research, as it is done, for example, in research by Karpava et al. 2019. Otherwise, the scholars remain both without a terminological apparatus, and, as J. MacSwan rightly notes, without empirical case and are left only with an ideological one that denies languages as social and political constructs (MacSwan 2017, 169, 177). I generally support MacSwan's multilingual perspective on translanguaging, which acknowledges the existence of discrete languages and multilingualism reconciling educational (that is, translanguaging) and linguistic research. The reconciliation becomes possible through distinguishing mental grammars from linguistic repertoires, arguing that bilinguals, like monolinguals, have a single linguistic repertoire but a richly diverse mental grammar (MacSwan 2017, 167).

Since, speaking of translanguaging, we talk primarily about individual bi- and multilingualism,[4] a look at what motivates people to switch and mix codes sheds additional light on the mechanism of translanguaging. Factors that contribute to the dynamic use of two or multiple languages in a particular instance can be broadly divided into three groups – external, internal, and linguistic:

1. *external* factors, i.e. independent of a speaker (such as history, politics or demography);
2. *internal* factors, i.e. attached to a speaker, both as an individual and as a group member (such as psycholinguistic, pragmatic, sense of ethnic identity);

4 Although, as J. MacSwan rightly notes, a language is the product of a community of speakers, a collection of overlapping individual languages (MacSwan 2017, 174).

3. *linguistic* factors (such as lack of a concept in a minority language-L1, insufficient linguistic competence of speakers in L1).

Although all the factors are intertwined and act in a complex way in a real language contact situation, translanguaging and CS from the point of view of this scheme are motivated primarily by *internal* factors. External or linguistic factors also contribute to "multiple discursive practices in which bilinguals engage in order to make sense of their bilingual worlds" (García (2009, 45), but in a more indirect way. In some way, only unconscious or, vice versa, conscious, purposeful cases of CS driven from within can be considered truly translingual strategies. The first ones are valuable for sociolinguistics because 'automatic' CSs, when a speaker would like to, but cannot speak without switching the code, or when a mixed code has become a kind of norm for the speaker (or for the entire language community) are the true indicators of real processes occurring in the contacting languages. Automatic 'slips of the tongue' in another language can be caused by such psycholinguistic factors as (i) insufficient linguistic competence in L1 combined with higher linguistic competence in a majority language (L2); (ii) the habit of speaking one of the languages from the bilingual repertoire (a stereotype of speaking behaviour); (iii) linguistic economy, or economy of speaking efforts.

Vice versa, a deliberate, goal-oriented CS discovers various speakers' intentions: for instance, the translingual strategy of mixing languages can be chosen consciously – as a special slang – to construct one's group identity, which also includes ethnicity (Moustaoui, Prego, Zas 2019). Vivid examples of 'protest' motivation can be found in translingual/transcultural literature created by bi- or polylingual writers. This is, for instance, the book by the German-Turkish author Feridun Zaimogly *Kanak Sprak*. The book's language is the language of protest, violence, obscenity, 'dirty' metaphors and slang, but it is also some kind of artificial language containing words and expressions that are not present in either German or Turkish (Yildiz 2004). Similarly developed and used is the pidginised German of 'Russian Germans' called 'kvelya' and 'legitimised' by Vladimir Kaminer's book *Russendisko* (Kaminer 2000; Kirilina 2011). My research, however, focuses not on professional creative writing 'at the crossroads of cultures' but on 'ordinary' multilingual migrant speakers and their translanguaging practices through which they represent their experiences and relationships in their new localities.

As already mentioned, translanguaging is an integral feature of increasing migration flows in the modern world. In the sociolinguistic literature, quite well studied are the 'big' languages involved in the translingual process in a migration situation. These are, for example, new transnational forms on the border between German and Turkish languages (Zaimogly 2004), translanguaging practices in the

home among bilingual/multilingual Russian-speaking children and their parents in Cyprus, Sweden and Estonia (Karpava, Ringblom, Zabrodskaja 2019). But the most widespread translingual forms are (for obvious reasons) the so-called World Englishes (including Runglish, Spanglish, Chinglish etc.) (Li 2018; Pennycook 2008). From the point of view of translanguaging theory, these combined variants that have arisen as a result of cultural and linguistic contact and carry the features of both English and other languages (sometimes serving as a means of communication for entire countries) are not defective speech, but specific translingual formations.

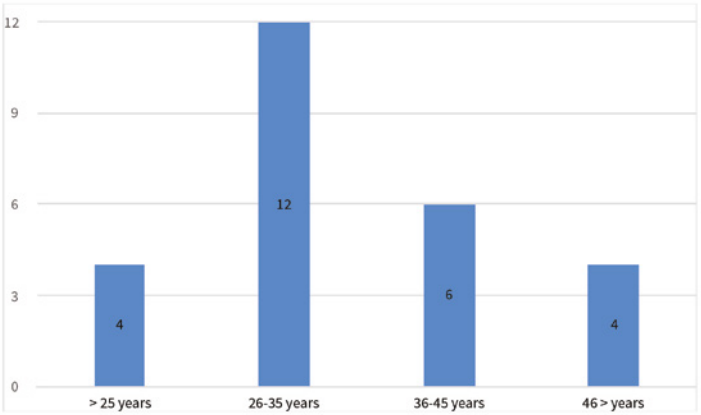
With the approaches, definitions and theoretical discussions presented above in mind, let us now turn to the methodological considerations and the description of data collection procedures.

3 Methodological Considerations

The data include a corpus of spoken materials, collected as part of the field research conducted in the period from July to October 2016 in four Western European countries (Germany, France, Netherlands, and Great Britain with the focus on Germany).[5] A total of twenty-six in-depth, semi-structured interviews lasting from one to several hours were collected. The interviews were conducted, recorded and transcribed by the author in person. Apart from the interviews, the data were collected employing extensive ethnographic observation. From an ethnic point of view, the pool of informants included Buryats, Yakuts and Kazakhs from two countries, i.e. Russia (twenty-four people) and Kazakhstan (two people). Of course, there is a bias towards the Russian area, which is explained both by the complexity of data collection among migrants and by a rather short field study duration. Therefore, the conclusions are made mainly on the Russian material, although the commonality of background and linguistic strategies of post-Soviet migrants are very much alike, as evidenced in other studies on post-Soviet migration (Kasatkina 2011). Interviews typically lasted one to two, sometimes more hours, and were usually held in public venues, such as restaurants, local businesses, and private homes.

The diagrams below **[figs 1-2]** show two important characteristics of the sample: age and the fields of professional activity/status of informants. I divided the study participants into four age groups, and

5 All subjects gave their informed consent for inclusion before they participated in the study. Transcripts are fully anonymised, with each participant being ascribed a pseudonym. Any information that could reveal the informants' identity such as names of localities, places of study etc. is removed from the published interview fragments. Information about the countries of residence is, however, retained, since it is impossible to determine the participants' identity from it, and yet this is important for analysis.



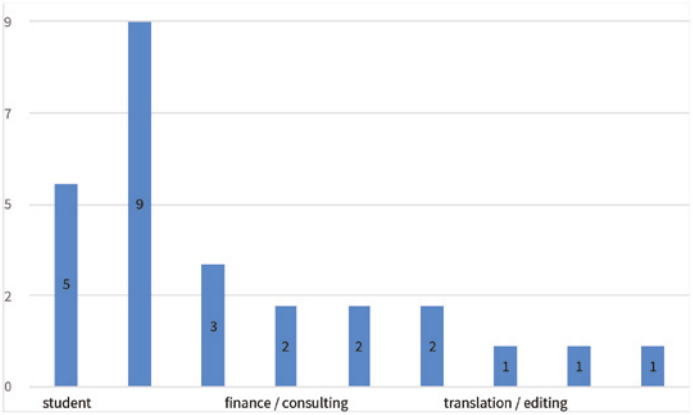


Figure 1 Age characteristics of informants

Figure 2 Professional characteristics / status of informants

we can see that the most numerous (almost half of the respondents) was the group of migrants, whose ages ranged from 26 to 35 years.

The informants' educational level and social status are reflected in their fields of activity **[fig. 2]**. Note that almost all participants worked or studied, except for three who were housewives at that time. Figure 2 shows that the majority of respondents work in the field of education and science (32%) or are college students (23%), mainly enrolled in master programmes.

