



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2021

Introduction - Clausal complementation in South Slavic

Sonnenhauser, Barbara ; Wiemer, Björn

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-208166>

Book Section

Published Version

Originally published at:

Sonnenhauser, Barbara; Wiemer, Björn (2021). Introduction - Clausal complementation in South Slavic. In: Wiemer, Björn; Sonnenhauser, Barbara. Clausal Complementation in South Slavic. Berlin, Boston: de Gruyter, 3-26.



Chapter I: **Introduction**

Barbara Sonnenhauser and Björn Wiemer

Clausal complementation in South Slavic: Introduction

1 South Slavic as a test case for variation

The South Slavic area forms part of different zones of convergence and divergence, with intersecting isoglosses reflecting multiple language contacts: Bulgarian, Macedonian and East Serbian dialects belong to the so-called Balkan League (the *Balkan Sprachbund*, including also Albanian, Greek, and Rumanian), while Slovene has been related to a Central European Linguistic Area, which includes also Slovak, Czech, German, and Hungarian (cf. Newerkla 2007, Kurzová 2019); in-between we have the varieties of the Serbian-Croatian (mainly Štokavian) dialect continuum which are in many respects transitional. Whereas the *Sprachbund* notion is troublesome,¹ and will therefore be avoided, the complicated overlap of isogloss bundles testifies to a complex layering of intersecting dialectal clines and convergence with non-related contact languages. This situation is probably best characterized as a contact superposition zone (Koptjevskaja-Tamm and Wälchli 2001), which implies an intersection of convergence zones resulting from multiple overlaps of different continua (or smaller clusters).

Given their dialectal differentiation and the multitude of contact situations, the South Slavic languages are a particularly interesting test case for the analysis and systematization of linguistic diversity encountered within one language subfamily. One case in point is the variability in the patterns of clausal complementation. On the basis of exemplary studies, this volume gives insights into the systematicity of this variability from different theoretical, empirical and methodological points of view.²

1 Cf. Nau (2012), Wiemer (2019a; 2021: 282) for critical discussions of ‘linguistic areas’ (with further references), Dedio, Ranacher, and Widmer (2019) for a methodological approach of how to gain evidence for linguistic areas. This approach has been applied to the Balkans – with a specific focus on Romance (within and outside the Balkans) – in Widmer, Dedio, and Sonnenhauser (2021).

2 The basic ideas that are discussed in the papers gathered in this volume were first presented at the workshop “Variation in space and time: clausal complementation in South Slavic” (University

Barbara Sonnenhauser, University of Zurich, Switzerland, Slavisches Seminar, Plattenstrasse 43, CH-8032, Zürich, e-mail: barbara.sonnenhauser@uzh.ch

Björn Wiemer, Johannes-Gutenberg-Universität (Mainz, Germany), Institut für Slavistik, Turkologie und zirkumbaltische Studien (ISTziB), Jakob-Welder-Weg 18, 55099, Mainz, e-mail: wiemberb@uni-mainz.de

The highly complex and diverse linguistic picture of South Slavic brings with it two challenges for a description of clausal complementation: 1) capturing the variation on a systematic, empirical basis that allows for meaningful comparisons (see § 2), 2) assessing the potential causes that might have been contributing to this multifaceted picture (see § 3).

2 Patterns of clausal complementation

Clausal complementation is represented by “biclausal syntactic constructions in which the predicate of one clause “entails reference to another proposition or state of affairs” (Cristofaro 2003: 95), expressed in a second clause” (Schmidtke-Bode 2014: 7). This is certainly one of the shortest definitions given in the literature on clausal complementation. It contains some notions which are neither self-explaining nor unproblematic. For instance, the opposition between bi- and monoclausal units is anything but clear-cut (and probably forms a continuum) and no less dependent on theoretical premises than the distinction between states of affairs (SoAs) and propositions. Moreover, speaking of ‘entailment of reference’ has its syntactic corollary in the distinction of ‘governed by’ (argument) vs. ‘supplementary’ (adjunct), which is equally fuzzy.

In any case, and independent of discussions about categories, clausal complementation (CC) is a particularly complex phenomenon, since it consists of many components whose interaction is, to a large extent, still poorly understood. The main components to be considered for a comprehensive understanding of this interaction are captured by the template in Figure 1.

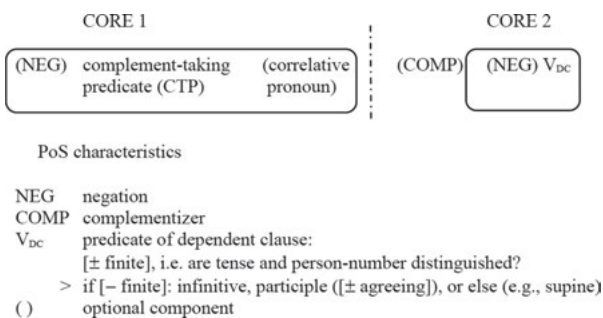


Figure 1: Maximum template of clausal complementation.

of Zurich, March 2016), funded by the Swiss National Science Foundation SNSF (grant number IZ32Z0_166419/1) and the University of Zurich *Hochschulstiftung* foundation.

This general template and the parameters included allow for a systematic, empirically based investigation of the structural conditions as well as the development and changes of CC as observed in the contemporary and historic varieties of South Slavic. It builds on the assumption that there are some necessary (or definitional) properties, which correspond to Schmidtke-Bode's definition quoted above, and some additional properties which are typical, but not necessary, or which are necessary only under specific conditions or for specific phenomena. Thus, the template also includes properties which have been neglected in research, such as the PoS-characteristics of the complement-taking predicate. Moreover, it accounts for the often pivotal role of complementizers in indicating meaning differences that correlate with meaning alternations of complement-taking predicates, while simultaneously complementizers cannot be claimed to be a necessary component of clausal complementation. These considerations do not only make zero complementizers a questionable notion, but they leave space for the study of emergent complementizers and their variation (together with other complementizing structures) along a gradient from emergence to entrenchment. This template is elaborated and commented on systematically in Wiemer (this volume).

The internal diversity of CC has grown as a consequence of the stepwise replacement of infinitives by finite structures, with the diachronic stages of this process becoming manifest in the areal differentiation of contemporary South Slavic. The areal continuum stretches from the southeast (= Balkan Slavic), where the infinitive has disappeared as a form,³ via the Štokavian (Serbian-Croatian) dialect continuum, where the infinitive as a form still exists, but the preferences for choosing infinitival or finite complements (plus an additional connective) differs (and the conditions are still poorly understood), up to the northwestern (Kajkavian) corner, i.e. mainly Slovene, where non-finite complements – including the supine – are still better attested than in the rest of South Slavic. This areal continuum corresponds to the diachronic direction of infinitive loss, which started in the southeast (cf. Joseph 1983; 2019). In general, the basic morphosyntactic isoglosses across South Slavic run on a southeast – northwest axis, although the hotbeds and directions of spread differ.⁴ This is indicative of a more intricate

³ Even if some Bulgarian varieties still show remnants of an abridged infinitive, this form is attested only in complex predicates, but not in clausal complementation proper (Wiemer, this volume: § 2.3.2).

⁴ South Slavic dialect continua reflecting inner-Slavic development have been established on the basis of phonological and morphological isoglosses smoothly running in parallel (e.g., Ivić 1958: 32). It would be intriguing to establish whether these accepted isogloss bundles at any degree correspond to clines and the development of properties that are related to clausal complementation. Since infinitives and *da*-clauses play a prominent role in clausal complementation,

overlap between the regions for which particular areal features have been discussed at least since Sandfeld (1926), and internal South Slavic dialect continua, which started forming since, by the end of the 6th century AD, Slavic-speaking population settled on the Balkans and reached the Adriatic coast up to the south-eastern branches of the Alps (cf. Holzer 2014: 1122); see further § 3.

In turn, the different degrees at which the infinitive has retreated (from change of preferences in favor of finite clauses to entire loss) are probably only the most prominent part of a general retreat of non-finite verb forms. Admittedly, Slavic languages have never frequently employed participles in clausal complementation; likewise, the supine, which is anyway preserved only in Slovene (apart from Lower Sorbian in the North Slavic area), is by its very function rather restricted to purpose clauses after verbs of directed movement.