The interview toolkit consisted of (1) questions identifying the respondents' social characteristics (age, sex, marital status, number of years spent in Western European countries, information on education both in the home country and the host countries, professional background, migration status as well as reports on the ethnic self-identification of the participants), (2) questions aimed at identifying the interaction of language, culture and ethnic identity, (3) questions that solicited participants' reports on their language proficiency in Russian, ethnic language and L3-n, and the use of these languages in the migration context. In addition, attention was paid to the social communication circle of post-Soviet migrants, i.e. social networks. An important goal was to find out how social networks (connections) of migrants affect their language practices and, vice versa, what this case adds to the already known models of interaction between the structure of social networks and the use of languages in migrant communities. Therefore, a part of the interview questionnaire focused specifically on eliciting information on participants' networks and the links between the personal social networks and language use.

Since the original purpose of the study was not related to translanguaging, I drew attention to the almost complete absence of a dynamic and functionally integrated use of different languages in the participants' speech later. Therefore, for a deeper study of this topic I decided to study samples of spontaneous speech, since it is in spontaneous speech that the creativity of multilingual speakers can manifest itself to the fullest (see also García 2009). Interviews, despite being held in a relaxed atmosphere and being close to a natural conversation, cannot by definition represent completely free speaking. Therefore, the field research materials were subsequently supplemented with data obtained from the analysis of virtual social communities of several migrant groups from post-Soviet Russia. Social media such as Facebook are a popular object of research on translingual practices (Halim, Maros 2014), because they represent

> a new, delocalized, multimodal space positioned on the boundary between orality and literacy in which translingual practices emerge spontaneously. (Moustaoui, Prego, Zas 2019, 3)

This method called 'virtual ethnography' was carried out on posts and informal discussions in social network groups of Russian-speaking migrants of non-Russian ethnicity residing in different western countries.

So, generally this study is based on a variety of methods to obtain and analyse empirical data, namely interviews and participant observations as central ones, in conjunction with a virtual ethnography of online interactions of Buryat, Yakut and Kalmyk migrants living in western countries and the sociolinguistic analysis of online social

media. Six social network groups in Facebook and Instagram became the object of my study. In total, I analysed the written speech practices of the following online communities: Kalmyks in USA / Kalmyki v SShA, which included 3,000 members at the time of writing this part of the article (June 2021), sakhadiaspora.official in Instagram with 2,695 subscribers and Sakha Diaspora community organisation in Facebook (171 followers), and three Buryat groups in Facebook (Vstrechi v Evrope; Buryat House, USA; and Buryat connection UK with total 1,600 members). All social media were viewed in the period from 2016 to June 2021, with the exception of the Sakha Diaspora Facebook page, which was created on December 11, 2020.

Using virtual ethnography poses additional ethical considerations, although most Internet communication is considered public due to its free availability. Like in other studies employing data from social media (Kliuchnikova 2016, 71), my approach was also to avoid using any material that could not be accessed freely, without the special permission of the author. Nevertheless, I still do not provide the full (nick)names of the writers I quote. Only the initials, names of the participant's group and the dates of publications are indicated.

For the comparative analysis of modern data with historically earlier information about the speech practices of non-Russian migrants from the Soviet Union, the paper employs secondary sources, i.e. interviews with Oirat-Kalmyks living in the United States. The interviews were published in the book by E. Guchinova who studied the history and ethnicity of the Kalmyk community in the USA in 1997-98 (Guchinova 2004).

It should be noted that the pool of informants is limited by two important and related characteristics: the high level of education of the studied migrants and the legality of migration channels. Almost all interviewed informants had a higher education (24 people), seven of them possessed PhD and Candidate of Sciences degree, and two people have finished professional (vocational) schools. The high level of education reflects the general picture of post-Soviet migration, which is closely related to the problem of 'brain drain' from Russia. For example, in the United States the 2000 census recorded a very high level of education among the Russians who arrived: more than half of those over 25 years old (52%) had a bachelor's and master's degree. Only one in four Americans had an equivalent level of education (24%) (Denisenko 2013). In the early 2000s, 43% of Russian migrants in OECD countries aged 15 years and over had higher education. For immigrants from most other countries, including the OECD, this indicator was lower (Dumont, Lemaître 2005).

Thus, the studied group represents a certain segment in the general emigration flow from FSU countries – namely intellectual migration. Currently, the concept of 'intellectual migration' mainly describes the processes of departure of scientists and qualified per-

sonnel to work outside their country. But there is no single meaningful definition of the phenomenon of intellectual migration in the scientific discourse. The broadest interpretation implies that intellectual migrants are all professionally qualified persons who have been abroad for more than one year (Ledeneva 2014, 107). This broad interpretation of intellectual migration is adopted in this paper. Accordingly, the characteristics and behavioural strategies of Buryat, Kazakh and Yakut migrants are determined by belonging to this category of migrants and, most likely, cannot be extrapolated to other groups, for example, unskilled migrants or illegal labour migration.

The next limitation was due to my ignorance of the Kazakh and Yakut languages, which also excluded the possibility of Kazakh-Russian and Yakut-Russian CS and CM. However, considering that there were 22 Buryat migrants among the Russian participants, the possibility of communicative practice that "blurs or breaks through apparent boundaries" (Blackledge, Creese 2016, 2) between Russian and Buryat remained open but was rarely noticed by the researcher during ethnographic observation and interviews. It should be noted that the possibility of translingual transitions between Buryat and Russian was initially excluded for seven Buryat informants who did not know the Buryat language (there were no such people among the Yakut and Kazakh informants). Despite my attempts to conduct a conversation in the Buryat language with Buryat informants, most of them could not maintain a lengthy conversation in it. As one of the participants put it, "I'm not as good [in Buryat] as you are" (Vlada, 32, Germany). So, the main language of the interview and communication beyond it was Russian.

In addition, all informants had a good (in many cases native-like) command in at least one foreign language, primarily English, but also German, French etc. They knew that the interviewer was also fluent in English and German, so they had a full opportunity to dynamically use at least these languages along Russian. However, despite their multilingual past and present, the speakers remained mostly within the monolingual Russian discourse, although the cases of translanguaging were also present. They are discussed in the following section.

4 Translanguaging Practices of Post-Soviet Non-Russian Migrants

In this section, I consider translingual transitions and creative, playful usage (or the lack thereof) of languages by migrants in oral speech (recorded during interviews and ethnographic observation) and informal written communication in social networks. The data obtained contain not so many examples when the speakers collaboratively engage in fluid, plurilingual communication practices.

The recorded cases of translanguaging were mostly *insertions* of single lexemes (sometimes longer speech segments) from L1 and L3-n into the predominantly Russian discourse. I interpret insertions as incidental cases of CS (see also Myers-Scotton 1993, 163). The identified CSs can be grouped into the following categories:

1. non-equivalence vocabulary (when describing cultural-specific realities) or words for which the speaker could not immediately find a translation equivalent or did not want to spend speaking efforts on it:[6]

> А самое-то интересное было, что дети иногда приезжали, и заходили вот так за порог, да, и мы, значит, с ними здороваемся, я в бурятском всегда была, *буряад дэгэлтэй, гутал, малгагай ху умдөд*. (Densema, 56, Germany)
> And the most interesting thing was that children sometimes came, and went like this over the threshold, yes, and we, then, say hello to them, I was always in Buryat, *in the Buryat clothes, shoes, cap and all on.*

> ALYONA Я учусь сейчас на Arzthelferin
> INTERVIEWER Но это не bachelor's?
> ALYONA Нет. Просто Ausbildung. Среднее (Alyona, 24, Germany)

> ALYONA I am currently studying to become medical assistant.
> INTERVIEWER But it's not bachelor's?
> ALYONA No. Just training. Vocational'.

> *Mendid Kalmyks*! In celebration of the year of the *Gaha*, there will be a *Tsagan* Dance next month in Philly. ALL PROCEEDS WILL BENEFIT THE PHILLY *HURUL*! (Facebook, Kalmyks in USA / Калмыки в США, T., 6 January 2019)
> *Hello Kalmyks*! In celebration of the *Pig* year, there will be a *White* Dance next month in Philly.
> ALL PROCEEDS WILL BENEFIT THE PHILLY *TEMPLE*!

2. Clichéd etiquette formulas (for example, greetings):

> Они заходят, я говорю: "*Сайн байна! Зэ орогты даа, дээшэ гарагты*", и потом на русском. Так дети ни бурятского, ни русского не знают! (Densema, 56, Germany)
> They come in, I say: "*Hello! Well, come in, come right in*", and then in Russian. So the children do not know neither Buryat nor Russian!

> *Хальмгуд, мендут*! кто в США, добавляйтесь! Давайте общаться (Facebook, Kalmyks in USA / Калмыки в США, A., 20 March 2019)
> *Kalmyks, hello*! Who is in the USA, please add! Let's communicate'

6 Transcripts' legend: *italics* for ethnic language (L1); standard font for Russian language (L2); underlined for English, German (L3-n); **bold** for mixed translingual forms. Examples from social networks are given completely in the authors' design. I only removed non-verbal symbols (such as emoticons) as uninformative in the context of this study.

3. Quotation:

… и я вдруг где-то там в шкафу нашла стихотворение "*шинии нюдэн хара мойһон шэнги*…" (Densema, 56, Germany)
… and I suddenly found a poem "*your eyes are like black currants*" somewhere in the closet…

Они говорят: Sie sind doch hier geboren, т.е. они себя чувствуют себя как дома. (Sara, 25, Netherlands)
They say: "They are born here", that is, they feel at home'.

4. Communication topic:

Могу *буряадар дуугарха* (Radzana, 48, Germany)
'I can *speak Buryat*'

Ухата эк-эцкнр! Даже в американской реальности надо попытаться сделать маленький шаг навстречу калмыцкому языку!
(Facebook, Kalmyks in USA / Калмыки в США, D., 14 August 2020)
Smart moms and dads! Even in the American reality, it is necessary to try to take a small step towards the Kalmyk language!

5. Externally unmotivated intersentential CSs:

И там я два года в школу ходила, *тиихэдэ һургуулимнэ буряад байга хаим*. (Densema, 56, Germany)
'And there I went to school for two years, *at that time the school was Buryat*'

Слушайте до конца. *Сертн, серцхәтн хальмгудм*. Просыпайтесь, пробуждайтесь калмыки![7] (Facebook, Kalmyks in USA / Калмыки в США, E., 11 December 2020)
'Listen to the end. *Wake up, wake up Kalmyks*. Wake up, wake up Kalmyks!'

6. Externally unmotivated intrasentential CSs:

Тут можешь сесть в ресторан **рандомно** и у тебя за соседним столом будут сидеть русские, в пабе будут русские, т.е. чувствовать homesick – нету здесь. (Kira, 28, Great Britain)
'Here you can sit in a restaurant randomly and you will have Russians at the next table. There will be Russians in the pub, that is, to feel homesick – there is no such thing here'

Минус в том что постоянно prejudice. Постоянно misunderstanding, все дела. (Sara, 25, Netherlands)
'The disadvantage is that there is always prejudice. Constantly misunderstanding, all that stuff'

7 Here the translation from Kalmyk into Russian is given by the author of this post.

Какие desired outcomes? (Dari, 32, Switzerland)
'What are desired outcomes?'

Не то чтобы она [английская культура] так ярко выражена, она довольно plain по сравнению с другими культурами. (Dari, 32, Switzerland)
'It's not that it [English culture] is so pronounced, it's quite plain compared to other cultures'

Есть предложение, скинуться по £15 с каждого взрослого и накрыть самим стол, а самое главное налепить as much as we want наши БУУЗЫ. (Facebook, Buryat connection UK, D., 23 February 2017)
'There is an offer to chip in £15 from each adult and cover the table ourselves, and most importantly, to make as much as we want our BUUZAS'.

7. Language creativity, language game:

А ты **хамагуист**? :) (Facebook, Buryat connection UK, B., 12 October 2020)
Are you a **khamaguist ('don't-care-at-all loafer')**?

8. Discourse markers:

И вот сейчас, смотря, да, обратно, also те два года которые я в [mentions the name of the village] провела, были самыми тяжелыми в моей жизни. ('Radzana, 48, Germany')
'And now, looking back, yes, well those two years that I spent in [mentions the name of the village] were the hardest in my life'

When looking at these CS examples, we see that they occur under the influence of external, internal and linguistic factors, which were discussed in section 2. Quotations and communication topic (examples 3 and 4) are well-known and frequent linguistic factors of CS (see, e.g., Gumperz 1982). Similarly, a person has to resort to the resources of another language while referring to the realities of the corresponding culture and using clichéd etiquette formulas, as in examples 1 and 2 (for more detail, see Khilkhanova 2009). As J. Fishman explained it,

> That language which has traditionally been linked with a given ethnoculture is, at any time during which that linkage is still intact, best able to name the artefacts and to formulate or express the interests, values and world-views of that culture. (Fishman 1991, 20)

However, as stated in section 2, primarily unmotivated, unconscious (examples 5, 6, 8) or, on the contrary, conscious, purposeful CSs (example 7) can be considered truly translingual strategies. For instance, in example 7 the popular Buryat expression "khama ugy" ('all the same, no care, no difference') is played with. The addition of the productive Russian suffix -ist creates an occasionalism with the meaning

'a person who doesn't care' (in Russian there is a similar slang equivalent, *pofigist*, formed according to the same model from the non-literary word *pofig*). The morphosyntactic basis of the phrase "Are you a khamaguist?" is provided by the Russian language, serving here as a *matrix language* (Myers-Scotton 1993). In the examples of intrasentential switching, we see that the speakers use different strategies, sometimes mixing, sometimes switching codes. Among the insertions, forms without matrix language affixes are quite frequent – the so-called *bare forms* (Myers-Scotton 2006): чувствовать (homesick); постоянно (prejudice); постоянно (misunderstanding); она довольно (plain). The frequent use of bare forms is noteworthy, because they do not violate the grammar of CS. Other cases, such as the word 'randomly', where the English root morpheme is supplied with the affix of the matrix language – the Russian adverbial suffix -o – are very rare.

The example with a German discursive marker 'also' deserves a separate comment. Discursive markers (hereinafter – DM) in general and in bilingual speech in particular are a separate topic that I can only touch on briefly here. As a rule, DMs perform modal, deictic and metalinguistic functions. In this example, we are dealing with the speaker's automatic choice of the DM 'also', which happens when language units are more active in the speaker's cognitive base due to the more frequent use. In this case, objective, propositional information is expressed in Russian, while subjective, evaluative, additional information is expressed using DM from German. Perhaps this is the first step towards understanding, evaluating, and commenting on reality in another language, i.e. one of the first links in the mechanism of language shift (from Russian to German) by this particular speaker.

It is noteworthy that there are more creative translingual practices and language games in the spontaneous online communication of migrants from the FSU. Nevertheless, the analysis of social networks shows that the language of communication there is mainly Russian, especially in Kalmyk and Buryat social networks. The situation with the Yakut online communication is somewhat different. There are more instances of spontaneous Yakut speech in posts, responses and comments to them, for example:

> Скоро День якутских оладушек! *Күндү Саха Сирин олохтоохторо, болҕойун!*
> *Муус устар 7 күнүгэр Алгыстаах АЛААДЬЫ күнэ бэлиэтэнэр буолла.*
> *Бу күннэ мааны алаадьыбытын астаан аймахтарбытын, доҕотторбутун ыныран амсатабыт, күндүлүүбүт.*
> *Атын группаларга тарҕат, биллэр.*
> Готовьте *алаадьы* #якутскиеоладьи, публикуйте в Инстаграм и не забудьте поставить хэштег #алаадьыкунэ!
>
> The Day of Yakut pancakes is coming soon! *Dear Sakha residents*, *attention! April 7 is the Day of the Blessed Pancakes*.

On this day, we fry pancakes and invite and give to try pancakes, we honour our relatives and friends with pancakes.
Report on other groups.
Prepare *alaadyi* #yakutskieoladyi, post on Instagram and do not forget to put the hashtag #*alaadyikune*!

The activity of the Yakut online and offline community, related to the relatively good state of the Yakut language, is a well-known fact for Turkologists and sociolinguists. According to comparative statistics on the presence of minority languages of the Russian Federation in the global network (Map of Languages 2016), the Yakut language is best represented online along with Bashkir, Udmurt and Tatar. These are the languages with the greatest linguistic vitality among all minority languages of Russia. This indicates that the Internet does not create a new reality, it only reflects the current situation with multilingualism in the country. Languages that are safe without the Internet, with a high level of language activism and national consciousness of native speakers, are widely used in the global network as well.