Taken together, these considerations give an illustration why South Slavic is highly relevant for the complex question of how to differentiate the effects of language contact from those of family-internal developments, i.e. shared ancestry (cf. Widmer, Dedio, and Sonnenhauser 2021), and of universal tendencies; see further § 3.

Another well-known process which has led to an increase of internal differentiation in CC is the loss of irrealis restrictions of the central and most versatile connective element, namely *da*, in the western part of South Slavic. In areal terms, the preservation of the original irrealis function of *da* appears to be complementary to infinitive loss, as Balkan Slavic is not only most conservative in retaining the irrealis function of *da*, but also more innovative than the western part of South Slavic in establishing *da* as an ubiquitous connective on practically all levels of constituency (cf. Wiemer 2018: 295–306, Sonnenhauser and Widmer 2020). This enormous spread has been concomitant to the restructuring of strict rules of clitic placement typical for all South Slavic languages. However, only in Balkan Slavic has *da* consistently been integrated into verb-oriented clusters of proclitics, by which it has participated in the partial morphologization of clitics (Spencer and Luís 2012: 59–65, 124–126, 153–157). Remarkably, one may even assume that by this very process the evolution of *da* as a complementizer has been precluded (cf. Wiemer, this volume: § 2.2.3, for a justification). By contrast, in the middle part of South Slavic (the Štokavian area) *da* shows a split of distributional patterns concerning not only its behavior as a clitic (vs. clause-initial connective),

it is plausible to assume that some such correlation should exist. However, in view of the lack of reliable data required for a sensible diachronic interpretation of contemporary isoglosses and clines such an endeavor would not be an easy task. Apart from that, there are a lot of understudied phenomena related to clausal complementation which until now have not made it into focal topics of research in South Slavic (or Balkan linguistics).

but also constraints (vs. their lack) on admissible TAM forms. In Slovene, in turn, even these constraints appear to be loosened (Wiemer, this volume: § 2.2.3, for a comprehensive treatment). In contrast to the rather “unidirectional” areal cline, it is less clear where the loss of irrealis restrictions of *da* began; the most probable “candidates” are the northwestern end (early Slovene, Kajkavian) and the Adriatic coast (Čakavian); cf. Grković-Major (2004; 2020) and Greenberg (this volume). In either case, language contact (with German, Vulgar Latin or Italian, respectively) can be assumed to have been a catalyst of this process.

Apart from these most prominent examples, an increase of internal South Slavic differentiation has ensued from other phenomena. The replacement of the old subjunctive (or conditional), based on *bi* (< Common Slavic *byti* ‘be’), by *da*-constructions and their interaction with the complex tense-aspect system in Balkan Slavic has yielded new preferred manners of expressing hypothetical or otherwise irreal situations (e.g., the Macedonian future/expectative marker *ќе* in combination with the imperfect) that are encountered in all kinds of complex sentences. This contrasts with the continued interaction of the old conditional marker with clause-initial connectives in the northwestern corner of South Slavic (compare *da+bi* in Slovene), which reminds us of North Slavic (Topolińska 2003). Another phenomenon to be considered is the different distribution and spread of specialized factive complementizers, i.e. of complementizers which introduce logically presupposed propositions (compare *što* in Serbian-Croatian and Macedonian). Such complementizers are intimately connected to causal conjunctions, both in their individual histories (Grković-Major, this volume) and in micro-areal patterns due to language contact (Breu, this volume), although not all causal connectives, while evolving into complementizers, keep a factivity requirement; for instance, Bulg. *če* most probably originated from a causal connective (Sonnenhauser 2015: 46), but it is now predominantly used as neutral (or default) complementizer. Furthermore, even though they have remained one of the most understudied domains in complementation in South Slavic, contrasts of factual vs. non-factual complements⁵ after predicates with propositional complements (i.e. those related to knowledge or belief, or related to speech acts) are particularly interesting, since patterns vary not only between most closely related languages (e.g., Bulgarian and Macedonian), but also for the means chosen and the

5 This contrast deals with the strength of epistemic support (for propositional complements) and the distinction between SoA (= non-factual, e.g. purpose or embedded deontic clauses) and propositional complements (as objects of epistemic assessment). Factuality must not be confused with factivity; the latter is here taken to refer only to complements with logically presupposed propositional content (e.g., *He regrets / doesn't regret that P*, P is presupposed). For clarification cf. Wiemer (this volume: § 2.2.2).

verbs (or other predicative units) that allow for a meaningful choice of clausal connectives. For instance, while Bulgarian and Macedonian basically share a possibility to distinguish between strong and weak epistemic support for the content of propositional complements, namely by choosing *da* (weak support) vs. Bulg. *če*/Mac. *deka* (strong support), the particular complement-taking predicates (CTPs) which favor (or disfavor) *da* in case of weak support can differ. In addition, one has to consider the relation of these connectives to Mac. *kako*/Bulg. *kak*, which can introduce either SoA or propositional complements, and the impact of negation. Negation generally weakens epistemic support for those CTPs which otherwise indicate strong support. Since complements are entailed by CTPs (see above), changes in the degree of epistemic support also have an impact on the assessment of the content of propositional complements of belief-related CTPs. This explains phenomena known as negative concord or NEG-raising (e.g., *She believes that P* vs. *She doesn't believe that P* \cong *She believes that not-P*).⁶ As a rule, differences of complementizer choice between languages (or varieties on a lower taxonomical level) become seizable not just by contrasting singular examples, but as a stochastic tendency (whose functional underpinnings need to be explained). This is shown by Mitkovska and Bužarovska's (this volume) analysis of variation among complements of verbs of visual perception, with their extensions into epistemic uses, in Macedonian and Bulgarian. Therefore, usage-based accounts (usually with corpus data) are indispensable to grasp differences that exist even between closely related languages (or varieties). They are also indispensable for "empirical checks" of assumptions made, for instance, in formal semantics, such as Giannakidou and Mari's (2016) proposal of (non)veridicality in the assertion or presupposition of the CTP to influence the choice of subjunctive or indicative complements and the different parametrization even among closely related languages. This, of course, is true not only for marking epistemic distance, but also in all concerns whenever variation matters. Such data-driven approaches in turn require a reasonable amount of data – which is still a challenge for South Slavic varieties – edited in a maximally theory-neutral way such that diatopic and diachronic comparison becomes possible (cf. Eckhoff, this volume).

From an areal point of view, the certainly most salient (though still little known) division in the choice of clause connectives for marking propositional complements against which the speaker distances themselves is Sln. *naj* (< **nehati* 'let') in relation to its cognate *neka* in the remainder of South Slavic (and *n(i)ech* in West Slavic).

⁶ Cf. Wiemer (this volume: § 2.1, § 2.2.4). For analogous findings in North Slavic cf. Dobrushina (2012, Russian), Wiemer (2015, Polish), Hansen, Letuchiy, and Błaszczyk (2016, Russian, Polish). For a general discussion cf. Kehayov and Boye (2016: 818–824).

While Sln. *naj* shows properties of an emerging complementizer which suspends propositional content (Sonnenhauser, this volume), its South Slavic cognates can hardly be treated as complementizers, although they also are used in contexts with suspended propositions. Moreover, the epistemic functions of these complementizers are probably diachronically secondary, as, in all likelihood, they derive from uses in directive speech acts, in deontic contexts and in purpose clauses. Therefore, a comparative analysis of Sln. *naj* and its South (and West) Slavic cognates promises to reveal reliable insights into the developmental paths from causative verbs ('let, allow', etc.) to clausal connectives, up to complementizers, and their diachronic relation to auxiliaries. This comparison should be extended to other causative constructions, among them those based on *da* + present indicative (on which see above).