One could assume that few cases of translanguaging in interviews between Russian and other languages are due to the fact that the interview is not a suitable place where personal linguistic creativity can be deployed. However, examples of other studies where people easily switch between different languages during interviews refute this assumption. For comparison, here is a fragment from the interview with a Kalmyk emigrant of the second emigration wave[8] taken from the book by E. Guchinova who studied the Kalmyk diaspora in the United States (Guchinova 2004, 168):

Главная разница – то, что те калмыки считают, что under коммунизм *бәәсн* и some kind *тиим* анти religion *тиим юмта*... хотя *тенд* были divided и были такие которые действительно against религия, но большинство были и сохраняли *тер*. *Тер* коммунистический режим *далта җилд бәгәд* ... abolish *кечксн – тиигҗ саннав*. Вот *тиим* suggestion. И *би* старался follow *эврәнь* **parents-*инь***. Я не скажу что был best студент, но average... *яһад*. И *эклəд, тигəд тер мана* old generation *Әдрхәрәнь йовад сурһульд* они были *манд* real mother. *Теднə* поведение *манд* понравилось. Они показали нам knowledge, intelligence и *книжкс дуудг*, и мне это понравилось. *Теднə* behaviar маднд influance *кеһəд* мы поехали *Адрхнюр*. Конечно, *мана* уровень **knowledge-*а*** *ик биш билə*, но все же мы могли читать, арифметик… Нас все же приняли в 30-м году в Калмыцкий педагогический техникум.

The main difference is that those Kalmyks believe that *being* under communism and some kind of *such* anti-religion *such things*... although *there* were divided and there were those who were really against religion, but most were and re-

8 The second wave of Kalmyk emigration occurred during the period of the construction of socialism, Stalin's repressions and the Second World War.

> tained *that*. *That* communist regime *having existed for seventy years*... *was* abolished – *so I think*. Here is *such* suggestion. And *I* tried to follow *my* **parents-*in'*.** I will not say that I was the best student, but *like* the average.... And *to begin, then that our* old generation *went to study in Astrakhan*, they were real mother *for us*. *Their* behaviour was liked *by us*. They showed us knowledge, intelligence and *reading books aloud*, and I loved it. *Their* behaviour *made* influence *on us*, we went to *Astrakhan*. Of course, *our* level of **knowledge-*a*** *was not high*, but still we could read, arithmetic… We were still accepted in the year of 1930 in the Kalmyk Pedagogical College.

This fragment, which is typical in the speech practices of non-Russian migrants of previous waves of the Russian emigration (at least the first and second waves), is an example of a translingual discourse that the author has rarely encountered not only in the migrant environment, but also in Russia. In this fragment, all three languages are almost equally involved: L1 (Kalmyk, in italics), L2 (Russian, standard font) and L3 (in this case, English, underlined). It presents a striking contrast to the speech of post-Soviet migrants (especially well-educated ones), which is an example of 'pure', grammatically correct and unmixed Russian speech. Despite the fact that constant bilingualism is often considered to create a specific context that is particularly conducive to developing linguistic creativity (Kellman 2000), speakers remain in a predominantly monolingual Russian-speaking mode both online and offline, even in 'ethnic' social networks and when communicating with co-ethnics.

In general, my data confirmed that for first-generation migrants from the FSU the Russian language plays the role of a collective unifier of the immigrant community, and its pragmatic value is similar to that in the USSR. At the same time, the data obtained indicate the heterogeneity of speech practices and language competencies of the studied migrants' group. For the majority, however, fluid, translingual practices employing the rich repertoire of their multilingual resources are not typical, which is especially noticeable in comparison with the speech of representatives of the previous 'waves' of the multilingual and multiethnic part of the Russian emigration. I believe that this is largely due to linguistic ideologies, which will be discussed in the next section.

5 Language Ideologies at Work

Language ideologies, like any mental formations, are difficult to study due to their implicit representation in speech and often hidden character for the speakers themselves. Sometimes they can be identified based on specific statements, sometimes only by analysing speech behaviour regarding the presence or, on the contrary, absence of relevant linguistic phenomena. Addressing language ide-

ologies presents a new turn in my research, that is I empirically arrived at this concept when trying to explain the rarity of translingual speech practices by the studied migrant group. I will begin my analysis with one piece of interview that is indicative in this regard:[9]

> INTERVIEWER What about the language?
> NORA Buryat? No, [mentions the name of town] is a Russian village anyway. More Russians live and... my parents did not speak Buryat with me. They didn't want people to make fun of me with my accent. They were ridiculed in Soviet times. Well, I myself no longer... did not try, or something, to learn the language. In principle, if I really wanted to, I could have learned it, but... but when everyone around me speaks Russian […]. But when I came here, I missed it a little, and I'm a little ashamed, because I don't know it. Because everyone here asks me first of all: do you know your language? When people ask me where I am from, I say: I am from Russia, from Buryatia. Where's that? I say: well, there … between Lake Baikal and Mongolia. Or they say to me for example: you don't look like a Russian. I say: well, we live there on the border with Mongolia. Oh, do you speak Mongolian? No, don't. Well, it's a little embarrassing, of course. When I moved here, for some reason I wanted to learn Buryat even more than when I was in Buryatia. (Nora, 30, Germany)

This fragment reflects both objective and subjective factors that determined the language situation and speech behaviour of people in the Soviet period through the prism of the migrant's personal perception. One of the objective factors is demography – by 1989, Russians made up 49.3% of the total population of the Soviet Union (All-Union Population Census 1989) and 79.8% in the Russian Federation eleven years after the collapse of the USSR (All-Russian Population Census 2002). Subjective factors refer to people's perception and handling the language inequality in the Soviet Union reflected in Nora's brief narrative. Traumatic experience and internalized ideological views about the status, hierarchy and pragmatic values of Russian and minority languages resulted in a certain linguistic behaviour, when the older generation did not pass on the ethnic languages to children and encouraged their study of the Russian language. In this way, they wanted to save their children from the negative experience they themselves went through and to ensure their social success. In that way, the 'monolingual ideology', i.e. Russian monolingualism, was formed in the mass consciousness of the Soviet people, Russians and non-Russians alike. As A. Burykin put it: "Bilingualism in general is not characteristic of the Russian language mentality. At the level of everyday consciousness, the command and use of other ethnic groups' languages is not encouraged and not welcomed by carri-

9 Since there are no cases of translanguaging here, the fragment is given in English translation at once.

ers of this mentality" (Burykin 2004a; see also Baranova, Fedorova 2018). This statement is confirmed statistically as well: only 0.6% of ethnic Russians spoke regional languages or the languages of the indigenous peoples of Russia, and the tendency of Russians to remain monolingual continues (Mikhalchenko, Trushkova 2003; Kharitonov, Stepina 2020). Speaking in a non-Russian language in the presence of Russian speakers was also considered 'indecent' in almost all regions of the former USSR (Burykin 2004b, 27). It is also a manifestation of 'monolingual ideology' in communicative behaviour. Other studies have also repeatedly stated that during the Soviet years, people were reprimanded or given pejorative assessments in cities because they spoke non-Russian languages in public places (Ferguson 2016b, 99; Khilkhanova 2020).

However, the above fragment reflects not only the ideology of (Russian) monolingualism. Nora also speaks about the shame for not knowing the native language and about her wish to learn Buryat "even more than when I was in Buryatia". Although such feelings and the desire to 'regain' an ethnic language rarely result in concrete actions to learn it and the majority of migrants remain with the same knowledge (or ignorance) of ethnic language with which they went abroad, their language attitudes, however, change. It can be also seen in another interview fragment:

> Я столько разных языков знаю, а своего родного не знаю. Поэтому missing something, ощущение потери. (Sara, 25, Netherlands)
> I know so many different languages, but I don't know my native one. Therefore, missing something, a sense of loss.

This quote illustrates that abroad, in a foreign cultural environment, the ethnic awareness of migrants raises, including a sense of the value of native language as an important component of identity. In addition, the Soviet ideas of language hierarchy and association of the Russian language with progress and civilisation, and other languages with backwardness and an uncivilised way of life are erased in the minds of migrants. Such changes in the perception and evaluation of languages are also determined by the migrants' new European localities. I mean that Western Europe is the birthplace of a new ideological paradigm built on the recognition of multilingual and multicultural values (Smokotin 2010, 4). This ideological turn was caused by the liberalisation of economic, political and social life in Western countries after the Second World War. In the USSR at that time, on the contrary, the early Soviet policy of supporting multilingualism was replaced by the policy of promoting the Russian language as the language of interethnic communication, so that it became the lingua franca of the entire country. Of course, even in Western Europe, the liberal approach is not universally recognised: there are enough

examples of preserving 'old' ideologies such as 'one nation-one language', 'language-of-the-past' (in relation to minority languages), and many others (Sallabank, Marquis 2018; Toivanen 2015). Nevertheless, in general, the liberation from language bias and the increased interest and pride in their language and culture are noted by many interview participants. This indicates that language ideologies as dynamic entities strongly depend on the social context.

The next interview fragment illustrates the attitude of a study participant to her CS between Russian and German:

> А в целом если брать, я чувствую себя свободнее, когда на русском, потому что я лучше могу сформулировать свои фразы и выразиться точнее. То, что переключаюсь – это печально. (Edita, 26, Germany)
> And in general, if you take it, I feel freer in Russian, because I can better formulate my phrases and express myself more accurately. The fact that I switch is sad.

We see that the informant assesses her CS as a disadvantage. In my view, this statement manifests the normative idea about language purity shared by many Russians. Purism and prescriptive language ideologies are very characteristic of the language culture of Russian society. As K. Pischlöger put it,

> Although language purism or 'prescriptivism' is not restricted to Russia, Russia belongs to a 'standard language culture' […] in which there […] is the view that one variety of language has an inherently higher value than others, and that this ought to be imposed on the whole of the speech community […]. This attitude is true particularly for the Russian language, but these ideas and standards are transferred by philologists and speakers of Udmurt and other minority languages in Russia to their own language(s). (Pischlöger 2016, 112)

These prescriptive and puristic attitudes have led to an almost exclusive scholarly interest in the standardised modern literary language (114) and in the 'high' spheres such as fiction, which is also explained by the authority of classical Russian literature. Therefore, 'low' spheres – such as the language of mass communication, spoken language and obscene vocabulary – have become objects of linguistic research in Russia only since the last decades of the twentieth century.

In many respects, such normative language attitudes originate in the Russian school where the normative approach to language dominates. The mode of duty, prescriptions, as well as 'orthographocentrism' are the most important features of Russian linguistic thinking (Golev 2002).

> The desire to 'abstain from mistakes' taken to its farthest limit becomes a brake on the speech development, which should lead to free, creative expression of thought (it is enough to recall the advice of experienced teachers and tutors to their pupils to avoid phrases in the unmistakable spelling of which they are not sure). (Golev 2002, 190)

This is true both in relation to Russian and foreign languages. In Soviet times, the grammar-based method was dominant in foreign language teaching at Russian schools and universities. This resulted in grammatically correct speech, on the one hand, and in the fear of speaking out of fear of making mistakes, on the other.

Therefore, it is perhaps not surprising that translingual freedom and creative treatment of languages, by definition violating their grammatical correctness and 'purity', are not widespread in the mass language consciousness of Russian residents. Of course, this is not an absolute statement, and the lingvocreative treatment of the Russian language is a frequent phenomenon both among professionals and lay people. Constant CS between Russian and minority languages is also widespread in Russia. However, for a large number of people, especially educated ones, correctness is absolutised, dogmatised, and sacralised (Golev 2002, 184), and language purism is one of the most common language ideologies in Russia. This explains the numerous debates about the 'corruption' of the Russian language under the influence of English. One of the Russian politicians V. Zhirinovsky expressed this mass concern proposing to do everything to "stop the Russians from using borrowed words and replace them with native Russian words" (Boguslavskaya, Kitanina 2016, 12). The prevalence of puristic ideologies in relation not only to Russian, but also towards minority languages was noted, apart from K. Pischlöger, by researchers of Sakha-Russian bilingualism (Ferguson 2016a), of Ukrainian and Russian languages (Bilaniuk 2005).

In the migration context, similar results were obtained by Kliuchnikova (2016) who studied linguistic practices, language attitudes and discourses on language(s) of Russian-speaking migrants in the North-East of England. She states that despite her initial assumptions, in her fieldwork she found very little evidence of any 'ludic' initiatives among the group under study, at least in the way that would explicitly operate both sets of linguistic resources – English and Russian. Kluchnikova explains rare linguistic games among Russian-speaking migrants by several reasons including low numbers, irregular contacts and general unpreparedness for creative experiments. However, as one of the cornerstones for maintaining linguistic practices, the author names the set of attitudes towards language(s) and normative linguistic behaviour by bilingual speakers themselves (Kliuchnikova 2016, 277). The most influential language attitude, in

Kluchnikova's view, is Russian speakers' admiration of highly standardised, 'classical', literary norm of Russian that dominates over any urge to explore bilingual limits. The sense of belonging to the 'Russian-speaking continuum' through shared norms and practices leaves little space for bilingual manoeuvres on the margins. Rare counterexamples illustrate the point providing a scarce contrast to wider creative passivity and inertness (Kliuchnikova 2016, 271).

6 Results and Discussion

This research examined the speech practices of a well-educated multiethnic and multilingual group of migrants from the former USSR in Western Europe from the translanguaging perspective. The field study participants were a heterogeneous group not only in terms of ethnicity, but also in terms of linguistic competence in L1, L2 and L3-n. Russian was the native language for one third of Buryat migrants who did not speak their ethnic language. At the same time, all migrants' Russian language proficiency was native or native-like. Also, all of them had high linguistic competence in foreign languages, primarily in English. Thus, the possibility of translanguaging (including CS as one of the translingual strategies or as an instance of translanguaging,) between Russian and foreign languages was open to all research participants, and between L1 and L2, for two-thirds of them.

Nevertheless, the analysis of interviews, ethnographic observation data and the language of migrant social networks revealed few translingual strategies and practices; in fact, instead of translanguaging, both conscious and unconscious language choices in favour of Russian were made. There are many reasons for this: external, internal and linguistic. Among the external ones are low numbers of Buryat, Kazakh and Yakut migrants, irregular contacts with co-ethnics and dispersed living of these small groups' members. However, the main reason for replacing translanguaging with linguistic assimilation is, in my opinion, that first-generation migrants from the FSU countries are products of the Soviet language policy. Therefore, in addition to 'exporting' their linguistic repertoire, they also export the relationships between L1 and L2 that have been established in their home country. Both in the country of origin and abroad, L2 (Russian) satisfies the need for mutual understanding, being the 'language of interethnic communication' of the multiethnic group of migrants from the FSU.

In post-Soviet Russia, language relations between the state (Russian) language and more than 150 other languages have not changed much. The short period of the 'parade of sovereignties' and "mobi-

lised linguicism'[10] in the nineties brought some changes: in almost all national republics of the Russian Federation, laws on languages and the concepts for the development of titular languages were adopted, and minority languages were introduced in school education. However, already shortly afterwards, the state returned to the Russian-dominated ethnolinguistic policy; as a result, the language shift towards Russian continues among most minority groups in Russia (see also Mikhal'chenko 2019).

In addition to the ideas of language hierarchy and the pragmatic value of languages ingrained in the consciousness of people, language purism and the (Russian) 'monolingual ideology' are also widespread in the mass consciousness of Soviet people and Russian citizens. Of course, the 'monolingual ideology' was primarily part of the state language policy, but, like with all ideologies, the mechanism of its formation and functioning is *reciprocal*. By reciprocity I mean that, on the one hand, this and other ideologies regulate people's social behaviour and dictate how they should think and behave. On the other hand, the language ideologies were implemented not only 'from above', but also 'from below', driven by ideological views about languages and their values internalised by the Soviet people. Since the language policy of Russia largely reproduces the Soviet one, as mentioned above, these language ideologies are also common among citizens of modern Russia.

Reciprocity is also manifested in the fact that language ideologies influence social and political reality and, in turn, are formed under the influence of situational, institutional and social contexts. Migration to Western European countries – where, according to informants, there is the rule of law and democratic values, non-discrimination, equality of career opportunities and conditions for self-realisation – changes language ideologies towards greater liberality and tolerance. This proves once again that language ideologies are dynamic entities and strongly depend on the social climate and societal expectations. Similar changes in language ideologies are taking place in Russia, but with a delay of several decades.