The relation between clausal connectives and auxiliaries brings us to the question of how clausal complementation relates to mood marking, in particular where both meet (see below). By a similar token, the investigation of *naj* and its cognates vis-à-vis the diachronic development of *da* as an ubiquitous connective of clauses and within complex predicates would probably yield insights allowing for more reliable generalizations concerning the relation between epistemic distancing, on the one hand, and directive speech acts and deontic modality, on the other. In the case of both *da* and *naj*, as indicated above, their employment to mark suspended propositions (epistemic distancing, weak epistemic support or indirect evidentiality) obviously only followed on more original employment in more "interactive" contexts of speech (and these more original usage types are still encountered frequently). This looks like another instance of subjectification (as introduced by Traugott 1989 and subsequent work); however, the tight entrenchment with syntax experienced by *da* and *naj* on different levels of constituency (and with different ranges of variability) as well as their obviously complementary areal distribution and different age provide us with an ideal starting point for the study of repeated similar (though not identical) developmental paths in the confines of one language subfamily. Moreover, *da* and *naj* show that, although uses in propositional complements appear to have emerged only at later stages, this development did not preclude their employment as complementizers, i.e. in tighter structures of clause combining. That is, we observe syntactic tightening in parallel to widening of semantic scope (from directive, or goal-directed, speech acts to subjective judgments regarding propositional content). This is remarkable inasmuch as widening of semantic scope is concomitant to, or even implied by, the Traugottian notion of subjectification, while simultaneously it seems to be unfavorable for the tightening of syntactic scope. This seeming contradiction has been the reason why the scope criterion has been neglected in grammaticalization research, to an extent that it recently has been excluded from an attempt at quantifying effects of change associated with grammaticalization in Bisang et al.

(2020). Obviously, there still is a principled point to be made for the relationship between syntactic and semantic scope (and associated changes), for which the study of clausal complementation can provide fruitful theoretical insights, and we find a particularly rich empirical basis for this kind of study in South Slavic.

However one may decide on the chronology and probable connections between the issues mentioned so far, we recognize that South Slavic, in particular its eastern part, supplies a playground for research in the relation between clausal complementation and non-indicative mood, provided the latter is conceived of as a clause-level distinction (not as a distinction of verbal morphology). Simultaneously, this issue is but a manifestation of a broader phenomenon characteristic especially of Balkan Slavic, namely the rise of a systematic architecture of periphrastic verb forms, based on the *l*-participle⁷ and various auxiliary elements (of which *da* is but one). These periphrastic forms have been bones of contention when it comes to delimiting from each other morphology and syntax. On the one hand, these forms can be shown to take part in the formation of complex paradigms; as such they are usually presented in standard grammars and handbooks on Balkan Slavic verb morphology, but see also analyses with a more far-reaching theoretical appeal as the one in Spencer (2003). On the other hand, since these complex forms are used productively and the composition of their parts (auxiliary elements + lexical verb) is to a large degree compositional, claims to treat them as part of syntax cannot be dismissed *a priori*; cf. Spencer and Popova (2015: 214–227) and Sims and Joseph (2019) for elaborate discussion. Again, whatever side one may take, the architecture of the morphology – syntax interface of South Slavic, in particular Balkan Slavic, languages is extraordinarily rich in “objects” to be considered with scrutiny, also for the relevant diachronic facts, and valuable for cross-linguistic comparison. Moreover, these objects intersect with phenomena belonging to clausal complementation, namely with those that demonstrate a tighter conceptual connection (or even dependence) between the involved predicative nuclei (as on the high-integration end of semantic integration scales; cf. Wiemer, this volume: § 2.1).

The selected examples discussed in this section surely show that South Slavic is extraordinarily rich in internally differentiated patterns of CC. Obviously, a thorough and methodologically well-considered examination of the different patterns, both in terms of structure and usage, can cast light on many issues that have been hotly debated in more general research on CC beyond this particular subgroup of languages (see § 5).

7 The *l*-participle originally marked the lexical verb in the periphrastic perfect (including the pluperfect and the future perfect). It still is used in this way for periphrastic tense forms, but only in Balkan Slavic has it preserved its original perfect functions, while in the remainder (again along a southeast – northwest cline) we observe a shift from perfect to general past. For the general background cf. Meermann and Sonnenhauser (2016), Arkadie vand Wiemer (2020).

3 Common heritage, areality, universal tendencies

Since here we are emphasizing the internal differentiation of South Slavic languages and already have pointed at some diachronic tendencies applying to this subfamily in general or to some part of it, we can hardly avoid raising the issue of what has caused the changes leading to this differentiation. Concomitantly, one may ask how these changes, and the variation behind them, can be systematized and which methods and tools should be employed to approach them.

The first issue, about factors and mechanisms of change, practically amounts to an equilibration of factors that favor (or impede) change and convergence between genealogically unrelated languages. Such an equilibration can be conceived of as methodological triangulation, inasmuch as three groups of factors and angles have been separated in areal typology and related disciplines, namely: (i) provenance from common ancestor languages, (ii) areal neighborhood (usually leading to language contact), and (iii) general cognitive and communicative principles, which have been considered responsible for Greenbergian markedness relations (Croft 2009) and which are often referred to as ‘universal tendencies’ (Heine and Kuteva 2005). Triangulation in this sense has been suggested and, to some extent, elaborated on in Wiemer, Seržant, and Erker (2014) and Wiemer (2019a), with respect to clausal complementation in Wiemer (2018).⁸

A major problem with triangulating phenomena related to CC arises from the fact that inner-South Slavic synchronic clines and isoglosses often cannot be interpreted diachronically in a straightforward manner. That is, it is particularly difficult to put particular phenomena in chronological order, all the more as usually these phenomena are gradable (in terms of spread on a diatopic or diastratic axis, in terms of variation with concurrent patterns, and probably in other regards). Moreover, change, as a rule, starts from minor or covert patterns, so that it is desirable to estimate increase in frequency and loss of constraints. Furthermore, we would like to know whether the spread of particular phenomena relates to non-linguistic factors, such as socio-cultural circumstances, e.g. more remarkable migration waves. In addition, we have to assess the likelihood that changes might not have occurred spontaneously, i.e. without some more remarkable “external” trigger, exactly because variation with minor patterns may always be assumed to exist. This applies in particular to spontaneous discourse in non-standardized varieties. For instance, consider the prominent example of the infinitive and its

⁸ Ranacher et al. (2021) propose a Bayesian clustering algorithm to detect geographic contact areas in language data in the presence of inheritance and universal preference as confounding effects. One of their case studies is the Balkans.

gradual replacement by finite structures, especially in complementation and other subordinate clauses: in order to assess the role of contact (against “internally” motivated change) we, strictly speaking, need to know how widespread finite structures have been prior to their spread in disfavor of the infinitive in the speech communities under consideration; and we need to know more about the overall distribution of these concurrent patterns of clause structures in a much larger, surrounding area.

This said, i.e. apart from the aforementioned three “angles”, we must not neglect the impact exerted by meta-linguistic factors. Such factors usually become more palpable not in a diatopic, but in a diastratic or diaphasic dimension (in Coseriu’s 1988 terms). They surface particularly in standardisation, accompanied by the influence of grammar writing following prestigious models. Prescriptive efforts aiming at differentiation at the expenses of variation or decisions motivated by considerations of language policy may have led to a differentiation on the level of the standard languages that is not necessarily observed for the spoken (urban and dialectal) varieties. In the Balkan Slavic context, this can be seen, e.g., on the discussions concerning the usage of the ‘long’ and ‘short’ article for masculine singular nouns (Fielder 2019). Another relevant factor is the dialectal basis chosen for the standard language. Consider, for instance, the Slovene relative pronoun *kdor*, which appears in the contemporary standard language as specialised marker for free relative clauses, but is unknown in the eastern dialects (Sonnenhauser 2019). This consideration makes the investigation of non-standard and historical data an even more urgent requirement. To our knowledge, clausal complementation has so far hardly been studied from this angle in South Slavic.

Meta-linguistic factors often have to do with dialect levelling under the influence of some roof variety, but also more fundamentally with speech habits guided by the awareness of implicit norms.⁹ More particularly, linguistic behavior can be tailored to the communicative needs of multilingual (or “multivarietal”) speaker communities which are highly heterogeneous as for individual speakers’ competence in the particular languages involved in multilateral contact. Such situations create particular communicative pressure to make one’s intention clear. On this background, Topolińska (2008 [1994]) even ventured a hypothesis according to which some of the so-called Balkanisms are more or less direct outcomes of communicative requirements posed by intensive multilateral contact. According to this hypothesis, these conditions favored the rise of means which enable speakers to ease disambiguation of reference or to mark the reality status

⁹ Cf. Wiemer (2021) for a survey of motives leading to structural convergence and research seeking to disclose these motives.

of denoted events. It concerns, first of all, such prominent Balkanisms like the definite article, clitic doubling and the evidential extensions of the perfect, but – provided such a hypothesis can be tested – one may well add choices of complementation devices triggered by realis vs. irrealis contrasts (see § 2). Convergence as for means that are sensitive to such functionally motivated choices never arise out of the blue; instead they must have existed at least as minor choices in an “inventory” of variants for the expression of particular communicative means, before they started being preferred for one reason or other. This reason can be the largest common denominator in a multilingual speech community, but the other two angles mentioned above (genealogical heritage and general tendencies) should not be dismissed. In fact, these different motives can work in conspiracy, so that it is often difficult or impossible to disentangle them.

The situation on the Balkans is additionally complicated by the fact that, despite Ottoman Turkic hegemony for many centuries, none of the “autochthonous” Balkan languages dominated over the others. There was, thus, no process of creolization (in contrast to so-called New Englishes; cf. Trudgill 2006, Ziegeler 2017). Rather, what we observe amounts to metatypy, i.e. “the wholesale restructuring of a language’s semantic and syntactic structures as a result of language contact” (Heine and Kuteva 2005: 180).¹⁰ Metatypy leads to isomorphism, which favors intertranslatability (cf. Gołąb 1990, Lindstedt 2000, among many others). It relies on the calquing of grammatical contrasts (PAT-borrowing) and often involves massive loan translations (including polysemy copying), but no re-lexification or MAT-borrowing.¹¹ However, below the surface, substantial differences in detail may appear concerning the extent to which these general structures apply and the specific morphosyntactic means employed, relating, among others, to specific language-internal conditions or different starting points of converging developments (Sonnenhauser and Widmer 2020, Joseph 2001). As concerns South Slavic in its entirety, the history of its varieties abounds in changes of political and economic authorities and of concomitant borders of states, or other administrative units. This brought about changes in non-Slavic and Slavic roof varieties, sometimes cutting across “indigenous” territories in the South Slavic dialect continuum. A good example in case is the Torlak variety, which was divided by the political boundaries established at the end of the 19th century and has since then been under the influence of different standard languages (Balkan Slavic Bulgarian/Macedonian and West South Slavic Serbian); cf. Vuković (2000; 2021). Such

¹⁰ The term was originally coined by Ross (1996; 2007, among other publications) to characterize complex contact situations (and their results) in New Guinea.

¹¹ On the distinction between PAT- and MAT-borrowing cf. Matras (2007) and Sakel (2007). Polysemy copying in South Slavic clausal complementation is analyzed by Breu (this volume).

an administrative-political “to and fro” also conditioned a rather late onset of standardization processes.

Thus, enormous structural diversity has been caused to a large degree by changing ethnographic and political circumstances during the last 1,500 years, which have left traces on formal and functional differentiation in patterns and the variability of CC in South Slavic languages. This differentiation has remained understudied, since, except for Balkanisms and the “*da-split*” caused by loss of irrealis features beyond Balkan Slavic (see § 2), phenomena related to CC have been treated rather promiscuously and unsystematically, if not for particular languages, so at least on a general South Slavic background, let alone in comparison to North Slavic. For instance, only recently have phenomena related to the “*da-split*” been considered another Balkanism by Topolińska (1997) and Ammann and van der Auwera (2004). As explained above, this split provides a neat illustration of how Balkan features intersect with the South Slavic dialect continuum (reflected in the standard languages), but it also shows the need for scrutiny in its diachronic assessment, not only on a South Slavic, but a general Slavic backdrop (cf. Wiemer 2017: 325–330; 2018: 293–306; 2019b: 120–127). To sum up, broader typological comparisons have been drawn only with respect to features considered Balkanisms. Apart from that, so far no serious attempts have been undertaken in triangulating the motives of change and the resulting differentiation on a diatopic, diastratic and a diaphasic level, and to relate them to meta-linguistic factors.

We thereby have come to the issue of systematisation. The high diversity of structures relevant for CC in South Slavic has remained a weak point in research, and this unsatisfying state of the arts is accompanied by problems in data collection and processing, predominantly in corpora. It is in general very difficult to find a sufficiently unified methodology and to apply a reasonably defined set of comparable categorial distinctions. One promising approach is a strictly decompositional one, as applied in Šimko (2020) for the development of definiteness marking in 17th to 19th century Damaskini texts. This point also raises the question of how corpora should be annotated (Eckhoff, this volume).

4 The contributions to this volume

The contributions to this volume cover the entire South Slavic territory. Most contributions focus on particular aspects of complementation, a large part deals with the contemporary standard languages Slovene, Serbian, Bulgarian, and Macedonian, while others focus on older stages and/or non-standard varieties and the impact of language contact, primarily with non-Slavic languages. They

all present in-depth studies and thus contribute to an overarching collective aim, which consists in a comprehensive picture of the patterns of CC on which South Slavic languages profile against a wider typological background, but also diverge internally if we look closer at details in the contemporary stage and in diachronic development. We thus hope to set some standards in methodological approaches to the empirical investigation of CC not only in South Slavic, but also beyond this subfamily. The general goals to be envisaged can be divided into the following three issues:

- (a) Determining the components and parameters which distinguish different patterns of CC and characterize their particular components. This implies a critical evaluation and treatment of notions involved in the analysis of CC.
- (b) Elaborating on the factors which have led to particular CC patterns in sub-areas of South Slavic and which thus have been contributing to the complex picture of clausal complementation in this subfamily.
- (c) Problems that need to be met when it comes to the classification of data on the basis of corpora and other resources.

The individual papers assembled in this volume make profound contributions to many of these issues by examining different parts of the template given in Figure 1. By taking this template as a common point of reference, the articles help establish a structured inventory of forms and functions of CC in South Slavic from a diachronic and diatopic perspective. The papers complement each other in their sources of data, which is partially explained by their different focus on diachronic or contemporary varieties, and on language contact or on accounts of single languages.

Apart from this introduction, the volume is divided into three parts whose contributions primarily focus, respectively, on the typology and parameters of CC structures (Part II), on the dispersion of CC patterns in geographical space (Part III), and on the dispersion of CC patterns in diachrony or in earlier stages (Part IV).

Part II starts with the article by **Björn Wiemer**. It is meant to set the stage for a comprehensive treatment of clausal complementation in South Slavic. It develops a general template for CC constructions reproduced above as Figure 1 and distinguishes its separate components. Since capturing CC implies a couple of notions which are complex themselves, these notions are surveyed first, before the different components of the template are assessed one after another on the basis of South Slavic material. Although the focus here is on the contemporary stage, some “excurses” into diachrony are inevitable, in particular as concerns the infinitive and other non-finite forms and the spread of *da*-constructions. As for the latter, some considerable part of the survey deals with the relation between complemen-

tation devices indicating a non-factual (= irrealis) status of the complements and non-indicative mood (in particular the so-called ‘analytical subjunctives’). The analysis includes a critical assessment of the state of the arts in this domain for South Slavic. It shows why clause-initial connectives and verb-oriented proclitics (united under the label of ‘complementation devices’)¹² fulfil a central role in CC especially in South Slavic and why, consequently, their differentiated treatment is of particular importance. This includes special attention to emergent patterns and the issue why some such patterns make it into entrenched structures (of the respective language), while others do not. Apart from metatheoretical discussion, the comprehensive data-driven step-by-step analysis of the components of the CC template is meant as a proposal of how to provide a consistent and systematic analysis of CC, for which the empirical situation in modern South Slavic is only a convenient starting point, but which might be applied further in broader cross-linguistic comparison.