All this explains the migrants' reluctance or lack of need to cross the language boundaries and to create hybrid forms. The above-de-

10 The 'parade of sovereignties' was a series of declarations of sovereignty of various degrees by the constituent republics of the Soviet Union from the late 1980s to the early 1990s. It followed a loosened power grip of the Soviet Communist Party as a result of the (democratisation) policy of the perestroika period. These declarations challenged the priority of the all-Union legislation over the republican ones (which started the so-called 'war of laws') and took actions to strengthen the republics' economic independencies. The 'parade of sovereignties' ended with the collapse of the USSR (Sintsov, Bityutskiy 2019). 'Mobilised linguicism' refers to those movements in the republics the USSR consisted of, which ended with the adoption of laws proclaiming the languages of the peoples (nationalities), after which the republics were named, as their state languages (Guboglo 1998).

scribed language ideologies block the potential translanguaging that is based on the recognition of all languages as equal. Or vice versa, translanguaging can also be driven by directly opposite motives: awareness of linguistic inequality and protest against it encourage a person to provocation, aggression and creation of artificial 'languages', as in the mentioned book by Zaimogly, *Kanak Sprak*. In the absence of such awareness and protest, a person, especially one who has passed all levels of the Russian-language education system, automatically reproduces the monolingual patterns of 'pure' speaking in a more prestigious and more frequently used language.

The control of consciousness that prohibits making mistakes and the normative linguistic behaviour are also due to the fact that all highly educated migrants grew up reading classical Russian literature[11] and absorbed the values of Russian 'standard language culture'. Thus, the general cultural level of the migrants also does not contribute to non-literary, profane, provocative treatment of languages. The little 'urge to explore bilingual limits' and to dynamically move between languages applies to both minority and foreign languages in contact with Russian.

To conclude, I hope that empirical findings and theoretical insights generated from the research presented in this paper can improve our understanding of translanguaging and language ideologies, particularly in relation to the FSU region. Hopefully, it also adds to an existing body of research on the Russian-speaking 'ethnic' migration and contributes to the disciplinary fields of sociolinguistics, linguistic and cultural anthropology.

11 Other ethnic literatures are not taught in Russian schools and universities, except for a few so-called 'national' faculties and schools.

Bibliography

Akhmetova, Ja.M. Ахметова Я.М. (2004). *Govor tatar Finljandii: V sravnitel'no-istoricheskom aspekte* Говор татар Финляндии: В сравнительно-историческом аспекте (The Patois of the Tatars of Finland: In the Comparative-historical Aspect) [PhD Dissertation]. Kazan: Kazan State Pedagogical University.

Aliev, R.T. Алиев Р.Т. (2017). "Izmenenie rjada jazykovyh patternov tatar na prigranichnyh prostranstvah: tatary Astrahanskogo kraja i tatary-migranty Kitaja" Изменение ряда языковых паттернов татар на приграничных пространствах: татары Астраханского края и татары-мигранты Китая (The Changes in Some Linguistic Patterns of the Tatars in the Frontier Spaces: The Tatars of the Astrakhan Region and the Tatar Migrants in China). *Journal of Frontier Studies*, 4, 58-71.

Anstatt, T. (2012). "In zwei Sprachen zu Hause: Slavisch-deutsche Mehrsprachigkeit im Ruhrgebiet". Osses, D. (Hrsg.), *Nach Westen. Zuwanderung aus Osteuropa ins Ruhrgebiet*. Essen: LWL-Industriemuseum, 47-53.

Artoni, D.; Pomarolli, G.; Urkhanova, R. Артони, Д.; Помаролли, Д.; Урханова, Р. (2021). "Kul'turnaya i yazykovaja identichnost' buryatskih, kalmytskikh, mongol'skikh immigrantov v Italii" Культурная и языковая идентичность бурятских, калмыцких, монгольских иммигрантов в Италии (Cultural and Linguistic Identity of Buryat, Kalmyk, and Mongol Immigrants in Italy). *Sociolinguistics*, 1(5), 91-112. http://doi.org/ 10.37892/2713-2951-2021-1-5-91-112.

Baranova, V.V.; Fedorova, K.S. Баранова, В.В.; Федорова, К.С. (2018). "Mnogoiazychie v gorode: iazykovaia politika, diskursy i praktika" Многоязычие в городе: языковая политика, дискурсы и практика (Multilingualism in the city: language policy, discourses and practice). *Acta Linguistica Petropolitana*, XIV, Ch. 3, 38-56. http://doi.org/10.30842/alp2306573714303.

Bedretdin, K. (2017). "Preserving of Mother Tongue, Religious and Cultural Identity of Tatar Migrants (the Case of the Tatar Community in Finland)". *Istoricheskaja jetnologija* Историческая этнология (Historical Ethnology), 2(2), 233-42.

Ben-Rafael, E.; Lyubansky, M.; Glöckner, O.; Harris, P.; Israel, Y.; Jasper W.; Schoeps, J. (2006). *Building a Diaspora. Russian Jews in Israel, Germany and the USA*. Leiden; Boston: Brill.

Bertagaev, T.A. Бертагаев Т.А. (1969). "K voprosu o vzaimovlijanii jazykov (o mongolizmah i dvujazychii)" К вопросу о взаимовлиянии языков (о монголизмах и двуязычии) [About the Mutual Influence of Languages (on Mongolisms and Bilingualism)]. *Vzaimodejstvie i vzaimoobogashhenie jazykov narodov SSSR* Взаимодействие и взаимообогащение языков народов СССР (Interaction and Mutual Enrichment of the Languages of the Peoples of the USSR). Moscow: Nauka, 121-5.

Besters-Dilger, J. (2013). "Russian in Germany – Intermediate Results on L1 Attrition." Moser, M.; Polinsky, M. (eds), *Slavic Languages in Migration*. Vienna: LIT, 189-203.

Bilaniuk, L. (2005). *Contested Tongues: Language Politics and Cultural Correction in Ukraine*. Ithaca (NY): Cornell University Press.

Bitkeeva, A.N. Биткеева А.Н. (2018). "Jazykovaja adaptacija diaspory: strategii i perspektivy razvitija" Языковая адаптация диаспоры: стратегии и перспективы развития (Language Adaptation of the Diaspora: Strategies

and Development Prospects). *Acta Linguistica Petropolitana*, 14, part 3, 57-81. http://doi.org/10.30842/alp2306573714304

Blackledge, A.; Creese, A. (2016). "Translanguaging as Cultural and Cosmopolitan Competence". IALIC 2016 Keynote presentation. *Book of Abstracts IALIC 2016 Barcelona* (Barcelona, 25-27 November 2016), 2.

Boguslavskaya, V.V.; Kitanina, E.A. Богуславская, В.В.; Китанина, Е.А. (2016). "Ideologicheskij purizm v kontekste patriotizma: sociokul'turnyj aspect" Идеологический пуризм в контексте патриотизма: социокультурный аспект (Ideological Purism in the Context of Patriotism: Socio-cultural aspect). *Obshhestvo: sociologija, psihologija, pedagogika* Общество: социология, психология, педагогика (Society: sociology, psychology, pedagogy), 3, 10-13.

Burykin, A.A. Бурыкин А.А. (2004a). *Jazyk malochislennogo naroda v ego pismennoy forme (na materiale evenskogo jazyka)* Язык малочисленного народа в его письменной форме (на материале эвенского языка) [The Language of a Small People in its Written Form (Based on the Material of the Even Language)]. St. Petersburg: Peterburgskoe Vostokovedenie.

Burykin, A.A. (2004b). *Jazyk menshinstva kak 'tajnyj jazyk' v otechestvennom sotsiokulturnom kontekste* Язык меньшинства как 'тайный язык' в отечественном социокультурном контексте (электронный ресурс) (Minority Language as a 'Secret Language' in the Russian Socio-Cultural Context). https://www.km.ru/referats/BD7FDF83F8E4453D8804C–216C325E380.

Canagarajah, S. (2002). "Multilingual Writers and the Academic Community: Towards a Critical Relationship". *Journal of English for Academic Purposes*, 1, 29-44.