Alexander Letuchiy presents a corpus-based case study of complement clauses with nominal attachment sites in Bulgarian. The central question behind this study is why the distribution of clausal complements with nominal heads differs from the distribution of the same type of complements after cognate verbs. Letuchiy considers a variety of syntactic and semantic factors, both in the complement and the complement-taking predicate. He observes that Grimshaw’s (1990) traditional account of action nouns based on syntactic and actional classes of verbs is insufficient to explain the encountered distributional biases in Bulgarian corpus data. His main conclusion is that, among a battery of heterogeneous factors, the main parameter appears to be the opposition of situation proper vs. occurrence, which is close to, but not identical with the distinction between state of affairs (SoA) and proposition (which has been pronounced, first and foremost, in Boye and Kehayov 2016). This parameter also closely interacts with the known parameter of actional classes, but differences in semantic role relations prove important as well. In general, Letuchiy’s analysis strongly supports the claim for multifactorial approaches to explaining different (dis)preferences and types of clausal complements with verbal vs. nominal attachment sites.

Iliyana Krapova analyzes the Bulgarian declarative complementizer *če* ‘that’ and the interrogative complementizers *dali* and *li* ‘if/whether’. She argues in favor of a Split CP approach, according to which the left periphery of embedded clauses (in particular of complement clauses) can provide slots for more than a single complementizer (C position). She thereby develops arguments of Rizzi (1997, and subsequent publications) in favor of an ordering hierarchy of

¹² The term is here adopted from Boye and Kehayov (2016).

complementizers within the same clause. In particular, Krapova finds evidence for a stable ordering of *dali/li* (COMP1) preceding *če* (COMP2); for the former she proposes an additional projection (VeridP) which takes the latter (FinP) into its scope. She argues that such an account is more suitable for explaining both this order and the interpretation of the involved connectives in view of some yet badly understood interaction between complementizers and Topic or Focus phrases. Concomitantly, Krapova discusses evidence in support of regarding *da* as a “modal particle” situated lower in the constituency tree and, thus, in a functional domain closer to the predicate of the dependent clause.

Liljana Mitkovska and **Eleni Bužarovska** inquire into the variation of complementation patterns for the basic verbs of visual perception in contemporary Standard Macedonian and Standard Bulgarian (Mac. *gleda/vidi*, Bulg. *gledam, viždam/vidja*). For this purpose they submit large representative samples to an in-depth analysis. Following Dik and Hengeveld (1991), the authors propose a threefold division capturing the relation between visual perception and knowledge states: immediate perception, mental perception triggered by visual experience (called ‘primary mental perception’), and knowledge obtained by inference from direct physical evidence (called ‘secondary mental perception’). These semantic distinctions are mapped onto the employment of Mac. *kako, deka, da* and Bulg. *kak, če, da* as complementation devices. The authors find that, in general, the threefold semantic division matches with the opposition between *kak(o)* and *deka/če*, but they also disclose a broader distribution for Mac. *kako* (vs. *deka*) in comparison to Bulg. *kak* (vs. *če*). This is interpreted as indicative of primary mental perception occupying an intermediary position between the two other conceptual configurations, but also as evidence that in Bulgarian the opposition between immediate and mental perception tends to be blurred. As for *da*, the data shows that in both languages this marker is used consistently not only after negated verbs of visual perception, but also in certain expressive contexts. However, unlike in Bulgarian, where *da* is frequently employed for immediate perception, in Macedonian its occurrence is restricted to mental perception.

The two contributions of Part III are united by their focus on locally restricted non-standard varieties and the impact of contact, mainly with non-Slavic varieties. Based on diverse textual sources and grammars available for different periods, **Marc L. Greenberg** investigates the evolution of the opposition between *da* and *ka* used as complementizers in Prekmurje Slovene, a non-standard variety of Slovene at the northern periphery of South Slavic which for a long time has been, and partially still is, under Hungarian rule. What strikes the eye is that, by the early 20th century, in Prekmurje Slovene *da* was used as an irrealis connective (vs. *ka* for realis contexts), which set it apart from the rest of western South Slavic (including the remaining Slovene varieties) and brought it close to Balkan

Slavic. More specifically, “there are at least two distinctions operating with the *ka:da* contrast, one opposing real vs. irreal and another opposing description vs. emotion/evaluation”; this very much resembles the situation in contemporary Bulgarian (see above). Nonetheless, the author concludes that the origin of the irrealis feature of *da* in Prekmurje Slovene must have been independent from Balkan Slavic, first of all because the loss of irrealis function for *da* must have reached the northern periphery of South Slavic already by the 11th century. Moreover, the later opposition between *da+bi* (IRR+conditional) vs. *ka* broke down in the second half of the 20th century in favor of *ka*, which thereby became the default complementizer. Given the fact that the remainder of Slovene has generalized *da* as default complementizer, the question arises why Prekmurje Slovene went another way, even if contact with central Slovene had not been interrupted. The author suggests that linguistic identity, i.e. the desire to distinguish it from Standard Slovene, may have affected the choice among alternatives in usage.

The contribution by **Walter Breu** presents a consistent analysis of complementizers and complementation structures in South Slavic (Molise Slavic, Resian) and Albanian (Italo-Albanian) minority dialects in Italy. It thus has a dedicated focus on recent and ongoing language contact outside the “indigenous” South Slavic and Albanian territory, with a due account of the diachronic backdrop; it thereby makes a pronounced contribution to methodological triangulation (see § 3). The main body of data derives from personal fieldwork, but data from other sources are accounted for as well, first of all since the author provides a systematic comparison with Standard Croatian (and other varieties of the BCMS-group), Standard Slovene, Standard Albanian and Italian (standard and dialects). Clausal complementation is analysed in connection with relative and causal clauses (and the interrogative pronoun ‘what’), as these are well-attested sources of complement clauses. In particular, the Italian connective *che* functions in all of these domains. The author shows this connective to have been the basis of different cases of PAT-borrowing, or of polysemy copying, in the aforementioned minority languages, although in every single variety this process has yielded slightly different results. Moreover, Molise Slavic and Resian have PAT-borrowed the Italian pattern of opposing factual and non-factual (more precisely: intentional) complements. Only in Molise have the Slavic and Albanian varieties MAT-borrowed It. *che*, however with different functional range. The consistent comparison of diverse replica varieties and their Italian model varieties brings to light also some implicational relationships. For instance, if the relativizer is borrowed, it also functions as a complementizer, but not vice versa.

Part IV unites articles with a focus on historical varieties and/or diachronic change. **Hanne Eckhoff** makes a corpus-driven contribution to the study of changes in dominant complementation patterns in Old Church Slavonic (OCS,

a South Slavic variety) compared to the earliest attested stages of East Slavic and of Middle Russian. Simultaneously, she contributes to an assessment of the explanatory potential of electronic diachronic corpora and of methodological pitfalls hidden in probably any kind of syntactically annotated corpus. The corpora used belong to the PROIEL and TOROT treebanks (see <https://proiel.github.io/>, Haug and Jøhndal 2008, Eckhoff and Berdicevskis 2015). The author addresses the problem of ambiguous dependency relations: how can clausal complementation be identified if the syntactic relation between a clause-initial connective and a preceding predicative expression (as a potential complement-taking predicate) and a verb in its own clause is not clear? A particular case in point is OCS *jako*, among other units which can be variably classified as “particle”, interrogative pronoun, adverb, or else. Eckhoff performs two case studies. The first is carried out on texts belonging to the canonical body of OCS, for which interannotator agreement and retrievability are shown to work well, provided one knows how to make use of annotation conventions. The second examines the history of East Slavic *čbto* ‘what, which, that’ and illustrates how one can tackle with a clause-combining element in the course of diachronic change.

Jasmina Grković-Major provides a comprehensive survey of complementation patterns of emotion predicates in the history of Serbian, with a particular concern for disentangling internal and external factors of change. She thereby makes a dedicated contribution to methodological triangulation (see § 3). Her considerations are embedded into the broader issue of how transitivity and configurationality got on shape in ancient stages of Indo-European languages and how the continued strengthening of configurational syntax became manifest in the history of Serbian (on a general Slavic background). Special attention is paid to complements of factive predicates, since there is a tight conceptual relation between presupposed propositions and their emotional evaluation, together with a link to causality. This explains the prominent role of causal conjunctions as sources of factive complementizers (e.g., Serb. *jer(e)*, *što*). Concomitantly, the author emphasizes the significance of different waves of innovation which overlapped in the Serbian territory and which led to its transitional properties between typical Balkan features (*da*-clauses as non-factual complements) and innovations which spread from the west and southwest (a factive – non-factive split in the complementizer system), together with a continued spread of realis uses of *da*-clauses at the expense of *što*-clauses (as specialized markers of factivity).

Barbara Sonnenhauser’s study on Sln. *naj* is devoted to a chapter in the most recent history of Slovene. The highly versatile element *naj* can function as an auxiliary in complex predicates denoting hortative or permissive speech acts, and also as a propositional particle marking epistemic distance and reportive evidentiality. In addition, it may function as a subordinative conjunction intro-

ducing purpose clauses, and it has been analysed as behaving like an emergent complementizer. Based both on systematic corpus queries and a perusal of contemporary and historical grammars of Slovene, Sonnenhauser traces the flexible semantic and syntactic behaviour of *naj* in order to eventually assess the degrees to which *naj* displays functions that can be considered typical of complementizers. Thereby, the paper also touches upon the question as to how such functions can be identified without ready-made categorical notions and distinctions that are in themselves problematic (see § 2). It proposes to start from the most general function complementizer-like elements assume, i.e. the marking of a particular element or structure as part of a larger syntactic unit, and decomposes it into smaller-scale features. The latter include an element's potential of building up hierarchical predicational structures, its ability to fill in a valency slot of a verbal or nominal structure, and its not being available as a target of agreement. While an elaboration of this proposal remains a task for the future, the results of the empirical analysis suggest that, while sharing its functional origin with its cognates in other South Slavic (and in West Slavic) languages, *naj* has advanced considerably further toward becoming a clausal subordinator (see § 2). In this way, the paper contributes to uncovering in more detail the variety of strategies of clausal complementation encountered in the South Slavic languages and the diversity in the underlying diachronic processes.

5 Relevance beyond Slavic

As suggested by our state-of-the-arts report and the survey of contributions, this volume not only provides insights for Slavic linguistics, but also contributes in at least three respects to more general debates in linguistics: (i) formal and functional aspects of clause linkage, (ii) factors involved in language change and areal clustering, and (iii) questions related to data processing.

(i) Clause linkage

One very general problem underlying theoretical approaches to the diversity of South Slavic CC, which is shared by virtually all types of analysis, consists in the fact that CC is not only a multifaceted phenomenon (Kehayov and Boye 2016), but in addition a still not very well-defined notion. As a consequence, research on CC makes use of diverse concepts and covers quite different structures, forms and functions. Among the main issues to be clarified are the semantic and syntactic features of CTPs, the morphosyntactic status and provenance of complementizers, the relation between matrix and dependent clause, the morphosyntactic shape of clausal complements, and the properties of the predicate in the depend-

ent clause. These questions are closely related to the delimitation of complementation structures from other types of clause linkage (such as relative or adverbial clauses) and more integrated structures (such as control and raising structures, or complex predicates). Such issues are primarily addressed in Wiemer (this volume) in connection with a general template of clausal complementation, which is also meant as a basis of crosslinguistic comparison.

(ii) Language change

The synchronic variation found among complementation structures is the result of manifold diachronic processes. As far as we can judge from the earliest written sources, the development started off from a restricted set of explicit connectives between clauses, and neither these nor distinctions of mood or finite – non-finite contrasts mapped onto more clear-cut functional distinctions along a coordination (or juxtaposition) – subordination cline; in particular, the status of clause connectives concerning this cline was diffuse. An increase in the formal inventory and the functional possibilities can be observed as the written documents get more and more diversified in terms of linguistic varieties and in terms of functional styles, with the emergence of vernacular patterns in written media playing an important role. One case in point is the employment of *če* alongside *kako* in 17th–18th century Balkan Slavic texts, which can be related to the development of literary styles (Sonnenhauser 2015). Similarly, emerging norms spreading with influential and widespread media might suggest a diachronic development that actually is a genre-based diffusion and stabilization of one specific option (cf. Sonnenhauser 2019 for the Slovene relative pronoun *kdor*). Therefore, in addition to triangulation (see § 3) an evaluation of the paths of development has to take into account standard and non-standard data and consider the possible impact of genres, registers and styles. For all these aspects, different kinds of contact scenarios need to be considered: inner-Slavic and areal, oral and literary contact as well as contact mediated by translation. Sketching this interaction of internal development, external contact and meta-linguistic influences is a challenge not only for an analysis of the formal and functional development of CC patterns, but for accounts of language change in general.

(iii) Corpus linguistics

The diversity encountered among complementation structures and the frequent polyfunctionality (or heterosemy) of complementizers is a challenge for the collection of data and their processing in corpora, in particular as concerns the annotation of (potential) complementizers and complement taking predicates, and the mapping of form and function (cf. Haug and Eckhoff 2011 and Eckhoff, this volume, for a discussion of Old Church Slavonic and early East Slavic). The development of a compatible format for approaching the data and the application of tools adequate for the investigation of CC tie up with current debates in

corpus linguistics, in particular on annotation schemas for non-standard and transitional varieties (cf., e.g., Vuković 2021 for Torlak, Vuković et al. 2019 for the development of an annotation scheme encompassing diachronic and diatopic transitional Balkan and South Slavic varieties) as well as in typology, in particular as concerns the status of linguistic categories as comparative or language-specific concepts (cf. Haspelmath 2010; 2015).

(iv) Additional topics

Including papers that discuss the degrees of clause integration and the position of the clause linking elements in the syntactic structure from a synchronic and diachronic point of view, this volume also contributes to topics in general syntactic theory such as cyclical change (van Gelderen 2009; cf. Sonnenhauser, this volume) or the structure of the left periphery and the CP domain (Rizzi 1997) and its relation to semantics, information structure and discourse (as depicted in the cartographic approach, e.g., by Shlonsky 2010); cf. Krapova (this volume). To the extent that contributions in this volume also discuss the semantics of CTPs and its impact on the choice of complementation formats, first of all on variation among connectives, the topics raised also relate to lexical typology, which is concerned with the interrelation of lexicon and grammar within a particular lexical field (cf. Koptjevskaja-Tamm 2012); cf. Mitkovska and Bužarovska (this volume), also Wiemer (this volume: § 2.1). From a diachronic perspective, here also belongs the development of complementizer functions exhibited by function words (compare all kinds of loosely defined particles) and other lexical elements such as adverbs, pronouns and other *wh*-elements.

References

- Ammann, Andreas & Johan van der Auwera. 2004. Complementizer-headed main clauses for volitional moods in the languages of south-eastern Europe. In Olga Mišeska Tomić (ed.), *Balkan Syntax and Semantics*, 293–314. Amsterdam & Philadelphia: Benjamins.
- Arkadijev, Peter & Björn Wiemer. 2020. Perfects in Baltic and Slavic. In Robert Crellin & Thomas Jügel (eds.), *Perfects in Indo-European Languages and Beyond*, 124–214. Amsterdam & Philadelphia: Benjamins.
- Bisang, Walter, Andrej Malchukov & the Mainz Grammaticalization Project team (Iris Rieder, Linlin Sun, Marvin Martiny, Svenja Lueli). 2020. Position paper: Universal and areal patterns in grammaticalization. In Walter Bisang & Andrej Malchukov (eds.), *Grammaticalization Scenarios: Cross-linguistic Variation and Universal Tendencies. Vol. 1: Grammaticalization Scenarios from Europe and Asia*, 1–87. Berlin & Boston: De Gruyter Mouton.
- Boye, Kasper & Petar Kehayov (eds.). 2016. *Complementizer Semantics in European Languages*. Berlin & Boston: De Gruyter Mouton.