Canagarajah, S. (2011). "Codemeshing in Academic Writing: Identifying Teachable Strategies of Translanguaging". *Modern Language Journal*, 95, 401-17.

Cavanaugh, J.R. (2009). *Living Memory: The Social Aesthetics of Language in a Northern Italian Town*. Oxford: Wiley Blackwell.

Denisenko, M. (2013). *Historical and Current Trends in Emigration From Russia*. http://russiancouncil.ru/en/inner/?id_4=2292#top (2015-06-10).

Dietz, B. (2000). "German and Jewish Migration from the Former Soviet Union to Germany: Background, Trends and Implications". *Journal of Ethnic and Migration Studies*, 26(4), 635-52. http://doi.org/10.1080/713680499.

Dumont, J.-C.; Lemaître, G. (2005). *Counting Immigrants and Expatriates: A New Perspective*. Paris: OECD.

Ferguson, J.K. (2016a). "Code-Mixing Among Sakha-Russian Bilinguals in Iakutsk: A Spectrum of Features and Shifting Indexical Fields". *Journal of Linguistic Anthropology*, 26(2), 141-61.

Ferguson, J.K. (2016b). "Language Has a Spirit: Sakha (Iakut) Language Ideologies and Aesthetics of Sustenance". *Arctic Anthropology*, 53(1), 95-111.

Gal, S.; Woolard, K. (1995). "Constructing Languages and Publics: Authority and Representation". *Pragmatics*, 5(2), 129-38.

García, O. (2009). "Education, multilingualism and translanguaging in the 21st century". Mohanty Minati Panda, A.; Phillipson, R.; Skutnabb-Kangas, T. (eds), *Multilingual Education for Social Justice: Globalising the Local*. New Delhi: Orient Blackswan, 128-45.

Glovinskaya, M. Гловинская, М. (2004). "Obshchiye tipy izmeneniy v yazyke pervogo pokoleniya emigratsii" Общие типы изменений в языке первого

поколения эмиграции (General Types of Changes in the Language of the First Generation of Emigration). Mustayoki, A.; Protasova, Y. (eds), *Russkoyazychnyy chelovek v inoyazychnom okruzhenii* Русскоязычный человек в иноязычном окружении (A Russian-Speaking Person in a Foreign-Speaking Environment). Helsinki: Helsinki University Press, 13-20. Slavica Helsingiensia 24.

Golev, N.D. Голев, Н.Д. (2003). "Sovremennaja mental'no-jazykovaja situacija v aspekte vzaimootnoshenija ee lingvisticheskogo i lingvodidakticheskogo aspektov" Современная ментально-языковая ситуация в аспекте взаимоотношения ее лингвистического и лингводидактического аспектов (The Modern Mental and Linguistic Situation in the Aspect of the Relationship Between its Linguistic and Linguodidactic Aspects). Golev, N.D. (ed), *Estestvennaja pis'mennaja russkaja rech': issledovatel'skij i obrazovatel'nyj aspekty* Естественная письменная русская речь: исследовательский и образовательный аспекты (Natural Written Russian speech: Research and Educational Aspects). Part. 2: *Teorija i praktika sovremennoj pis'mennoj rechi: materialy konf* Теория и практика современной письменной речи: материалы конф (Theory and Practice of Modern Written Speech: Conference Proceedings). Barnaul: Izd-vo Alt. gos. universiteta, 178-92.

Guboglo, M.N. Губогло, М.Н. (1998). *Yazyki etnicheskoi mobilizatsii* Языки этнической мобилизации (Languages of Ethnic Mobilization). Moscow: Iazyki russkoi kul'tury.

Guchinova E.-B. Гучинова Э.-Б. (2004). *Ulitsa Kalmuck road. Istorija, kul'tura I identichnosti kalmytskoj obsh'iny SShA* Улица Kalmuck road. История, культура и идентичности калмыцкой общины США (The Street Kalmyck road. History, Culture and Identities in the Kalmyk Community of the USA). St. Petersburg: Aleteiya.

Gumperz, J.J. (1982). *Discourse Strategies*. New York: Cambridge University Press.

Halim, N.S.; Maros, M. (2014). "The Functions of Code-switching in Facebook Interactions". *Procedia - Social and Behavioral Sciences*, 118, 126-33. http://doi.org/10.1016/j.sbspro.2014.02.017.

Hüppauf, B. (2004). "Globalization – Threats and Opportunities". Gardt, A.; Hüppauf, B. (eds), *Globalization and the Future of German*. Berlin: Mouton de Gruyter, 3-25.

Isurin, L. (2011). *Russian Diaspora: Culture, Identity, and Language Change*. Berlin; New York: Mouton de Gruyter. https://doi.org/10.1515/9781934078457.

Karpava, S.; Ringblom, N.; Zabrodskaja, A. (2019). "Translanguaging in the Family Context: Evidence from Cyprus, Sweden and Estonia". *Russian Journal of Linguistics*, 23(3), 619-41. http://doi.org/10.22363/2312–9182–2019–23–3–619–641.

Kasatkina, N. (2011). "Language Shift and Maintenance Among Russian Immigrants from the Former Soviet Union". *Arizona Working Papers in SLA & Teaching*, 18, 35-54.

Kellman, S. (2000). *The Translingual Imagination.* Lincoln (NE): University of Nebraska Press.

Kharitonov, V.; Stepina D. Харитонов, В.; Степина, Д. (2020). "Monolingvalnyy gabitus v mnogoyazychnom obshchestve: kak i chto govoryat o yazykakh v Rossii" Монолингвальный габитус в многоязычном обществе: как и что

говорят о языках в России (Monolingual Habitus in a Multilingual Society: How and What People Say About Languages in Russia). *Antropologiya. Folkloristika. Sotsiolingvistika. Konferentsiya molodykh uchenykh. Sbornik tezisov* (Anthropology. Folkloristics. Sociolinguistics. Conference of Young Scientists. Book of Abstracts). St. Petersburg: s.n., 92-4.

Khilkhanova, E.V. Хилханова, Э.В. (2009). *Faktory yazykovogo sdviga i sokhraneniya minoritarnykh yazykov: diskursnyy i sotsiolingvisticheskiy analiz (na primere yazykovoy situatsii v etnicheskoy Buryatii)* Факторы языкового сдвига и сохранения миноритарных языков: дискурсный и социолингвистический анализ (на примере языковой ситуации в этнической Бурятии) [Factors of Language Shift and Minority Languages Maintenance: Discourse and Sociolinguistic Analysis (on the Example of Language Situation in Ethnic Buryatia)] [Habil. thesis]. Barnaul: Altaj State University.

Khilkhanova, E.V. Хилханова Э.В. (2017). "Minoritarnye jazyki i migracija: na primere migrantov iz nacional'nyh regionov Rossii v Zapadnoj Evrope" Миноритарные языки и миграция: на примере мигрантов из национальных регионов России в Западной Европе (Minority Languages and Migration: on the Example of Migrants from the National Regions of Russia in Western Europe). *Vestnik Khakasskogo universiteta im. N.F. Katanova* Вестник Хакасского университета им. Н.Ф. Катанова (Bulletin of the N.F. Katanov Khakass University), 19, 54-6.

Khilkhanova, E.V. Хилханова Э.В. (2020). "Mnogoyazychie postsovetskoy migratsii: funktsional'nyy aspekt" Многоязычие постсоветской миграции: функциональный аспект (Multilingualism of Post-Soviet Migration: Functional Aspect). *Sociolinguistics*, 1, 63-78. http://doi.org/10.37892/2713-2951-2020-1-1-63-78.

Kirilina, A.V. Кирилина А.В. (2011). "Kontakt jazykov v nemeckom jazykovom prostranstve" Контакт языков в немецком языковом пространстве (Language Contacts in the German Language Space). *Vestnik Moskovskoj mezhdunarodnoj akademii* Вестник Московской международной академии (Bulletin of the Moscow International Academy), 2, 36-45.

Kliuchnikova, P. (2016). *Linguistic Biographies & Communities of Language of Russian Speakers in Great Britain* [PhD dissertation]. Durham: Durham University. http://etheses.dur.ac.uk/11374/.

Kroskrity, P.V. (2000). "Regimenting Languages: Language Ideological Perspectives". Kroskrity, P.V. (ed.), *Regimes of Language: Ideologies, Polities, and Identities*. Santa Fe: School of American Research Press, 1-34.