- Coseriu, Eugenio. 1988. Die Begriffe “Dialekt”, “Niveau” und “Sprachstil” und der eigentliche Sinn der Dialektologie. In J. Albrecht (ed.), *Energie und Ergon (Sprachliche Variation – Sprachgeschichte – Sprachtypologie)*. *Studia in honorem Eugenio Coseriu*, vol. 1, 15–43. Tübingen: Narr.
- Crostafaro, Sonia. 2003. *Subordination*. Oxford: Oxford University Press.
- Croft, William. 2009. *Typology and Universals*. 2nd ed. Cambridge: Cambridge University Press.
- Deđio, Stefan, Peter Ranacher & Paul Widmer. 2019. Evidence for Britain and Ireland as a linguistic area. *Language* 95 (3). 498–522.
- Dik, Simon & Kees Hengeveld. 1991. The hierarchical structure of the clause and the typology of perception verb complements. *Linguistics* 29. 231–259.
- Eckhoff, Hanne Martine & Aleksandrs Berdicevskis. 2015. Linguistics vs. digital editions: The Tromsø Old Russian and OCS Treebank. *Scripta & e-Scripta* 14–15. 9–25.
- Fielder, Grace E. 2019. The semiotics of ideology: the definite article rule in Bulgarian. *Balkanistica* 32 (2). 45–70.
- Gast, Volker & Holger Diessel. 2012. The typology of clause linkage: status quo, challenges, prospects. In Volker Gast & Holger Diessel (eds.), *Clause Linkage in Cross-Linguistic Perspective (Data-Driven Approaches to Cross-Clausal Syntax)*, 1–36. Berlin & Boston: De Gruyter Mouton.
- Giannakidou, Anastasia & Alda Mari. 2016. Mixed (non)veridicality and mood choice with emotive verbs. *Chicago Linguistics Society* 51. 181–196. (available at: https://jeannicod.ccsd.cnrs.fr/ijn_01181251, accessed April 28, 2012).
- Gołęb, Zbigniew. 1990. The ethnic background and internal linguistic mechanism of the so-called Balkanization of Macedonian. *Balkanistica* 10. 13–19.
- Grimshaw, Jane. 1990. *Argument Structure*. Cambridge, MA: MIT Press.
- Grković-Major [Grković-Mejdžor], Jasmina. 2004. Razvoj hipotaktičkog *da* u starosrpskom jeziku [Development of the hypotactic *da* in Old Serbian]. *Zbornik Matice srpske za filologiju i lingvistiku* 47 (1–2). 185–203.
- Grković-Major, Jasmina. 2020. *Da*-clauses (connectives). In Marc L. Greenberg (ed.), *Encyclopedia of Slavic Languages and Linguistics online*. Brill: Leiden. (available at <https://referenceworks.brillonline.com/browse/encyclopedia-of-slavic-languages-and-linguistics-online>).
- Hansen, Björn, Alexander Letuchiy & Izabela Błaszczczyk. 2016. Complementizers in Slavonic (Russian, Polish, and Bulgarian). In Kasper Boye & Petar Kehayov (eds.), *Complementizer Semantics in European Languages*, 175–223. Berlin & Boston: De Gruyter Mouton.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in cross-linguistic studies. *Language* 86 (3). 663–687.
- Haspelmath, Martin. 2015. Descriptive scales versus comparative scales. In Ina Bornkessel-Schlesewsky, Andrej L. Malchukov & Marc D. Richards (eds.), *Scales and Hierarchies (A Cross-Disciplinary Perspective)*, 45–58. Berlin & Boston: De Gruyter Mouton.
- Haug, Dag T. T. & Marius L. Jøhndal. 2008. Creating a Parallel Treebank of the Old Indo-European Bible Translations. In Caroline Sporleder & Kiril Ribarov (eds.), *Proceedings of the Second Workshop on Language Technology for Cultural Heritage Data (LaTeCH 2008)*, 27–34. (available at: <https://proiel.github.io/>)
- Haug, Dag & Hanne Eckhoff. 2011. *The PROIEL corpus as a source to Old Church Slavic: a practical introduction* (Presentation available at: http://www.hf.uio.no/ifikk/english/research/projects/proiel/Activities/proiel/publications/corpus_intro.pdf, accessed March 11, 2016)

- Heine, Bernd & Tania Kuteva. 2005. *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.
- Holzer, Georg. 2014. Vorhistorische Periode. In Tilman Berger, Karl Gutschmidt, Sebastian Kempgen & Peter Kosta (eds.), *Slavische Sprachen (Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung)*, 2. Halbband; Reihe HSK, 1117–1131. Berlin & Boston: De Gruyter Mouton.
- Ivić, Pavle. 1958. *Die serbokroatischen Dialekte. Ihre Struktur und Entwicklung. Erster Band: Allgemeines und die štokavische Dialektgruppe*. 's Gravenhage: Mouton.
- Joseph, Brian D. 1983. *The Synchrony and Diachrony of the Balkan Infinitive: A study in Areal, General, and Historical Linguistics*. Cambridge: Cambridge University Press.
- Joseph, Brian D. 2001. Is Balkan comparative syntax possible? In Maria Luisa Rivero & Angela Ralli (eds.), *Comparative Syntax of Balkan Languages*, 17–43. Oxford: Oxford University Press.
- Joseph, Brian D. 2019. Balkan infinitive loss, event structure and switch reference. *Balkanistica* 32 (2). 137–153.
- Kehayov, P. & K. Boye. 2016. Complementizer semantics in European languages: Overview and generalizations. In Kasper Boye & Petar Kehayov (eds.), *Complementizer Semantics in European Languages*, 809–878. Berlin & Boston: De Gruyter Mouton.
- Koptjevskaja-Tamm, Maria 2012. New directions in lexical typology. *Linguistics* 50 (3). 373–394.
- Koptjevskaja-Tamm, Maria & Bernhard Wälchli. 2001. The Circum-Baltic languages: An areal-typological approach. In Östen Dahl & Maria Koptjevskaja-Tamm (eds.), *The Circum-Baltic Languages. Typology and Contact*, vol. 2, 615–750. Amsterdam & Philadelphia: Benjamins.
- Kurzová, Helena. 2019. Defining the Central European convergence area. In Andrii Danylenko & Motoki Nomachi (eds.), *Slavic on the Language Map of Europe: Historical and Areal-Typological Dimensions*, 261–289. Berlin, Boston: De Gruyter Mouton.
- Lindstedt, Jouko. 2000. Linguistic Balkanization: contact-induced change by mutual reinforcement. In Dicky Gilbers, John Nerbonne & Jos Schaeken (eds.), *Languages in Contact*, 231–246. Amsterdam & Atlanta: Rodopi.
- Matras, Yaron. 2007. The borrowability of grammatical categories. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical Borrowing in Cross-Linguistic Perspective*, 31–74. Berlin & New York: Mouton de Gruyter.
- Meermann, Anastasia & Barbara Sonnenhauser. 2016. Das Perfekt im Serbischen zwischen slavischer und balkanslavischer Entwicklung. In Alena Bazhutkina & Barbara Sonnenhauser (eds.), *Linguistische Beiträge zur Slavistik. XXII. JungslavistInnen-Treffen in München, 12. bis 14. September 2013*, 83–110. Munich: Sagner.
- Nau, Nicole. 2012. Modality in an areal context: The case of a Latgalian dialect. In Björn Wiemer, Bernhard Wälchli & Björn Hansen (eds.), *Grammatical Replication and Borrowability in Language Contact*, 465–508. Berlin, Boston: De Gruyter Mouton.
- Newerkla, Stefan. 2007. Areály jazykového kontaktu ve střední Evropě a německo-český mikroareál ve východním Rakousku [Areas of language contact in central Europe and the German-Czech micro-area in eastern Austria]. *Slovo a slovesnost* 68. 271–286.
- Ranacher, Peter, Nico Neureiter, Rik van Gijn, Barbara Sonnenhauser, Anastasia Escher, Robert Weibel, Pieter Muysken and Balthasar Bickel. 2021. Contact-tracing in cultural evolution: a Bayesian mixture model to detect geographic areas of language contact. *Journal of the Royal Society Interface* (forthcoming).
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of Grammar*, 281–337. Dordrecht: Kluwer.

- Ross, Malcolm D. 1996. Contact-induced change and the comparative method: Cases from Papua New Guinea. In Mark Durie & Malcolm D. Ross (eds.), *The Comparative Method Reviewed*, 180–217. New York & Oxford: Oxford University Press.
- Ross, Malcolm. 2007. Calquing and metatypy. *Journal of Language Contact* 1. 116–143.
- Sakel, Jeanette. 2007. Types of loans: Matter and pattern. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical Borrowing in Cross-Linguistic Perspective*, 15–30. Berlin & New York: Mouton de Gruyter.
- Sandfeld, Kristian. 1926. *Balkanfilologien. En oversigt over dens resultater og problemer*. København: Bianco Lunos. [French translation 1968: *Linguistique balkanique. Problèmes et résultats*. Paris: Klincksieck.]
- Schmidtke-Bode, Karsten. 2014. *Complement Clauses and Complementation Systems: A Cross-Linguistic Study of Grammatical Organization*. Jena (unpubl. PhD thesis).
- Shlonsky, U. 2010. The cartographic enterprise in syntax. *Language and Linguistics Compass* 4 (6). 417–429.
- Šimko, Ivan. 2020. Definiteness markers in the *Life of St Petka*. *Zeitschrift für Slawistik* 65 (2). 272–307.
- Sims, Andrea D. & Brian D. Joseph. 2019. Morphology versus syntax in the Balkan verbal complex. In Iliyana Krapova & Brian D. Joseph (eds.), *Balkan Syntax and (Universal) Principles of Grammar*, 99–150. Berlin & Boston: De Gruyter Mouton.
- Sonnenhauser, Barbara. 2015. Functionalising syntactic variance: declarative complementation with *kako* and *če* in 17th to 19th century Balkan Slavic. *Wiener slavistisches Jahrbuch (Neue Folge)* 3. 41–72.
- Sonnenhauser, Barbara. 2019. Interrogative, indefinite, relative *kdo(r)*. Why Slovene is (not so) different. *Zeitschrift für Slavische Philologie* 75 (1). 151–181.
- Sonnenhauser, Barbara & Paul Widmer. 2020. Indeed, nothing lost in the Balkans. Assessing morphosyntactic convergence in an areal context. *Balkanistica* 33. 103–131.
- Spencer, Andrew. 2003. Periphrastic paradigms in Bulgarian. In Uwe Junghanns (ed.), *Syntactic Structures and Morphological Information*, 249–282. Berlin & New York: Mouton de Gruyter.
- Spencer, Andrew & Ana R. Luís. 2012. *Clitics. An Introduction*. Cambridge: Cambridge University Press.
- Spencer, Andrew & Gergana Popova. 2015. Periphrasis and inflection. In Matthew Baerman (ed.), *The Oxford Handbook of Inflection*, 197–230. Oxford: Oxford University Press.
- Topolińska, Zuzanna. 1997. The opposition [+/- factive] and its main exponents in Macedonian. *Lětopis* 44 (1). 57–66.
- Topolińska, Zuzanna. 2003. Means for grammatical accommodation of finite clauses: Slovenian between South and West Slavic. *Sprachtypologie und Universalienforschung* 56 (3). 306–322.
- Topolińska, Zuzanna. 2008. Factivity as a Grammatical Category in Balkan Slavic and Balkan Romance. In Zuzanna Topolińska: *Z Polski do Macedonii. Studia językoznawcze*, tom I: *Problemy predykcji*, 173–184. Kraków: Lexis. [Reprinted from *Slavia Meridionalis* I (1994). 105–121.]
- Traugott, Elizabeth Closs. 1989. On the rise of epistemic meanings in English: an example of subjectification in semantic change. *Language* 65 (1). 31–55.
- Trudgill, Peter. 2006. *New-Dialect Formation (The Inevitability of Colonial Englishes)*. Edinburgh: Edinburgh University Press.

- van Gelderen, Elly. 2009. Renewal in the left periphery: economy and the complementizer layer. *Transactions of the Philological Society* 107 (2). 131–195.
- Vuković, Teodora. 2020. Torlak. In Marc L. Greenberg (ed.), *Encyclopedia of Slavic Languages and Linguistics online*. Brill: Leiden. (available at: <https://referenceworks.brillonline.com/browse/encyclopedia-of-slavic-languages-and-linguistics-online>)
- Vuković, Teodora. 2021. Representing variation in a spoken corpus of an endangered dialect. The case of Torlak. *Language Resources and Evaluation* 55 (1). 1–26. (available at: <https://link.springer.com/article/10.1007/s10579-020-09522-4>)
- Vuković, Teodora, Nora Muheim, Olivier-Andreas Winistörfer, Anastasia Makarova, Ivan Šimko & Sanja Bradjan. 2019. Corpora and processing tools for non-standard contemporary and diachronic Balkan Slavic. *The 12th International Conference on Recent Advances in Natural Language Processing (RANLP 2019), Varna, Bulgaria, 2 September 2019–4 September 2019*, 62–68.
- Widmer, Paul, Stefan Dedio & Barbara Sonnenhauser. 2021. Convergence by shared ancestry in Romance. *Journal of Language Contact* 14. 53–71.
- Wiemer, Björn. 2015. Među naklonenim i fossilizacij: O mnogolikoj sud’be klitiki by [Between mood and fossilization: on the multi-faceted fate of the clitic by]. In Ljudmila Popović, Dojčil Vojvodić & Motoki Nomachi (eds.), *U prostoru lingvističke slavistike (Zbornik naučnih radova povodom 65 godina života akademika Predraga Pipera)*, 189–224. Belgrade: Univerzitet u Beogradu.
- Wiemer, Björn. 2017. Main clause infinitival predicates and their equivalents in Slavic: Why they are not instances of insubordination. In Łukasz Jędrzejowski & Ulrike Demske (eds.), *Infinitives at the Syntax-Semantics Interface: A Diachronic Perspective*, 265–338. Berlin & Boston: De Gruyter Mouton.
- Wiemer, Björn. 2018. On triangulation in the domain of clause linkage and propositional marking. In Björn Hansen, Jasmina Grković-Major & Barbara Sonnenhauser (eds.), *Diachronic Slavonic Syntax: The Interplay between Internal Development, Language Contact and Metalinguistic Factors*, 285–338. Berlin & Boston: De Gruyter Mouton.
- Wiemer, B. 2019a. “Matrěška” and areal clusters involving varieties of Slavic. On methodology and data treatment. In Andrii Danylenko & Motoki Nomachi (eds.), *Slavic on the Language Map of Europe*, 21–61. Berlin & Boston: De Gruyter Mouton.
- Wiemer, Björn. 2019b. On illusory insubordination and semi-insubordination in Slavic: Independent infinitives, clause-initial particles and predicates put to the test. In Karin Beijering, Gunther Kaltenböck & Mariá Sol Sansiñena (eds.), *Insubordination. Theoretical and Empirical issues*, 107–166. Amsterdam & Philadelphia: Benjamins.
- Wiemer, Björn. 2021. Convergence. In Evangelia Adamou & Yaron Matras (eds.), *The Routledge Handbook on Language Contact*, 276–299. London: Routledge.
- Wiemer, Björn, Il’ja Seržant & Aksana Erker. 2014. Convergence in the Baltic-Slavic contact zone (Triangulation approach). In Juliane Besters-Dilger, Cynthia Dermarkar, Stefan Pfänder & Achim Rabus (eds.), *Congruence in Contact-induced Language Change (Language Families, Typological Resemblance, and Perceived Similarity)*, 15–42. Berlin & New York: De Gruyter.
- Ziegeler, Debra. 2017. Historical replication in contact grammaticalization. In Daniël Van Olmen, Hubert Cuyckens & Lobke Ghesquière (eds.), *Aspects of Grammaticalization. (Inter)Subjectification and Directionality*, 311–352. Berlin & Boston: De Gruyter Mouton.