Ledeneva, V.Ju. Леденева, В.Ю. (2014). "Intellektual'naja migracija: mirovye i rossijskie tendencii" Интеллектуальная миграция: мировые и российские тенденции (Intellectual Migration: Global and Russian trends). *Vysshee obrazovanie v Rossii* Высшее образование в России (Higher Education in Russia), 2, 106-13.

Lee, J.W. (2014). "Transnational Linguistic Landscapes and the Transgression of Metadiscursive Regimes of Language". *Critical Inquiry in Language Studies*, 11(1), 50-74. http://doi.org/10.1080/15427587.2014.871623.

Li, W. (2018). "Translanguaging as a Practical Theory of Language". *Applied Linguistics*, 39(1), February, 9-30. https://doi.org/10.1093/applin/amx039.

MacSwan, J. (2017). "A Multilingual Perspective on Translanguaging". *American Educational Research Journal*, 54(1), 167-201. https://doi.org/10.3102/0002831216683935.

Mikhal'chenko, V. Yu. Михальченко, В. Ю. (2019). "Dinamika yazykovoy situatsii v Rossiyskoy Federatsii" Динамика языковой ситуации в Российской Федерации (Dynamics of the language situation in the Russian Federation). Wingender, M.; Bitkeeva, A.N.; Khilkhanova, E.V. (eds), *Perspektivy i vyzovy yazykovoy politiki i yazykovoy situatsii v Krymu. Mezhdunarodnyy kruglyy stol (Moskva, 05-06 noyabrya 2019 g.): doklady i soobshcheniya* Перспективы и вызовы языковой политики и языковой ситуации в Крыму. Международный круглый стол (Москва, 05-06 ноября 2019 г.): доклады и сообщения (Prospects and Challenges of Language Policy and Language Situation in the Crimea. International Round Table (Moscow, 05-06 November 2019): Reports and Presentations). Moscow: Russian Academy of Sciences, Institute of Linguistics of RAS/Institute for Slavic Studies, Justus Liebig University Giessen, The Giessen Centre for Eastern European Studies (GiZo), 8-24.

Mikhalchenko, V.Ju.; Trushkova, J. (2003). "Russian in the Modern World". Maurais, J.; Morris, M.A. (eds), *Languages in a globalizing world*. Cambridge; New York: Cambridge University Press, 260-90.

Moustaoui Srhir, A.; Prego Vázquez, G.; Zas Varela, L. (2019). "Translingual Practices and Reconstruction of Identities in Maghrebi Students in Galicia". *Languages*, 4(3), 63. https://doi.org/10.3390/languages4030063.

Myers-Scotton, C. (1993). *Duelling Languages. Grammatical Structure in Code-Switching*. Oxford: Clarendon Press.

Myers-Scotton, C. (2006). *Multiple Voices: An Introduction to Bilingualism*. Oxford: Blackwell Publishing.

Nabiullina, G.; Yusupova, A. (2015). "On the Linguistic Situation of the Tatar Diaspora in the USA". *Mediterranean Journal of Social Sciences*, 6(5), 298-302. https://doi.org/10.5901/mjss.2015.v6n5s3p298.

Nabiullina, G.A.; Yusupova, A.Sh. (2014). "Tatar-Chinese Language Interactions (Based on Chinese Loanwords in the Speech of the Tatar Diaspora Living in Modern China)". *Journal of Language and Literature*, 5(3), 241-5.

Pennycook, A. (2006). "Postmodernism in Language Policy". Ricento, T. (ed.), *An Introduction to Language Policy: Theory and Method*. London: Blackwell, 60-7.

Pennycook, A. (2008). "Translingual English". Sharifian, F.; Clyne, M. (eds), *Australian Review of Applied Linguistics, International Forum on English as an International Language*, 31(3), 30.1-30.9. https://doi.org/10.2104/aral0830.

Pischlöger, C. (2016). "Udmurt on Social Network Sites: A Comparison with the Welsh Case". Toivanen, R.; Saarikivi, J. (eds), *Linguistic Genocide Or Superdiversity? New and Old Language Diversities*. Bristol; Buffalo: Multilingual Matters, 14, 108-32.

Proshina, Z.G. Прошина З.Г. (2017). "Translingvizm i ego prikladnoe znachenie" Транслингвизм и его прикладное значение (Translingualism and its Application). *RUDN Journal of Language Education and Translingual Practices*, 14(2), 155-70.

Sallabank, J.; Marquis, Y. (2018). "'We Don't Say It Like That': Language Ownership and (De)Legitimising the New Speaker". Smith-Christmas, C.; Hornsby, M.; Moriarty, M.; Ó Murchadha, N. (eds), *New Speakers of Minority Languages: Linguistic Ideologies and Practices*. Basingstoke: Palgrave McMillan, 67-90.

Sintsov, G.; Bityutskiy, E. Синцов, Г.; Битюцкий, Е. (2019). "Rol' i posledstvija 'parada suverenitetov' 1990-1991 gg. v stanovlenii i razvitii rossijskoj gosudarstvennosti" Роль и последствия "парада суверенитетов" 1990-1991 гг. в становлении и развитии российской государственности (The Role and Consequences of the 1990-1991 "Parade of Sovereignties" in the Formation and Development of Russian Statehood). *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Obshhestvennye nauki* Известия высших учебных заведений. Поволжский регион. Общественные науки (News of Higher Educational Institutions. Volga Region. Social Sciences), 1(49), 59-66. http://doi.org/10.21685/2072-3016-2019-1-7.

Smokotin, V.M. Смокотин, В.М. (2010). *Evropeyskoe mnogoyazychie: ot gosudarstv-natsiy k mnogoyazychnoy i polikulturnoy Evrope* Европейское многоязычие: от государств-наций к многоязычию и поликультурной Европе (European Multilingualism: From Nation-states to a Multilingual and Multicultural Europe). Tomsk: Tomsk State University Press.

Toivanen, R. (2015). "Obstacles and Successes Comparing Minority Language Activism Among the Sorbs in Germany and the Sámi in Finland". Marten, H.F.; Riessler, M.; Saarikivi, J.; Toivanen, R. (eds), *Cultural and Linguistic Minorities in the Russian Federation and the European Union. Comparative Studies on Equality and Diversity*. London: Springer International Publishing, 83-105. http://doi.org/10.1007/978-3-319-10455-3_4. Series Multilingual Education 13.

Woolard, K.A. (2003). "We Don't Speak Catalan Because We Are Marginalized: Ethnic and Class Connotations of Language in Barcelona". Blot, R.K. (ed), *Language and Social Identity*. Westport (CT): Praeger Publishers, 85-103.

Woolard, K.A. (2020). "Language Ideology". Stanlaw, J. (ed.), *The International Encyclopedia of Linguistic Anthropology*. https://doi.org/10.1002/9781118786093.iela0217.

Yildiz, Y. (2004). "Critically 'Kanak': A Reimagination of German Culture German". Gardt, A.; Hüppauf, B. (eds), *Globalization and the Future of German*. Berlin: Mouton de Gruyter, 319-40.

Yusupova, A.Sh.; Nabiullina, G.A. Юсупова, А.Ш.; Набиуллина, Г.А. (2016). "Jazyk i etnicheskaja identichnost' tatarskoj diaspory v SShA" Язык и этническая идентичность татарской диаспоры в США (Language and Ethnic Identity of the Tatar Diaspora in the United States). *Vestnik Chuvashskogo universiteta* Вестник Чувашского университета (Bulletin of the Chuvash University), 2, 158-64.

Yusupova, A.Sh. et. al. (2013). "Features of the Language of Tatars Living in China". *Middle-East Journal of Scientific Research*, 17(2), 168-72.

Zemskaya, Ye.A. Земская, Е.А. (ed.) (2001). *Yazyk russkogo zarubezh'ya: Obshchiye protsessy i rechevyye portrety* Язык русского зарубежья: Общие процессы и речевые портреты (The Language of the Russian Abroad: General Processes and Speech Portraits). Moscow; Vienna: Yazyki slavyanskoy kul'tury: Wiener Slawistischer Almanach, Sonderband 53.

Rivista semestrale | **Semestral journal**
Dipartimento di Studi Linguistici
e Culturali Comparati
Università Ca' Foscari Venezia

Università
Ca'Foscari
Venezia